University of the Free State
Department of Mathematical Statistics

Bayesian Analysis of Process Capability Indices for Single and Multiple Sources of Variability

Delson Chikobvu

Bayesian Analysis of Process Capability Indices for Single and Multiple Sources of Variability

By

Delson Chikobvu

Dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Mathematical Statistics

in the

Faculty of Agriculture and Natural Sciences

Department of Mathematical Statistics

at the

University of the Free State

June 2008

Supervisor

Professor AJ van der Merwe

Declaration

I declare that the thesis hereby submitted by me for the Doctor of Philosophy in Mathematical Statistics degree at the University of the Free State is my own independent work and has not previously been submitted by me at another university or faculty. I further more cede copyright of the thesis in favour of the University of the Free State.

.

…………………. ………………….

Signature Date

Acknowledgements

I wish thank and express my appreciation to all people who contributed in so many ways to facilitate the completion of my dissertation.

In particular I am forever thankful to:

- My supervisor Professor AJ van der Merwe for his patience, motivation, professional guidance and intellectual support;

- My wife Perpetual and two daughters, Irene and Emily for their time sacrifice and support during the compilation of this thesis.

I also wish to express my gratitude to the Almighty for giving me strength and courage to eventually finish this thesis.

SUMMARY

Process capability index (process performance index) -relates the specification limits to the performance of a process, it reduces complex information about the performance of a process to a single number. A capability index is a dimensionless measure of relative variability. In this thesis, Bayesian statistics is employed to simulate and estimate most of the widely used process capability indices.

In Bayesian analysis, we assume that we have prior knowledge or information or opinion about parameters of a statistical distribution and very often in practice we do. We then attach a distribution to this belief. Parameters do not really have a distribution, parameters are constants, and so a prior distribution is a way of expressing our belief or opinion on our parameters. A posterior distribution is the belief distribution of the parameters after the outcomes of experiments (data) have been observed. There is now an updated belief distribution in light of the information from the data.

Bayesian inference is shown to have a number of advantages. A full Bayesian analysis provides a natural way of taking into account all sources of uncertainty in the estimation of the parameters. Uncertainty about the true value of the process capability index is incorporated into the analysis through the choice of a prior distribution. The most familiar element of the Bayesian school is the use of the non-informative (objective) prior distribution, designed to be minimally informative in some sense. The most famous of these is the Jeffrey's-rule prior and is utilised throughout the thesis. Scientists hold up objectivity as the ideal of science. Reference priors are a refinement of the Jeffrey's-rule priors for higher dimensional problems that have proven to be remarkably successful. The probability matching prior is recommended because it is designed to produce posterior credible intervals which are asymptotically identical to their frequentist counterparts.

The Bayesian simulation procedure employs the posterior distribution of the parameters in doing the simulations. The procedure is also shown to be useful and comparable to existing classical statistical procedures in solving the supplier selection problem.

Data arising from multiple sources of variability are very common in practice. Virtually all industrial processes exhibit between-batch and within-batch components of variation. In some cases the between-batch (or between subgroup) component is viewed as part of the common-cause-system for the process. A process capability index in more general settings is developed using C_{pl} as a point of reference. C_{pl} is a single variance index and is adapted to give 2 and 3 variance components indices. The variance component model proves to be suitable for handling multiple sources of variability capability indices. Again, Bayesian simulation methods prove to be useful in handling these multiple sources of variability indices.

Results show that the Bayesian simulation approach is just as good if not better than the standard classical statistics approach in assessing the capability of an industrial process. The added advantage of the Bayesian approach is that, from the posterior distribution of the capability indices, we are in a position to obtain quantiles, credible regions and perform other inferential tasks.

KEY WORDS: Bayesian analysis, Moments, Monte Carlo simulation, Non-informative prior, Pearson’s curve, Posterior distribution, Probability matching prior, Process capability index, Reference prior, Variance components

OPSOMMING

Prosesgeskiktheidsanalise verwys na die moontlikheid om die Bayes-simulasiebenadering toe te pas op prosesgeskiktheidsindekse soos onder andere C_p , C_{pk} , P_p en P_{pk} . In hierdie verhandeling word Bayes-statistiek gebruik om die meeste van die prosesgeskiktheidsindekse te simuleer en te beraam.

In Bayes-analise neem ons aan dat ons prior kennis of inligting of ‘n opinie het aangaande parameters van ‘n statistiese verdeling, soos die geval dikwels in die praktyk is. ‘n Verdeling kan dan aan hierdie oortuiging gekoppel word. Parameters is konstantes en het nie regtig ‘n verdeling nie, dus is ‘n priorverdeling ‘n manier om ons opinie of oortuiging aangaande parameters uit te druk. ‘n Posteriorverdeling is ‘n oortuigingsverdeling van die parameters nadat die uitkomste of eksperimente (data) waargeneem is. Daar is nou ‘n opgedateerde oortuigingsverdeling in die lig van die inligting uit die data bekom.

Bayes-inferensie het ‘n hele aantal voordele. ‘n Volledige Bayes-analise voorsien ‘n natuurlike manier om alle bronne van onsekerheid met die beraming van die parameters in ag te neem. Onsekerheid oor die werklike waarde van die prosesgeskiktheidsindeks word in die analise ingesluit deur middel van die keuse van ‘n priorverdeling. Die mees bekende element van die Bayesskool is die gebruik van die objektiewe priorverdeling, wat ontwerp is om minimale inligting in ‘n sekere sin te gee. Die mees gewildste een is die Jeffreys-reël prior wat deurgaans in die verhandeling gebruik word. Wetenskaplikes hou objektiwiteit as die ideaal van wetenskap voor. Verwysingspriors is ‘n verfyning van die Jeffreys-reël priors vir hoër dimensionele probleme wat reeds as suksesvol beskou word. Die waarskynlikheidsgepaste prior word aanbeveel omdat dit ontwerp is om posterior kredietwaardigheidsintervalle te lewer wat assimptoties identies is aan hulle frekwentistiese teenpartye.

Die Bayes-simulasieprosedure gebruik die posteriorverdeling om die simulasies uit te voer. Die prosedure het getoon dat dit geskik en vergelykbaar is met bestaande klassieke statistiese procedures om die verskaffer-seleksieprobleem op te los.

Data wat uit meervoudige bronne van variasie voortspruit is baie algemeen in die praktyk. Letterlik alle industriële prosesse toon tussengroep en binnegroep komponente van variasie. In sommige gevalle word die tussengroepkomponent beskou as deel van die algemeen-oorsaak-sisteem van die proses. ‘n Prosesgeskiktheidsindeks in meer algemene omstandighede is ontwikkel deur C_{pl} as ‘n puntverwysing te gebruik. C_{pl} is ‘n enkel variansie-indeks en is aangepas om 2 en 3 variansiekomponentindekse te gee. Daar is bewys dat die variansiekomponentmodel geskik is vir die hantering van meervoudige bronne van variasiegeskiktheidsindekse. Weereens kan bewys word dat Bayes-simulasiemetodes geskik is vir die hantering van hierdie meervoudige bronne van variasie-indekse.

Resultate toon dat die Bayes-simulasiebenadering net so goed, indien nie beter nie, is as die standaard klassieke statistiekbenadering om die vermoë van die industriële proses te assesseer. ‘n Bykomende voordeel van die Bayesbenadering is dat, vanuit die priorverdeling van die geskiktheidsindekse, die moontlikheid geskep word om kwantiele en kredietwaardigheidsintervalle te bekom, asook om ander inferensiële take uit te voer.

TABLE OF CONTENTS

NOTATION AND TERMINOLOGY

Y -some characteristic of interest of a manufactured product.

T -‘nominal’ or ‘target value’ of Y which will satisfy the design engineer’s criteria for the optimum performance of a product.

Specification limits -upper and lower specification limits denoted as USL and LSL, or simply ℓ_1 and ℓ_0 respectively, and to require that Y be within these limits.

$(USL-LSL)$ –length of the specification interval (tolerance interval).

d -half the length of the specification interval i.e. $d=\dfrac{USL-LSL}{2}$.

M -the midpoint of the specification interval i.e. $M=\dfrac{USL+LSL}{2}$

μ -the mean of a production process

σ -a measure of variability of the production process as measured by the standard deviation.

Process capability index (process performance index) -relates the specification limits to the performance of a process, it reduces complex information about the performance of a process to a single number. A capability index is a dimensionless measure of relative variability. Various indices have been proposed by various authors. The table below summarises the main indices discussed in this research.

Approach	**Formula**
Process capability index	$C_p = \frac{USL - LSL}{6\sigma}$
Process capability index	$C_{pk} = \min\left(\frac{USL-\mu}{3\sigma}, \frac{\mu - LSL}{3\sigma}\right)$
Process capability index	$C_{pl} = \frac{\mu - LSL}{3\sigma}$
Process capability index	$C_{pu} = \frac{USL-\mu}{3\sigma}$
Process performance index	$P_p = \frac{USL-LSL}{\sigma_{total}}$
Process performance index	$P_{pk} = \min\left(\frac{USL-\mu}{3\sigma_{total}}, \frac{\mu-LSL}{3\sigma_{total}}\right)$
Process performance index	$C_pT = \min\left(\frac{USL-T}{3\sigma}, \frac{T-LSL}{3\sigma}\right)$
Process performance index	$C_{pm} = \frac{USL-LSL}{6\sqrt{\sigma^2 + (\mu - T)^2}}$
Process performance index	$C_{pmk} = \frac{\min(USL-\mu, \mu - LSL)}{3\sqrt{\sigma^2 + (\mu-T)^2}}$
Process performance index	$C_{pm}^{\#} = \frac{\min(USL-T, T-LSL)}{3\sqrt{\sigma^2+(\mu-T)^2}}$
Unified index	$C_p(u,v) = \frac{d - u\lvert\mu - M\rvert}{3\sqrt{\sigma^2 + v(\mu-T)^2}}$
Normative index	$C_B(D) = \frac{1}{3}\Phi^{-1}\{\Pr(y \in A \mid D)\}$

CHAPTER 1

OVERVIEW OF CAPABILITY INDICES

1.1 INTRODUCTION

Capability indices are tricky to interpret, controversial to apply and often misunderstood by many practitioners. Unless the properties of an index are clearly understood, making major capital improvements may not be the most prudent way to fix an unacceptable capability. Understanding the meaning of a particular index can have a profound impact on the cost of manufacturing. Process improvement must be driven by more than the need to improve an index number, otherwise management may be wasting time and money. This chapter provides a broad overview of capability indices and what they measure.

A capability index relates the voice of the customer (specification limits) to the voice of the process. A capability index is convenient because it reduces complex information about the process to a single number.

This research discusses versions of the process capability or performance index including indices derived from hierarchical models with more than one variance component. The term capability index will be used as a generic term.

This chapter concentrates on what is measured by a process capability index. Some problems associated with application of process capability indices are discussed. One process capability index C_{pk} is widely used to determine whether manufacturing processes are capable of meeting specifications. Therefore a critical look is taken at what C_{pk}, together with various other indices, measure. Two components of C_{pk} namely C_{pl} and C_{pu} will specifically be investigated. Various capability indices such as C_p, C_{pk}, C_{pm}, to name a few, are investigated. The inter-relationships between the indices are also examined. It is rather surprising and interesting to see the complexities of connections among many of the capability indices.

Most literature would simply suggest that management must choose the 'correct' index for their application or process. Each index states something different and unless one knows what they measure, one may end up using the wrong index and making the wrong decisions.

1.2 DEFINITIONS AND NOTATIONS

Let Y be some characteristics of interest of a manufactured product. The engineering or design specifications for Y are generally stated in terms of a 'nominal-' or 'target value', say T. That is, T is the value of Y which will satisfy the design engineer's criteria for the optimum performance of the product. Manufacturing the product so that Y exactly equals T is prohibitively expensive, and thus it is common practice to specify upper and lower 'specification' limits, USL and LSL, or simply ℓ_1 and ℓ_0 respectively, and to require that Y be within these limits.

Tolerance limits (specification limits) are limits that define the conformance boundaries for an individual unit of a manufacturing or service unit. An upper tolerance limit (upper specification limit) is a tolerance limit applicable to the upper conformance boundary for an individual unit of a manufacturing or service operation. A lower tolerance limit (lower specification limit) is a tolerance limit that defines the lower conformance boundary for an individual unit of a manufacturing or service operation.

The physical processes that manufacture the parts are generally subject to many sources of variation, starting from the quality of raw material to the aging and wear-out of the manufacturing equipment. Consequently, Y is a random quantity (or a random variable), whose distribution is often assumed to be Gaussian with mean, say μ, and a variance, say σ^2. In manufacturing parlance, the variance is referred to as the natural tolerance of Y. When working with process capability indices it is common practice to assume that both μ and σ^2 do not change with time; i.e. the process is stable, or what is known in quality control as being in statistical control.

Statistical tolerance limits are the limits of the interval for which it can be stated with a given level of confidence that it contains at least a specified proportion of the population of production.

There is no direct connection or relationship between the statistical tolerance limits (control limits) on a process and the specification limits on a product. The control limits are driven by the natural variability of the process (measured by the standard deviation σ). That is, the statistical tolerance limits are driven by the natural tolerance of the process. It is customary to define the upper and lower natural tolerance limits, say $UNTL$ and $LNTL$, as 3σ above and below the process mean, i.e. $\mu \pm 3\sigma$. The specification limits, on the other hand, are determined externally. They may be set by management, the manufacturing engineers, the customer, the standards authority, or by the product developers/designers. However, one should have knowledge of inherent process variability when setting specifications, but there is no mathematical or statistical relationship between the control limits and the specification limits.

The question which arises is whether the design engineer's compromise in going from the ideal T to the upper and lower specifications limits (the USL and the LSL), is matched by the manufacturer's ability to meet such a compromise vis-à-vis the assumed μ and σ^2 mentioned above. Process capability indices are introduced to address this matter. The quantity ($USL - LSL$) is known as the specification interval (or tolerance interval); it will be denoted by $2d$, where d is the half length of the specification interval. The midpoint of the specification interval, which will be denoted by M, is equal to ($M = \frac{USL + LSL}{2}$) (see figure 1.1).

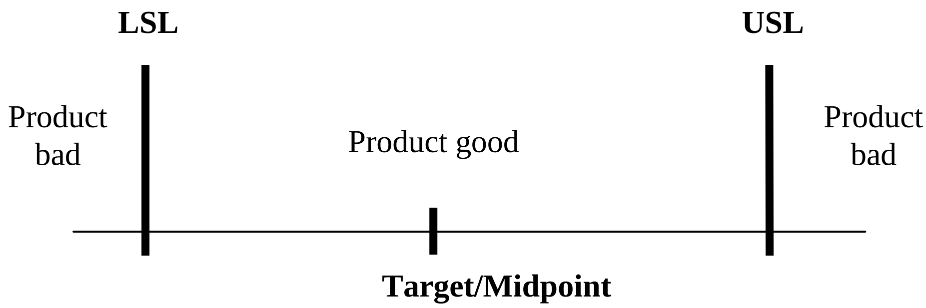


Figure 1.1: "USL and LSL"

1.3 BACKGROUND

How does one go about defining an index? Capability indices, similar to coefficients of variation, are dimensionless measures of relative variability. It is a ratio - a number without units of measurement - that compares specification range to natural tolerance and results in a single number. That number is then judged acceptable or unacceptable by some arbitrary standard. An index can also be used to compare one process to another or set a minimum acceptable quality standard for processes.

A capability index should be computed using data from a stable process. Typically, process stability is assessed by collecting sub-samples at regular intervals and plotting sub-sample statistics on control charts. Once the charts show a reasonable degree of stability, process capability can be assessed.

Capability analysis is used in many facets of industrial processes and is beginning to be used in business processes as well. Capability analysis and thresholds for capability indices are used in the quantification of processes, acceptance of equipment, purchase parts approval activities, continuous improvement efforts, problem solving activities and for many other purposes. It is the backbone of measuring processes' ability to produce product that falls within a desired specification through the enumeration of variation. Capability indices provide a yardstick for measuring improvement. The accuracy of capability indices is dependent on proper understanding of the theory behind the indices as well as an understanding of variation.

All the indices considered in this chapter have their individual merits and demerits, which helps in coming up with characteristics that are crucial in the process of quality assessment. Each one of these takes into account at least one, but not all, of these characteristics needed for the full quality picture of the process under consideration. No one approach has taken into account all these characteristics. The quality characteristics of a production process that are usually considered are listed below:

- Inherent variation
- Total variation
- Normality of distribution
- Stability
- Target value
- Bias
- Potential
- Sensitivity to variation from target
- Symmetric tolerance
- Asymmetric tolerance
- Proactive (Predictive & control)
- Retroactive (assessment and monitoring)

The inherent process variation is the variation caused by common causes only and is used to represent the true process capability and potential. Inherent variation is often estimated from control charts after verifying stability. In the absence of a computer package, the inherent variation is estimated by:

$$\hat{\sigma} = \frac{\bar{R}}{d_2}$$

$\hat{\sigma}$ is the inherent process variation, $\bar{R}$ is the average of sample ranges taken from the R chart and d_2 is a constant that depends on the sample size of subgroups taken in the chart.

The total variation is the variation caused by common cause and special cause variation and represents the current process performance. Total process variation is estimated directly from the process data by the following formula:

$$\hat{\sigma}^2_{total} = \frac{1}{n-1}\sum_{j=1}^{n}(Y_j - \overline{Y})^2 \text{ where } \overline{Y} = \frac{1}{n}\sum_{j=1}^{n} Y_j \,.$$

Total process variation is meant to take into consideration all forms of variation including inherent variation, special cause, mean variation between groups (referred to as shift and drift) and mean deviation of an entire population (referred to as target variation). Proper characterisation of this type of variation in capability analysis is dependent upon the total process variation encompassing all of the potential types of variation. The total process variation encompassing all of the potential types of variation is dependent upon a well developed rational sampling strategy.

An important assumption is that of normality of distribution of the data from the process. Normality is typically assessed visually through a histogram or a normal probability plot and quantitatively through a normal hypothesis test. The process must be 'normal' prior to calculation of capability indices. If the process is not normally distributed, a practitioner can apply a transformation to the data to make it normal.

Another important assumption is statistical stability. Statistical stability is an underlying assumption regarding the process and hence the data in the capability analysis. Stability implies that within any two rating periods, the underlying distribution, which generates the indices, does not change (over time). Thus gradual drifts in the process mean and/or the process variance are not allowed. The assumption is a precursor in that any results with capability indices are only valid if the assumption holds. If the process is not in statistical control, special causes must be identified and corrective action taken prior to reporting of capability results.

As mentioned earlier, it is often required that each quality characteristic have a target or nominal value. The objective is to reduce variability around this target. Any difference between the location of the process and the target value gives rise to bias $(\mu \neq T)$. The mean bias is the absolute value of the average deviation from the target $|\mu - T|$. Ideally, manufacturers would want to produce components in such a way that each dimension is, on average, at the specification target. If location of the process average and the target value coincides, then there is no bias $(\mu = T)$.

Some indices ignore the current estimate of the process mean and relate the specification range directly to the process variation. In effect, such indices can be considered measures that suggest how capable (potential rather than actual capability) the process could be if the process mean was centred midway between specification limits in the case of symmetrical tolerance intervals. If an index is not sensitive to the distance between process mean and target value then it is essentially a measure of process "potential only" (Lynch, 2004).

If any index is robust against departures from the target, then it is not sensitive to variation from the target, otherwise it is sensitive to this variation. Some process capability indices do not evaluate where the process average is, or if it is centred with respect to the nominal (target) of the specifications. These indices are insensitive to deviation from the target. It is actually possible to have a process producing product that is 100% out of specification but associated with an acceptably high value of the index.

In symmetrical tolerance, the target is the midpoint of the tolerance interval (midpoint of the specification range). Asymmetrical tolerance intervals appear when the target is not centred at the midpoint of the specification interval. Symmetrical tolerances imply that $T = M$, and in the case of asymmetrical tolerances $T \neq M$.

Often, calculated process capability measures from different processes within a given plant cannot be averaged, even when using the same measure of capability. However, some measures can, as when it is desired to evaluate the overall quality for the entire plant.

A much more useful role can be served by the process capability indices if they can also be used to perform the proactive functions of prediction and control of the quality of future output as opposed to the traditional passive retroactive role of assessing and monitoring current quality. A few of the process capability indices can be used to predict and to control the quality of future output. Here, one monitors the observable Y (rather than the unobservable mean μ), and makes a decision to continue production, to modulate it, or stop it, based on the consequences of the deviation of Y from T. The decision is proactive and is dictated by the predictive distribution of Y and the utilities associated with the deviation of Y from T, and also the utilities associated with the control of the process. Most of the traditional indices are, however, passive devices whose main role is to retroactively monitor and assess process capability. Their purpose is to ensure (but only retrospectively) that the number of non-conforming items in a batch is below a specific limit. The functions of assessing and monitoring are not predictive, nor are they proactive, and thus these indices mainly serve as policing devices (Singpurwalla, 1998).

The list of indices considered in the next sections is as follows:

$C_{p,}$ C_{pk}, $C_{pl,}$ C_{pu}, $P_{p,}$ P_{pk}, C_pT, $C_{pm,}$ C_{pmk}, $C^{\#}{}_{pm.}$

1.3.1 PROCESS CONTROL AND PROCESS CAPABILITY INDICES

Statistical process control (SPC) and quality improvement methods are generally based on control charts which are used for monitoring relevant process characteristics, like process capability indices (PCI) which were developed for measuring uniformity of the process. The main goal of SPC consists of keeping small process variation around a given target value and thus guaranteeing a small number of nonconforming items produced and a large PCI value. Process capability analysis includes substantially more than just the computation of any index. After process control has been established, capability is assessed.

The use of classical univariate PCI is based on the following assumptions:

1. There is only one quality characteristic to be considered;
2. The distribution of the quality characteristic is approximately normal;
3. The quality characteristics of different items are stochastically independent;

4. The process is under statistical control, that is, the process mean and process variability are constant;
5. The sample size is large enough so that calculation of standard deviation is rational.

After process control has been established, capability is assessed. Assessment is essentially the act of comparing the distribution of data, or a model, to the engineering requirements, typically in the form of engineering specifications. If the process is deemed capable, then the process will be maintained using statistical process control methods. If, on the other hand, the process is deemed not capable, i.e. it is producing an unacceptable level of non-conforming product, then the process will undergo a process improvement stage and work toward an acceptable level of capability and control.

Other researchers Kane (1986), Chan *et al.* (1988), Choi and Owen (1990), Pearn *et al.* (1992) and Greenwich and Jahr-Schaffrath (1995) address different process capability indices for providing measures for process potential and process performance. The initially proposed process capability index (PCI) is C_p. It is proposed by Juran (1974) has its foundation in a fundamental result of probability, namely Chebyshev's inequality, which states:

Theorem 1.3.1

Chebyshev's inequality

Let Y be a random variable with mean μ and variance σ^2. Then, for any c >0,

$$P|Y-\mu| > c) \le \frac{\sigma^2}{c^2}$$

<u>Proof</u>

The proof is given in Appendix A1.

And for $c = 3\sigma$

$$P(|Y-\mu|>c)\le\frac{\sigma^2}{c^2}$$
$$=\frac{\sigma^2}{(3\sigma)^2}$$
$$=\frac{1}{9}\approx 0.1.$$

The essence of this inequality is the result that the probability of any random variable deviating from its mean by more than three times its standard deviation is small, at most 0.1. This inequality, though sharp, is too broad and too general to be of much practical value and demands of the user only the knowledge of the variance.

Suppose that we wish to find a more exact probability $P(|Y-\mu|>c)$ for some constant c. We can then use the central limit theorem to approximate this probability, we first standardise, using mean $E\left(Y\right)=\mu$ and variance $Var(Y)=\sigma^2$.

$$P(|Y-\mu|>c)=1-P(-c<Y-\mu<c)$$
$$=1-P(\frac{-c}{\sigma}<\frac{Y-\mu}{\sigma}<\frac{c}{\sigma})$$

$P(|Y-\mu|>c)\approx 1-P(\frac{-c}{\sigma}<Z<\frac{c}{\sigma})$, where $Z\sim N(0,1)$ i.e. standard normal distribution.

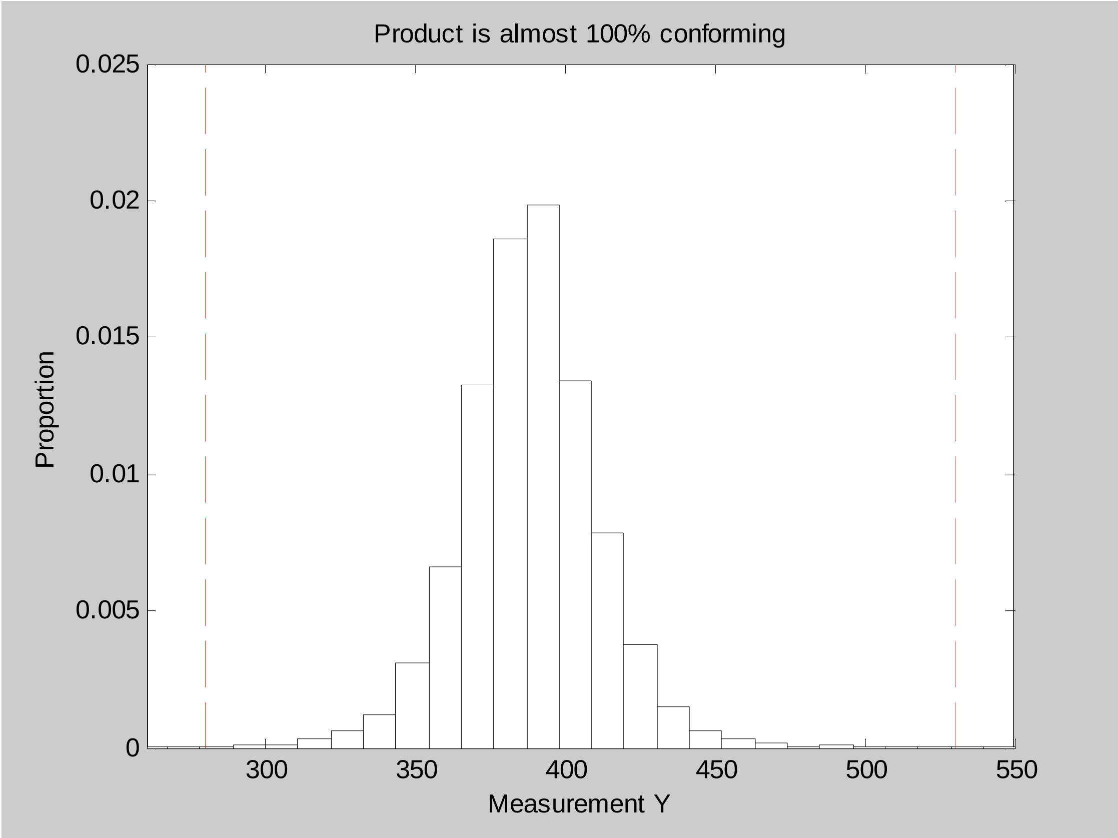
And for $c=USL-\mu=\mu-LSL=3\sigma$

$$P(|Y-\mu|>c)=1-\left(\Phi(\frac{3\sigma}{\sigma})-\Phi(\frac{-3\sigma}{\sigma})\right)\text{, where } Z\sim N(0,1)$$
$$=1-\left(\Phi(3)-\Phi(-3)\right)=0.0027,$$

where Φ denotes standard normal cumulative distribution function.

This probability will later be described as the expected proportion of product that is non-conforming to the specifications.

C_p is a powerful index that provides a quick observation to determine whether the process is capable of meeting specification. You could also say that C_p is the ratio between what you want the process to do (management's hope or allowable spread) versus what the process is actually doing (reality or actual).



LSL=280 and USL=530

Figure 1.2: Production from a process which is capable

$$C_p = \frac{Hope}{Reality}.$$

It was initially known as the capability ratio (Kotz and Johnson, 2002). It is a measure of tolerance spread to process spread (see figure 1.2) and is calculated as:

$$C_p = \frac{Tolerance\,Spread}{6\sigma\ \ Spread},$$

$$C_p = \frac{USL - LSL}{6\sigma} = \frac{d}{3\sigma}, \qquad (1.3.1)$$

where USL and LSL are the upper and lower specification limits respectively and $d = \frac{USL - LSL}{2}$. σ is the within subgroup standard deviation.

It is often required that for acceptance we should have $C_p \geq c$ with c=1, 1.33, 1.5 or 1.67 corresponding to $USL - LSL = 6\sigma$, 8σ 9σ or 10σ. Large values of C_p are desirable and small values undesirable (because a large standard deviation is undesirable).

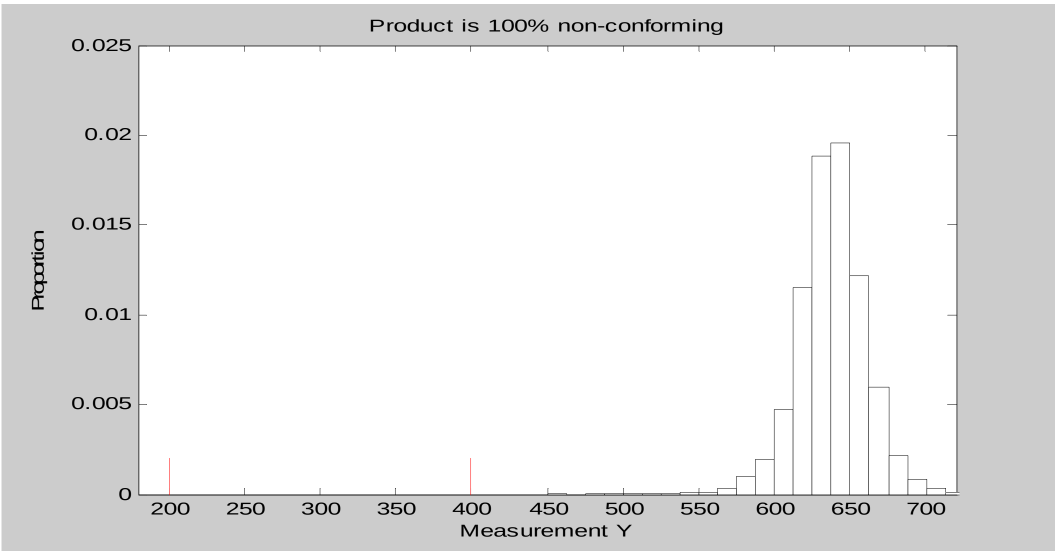
Depending on the index value, target centred processes can be classified into five different categories:

- $C_p < 1.00$ inadequate/incapable
- $1.00 \leq C_p < 1.33$ capable
- $1.33 \leq C_p < 1.50$ satisfactory
- $1.50 \leq C_p < 1.67$ excellent
- $C_p \geq 1.67$ superb.

If C_p=1 then $USL = \mu + 3\sigma$ and $LSL = \mu - 3\sigma$, the expected proportion non-conforming product when the process is centred is 0.27%, which is regarded as 'acceptably small'. As long as μ coincides with the target T, any value of C_p greater than 1 will decrease the above probability, making the process more efficient. Since σ^2 is unknown, it has to be inferred from the data, and to compensate for the uncertainties of estimation, industrial practice follows the dictum that C_p must be a minimum of 1.33 (instead of the aforementioned 1).

The choice 1.33 is completely ad hoc; indeed for pilot (or qualifications) runs, C_p is sometimes required to be in excess of values as high as 1.5 and 1.6. A possible explanation for the value of c=1.33 is that the formula for C_p quantifies a rule of thumb quality engineers have used for decades: the process spread should be no more than 75% of the specification interval ($1.33 = \frac{1.00}{0.75}$). A C_p of 1.33 or greater would yield a good process. Quality engineers know that a 75% ratio allows the process average to drift naturally while still controlling the overall process average within acceptable boundaries. A possible explanation for a $C_p >$ 1.67 is that 1.67 corresponds, approximately, to a rejection-rate of one unit per million; see for example Spiring *et al* (2002). Large values of C_p increase the cost of manufacturing. The data needed to estimate σ^2, mentioned above, is taken at certain specified points in time called rating periods. The specification of the rating periods also appears to be based on arbitrary considerations.

C_p compares one process spread to another. It does not, for instance, evaluate where the process average is or if it is centred with respect to the nominal (target) of the specifications. It is actually possible to have a process producing product that is 100% out of specification but associated with an acceptably high value of the index (see figure 1.3 below). Figure 1.3 shows production from a potentially capable process which is currently producing product that is 100% non-conforming, and C_p will not detect this. This is because it relates the specification range directly to the process variation, without worrying about the location of the process.



LSL=200 and USL=400

Figure 1.3: Production from a potentially capable process which is currently producing product that is 100% non-conforming

Therefore C_p has its limitations, but it can serve as a powerful tool once you understand its strengths and weaknesses.

Despite its common use in industry, enhancements and refinements of C_p have been proposed.

Kane (1986) proposed C_{pk} as a PCI.

$$C_{pk} = \min\left(\frac{USL-\mu}{3\sigma}, \frac{\mu-LSL}{3\sigma}\right) = \frac{d-|\mu-M|}{3\sigma} \tag{1.3.2}$$

μ is the process mean, $M = \dfrac{USL+LSL}{2}$ and $d = \dfrac{USL-LSL}{2}$.

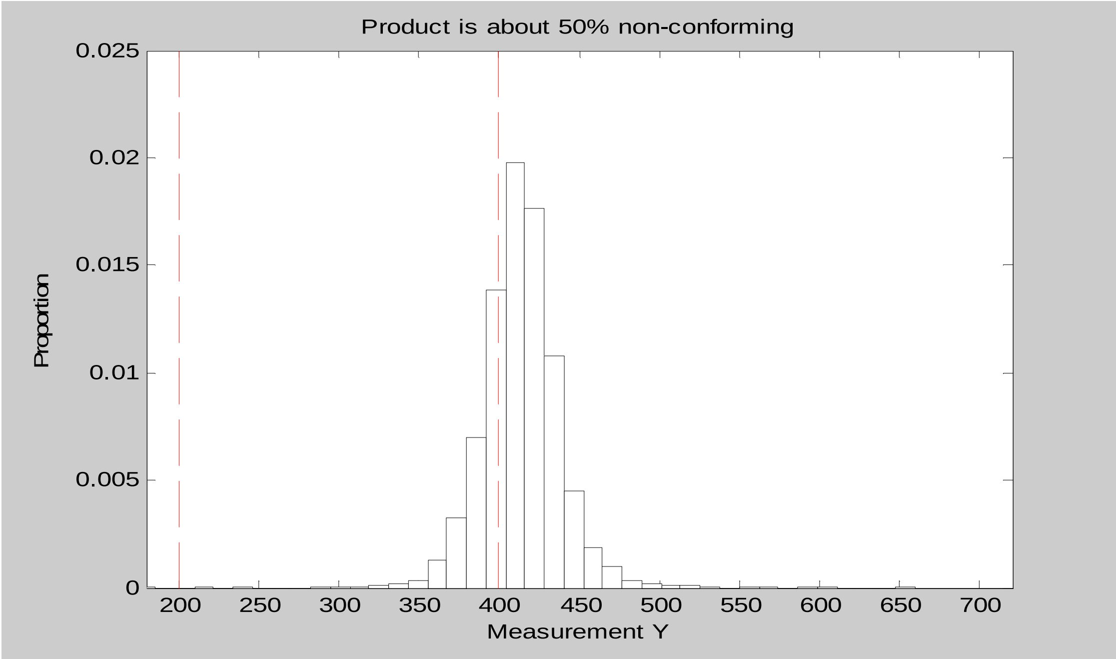
Notice that C_{pk} is made up of two indices namely

$$C_{pu} = \frac{USL-\mu}{3\sigma} \tag{1.3.3}$$

and

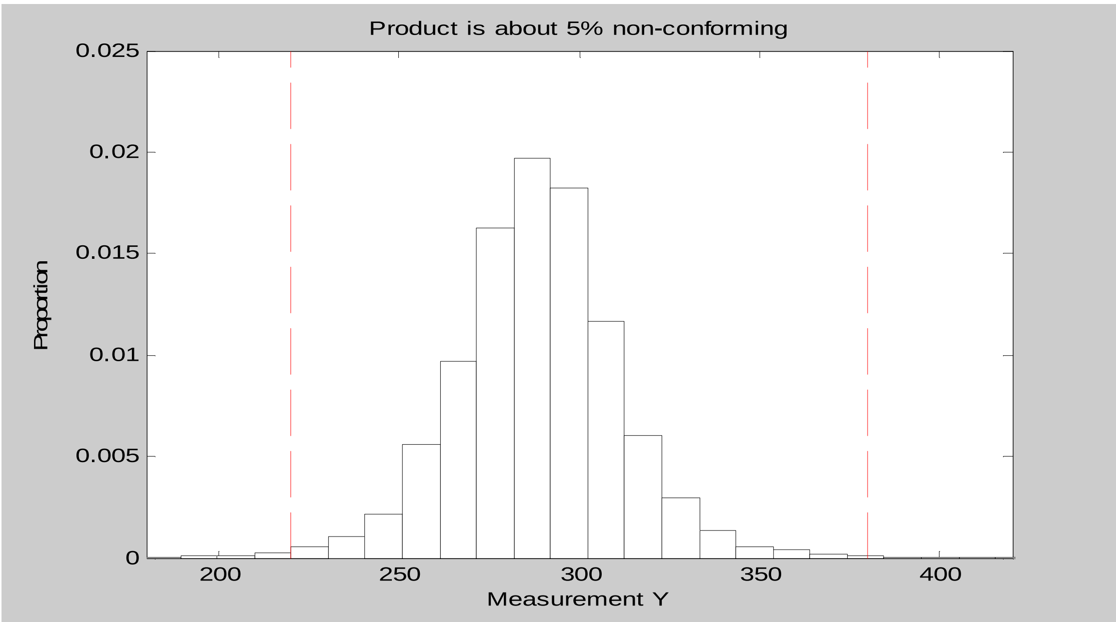
$$C_{pl} = \frac{\mu-LSL}{3\sigma}, \tag{1.3.4}$$

hence can be written as $C_{pk} = \min\left(C_{pu}, C_{pl}\right)$. Negative values of C_{pk} occur when the process average is positioned outside of the specification interval (see figure 1.4). Whenever C_p is “large” and C_{pk} is “small,” then μ is not centred at the middle of the tolerance interval.



LSL=200 and USL=400

Figure 1.4: A process with an average positioned outside of the specification interval



LSL=200 and USL=400

Figure 1.5: A process which is centred at the middle of the specification interval but with a wide spread

In situations where both C_p and C_{pk} are "small," μ is centred near the middle of the specification interval but the process spread is too wide (see figure 1.5).

If $C_{pk} = 1$, it can be shown that $M - d < \mu < M + d$.

Because the process average is part of the calculation, some believe the C_{pk} formula incorporates process centring. This is an erroneous assumption, because you do not know how far the process average is from the target. C_{pk} indicates where the process average is, but does not cover process centring. The C_{pk} index evaluates half the process spread with respect to where the process is actually located (some point in space). C_{pk} offers the most information about the proportion non-conforming, say p, and it will be shown later that it provides the least insight about the location μ.

C_{pk} is inappropriate for product features with asymmetric tolerances, i.e., $T \neq M$ where T is the target and M is the midpoint of the tolerance interval. Assuming a normal distribution for the process output, the C_{pk} index will achieve its highest value when the mean, μ is located at M. However, optimal product performance occurs when μ is positioned at T.

C_{pk} is not meaningful for a process which is not in statistical control, and PPM (parts per million nonconforming), as it is often estimated, can be grossly wrong unless the process of interest is in statistical control. Furthermore, C_{pk} and PPM are questionable when they are applied to populations that are not normally distributed. For such processes, the capability indices do not describe what fraction of the process output will fall between specification limits and the PPM estimates can be severely in error.

Boyles (1991) points out that "the C_p and C_{pk} do not say anything about the distance between process mean and target value" and "are essentially a measure of process potential only". Boyles showed that C_{pk} becomes arbitrarily large as σ approaches 0, irrespective of where the process is centred and this characteristic makes C_{pk} unsuitable as a measure of process centring (Boyles, 1991). The same is true for C_p. Notice also that, since

$$
\begin{aligned}
C_p &= \frac{USL - LSL}{6\sigma} \\
&= \frac{1}{2}\left(\frac{USL-\mu}{3\sigma} + \frac{\mu - LSL}{3\sigma}\right) \\
\therefore C_p &= \frac{1}{2}\left(C_{pu} + C_{pl}\right)
\end{aligned}
$$

if the process is centred within the specification range.

Herman (1989) provides a thought-provoking criticism of the PCI concept. The σ is intended to represent process variability when production is 'in control'. But usually variation has two components – from within-lots and among-lots variation. The σ in the denominator of C_p is intended to refer to within-lot process variation. This σ can be considerably less than, say, the overall standard deviation ,σ_{total}.

Herman suggests that a different index, the 'process performance index' (PPI), P_p might 'have more value to a customer than C_p' (Herman, 1989) and P_p is defined as:

$$P_p = \frac{USL - LSL}{6\sigma_{total}}. \qquad (1.3.5)$$

An analogy to C_{pk} is:

$$P_{pk} = \min\left(\frac{USL-\mu}{3\sigma_{total}}, \frac{\mu - LSL}{3\sigma_{total}}\right). \qquad (1.3.6)$$

Other literature refers to P_{pk} as the preliminary process capability. It is used whenever a new process is started or a major revision to an existing process is resumed. This is why some practitioners mistakenly assume P_{pk} is for short-term data and is to be used on an unstable process. Both assumptions are false. P_{pk} is an initial production run of a new process (less than 30 production days), and C_{pk} is everything thereafter. The Automotive Industry Action Group (AIAG) specifies a P_{pk} value of greater than 1.67. P_p is the companion automotive index to P_{pk}. Like the P_p index, this preliminary process capability index is used when a new process is started or a major process modification is initially resumed and does not result in short-term data.

The two capability indices, C_{pk} and P_{pk}, are widely used. The index P_p, and related C_p, are similar to C_{pk} and P_{pk}. However, P_p and C_p ignore the current estimate of the process mean and relate the specification range directly to the process variation. In effect, C_p and P_p can be considered measures that suggest how capable the process could be if the process mean were centred midway between specification limits. The indices P_p and C_p are not recommended for reporting purposes, since the information they provide to supplement C_{pk} and P_{pk} is also easily obtained from a histogram of the data. The measures C_{pk} and P_{pk} again differ only in the estimate of the process standard deviation used in the denominator. Since σ is usually calculated based on subgroup ranges, it uses only the variability within each group to estimate the process standard deviation. The simple standard deviation-based estimate (*s)* combines all the data together, and thus uses both the within and between subgroup variability.

As a result, the capability of a process should be based on the process' total variation, that is, we should use the capability index P_{pk} rather than C_{pk}. C_{pk} seriously underestimates the total variation if the between subgroup variability is substantial. In all cases of practical interest, the estimates are larger than σ since it includes the between subgroup variability in the calculations. Thus, P_{pk} tends to be smaller and using P_{pk} rather than C_{pk} makes the process 'look worse'. For this reason, suppliers may be reluctant to use P_{pk} though it is beneficial to both parties to obtain a realistic view of the capability of the process to produce parts that are within specification limits.

If the process is stable, C_{pk} is approximately equal to P_{pk}, since a stable process has little between subgroup variability. If the process is in statistical control and there is nothing present but the inherent variation, C_{pk} is the best measure of long-term capability. If the process is out of statistical control, with some potential special cases and/or shift and drift present, P_{pk} is the best indicator of long-term capability.

The information that is derived from a well conceived capability analysis provides a solid baseline for process potential and opportunity. If C_p is equal to C_{pk}, then the process is mean-centred (no mean deviation issues).

If C_{pk} is negative, this is an indication that the mean lies outside one of the specification limits and over 50% of the distribution is outside the specifications. If both P_p and C_p are less than 1.33, the inherent variation is high and the process has inadequate capability. If P_{pk} is less than P_p and C_{pk} less than C_p, there is an indication of a mean deviation (targeting issue). Finally, if P_p is much less than C_p, there is mean variation between subgroups (shifts and drifts) present.

Because C_p is independent from the target value *T*, it is robust against departures of the process mean μ from *T* and this is its drawback.

One variation of C_{pk} is a relatively new index called C_pT, in which the T represents a target value. It allows one to select a target dimension and calculate capability from the target. C_pT calculations are the same as C_{pk} calculations, except that one substitutes a target dimension for the process average.

$$C_pT = \min\left(\frac{USL - T}{3\sigma}, \frac{T - LSL}{3\sigma}\right) = \frac{d - |T - M|}{3\sigma}. \qquad (1.3.7)$$

Like the C_{pk} index, both parts of the C_pT index are calculated, but only the minimum is used. The target dimension is usually the nominal of the specification, and some call it the true process centring of an index. In reality, however, the C_pT index is the same as the C_p index and has nothing to do with process centring. If the target T is set as the midpoint of the specification interval, i.e. $T = M$, C_pT yields the same ratio as C_p. C_pT, therefore, will not be discussed in any further detail.

The concept of variation has recently undergone a paradigm shift in the industry. This shift has occurred in the interpretation of the quality of product varying within the allowable process specification. All the indices discussed so far uses the historical perspective of variation. A historical perspective of variation is that product has the same quality; that is to say that the product is equally good, regardless of where it falls within the specification limits. Product is considered bad, lacking in quality, only if it falls outside of the specification limits. Engineers are comfortable with this notion of variation, which is sometimes referred to as "goal post mentality" and is displayed graphically in figure 1.6 below.

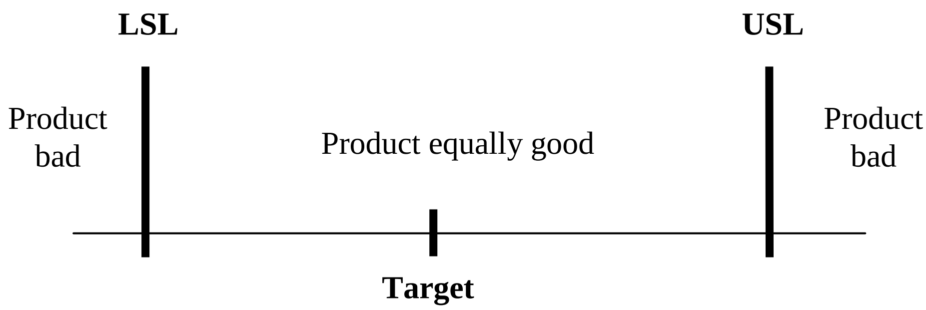


Figure 1.6:*Goal Post Mentality*

The problem with this mentality is the step function that occurs directly at the specification limits. A product is perfectly good up to some specification limit and completely bad beyond that limit. In regard to a process, the quality of a part falling just within the specification limit has little practical difference from the quality of a product falling just outside the specification limit. This model of quality variation has little relevance to industry.

Figure 1.7 below shows a model proposed by statisticians. This model is more practical in that the loss in quality and therefore value loss to an organisation increases as the quality varies from a process target.

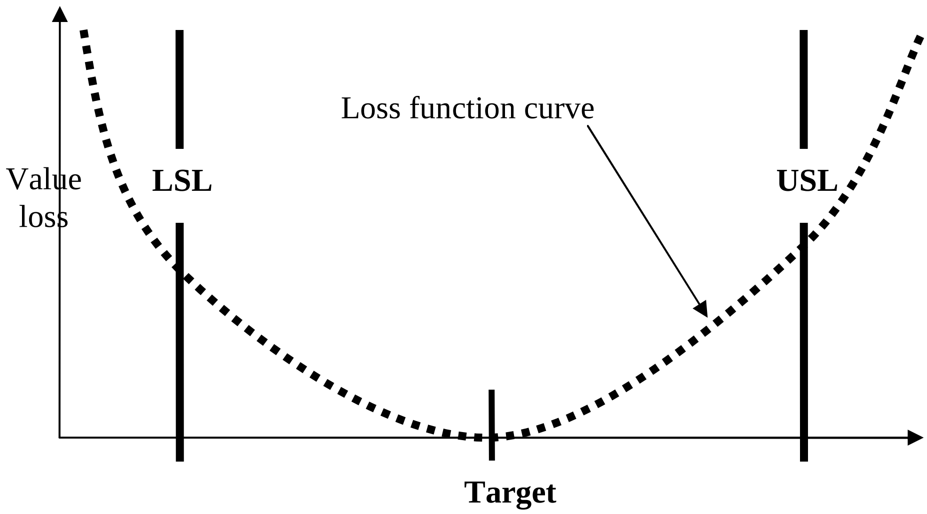


Figure 1.7: Loss function mentality

This notion of variation, referred to as "Loss function mentality", states that there is a quadratic relationship between the loss and the distance from the target, as proposed by Taguchi (Spiring *et a*, 2002). This function is called the loss function curve and it ties variation to the loss in a process. This notion is what capability is now based on. Capability indices enumerate a process' ability to minimise the loss function curve.

Hsiang and Taguchi (1985) (and also Chan, Cheng, and Spiring (1988)) developed the index C_{pm} in order to take into account the process target being defined as follows:

$$
\begin{aligned}
C_{pm} &= \frac{USL - LSL}{6\sqrt{\sigma^2 + (\mu - T)^2}} \qquad (1.3.8) \\
&= \frac{d}{3\sqrt{\sigma^2 + (\mu - T)^2}} \\
&= \frac{d}{3\sqrt{E(Y-T)^2}} \\
&= \frac{d}{3\sqrt{E(L(Y))}}
\end{aligned}
$$

where $L(Y)=(Y-T)^2$ is the loss function. $L(Y)$ is the loss associated with a characteristic Y not produced at the target. This implies the loss is zero when the process is on target and positive for any deviation from the target.

The expected loss becomes:

$$E(Y-T)^2 = E(Y-\mu+\mu-T)^2$$

$$= E\left((Y-\mu)+(\mu-T)\right)^2$$

$$= E(Y-\mu)^2 + E(\mu-T)^2$$ since $E(Y-\mu)=0$ making cross product is zero

$$= \sigma^2 + (\mu-T)^2,$$

which is the denominator in equation (1.3.8).

Notice that:

$$C_{pm} = \frac{\dfrac{USL-LSL}{6\sigma^2}}{\sqrt{1+\left(\dfrac{\mu-T}{\sigma}\right)^2}}$$

$$i.e. C_{pm} = \frac{C_p}{\sqrt{1+\left(\dfrac{\mu-T}{\sigma}\right)^2}}. \qquad (1.3.9)$$

If $\mu = T$ (process is on target) then $C_{pm} = C_p$.

The C_{pm} index is similar to the C_p index in calculation, except that the standard deviation is defined as $\sqrt{E(Y-T)^2}$ instead of $\sqrt{E(Y-\mu)^2}$. The target dimension is substituted for the process mean in the formula for the standard deviation.

The index C_{pm} does not directly relate to the percentage of non-conforming product, p. If p is regarded as the most important quality aspect of the process, this is definitely the wrong capability index to use.

Boyles (1991) showed that for fixed μ, the index C_{pm} is bounded from above when σ tends to 0 and furthermore, that $C_{pm} < \frac{d}{(3|\mu - T|)}$ and hence $|\mu - T| < \frac{d}{3C_{pm}}$.

This inequality can be interpreted as, a C_{pm}-value of 1 which implies that the process mean, μ, lies within the middle third of the specification interval, and in general, it lies within the middle $1/3C_{pm}$ of the specification interval i.e. $d/3C_{pm}$. Therefore, given a C_{pm} index of 1.00, we know that $M - \frac{d}{3} < \mu < M + \frac{d}{3}$. This interval is much smaller than the one for C_{pk} equals to 1.00, which is equal to $M - d < \mu < M + d$.

Parlar and Wesolowsky (1999) notes that if $T = \mu$, then the three basic PCIs, C_{pk}, C_p, C_{pm}, are connected by the relationship

$$C_{pk} = C_p - \frac{1}{3}\sqrt{\left(\frac{C_p}{C_{pm}}\right)^2 - 1}.$$

Whereas the index C_{pm} has the attractive features that it incorporates the parameters d, μ, T, and σ, it has an important omission, namely the parameter M. The index C_{pmk} rectifies this deficiency. To devise an index that is more sensitive to departures of μ from T, Pearn *et al.* (1992) introduced another process capability index, C_{pmk}. The index takes its numerator from C_{pk} and its denominator from C_{pm}, hence it is a hybrid.

$$C_{pmk} = \frac{\min(USL - \mu, \mu - LSL)}{3\sqrt{\sigma^2 + (\mu - T)^2}} \qquad (1.3.10)$$

$$C_{pmk} = \frac{d - |\mu - M|}{3\sqrt{\sigma^2 + (\mu - T)^2}}. \qquad (1.3.11)$$

When μ is equal to M, C_{pmk} is equal to C_{pm}.

Also notice that

$$C_{pmk} = \frac{\min(\frac{USL-\mu}{3\sigma}, \frac{\mu - LSL}{3\sigma})}{\sqrt{1+\left(\frac{\mu - T}{\sigma}\right)^2}}$$

$$C_{pmk} = \frac{C_{pk}}{\sqrt{1+\left(\frac{\mu - T}{\sigma}\right)^2}}$$

and when μ is equal to T, C_{pmk} is equal to C_{pk}.

When $T = M$ and for fixed μ, the index C_{pmk} is bounded from above when σ tends to 0 and that $C_{pmk} < \frac{d}{(3|\mu - T|)} - \frac{1}{3}$ or $|\mu - T| < \frac{d}{(1+3C_{pmk})}$.

This inequality can be interpreted as a C_{pmk}-value of 1 implies that the process mean, μ, lies within the middle fourth of the specification range i.e. $M - \frac{d}{4} < \mu < M + \frac{d}{4}$. In general, the process mean, μ, lies within the middle $\frac{1}{(1+3C_{pmk})}$ of the specification range i.e. $\frac{d}{(1+3C_{pmk})}$, when $T = M$.

C_{pmk} is certainly worse than C_{pk} for being associated with a certain percentage of non-conforming product, but again, one should not choose this index if p is the main interest. C_{pmk} (and usually C_{pm}) is much more sensitive than other capability indices to movements in the process average relative to T. As seen in (1.3.10), when μ is equal to T, C_{pmk} is equal to C_{pk}. If μ moves away from T, however, C_{pmk} decreases more rapidly than does C_{pk}, although both are zero when μ equals one of the specification limits. Conversely, when μ is brought closer to M, C_{pmk} increases much faster than does C_{pk}. C_{pmk} reveals the most information about the location of the process mean and the least about the proportion non-conforming p.

Vannman (1995) shows that among all the indices presented thus far, C_{pmk} is the most sensitive to departures of μ from T. The ranking of the following four basic indices discussed thus far in terms of sensitivity to departure of the process mean from the target value, from the most sensitive to the least sensitive are (1) C_{pmk}, (2) C_{pm}, (3) C_{pk} and (4) C_p.

A further interesting relationship among the indices given in Kotz and Johnson (2002) is derived as follows:

Since

$$C_{pmk} = \frac{C_{pk}}{\sqrt{1+\left(\frac{\mu-T}{\sigma}\right)^2}}$$

$$= \frac{\frac{USL-LSL}{6\sigma^2}}{\sqrt{1+\left(\frac{\mu-T}{\sigma}\right)^2}} \frac{C_{pk}}{\frac{USL-LSL}{6\sigma^2}}$$

$$\therefore C_{pmk} = C_{pm}\frac{C_{pk}}{C_p} = \frac{C_{pm}C_{pk}}{C_p}.$$

Another hybrid index is $C_{pm}^{\#}$, proposed by Chan, Cheng and Spiring (1988). The index takes its numerator from C_pT and its denominator from C_{pm}.

$$C_{pm}^{\#} = \frac{\min(USL-T, T-LSL)}{3\sqrt{\sigma^2+(\mu-T)^2}} \tag{1.3.12}$$

$$C_{pm}^{\#} = \frac{d-|T-M|}{3\sqrt{\sigma^2+(\mu-T)^2}} \quad .$$

When T is equal to M, $C_{pm}^{\#}$ is equal to C_{pm}.

Notice also that

$$C_{pm}^{\#} = \frac{\min(\frac{USL-T}{3\sigma}, \frac{T-LSL}{3\sigma})}{\sqrt{1+\left(\frac{\mu-T}{\sigma}\right)^2}}$$

$$C_{pm}^{\#} = \frac{C_p T}{\sqrt{1+\left(\frac{\mu-T}{\sigma}\right)^2}}.$$

When μ is equal to T, $C_{pm}^{\#}$ is equal to $C_p T$.

1.4 THE UNIFIED APPROACH

The unified approach is proposed by Vannman (1995). Vannman constructs a superstructure class to include the four basic indices, C_p, C_{pk}, C_{pm} and C_{pmk} as special cases. By varying the parameters of this class, we can find indices with different desirable properties. The proposed new indices depend on two non-negative parameters, u and v, as

$$C_p(u,v) = \frac{d-u|\mu-M|}{3\sqrt{\sigma^2+v(\mu-T)^2}}. \qquad (1.4.1)$$

It is easy to verify that:

$C_p(0,0) = C_p$; $C_p(1,0) = C_{pk}$; $C_p(0,1) = C_{pm}$; $C_p(1,1) = C_{pmk}$.

From the study of $C_p(u,v)$, large values of u and v will make the index $C_p(u,v)$ more sensitive to departures from the target value. A slight modification gives the even more general index class which includes $C_{pm}^{\#}$ as a special case as well.

$$C_p(u_1,u_2,v) = \frac{d-u_1|\mu-M|-u_2|T-M|}{3\sqrt{\sigma^2+v(\mu-T)^2}} \qquad (1.4.2)$$

$C_p(0,1,1) = C_{pm}^{\#}$.

The five C_p, C_{pk}, C_{pm}, C_{pmk} and $C^{\#}_{pm}$, are equal when $\mu = T = M$, but differ in behaviour when $\mu \neq T$.

1.5 THE NORMATIVE APPROACH

It is important to note that the process capability indices discussed thus far plays only a passive role in the manufacturing sciences. The functions of assessing and monitoring are not predictive, nor are they proactive, and thus the available process capability indices mainly serve as policing devices. Whereas this by itself is a necessary activity, a much more useful role can be served by the process capability indices if they can also be used to predict and to control the quality of the future output (Singpurwalla, 1998).

The normative approach for the control of quality is based on decision-theoretic considerations. It provides a vehicle for accomplishing both the retroactive function of assessment and monitoring as well as the proactive function of prediction and control. Furthermore, the normative approach is able to integrate the three tasks of assessment, prediction and control within an interactive and unifying framework. Here, one monitors the observable Y (rather than the unobservable μ), and makes a decision to continue production, to modulate it or to stop it, based on the consequences of the deviation of Y from T. The decision is proactive and is dictated by the predictive distribution of Y and the utilities associated with a control of the process.

The work of Bernardo and Telba (1996) appears to be first to have introduced the normative approach in the context of process capability indices (Singpurwalla, 1998).

1.5.1 BAYES CAPABILITY INDEX

In the manufacturing industry, process capability analysis is used to flag high values of the proportion of non-conforming parts in order to prevent further production of unacceptable output. The analysis assumes existence of engineering specifications, that the process is normally distributed and that the process is in statistical control.

However, the abundance of outputs from skewed distributions and the censoring effect induced by the finite precision of actual measures often makes rather unreasonable the normality assumption on which traditional capability indices are intuitively based. Moreover, the sampling distributions of the estimators of the capability indices are often intricate, even under normality assumption (Singpurwalla, 1998). Consequently, point estimators of the capability indices, with no reference to their precision, are usually quoted. This is misleading practice, for even large samples may produce rather unreliable estimators.

A Bayesian index is proposed to evaluate process capability which, within a decision-theoretical framework, directly assesses the proportion of future parts which may be expected to lie outside the tolerance limits. This results in a new general capability index which:

i. Has a solid decision-theoretical foundation
ii. Does not require the process to be normal
iii. May be used for multivariate observations
iv. May accommodate measurements with error
v. Contains the conventional index as a limiting case.

The proposed capability index is a direct function of the data, whose value is sufficient to solve the relevant decision problem.

The Bayes capability index $C_B(D)$ (Bernardo and Irony, 1996) is given by:

$$C_B(D) = \frac{1}{v}\Phi^{-1}\left\{\Pr\left(y \in A \mid D\right)\right\} \qquad (1.5.1)$$

where v will be set equal to 3 or 6 and A is the tolerance region, Φ is the distribution function of the standard normal distribution, and D the available data.

Accept that the process is capable if and only if:

$$C_B(D) \geq c_0 \qquad (1.5.2)$$

where c_0 is a threshold value.

The precise relationship between the general Bayesian index $C_B(D)$ and the traditional (conventional) index C_p and C_{pk} is now derived below (see Singpurwalla (1998) for a similar analysis).

If $\{\Pr(y \in A \mid D)\}$ is set greater or equal to p_0, *i.e.*:

$$\{\Pr(y \in A \mid D)\} \geq p_0, \tag{1.5.3}$$

then a strategy would then be to take a monotonic transformation of the terms on both sides of inequality (1.5.3). A possibility, and the one proposed by Bernardo and Irony (1996), is the probit transformation:

$$\Phi^{-1}\{\Pr(y \in A \mid D)\} \geq \Phi^{-1}(p_0).$$

Dividing both sides by v gives the Bayes capability index as defined by Bernardo and Irony (1996), since:

$$\frac{1}{v}\Phi^{-1}\{\Pr(y \in A \mid D)\} \geq \frac{1}{v}\Phi^{-1}(p_0)$$

implies that

$$C_B(D) \geq c_0$$

with

$$c_0 = \frac{1}{v}\Phi^{-1}(p_0).$$

Instead of the probit transformation one may also consider the log-odds transformation of the type $\log\frac{p}{1-p}$. However, if the predictive distribution of Y is also a Gaussian with mean μ and variance σ^2 - which can occur under some very general conditions involving n large - then

$$\left\{\Pr\left(x \in A \mid D\right)\right\} = \Phi\left(\frac{USL-\mu}{\sigma}\right) - \Phi\left(\frac{LSL-\mu}{\sigma}\right)$$

and now in the case probit transformation the above reduces statement to:

$$\Phi^{-1}\left\{\Phi\left(\frac{USL-\mu}{\sigma}\right) - \Phi\left(\frac{LSL-\mu}{\sigma}\right)\right\}.$$

Dividing by v, the Bayes capability index now states the process is capable if:

$$\frac{1}{v}\Phi^{-1}\left\{\Phi\left(\frac{USL-\mu}{\sigma}\right) - \Phi\left(\frac{LSL-\mu}{\sigma}\right)\right\} \geq c_0 \qquad (1.5.4)$$

$$\frac{1}{v}\left\{\left(\frac{USL-\mu}{\sigma}\right) - \left(\frac{LSL-\mu}{\sigma}\right)\right\} \geq c_0$$

$$\frac{USL-LSL}{v\sigma} \geq c_0,$$

and with $v = 6$ the Bayes capability index now states the process is capable if:

$$\frac{USL-LSL}{6\sigma} \geq c_0 \qquad (1.5.5)$$

$$C_p \geq c_0.$$

The left hand side of (1.5.5) is precisely C_p, the PCI introduced by Juran (1974).

Going back to (1.5.4)

Assume $v = 3$, (1.5.4) results in

$$\frac{1}{3}\Phi^{-1}\left\{\Phi\left(\frac{USL-\mu}{\sigma}\right) - \Phi\left(\frac{LSL-\mu}{\sigma}\right)\right\} \geq c_0. \qquad (1.5.6)$$

Suppose now that μ is not centred at M, but is in the vicinity of LSL in such a way that

$$\left(USL-\mu\right)\gg\left(\mu-LSL\right) \text{ and again} \left(USL-\mu\right)\gg\sigma.$$

That is, the process is said to be non-centred but potentially capable, then,

$\Phi\left(\dfrac{USL-\mu}{\sigma}\right)\approx 1$, and inequality (1.5.6) becomes

$$\frac{1}{3}\Phi^{-1}\left\{1-\Phi\left(\frac{LSL-\mu}{\sigma}\right)\right\}\geq c_0$$

$$\frac{1}{3}\Phi^{-1}\left\{\Phi\left(\frac{\mu-LSL}{\sigma}\right)\right\}\geq c_0$$

$$\frac{\mu-LSL}{3\sigma}\geq c_0 \tag{1.5.7}$$

$$C_{pl}\geq c_0 .$$

Conversely, suppose now that μ is not centred at M, but is in the vicinity of USL in such a way that

$$\left(\mu-LSL\right)\gg\left(USL-\mu\right) \text{ and again} \left(\mu-LSL\right)\gg\sigma.$$

That is, the process is said to be non-centred but potentially capable, then,

$\Phi\left(\dfrac{\mu-LSL}{\sigma}\right)\approx 0$, and inequality (1.5.6) becomes

$$\frac{1}{3}\Phi^{-1}\left\{\Phi\left(\frac{USL-\mu}{\sigma}\right)-0\right\}\geq c_0$$

$$\frac{1}{3}\Phi^{-1}\left\{\Phi\left(\frac{USL-\mu}{\sigma}\right)\right\}\geq c_0$$

$$\frac{USL-\mu}{3\sigma}\geq c_0 \tag{1.5.8}$$

$$C_{pu}\geq c_0 .$$

The indices C_{pk} and C_{pm} would lead to decisions similar to those of C_{pmk}, but the decisions based on the index C_p would be closer in tune with those based on the Bayes capability index. This point is discussed in detail in chapter 3. Why this disparity of decisions based on the index C_{pmk} and the Bayes capability index? The answer lies in the fact that the index C_{pmk} inflicts penalties whenever the sample mean of a rating period deviates from the target, T, and also from the mid-point, M, whereas the Bayes index penalises whenever a unit fails to belong to its specification limits.

One advantage of the Bayes index is that it automatically takes into account the possibility of having non-conforming units in both sides of the interval defined by the specification limits, whereas C_{pk} will only consider that side in which μ is closer to the specification limits.

1.6 EXPECTED PROPORTION NON-CONFORMING

The capability indices discussed thus far are designed to measure the process capability when the studied characteristic is normal. In such a case, the index C_p can be interpreted using the probability of non-conformance, that is, the probability of obtaining a value that is outside the specification limits. Elementary probability arguments show that the probability of non-conformance:

$$1-P(LSL<Y<USL)=1-P(\frac{LSL-\mu}{\sigma}\leq\frac{Y-\mu}{\sigma}\leq\frac{USL-\mu}{\sigma})$$

$$= 1\text{-}(\Phi\left(3C_{pu}\right)\text{-}\Phi\left(-3C_{pl}\right))$$

$$= (\Phi\left(-3C_{pu}\right)+\Phi\left(-3C_{pl}\right))$$

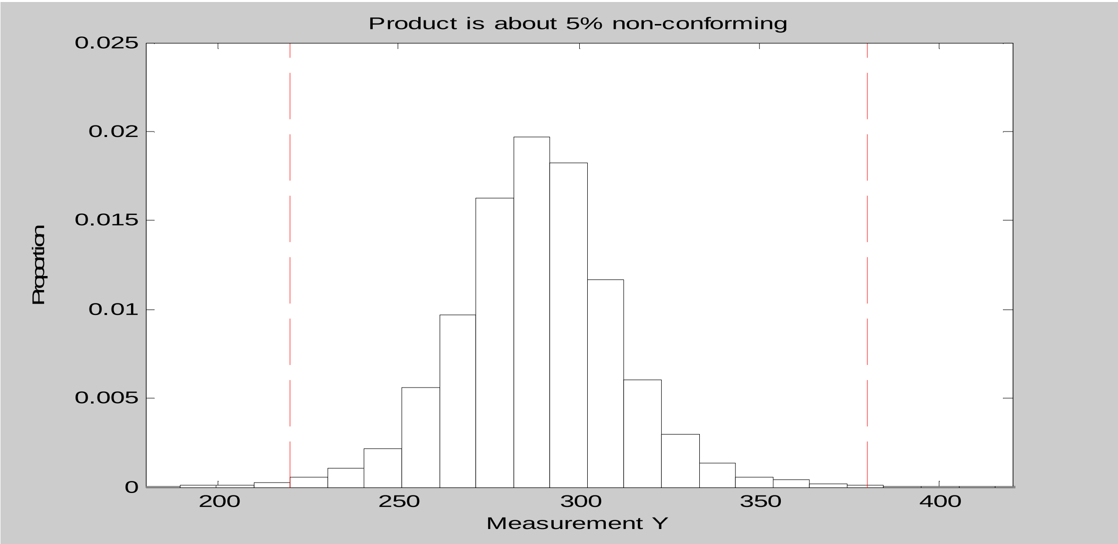
$$= 2\Phi\left(-3C_{p}\right) \text{ if } (\mu=M) \text{ (centred state),}$$

where Φ denotes the standard normal cumulative distribution function. If C_p=1 then USL=$\mu+3\sigma$ LSL=$\mu-3\sigma$, the expected proportion non-conforming product when the process is centered is 0.27%, which is regarded as 'acceptably small'.

The value of C_{pk} does not determine the probability of non-conformance, but limits it, and the probability of non-conformance is never more than $2\Phi\left(-3C_{pk}\right)$. The corresponding is true for both C_{pm} and C_{pmk}, and the probability of non-conformance is never more than $2\Phi\left(-3C_{pm}\right)$ and $2\Phi\left(-3C_{pmk}\right)$ respectively.

Several authors recommend reporting just the percentage of non-conforming product, say p, as an indication of capability. Although p provides a simple and concise summary of process quality that is easy for customers and top-level managers to understand, it is of limited use for practitioners who are working to improve the process as illustrated in the figures below.

Given that p is 5 percent for a process, how would one improve quality?

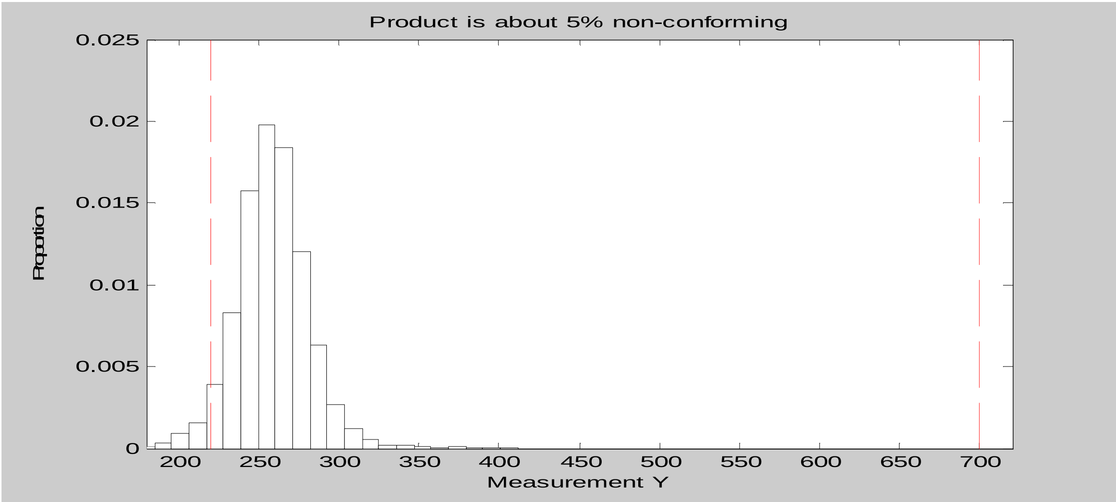


LSL=220 and USL=380

Figure 1.8: A process which can be improved by reducing process standard deviation

If 2.5 percent of the non-conforming parts are below the LSL and the other 2.5 percent are above the USL, the process standard deviation must be reduced (see figure 1.8). This is much harder to achieve without changing processes or machinery. Using superscripts ℓ_0, ℓ_1 to denote the p-values applicable to non-conforming by reason of

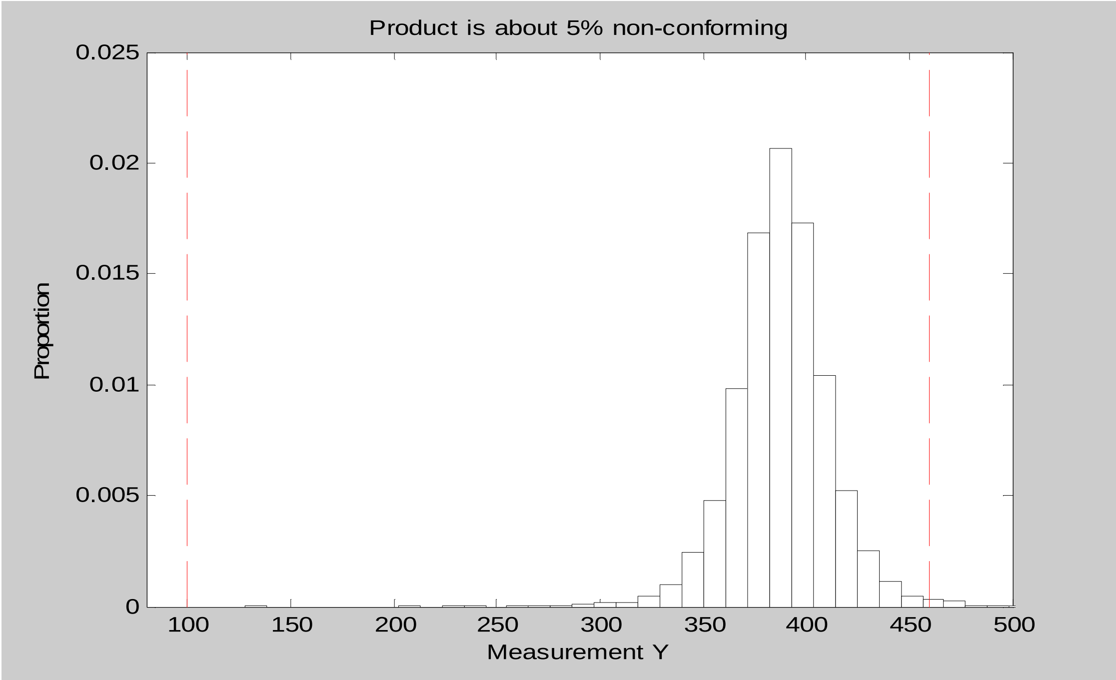
Y being less than ℓ_0, or greater than ℓ_1, both p^{ℓ_0} and p^{ℓ_1} being “large”, improvement efforts must focus on reducing σ.

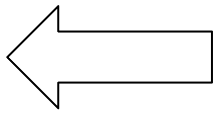


LSL=220 and USL=700

Figure 1.9: A process which can be improved by raising the process average with standard deviation unchanged

If all 5 percent of the non-conforming parts are below the LSL, the process average must be raised (see figure 1.9). If p^{ℓ_0} is greater than p^{ℓ_1}, the practitioner knows that μ should be shifted higher.





LSL=100 and USL=460

Figure 1.10: A process which can be improved by lowering the process average with standard deviation unchanged

If the non-conforming parts are all above the USL, then efforts must be made to lower the mean of the production with the standard deviation unchanged (see figure 1.10). If p^{ℓ_0} is less than p^{ℓ_1}, the practitioner knows μ should be shifted lower.

Based on just p, which course of action should one pursue to improve process performance? The capability index measured by just p provides few answers to these vital questions.

Yeh and Bhattacharya (1998) proposed the use of a PCI based on the ratios of expected proportion non-conforming to actual observed or estimated proportion non-conforming. In itself this is simply the ratio

$$\frac{p_0}{p},$$

where p_0 is the desired proportion of non-conforming output and p is the actual proportion. This PCI is simply estimated by

$$\frac{p_0}{\hat{p}}.$$

Another PCI, suggested by Yeh and Bhattacharya (1998), distinguishes between non-conforming items for which Y is less than ℓ_0, and those for which Y are greater than ℓ_1. They suggest using the PCI index as

$$\min\left(\frac{p_0{}^{\ell_0}}{p^{\ell_0}}, \frac{p_0{}^{\ell_1}}{p^{\ell_1}}\right)$$

Using this latest notation, $p = p^{\ell_0} + p^{\ell_1}$.

1.7 ESTIMATION OF THE INDICES

The most prevalent methods of estimating the basic PCIs are to replace μ by

$\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$ and σ by $S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \overline{Y})^2}$,

where $Y_1, Y_2, \cdots, Y_n$ are independent values of Y or by an appropriate multiple of the range $(\mathrm{R} = \mathrm{maximum}(Y) - \mathrm{minimum}(Y))$. Very often, estimation is based on values from a series of samples, combining the individual sample estimates into a single estimate.

The assumptions are that Y has a $N\left(\mu, \sigma^2\right)$ distribution, hence $\overline{Y}$ has a $N\left(\mu, \frac{\sigma^2}{n}\right)$ distribution and S^2 has a $\frac{\sigma^2 \chi^2}{(n-1)}$ distribution, and $\overline{Y}$ and S are mutually independent. The estimates of the capability indices become:

$\widehat{C}_{p}, \widehat{C}_{pk}, \widehat{C}_{pl}, \widehat{C}_{pu}, \widehat{P}_{p}, \widehat{P}_{pk}, \widehat{C}_{p}T, \widehat{C}_{pm}, \widehat{C}_{pmk}, \widehat{C^{\#}}_{pm}$.

Traditionally, practitioners estimate

C_p and C_{pk} with $\widehat{\sigma}_{\overline{R}}$ (AIAG (1995)). For example,

$$\widehat{C}_p = \frac{USL - LSL}{6\widehat{\sigma}_{\overline{R}}} \text{ where } \widehat{\sigma}_{\overline{R}} = \frac{\overline{R}}{d_2}.$$

$\widehat{\sigma}_{\overline{R}}$ is a popular measure of within–subgroup estimate for the process standard deviation in the manufacturing industry (see discussion papers by Bothe (2002), Rodriguez (2002), Lu and Rudy, (2002)), where $\overline{R}$ is the average range for a set of consecutive subgroup samples and d_2 is the control chart constant used to determine control limits for a range chart. It is a constant to achieve an unbiased estimate of σ.

When the indices are computed using s as an estimate of σ based on individual measurements, they are denoted as $\widehat{P}_p$ and $\widehat{P}_{pk}$.

For example,

$$\widehat{P_p} = \frac{USL - LSL}{6s}.$$

$\widehat{P}_p$ and $\widehat{P}_{pk}$ are designated as estimators of “process performance indices”.

If a process is stable, P_p and P_{pk} are approximately equal to C_p and C_{pk} respectively, since a stable process has little between subgroup variability. If the process is in statistical control and there is nothing present but the inherent variation, there is therefore no distinction between P_p and C_p on the one hand and P_{pk} and C_{pk} on the other and the two estimates of the indices would yield the same result.

1.8 ORGANISATION OF THE THESIS

Process capability analysis alludes to the possibility of applying the Bayesian simulation approach to process capability indices such as C_p, C_{pk}, P_p and P_{pk}. By putting a prior on μ and σ^2. Simulation of most of the indices covered in this introductory chapter is dealt in chapter 2.

Bayesian simulation procedures are used to solve the supplier selection problem in chapter 3.

In chapter 4, Bayesian simulation estimation of the single variance component index $C_{pl,}$ including estimation of the distribution of the index from the first four moments about the origin and other methods are discussed in depth. Differences between two such indices are also discussed.

Chapter 5 introduces a process capability index for averages of observations from new or unknown batches in the case of the balanced random effects model with two variance components.

Chapter 6 extends the capability index in chapter 5 for the balanced random effects model to three variance components.

Chapter 7 summaries, concludes and suggests areas for further research.

Appendix A1

Proof of theorem 1.3.1

Let $R\left\{y:|y-\mu|>c\right\}$, then

$$P(|Y-\mu|>c)=\int_R f(y)dy$$

$if \quad y\in R,$

$$\frac{|y-\mu|^2}{c^2}\geq 1.$$

$Thus,$

$$\int_R f(y)dy \leq \int_R \frac{|y-u|^2}{c^2} f(y)dy \leq \int_{-\infty}^{\infty} \frac{|y-\mu|^2}{c^2} f(y)dy = \frac{\sigma^2}{c^2}$$

$Hence$

$$P(|Y-\mu|>c)\leq \frac{\sigma^2}{c^2}.$$

CHAPTER 2

BAYESIAN SIMULATION IN PROCESS CAPABILITY ANALYIS

2.1 INTRODUCTION

Process capability analysis alludes to the possibility of applying Bayesian simulation procedures to these indices since a prior distribution can easily be attached to the mean and variance of a process. This chapter demonstrates how Bayesian methods for making inferences about the proportion of non-conforming units in a quality control setting can be implemented using simulation techniques. Random draws from the posterior distribution of the quantities of interest are used to construct the needed inferences. Although Bayesian simulation will be discussed in all subsequent chapters, it is formally discussed in this chapter. The theory behind a Bayesian simulation technique is discussed and a complete algorithm for doing the simulation is presented for each of the capability indices: C_p, C_{pl}, C_{pu}, , C_{pk}, P_p, P_{pk}, C_pT, C_{pm}, C_{pmk} and $C_{pm}{}^{\#}$. The methods are illustrated using the aircraft data collected by the Pratt and Whitney Company (Niverthi and Dey, 2000). The Bayesian simulation results will be compared with the frequentist classical methods.

Histograms of the simulated data obtained from the example will be illustrated and where possible compared with the theoretical curve plot. The methods can be generalised to more complicated situations which are illustrated by discussing the extensions to two and three variance component models in chapters 5 and 6 respectively.

Bayesian data analysis essentially involves:

(1) A parametric model and the assignment of probability distributions to all parameters. Setting up a probability model, which suggests a distribution for observables (measurements or attributes) conditional on unobservables (the

parameter $\underline{\theta}$ vector) where a function, $t(\underline{\theta})$ say, of some or all of the parameters and observables is the object of inference.

(2) A prior distribution $p(\underline{\theta})$ summarises a priori uncertainty about the likely values of the parameters. The prior distribution for $\underline{\theta}$ needs to be formulated based on prior knowledge. This is usually a difficult task because such prior knowledge may not be available. In such situations, usually a "*non-informative*" prior distribution is used. The basic idea behind formulating such a vague prior distribution is that it should be flat over all $\underline{\theta}$ such that the likelihood (the density $p(\underline{Y}\,|\,\underline{\theta})$, evaluated at the observed value of $\underline{Y}$) plays a dominant role in the construction of the posterior density. Jeffreys (1961) formulates such prior distribution based on certain invariance arguments.

(3) Upon observing data, inference is based on the posterior density of the parameters, given the data. The recent increase in the availability of computational resources and the development of computational techniques has led to great advances in the application methods to complicated problems in various disciplines. The posterior is computed in unnormalised form using Bayes theorem by product of the likelihood function and the prior density of the parameters. However, explicitly suppose that $p(\underline{Y}\,|\,\underline{\theta})$ is the density function of the measurements $\underline{Y}$, conditional on the parameter $\underline{\theta}$ vector, then the posterior density function of $\underline{\theta}$ is given by:

$$p(\underline{\theta}\,|\,\underline{Y}) = \frac{p(\underline{Y}\,|\,\underline{\theta})\,p(\underline{\theta})}{\int p(\underline{Y}\,|\,\underline{\theta})\,p(\underline{\theta})\,d\underline{\theta}}. \qquad (2.1.1)$$

Draws are simulated from the above posterior distribution by drawing $\underline{\theta}$ from $p(\underline{\theta}\,|\,\underline{Y})$.

The above equation is the essential ingredient of inference about the quantity of interest $t(\underline{\theta})$, a known function of $\underline{\theta}$, such as a process capability index.

2.2 ADVANTAGES AND DISADVANTAGES OF THE BAYESIAN APPROACH

Some of the advantages of the Bayesian approach are:

(i) The Bayesian practitioner does not need to commit himself/herself only to a point estimate of the parameters. Credibility intervals can easily be obtained.

(ii) Uncertainty about the true values of the parameters is formally incorporated into the analysis through the choice of the appropriate prior distribution.

(iii) Given the data, the prior information about the unknown parameters and a well defined loss function, there exits an optimal Bayes predictor.

(iv) All the available information about the quantity of interest to be predicted is contained in the posterior distribution. The practitioner can, therefore, base all of the inferences on this distribution.

(v) The Bayesian approach is conceptually more appealing than the classical approach.

Critics of the Bayesian approach have most often cited the following points:

(i) The Bayesian practitioner must formally express his/her prior beliefs about the unknown parameters in the form of a probability distribution.

(ii) The Bayesian methodology is computer intensive. In many situations, integrations in several dimensions are required to obtain the required posterior distributions.

These might have been valid criticisms in the past but by using (a) non-informative priors like the Jeffreys prior, probability matching and reference priors, and (b) numerical integration techniques like Markov Chain Monte Carlo methods and more specifically Gibbs sampling, these problems can be overcome.

In Monte Carlo simulation, random draws from the posterior distribution of the quantities of interest are used to construct the needed inferences. Histograms of the simulations can be constructed. This is precisely the advantage of the sampling based approach, where one can create the posterior distributions (realisations of the parameter in the form of a histogram) based on the samples and hence can do inferences from the posterior distribution without going through the exact distribution. From distributions of the capability indices one is in a position to obtain quantiles, credible regions and to perform other inferential tasks. The methods can be generalised to more complicated situations. This however requires computational resources.

2.3 THE BAYES STRUCTURE FOR NORMAL DISTRIBUTION WITH BOTH PARAMETERS, MEAN AND VARIANCE, UNKNOWN

Denote the likelihood function by $l(\underline{Y}\,|\,\underline{\theta})$. A class Ω of prior distributions is said to form a conjugate family if the posterior density $p(\underline{\theta}\,|\,\underline{Y}) \propto p(\underline{\theta})l(\underline{Y}\,|\,\underline{\theta})$ is in the class Ω for all $\underline{Y}$ whenever the prior density is in Ω. Conjugate priors are often very precise and represent a very strong belief about the value of some unknown parameter and this belief ends up dominating the likelihood.

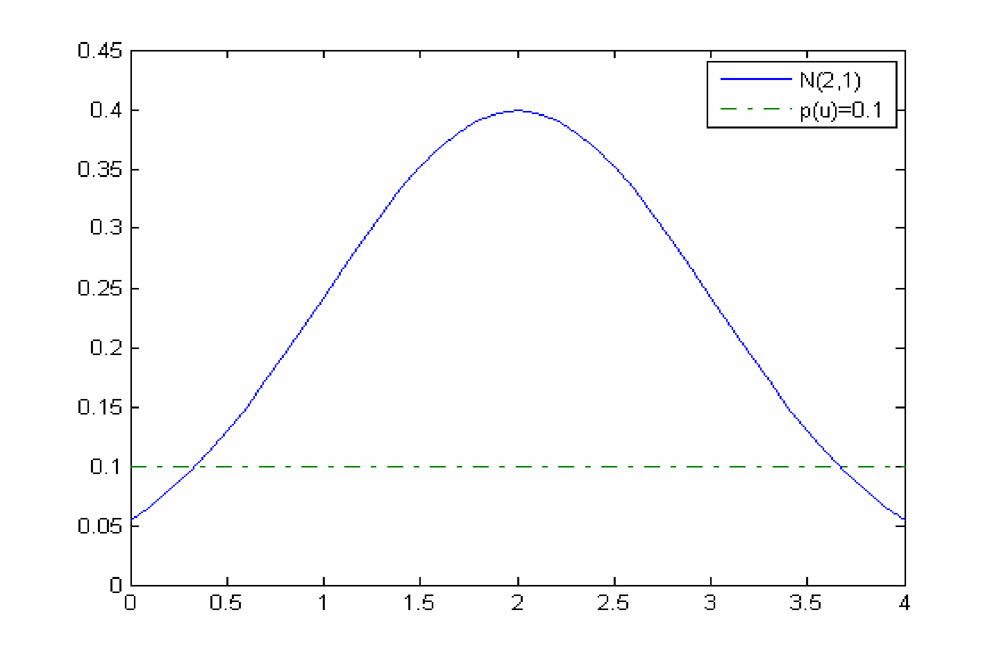
It is often sensible to analyse scientific data on the assumption that the likelihood dominates the prior. Two important reasons for this are as follows.

i. Even if two people both have strong prior beliefs about the value of some unknown quantity, they might not agree, and it seems sensible to use a neutral prior which is dominated by the likelihood. This kind of prior could be said to represent the views of someone who had no strong beliefs a priori.
ii. In many scientific contexts one thinks that the experiment will increase one's knowledge significantly, and if that is the case then presumably the likelihood will dominate the prior.

A vague prior distribution for the normal parameter, mean μ, is taken as $p(\mu) \propto c, -\infty < \mu < \infty,$ a constant. This density can be regarded as representing a normal density of infinite variance. It should be noted that this is not a proper density function. Nevertheless, one often finds it useful to extend the concept of a probability density function to some cases like this which are called improper 'densities'. It turns out that sometimes when one takes an improper prior density, it can combine with an ordinary likelihood to give a posterior which is proper.

Suppose that $p(\mu) \propto c, -\infty < \mu < \infty$ is the improper or non-informative prior distribution and an informative prior is represented by the conjugate prior μ~N(μ_0 ,σ_0^2) with μ_0=2 and $\sigma_0^2 = 1$. The next figure is a plot of a conjugate and an improper prior. For
illustration purposes c will be taken as 0.1.

$p(\mu)$



μ

Figure 2.1: A plot of a conjugate and an improper prior for the mean

If a prior distribution is quite flat, then the prior distribution has little influence on the posterior. If one does not have strong prior opinions as required by the conjugate prior, a vague or non-informative prior is the best alternative. Under such circumstances, one's posterior opinion is mainly determined by the data one observes.

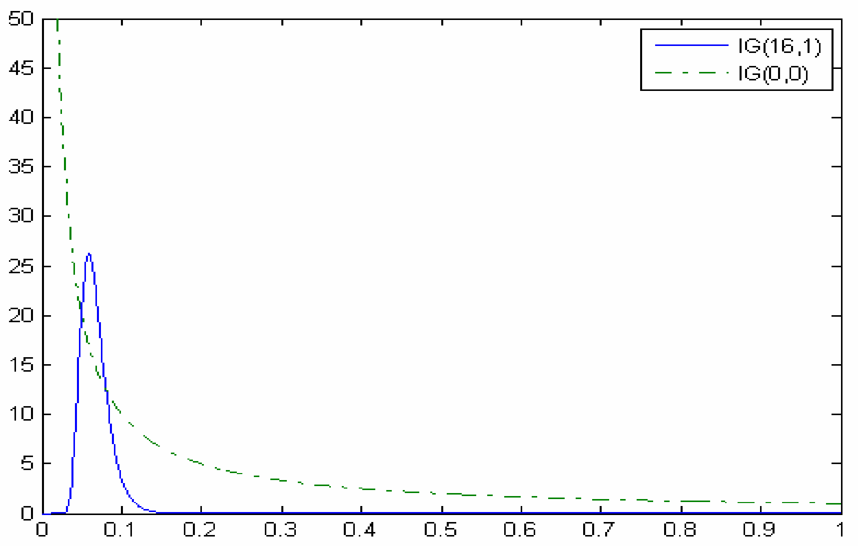
Suppose the informative prior for a parameter is represented by the conjugate prior as $\phi \sim IG(\alpha, \beta)$ i.e. an inverse gamma distribution with hyper-parameters $\alpha, \beta > 0$, where ϕ denotes the variance parameter. The vague prior used for ϕ is a particular case of the prior $IG(\alpha,\beta)$ in which $\alpha=0$ and $\beta=0$. Thus

$$p_Y(\phi)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}\phi^{-(\alpha+1)}e^{-\beta/\phi}\text{, where }\phi>0\,.$$

$$p(\phi)\propto\phi^{-(\alpha+1)}e^{-\beta/\phi}=\phi^{-(0+1)}e^{-0/\phi}\propto\frac{1}{\phi}.$$

The next figure is a plot of a conjugate and vague prior for ϕ .

$p(\phi)$



ϕ

Figure 2.2: A plot of a conjugate and an improper prior for ϕ

Theorem 2.3.1

If X has gamma distribution with parameters $\alpha>0$ and $\beta>0$, denoted as $X\sim G(\alpha,\beta)$ and density function

$$p_X(x)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}x^{(\alpha-1)}e^{-\beta x},\quad x>0\,,$$

then $Y=\dfrac{1}{X}$ has an inverse gamma distribution denoted by

$Y\sim IG(\alpha,\beta)$

The density function of Y is

$$p_Y(y) = \frac{\beta^\alpha}{\Gamma(\alpha)} y^{-(\alpha+1)} e^{-\beta/y}, \quad y > 0.$$

Proof

The proof is given in Appendix A2.

Remarks

1. If $X \sim IG(\alpha,\beta)$, then $Y = \frac{1}{X} \sim IG(\alpha,\beta)$.

2. A gamma density with $\alpha = \frac{n}{2}$ and $\beta = \frac{1}{2}$ is a Chi-square density since

$$p_X(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{(\alpha-1)} e^{-\beta x} = \frac{\left(\frac{1}{2}\right)^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)} x^{(\frac{n}{2}-1)} e^{-\left(\frac{1}{2}\right)x} = \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} x^{(\frac{n}{2})-1} e^{-\left(\frac{x}{2}\right)}$$ which is the Chi-square density with n degrees of freedom.

Theorem 2.3.2

If $X \sim IG(\alpha,\beta)$, then $\frac{2\beta}{X} \sim \chi^2_{2\alpha}$.

Proof

The proof is given in Appendix A2.

The Bayes structure for the normal distribution with both parameters μ and σ^2 unknown will now be introduced. Since both the mean μ and variance σ^2 are unknown, a joint prior is put on them. It is much more realistic to suppose that both parameters of a normal distribution are unknown rather than just one.

Let $Y_i = \mu + \varepsilon_i$

where $\varepsilon_i \sim N(0,\sigma^2)$ independent of each other.

or simply $Y_i \sim N\left(\mu,\sigma^2\right)$.

A prior on $\underline{\theta} = \left[\mu,\sigma^2\right]'$ is now needed. i.e., a joint prior on μ and σ^2 must be specified.

The non-informative Jeffreys' joint prior is

$$p(\mu,\sigma^2) \propto \frac{1}{\sigma^2} = \sigma^{-2}.$$

It is usual to take $p(\mu,\sigma^2) \propto \frac{1}{\sigma^2}$ as a joint vague prior. This is the product of the vague prior $p(\mu \mid \sigma^2) \propto c$ for μ and $p(\sigma^2) \propto \frac{1}{\sigma^2}$ for $\phi = \sigma^2$ in the gamma vague prior.

$$p(\mu,\sigma^2) = p(\mu \mid \sigma^2)\, p(\sigma^2) \propto c \times \frac{1}{\sigma^2} \propto \sigma^{-2}.$$

This then gives rise to Jeffereys' prior (Jeffereys, 1961).

$$p(\mu,\sigma^2) \propto \sigma^{-2}.$$

Consider the case where we have a set of observations $\underline{Y} = [Y_1, Y_2, \cdots Y_n]'$ which are normally distributed with mean μ and variance σ^2, i.e. $Y_i \sim N(\mu,\sigma^2)$. Both μ and σ^2 are unknown. The following theorem can now be stated:

Theorem 2.3.3

Suppose one has a set of observations $\underline{Y} = [Y_1, Y_2, \cdots Y_n]'$ which are normally distributed with parameters μ and σ^2 , both unknown and prior distribution $p(\mu,\sigma^2) \propto \sigma^{-2}$.

The marginal posterior distribution for σ^2 is:

$$\sigma^2 \big| \underline{Y} \sim IG\left(\tfrac{v}{2}, \tfrac{S}{2}\right)$$

and that for μ given σ^2:

$$\mu \mid \sigma^2, \underline{Y} \sim N\left(\overline{Y}, \sigma^2/n\right),$$

where $S = \sum \left(Y_i - \overline{Y}\right)^2$, $v = n-1$ and $s^2 = \frac{S}{v}$.

<u>Proof</u>

The proof is given in Appendix A2.

In theorem 2.3.3, the following is proven:

The conditional posterior density of μ is normally distributed

$$\mu \mid \sigma^2, \underline{Y} \sim N\left(\overline{Y}, \frac{\sigma^2}{n}\right)$$

and the posterior density for the variance component σ^2, is given by

$$p(\sigma^2 \mid \underline{Y}) = C(\sigma^2)^{-\frac{(n-1)}{2}-1} \exp\left\{-\tfrac{1}{2}(n-1)s^2/\sigma^2\right\} \quad \sigma^2 > 0,$$

i.e. an inverted gamma density, where $\underline{Y} = [Y_1, Y_2, \ldots Y_n]'$, $\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$, the sample mean, $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \overline{Y})^2$ the sample variance and the normalising constant is

$$C = \left\{\frac{(n-1)s^2}{2}\right\}^{\frac{1}{2}(n-1)} \frac{1}{\Gamma\left(\frac{n-1}{2}\right)}.$$

$p(\sigma^2 \mid \underline{Y})$ is an inverted gamma density of the form

$$IG(x \mid \alpha, \beta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{-\alpha-1} \exp(-\beta/x)$$

with

$$\alpha = \frac{1}{2}(n-1) \text{ and } \beta = \frac{(n-1)s^2}{2}.$$

Also $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{(n-1)}$.

Aircraft Data Example

To illustrate how to simulate process capability indices using the Bayesian techniques discussed, the following historical air–craft data collected by Pratt and Whitney Company is discussed. The data set provided in Table 2.1 is the first twenty of fifty observations of an aircraft feature (MQI 128) from a component hub which is part of the engine and is given in the first column of Table 1 on page 672 of Niverthi and Dey (2000). The unit of measurement is centimetres. The aircraft component, being very critical, requires the manufacturing process and subsequent steps to have a high degree of precision ($USL = 6.397$, $LSL = 6.393$ and $T = 6.395$). Both of the above limits were selected solely for illustrative purposes. In practice, fixed in advance limits are often determined from engineering or regulatory considerations. The target T is chosen as the centre of the specification limits.

Table 2.1 Aircraft data (MQI 128)

6.3950	6.3952	6.3950	6.3958	6.3950
6.3952	6.3952	6.3948	6.3952	6.3950
6.3950	6.3952	6.3946	6.3954	6.3952
6.3950	6.3952	6.3950	6.3952	6.3952

Source: (Niverthi and Dey, 2000)

2.4 SIMULATION OF THE VARIANCE, THE MEAN AND A FUNCTION OF THE MEAN AND VARIANCE

The simulation is performed with Matlab package version 7.1.0.246 (R14) service pack 3, August 02 2005, and subsequent calculations are all performed with the Matlab command prompt.

It has been shown that $\tau = \frac{2\beta}{X} \sim \chi^2_{2\alpha}$ and for $\alpha = \frac{(n-1)}{2}$, $\beta = \frac{(n-1)s^2}{2}$, $X = \sigma^2$

then $\tau = \frac{2\left(\frac{(n-1)s^2}{2}\right)}{\sigma^2} = \frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{(n-1)}$, and therefore $\sigma^2 = \frac{(n-1)s^2}{\tau}$.

1. Simulation of σ^2 can be obtained in the following way:

   (a) Simulate τ from a χ^2_{n-1} distribution, as a sum of $(n-1)$ squared independent standard normal random deviates i.e. $\tau = \sum_{i=1}^{n-1} Z_i^2$ where $Z_i \sim N(0,1)$.

   (b) Calculate $\sigma^{2*} = \frac{(n-1)s^2}{\tau}$ where * indicates a simulated value.

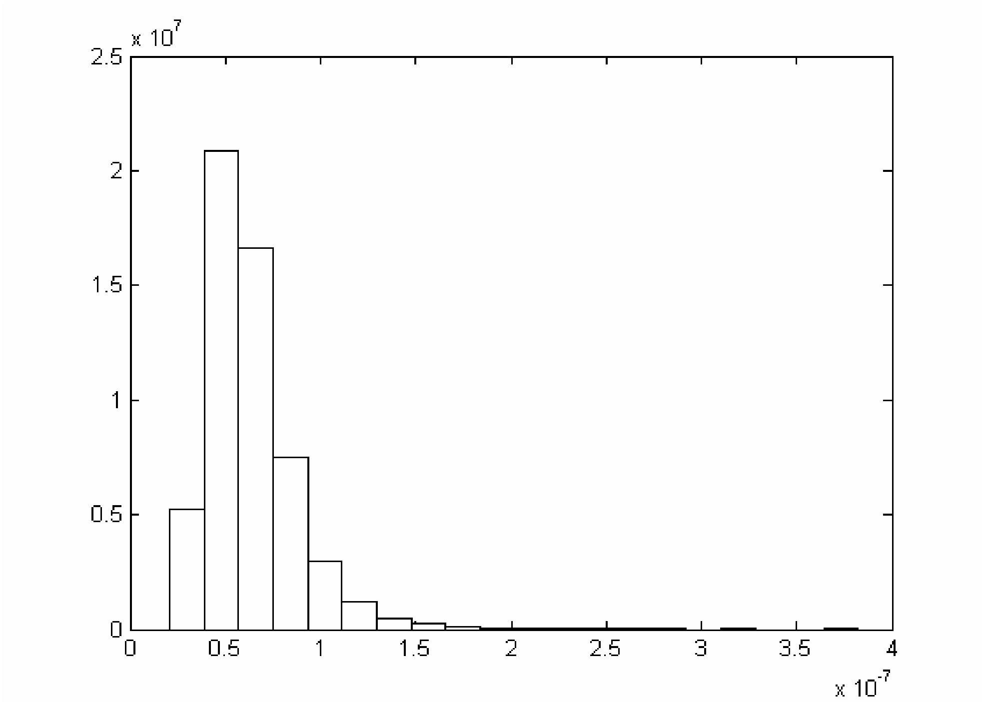
   (c) $\sigma^* = \sqrt{\sigma^{2*}}$.

2. $\mu^* \mid \sigma^{2*}, \underline{Y}$ is simulated from $N\left(\overline{Y}, \frac{\sigma^{2*}}{n}\right)$.

3. Compute $t\left(\underline{\theta}^*\right)$ where $\underline{\theta}^* = [\mu^*, \sigma^*]'$.

Repeat steps (1-3) for ℓ =1 to 10 000 times and a histogram of $t\left(\underline{\theta}^{*}\right)$ is then constructed. For each simulation one ends up with 10 000 pairs of $\left(\mu^{*(\ell)},\sigma^{*(\ell)}\right)$. This is precisely the advantage of the sampling based approach, where one can create the posterior distributions (in the form of a histogram) based on the samples and hence one can do inference from the posterior distribution without going through the exact distribution. From the distributions of the capability indices one is in a position to obtain quantiles, credible regions and perform other inferential tasks e.g. single summary measures of the process capability indices. Figure 2.1 below shows a histogram of simulated $p(\sigma^2|\underline{Y})$ for the aircraft example.

$p(\sigma^2|\underline{Y})$



$\sigma^2|\underline{Y}$

Figure 2.3: A histogram of simulated $p(\sigma^2|\underline{Y})$

Bayesian simulation is started with the simplest capability index, namely C_p.

2.5 SIMULATION OF C_p

Simulation of C_p can be obtained in the following way:

1. Simulation of σ:

   (a) Simulate τ from a χ^2_{n-1} distribution, as a sum of $(n-1)$ squared independent standard normal random deviates i.e. $\tau = \sum_{i=1}^{n-1} Z_i^{\,2}$ where $Z_i \sim N(0,1)$.

   (b) Calculate $\sigma^{2*} = \dfrac{(n-1)s^2}{\tau}$ where * indicates a simulated value.

   (c) $\sigma^* = \sqrt{\sigma^{2*}}$.

2. Compute $C_p^{\,*} = \dfrac{USL - LSL}{6\sigma^*}$.

Repeat steps (1 and 2) for ℓ =1 to 10 000 to get a series of values of C_p and plot a histogram of the index.

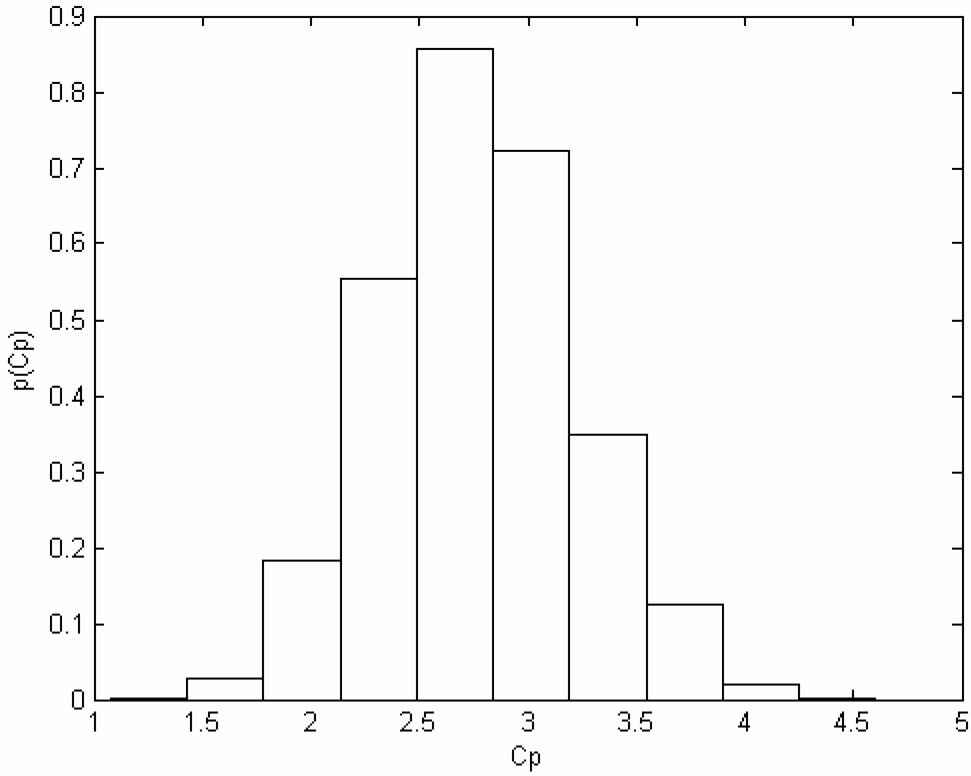


Figure 2.4: A histogram of simulated $p(C_p|\underline{Y})$

Table 2.2: Estimates of index value and 95% credible region for C_p

Index	Classical estimate	Bayesian estimate	
		Mean	95% Credible region
C_p	2.8066	2.7689	(1.9156; 3.6863)

The classical index estimate is $\widehat{C}_p = \dfrac{USL - LSL}{6s}$ where $s^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(Y_i - \overline{Y})^2$.

The Bayesian mean estimate is simply the average of the 10 000 realisations. $\overline{C}_p^{\,*} = \dfrac{1}{10000}\sum_{l=1}^{10000} C_p^{*(\ell)}$.

To construct the percentile credibility interval for C_p we sort the 10 000 C_p^{*} values in ascending order so that:

$$C_p^{*(1)} \le C_p^{*(2)} \le \cdots \le C_p^{*(10000)} \, .$$

In this application, 10 000 values of C_p^{*} are sorted from smallest to greatest and the critical values are found by selecting the value in the position $(\frac{\alpha}{2})\times 10\,000$ as the lower bound and the value in the position $(1-\frac{\alpha}{2})\times 10\,000$ as the upper bound. The credibility interval is then constructed as $C_p^{*}\left((\frac{\alpha}{2})\times 10\,000\right) - C_p^{*}\left((1-\frac{\alpha}{2})\times 10\,000\right)$. The 95% credibility interval is $C_p^{*}(250) - C_p^{*}(9750)$.

2.6 CHECKING THE SIMULATIONS USING THE TRUE DISTRIBUTIONS OF THE VARIANCE AND C_p

Once computation of the posterior distributions of all the estimates (using simulation) is accomplished, the relatively easy step of assessing the fit of the model to the data should not be ignored. Checking the model is crucial in statistical analysis. Bayesian prior-to-posterior inferences assume the whole structure of a probability model and can yield false inferences when the model is invalid.

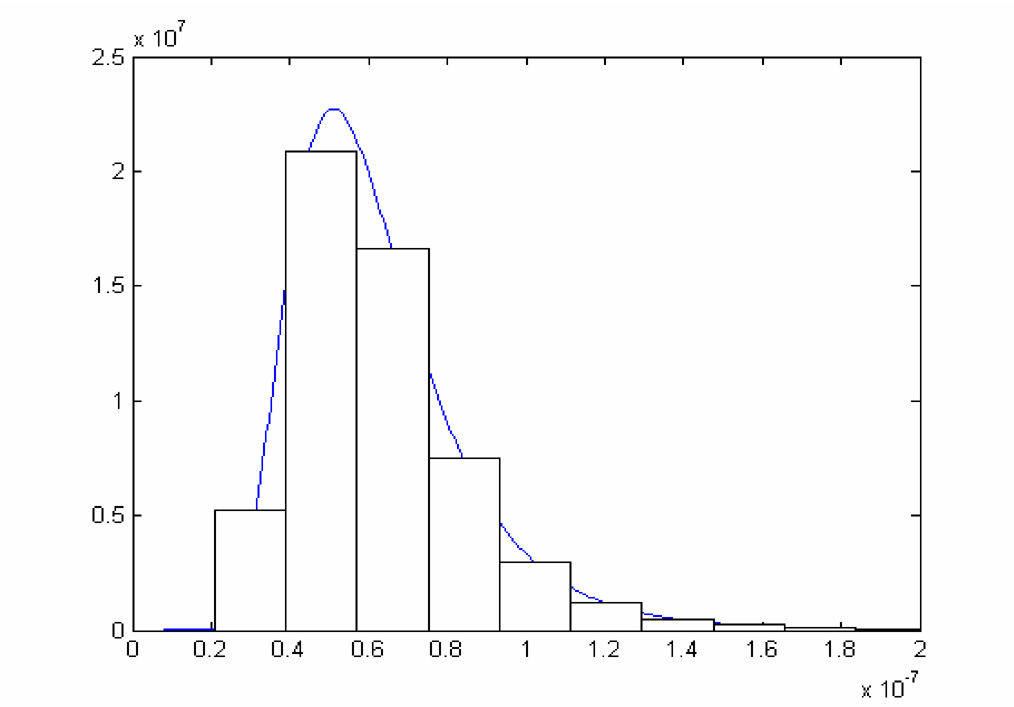
In theorem 2.3.3 one observes that for a given set of observations $\underline{Y} = [Y, Y_2, \cdots Y_n]'$ which are $N(\mu, \sigma^2)$, with μ and σ^2 both unknown and for the prior distribution $p(\mu, \sigma^2) \propto \sigma^{-2}$.

The marginal posterior distribution for σ^2 is:

$\sigma^2 | \underline{Y} \sim IG\left(\frac{v}{2}, \frac{S}{2}\right)$ where $S = \sum \left(Y_i - \overline{Y}\right)^2$, $v = n - 1$ and $s^2 = \dfrac{S}{v}$.

This distribution can now be plotted and contrasted with the simulated values of σ^2. Below is a diagrammatic representation of the distribution of $(\sigma^2 | \underline{Y})$ over-laid on a histogram of the 10000 simulations.

$p(\sigma^2 | \underline{Y})$



$\sigma^2 | \underline{Y}$

Figure 2.5: A posterior of distribution σ^2 and histogram of simulated $\sigma^2 | \underline{Y}$

The next theorem also illustrates the idea of checking simulations by plotting the true distribution of the process capability index.

Theorem 2.3.4

If X has an inverse gamma distribution with parameters $\alpha > 0$ and $\beta > 0$, denoted as $X \sim IG(\alpha, \beta)$ to mean

$$p_X(x) = \frac{\beta^{\alpha}}{\Gamma(x)} x^{-(\alpha+1)} e^{-\beta/x},\ x > 0$$

then $Y = k\sqrt{\frac{1}{X}}$ has the following distribution: $p_Y(y) = 2\frac{\left(\beta k^{-2}\right)^{\alpha}}{\Gamma(\alpha)}\left(y^2\right)^{\alpha-\frac{1}{2}} e^{-\left(\beta k^{-2}\right)y^2}$.

<u>Proof</u>

The proof is given in Appendix A2.

If we now take $k = \frac{USL - LSL}{6} = \frac{d}{3}$, were $d = \frac{USL - LSL}{2}$ and $X = \sigma^2$ in theorem 2.3.4, then $Y = k\sqrt{\frac{1}{X}} = \frac{USL - LSL}{6\sigma} = C_p$. The density in theorem 2.3.4 is then the density of C_p and it can be plotted and compared with the histogram of simulated C_p values.

Below is a diagrammatic representation of the distribution of C_p over-laid on a histogram of the10000 simulations.

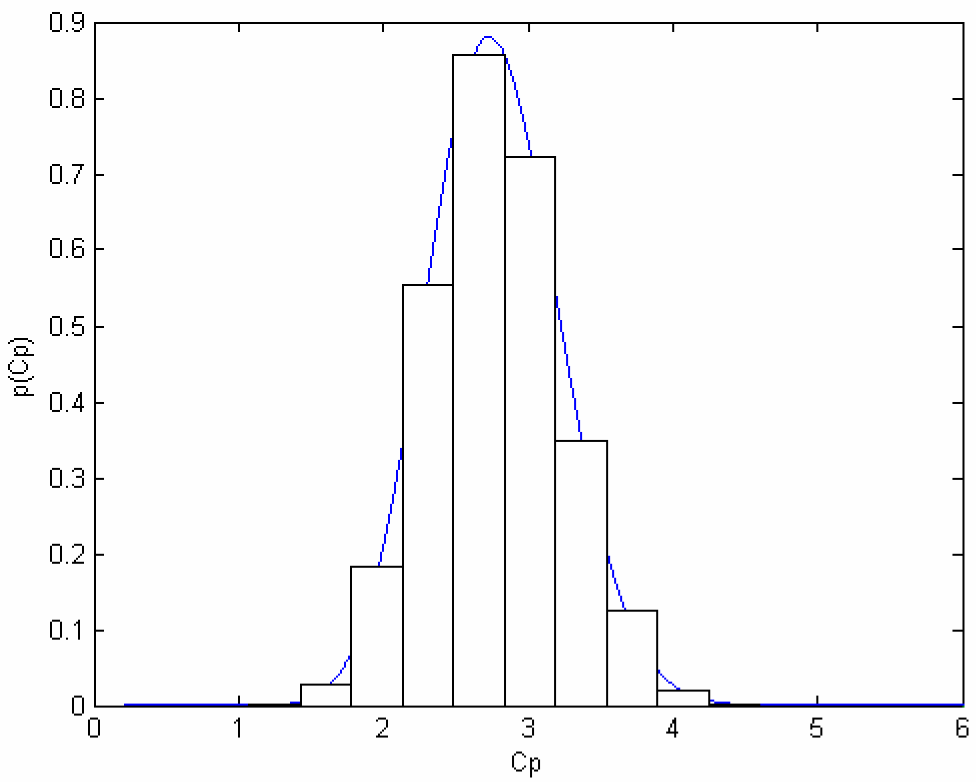


Figure 2.6: A posterior distribution of C_p and histogram of simulated $C_p|\underline{Y}$

The Bayesian simulation results using the non-informative priors look plausible. The curves do go through the mid-points of the histograms for a good fit.

2.7 SIMULATION OF C_{pl} AND C_{pu}

Simulation of C_{pl} and C_{pu} can be obtained in the following way:

1. Simulation of σ:

   (a) Simulate τ from a χ^2_{n-1} distribution, as a sum of $(n-1)$ squared independent standard normal random deviates i.e. $\tau = \sum_{i=1}^{n-1} Z_i^{\ 2}$ where $Z_i \sim N\left(0,1\right)$.

   (b) Calculate $\sigma^{2*} = \dfrac{(n-1)s^2}{\tau}$ where * indicates a simulated value.

(c) $\sigma^* = \sqrt{\sigma^{2*}}$.

2. $\mu^* \mid \sigma^{2*}, \underline{Y}$ is simulated from $N\left(\overline{Y}, \frac{\sigma^{2*}}{n}\right)$.

3. Compute $C_{pl}{}^* = \frac{\mu^* - LSL}{3\sigma^*}$ and $C_{pu}{}^* = \frac{USL - \mu^*}{3\sigma^*}$.

Repeat steps (1-3) for ℓ=1 to 10 000 to get a series of simulated values of C_{pl} and C_{pu} . A histogram of each index is then constructed in Matlab.

Figures 2.7 and 2.8 show the histograms. Summary measures from the Bayesian simulation approach are given in tables 2.3 and 2.4.

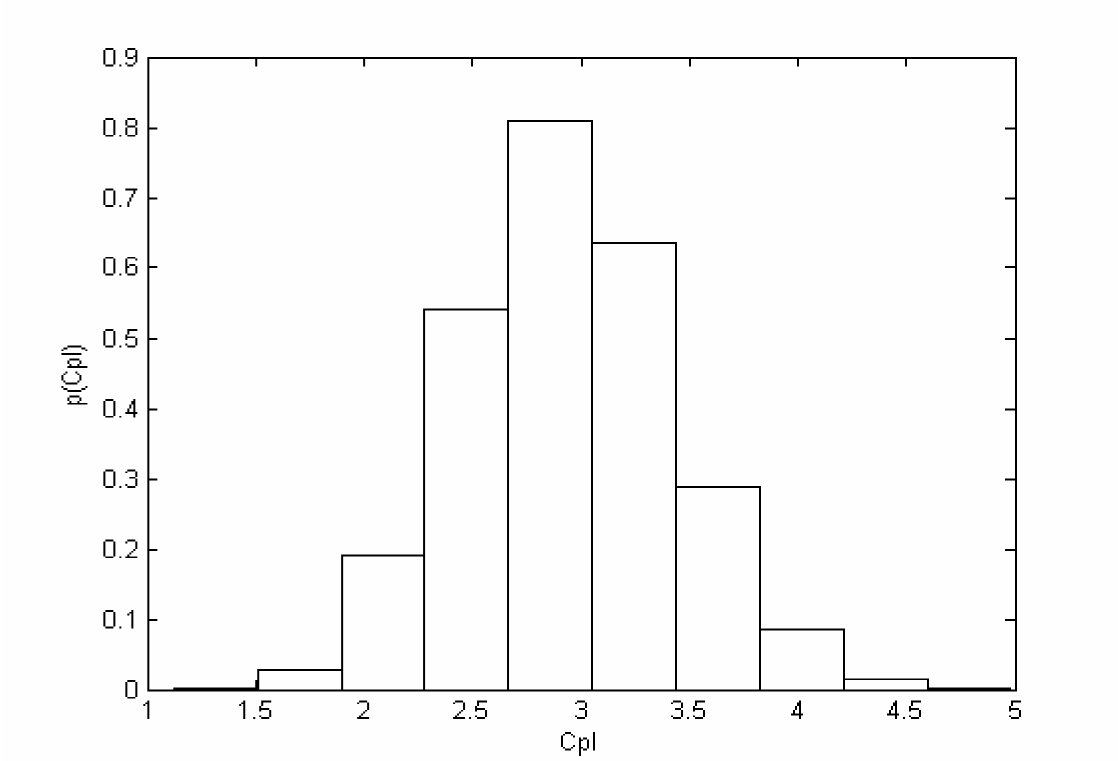


Figure 2.7: A histogram of simulated $C_{pl} \mid \underline{Y}$

Table 2.3: Estimates of index value and 95% credible region for C_{pl}

Index	Classical estimate	Bayesian estimate	
		Mean	95% Credible region
C_{pl}	2.9750	2.9349	(2.0185; 3.9118)

The classical index is estimated by $\widehat{C}_{pl} = \dfrac{\overline{Y} - LSL}{3s}$.

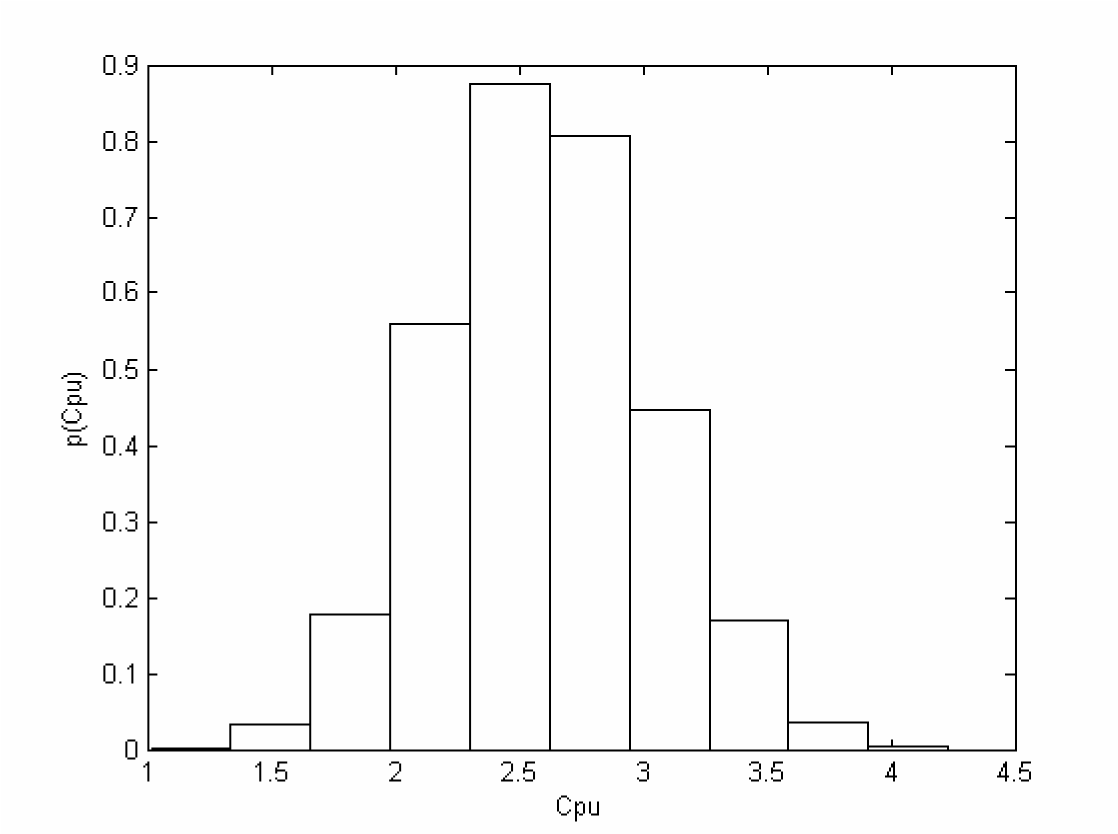


Figure 2.8: A histogram of simulated $C_{pu} \mid \underline{Y}$

Table 2.4: Estimates of index value and 95% credible region for C_{pu}

Index	Classical estimate	Bayesian estimate	
		Mean	95% Credible region
C_{pu}	2.6383	2.6029	(1.7891; 3.4800)

The classical index is estimated by $\widehat{C}_{pu} = \dfrac{USL - \overline{Y}}{3s}$.

Since the process is centred on the midpoint (target) of the specification range ($\overline{Y} = 6.3951$ and T=6.395), and the symmetry properties, the histograms of C_{pl} and C_{pu} look the same.

Chapter 4 will explore the index $C_{pl} = \frac{\mu - LSL}{3\sigma}$ further. An exact posterior distribution of the index will be derived. Because of symmetry, the same results distribution would apply to the index C_{pu}. The index C_{pl} is extended to two variance components in chapter 5 and to three variance components in chapter 6.

2.8 SIMULATION OF C_{pk}

Simulation of C_{pk} can be obtained in the following way:

1. Simulation of σ:

   (a) Simulate τ from a χ^2_{n-1} distribution, as a sum of $(n-1)$ squared independent standard normal random deviates i.e. $\tau = \sum_{i=1}^{n-1} Z_i^{\,2}$ where $Z_i \sim N(0,1)$.

   (b) Calculate $\sigma^{2*} = \frac{(n-1)s^2}{\tau}$ where * indicates a simulated value.

   (c) $\sigma^* = \sqrt{\sigma^{2*}}$.

2. $\mu^* \mid \sigma^{2*}, \underline{Y}$ is simulated from $N\left(\overline{Y}, \frac{\sigma^{2*}}{n}\right)$.

3. Compute $C_{pl}{}^* = \frac{\mu^* - LSL}{3\sigma^*}$, $C_{pu}{}^* = \frac{USL - \mu^*}{3\sigma^*}$ and

   $C_{pk}{}^* = \min\left(C_{pu}{}^* \, , C_{pl}{}^* \right)$.

Repeat steps (1-3) for ℓ =1 to 10 000 to get a series of values of C_{pk} and a histogram of the index.

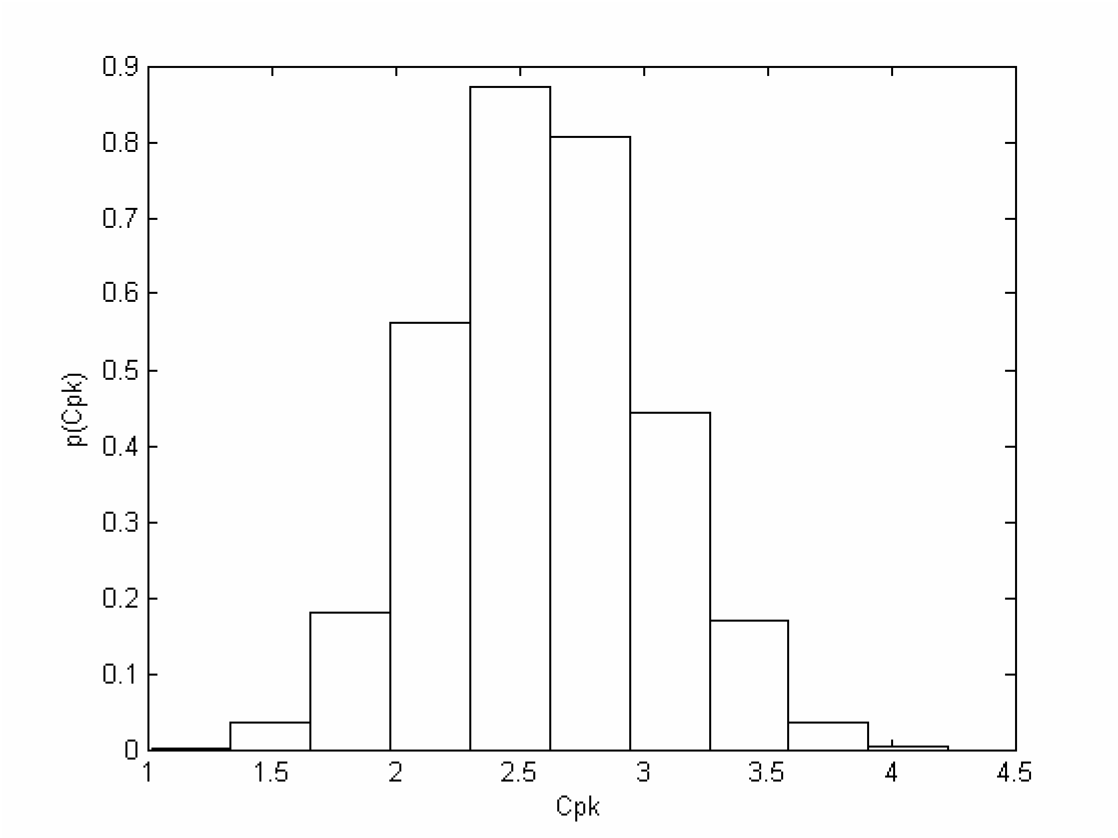


Figure 2.9: A histogram of simulated $C_{pk} \left| \underline{Y} \right.$

Table 2.5: Estimates of index value and 95% credible region for C_{pk}

Index	Classical estimate	Bayesian estimate	
		Mean	95% Credible region
C_{pk}	2.6383	2.6017	(1.7859; 3.4800)

The classical index is estimated by minimum$\left(\widehat{C}_{pl}, \widehat{C}_{pu}\right)$.

2.9 SIMULATION OF P_{pl}

When the index C_p is computed using $s = \sqrt{\dfrac{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}{n-1}}$, as an estimate of σ based on individual measurements, then the indices C_p and P_p are one and the same index.

2.10 SIMULATION OF P_{pk}

Following the same arguments for C_p and P_p, when the index C_{pk} is computed using $s=\sqrt{\frac{\sum_{i=1}^{n}(Y_i-\overline{Y})^2}{n-1}}$ as an estimate of σ based on individual measurements, then the indices C_{pk} and P_{pk} are one and the same index.

2.11 SIMULATION OF C_pT

Simulation of C_pT can be obtained in the following way:

1. Simulation of σ:

   (a) Simulate τ from a χ^2_{n-1} distribution, as a sum of $(n-1)$ squared independent standard normal random deviates i.e. $\tau=\sum_{i=1}^{n-1} Z_i^2$ where $Z_i \sim N\left(0,1\right)$.

   (b) Calculate $\sigma^{2*}=\frac{(n-1)s^2}{\tau}$ where * indicates a simulated value.

   (c) $\sigma^*=\sqrt{\sigma^{2*}}$.

2. Compute $C_pT^*=\min\left(\frac{USL-T}{3\sigma^*},\frac{T-LSL}{3\sigma^*}\right)$, where the target value is T .

Repeat steps (1-2) for $\ell=1$ to 10 000 to get a series of values of C_pT and plot a histogram of the index.

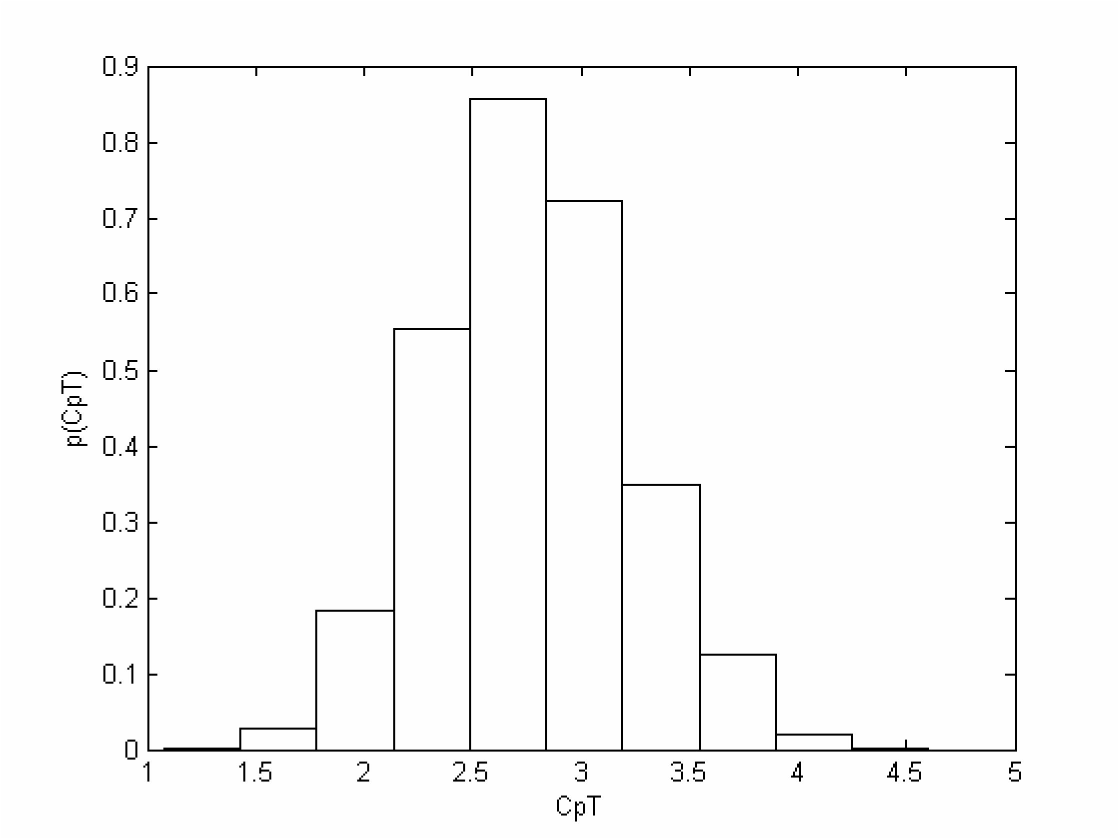


Figure 2.10: A histogram of simulated $C_pT \,|\, \underline{Y}$

Table 2.6: Estimates of index value and 95% credible region for C_pT

Index	Classical estimate	Bayesian estimate	
		Mean	95% Credible region
C_pT	2.8066	2.7689	(1.9156; 3.6863)

The classical index is estimated by $\widehat{C}_pT = \min\left(\frac{USL-T}{3s}, \frac{T-LSL}{3s}\right)$

$C_pT = \min\left(\frac{USL-T}{3\sigma}, \frac{T-LSL}{3\sigma}\right) = \frac{d-|T-M|}{3\sigma}$ and when $T = M$, $C_pT = C_p$ which explains why the histograms and estimates of C_pT are the same as C_p.

If we now take $k = \dfrac{d - |T - M|}{3}$ were $d = \dfrac{USL - LSL}{2}$ and $X = \sigma^2$ in theorem 2.3.4, then $Y = k\sqrt{\dfrac{1}{X}} = \dfrac{d - |T - M|}{3\sigma} = C_p T$. The density in theorem 2.3.4 is then the density of $C_p T$ and it can be plotted and compared with the histogram of simulated $C_p T$ values.

Below is a diagrammatic representation of the distribution of $C_p T$ over-laid on a histogram of the10 000 simulations.

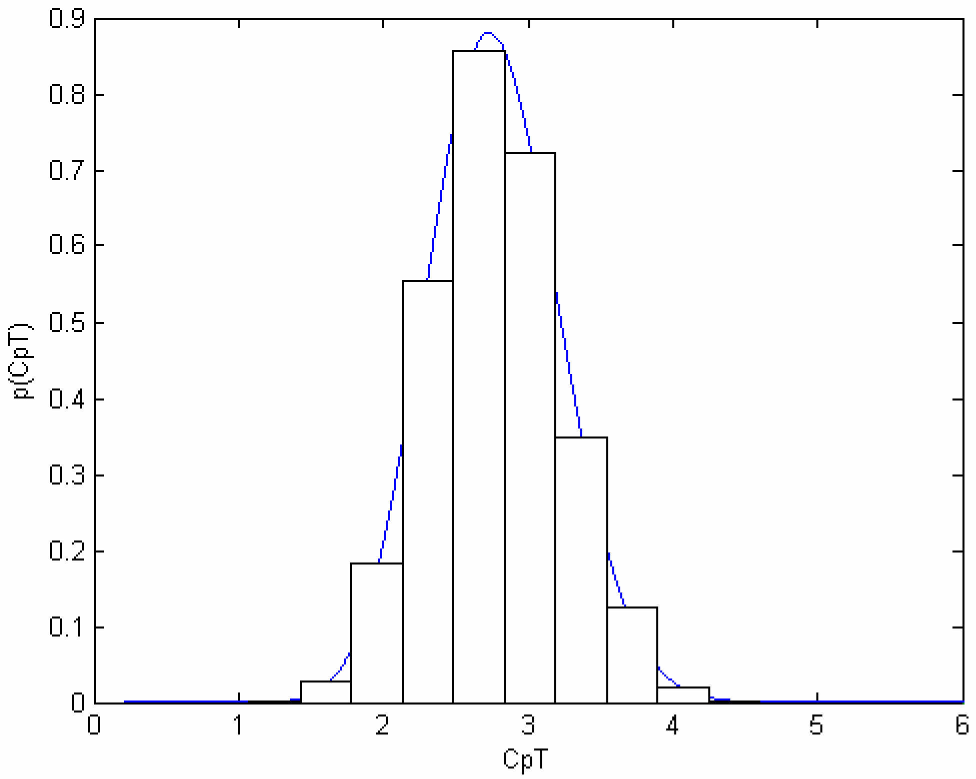


Figure 2.11: A posterior distribution of $C_p T$ and histogram of simulated $C_p T | \underline{Y}$

2.12 SIMULATION OF C_{pm}

Simulation of C_{pm} can be obtained in the following way:

1. Simulation of σ:

   (a) Simulate τ from a χ^2_{n-1} distribution, as a sum of $(n-1)$

squared independent standard normal random deviates i.e. $\tau = \sum_{i=1}^{n-1} Z_i^2$ where $Z_i \sim N(0,1)$.

(b) Calculate $\sigma^{2*} = \frac{(n-1)s^2}{\tau}$ where * indicates a simulated value.

(c) $\sigma^* = \sqrt{\sigma^{2*}}$.

2. $\mu^* \mid \sigma^{2*}, \underline{Y}$ is simulated from $N\left(\overline{Y}, \frac{\sigma^{2*}}{n}\right)$.

3. Compute $C_{pm}{}^* = \frac{USL - LSL}{6\sqrt{\sigma^* + \left(\mu^* - T\right)^2}}$.

Repeat steps (1-3) for ℓ =1 to 10 000 to get a series of values of C_{pm} and plot a histogram of the index.

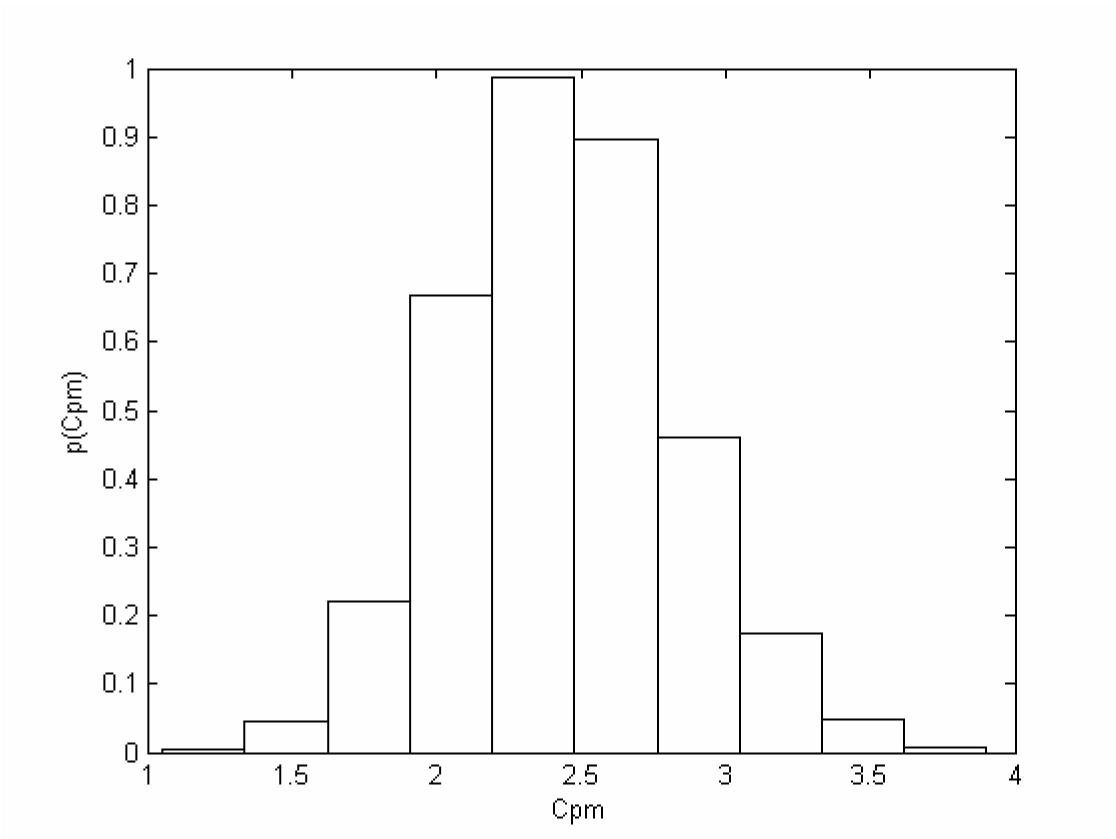


Figure 2.12: A histogram of simulated $C_{pm} | \underline{Y}$

Table 2.7: Estimates of index value and 95% credible region for C_{pm}

Index	Classical estimate	Bayesian estimate	
		Mean	95% Credible region
C_{pm}	2.5051	2.4419	(1.7199; 3.2467)

The classical index is estimated by $\widehat{C}_{pm} = \dfrac{USL - LSL}{6\sqrt{s + \left(\overline{Y} - T\right)^2}}$.

2.13 SIMULATION OF C_{pmk}

Simulation of C_{pmk} can be obtained in the following way:

1. Simulation of σ:

   (a) Simulate τ from a χ^2_{n-1} distribution, as a sum of $(n-1)$ squared independent standard normal random deviates i.e. $\tau = \sum_{i=1}^{n-1} Z_i^{\ 2}$ where $Z_i \sim N\left(0,1\right)$.

   (b) Calculate $\sigma^{2*} = \dfrac{(n-1)s^2}{\tau}$ where * indicates a simulated value.

   (c) $\sigma^* = \sqrt{\sigma^{2*}}$.

2. $\mu^* \mid \sigma^{2*}, \underline{Y}$ is simulated from $N\left(\overline{Y}, \dfrac{\sigma^{2*}}{n}\right)$

3. Compute $C_{pmk}{}^* = \dfrac{\min(USL - \mu^*, \mu^* - LSL)}{3\sqrt{\sigma^* + \left(\mu^* - T\right)^2}}$.

Repeat steps (1-3) for ℓ =1 to 10 000 to get a series of values of C_{pmk} and plot a histogram of the index.

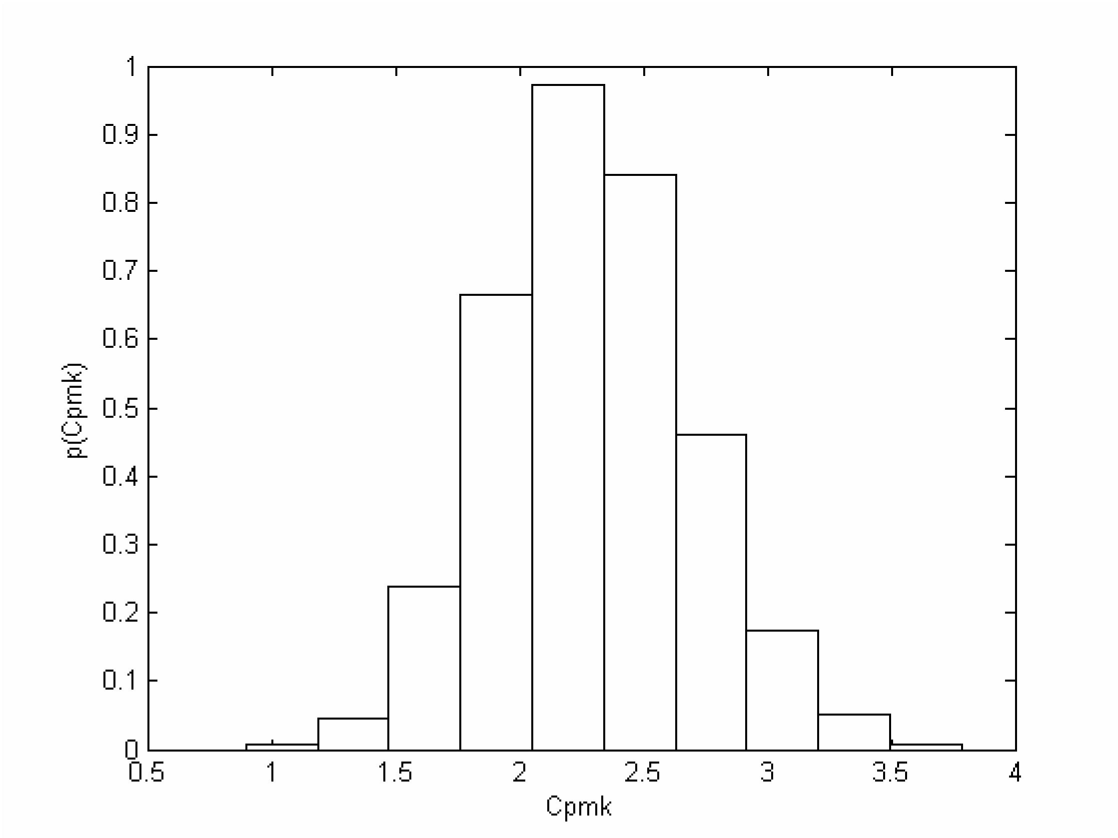


Figure 2.13: A histogram of simulated $C_{pmk} \mid \underline{Y}$

Table 2.8: Estimate of index value and 95% credible region for C_{pmk}

Index	Classical estimate	Bayesian estimate	
		Mean	95% Credible region
C_{pmk}	2.3548	2.2996	(1.5572;3.1352)

$$\widehat{C}_{pmk} = \frac{\min(USL-\overline{Y},\overline{Y}-LSL)}{3\sqrt{s+\left(\overline{Y}-T\right)^2}} \quad .$$

2.14 SIMULATION OF $C^{\#}_{pm}$

Simulation of $C^{\#}_{pm}$ can be obtained from in the following way:

1. Simulation of σ:

   (a) Simulate τ from a χ^2_{n-1} distribution, as a sum of $(n-1)$ squared independent standard normal random deviates i.e. $\tau = \sum_{i=1}^{n-1} Z_i^{\,2}$

where $Z_i \sim N(0,1)$.

(b) Calculate $\sigma^{2*} = \frac{(n-1)s^2}{\tau}$ where * indicates a simulated value.

(c) $\sigma^* = \sqrt{\sigma^{2*}}$.

2. $\mu^* \mid \sigma^{2*}, \underline{Y}$ is simulated from $N\left(\overline{Y}, \frac{\sigma^{2*}}{n}\right)$.

3. Compute $C_{pm}^{\#\,*} = \frac{\min(USL - T, T - LSL)}{6\sqrt{\sigma^* + \left(\mu^* - T\right)^2}}$.

Repeat steps (1-3) ℓ=1 to 10 000 to get a series of values of $C_{pm}^{\#}$ and plot a histogram of the index.

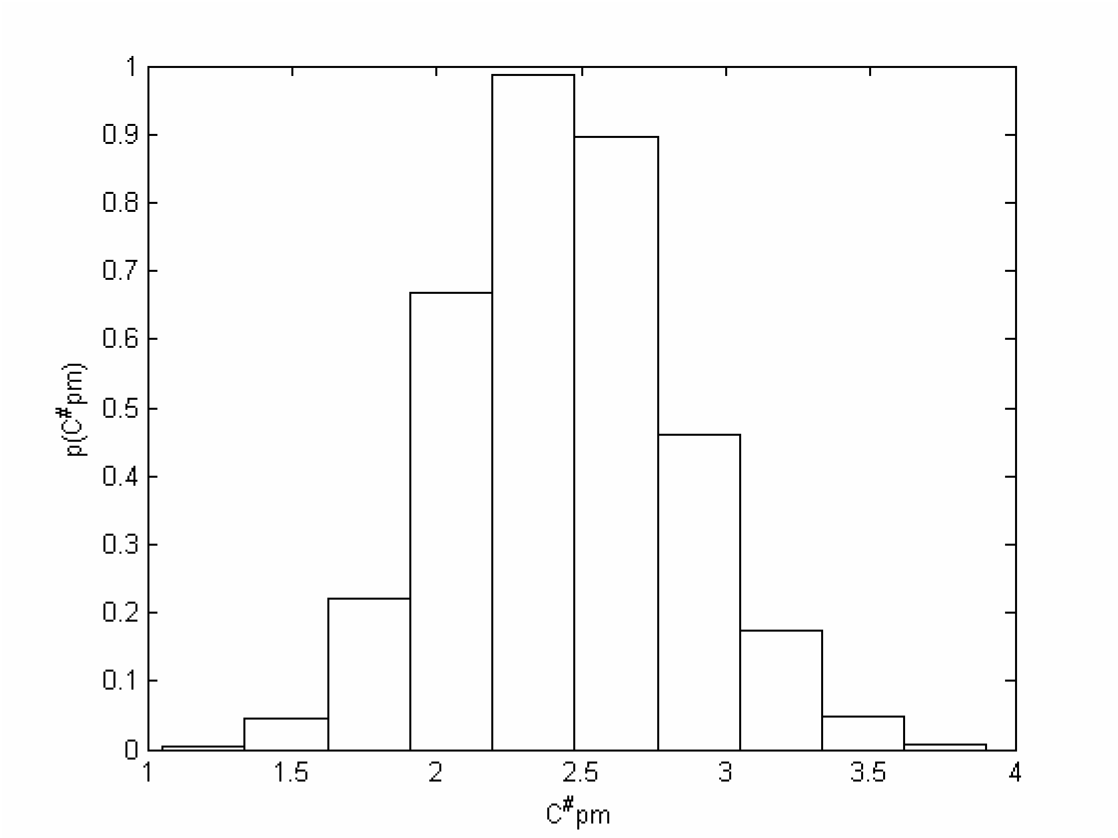


Figure 2.14: A histogram of simulated $C_{pm}^{\#} \mid \underline{Y}$

Table 2.9: Estimates of index value and 95% credible region for $C^{\#}_{pm}$

Index	Classical estimate	Bayesian estimate Mean	95% Credible region
$C^{\#}_{pm}$	2.5051	2.4419	(1.7199; 3.2467)

$$\hat{C}^{\#}_{pm} = \frac{\min(USL-T, T-LSL)}{3\sqrt{s+\left(\overline{Y}-T\right)^2}}.$$

When T is equal to M, $C^{\#}_{pm}$ is equal to C_{pm}.

2.15 COMPARING THE RESULTS FOR THE AIRCRAFT DATA

The numerical results of the indices are summarised in table 2.10 below.

Table 2.10: Summary table of the values of the indices

Index	Classical index estimate	Bayesian simulations mean estimate	Bayesian 95% credible region
C_p	2.8066	2.7689	(1.9156; 3.6863)
C_{pl}	2.9750	2.9349	(2.0185; 3.9118)
C_{pu}	2.6383	2.6029	(1.7891; 3.4800)
C_{pk}	2.6383	2.6017	(1.7859; 3.4800)
C_pT	2.8066	2.7689	(1.9156; 3.6863)
C_{pm}	2.5051	2.4419	(1.7199; 3.2467)
C_{pmk}	2.3548	2.2996	(1.5572; 3.1352)
$C^{\#}_{pm}$	2.5051	2.4419	(1.7199; 3.2467)

The means of the indices using Bayesian simulation techniques are compared to their more commonly used frequentist estimates. All the estimated process capability indices (PCIs) are larger than 1, indicating that this process is capable; in fact all the estimated PCIs are greater than 1.67, indicating that this process is super. This conclusion is true regardless of whether the frequentist or the Bayesian method is under consideration. Since the process is centred on the midpoint (target) of the specification range ($\overline{Y} = 6.3951$and T=M=6.395), and the symmetry properties, the histograms of the indices look similar. This is not surprising since most of the indices are all equal when $\mu = T = M$.

Since probability models in most data analysis will not be perfectly true, and looking at columns 2 and 3 in table 2.10, the Bayesian results look plausible. The mean values of the Bayesian approach are nearly the same to one decimal place to their frequentist analogue. From the simulation results it is clear that the prior $\pi(\mu,\sigma^2) \propto \sigma^{-2}$ works quite well. It seems, therefore, that the frequentist properties of Bayesian inferences of capability indices based on the prior $\pi(\mu,\sigma^2) \propto \sigma^{-2}$ are adequate.

The added advantage of the Bayesian approach is that, from the distribution (represented by the histogram) of the capability indices, one is in a position to obtain quantiles, credible regions and perform other inferential tasks. In the case of C_p and C_pT , the exact distribution are derived.

In Van der Merwe and Chikobvu (2004), the reference prior for the capability index $C_{pl} = \dfrac{\mu - LSL}{3\sigma}$ (which is applicable when there is no upper limit) is derived. The reference prior relative to the ordered parameterisation (μ,σ^2) is given by

$$\pi^R(\mu,\sigma^2) \propto \sigma^{-3}\left\{1 + \frac{(\mu - LSL)^2}{2\sigma^2}\right\}^{-\frac{1}{2}}$$

and is also a probability-matching prior. The derivations of the results are also discussed in chapter 4.

In a similar way the reference prior for the capability index $C_{pu} = \frac{USL-\mu}{3\sigma}$ can be derived. However, the reference prior for $C_{pk} = \min\left(\frac{USL-\mu}{3\sigma}, \frac{\mu-LSL}{3\sigma}\right)$ and other indices cannot be derived.

In chapter 3, Bayesian simulation techniques are used to solve the supplier selection problem using process capability indices.

Appendix A2

Proof of theorem 2.3.1

$p_Y(y) = p_X(g^{-1}(y))\left|\frac{d}{dy}g^{-1}(y)\right|$ let $y = g(x) = \frac{1}{x}$

then $x = g^{-1}(y) = \frac{1}{y}$ and $\frac{d}{dy}g^{-1}(y) = -y^{-2}$

$$p_Y(y) = \frac{\beta^\alpha}{\Gamma(\alpha)}\left(\frac{1}{y}\right)^{(\alpha-1)} e^{-\beta\left(\frac{1}{y}\right)}\left|-y^{-2}\right|$$

$$p_Y(y) = \frac{\beta^\alpha}{\Gamma(\alpha)}(y)^{-\alpha-1} e^{-\beta\left(\frac{1}{y}\right)}$$

$$p_Y(y) = \frac{\beta^\alpha}{\Gamma(\alpha)} y^{-(\alpha+1)} e^{-\beta/y}$$

as required.

Proof of theorem 2.3.2

Using moment generating functions

$Y = \frac{1}{X} \sim G(\alpha,\beta)$ and $M_Y(t) = \left(\frac{\beta}{\beta-t}\right)^\alpha$

If we let $\tau = \frac{2\beta}{X} = 2\beta Y$ and remembering that if $\tau = a + bY$

then $M_\tau(t) = e^{at}M_Y(bt)$, with $a = 0$ *and* $b = 2\beta$

therefore $M_\tau(t) = e^0 M_Y(2\beta t) = \left(\frac{\beta}{\beta-2\beta t}\right)^\alpha = \left(\frac{1}{1-2t}\right)^\alpha = (1-2t)^{-\alpha}$

This is the moment generating function of a Chi-square distribution with $2a$ degrees of freedom. Hence $\tau = \dfrac{2\beta}{X} = 2\beta Y \sim \chi^2_{2\alpha}$.

Proof of theorem 2.3.3

The likelihood of n independent and identically normally distributed random variables is:

$$l(\mu,\sigma^2 | \underline{Y}) \propto p(\underline{Y} | \mu,\sigma^2)$$

$$= \prod_{i=1}^{n} \left(2\pi\sigma^2\right)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(Y_i-\mu)^2 \big/ \sigma^2\right\}$$

$$= \left(\sigma^2\right)^{-n/2} \exp\left(-\frac{1}{2}\sum\left(Y_i - \overline{Y} + \overline{Y} - \mu\right)^2 \big/ \sigma^2\right)$$

$$= \left(\sigma^2\right)^{-n/2} \exp\left(-\frac{1}{2}\sum\left((Y_i-\overline{Y})^2 + 2(Y_i-\overline{Y})(\overline{Y}-\theta) + (\overline{Y}-\mu)^2\right) \big/ \sigma^2\right)$$

$$= \left(\sigma^2\right)^{-n/2} \exp\left(-\frac{1}{2}\left\{\sum\left(Y_i-\overline{Y}\right)^2 + n\left(\overline{Y}-\mu\right)^2\right\} \big/ \sigma^2\right)$$

$$= \left(\sigma^2\right)^{-n/2} \exp\left(-\frac{1}{2}\left\{S + n\left(\overline{Y}-\mu\right)^2\right\} \big/ \sigma^2\right)$$

where $S = \sum\left(Y_i - \overline{Y}\right)^2$.

It is convenient to define $s^2 = \dfrac{S}{n-1}$.

If we take the vague prior, then

$$p(\mu,\sigma^2 | \underline{Y}) \propto p(\mu,\sigma^2)\, p(\underline{Y} | \mu,\sigma^2)$$

$$\propto \left(\sigma^2\right)^{-1} \left(\sigma^2\right)^{-n/2} \exp\left(-\frac{1}{2}\left\{S + n\left(\overline{Y}-\mu\right)^2\right\} \big/ \sigma^2\right)$$

$$\propto \left(\sigma^2\right)^{-n/2-1} \exp\left(-\frac{1}{2}\left\{S + n\left(\overline{Y}-\mu\right)^2\right\} \big/ \sigma^2\right)$$

For reasons which will appear later it is convenient to set

$$v = n-1$$

in the power of σ^2, but not in the exponential, so that

$$p(\mu,\sigma^2 \mid \underline{Y}) \propto \left(\sigma^2\right)^{-(v+1)/2-1} \exp\left(-\frac{1}{2}\left\{S+n\left(\overline{Y}-\mu\right)^2\right\}\Big/\sigma^2\right).$$

Marginal distribution of the variance

If knowledge about σ^2 rather than μ is required, μ is integrated out from the posterior distribution.

$$p(\sigma^2 \mid \underline{Y}) = \int p(\mu,\sigma^2 \mid \underline{Y})\, d\mu$$

$$\propto \int_{-\infty}^{\infty} \left(\sigma^2\right)^{-\frac{n}{2}-1} \exp\left(-\frac{1}{2}\left\{S+n\left(\mu-\overline{Y}\right)^2\right\}\Big/\sigma^2\right) d\mu$$

$$\propto \left(\sigma^2\right)^{-(\frac{n}{2}+\frac{1}{2})} \exp\left(-\frac{1}{2}S/\sigma^2\right) \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2/n}} \exp\left\{-\frac{1}{2}\left(\mu-\overline{Y}\right)^2\Big/\left(\sigma^2/n\right)\right\} d\mu$$

$$\propto \left(\sigma^2\right)^{-(\frac{n-1}{2}+1)} \exp\left(-\tfrac{S}{2}\big/\sigma^2\right)$$

$$= \left(\sigma^2\right)^{-(v/2+1)} \exp\left(-\tfrac{S}{2}\big/\sigma^2\right), \text{ where } v = n-1$$

as the last integral is that of a normal density.

It follows that the posterior density of the variance is $Inv-G\left(\frac{v}{2},\frac{S}{2}\right)$. Except for the fact that n is replaced by v = n-1.

Conditional density of the mean for given variance

The joint posterior can be written in the following form

$$p(\mu,\sigma^2 \mid \underline{Y}) = p(\sigma^2 \mid \underline{Y})\, p(\mu \mid \sigma^2,\underline{Y}).$$

Thus $$p(\mu \mid \sigma^2,\underline{Y}) = \frac{p(\mu,\sigma^2 \mid \underline{Y})}{p(\sigma^2 \mid \underline{Y})}$$

$$p(\mu,\sigma^2 \mid \underline{Y}) \propto \left(\sigma^2\right)^{-(v+1)/2-1} \exp\left(-\frac{1}{2}\left\{S+n\left(\overline{Y}-\mu\right)^2\right\}\Big/\sigma^2\right)$$

$$p(\sigma^2 \mid \underline{Y}) \propto \left(\sigma^2\right)^{-v/2-1} \exp\left(-\frac{1}{2}S/\sigma^2\right).$$

This implies that

$$p(\mu \mid \sigma^2, \underline{Y}) \propto \left(\sigma^2\right)^{-1/2} \exp\left(-\frac{1}{2} n(\mu - \overline{Y})^2 \big/ \sigma^2 \right)$$

which as the density integrates to unity implies that

$$p(\mu \mid \sigma^2, \underline{Y}) = \frac{1}{\sqrt{2\pi\sigma^2/n}} \exp\left(-\frac{1}{2} (\mu - \overline{Y})^2 \big/ \left(\sigma^2/n\right) \right)$$

that is, for given σ^2 and $\underline{Y}$, the distribution of the mean μ is $N\left(\overline{Y}, \sigma^2/n\right)$.

Proof of theorem 2.3.4

$p_Y(y) = p_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right|$ let $y = g(x) = k\sqrt{\frac{1}{x}}$

then $x = g^{-1}(y) = \frac{k^2}{y^2}$ and $\frac{d}{dy} g^{-1}(y) = -2k^2 y^{-3}$

$$p_Y(y) = \frac{\beta^\alpha}{\Gamma(\alpha)} \left(\frac{k^2}{y^2} \right)^{-(\alpha+1)} e^{-\beta/\frac{k^2}{y^2}} \left| -2k^2 y^{-3} \right|$$

$$p_Y(y) = 2 \frac{\left(\beta k^{-2}\right)^\alpha}{\Gamma(\alpha)} (y)^{2\alpha-1} e^{-\beta y^2/k^2}$$

$$p_Y(y) = 2 \frac{\left(\beta k^{-2}\right)^\alpha}{\Gamma(\alpha)} \left(y^2\right)^{\alpha-\frac{1}{2}} e^{-\left(\beta k^{-2}\right) y^2}$$

as required.

CHAPTER 3

A BAYESIAN SIMULATION SOLUTION TO THE SUPPLIER SELECTION PROBLEM USING CAPABILITY INDICES

3.1 INTRODUCTION

Supplier selection is an important part of supply chain management. It is sometimes of interest to compare capability indices for two different suppliers or the same supplier before and after an adjustment (assessing the impact of process improvement). This chapter investigates a procedure for testing the equality of $N \geq 2$ process capability indices from a Bayesian simulation point of view. The problem naturally occurs when comparing suppliers or product performance from N possible suppliers of a critical component. The statistical comparison of competing manufacturing suppliers is an important aspect of quality control that aids quality managers in the choice of potential suppliers of a product. Chou (1994) applies hypothesis testing to select a supplier by testing process capability indices. The method is not difficult to apply, but is only suitable for comparing two suppliers. Tong *et al.* (1998) applies the bootstrap method to construct the bootstrap confidence interval to distinguish between the C_{pk} values of two suppliers.

Suppose a component has a univariate quality characteristic Y with upper and lower specification limits USL and LSL respectively. Manufacturing processes are assumed to be normally distributed with mean μ and variance σ^2. It is assumed that the cost of the component is the same regardless of the supplier. Thus, the quality control manager can choose the supplier that is best able to provide the manufacturer with components that are within specification. Hence, the quality control manager wishes to choose the supplier with a process that has the highest capability. The quality control manager obtains data from each of the supplier processes and estimates the capability of each process. The objective of this problem then consists of choosing the process that has the highest capability based on the observed estimates of process capability.

The following hypotheses are specifically of interest should be considered:

$$H_o : C_{pk_1} = C_{pk_2} = \cdots = C_{pk_N}$$
$$H_1 : C_{pk_i} \neq C_{pk_j}$$

$i \neq j$ for at least one pair of (i,j) where $i<j \in \{1, 2, \ldots, N\}$

and $C_{pk_i} = \min\left(\dfrac{USL - \mu_i}{3\sigma_i}, \dfrac{\mu_i - LSL}{3\sigma_i}\right)$ (3.1.1)

μ_i and σ_i are the mean and the standard deviation, respectively, of each of the supplier processes for measurement $\underline{Y_{i.}} = [Y_{i1}, Y_{i2}, \ldots, Y_{in_i}]'$. As mentioned in chapter 1, negative values of C_{pk_i} occur when the process average is positioned outside of the specification interval. This is an indication that the mean of process *i* lies outside one of the specification limits and over 50% of the distribution is outside the specifications.

C_{pk} indicates the location of the process average with respect to the specification limits. The C_{pk} index evaluates half of the process spread with respect to where the process is actually located (a point in space). As mentioned in Chapter 1, C_{pk} is not meaningful for a process which is not in statistical control. The PPM (parts per million nonconforming), which is often estimated, can be grossly wrong unless the process of interest is in statistical control.

Let Y_{ij} be observation *j* from a random sample of size n_i items from the i^{th} supplier, and define

$\overline{Y}_{i.} = \dfrac{\sum_{j=1}^{n_i} Y_{ij}}{n_i}$ and $s_i = \sqrt{\dfrac{\sum_{i=1}^{n_i} (Y_{ij} - \overline{Y}_{i.})^2}{n_i - 1}}$ as estimates of μ_i and σ_i respectively.

It is assumed that there is no restriction on sample sizes drawn from the N processes. A random sample in these cases is assumed to imply that the random variables observed from each process are mutually independent and identically distributed. Therefore, it is also assumed that all of the processes to be compared are in a state of statistical control.

In the example that follows, four suppliers are investigated. The assumption is that all four of the processes which are to be compared are in a state of statistical control. A further assumption is that all the samples are mutually independent of one another.

Piston ring example

Consider a company with $N = 4$ suppliers representing the four processes that produce piston rings for automobile engines (as studied by Chou, 1994). The edge width of a piston ring after the preliminary disk grind is a crucial quality characteristic in automobile engine manufacturing. Suppose that the automotive engineers have set the lower and upper specification limits of this quality characteristic to be $LSL = 2.6795mm$ and $USL = 2.7205mm$ respectively. The target (T) is assumed to be the midpoint (M) of this range (therefore $T = 2.7mm$). The four potential suppliers (supplier 1 to supplier 4) for such rings are under consideration by one quality control manager. There is no restriction on sample size and convenient sample sizes of sizes of $n_1 = 50$, $n_2 = 75$, $n_3 = 70$ and $n_4 = 75$ are taken from the manufacturing processes of suppliers 1 to 4, respectively. A summary of the results from the samples, calculated C_{pk} values and other statistics are given in table 3.1

Table 3.1: Summary of the observed process data obtained from the four potential suppliers of Piston rings

Supplier (i)	**1**	**2**	**3**	**4**
Sample size (n_i)	**50**	**75**	**70**	**75**
Estimated mean ($\overline{Y}_i$)	**2.7048**	**2.7019**	**2.6979**	**2.6972**
Estimated standard deviation (s_i)	**0.0034**	**0.0055**	**0.0046**	**0.0038**
Estimated classical index (C_{pk})	**1.5392**	**1.1273**	**1.3333**	**1.5526**

Source: (Polansky A.M., 2006).

Polansky (2006) only provides the sufficient statistics as displayed in table 3.1 and the histograms of the real data but not the data itself. To get a better understanding of the processes, the information in table 3.1 below is used to simulate data from each of the suppliers. The estimates of table 3.1 are taken as the population parameters and an assumption of normality is made. Figure 3.1 below is therefore only for illustration purposes and compares well with the histograms of the original data in figure 1 on page 265 of Polansky (2006).

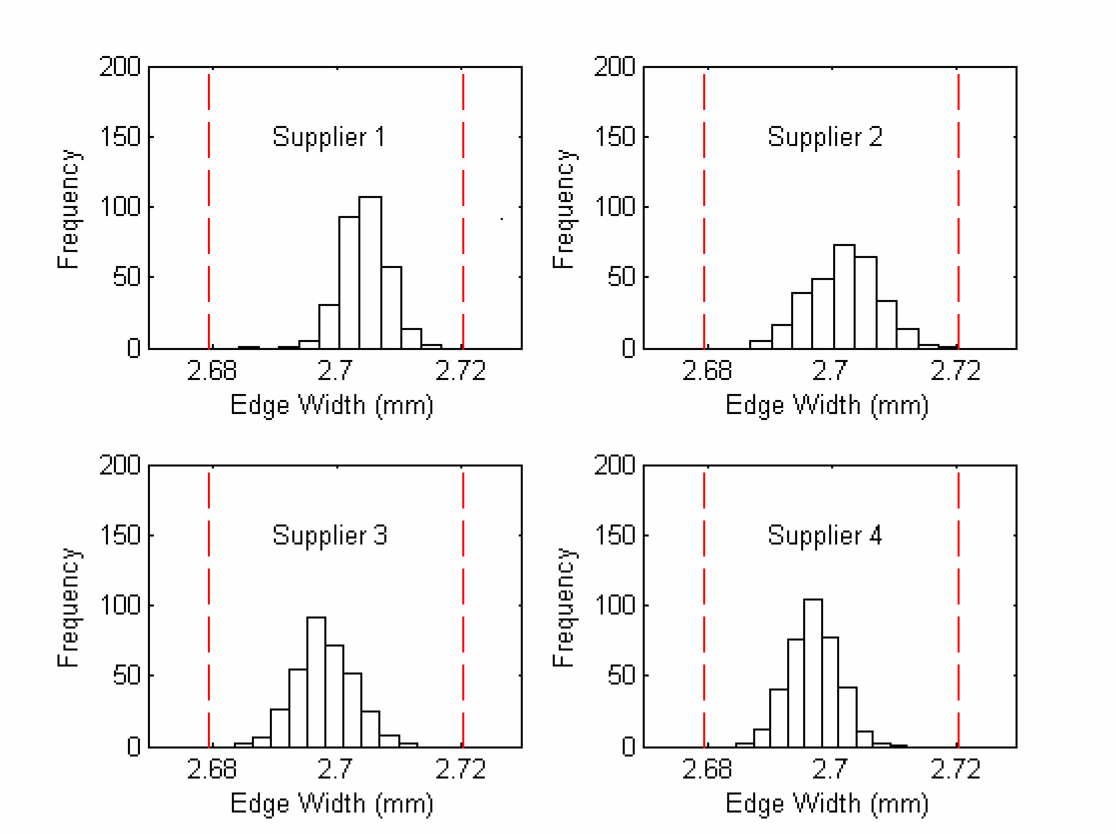


Figure 3.1: Frequency histograms of 1000 simulated edge width values from the four potential suppliers of piston rings

Looking at table 3.1 and figure 3.1 it is clear that suppliers 4 and 1 give the largest and second largest values of C_{pk} suggesting that they are the most capable. This may be because the data from suppliers 4 and 1 seems to have the smallest variation within the specification limits. In fact supplier 1's output has the smallest variance (0.0034) among the four suppliers, but the distribution of the output is not centred within the specification limits. Supplier 4's variance is slightly greater (0.0038) and the distribution of the output is centred within the specification limits. Looking at the index C_{pk}, suppliers 1 and 4 therefore represent the two best choices of suppliers. Suppliers 3 and 2 are not as capable as the former because of their greater variability. Because the estimated C_{pk} index for supplier 1 is close to that of supplier 4, one might feel that the difference in capability of the processes between these suppliers is not significant. The same statement may hold true of suppliers 2 and 3. Statistical methods for the comparison of the suppliers' process capability indices are required for the quality control manager to draw intelligent conclusion from this data.

A Bayesian simulation procedure is now considered to determine which of the supplier processes has the largest process capability index and which processes are significantly different from one another. The potential performance of the proposed method is compared with the permutation approach by Polansky (2006).

In this chapter, the Bayesian simulation based strategy for solving the supplier selection problem when there are two or more suppliers using Monte Carlo simulation is developed or adapted.

In the next section the Bayesian model used for the simulation and some commonly used terminology is reviewed. In section 3.3, a Bayesian simulation approach to the supplier selection problem is considered for C_{pk} as the process capability index. Section 3.4 investigates model checking using credibility intervals. Section 3.5 considers a Bayesian simulation approach to the supplier selection problem in the case of C_{pm} as the process capability index. Section 3.6 considers the index C_{pmk} in the context of the supplier selection problem. In Section 3.7 we consolidate the results from all the three indices. The piston rings example is used throughout the sections.

3.2 MODEL

As mentioned previously, capability analysis is designed to monitor the proportion of items which is expected to fall outside the engineering specification to prevent an excessive production of non-conforming output. According to Bernardo and Irony (1996) this is usually done at specified rating periods, to obtain a random sample $\underline{Y_{i.}} = [Y_{i1}, Y_{i2,} \ldots, Y_{in_i}]'$ of size n_i from the i^{th} process. In this investigation it is assumed that $\underline{Y_{i.}}\,(i = 1, \ldots, N)$ are independently identically normally distributed with mean μ_i and variance $\sigma_i^{\,2}$. Since both μ_i and $\sigma_i^{\,2}$ are unknown and no prior information is available, the conventional 'non–informative' default joint prior

$$p(\mu_i, \sigma_i^{\,2}) \propto \sigma_i^{\,-2} \tag{3.2.1}$$

will be specified.

Using (3.2.1), it is well known (see for example Zellner,1971) and also proof in chapter 2) that the conditional posterior density of μ_i is normal:

$$\mu_i \mid \sigma_i^{\,2}, \underline{Y_{i.}} \sim N\left(\overline{Y}_{i.}, \frac{\sigma_i^{\,2}}{n_i}\right) \tag{3.2.2}$$

and the posterior density for the variance component $\sigma_i^{\,2}$, is given by

$$p(\sigma_i^{\,2} \mid \underline{Y_{i.}}) = C(\sigma_i^{\,2})^{-\frac{1}{2}(n_i-1)-1} \exp\left\{-\tfrac{1}{2}(n_i-1)s_i^{\,2}/\sigma_i^{\,2}\right\} \quad \sigma_i^{\,2} > 0 \tag{3.2.3}$$

$$= IG(\sigma_i^2 \mid \frac{1}{2}(n_i-1)s_i^{\,2}, \frac{1}{2}(n_i-1))$$

an inverted gamma density, where $\underline{Y_{i.}} = [Y_{i1}, Y_{i2} \ldots, Y_{in_i}]'$, $\overline{Y_{i.}} = \frac{\sum_{j=1}^{n_i} Y_{ij}}{n_i}$, the sample mean,

$s_i^2 = \frac{\sum_{i=1}^{n_i} (Y_{ij} - \overline{Y}_{i}.)^2}{n_i - 1}$, the sample variance and the normalizing constant is

$C = \left\{\frac{(n_i-1)s_i^{\,2}}{2}\right\}^{\frac{1}{2}(n_i-1)} \frac{1}{\Gamma\left(\frac{n_i-1}{2}\right)}$ and

$$IG(y \mid \beta_i, \alpha_i) = \frac{\alpha^{\beta_i}}{\Gamma(\beta_i)} y^{-\beta_i-1} \exp(-\alpha_i/y)\,. \tag{3.2.4}$$

i.e. the inverted gamma density with positive parameters $\beta_i = \frac{1}{2}(n_i-1)s_i^2$ and $\alpha_i = \frac{1}{2}(n_i-1)$.

3.3 SIMULATION OF THE C_{pk} INDEX FOR THE DIFFERENT SUPPLIERS

Standard routines are used in the simulation procedure for $i = 1, \ldots, N$

1. By using the Matlab package, simulation of $\sigma_i^{\,2}$ can be obtained in the following way:

(a) Simulate τ from a $\chi^2_{n_i-1}$ distribution, as a sum of (n_i-1) squared independent standard normal random variates.

(b) Calculate $\sigma_i^{\,2*} = \dfrac{(n_i-1)s_i^{\,2}}{\tau}$. where (*) indicates a simulated value.

(c) $\sigma_i^{\,*} = \sqrt{\sigma_i^{\,2*}}$.

2. By making use of the fact that $\mu_i \,|\, \underline{Y_{i.}}, \sigma_i^{\,2} \sim N(\overline{Y}_{i.}, \dfrac{\sigma_i^{\,2}}{n_i})$, where $\underline{Y_{i.}}$ is the data drawn from process/supplier i, simulate μ_i^* and from the definition of the index, it follows that C_{pk} can be simulated as

$$C_{pk_i}^{\;*} = \min\left(\frac{USL-\mu_i^*}{3\sigma_i^{\,*}}, \frac{\mu_i^*-LSL}{3\sigma_i^{\,*}}\right).$$

3. Repeat steps (1-2) $\tilde{\ell}$ times. For this example, $\tilde{\ell}$ is taken as 1 000.

Note: This process simulates (or repeats) a simulation process!

The above procedure is used to simulate C_{pk} for each of the four suppliers. Part of the simulated values is presented in the table 3.2 below.

Table 3.2: Part of the 1000 simulated C_{pk} values from the four potential suppliers of piston rings

ℓ	**Supplier 1**	**Supplier 2**	**Supplier 3**	**Supplier 4**
1	**1.6629**	**1.0464**	**1.1783**	**1.2944**
2	**1.4689**	**0.9420**	**1.4004**	**1.6970**
3	**1.5118**	**1.0357**	**1.4988**	**1.4324**
4	**1.5221**	**1.0586**	**1.3420**	**1.5062**
5	**1.8913**	**1.0485**	**1.3220**	**1.4597**
6	**1.8638**	**1.2141**	**1.3394**	**1.5415**
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
1000	**1.3831**	**1.0732**	**1.3282**	**1.3914**

Histograms of the simulated C_{pk} values are drawn to illustrate the distribution of the capability index for each supplier. The mean index value is also calculated as

$$\overline{C}_{pk_i}{}^{*} = \frac{1}{1000}\sum_{l=1}^{1000} C_{pk_i}{}^{*(\ell)}.$$

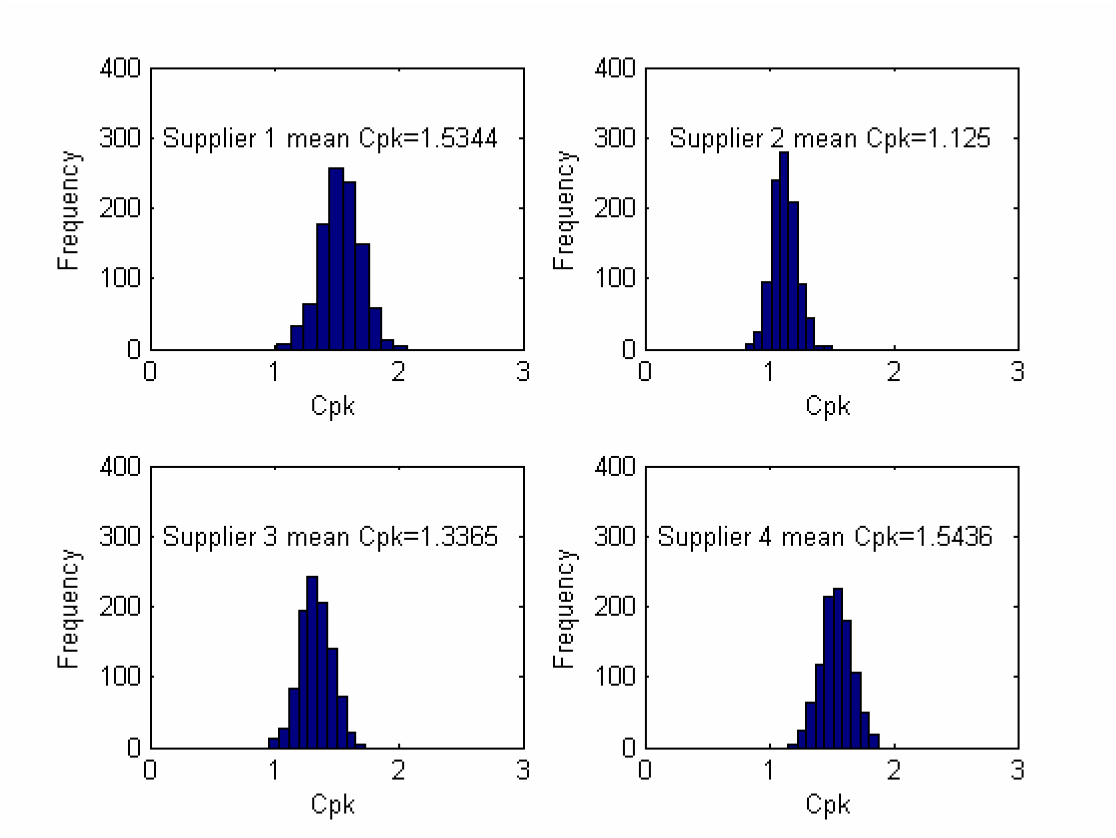


Figure 3.2: Frequency histograms of 1000 simulated C_{pk} values from the four potential suppliers of piston rings

Investigating Figure 3.2, it is again clear that suppliers 4 and 1 give the largest values of C_{pk}. Suppliers 3 and 2 are not as capable as the former. The mean values of the simulated indices are of the same magnitude as the classical results presented in table 3.1.

To calculate the probability that supplier 4, say, is the best or second best and so forth, we assign ranks to each simulation in a row in table 3.2. The highest $C_{pk_i}{}^{*(\ell)}$ in a row (row ℓ for example) is assigned the rank of 1, the second highest gets a rank of 2, the third highest value is assigned rank 3 and the least value gets rank 4. The table of ranks for C_{pk} values in table 3.2 would then look like table 3.3.

Table 3.3: Part of the 1000 rankings of Bayesian simulated C_{pk} values for the piston rings data

ℓ	Supplier 1	Supplier 2	Supplier 3	Supplier 4
1	1	4	3	2
2	2	4	3	1
3	1	4	2	3
4	1	4	3	2
5	1	4	3	2
6	1	4	3	2
⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮
1000	2	4	3	1

The results in table 3.3 can now be summarised to give table 3.4.

Table 3.4: Summary of the 1000 rankings of Bayesian simulated C_{pk} values for the piston rings data

Frequency	Supplier 1	Supplier 2	Supplier 3	Supplier 4
1's	455	0	52	493
2's	403	4	177	416
3's	131	103	678	88
4's	11	893	93	3

The required probabilities can now be calculated by dividing the above frequencies by 1 000. The probabilities are given in table 3.5.

Table 3.5: Probabilities that a given supplier is ranked 1st, 2nd, 3rd or 4th according to Bayesian simulated C_{pk} values for the piston rings data

Probability	Supplier 1	Supplier 2	Supplier 3	Supplier 4
Prob(Supplier$_i$=1)	0.455	0.000	0.052	0.493
Prob(Supplier$_i$=2)	0.403	0.004	0.177	0.416
Prob(Supplier$_i$=3)	0.131	0.103	0.678	0.088
Prob(Supplier$_i$=4)	0.011	0.893	0.093	0.003
Total	1.000	1.000	1.000	1.000

Suppliers 1 and 4 have the highest probabilities of being ranked 1st or 2nd by the quality control manager. Supplier 2 has a zero probability of being ranked 1st and a high probability of being ranked 4th (0.893). Supplier 3 is likely to be ranked 3rd (with probability 0.678).

3.3.1 MULTIPLE COMPARISON OF DIFFERENCES IN INDICES (C_{pk})

Polansky (2006) develops a strategy for solving the supplier selection problem when there are two or more suppliers using a methodology based on permutation tests. In the case of two processes, the method is based on a single simple permutation test. For the case of more than two processes, multiple comparison techniques are used in conjunction with permutation tests. The multiple comparison techniques that are used are:

i. The Bonferonni method, which adjusts the significance level of the pair-wise tests; and
ii. The protected multiple comparison method, which requires that an omnibus test of equality between all of the process capability indices be rejected before pair-wise tests are performed and does not require adjustment of the significance level of the pair-wise tests.

In the Bonferonni method, Polansky (2006) uses the test statistic

$D = \left| C_{pk_1} - C_{pk_2} \right|$ for, say, processes 1 and 2.

The author calculates the absolute difference between the process-capability indices estimated from the process data leading to table 3.1. The p-value for this test is given as

$$\tilde{p} = \frac{\#\{\tilde{D}_i > D\}}{n!}$$

where $n!$ is the number of distinct permutations of the n observations. $\#\{\tilde{D}_i > D\}$ is the number of times that $\tilde{D}_i$ exceeds D in the sequence $\tilde{D}_1, \tilde{D}_2, \ldots, \tilde{D}_{n!}$ and $\tilde{D}_i = \left| C_{pk_{1i}} - C_{pk_{2i}} \right|$ is the absolute difference between the process-capability indices computed from the i^{th} permutation of the observations in a combined sample from, say, processes 1 and 2, so that $n = n_1 + n_2$. $C_{pk_{1i}}$ is computed from the first n_1 observations in the i^{th} permutation and $C_{pk_{2i}}$ is computed from the remaining n_2 observations in the same permutation. Polansky then approximates the p-value for this test by a method that requires far fewer computations and is based on randomly generating b combinations of n_1 observations. For each randomly generated combination, the $\tilde{D}$ statistic is computed. Denote these values as $\tilde{D}_1^*, \tilde{D}_2^*, \cdots, \tilde{D}_b^*$. The resulting p-value for the test can then be approximated by

$$\tilde{p} = \tilde{p}_D^* = \frac{\#\{\tilde{D}_i^* > D\}}{b}.$$

For N processes the permutation is on $n = n_1 + n_2 + \cdots + n_N$ observations. Detailed discussions on the calculation of the p-value can be found in Polansky (2006).

A total of $v = \left(\frac{1}{2}\right) N(N-1)$ distinct pair wise comparisons of process capability indices among the N suppliers are possible. If each of the tests is performed with a significance level equal to α, it will result in at least one rejection in the v pair-wise comparisons with a probability of at most $v\alpha$. To control this error rate, the Bonferroni method uses a significance level equal to $\frac{\alpha}{v}$ for each pair-wise comparison. Therefore, the probability of at least one rejection in v pair-wise comparisons is at most α, when the null hypothesis $H_o : C_{pk_1} = C_{pk_2} = \cdots = C_{pk_N}$ is true.

In the protected multiple-comparison method for controlling overall error rate of the v pair-wise comparisons, the method is based on first performing a preliminary test of the null hypothesis $H_o : C_{pk_1} = C_{pk_2} = \cdots = C_{pk_N}$ versus the alternative hypothesis $H_1 : C_{pk_i} \neq C_{pk_j}$ for $i \neq j$ for at least one pair of (i,j) where $i<j \in \{1,2,\ldots, N\}$. The v pair-wise comparisons are performed only if the preliminary test rejects the null hypothesis at a significant level equal to α. If the preliminary test fails to reject the null hypothesis, then no comparisons are performed. Because the pair-wise comparisons are only performed when the preliminary test rejects the null hypothesis, the maximum probability of rejecting a pair-wise comparison when the null hypothesis $H_o : C_{pk_1} = C_{pk_2} = \cdots = C_{pk_N}$ is true is at most α.

In this chapter, the Bayesian simulation based on pair-wise multiple-comparison tests of equality between all of the suppliers' process capability indices is performed to solve the supplier selection problem. For the simulated C_{pk_i} (as illustrated in table 3.2) define $D_{ij}^{*(\ell)} = C_{pk_i}{}^{*(\ell)} - C_{pk_j}{}^{*(\ell)}$ for $\ell = 1:1000$ and $i \neq j$ and $i<j \in \{1,2,\ldots, N=4\}$. Below are histograms of simulated D_{ij} values for all pairs of suppliers. The largest differences are between suppliers 2 and 4 as well as 1 and 2.

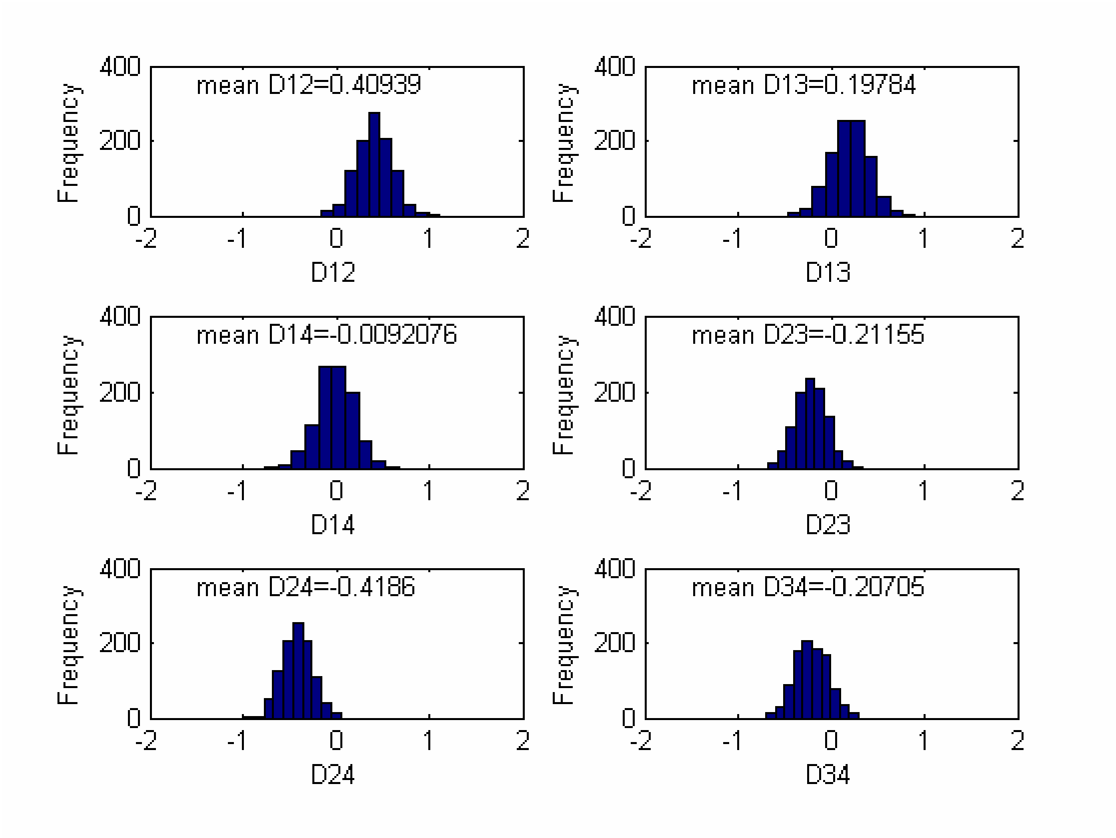


Figure 3.3: Frequency histograms of 1000 simulated differences $D_{ij}^{*} = C_{pk_i}^{*} - C_{pk_j}^{*}$ *values from the four potential suppliers of piston rings*

To construct the percentile credibility interval for the differences $D_{ij} = C_{pk_i} - C_{pk_j}$ we sort the $D_{ij}^{*(\ell)} = C_{pk_i}^{*(\ell)} - C_{pk_j}^{*(\ell)}$ values in ascending order so that:

$D_{ij}^{*(1)} \le D_{ij}^{*(2)} \le \cdots \le D_{ij}^{*(1000)}$ $i \neq j$ and $i<j \in \{1,2,\ldots,N=4\}$,

In this application, 1000 values of D_{ij} are sorted from smallest to greatest and the critical values are found by selecting the value in the position $(\frac{\alpha}{2})\times 1000$ as the lower bound and the value in the position $(1-\frac{\alpha}{2})\times 1000$ as the upper bound. The credibility interval is then constructed as $D_{ij}\left((\frac{\alpha}{2})\times 1000\right) - D_{ij}\left((1-\frac{\alpha}{2})\times 1000\right)$. The

95% credibility interval is $D_{ij}(25) - D_{ij}(975)$. The decision rule using these intervals to test $H_o : C_{pk_i} = C_{pk_j}$ vs. $H_1 : C_{pk_i} \neq C_{pk_j}$ is as follows:

Reject H_o if zero is contained in the interval mentioned above, otherwise accept.

The results of permutation based methods from the empirical study by Polansky (2006) are compared with the Bayesian simulation based methods using the same data set. The results are presented in table 3.6 below.

Table 3.6: Results of the six pair-wise comparisons of the C_{pk} indices of the piston rings example

Test Pair	**Observed D**	**p-value (Polansky method)**	**Simulated Bayesian mean D**	**95% credibility interval**
Supplier 1 vs. Supplier 2	**0.4119**	**0.0069**	**0.4094**	**(0.0385;0.7730)**
Supplier 1 vs. Supplier 3	**0.2059**	**0.1810**	**0.1978**	**(-0.2015;0.5738)**
Supplier 1 vs. Supplier 4	**0.0134**	**0.9323**	**-0.0092**	**(-0.4171;0.3879)**
Supplier 2 vs. Supplier 3	**0.2060**	**0.1333**	**-0.2116**	**(-0.5283;0.1083)**
Supplier 2 vs. Supplier 4	**0.4253**	**0.0034**	**-0.4186**	**(-0.7267;-0.1067)**
Supplier 3 vs. Supplier 4	**0.2193**	**0.2577**	**-0.2071**	**(-0.5517;0.1461)**

In interpreting the 95% credibility intervals, seen in table 3.6 above, one can observe that suppliers 1, 3, and 4 have process capabilities that are not significantly different since the intervals contain zero. Similarly, suppliers 2 and 3 are not significantly different from one another, but supplier 2 is significantly different from suppliers 1 and 4. The same conclusion is made using the p-values calculated by Polansky (2006) using $\alpha = 0.05$.

3.4 MODEL CHECKING USING RANDOM CREDIBILITY INTERVALS FOR C_{pk}

Once computing (using simulation) the posterior distributions of all our estimates is completed, the relatively easy step of assessing the fit of the model to the data should not be ignored. Checking the model is crucial in statistical analysis. Bayesian prior-to-posterior inferences assume the whole structure of a probability model and can yield false inferences when the model is invalid. The simulation procedures in this section illustrate the idea of constructing credibility intervals for the true value of the process capability indices.

A $100\times(1-\alpha)\%$ credibility interval for C_{pk_i} (for supplier i) is a random interval that contains C_{pk_i} with probability $(1-\alpha)$. The simulation will be done in the following manner.

Let the figures in table 3.1 be the true parameters for Supplier i, where the mean is μ_i and the standard deviation is σ_i and the index is calculated as $C_{pk_i} = \min\left(\dfrac{USL-\mu_i}{3\sigma_i}, \dfrac{\mu_i - LSL}{3\sigma_i}\right)$ as given in the table.

<u>First method</u>

In this section, x_{ij} $\{i=1,2,\ldots,N \quad j=1,2,\ldots,n_i\}$ will be used to denote a simulated value as opposed to the observed value y_{ij}.

Draw a sample x_{ij} $\{i=1,2,\ldots,N \quad j=1,2,\ldots,n_i\}$ of size n_i observations from a $N\left(\mu_i, \sigma_i^2\right)$ distribution.

a. Calculate $\overline{x_{i.}} = \dfrac{\sum_{j=1}^{n_i} x_{ij}}{n_i}$ and $s_i^2 = \dfrac{\sum_{i=1}^{n_i} (x_{ij} - \overline{x}_{i}.)^2}{n_i - 1}$.

b. At this stage we pretend that we do not know the value of μ_i, σ_i and C_{pk_i}.

c. Although we do not know what the parameters values (μ_i, σ_i and C_{pk_i}) are, we can simulate them (as before):

(i) $p(\sigma_i^2 \mid \underline{x_{i.}}) = C(\sigma_i^2)^{-\frac{1}{2}(n_i+1)} \exp\left\{-\tfrac{1}{2}(n_i-1)s_i^2/\sigma_i^2\right\} \quad \sigma_i^2 > 0$.

Simulate a τ from a $\chi^2_{n_i-1}$ distribution as a sum of (n_i-1) squared independent normal random deviates and calculate

$$\sigma_i^{2*} = \frac{(n_i-1)s_i^2}{\tau}.$$

(ii) Given σ_i^{2*}, simulate $\mu_i^* \sim N\left(\overline{x}_{i.}, \frac{\sigma_i^{2*}}{n_i}\right)$.

(iii) Substitute μ_i^* and σ_i^* in

$$C_{pk_i}^{\ *} = \min\left(\frac{USL-\mu_i^*}{3\sigma_i^*}, \frac{\mu_i^*-LSL}{3\sigma_i^*}\right)$$

and calculate it, also calculate $C_{pk_i}^{\ *} - C_{pk_j}^{\ *}$ for all $i \neq j$.

d. (i) Repeat steps (c) (i), (ii) and (iii) 1000 times. In other words you will get 1000 $C_{pk_i}^{\ *}$'s and hence1000 $C_{pk_i}^{\ *} - C_{pk_j}^{\ *}$'s for each $\overline{x}_{i.}$ and s_i^2.

(ii) from the simulated distributions calculate 95% credibility intervals for both C_{pk_i} and $C_{pk_i} - C_{pk_j}$ and see if it covers the **true** C_{pk_i} and $C_{pk_i} - C_{pk_j}$ according to table 3.1.

e. Draw another sample of n_i observations from a $\mathrm{N}\left(\mu_i, \sigma_i^2\right)$ distribution and repeat the whole procedure.

f. Draw 1000 samples. About 950 of these credibility intervals should cover the **true** capability index or the **true** difference in the indices (i.e. C_{pk_i} and $C_{pk_i} - C_{pk_j}$). It can also be stated in another way by saying that about 50 of these intervals should not cover the true parameter values.

Second method

In the first method the complete sample was simulated. It is not necessary to simulate the complete sample. The alternative to the simulation procedure above is to simulate the necessary sufficient statistics $\overline{x_{i.}}$ and s_i^2 only as follows:

Simulate $\overline{x_{i.}} \sim N\left(\mu_i, \frac{\sigma_i^2}{n_i}\right)$ and

$$\frac{(n_i - 1)s_i^2}{\sigma_i^2} \sim \chi^2_{n_i - 1}$$

where μ_i and σ_i^2 are assumed to be the true parameter values obtained from the statistics given in table 3.1.

Therefore $(n_i - 1)s_i^2 = \chi^2_{n_i-1}\sigma_i^2$ and $s_i^2 = \frac{\chi^2_{n_i-1}\sigma_i^2}{(n_i - 1)}$.

The results of the simulations from the first method are presented in tables 3.7 and 3.8 below. The simulations are done 1000 times and represent credibility intervals for both C_{pk_i} and $C_{pk_i} - C_{pk_j}$ from the simulated distributions.

Table 3.7: Summary of parameters from the four potential suppliers of piston rings and results of the proportion of credibility Intervals that do not contain the true index for the 1000 simulations

Supplier	μ_i	σ_i	**True value of the index** C_{pk_i}	**Proportion of credibility intervals that do not contain the true index**
1	2.7048	0.0034	1.5392	$\frac{52}{1000}$
2	2.7019	0.0055	1.1273	$\frac{53}{1000}$
3	2.6979	0.0046	1.3333	$\frac{42}{1000}$
4	2.6972	0.0038	1.5526	$\frac{58}{1000}$

Table 3.8: Results of the proportion of credibility intervals that do not contain the true difference for the 1000 simulations

Test pair	**True value of the difference parameter**	**Proportion of credibility intervals that do not contain the true difference**
$C_{pk_1} - C_{pk_2}$	0.4119	$\frac{51}{1000}$
$C_{pk_1} - C_{pk_3}$	0.2059	$\frac{46}{1000}$
$C_{pk_1} - C_{pk_4}$	-0.0134	$\frac{55}{1000}$
$C_{pk_2} - C_{pk_3}$	-0.2060	$\frac{50}{1000}$
$C_{pk_2} - C_{pk_4}$	-0.4253	$\frac{55}{1000}$
$C_{pk_3} - C_{pk_4}$	-0.2193	$\frac{48}{1000}$

Since probability models in most data analysis will not be perfectly true, the results look plausible. It therefore seems that the frequentist properties of Bayesian inferences of capability indices based on the prior $p(\mu_i, \sigma_i^{\,2}) \propto \sigma_i^{\,-2}$ are adequate for frequentist properties.

Below is a diagrammatic representation of one of the credibility intervals plot with the horizontal solid line representing the true value of the parameter (C_{pk_1}) for 100 simulations only.

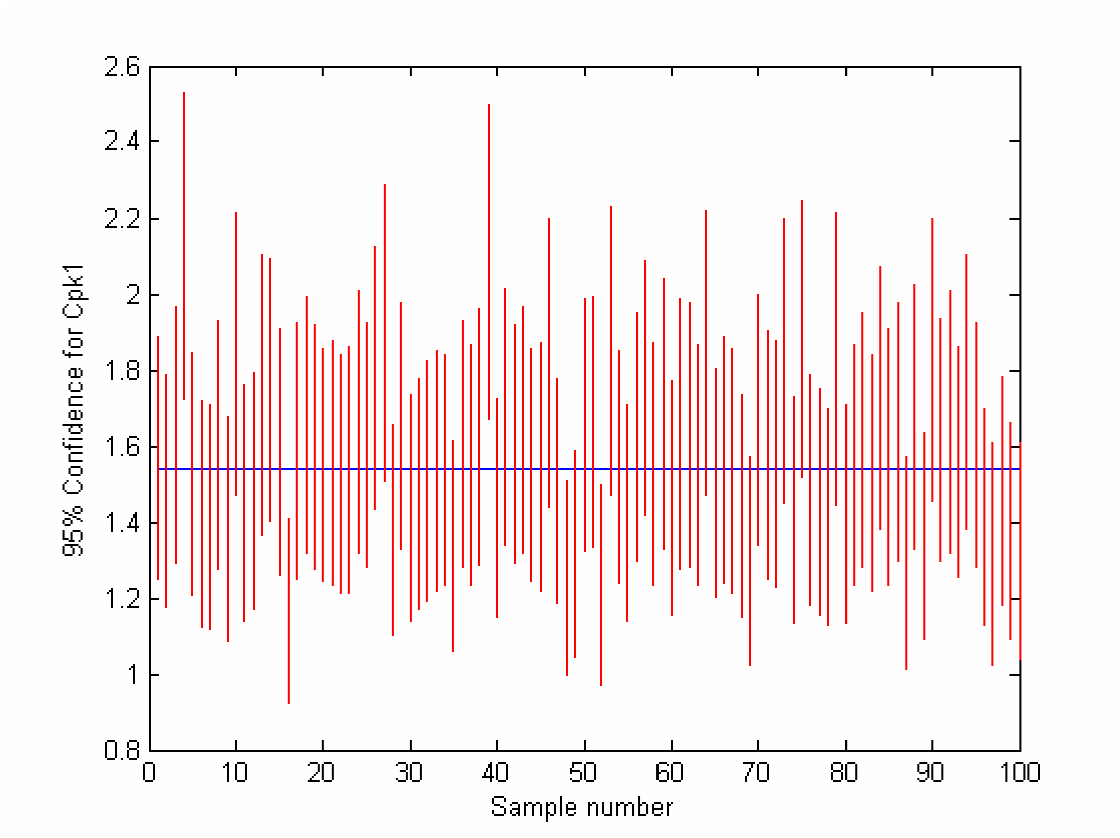


Figure 3.4: Results of the credibility intervals and the true index for 100 simulations

From the simulation studies (and because the piston rings example consists of large samples) it is clear that the prior $\pi(\mu,\sigma^2)\propto\sigma^{-2}$ works well.

3.5 C_{pm} INDEX

As noted in chapter 1, the quantification of process variation and target location is central to understanding the quality of units from a process. The concept of variation has recently undergone a paradigm shift in industry. This shift has occurred in the interpretation of the quality of product varying within the allowable process specification. The C_{pk} index discussed thus far uses the historical perspective of quality variation. A historical perspective of quality variation is that product has the same quality; that is to say that the product is equally good, regardless of where it falls within the specification limits. Product is considered bad, or lacking in quality, only if product falls outside of the specification limits. Engineers are comfortable with this notion of variation, which is sometimes referred to as the “goal post mentality”, and is displayed graphically in the figure 1.6 in chapter 1. As noted, the problem with this mentality is the step function that occurs directly at the specification limits. In regard to a process, the quality of a part falling just within the specification limit has little practical difference from the quality of a product falling just outside the specification limit. This model of quality variation has little relevance to industry.

Figure 1.7 in chapter 1 shows a model proposed by statisticians. This model is more practical in that the loss of quality and thus value loss to an organisation increases as the quality varies from a process target.

As noted in Chapter 1, this notion of variation referred to as “loss function mentality” states that there is a quadratic relationship between the loss and the distance from the target and it was proposed by Taguchi. This function is called the loss function curve and it ties variation to the loss in a process. This notion is what capability is based on. A capability index, according to Taguchi, enumerates a process’ ability to minimise the loss function curve.

As mentioned in chapter 1, Hsiang and Taguchi (1985) (and also Chan, Cheng, and Spiring, 1988) developed the index C_{pm} in order to take into account the process target and defined it as follows from (1.3.8)

$$C_{pm} = \frac{USL - LSL}{6\sqrt{\sigma^2 + (\mu - T)^2}}$$

$$= \frac{USL - LSL}{6\sqrt{E\{L(Y)\}}} \ . \qquad (3.5.1)$$

In the next section the Bayesian simulation tests are repeated, but this time using C_{pm} as the index of interest to underlie the importance of understanding the meaning of an index and what it measures. Specifically, the following hypothesis of interest in investigated:

$$H_o : C_{pm_1} = C_{pm_2} = \cdots = C_{pm_N}$$
$$H_1 : C_{pm_i} \neq C_{pm_j}$$

$i \neq j$ for at least one pair of (i,j) where $i<j$ $\{1, 2, \ldots, N\}$, where

$$C_{pm_i} = \frac{USL - LSL}{6\sqrt{\sigma_i^{\ 2} + (\mu_i - T)^2}} \qquad (3.5.2)$$

3.5.1 SIMULATION OF C_{pm} INDEX FOR THE DIFFERENT SUPPLIERS

Consider a sample y_{ij} $\{i = 1, 2, \ldots, N \quad j = 1, 2, \ldots, n_i\}$ of size n_i observations from a N(μ_i, σ_i^2) distribution.

Calculate $\overline{Y_{i.}} = \dfrac{\sum_{j=1}^{n_i} Y_{ij}}{n_i}$ and $s_i^2 = \dfrac{\sum_{i=1}^{n_i} (Y_{ij} - \overline{Y}_i.)^2}{n_i - 1}$.

Standard routines are used in the simulation procedure:

1. By using the Matlab package, simulation of $\sigma_i^{\ 2}$ can be obtained in the following way:

   Simulate τ from a $\chi^2_{n_i - 1}$ distribution, as a sum of $(n_i - 1)$ squared independent standard normal random variates.

   Calculate $\sigma_i^{\ 2*} = \dfrac{(n_i - 1)s_i^{\ 2}}{\tau}$, where (*) indicates a simulated value.

2. By making use of the fact that $\mu_i \mid \underline{Y_i}, \sigma_i^{\ 2} \sim N(\overline{Y}_i, \dfrac{\sigma_i^{\ 2}}{n_i})$, where $\underline{Y_i}$ is data drawn from process/supplier i, simulate μ_i^* and from the definition of the capability index, it follows that C_{pm_i} can be simulated as $C_{pm_i}^{\ *} = \dfrac{USL - LSL}{6\sqrt{\sigma_i^{\ 2*} + (\mu_i^* - T)^2}}$.

3. Repeat steps (1-2) $\tilde{\ell}$ times. For our example, $\tilde{\ell}$ was taken as 1 000.

Histograms of the simulated C_{pm} values are drawn to illustrate the distribution of this capability index for each supplier. See figure 3.5.

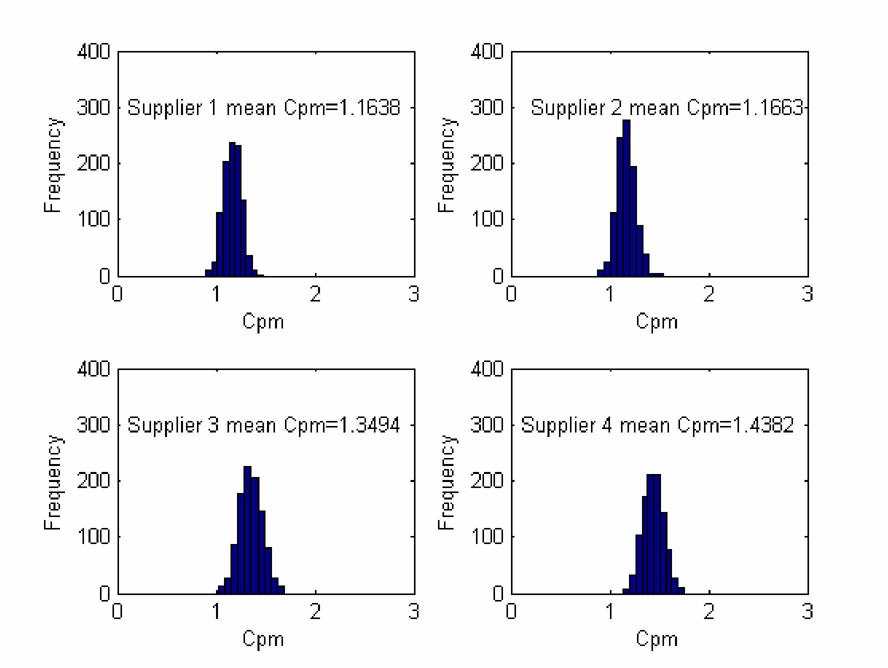


Figure 3.5: Frequency histograms of 1000 simulated C_{pm} values from the four potential suppliers of piston rings

Looking at figure 3.5, suppliers 4 and now 3 give the largest values for C_{pm}. The mean Bayesian simulated C_{pm} index for suppliers 4 and 3 are 1.4382 and 1.3494 respectively. It is previously mentioned that suppliers 4 and 1 have the largest C_{pk} index value. Suppliers 2 and 1 are not as capable as 3 and 4 according to the index C_{pm}. In fact, supplier 1 is the least capable according to C_{pm} index.

The probability each supplier is ranked as the best or second best and so forth, are given in table 3.9 below. To calculate the probability that supplier 4, say, is the best or second best and so forth according to this index, we assign ranks to each simulation in a row as in table 3.2. The highest $C_{pm_i}^{*(\ell)}$ in a row (row ℓ for example) is assigned the rank of 1, the second highest gets a rank of 2, the third highest value is assigned rank 3 and the least value gets rank 4. The total frequencies for the ranks of C_{pm_i} are then divided by 1000 to give probability values in table 3.9.

Table 3.9: Probabilities that a given supplier is ranked 1^{st}, 2^{nd}, 3^{rd} or 4^{th} according to simulated C_{pm} values for the piston rings data

Probability	Supplier 1	Supplier 2	Supplier 3	Supplier 4
Prob(Supplier$_i$=1^{st})	**0.004**	**0.011**	**0.291**	**0.694**
Prob(Supplier$_i$=2^{nd})	**0.078**	**0.097**	**0.550**	**0.275**
Prob(Supplier$_i$=3^{rd})	**0.443**	**0.413**	**0.114**	**0.030**
Prob(Supplier$_i$=4^{th})	**0.475**	**0.479**	**0.045**	**0.001**
Total	**1.000**	**1.000**	**1.000**	**1.000**

Supplier 4 has the highest probabilities (0.694) of being ranked 1^{st} by the quality control manager. Supplier 3 has the second highest probabilities (0.291) of being ranked 1^{st} and the highest probability (0.550) of being ranked 2^{nd} by the quality control manager. Supplier 1 and 2 are likely to be ranked 3^{rd} and 4^{th} according to this index.

3.5.2 MULTIPLE-COMPARISONS OF DIFFERENCES IN INDICES(C_{pm})

The Bayesian simulation based on pair-wise multiple-comparison tests of equality between all of the suppliers' process capability indices is also performed to solve the supplier selection problem. For the simulated C_{pm_i}, we now define $D_{ij}^{*(\ell)} = C_{pm_i}^{*(\ell)} - C_{pm_j}^{*(\ell)}$ for $\ell = 1:1000$ and $i \neq j$ and $i<j \in \{1, 2, \ldots, N = 4\}$ Histograms of the simulated results are given in figure 3.6.

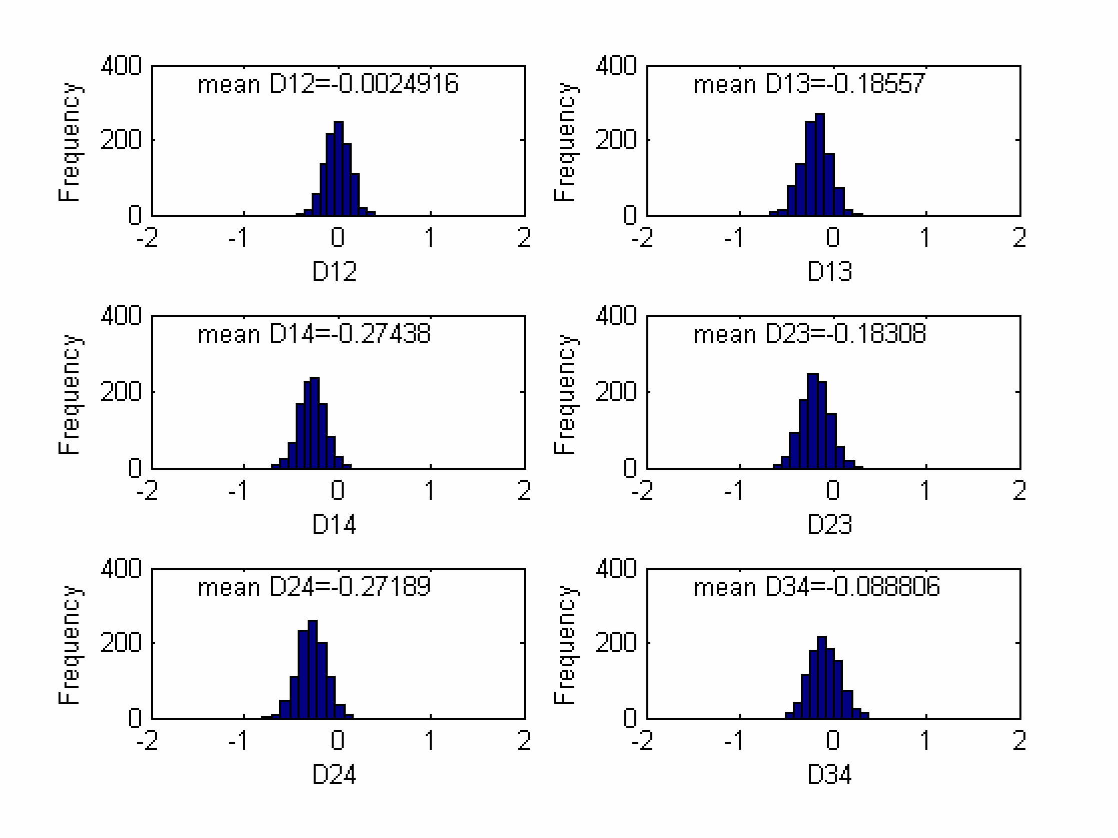


Figure 3.6: Frequency histograms of 1000 simulated differences $D_{ij}^* = C_{pm_i}^* - C_{pm_j}^*$ *values from the four potential suppliers of piston rings*

The largest difference is now between suppliers 1 and 4. The percentile credibility interval for the differences $D_{ij} = C_{pm_i} - C_{pm_j}$ is given in table 3.10.

Table 3.10: Results of the six pair-wise comparisons of the C_{pm} *indices of the piston rings example*

Test pair	**Simulated Bayesian mean** D_{ij}	**95% credibility Interval**
Supplier 1 vs. Supplier 2	**- 0.0025**	**(-0.2590;0.2326)**
Supplier 1 vs. Supplier 3	**-0.1856**	**(-0.4766;0.0981)**
Supplier 1 vs. Supplier 4	**-0.2744**	**(-0.5370;-0.0174)**
Supplier 2 vs. Supplier 3	**-0.1831**	**(-0.4880;0.1202)**
Supplier 2 vs. Supplier 4	**-0.2719**	**(-0.5467;0.0019)**
Supplier 3 vs. Supplier 4	**-0.0888**	**(-0.3814;0.2172)**

In interpreting the 95% credibility intervals, seen in table 3.10 above, one can observe that suppliers 1, 2 and 3 have process capabilities that are not significantly different since the intervals contain zero. Similarly, suppliers 2 and 4, 3 and 4 are not significantly different from one another, but supplier 1 is significantly different from supplier 4.

3.6 C_{pmk} INDEX

As discussed in chapter 1, whereas the index C_{pm} has the attractive features that it incorporates the parameters d, μ, T, and σ, it has an important omission, namely, the parameter M (the midpoint of the specification range). The index C_{pmk} rectifies this deficiency. To devise an index that is more sensitive to departures of μ from T, Pearn *et al.* (1992) introduces another process capability index, C_{pmk}. The index takes its numerator from C_{pk} and its denominator from C_{pm}, hence it is a hybrid.

$$C_{pmk} = \frac{\min(USL-\mu, \mu-LSL)}{3\sqrt{\sigma^2+(\mu-T)^2}} \quad (3.6.1)$$
.

It can be shown that this index can be re-written as

$$C_{pmk} = \frac{d-|\mu-M|}{3\sqrt{\sigma^2+(\mu-T)^2}} \quad (3.6.2)$$

where

$M = \dfrac{USL+LSL}{2}$ and $d = \dfrac{\text{USL-LSL}}{2}$

3.6.1 SIMULATION OF C_{pmk} INDEX FOR THE DIFFERENT SUPPLIERS

Standard routines are used in the simulation procedure:

1. Simulate τ from a $\chi^2_{n_i-1}$ distribution, as a sum of (n_i-1) squared

independent standard normal random variates.

Calculate $\sigma_i^{2*} = \dfrac{(n_i - 1)s_i^2}{\tau}$. where (*) indicates a simulated value.

2. By making use of the fact that $\mu_i \mid \underline{Y_i}, \sigma_i^2 \sim N(\overline{Y}_i, \dfrac{\sigma_i^2}{n_i})$, where $\underline{Y_i}$ is data drawn from process/supplier i, simulate μ_i^* and from the definition of the performance index, it follows that C_{pm_i} can be

Simulated as $C_{pmk_i}^{\ *} = \min\left(\dfrac{USL - \mu_i^*}{3\sqrt{\sigma_i^{2*} + \left(\mu_i^* - T\right)^2}}, \dfrac{\mu_i^* - LSL}{3\sqrt{\sigma_i^{2*} + \left(\mu_i^* - T\right)^2}}\right)$.

3. Repeat steps (1-2) $\tilde{\ell}$ times. For our example, $\tilde{\ell}$ was taken as 1 000.

Histograms of the simulated C_{pmk} values are drawn to illustrate the distribution of this capability index for each supplier. See figure 3.7.

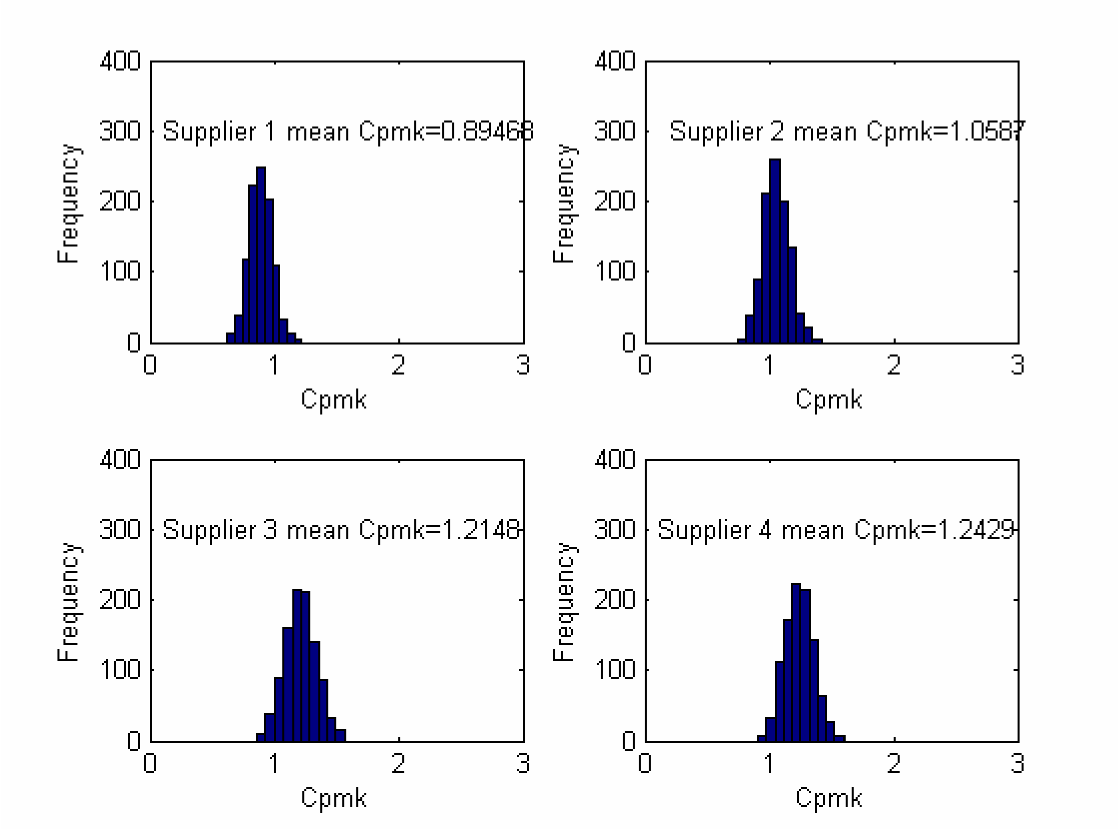


Figure 3.7: Frequency histograms of 1000 simulated C_{pmk} values from the four potential suppliers of piston rings

From figure 3.7, suppliers 4 and 3 give the largest and second largest values of C_{pmk} respectively. Supplier 2 is not as capable as the former. Supplier 1 is incapable according to this index. It gives a mean index value much less than 1. The probabilities each supplier is ranked as the best or second best and so forth, using the C_{pmk} index, are given in table 3.11 below.

Table 3.11: Probabilities that a given supplier is ranked 1^{st}, 2^{nd}, 3^{rd} or 4^{th} according to simulated C_{pmk} values for the piston rings data

Probability	**Supplier 1**	**Supplier 2**	**Supplier 3**	**Supplier 4**
Prob(Supplier$_i$=1)	**0.000**	**0.049**	**0.404**	**0.547**
Prob(Supplier$_i$=2)	**0.009**	**0.199**	**0.444**	**0.348**
Prob(Supplier$_i$=3)	**0.135**	**0.625**	**0.140**	**0.100**
Prob(Supplier$_i$=4)	**0.856**	**0.127**	**0.012**	**0.005**
Total	**1.000**	**1.000**	**1.000**	**1.000**

Undoubtedly, supplier 1 is now ranked as the worst performer. The probability that supplier 1 is ranked first and second is now zero and nearly zero (0.009) respectively. Supplier 4 still has the highest probabilities of being ranked 1^{st}. Supplier 3 has the second highest probabilities of being ranked 1^{st} and the highest probability (0.444) of being ranked 2^{nd} by the quality control manager. Supplier 2 is likely to be ranked 3^{rd} according to this index.

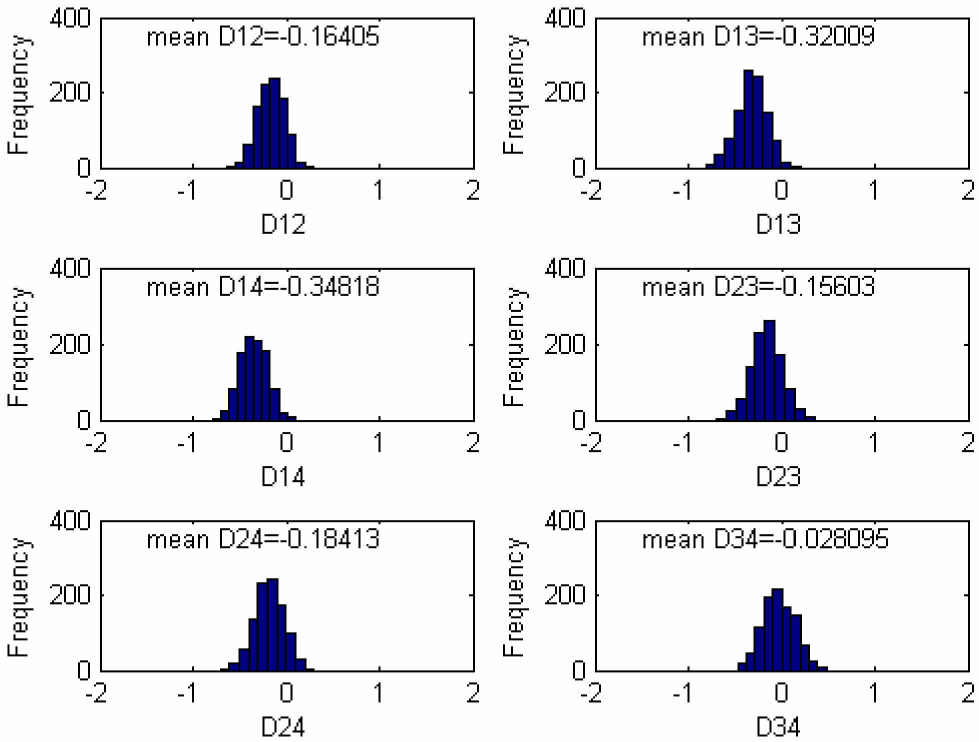


Figure 3.8: Frequency histograms of 1000 simulated differences $D_{ij}^{*} = C_{pmk_i}^{*} - C_{pmk_j}^{*}$ *values from the four potential suppliers of piston rings*

The differences are defined as $D_{ij} = C_{pmk_i} - C_{pmk_j}$. The largest difference is among suppliers 1 and 3, 1 and 4.

Table 3.12: Results of the six pair-wise comparisons of the C_{pmk} *indices of the piston rings example*

Test pair	**Simulated Bayesian mean D**	**95% credibility interval**
Supplier 1 vs. Supplier 2	**-0.1641**	**(-0.4308; 0.0917)**
Supplier 1 vs. Supplier 3	**-0.3201**	**(-0.6425; -0.0266)**
Supplier 1 vs. Supplier 4	**-0.3482**	**(-0.6173; -0.0718)**
Supplier 2 vs. Supplier 3	**-0.1560**	**(-0.4990; 0.1711)**
Supplier 2 vs. Supplier 4	**-0.1841**	**(-0.4856; 0.1228)**
Supplier 3 vs. Supplier 4	**-0.0281**	**(-0.3680; 0.3023)**

The 95% credibility intervals given in table 3.12 above leads one to conclude that supplier 1 is significantly different from suppliers 3 and 4 since the intervals do not contain zero. All the other intervals contain zero.

3.7 WHY DO SUPPLIER SELECTION RESULTS FOR C_{pk}, C_{pm} AND C_{pmk} DIFFER?

It would seem that the solution to the supplier selection problem depends on which index is selected. To understand why the results differ, it is necessary to look again at the indices C_{pk}, C_{pm} and C_{pmk} from a classical point of view.

$$C_{pk} = \min\left(\frac{USL-\mu}{3\sigma}, \frac{\mu - LSL}{3\sigma}\right).$$

Because the process average is part of the calculation, some believe this formula incorporates process centring. This is an erroneous assumption, because you do not know how far the process average is from the target. C_{pk} indicates where the process average is, but does not cover process centring (targeting). The C_{pk} index evaluates half the process spread with respect to where the process is actually located (some point in space). C_{pk} offers the most information about the proportion non-conforming, say p, and it will be shown later that it provides the least insight about the location of μ.

Boyles (1991) pointed out that C_{pk} does not say anything about the distance between process mean and target value and is essentially a measure of process potential only. He showed that C_{pk} becomes arbitrarily large as σ approaches 0, irrespective of where the process is centred and this characteristic makes C_{pk} unsuitable as a measure of process centring.

C_{pm} is related to the index C_p which is defined as $C_p = \dfrac{USL-LSL}{6\sigma^2}$.

Notice that:

$$C_{pm} = \frac{\frac{USL - LSL}{6\sigma^2}}{\sqrt{1+\left(\frac{\mu - T}{\sigma}\right)^2}}$$

$$i.e.\, C_{pm} = \frac{C_p}{\sqrt{1+\left(\frac{\mu - T}{\sigma}\right)^2}} \quad (3.7.1)$$

.

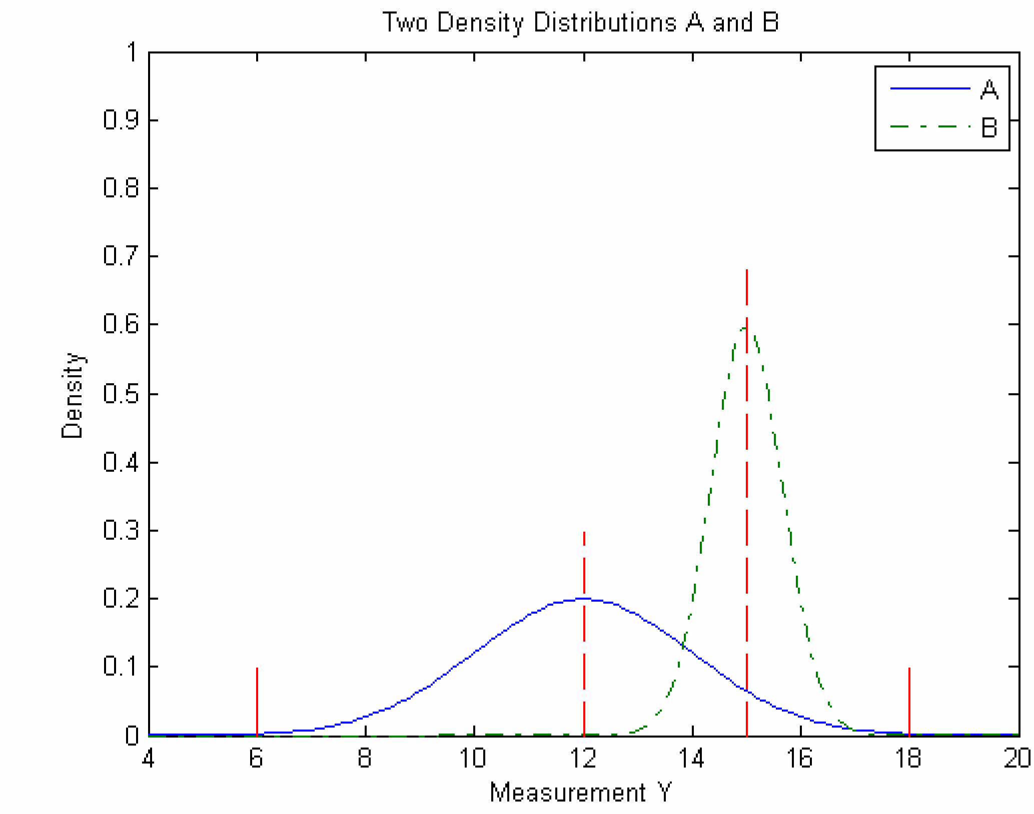
If $\mu = T$ (process is target centred) then $C_{pm} = C_p$.

$$C_{pmk} = \frac{\min(\frac{USL - \mu}{3\sigma}, \frac{\mu - LSL}{3\sigma})}{\sqrt{1+\left(\frac{\mu - T}{\sigma}\right)^2}}$$

$$C_{pmk} = \frac{C_{pk}}{\sqrt{1+\left(\frac{\mu - T}{\sigma}\right)^2}} \quad . \quad (3.7.2)$$

And when μ is equal to T, C_{pmk} is equal to C_{pk}.

The index C_{pm} does not directly relate to the percentage of non-conforming product, p, If p is regarded as the most important quality aspect of the process, this is definitely the *wrong* capability index to use. To illustrate this point we consider the following example involving C_{pk}, C_{pm} and C_{pmk} given in Bothe‘s (2002) discussion paper using classical statistics.



LSL=6 and USL=18

Figure 3.9: A change in process output from A to B

Consider a process (A) with an average $\mu = 12$, standard deviation of $\sigma = 2$ (distribution represented by solid line in figure 3.9) and target T =12. The lower and upper specifications limits are respectively 6 and 18.

$$
\begin{aligned}
C_{pk} &= \min\left(\frac{USL-\mu}{3\sigma}, \frac{\mu-LSL}{3\sigma}\right) \\
&= \min\left(\frac{18-12}{3(2)}, \frac{12-6}{3(2)}\right) \\
&= \min\left(\frac{6}{6}, \frac{6}{6}\right) \\
&= 1.00 .
\end{aligned}
$$

$$C_{pm} = \frac{USL - LSL}{6\sqrt{\sigma^2 + (\mu - T)^2}}$$

$$= \frac{18 - 6}{6\sqrt{2^2 + (12 - 12)^2}}$$

$$= \frac{12}{12}$$

$$= 1.00 \, .$$

$$C_{pmk} = \min\left(\frac{USL - \mu}{3\sqrt{\sigma^2 + (\mu - T)^2}}, \frac{\mu - LSL}{3\sqrt{\sigma^2 + (\mu - T)^2}} \right)$$

$$= \min\left(\frac{18 - 12}{3\sqrt{2^2 + (12 - 12)^2}}, \frac{12 - 6}{3\sqrt{2^2 + (12 - 12)^2}} \right)$$

$$= \min\left(\frac{6}{6}, \frac{6}{6} \right)$$

$$= 1.00 \, .$$

All the above indices indicate that the process is just capable.

A modification is made to this process and the output changes to a new distribution B, with the following parameters: an average $\mu = 15$ and a standard deviation of $\sigma = 0.667$ (distribution represented by broken line in figure 3.9). There are now fewer non-conforming parts, but also fewer parts produced at the target value. Did this change improve process capability? C_{pk}, C_{pm} and C_{pmk} are re-calculated.

$$C_{pk} = \min\left(\frac{USL - \mu}{3\sigma}, \frac{\mu - LSL}{3\sigma} \right)$$

$$= \min\left(\frac{18 - 15}{3(0.667)}, \frac{15 - 6}{3(0.667)} \right)$$

$$= \min\left(\frac{3}{2}, \frac{9}{2} \right)$$

$$= 1.50 \, .$$

$$C_{pm} = \frac{USL - LSL}{6\sqrt{\sigma^2 + (\mu - T)^2}}$$

$$= \frac{18-6}{6\sqrt{(0.667)^2 + (15-12)^2}}$$

$$= 0.65 \, .$$

$$C_{pmk} = \min\left(\frac{USL - \mu}{3\sqrt{\sigma^2 + (\mu - T)^2}}, \frac{\mu - LSL}{3\sqrt{\sigma^2 + (\mu - T)^2}} \right)$$

$$= \min\left(\frac{18-15}{3\sqrt{(0.667)^2 + (15-12)^2}}, \frac{15-6}{3\sqrt{(0.667)^2 + (15-12)^2}} \right)$$

$$= \min\left(\frac{3}{9.220}, \frac{9}{9.220} \right)$$

$$= 0.325 \, .$$

The C_{pk} index indicates that the process improved by jumping from 1.00 to 1.50. However, the C_{pm} and C_{pmk} indices indicate that the process has worsened by dropping from 1.00 to 0.65 and 0.325 respectively. The index C_{pmk} inflicts penalties whenever the sample mean of a rating period deviates from the target, T, and also from the mid-point M .

Was the change helpful or harmful to the quality level of this process? The answer depends on whether the historical perspective of variation or the loss function approach is used. Those companies concerned mainly with making parts and reducing p would claim that B's performance is better than A's. Those whose primary concern is making every part on target will believe that B is worse than A. It can now be appreciated that process improvement must be driven by more than the need to improve an index number, otherwise management may be wasting time and money. Most literature would simply suggest that management must choose the 'correct' index for their application or process. Each index tells you something different and unless you know what they measure, you may end up using the wrong index and making the wrong decisions.

Boyles (1991) shows that for fixed μ, the index C_{pm} is bounded above when σ tends to 0 and, furthermore, that $C_{pm} < \frac{d}{(3|\mu - T|)}$ and hence $|\mu - T| < \frac{d}{3C_{pm}}$ and where $d = \frac{USL - LSL}{2}$. This inequality can be interpreted as, a C_{pm}-value of 1 implies that the process mean, μ, lies within the middle third of the specification interval, and in general, it lies within the middle $1/3C_{pm}$ of the specification interval i.e. $d/3C_{pm}$.

Therefore, given a C_{pm} index of 1.00, it becomes known that $M - \frac{d}{3} < \mu < M + \frac{d}{3}$ where $M = \frac{USL + LSL}{2}$. This interval is much smaller than the one for C_{pk} equal to 1.00 which is equal to $M - d < \mu < M + d$.

Parlar and Wesolowsky (1999) notes that if $\mu = T$, then the three basic PCIs C_{pk}, C_p, C_{pm} are connected by the relationship

$$C_{pk} = C_p - \frac{1}{3}\sqrt{\left(\frac{C_p}{C_{pm}}\right)^2 - 1}.$$

When $T = M$ and for fixed μ, the index C_{pmk} is bounded above when σ tends to 0 and that $C_{pmk} < \frac{d}{(3|\mu - T|)} - \frac{1}{3}$ or $|\mu - T| < \frac{d}{(1 + 3C_{pmk})}$.

This inequality can be interpreted as follows: a C_{pmk}-value of 1 implies that the process mean μ, lies within the middle fourth of the specification range i.e. $M - \frac{d}{4} < \mu < M + \frac{d}{4}$. In general, the process mean, μ, lies within the middle $\frac{1}{(1 + 3C_{pmk})}$ of the specification range i.e. $\frac{d}{(1 + 3C_{pmk})}$, when $T = M$.

C_{pmk} is certainly worse than C_{pk} for being associated with a certain percentage of non-conforming products, but again, one should not choose this index if p is the main interest. C_{pmk} (and usually C_{pm}) is much more sensitive than other capability indices to movements in the process average relative to M. As seen in (3.7.2), when μ is equal to T, C_{pmk} is equal to C_{pk}. If μ moves away from T, however, C_{pmk} decreases more rapidly than does C_{pk} (although both are zero when μ is equal to one of the specification limits). Conversely, when μ is brought closer to T, C_{pmk} increases much faster than does C_{pk}. C_{pmk} reveals the most information about the location of the process average and the least about the proportion non-conforming p.

Vannman (1995) shows that among all the indices presented thus far, C_{pmk} is the most sensitive to departures of μ from T. The ranking of the following four basic indices discussed thus far in terms of sensitivity to departure of the process mean from the target value, from the most sensitive to the least sensitive are (1) C_{pmk}, (2) C_{pm}, (3) C_{pk} and (4) C_p.

A further interesting relationship among the indices given in Kotz and Johnson (2002) is derived as follows:

Since $C_{pmk} = \dfrac{C_{pk}}{\sqrt{1+\left(\dfrac{\mu-T}{\sigma}\right)^2}}$

$$= \frac{\dfrac{USL-LSL}{6\sigma^2}}{\sqrt{1+\left(\dfrac{\mu-T}{\sigma}\right)^2}} \frac{C_{pk}}{\dfrac{USL-LSL}{6\sigma^2}}$$

$$\therefore C_{pmk} = C_{pm}\frac{C_{pk}}{C_p} = \frac{C_{pm}C_{pk}}{C_p} \qquad (3.7.3)$$

The question is: why do the results depend on which index was used? The answer lies in the fact that the index C_{pmk} inflicts penalties whenever the sample mean of a rating period deviates from the target, T, and also from the mid-point M. The results suggest that supplier 1's process is not target centred. This is evidenced by the dramatic change in the index values from C_{pk} to C_{pm} and C_{pmk}.

The current chapter focused on C_{pk}, C_{pm} and C_{pmk} process capability indices, but virtually any other normally based index could have been used. The accuracy of capability indices is dependent on proper understanding of the theory behind the indices as well as an understanding of variation and centring.

CHAPTER 4

BAYESIAN ESTIMATION OF THE LOWER PROCESS CAPABILITY INDEX C_{pl}

4.1 INTRODUCTION

In this chapter, theoretical and simulation results are derived for the posterior distribution of the process capability index C_{pl} for two different but related prior distributions. For the conventional prior the exact posterior moments of the index are calculated. By using these moments, Pearson curve and Cornish-Fisher approximations of the posterior distribution are obtained for a real problem. Gibbs sampling is used in the case of the reference (probability matching) prior to obtain the unconditional posterior distribution of C_{pl}. Finally it is shown that the probability matching prior for the distribution function (a function of the index) is the same as the probability matching prior of the index itself.

4.2 INDEX AND NOTATION

One of the commonly used process capability indices is

$C_{pk} = \min\left(\frac{\ell_1 - \mu}{3\sigma}, \frac{\mu - \ell_0}{3\sigma}\right)$ with $\ell_1 = USL \equiv$ upper specification limit,

$\ell_0 = LSL \equiv$ lower specification limit, μ is the process mean and, σ is the process standard deviation.

According to Bernardo and Irony (1996) C_{pk} is the normalised distance between the process mean and its closest specification limit.

The definition of C_{pk} includes, as special cases, those processes where only one limit exists, by setting either $\ell_0 \to -\infty$ or $\ell_1 \to \infty$, in which case it reduces to the appropriate standardised measure. Thus if there is no lower specification limit, we obtain $C_{pu} = \frac{(\ell_1 - \mu)}{3\sigma}$ by simply letting $\ell_0 \to -\infty$ in the original definition. Similarly, if there is no upper specification limit, $C_{pl} = \frac{(\mu - \ell_0)}{3\sigma}$ (see for example Boyles, 2002).

The next section shows how the exact posterior moments (mean, variance, third and fourth central moments) are derived for the process capability index C_{pl}. By applying these moments and Pearson curves or Cornish–Fisher expansions, it will be shown in section 4.4 that approximations of the exact posterior distribution of C_{pl} can be obtained. The Monte Carlo simulation procedure for estimating the posterior distribution of C_{pl} is dealt with in section 4.5. Section 4.6 deals with a difference between two lower process capability indices. Section 4.7 investigates the probability matching and the reference priors for C_{pl}. Gibbs sampling is discussed in section 4.8. Section 4.9 looks at a probability matching prior for the parameter $\delta = \Phi\left(\frac{\mu - \ell_0}{\sigma}\right)$ while section 4.10 looks at an application, section 4.11 looks at another application and section 4.12 concludes.

4.3 EXACT POSTERIOR MOMENTS OF THE LOWER PROCESS CAPABILITY INDEX C_{pl}

As mentioned previously, capability analysis is designed to monitor the proportion of items which are expected to fall outside the engineering specification to prevent an excessive production of non-conforming output. According to Bernardo and Irony (1996) this is usually done at specified rating periods to obtain a random sample $Y_1, Y_2, \ldots, Y_n$ from the process. In this chapter, it will be assumed that $Y_i (i = 1, \ldots, n)$ are independently identically normally distributed with mean μ and variance σ^2. Since both μ and σ^2 are unknown and no prior information is available, the conventional 'non–informative' default prior

$$\pi(\mu, \sigma^2) \propto \sigma^{-2} \tag{4.3.1}$$

will be specified for them in this section. In chapter 2, while using (4.3.1), it becomes clear that the conditional posterior density of μ is normal (see for example Zellner, 1971):

$$\mu \mid \sigma^2, \underline{Y} \sim N\left(\overline{Y}, \frac{\sigma^2}{n}\right) \tag{4.3.2}$$

and the posterior density for the variance component σ^2, is given by

$$p(\sigma^2 \mid Y_1, Y_2, \ldots Y_n) = C(\sigma^2)^{-\frac{1}{2}(n-1)-1} \exp\left\{-\tfrac{1}{2}(n-1)s^2 / \sigma^2\right\} \tag{4.3.3}$$

an inverted gamma density, where $\underline{Y} = \left[Y_1, Y_2, \cdots Y_n\right]'$, $\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$ the sample mean,

$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \overline{Y})^2$, the sample variance and the normalizing constant

$$C = \left\{\frac{(n-1)s^2}{2}\right\}^{\frac{1}{2}(n-1)} \frac{1}{\Gamma\left(\frac{n-1}{2}\right)}. \tag{4.3.4}$$

Equation (4.3.3) is an inverted gamma density of the form

$$IG(x \mid \alpha, \beta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{-\alpha-1} \exp(-\beta/x)$$

with

$$\alpha = \frac{1}{2}(n-1) \text{ and } \beta = \frac{(n-1)s^2}{2}.$$

Theorem 4.3.1

$$C_{pl} \mid \sigma^2, \underline{Y} \sim N\left\{\frac{\overline{Y} - \ell_o}{3\sigma}, \frac{1}{9n}\right\}.$$

Proof

The proof is given in Appendix A4.

Before stating the moments of C_{pl}, we will remind ourselves of the relationship between moments for a general random variable X. Denote the first four posterior moments about the origin for X by μ'_1, μ'_2, μ'_3 and μ'_4 and the central moments by μ_2, μ_3 and μ_4, then:

Theorem 4.3.2

$$E(X) = \mu'_1$$

$$\mu'_2 = \mu_2 + (\mu'_1)^2$$

$$\mu'_3 = \mu_3 + 3\mu_2\mu'_1 + (\mu'_1)^3$$

$$\mu'_4 = \mu_4 + 4\mu'_1\mu_3 + 6(\mu'_1)^2\mu_2 + (\mu'_1)^4 .$$

Proof

The proof is given in Appendix A4.

The following two theorems can now be stated.

Theorem 4.3.2

Denote the first four posterior moments about the origin for C_{pl} (conditional on σ^2) by $\mu'_1, , \mu'_2, \mu'_3$ and μ'_4 therefore

$$\mu'_1 = \frac{\overline{Y} - \ell_0}{3\sigma}$$

$$\mu'_2 = \frac{1}{9}\left\{\frac{1}{n} + \frac{(\overline{Y} - \ell_0)^2}{\sigma^2}\right\}$$

$$\mu'_3 = \frac{\left(\overline{Y} - \ell_0\right)}{27}\left\{\frac{3}{n\sigma} + \frac{(\overline{Y} - \ell_0)^2}{\sigma^3}\right\}$$

$$\mu'_4 = \frac{1}{27}\left[\frac{1}{n^2} + (\overline{Y} - \ell_0)^2\left\{\frac{2}{n\sigma^2} + \frac{(\overline{Y} - \ell_0)^2}{3\sigma^4}\right\}\right].$$

Proof

The proof is given in Appendix A4.

Theorem 4.3.3

The r^th posterior moment of σ^{-2} about the origin is given by

$$E\left((\sigma^{-2})^r \mid \underline{Y}\right) = \left(\frac{2}{(n-1)s^2}\right)^r \frac{\Gamma(\frac{1}{2}(n+2r-1))}{\Gamma\left(\frac{n-1}{2}\right)}.$$

Proof

The proof is given in Appendix A4.

Theorem 4.3.4

Denote the first four posterior moments about the origin for C_{pl} (unconditional) by $m_1^{'}, m_2^{'}, m_3^{'}$ and m_4^{1} and also denote the (unconditional) variance, third and fourth central moments of the process capability index C_{pl} by m_2, m_3 *and* m_4, therefore

$$m'_1 = \left(\frac{\bar{Y}-\ell_0}{3}\right)\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{1}{2}} \tag{4.3.5}$$

$$m_2 = \frac{1}{9n} + \frac{2}{9}\frac{(\bar{Y}-\ell_0)^2}{(n-1)s^2}\left\{\frac{(n-1)}{2} - \frac{\Gamma^2\left(\frac{n}{2}\right)}{\Gamma^2\left(\frac{n-1}{2}\right)}\right\} \tag{4.3.6}$$

$$m_3 = \frac{(\bar{Y}-\ell_0)^3}{27}\left\{\frac{2}{(n-1)s^2}\right\}^{3/2}\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\left\{\frac{2\Gamma^2\left(\frac{n}{2}\right)}{\Gamma^2\left(\frac{n-1}{2}\right)} - \frac{(2n-3)}{2}\right\} \tag{4.3.7}$$

and

$$m_4 = \frac{1}{27n^2} + \frac{(\bar{Y}-\ell_0)^2}{27n}\frac{2}{(n-1)s^2}\left\{(n-1) - \frac{2\Gamma^2\left(\frac{n}{2}\right)}{\Gamma^2\left(\frac{n-1}{2}\right)}\right\}$$

$$+\frac{(\bar{Y}-\ell_0)^4}{27(n-1)^2(s^2)^2}\left[\frac{1}{3}(n-1)(n+1) + \frac{4\Gamma^2\left(\frac{n}{2}\right)}{\Gamma^2\left(\frac{n-1}{2}\right)}\left\{\frac{(n-3)}{3} - \frac{\Gamma^2\left(\frac{n}{2}\right)}{\Gamma^2\left(\frac{n-1}{2}\right)}\right\}\right]. \tag{4.3.8}$$

Proof

The proof is given in Appendix A4.

4.4 APPROXIMATE POSTERIOR DISTRIBUTIONS OF THE LOWER PROCESS CAPABILITY INDEX C_{pl}

In this section Pearson curve and Cornish–Fisher approximations of the posterior distribution of C_{pl} are derived. For details of how to determine the parameters of a Pearson curve, given the values of its moments, see for example Elderton (1953) or Elderton and Johnson (1969). As will be shown later, a type I Pearson curve can be used to approximate the posterior distribution of C_{pl} for our data set.

The density of a type I curve is given by

$$f(C_{pl}) = \tilde{K}\left(1+\frac{C_{pl}}{a_1}\right)^{M_1}\left(1-\frac{C_{pl}}{a_2}\right)^{M_2} \quad -a_1 < C_{pl} < a_2$$

where

$$\frac{M_1}{a_1} = \frac{M_2}{a_2}$$

$$M_1 = \tfrac{1}{2}\left\{r-2-r(r+2)\sqrt{\frac{\beta_1}{\beta_1(r+2)^2+16(r+1)}}\right\}$$

$$M_2 = \tfrac{1}{2}\left\{r-2+r(r+2)\sqrt{\frac{\beta_1}{\beta_1(r+2)^2+16(r+1)}}\right\}$$

$$r = \frac{6(\beta_2-\beta_1-1)}{(6+3\beta_1-2\beta_2)}$$

$$a_1+a_2 = \tfrac{1}{2}\sqrt{m_2}\sqrt{\left\{\beta_1(r+2)^2+16(r+1)\right\}}$$

$$\tilde{K}^{-1} = \int_{-\infty}^{\infty} f(C_{pk})\,dC_{pk}$$

$$\kappa = \frac{\beta_1(\beta_2+3)^2}{4(4\beta_2-3\beta_1)(2\beta_2-3\beta_1-6)};$$

with

$$\beta_1 = \frac{m_3^2}{m_2^3} \text{ and } \beta_2 = \frac{m_4}{m_2^2}$$

where $m'_1; m_2, m_3$ and m_4 are defined in (4.3.5) – (4.3.8).

The Cornish–Fisher percentage points can be calculated in the following way:

The standardised version of the capability index C_{pl} is $T = \dfrac{C_{pl} - m'_1}{\sqrt{m_2}}$ and the percentage point of level α of T is defined as t_α. With this definition it follows that a Cornish–Fisher expansion (see Cornish and Fisher (1937) and Fisher and Cornish (1960)) for the percentage point t_α of T is given by:

$$\tilde{t}_\alpha = z_\alpha + \frac{1}{6}\ell_3(z_\alpha^2 - 1) + \frac{1}{24}\ell_4(z_\alpha^3 - 3z_\alpha) - \frac{1}{36}\ell_3^2(2z_\alpha^3 - 5z_\alpha)$$

where z_α is the corresponding percentage point of the standard normal distribution, $\ell_r = \dfrac{\kappa_r}{(\kappa_2)^{r/2}} (r = 3, 4)$ and κ_r is the r–th cumulant of C_{pl}. Also

$$\kappa_2 = m_2, \quad \kappa_3 = m_3 \quad \text{and} \quad \kappa_4 = m_4 - 3m_2^2.$$

The percentage point of level α of C_{pl} is then given by $t_\alpha \sqrt{m_2} + m'_1$.

4.5 MONTE CARLO SIMULATION PROCEDURE FOR ESTIMATING THE POSTERIOR DISTRIBUTION OF C_{pl}

By making use of the fact that $\mu | \sigma^2, \underline{Y} \sim N\left(\bar{Y}, \dfrac{\sigma^2}{n}\right)$ (in (4.3.2)) and from the definition of the capability index, it follows that

$C_{pl} \left| \sigma^2, \underline{Y} \sim N\left(\dfrac{\bar{Y} - \ell_0}{3\sigma}, \dfrac{1}{9n}\right)\right.$ (theorem 4.3.1).

(4.5.1)

From (4.3.3) it follows that

$$p(\sigma^2 \mid \underline{Y}) = C(\sigma^2)^{-\frac{1}{2}(n-1)-1} \exp\left\{-\tfrac{1}{2}(n-1)s^2 / \sigma^2\right\}$$

which is an inverted gamma distribution and the normalising constant C is defined in (4.3.4).

Standard routines are again used in the simulation procedure:

1. By using the Matlab package, simulation of σ^2 can be obtained from (4.3.3) in the following way:

   (a) Simulate τ from a χ^2_{n-1} distribution, as a sum of $(n-1)$ squared independent standard normal random variates.

   (b) Calculate $\sigma^{2*} = \dfrac{(n-1)s^2}{\tau}$ where * indicates a simulated value

   (c) $\sigma^* = \sqrt{\sigma^{2*}}$.

2. Given σ^{2*}, calculate the conditional posterior density function $p(C_{pl} \mid \sigma^{2*}, \underline{Y})$ which is defined in equation 4.5.1.

3. Repeat steps (i) $\tilde{\ell}$ times. For our example $\tilde{\ell}$ was taken as 10 000. Using a Rao–Blackwell argument (see Gelfand and Smith, 1991) a density estimate of the unconditional posterior distribution of $C_{pl} = \dfrac{\mu - \ell_0}{3\sigma}$ can be obtained by averaging $p(C_{pl} \mid \sigma^2, \underline{Y})$ over the $\tilde{\ell}$ repetitions.

4.6 DIFFERENCES BETWEEN TWO LOWER PROCESS CAPABILITY INDICES

Consider two processes which are suspected to differ only in the mean setting and variance. For each process the posterior mean and precision have the following distributions:

$$\mu_i \mid \underline{Y_i}, \sigma_i^{\,2} \sim N(\overline{Y}_i, \frac{\sigma_i^{\,2}}{n_i}) \;\; \textbf{and} \;\; \frac{\nu_i s_i^{\,2}}{\sigma_i^{\,2}} \sim \chi^2_{\nu_i} \quad \textbf{for} \quad i = 1,2$$

where $\nu_i = (n_i - 1)$ and $\nu_i s_i^{\,2} = \sum_{j=1}^{n_i} (Y_{ij} - \overline{Y}_i)^2$

The standard routines are used to simulate each of the lower capability index as:

1. By using the Matlab package, simulation of $\sigma_i^{\,2}$ can be obtained in the following way:

   (a) Simulate a $\chi^2_{\nu_i}$ variate, as a sum of ν_i squared independent standard normal random variates.

   (b) Calculate $\sigma_i^{\,2*} = \frac{\nu_i s_i^{\,2}}{\chi^2_{\nu_i}}$, where (*) indicates a simulated value.

   (c) $\sigma_i^{\,*} = \sqrt{\frac{\nu_i s_i^{\,2}}{\chi^2_{\nu_i}}}$.

2. By making use of the fact that $\mu_i \mid \underline{Y_i}, \sigma_i^{\,2} \sim N(\overline{Y}_i, \frac{\sigma_i^{\,2}}{n_i})$, where $\underline{Y_i}$ is data drawn from process/supplier *i*, simulate μ_i^* and from the definition of the index, it follows that C_{pl_i} can be simulated as

$$C_{pl_i}^{\;*} = \frac{1}{3}\left(\frac{\mu_i^* - \ell_0}{\sigma_i^{\,*}}\right) = \frac{1}{3}\left(\frac{\overline{Y}_i + Z_i\sqrt{\frac{\sigma_i^{\,*}}{n_i}} - \ell_0}{\sigma_i^{\,*}}\right) = \frac{1}{3}\left(\frac{(\overline{Y}_i - \ell_0) + Z_i\sqrt{\frac{\sigma_i^{\,*}}{n_i}}}{\sigma_i^{\,*}}\right)$$

$$= \frac{1}{3}\left((\overline{Y}_i - \ell_0)\frac{1}{\sigma_i^{\,*}} + Z_i\sqrt{\frac{1}{n_i}}\right) = \frac{1}{3}\left((\overline{Y}_i - \ell_0)\sqrt{\frac{\chi^2_{\nu_i}}{\nu_i s_i^{\,2}}} + Z_i\sqrt{\frac{1}{n_i}}\right)$$

where $Z_i \sim N(0,1)$.

Let $d_{12} = C_{pl_1} - C_{pl_2}$ be the difference between the two lower process capability indices then:

$d_{12} = C_{pl_1} - C_{pl_2} = \left(\frac{\mu_1 - \ell_0}{3\sigma_1}\right) - \left(\frac{\mu_2 - \ell_0}{3\sigma_2}\right)$ and d_{12} can now be simulated as follows

$$d_{12}^{\,*} = \frac{1}{3}\left((\overline{Y}_1 - \ell_0)\sqrt{\frac{\chi^2_{\nu_1}}{\nu_1 s_1^{\,2}}} - (\overline{Y}_2 - \ell_0)\sqrt{\frac{\chi^2_{\nu_2}}{\nu_2 s_2^{\,2}}} + Z_1\sqrt{\frac{1}{n_1}} - Z_2\sqrt{\frac{1}{n_2}} \right).$$

3. Repeat steps (1-2) $\tilde{\ell}$ times for i=1,2

The distribution of

$$d_{12} \mid \chi^2_{\nu_1}, \chi^2_{\nu_2}, \underline{Y_i} \sim N\left((\overline{Y}_1 - \ell_0)\frac{1}{3}\sqrt{\frac{\chi^2_{\nu_1}}{\nu_1 s_1^{\,2}}} - (\overline{Y}_2 - \ell_0)\frac{1}{3}\sqrt{\frac{\chi^2_{\nu_2}}{\nu_2 s_2^{\,2}}},\quad \frac{1}{9}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right).$$

The following theorem can now be stated.

Theorem 4.6.1

Let $d_{12} \mid \chi^2_{\nu_1}, \chi^2_{\nu_2}, \underline{Y_i} \sim N\left((\overline{Y}_1 - \ell_0)\frac{1}{3}\sqrt{\frac{\chi^2_{\nu_1}}{\nu_1 s_1^{\,2}}} - (\overline{Y}_2 - \ell_0)\frac{1}{3}\sqrt{\frac{\chi^2_{\nu_2}}{\nu_2 s_2^{\,2}}},\quad \frac{1}{9}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right)$ for

i=1,2

then

$$E(d_{12} \mid \underline{Y_i}) = \frac{\sqrt{2}}{3}\left(\frac{(\overline{Y}_1 - \ell_0)}{\sqrt{\nu_1 s_1^{\,2}}}\frac{\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)} - \frac{(\overline{Y}_2 - \ell_0)}{\sqrt{\nu_2 s_2^{\,2}}}\frac{\Gamma(\left(\frac{\nu_2+1}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)} \right)$$

and

$$Var(d_{12} \mid \underline{Y_i}) = \frac{1}{9}\left(\frac{1}{n_1} + \frac{1}{n_2}\right) + \frac{1}{9}\left(\frac{(\overline{Y}_1 - \ell_0)^2}{\nu_1 s_1^{\,2}}\left\{\nu_1 - \frac{2\Gamma^2\left(\frac{\nu_1+1}{2}\right)}{\Gamma^2\left(\frac{\nu_1}{2}\right)}\right\} + \frac{(\overline{Y}_2 - \ell_0)^2}{\nu_2 s_2^{\,2}}\left\{\nu_2 - \frac{2\Gamma^2\left(\frac{\nu_2+1}{2}\right)}{\Gamma^2\left(\frac{\nu_2}{2}\right)}\right\} \right)$$

<u>Proof</u>

The proof is given in Appendix A4.

If $n_1 = n_2 = n$ and $\nu = (n-1)$, then

$$E(d_{12} \mid \underline{Y_i}) = \frac{\sqrt{2}}{3}\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\left(\frac{(\overline{Y}_1 - \ell_0)}{\sqrt{\nu s_1^{\,2}}} - \frac{(\overline{Y}_2 - \ell_0)}{\sqrt{\nu s_2^{\,2}}} \right)$$

and

$$Var(d_{12} \mid \underline{Y_i}) = \frac{2}{9n} + \frac{1}{9}\left\{\nu - \frac{2\Gamma^2\left(\frac{\nu+1}{2}\right)}{\Gamma^2\left(\frac{\nu}{2}\right)}\right\}\left(\frac{(\overline{Y}_1-\ell_0)^2}{\nu s_1^{\,2}} + \frac{(\overline{Y}_2-\ell_0)^2}{\nu s_2^{\,2}}\right).$$

Theorem 4.6.2

For given $\chi^2_{\nu_1}$ and $\chi^2_{\nu_2}$, denote the first four posterior moments about the origin for d_{12} by μ_1, μ'_2, μ'_3 and μ'_4 and the central moments by μ_2, μ_3 and μ_4, then:

$$\mu'_1 = (\overline{Y}_1-\ell_0)\frac{1}{3}\sqrt{\frac{\chi^2_{\nu_1}}{\nu_1 s_1^{\,2}}} - (\overline{Y}_2-\ell_0)\frac{1}{3}\sqrt{\frac{\chi^2_{\nu_2}}{\nu_2 s_2^{\,2}}}$$

$$\mu'_2 = \frac{1}{9}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)$$

$$\mu'_3 = \frac{3}{9}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\left(\frac{(\overline{Y}_1-\ell_0)}{3}\sqrt{\frac{\chi^2_{\nu_1}}{\nu_1 s_1^{\,2}}} - \frac{(\overline{Y}_2-\ell_0)}{3}\sqrt{\frac{\chi^2_{\nu_2}}{\nu_2 s_2^{\,2}}}\right) + \frac{(\overline{Y}_1-\ell_0)^3(\chi^2_{\nu_1})^{\frac{3}{2}}}{27(\nu_1 s_1^{\,2})^{\frac{3}{2}}}$$

$$-\frac{3(\overline{Y}_1-\ell_0)^2(\chi^2_{\nu_1})}{9(\nu_1 s_1^{\,2})}\frac{(\overline{Y}_2-\ell_0)(\chi^2_{\nu_2})^{\frac{1}{2}}}{3(\nu_2 s_2^{\,2})^{\frac{1}{2}}} + \frac{3(\overline{Y}_1-\ell_0)(\chi^2_{\nu_1})^{\frac{1}{2}}}{3(\nu_1 s_1^{\,2})^{\frac{1}{2}}}\frac{(\overline{Y}_2-\ell_0)^2(\chi^2_{\nu_2})}{9(\nu_2 s_2^{\,2})} - \frac{(\overline{Y}_2-\ell_0)^3(\chi^2_{\nu_2})^{\frac{3}{2}}}{27(\nu_2 s_2^{\,2})^{\frac{3}{2}}}$$

$$\mu'_4 = \frac{1}{27}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)^2 + \frac{2}{27}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\left(\frac{(\overline{Y}_1-\ell_0)^2(\chi^2_{\nu_1})}{(\nu_1 s_1^{\,2})} - 2(\overline{Y}_1-\ell_0)(\overline{Y}_2-\ell_0)\sqrt{\frac{\chi^2_{\nu_1}\chi^2_{\nu_2}}{\nu_1 s_1^{\,2}\nu_2 s_2^{\,2}}} + \frac{(\overline{Y}_2-\ell_0)^2(\chi^2_{\nu_2})}{(\nu_2 s_2^{\,2})}\right)$$

$$+\frac{(\overline{Y}_1-\ell_0)^4(\chi^2_{\nu_1})^2}{81(\nu_1 s_1^{\,2})^2} + 4\frac{(\overline{Y}_1-\ell_0)^3(\chi^2_{\nu_1})^{\frac{3}{2}}}{81(\nu_1 s_1^{\,2})^{\frac{3}{2}}}(\overline{Y}_2-\ell_0)\sqrt{\frac{\chi^2_{\nu_2}}{\nu_2 s_2^{\,2}}} + 6\frac{(\overline{Y}_1-\ell_0)^2(\overline{Y}_2-\ell_0)^2(\chi^2_{\nu_1})(\chi^2_{\nu_2})}{81(\nu_1 s_1^{\,2})(\nu_2 s_2^{\,2})}$$

$$+4(\overline{Y}_1-\ell_0)(\overline{Y}_2-\ell_0)^3\frac{(\chi^2_{\nu_1})^{\frac{1}{2}}(\chi^2_{\nu_2})^{\frac{3}{2}}}{81(\nu_1 s_1^{\,2})^{\frac{1}{2}}(\nu_2 s_2^{\,2})^{\frac{3}{2}}} + \frac{(\overline{Y}_2-\ell_0)^4(\chi^2_{\nu_2})^2}{81(\nu_2 s_2^{\,2})^2}.$$

Proof

The proof is given in Appendix A4.

Theorem 4.6.3

Denote the first four posterior moments about the origin for d_{12} (unconditional on the variance components or unconditional on the chi-square distribution) by $m_1^{'}, m_2^{'}, m_3^{'}$ *and* $m_4^{'}$ and also denote the (unconditional) variance, third and fourth central moments of the difference in process capability indices $d_{12} = C_{pl_1} - C_{pl_2}$ by m_2, m_3 *and* m_4, therefore

$$m'_1 = E(d_{12} \mid \underline{Y_i}) = \frac{\sqrt{2}}{3}\left(\frac{(\overline{Y}_1 - \ell_0)}{\sqrt{\nu_1 s_1^{\,2}}} \frac{\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)} - \frac{(\overline{Y}_2 - \ell_0)}{\sqrt{\nu_2 s_2^{\,2}}} \frac{\Gamma\left(\frac{\nu_2+1}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)} \right)$$

$$m_2 = \frac{1}{9}\left(\frac{1}{n_1} + \frac{1}{n_2}\right) + \frac{1}{9}\left(\frac{(\overline{Y}_1 - \ell_0)^2}{\nu_1 s_1^{\,2}} \left\{ \nu_1 - \frac{2\Gamma^2\left(\frac{\nu_1+1}{2}\right)}{\Gamma^2\left(\frac{\nu_1}{2}\right)} \right\} + \frac{(\overline{Y}_2 - \ell_0)^2}{\nu_2 s_2^{\,2}} \left\{ \nu_2 - \frac{2\Gamma^2\left(\frac{\nu_2+1}{2}\right)}{\Gamma^2\left(\frac{\nu_2}{2}\right)} \right\} \right)$$

$$m_3 = \frac{(\overline{Y}_1 - \ell_0)^3}{27}\left(\frac{2}{\nu_1 s_1^{\,2}}\right)^{\frac{3}{2}} \frac{\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)} \left\{ \frac{2\Gamma^2\left(\frac{\nu_1+1}{2}\right)}{\Gamma^2\left(\frac{\nu_1}{2}\right)} - \frac{(2\nu_1 - 1)}{2} \right\} - \frac{(\overline{Y}_2 - \ell_0)^3}{27}\left(\frac{2}{\nu_2 s_2^{\,2}}\right)^{\frac{3}{2}} \frac{\Gamma\left(\frac{\nu_2+1}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)} \left\{ \frac{2\Gamma^2\left(\frac{\nu_2+1}{2}\right)}{\Gamma^2\left(\frac{\nu_2}{2}\right)} - \frac{(2\nu_2 - 1)}{2} \right.$$

$$m_4 = \frac{1}{27}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)^2 + \frac{2}{27}(\overline{Y}_1 - \ell_0)^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right) \frac{1}{(\nu_1 s_1^{\,2})} \left\{ \nu_1 - \frac{2\Gamma^2\left(\frac{\nu_1+1}{2}\right)}{\Gamma^2\left(\frac{\nu_1}{2}\right)} \right\}$$

$$+ \frac{2}{27}(\overline{Y}_2 - \ell_0)^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right) \frac{1}{(\nu_2 s_2^{\,2})} \left\{ \nu_2 - \frac{2\Gamma^2\left(\frac{\nu_2+1}{2}\right)}{\Gamma^2\left(\frac{\nu_2}{2}\right)} \right\}$$

$$+ \frac{(\overline{Y}_1 - \ell_0)^4}{27(\nu_1 s_1^{\,2})^2} \left\{ \frac{\nu_1(\nu_1 + 2)}{3} + 4\frac{\Gamma^2\left(\frac{\nu_1+1}{2}\right)}{\Gamma^2\left(\frac{\nu_1}{2}\right)} \left[\frac{1}{3}(\nu_1 - 2) - \frac{\Gamma^2\left(\frac{\nu_1+1}{2}\right)}{\Gamma^2\left(\frac{\nu_1}{2}\right)} \right] \right\}$$

$$+ \frac{(\overline{Y}_2 - \ell_0)^4}{27(\nu_2 s_2^{\,2})^2} \left\{ \frac{\nu_2(\nu_2 + 2)}{3} + 4\frac{\Gamma^2\left(\frac{\nu_2+1}{2}\right)}{\Gamma^2\left(\frac{\nu_2}{2}\right)} \left[\frac{1}{3}(\nu_2 - 2) - \frac{\Gamma^2\left(\frac{\nu_2+1}{2}\right)}{\Gamma^2\left(\frac{\nu_2}{2}\right)} \right] \right\}$$

$$+ \frac{2(\overline{Y}_1 - \ell_0)^2(\overline{Y}_2 - \ell_0)^2}{27(\nu_1 s_1^{\,2})(\nu_2 s_2^{\,2})} \left\{ \nu_1\nu_2 + 4\frac{\Gamma^2\left(\frac{\nu_1+1}{2}\right)}{\Gamma^2\left(\frac{\nu_1}{2}\right)} \frac{\Gamma^2\left(\frac{\nu_2+1}{2}\right)}{\Gamma^2\left(\frac{\nu_2}{2}\right)} - \frac{2\Gamma^2\left(\frac{\nu_1+1}{2}\right)\nu_2}{\Gamma^2\left(\frac{\nu_1}{2}\right)} - \frac{2\Gamma^2\left(\frac{\nu_2+1}{2}\right)\nu_1}{\Gamma^2\left(\frac{\nu_2}{2}\right)} \right\}$$

where

$\nu_1 = (n_1 - 1)$ and $\nu_2 = (n_2 - 1)$.

<u>Proof</u>

The proof is given in Appendix A4.

If $n_1 = n_2 = n$ and $\nu_1 = \nu_2 = \nu = (n-1)$, then

$$m_3=\frac{2^{\frac{3}{2}}}{27}\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\left\{\frac{2\Gamma^2\left(\frac{\nu+1}{2}\right)}{\Gamma^2\left(\frac{\nu}{2}\right)}-\frac{(2\nu-1)}{2}\right\}\times\left\{\frac{(\overline{Y}_1-\ell_0)^3}{\left(\nu s_1{}^2\right)^{\frac{3}{2}}}-\frac{(\overline{Y}_2-\ell_0)^3}{\left(\nu s_2{}^2\right)^{\frac{3}{2}}}\right\}$$

and

$$m_4=\frac{1}{27}\left(\frac{2}{n}\right)^2+\frac{2}{27}(\overline{Y}_1-\ell_0)^2\left(\frac{2}{n}\right)\frac{1}{(\nu s_1{}^2)}\left\{\nu-\frac{2\Gamma^2\left(\frac{\nu+1}{2}\right)}{\Gamma^2\left(\frac{\nu}{2}\right)}\right\}$$

$$+\frac{2}{27}(\overline{Y}_2-\ell_0)^2\left(\frac{2}{n}\right)\frac{1}{(\nu s_2{}^2)}\left\{\nu-\frac{2\Gamma^2\left(\frac{\nu+1}{2}\right)}{\Gamma^2\left(\frac{\nu}{2}\right)}\right\}$$

$$+\frac{(\overline{Y}_1-\ell_0)^4}{27(\nu s_1{}^2)^2}\left\{\frac{\nu(\nu+2)}{3}+4\frac{\Gamma^2\left(\frac{\nu+1}{2}\right)}{\Gamma^2\left(\frac{\nu}{2}\right)}\left[\frac{1}{3}(\nu-2)-\frac{\Gamma^2\left(\frac{\nu+1}{2}\right)}{\Gamma^2\left(\frac{\nu}{2}\right)}\right]\right\}$$

$$+\frac{(\overline{Y}_2-\ell_0)^4}{27(\nu s_2{}^2)^2}\left\{\frac{\nu(\nu+2)}{3}+4\frac{\Gamma^2\left(\frac{\nu+1}{2}\right)}{\Gamma^2\left(\frac{\nu}{2}\right)}\left[\frac{1}{3}(\nu-2)-\frac{\Gamma^2\left(\frac{\nu+1}{2}\right)}{\Gamma^2\left(\frac{\nu}{2}\right)}\right]\right\}$$

$$+\frac{2(\overline{Y}_1-\ell_0)^2(\overline{Y}_2-\ell_0)^2}{27(\nu s_1{}^2)(\nu s_2{}^2)}\left\{\nu^2+4\frac{\Gamma^4\left(\frac{\nu+1}{2}\right)}{\Gamma^4\left(\frac{\nu}{2}\right)}-\frac{4\Gamma^2\left(\frac{\nu+1}{2}\right)\nu}{\Gamma^2\left(\frac{\nu}{2}\right)}\right\}.$$

4.7 PROBABILITY MATCHING AND REFERENCE PRIORS FOR C_{pl}

The Bayesian paradigm becomes attractive in many types of statistical problems – especially in capability index problems but the choice of an appropriate non–informative prior distribution has been controversial. Common non–informative priors in multi-parameter problems, such as Jeffreys' prior, can have features that have an unexpectedly dramatic effect on the posterior. Recently Datta and Ghosh (1995) derived the differential equation that a prior must satisfy if the posterior probability of a one-sided credibility interval for a parametric function and its frequentist probability agree up to $0(n^{-1})$ where n is the sample size. They prove that the agreement between the posterior probability and the frequentist probability holds if and only if

$$\sum_{\alpha=1}^{m}\frac{\partial}{\partial\theta_\alpha}\{\eta_\alpha(\underline{\theta})\pi(\underline{\theta})\}=0 \qquad (4.7.1)$$

where $\pi(\underline{\theta})$ is the probability–matching prior for $\underline{\theta}$ the vector of unknown parameters.

Also

$$\nabla_t = \left[\frac{\partial}{\partial \theta_1} t(\underline{\theta}), \ldots, \frac{\partial}{\partial \theta_m} t(\underline{\theta}) \right]' \tag{4.7.2}$$

and

$$\eta(\underline{\theta}) = \frac{F^{-1}(\underline{\theta}) \nabla_t(\underline{\theta})}{\sqrt{\nabla'_t(\underline{\theta}) F^{-1}(\underline{\theta}) \nabla_t(\underline{\theta})}} = [\eta_1(\underline{\theta}), \ldots, \eta_m(\underline{\theta})]'. \tag{4.7.3}$$

It is clear that $\eta'(\underline{\theta}) F(\underline{\theta}) \eta(\underline{\theta}) = 1$ for all $\underline{\theta}$ where $F^{-1}(\underline{\theta})$ is the inverse of $F(\underline{\theta})$, the Fisher information matrix of $\underline{\theta}$ per unit observation and $t(\underline{\theta})$ the parameter of interest. The following theorem can now be stated.

Theorem 4.7.1

Let $\{Y_i, i = 1, \ldots, n\}$ be independently normally distributed with parameter vector $\underline{\theta} = [\mu, \sigma]'$. Suppose the parameter of interest is the process capability index $C_{pl} = \dfrac{\mu - \ell_0}{3\sigma} = t(\underline{\theta})$, then the probability matching prior for C_{pl} is given by

$$\pi^M(\mu, \sigma) = \left\{ 1 + \frac{(\ell_0 - \mu)^2}{2\sigma^2} \right\}^{-\frac{1}{2}} \sigma^{-2}. \tag{4.7.4}$$

where $\sigma^2 > 0$ and $-\infty \leq \mu \leq \infty$

Proof

The proof is given in Appendix A4.

The following corollary involves a transformation from σ to σ^2.

Corollary 4.7.1.1

Let $\{Y_i, i = 1, \ldots, n\}$ be independently normally distributed with parameter vector $\underline{\theta} = [\mu, \sigma^2]'$. Suppose the parameter of interest is the process capability index $C_{pl} = \dfrac{\mu - \ell_0}{3\sigma} = t(\underline{\theta})$, then the probability matching prior for C_{pl} is given by

$$\pi^M(\mu,\sigma^2)=\left\{1+\frac{(\mu-\ell_0)^2}{2\sigma^2}\right\}^{-\frac{1}{2}}\sigma^{-3}. \tag{4.7.5}$$

Proof

The proof is given in Appendix A4.

Corollary 4.7.1.2

Let $\{Y_i, i=1,\ldots,n\}$ be independently normally distributed with parameter vector $\underline{\theta}=[\mu,\sigma]'$. Suppose the parameter of interest is the process capability index $C_{pl}=\dfrac{\mu-\ell_0}{3\sigma}=t(\underline{\theta})$, then the probability matching prior for C_{pl} is also given by

$$\pi^M(\mu,\sigma)=\left\{1+\frac{(\mu-\ell_0)^2}{2\sigma^2}\right\}^{\frac{1}{2}}. \tag{4.7.6}$$

Proof

The proof is given in Appendix A4.

As mentioned earlier, the Jeffreys' prior is not always suitable for multi-parameter problems. In recognition of this problem Berger and Bernardo (1992), propose the reference prior approach to the development of non–informative priors, the key feature of which was a possible dependence of the reference prior on specification of parameters of interest and nuisance parameters. In this section the reference prior of Berger and Bernardo (1992) is derived for the process capability index and the solution depends on the ordering of the parameters and how the parameter vector is divided into sub–vectors. In spite of these difficulties, there is growing evidence, mainly through examples that reference priors provide "sensible" answers from a Bayesian point of view and some more limited evidence that frequentist properties of inference from reference posteriors are asymptotically "reasonable". It will also be examined whether the reference priors satisfy the probability–matching criterion.

The following theorem can now be stated.

Theorem 4.7.2

For the process capability index $C_{pl} = \dfrac{\mu - \ell_0}{3\sigma}$, the reference prior relative to the ordered parameterisation (μ, σ^2) is given by

$$\pi^R(\mu, \sigma^2) = \sigma^{-3}\left\{1 + \frac{(\mu - \ell_0)^2}{2\sigma^2}\right\}^{-\frac{1}{2}}. \qquad (4.7.7)$$

Proof

The proof is given in Appendix A4.

Corollary 4.7.2

The reference prior is also a probability matching prior. This follows from (4.7.5) and (4.7.7). Mukerjee and Dey (1993), on the other hand, derived priors ensuring up to $0(n^{-1})$ frequentist validity of the posterior quantiles of the parameter of interest. If the reference prior and the probability matching prior produce the same results, they suggest higher order asymptotics to obtain the appropriate prior.

4.8 GIBBS SAMPLING

For the reference (probability matching) prior of the capability index, the joint posterior density of μ and σ^2 is given by

$$\begin{aligned} p(\mu, \sigma^2 \mid \underline{Y}) \;\propto\; & (\sigma^2)^{-\frac{1}{2}(n-1)} e^{-\frac{1}{2}(n-1)s^2/\sigma^2} \left(\frac{n}{\sigma^2}\right)^{\frac{1}{2}} \frac{e^{-\frac{n}{2}(\bar{Y}-\mu)^2/\sigma^2}}{\sqrt{2\pi}} \times \\ & \sigma^{-3}\left\{1 + \frac{(\mu - \ell_0)^2}{2\sigma^2}\right\}^{-\frac{1}{2}}. \end{aligned} \qquad (4.8.1)$$

Technical difficulties arising in the calculation of the marginal posterior densities needed for Bayesian inference in the case of reference and probability matching priors have long served as an impediment to the wider application of the Bayesian framework to data analysis. The reason for this is that the integration operation plays a fundamental role in Bayesian statistics.

Recently, due to work of Gelfand and Smith (1991), Gelfand *et al.* (1990), Carlin *et al.* (1992), and Gelfand *et al.* (1992), the Gibbs sampler has been shown to be a useful tool for applied Bayesian inference in a broad variety of statistical problems. The typical objective of the sampler is to collect a sufficiently large number of parameter realisations from conditional posterior densities in order to obtain accurate estimates of the marginal posterior densities.

For the sake of convenience, define $C_{pl} = t(\underline{\theta})$.

In the $(t(\underline{\theta}), \sigma^2)$ parameterisation the conditional densities are given by

$$p(t(\underline{\theta}) \mid \sigma^2, \underline{Y}) \propto \left\{1 + \tfrac{9}{2} t^2(\underline{\theta})\right\}^{-\frac{1}{2}} e^{-\frac{n}{2}(\bar{Y} - 3\sigma t(\underline{\theta}) + \ell_o)^2/\sigma^2} \qquad -\infty < t(\underline{\theta}) < \infty \qquad (4.8.2)$$

and

$$p(\sigma^2 \mid t(\underline{\theta})) \propto (\sigma^2)^{-\frac{1}{2}(n+2)} e^{-\frac{n}{2}(\bar{Y} - 3\sigma t(\underline{\theta}) + \ell_0)^2/\sigma^2} \qquad \sigma^2 > 0. \qquad (4.8.3)$$

The Gibbs–sampler can now be implemented. The iterative process starts by using arbitrary starting values $\sigma^{2(0)}, t^0(\underline{\theta})$, to calculate the first iteration using (4.8.2) and (4.8.3). As burn in period we used 5000 iterations. After k iterations, in which the conditional densities are updated at each iteration, the Gibbs–sampler generates the values $\sigma^{2(k)}, t^{(k)}(\underline{\theta})$. Given σ^2, calculate the conditional posterior density function $p(t(\underline{\theta}) \mid \sigma^2, \underline{Y})$ as defined in (4.8.2). The process is repeated $\tilde{\ell}$ times. The convergence of the algorithm is discussed in Gelfand and Smith (1991). The problem in this chapter was done with $k = 50$ and $\tilde{\ell} = 10000$. Using a Rao–Blackwell argument as discussed in Van der Merwe and Chikobvu (2004) a density estimate of the unconditional posterior distribution of $t(\underline{\theta}) = C_{pl} = \dfrac{\mu - \ell_0}{3\sigma}$ can be obtained.

It is less expensive to use successive values in Gibbs sampling procedure. Gelman and Rubin (1992), however, warned “that particularly during the first tentative examination of a new problem, it can be argued that monitoring the evolutionary behaviour of several runs of the chain starting from a wide range of interval values is necessary”.

4.9 PROBABILITY MATCHING PRIOR FOR A SPECIAL FUNCTION

A parameter that is often of interest in capability index theory is $\delta = \Phi\left(\frac{\mu - \ell_0}{\sigma}\right)$.

Bernardo and Irony (1996) use the conventional non-informative prior $\pi(\mu,\sigma^2) \propto \sigma^{-2}$ but mentions on page 15 of their article that this prior is not necessarily the most appropriate "default" prior for δ. This section will therefore derive the probability matching prior for δ. The following theorem can now be stated.

Theorem 4.9.1

Let $\{Y_i, i = 1,\ldots,n\}$ be independently normally distributed with parameter vector $\theta = [\mu,\sigma]'$. Suppose the parameter of interest is $\delta = \Phi\left(\frac{\mu - \ell_0}{\sigma}\right)$ where Φ is the distribution function of the standard normal distribution, then the probability matching prior for δ is given by

$$\pi^M(\mu,\sigma^2) = \left\{1 + \frac{(\ell_0 - \mu)^2}{2\sigma^2}\right\}^{-\frac{1}{2}} \sigma^{-3} \qquad (4.9.1)$$

Here, (4.9.1) is exactly the same as the probability matching prior for C_{pl}, (4.7.5).

4.10 APPLICATION ONE

To illustrate how to calculate process capability index (C_{pl}) using the Bayesian techniques discussed, we refer to the Aircraft data example in chapter 2.

The definition of C_{pk} includes, as special cases, those processes where only one limit exists, in which case it reduces to the appropriate one sided standardized measure. Thus if there is no upper specification limit, we obtain $C_{pl} = \frac{(\mu - \ell_0)}{3\sigma}$ by simply letting $\ell_1 \to \infty$ in the original definition.

Results

The non-informative prior is used in the Pearson's curve (Monte Carlo simulation), and the Rao–Blackwell as well as the Gibbs–sampling methods are used in the case of the reference prior (the same as the probability matching prior in Corollary 4.7.1.1). Figures 4.1 and 4.2 illustrate the posterior distributions of C_{pl} for the data set given in Table 4.1. The graphs do not differ much because of the large data set used. The histogram results are also shown in Figures 4.1. It is further clear that the Pearson curve approximation is quite good. Under Jeffreys's prior the mean C_{pl} is 2.54. A 95% credibility interval for C_{pl} is (2.04, 3.06). Under the reference prior (which is the same as the probability matching prior), the mean C_{pl} is 2.63. A 95% credibility interval for C_{pl} is (2.11, 3.11).

Figures 4.3 and 4.4 illustrate the posterior distribution of C_{pl} for the first 20 observations given in Table 4.1. Because of the smaller number of observations used the posterior distributions in figures 4.3 and 4.4 have a larger variance when compared to the corresponding figures 4.1 and 4.2. The Jeffrey's prior gives rise to a mean C_{pl} of 2.94 with a 95% credibility interval for C_{pl} of (2.02, 3.94). The reference prior (which is the same as the probability matching prior) gives rise to a mean C_{pl} of 3.15 and a 95% credibility interval for C_{pl} of (2.21, 4.11).

From the results it is clear that the different priors give different estimates. The probability matching (reference) prior is recommended, because it is designed to produce posterior credible intervals which are asymptotically identical to their frequentist counterparts.

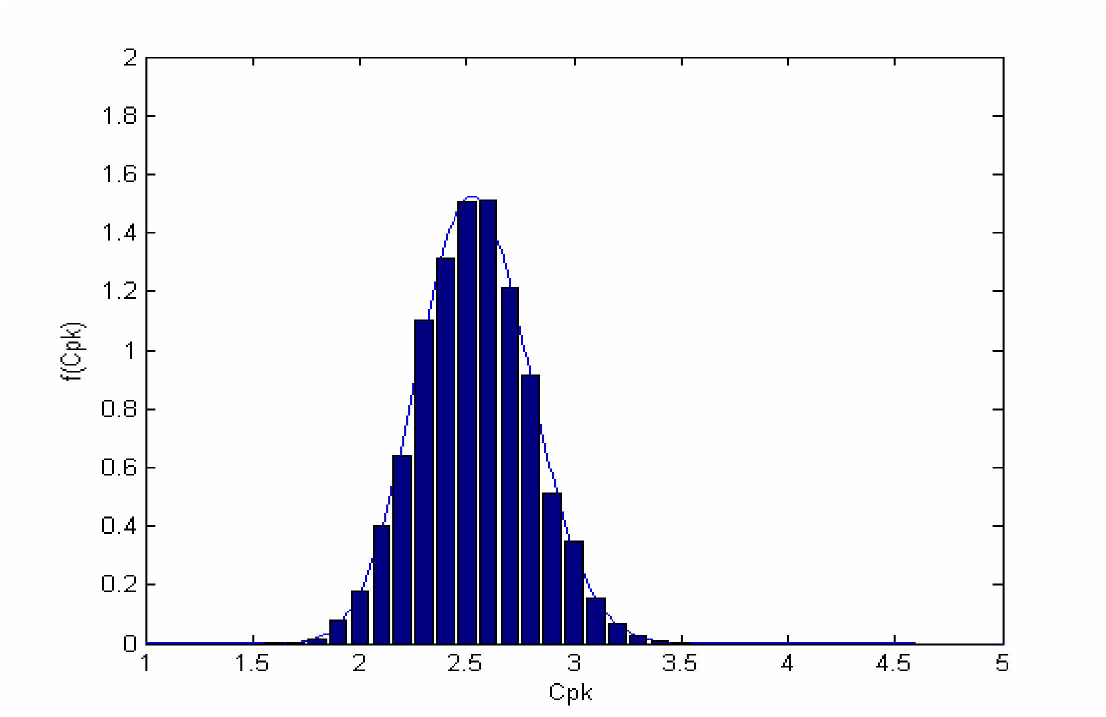


Figure 4.1: Histogram and Pearson's type 1 curve for the process capability index C_{pl}

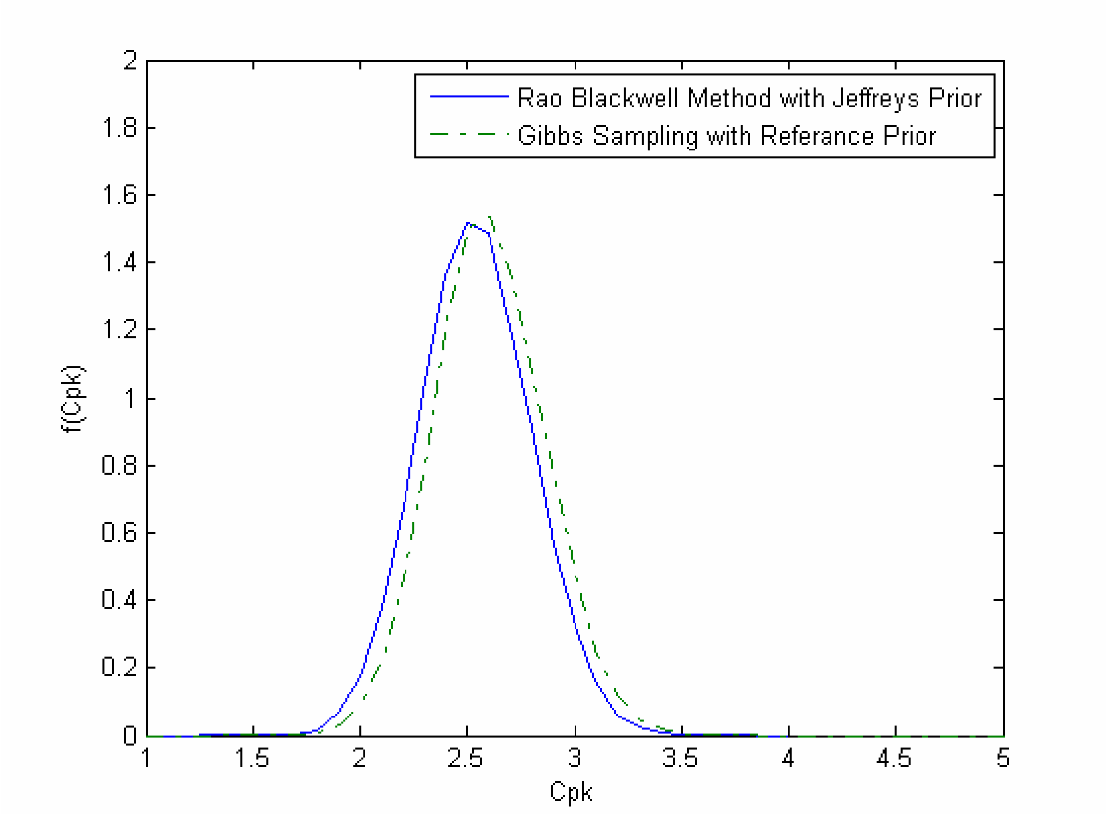


Figure 4.2: Distribution curves for the process capability index C_{pl}

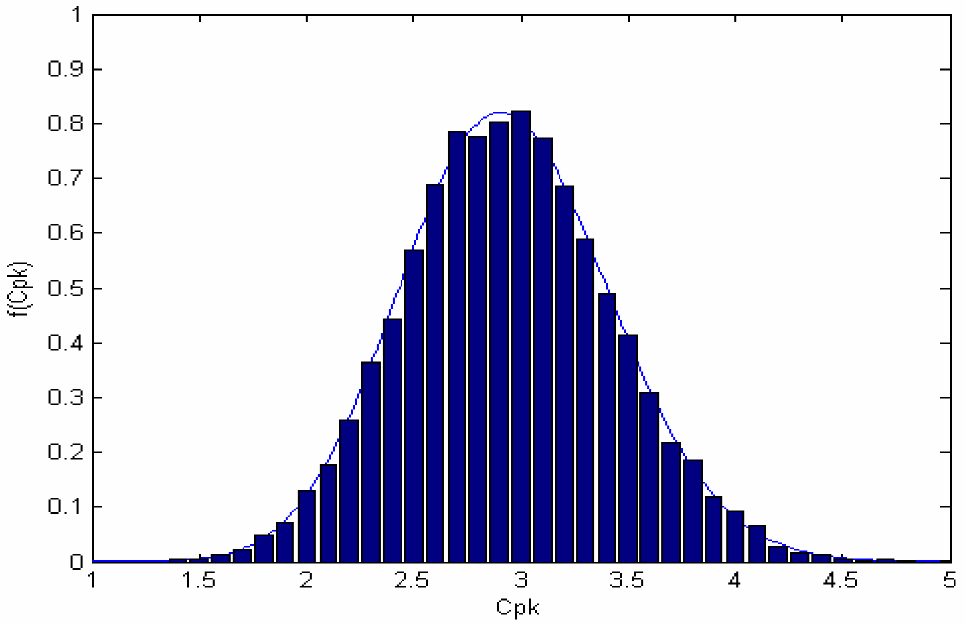


Figure 4.3: Histogram and Pearson's type 1 curve for the process capability index C_{pl} (Data from first 20 observations only)

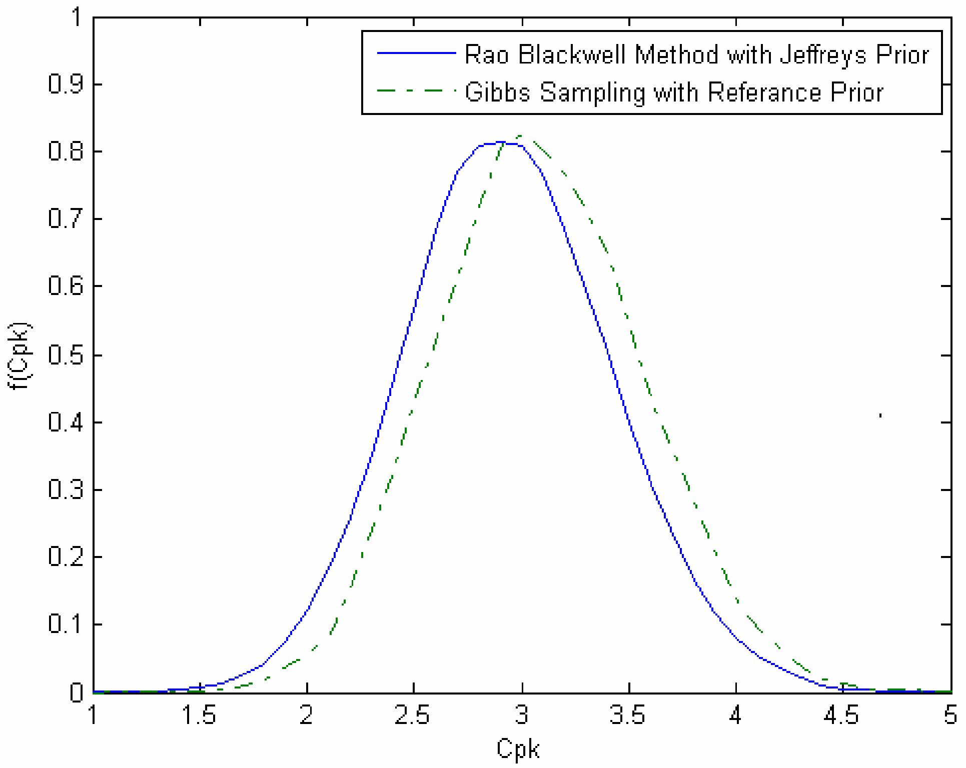


Figure 4.4: Distribution curves for the process capability index C_{pl} (Data from first 20 observations only)

4.11 APPLICATION TWO

In this chapter $C_{pl} = \dfrac{\mu - LSL}{3\sigma}$ is used in defense of the Bayesian approach. Some applications require an upper limit rather than a lower limit. In a similar way to C_{pl}, the moments for the difference between two upper capability indices of the form $C_{pu} = \dfrac{USL - \mu}{3\sigma}$ could be derived without much extra effort. The Bayesian results derived, will still hold in the case of C_{pu} since $C_{pu} = -\left(\dfrac{\mu - USL}{3\sigma}\right)$ which is now of the same form as C_{pl} except for the negative sign.

The following is taken from actual flatness measurement data obtained from industrial processes. The data sets passed a goodness-of-fit test for normality and are collected from stable processes. The summary statistics and specifications are given in table 4.2.

In this example, we consider three samples from an aluminum part. The null hypothesis for this case is

$$H_0 : \begin{pmatrix} d_{12} \\ d_{13} \\ d_{23} \end{pmatrix} = 0$$

here $d_{ij} = C_{pu_i} - C_{pu_j} \;\; i < j = 2,3$

Table 4.2: Summary data from flatness measurements

i	n_i	$\overline{Y}_i$	s_i	USL	$\widehat{C}_{pu_i}$
1	**20**	**0.00045**	**0.00012**	**0.001**	**1.619**
2	**20**	**0.00045**	**0.00009**	**0.001**	**2.061**
3	**20**	**0.00073**	**0.00010**	**0.001**	**0.946**

Source: Hubele, Berrado and Gel (2005)

Ten thousand simulations of each d_{ij} for $i < j = 2,3$ are performed. Pair-wise comparisons listed in table 4.3 below indicate that the capability index of sample 3 is significantly different from the capability indices of samples 1 and 2, whereas the indices from samples 1 and 2 are comparable. Zero is not included in the last two intervals.

Table 4.3: Summary of pair-wise tests

Statistic	**95% credibility interval**	**Test Conclusion**
d_{12}^{*}	**(-1.3422;0.3123)**	**fail to reject** H_0
d_{13}^{*}	**(0.0336;1.2200)**	**reject** H_0
d_{23}^{*}	**(0.4251;1.8754)**	**reject** H_0

Hubele, Berrado and Gel (2005) came to a similar conclusion using the same data set but using a Wald test for comparing multiple capability indices. Their results are summarised in table 4.4 below.

Table 4.4: Summary of pair-wise tests using Wald test

Samples Compared	**Statistic W**	$\chi_1^2(0.05)$	**Test Conclusion**
1 & 2	**1.07**	**3.84**	**fail to reject** H_0
1 & 3	**4.56**	**3.84**	**reject** H_0
2 &3	**8.89**	**3.84**	**reject** H_0

Plots of the histograms and Pearson's Type 1 distribution curves for d_{12}, d_{13} and d_{23} are done in figures 4.5 through to 4.7.

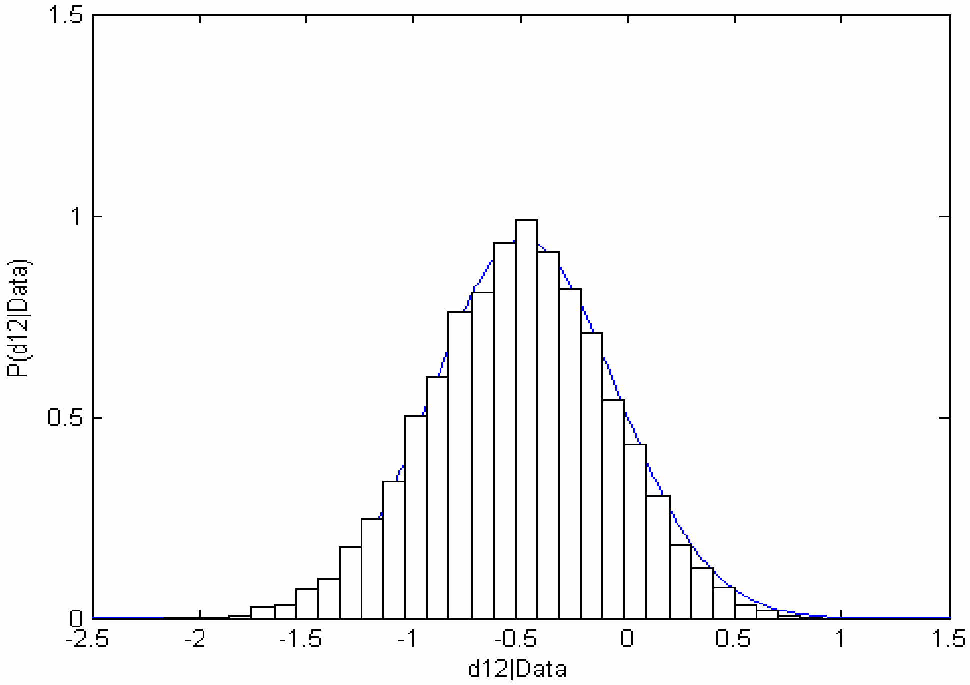


Figure 4.5: Histogram and Pearson's type 1 curve for the differences in process capability indices (C_{pu}'s)

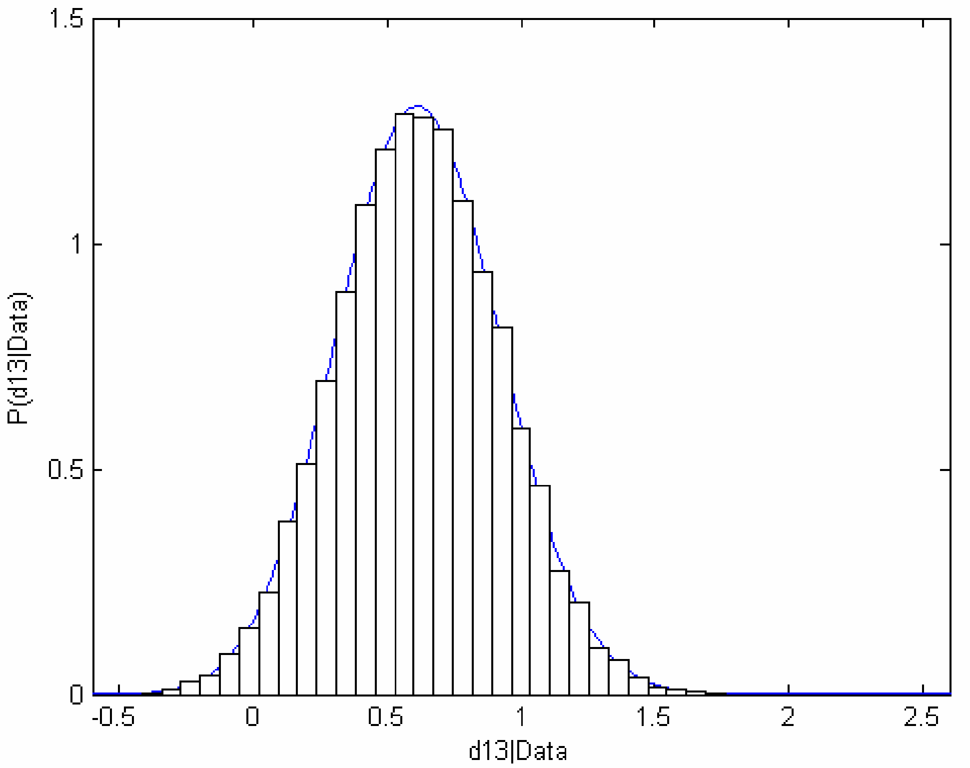


Figure 4.6: Histogram and Pearson's type 1 curve for the differences in process capability Indices (C_{pu}'s)

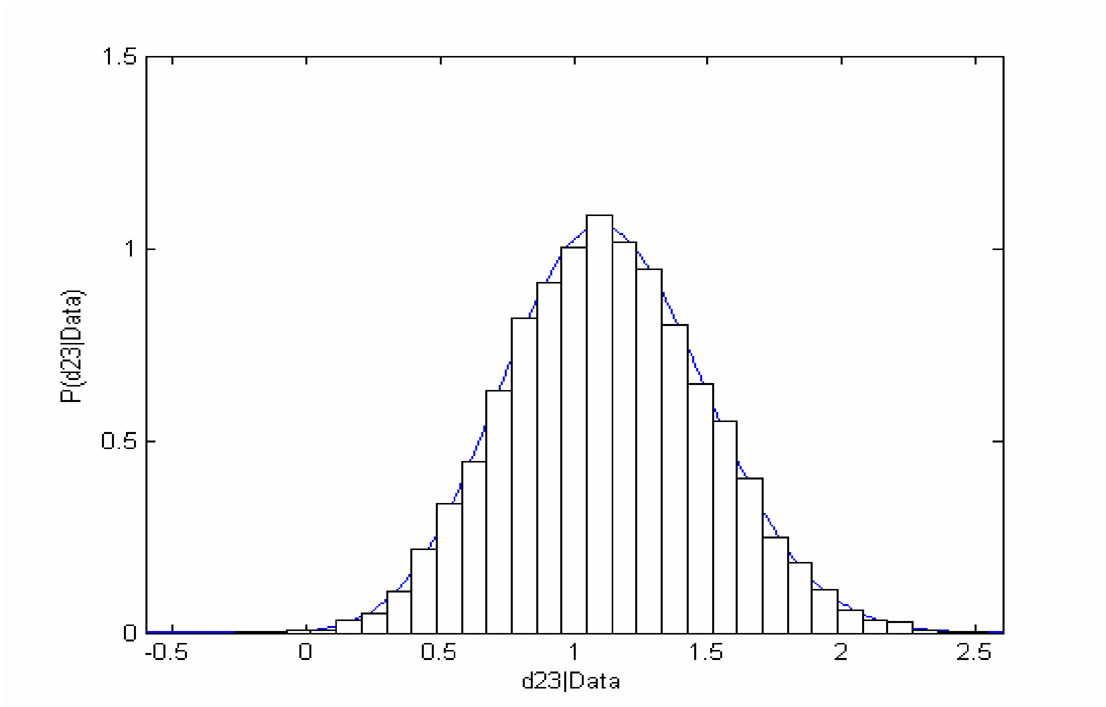


Figure 4.7: Histogram and Pearson's type 1 curve for the differences in process capability Indices (C_{pu} 's)

This index C_{pu} is recommended for quality features with a unilateral specification of USL. Linearity, flatness, circularity and cylindricity are examples of such quality features.

4.12 CONCLUSION

Bayesian inference has a number of advantages. A full Bayesian analysis provides a natural way of taking into account all sources of uncertainty in the estimation of the parameters. In this chapter uncertainty about the true value of the process capability index is incorporated into the analysis through the choice of a non informative prior distribution. The probability matching (reference) prior is recommended because it is designed to produce posterior credible intervals which are asymptotically identical to their frequentist counterparts.

Appendix A4

Proof of theorem 4.3.1

By making use of the fact that $\mu \mid \sigma^2, \underline{Y} \sim N\left(\overline{Y}, \frac{\sigma^2}{n}\right)$ (equation (4.3.2))

Proof:

$$\mu \mid \sigma^2, \underline{Y} \sim N\left(\overline{Y}, \frac{\sigma^2}{n}\right)$$

$$\mu - l_o \mid \sigma^2, \underline{Y} \sim N\left(\overline{Y} - l_o, \frac{\sigma^2}{n}\right)$$

$$\frac{\mu - l_o}{\sigma} \mid \sigma^2, \underline{Y} \sim N\left(\frac{\overline{Y} - l_o}{\sigma}, \frac{1}{n}\right)$$

$$\frac{\mu - l_o}{3\sigma} \mid \sigma^2, \underline{Y} \sim N\left(\frac{\overline{Y} - l_o}{3\sigma}, \frac{1}{9n}\right)$$

$$C_{pl} \mid \sigma^2, \underline{Y} \sim N\left\{\frac{\overline{Y} - \ell_o}{3\sigma}, \frac{1}{9n}\right\} \text{since } C_{pl} = \frac{\mu - l_o}{3\sigma} \ .$$

Proof of theorem 4.3.2

$E(X) = \mu'_1$ by definition.

$$E(X - \mu'_1)^2 = E(X^2 - 2X\mu'_1 + (\mu'_1)^2)$$

$$\mu_2 = E(X^2) - 2E(X)\mu'_1 + (\mu'_1)^2$$

$$= \mu'_2 - 2(\mu'_1)^2 + (\mu'_1)^2$$

$$= \mu'_2 - (\mu'_1)^2$$

$$\therefore \mu'_2 = \mu_2 + (\mu'_1)^2 .$$

$$E(X - \mu'_1)^3 = E(X^3 - 3X^2\mu'_1 + 3X(\mu'_1)^2 - (\mu'_1)^3)$$

$$\mu_3 = E(X^3) - 3E(X^2)\mu'_1 + 3(X)(\mu'_1)^2 - (\mu'_1)^3)$$

$$= \mu'_3 - 3\mu'_2\mu'_1 + 3\mu'_1(\mu'_1)^2 - (\mu'_1)^3)$$

$$= \mu'_3 - 3\mu'_2\mu'_1 + 3(\mu'_1)^3 - (\mu'_1)^3)$$

$$= \mu'_3 - 3\mu'_2\mu'_1 + 2(\mu'_1)^3$$

$$= \mu'_3 - 3(\mu_2 + (\mu'_1)^2)\mu'_1 + 2(\mu'_1)^3 \quad \text{since} \quad \mu'_2 = \mu_2 + (\mu'_1)^2$$

$$= \mu'_3 - 3\mu_2\mu'_1 - 3(\mu'_1)^3 + 2(\mu'_1)^3$$

$$= \mu'_3 - 3\mu_2\mu'_1 - (\mu'_1)^3$$

$$\therefore \mu'_3 = \mu_3 + 3\mu_2\mu'_1 + (\mu'_1)^3 .$$

$$E(X - \mu'_1)^4 = E(X^4 - 4X^3\mu'_1 + 6X^2(\mu'_1)^2 - 4X(\mu'_1)^3 + (\mu'_1)^4)$$

$$\mu_4 = E(X^4) - 4E(X^3)\mu'_1 + 6E(X^2)(\mu'_1)^2 - 4E(X)(\mu'_1)^3 + (\mu'_1)^4$$

$$= \mu'_4 - 4\mu'_3\mu'_1 + 6\mu'_2(\mu'_1)^2 - 4\mu'_1(\mu'_1)^3 + (\mu'_1)^4$$

$$= \mu'_4 - 4\mu'_3\mu'_1 + 6\mu'_2(\mu'_1)^2 - 4(\mu'_1)^4 + (\mu'_1)^4$$

$$= \mu'_4 - 4\mu'_3\mu'_1 + 6\mu'_2(\mu'_1)^2 - 3(\mu'_1)^4$$

$$but \quad \mu'_3 = \mu_3 + 3\mu_2\mu'_1 + (\mu'_1)^3 \quad and \quad \mu'_2 = \mu_2 + (\mu'_1)^2$$

$$\therefore \qquad \mu_4 = \mu'_4 - 4(\mu_3 + 3\mu_2\mu'_1 + (\mu'_1)^3)\mu'_1 + 6(\mu_2 + (\mu'_1)^2)(\mu'_1)^2 - 3(\mu'_1)^4$$

$$= \mu'_4 - 4\mu_3\mu'_1 - 12\mu_2(\mu'_1)^2 - 4(\mu'_1)^4 + 6\mu_2(\mu'_1)^2 + 6(\mu'_1)^4 - 3(\mu'_1)^4$$

$$= \mu'_4 - 4\mu_3\mu'_1 - 6\mu_2(\mu'_1)^2 - (\mu'_1)^4$$

$$\therefore \mu'_4 = \mu_4 + 4\mu'_1\mu_3 + 6(\mu'_1)^2\mu_2 + (\mu'_1)^4 .$$

Proof of Theorem 4.3.2

Since $\mu \mid \sigma^2, \underline{Y} \sim N\left(\overline{Y}, \frac{\sigma^2}{n}\right)$, the posterior density of $C_{pl} = \frac{\mu - \ell_0}{3\sigma}$ conditionally on σ^2 is normal with mean $\frac{\overline{Y} - \ell_0}{3\sigma}$ and variance $\frac{1}{9n}$. Denote the first four posterior moments about the origin for C_{pl} (conditionally on σ^2) by μ'_1, μ'_2, μ'_3 and μ'_4 and the central moments by μ_2, μ_3 and μ_4, then

$$\mu'_1 = \frac{\overline{Y} - \ell_0}{3\sigma} . \qquad \text{(A4.1)}$$

$$\mu'_2 = \mu_2 + (\mu'_1)^2$$

$$= \frac{1}{9n} + \frac{(\overline{Y} - \ell_0)^2}{9\sigma^2}$$

$$= \frac{1}{9}\left\{\frac{1}{n} + \frac{(\overline{Y} - \ell_0)^2}{\sigma^2}\right\} . \qquad \text{(A4.2)}$$

$$\mu'_3 = \mu_3 + 3\mu_2\mu'_1 + (\mu'_1)^3$$

$$= 0 + 3\frac{1}{9n}\frac{\overline{Y} - \ell_0}{3\sigma} + \left(\frac{\overline{Y} - \ell_0}{3\sigma}\right)^3$$

$$= \frac{1}{3n}\left(\frac{\overline{Y} - \ell_0}{3\sigma}\right) + \left(\frac{\overline{Y} - \ell_0}{3\sigma}\right)^3$$

$$= 3\left(\frac{\overline{Y} - \ell_0}{27n\sigma}\right) + \frac{\left(\overline{Y} - \ell_0\right)^3}{27\sigma^3}$$

$$= \frac{\left(\overline{Y} - \ell_0\right)}{27}\left\{\frac{3}{n\sigma} + \frac{(\overline{Y} - \ell_0)^2}{\sigma^3}\right\}. \qquad \text{(A4.3)}$$

and

$$\mu'_4 = \mu_4 + 4\mu'_1\mu_3 + 6(\mu'_1)^2\mu_2 + (\mu'_1)^4$$

$$= 3\left(\frac{1}{9n}\right)^2 + 0 + 6\left(\frac{\overline{Y} - \ell_0}{3\sigma}\right)^2\frac{1}{9n} + \left(\frac{\overline{Y} - \ell_0}{3\sigma}\right)^4$$

$$= \left(\frac{1}{27n^2}\right) + 6\frac{\left(\overline{Y} - \ell_0\right)^2}{81n\sigma^2} + \frac{\left(\overline{Y} - \ell_0\right)^4}{81\sigma^4}$$

$$= \left(\frac{1}{27n^2}\right) + \frac{2}{3n}\left(\frac{\overline{Y} - \ell_0}{3\sigma}\right)^2 + \left(\frac{\overline{Y} - \ell_0}{3\sigma}\right)^4$$

$$= \frac{1}{27}\left[\frac{1}{n^2} + (\overline{Y} - \ell_0)^2\left\{\frac{2}{n\sigma^2} + \frac{(\overline{Y} - \ell_0)^2}{3\sigma^4}\right\}\right]. \qquad \text{(A4.4)}$$

Proof of Theorem 4.3.3

Since the inverse gamma distribution is given by

$$IG(x\,|\,\alpha,\beta)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}x^{-\alpha-1}\exp(-\beta/x)$$

with

$$\int_0^{\infty}\frac{\beta^{\alpha}}{\Gamma(\alpha)}x^{-\alpha-1}\exp(-\beta/x)dx=1$$

and therefore

$$\int_0^{\infty}x^{-\alpha-1}\exp(-\beta/x)=\frac{\Gamma(\alpha)}{\beta^{\alpha}}$$

$$E\{(\sigma^{-2})^{r}\,|\,\underline{Y}\}=E_{\sigma^2}\{(\sigma^2)^{-r}\,|\,\underline{Y}\}=C\int_0^{\infty}(\sigma^2)^{-\frac{1}{2}(n-1)-1}(\sigma^2)^{-r}\exp\left\{-\tfrac{1}{2}(n-1)s^2/\sigma^2\right\}d\sigma^2$$

$$=C\int_0^{\infty}(\sigma^2)^{-\frac{1}{2}(n+2r-1)-1}\exp\left\{-\tfrac{1}{2}(n-1)s^2/\sigma^2\right\}d\sigma^2$$

$$=C\frac{\Gamma(\frac{1}{2}(n+2r-1))}{\left(\frac{1}{2}(n-1)s^2\right)^{\frac{1}{2}(n+2r-1)}}\ \text{ and with } C=\left\{\frac{(n-1)s^2}{2}\right\}^{\frac{1}{2}(n-1)}\frac{1}{\Gamma\left(\frac{n-1}{2}\right)}$$

$$=\frac{\left\{\frac{(n-1)s^2}{2}\right\}^{\frac{1}{2}(n-1)}}{\left(\frac{1}{2}(n-1)s^2\right)^{\frac{1}{2}(n+2r-1)}}\frac{\Gamma(\frac{1}{2}(n+2r-1))}{\Gamma\left(\frac{n-1}{2}\right)}$$

$$=\frac{1}{\left(\frac{1}{2}(n-1)s^2\right)^{\frac{1}{2}(2r)}}\frac{\Gamma(\frac{1}{2}(n+2r-1))}{\Gamma\left(\frac{n-1}{2}\right)}$$

$$=\left(\frac{2}{(n-1)s^2}\right)^{r}\frac{\Gamma(\frac{1}{2}(n+2r-1))}{\Gamma\left(\frac{n-1}{2}\right)}.$$

Proof of Theorem 4.3.4

By substituting for $r=\frac{1}{2},1,\frac{3}{2}$ and 2 in equations (A4.1) to (A4.4) after integrating out σ and using the relationships between moments about the origin and central moments, equations (4.3.5) to (4.3.8) will be obtained as follows:

Let $Z=\dfrac{\overline{Y}-\ell_0}{3\sigma}$

$$m'_1 = E_{\sigma^2}(\frac{\overline{Y}-\ell_0}{3\sigma})$$
$$= E_{\sigma^2}(Z)$$
$$= \frac{(\overline{Y}-\ell_0)}{3}E(\frac{1}{\sigma})$$

but

$$E\{(\sigma^2)^{-r} \mid \underline{Y}\} = \left(\frac{2}{(n-1)s^2}\right)^r \frac{\Gamma(\frac{1}{2}(n+2r-1))}{\Gamma\left(\frac{n-1}{2}\right)}$$

$$r = \frac{1}{2} \qquad E\{(\sigma^2)^{-\frac{1}{2}} \mid \underline{Y}\} = \left(\frac{2}{(n-1)s^2}\right)^{\frac{1}{2}} \frac{\Gamma(\frac{1}{2}(n+1-1))}{\Gamma\left(\frac{n-1}{2}\right)} = \left(\frac{2}{(n-1)s^2}\right)^{\frac{1}{2}} \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}.$$

Therefore

$$m'_1 = \frac{(\overline{Y}-\ell_0)}{3}\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\left(\frac{2}{(n-1)s^2}\right)^{\frac{1}{2}}$$

which proves the result in equation (4.3.5).

$$m'_2 = E_{\sigma^2}\left(\frac{1}{9n}\right) + E_{\sigma^2}\left(\frac{(\overline{Y}-\ell_0)^2}{9\sigma^2}\right)$$
$$= E_{\sigma^2}\left(\frac{1}{9n}\right) + E_{\sigma^2}(Z^2)$$

$$= \frac{1}{9n} + \frac{(\overline{Y}-\ell_0)^2}{3^2} C \frac{\Gamma(\frac{n+1}{2})}{\left\{\frac{(n-1)s^2}{2}\right\}^{\frac{1}{2}(n+1)}}$$

$$= \frac{1}{9n} + \frac{(\overline{Y}-\ell_0)^2}{9}\left\{\frac{(n-1)s^2}{2}\right\}^{\frac{1}{2}(n-1)} \left\{\frac{2}{(n-1)s^2}\right\}^{\frac{1}{2}(n+1)} \frac{\Gamma(\frac{n+1}{2})}{\Gamma\left(\frac{n-1}{2}\right)}$$
$$= \frac{1}{9n} + \frac{(\overline{Y}-\ell_0)^2}{9}\left\{\frac{2}{(n-1)s^2}\right\}\frac{\Gamma(\frac{n+1}{2})}{\Gamma\left(\frac{n-1}{2}\right)}$$
$$= \frac{1}{9n} + \frac{(\overline{Y}-\ell_0)^2}{9}\left\{\frac{2}{(n-1)s^2}\right\}\frac{(n-1)}{2}$$
$$= \frac{1}{9n} + \frac{(\overline{Y}-\ell_0)^2}{9}\left\{\frac{1}{s^2}\right\}$$

$$=\frac{1}{9n}+\frac{(\overline{Y}-\ell_0)^2}{9s^2}$$

$$=\frac{1}{9}\left(\frac{1}{n}+\frac{(\overline{Y}-\ell_0)^2}{s^2}\right).$$

$$m'_3=\frac{1}{3n}E_{\sigma^2}\left(\frac{\overline{Y}-\ell_0}{3\sigma}\right)+E_{\sigma^2}\left(\frac{\overline{Y}-\ell_0}{3\sigma}\right)^3$$

$$=\frac{1}{3n}E_{\sigma^2}\left(Z\right)+E_{\sigma^2}\left(Z^3\right)$$

$$=\frac{1}{3n}\frac{(\overline{Y}-\ell_0)}{3}\frac{\Gamma(\frac{n}{2})}{\Gamma\left(\frac{n-1}{2}\right)}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{1}{2}}+\frac{(\overline{Y}-\ell_0)^3}{3^3}C\frac{\Gamma(\frac{1}{2}(n+3-1))}{\left\{\frac{(n-1)s^2}{2}\right\}^{\frac{1}{2}(n+3-1)}}$$

$$=\frac{(\overline{Y}-\ell_0)}{9n}\frac{\Gamma(\frac{n}{2})}{\Gamma\left(\frac{n-1}{2}\right)}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{1}{2}}+\frac{(\overline{Y}-\ell_0)^3}{27}\left\{\frac{2}{(n-1)s^2}\right\}^{-\frac{1}{2}(n-1)}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{1}{2}(n+2)}\frac{\Gamma(\frac{n+2}{2})}{\Gamma\left(\frac{n-1}{2}\right)}$$

$$=\frac{(\overline{Y}-\ell_0)}{9n}\frac{\Gamma(\frac{n}{2})}{\Gamma\left(\frac{n-1}{2}\right)}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{1}{2}}+\frac{(\overline{Y}-\ell_0)^3}{27}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{3}{2}}\frac{\Gamma(\frac{n+2}{2})}{\Gamma\left(\frac{n-1}{2}\right)}$$

$$=\frac{(\overline{Y}-\ell_0)}{9}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{1}{2}}\left(\frac{\Gamma(\frac{n}{2})}{n\Gamma\left(\frac{n-1}{2}\right)}+\frac{(\overline{Y}-\ell_0)^2}{3}\left\{\frac{2}{(n-1)s^2}\right\}\frac{\Gamma(\frac{n+2}{2})}{\Gamma\left(\frac{n-1}{2}\right)}\right).$$

$$m'_4=E_{\sigma^2}\left(\frac{1}{27n^2}\right)+\frac{2}{3n}E_{\sigma^2}\left(\frac{\overline{Y}-\ell_0}{3\sigma}\right)^2+E_{\sigma^2}\left(\frac{\overline{Y}-\ell_0}{3\sigma}\right)^4$$

$$=\left(\frac{1}{27n^2}\right)+\frac{2}{3n}E_{\sigma^2}\left(Z^2\right)+E_{\sigma^2}\left(Z^4\right)$$

$$=\left(\frac{1}{27n^2}\right)+\frac{2}{3n}\frac{(\overline{Y}-\ell_0)^2}{3^2}C\frac{\Gamma(\frac{1}{2}(n+2-1))}{\left\{\frac{(n-1)s^2}{2}\right\}^{\frac{1}{2}(n+2-1)}}+\frac{(\overline{Y}-\ell_0)^4}{3^4}C\frac{\Gamma(\frac{1}{2}(n+4-1))}{\left\{\frac{(n-1)s^2}{2}\right\}^{\frac{1}{2}(n+4-1)}}$$

$$=\left(\frac{1}{27n^2}\right)+\frac{2}{3n}\frac{(\overline{Y}-\ell_0)^2}{9}C\frac{\Gamma(\frac{1}{2}(n+1))}{\left\{\frac{(n-1)s^2}{2}\right\}^{\frac{1}{2}(n+1)}}+\frac{(\overline{Y}-\ell_0)^4}{81}C\frac{\Gamma(\frac{1}{2}(n+3))}{\left\{\frac{(n-1)s^2}{2}\right\}^{\frac{1}{2}(n+3)}}$$

$$=\left(\frac{1}{27n^2}\right)+\frac{2}{3n}\frac{(\bar{Y}-\ell_0)^2}{9}\left\{\frac{(n-1)s^2}{2}\right\}^{\frac{1}{2}(n-1)}\frac{1}{\Gamma\left(\frac{n-1}{2}\right)}\frac{\Gamma(\frac{1}{2}(n+1))}{\left\{\frac{(n-1)s^2}{2}\right\}^{\frac{1}{2}(n+1)}}+$$

$$\frac{(\bar{Y}-\ell_0)^4}{81}\left\{\frac{(n-1)s^2}{2}\right\}^{\frac{1}{2}(n-1)}\frac{1}{\Gamma\left(\frac{n-1}{2}\right)}\frac{\Gamma(\frac{1}{2}(n+3))}{\left\{\frac{(n-1)s^2}{2}\right\}^{\frac{1}{2}(n+3)}}$$

$$=\left(\frac{1}{27n^2}\right)+\frac{2}{3n}\frac{(\bar{Y}-\ell_0)^2}{9}\left\{\frac{2}{(n-1)s^2}\right\}^{-\frac{1}{2}(n-1)}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{1}{2}(n+1)}\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}+$$

$$\frac{(\bar{Y}-\ell_0)^4}{81}\left\{\frac{2}{(n-1)s^2}\right\}^{-\frac{1}{2}(n-1)}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{1}{2}(n+3)}\frac{\Gamma\left(\frac{n+3}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}$$

$$=\left(\frac{1}{27n^2}\right)+\frac{2(\bar{Y}-\ell_0)^2}{27n}\left\{\frac{2}{(n-1)s^2}\right\}\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}+\frac{(\bar{Y}-\ell_0)^4}{81}\left\{\frac{2}{(n-1)s^2}\right\}^2\frac{\Gamma\left(\frac{n+3}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}$$

$$=\frac{1}{27}\left(\frac{1}{n^2}+(\bar{Y}-\ell_0)^2\left\{\frac{2}{(n-1)s^2}\right\}\left\{\frac{2}{n}\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}+\frac{(\bar{Y}-\ell_0)^2}{3}\left\{\frac{2}{(n-1)s^2}\right\}\frac{\Gamma\left(\frac{n+3}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\right\}\right).$$

$$m_2=m'_2-(m'_1)^2$$

$$=\frac{1}{9n}+\frac{(\bar{Y}-\ell_0)^2}{9s^2}-\left(\frac{(\bar{Y}-\ell_0)}{3}\frac{\Gamma(\frac{n}{2})}{\Gamma\left(\frac{n-1}{2}\right)}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{1}{2}}\right)^2$$

$$=\frac{1}{9n}+\frac{(\bar{Y}-\ell_0)^2}{9s^2}-\left(\frac{(\bar{Y}-\ell_0)^2}{9}\frac{\left(\Gamma(\frac{n}{2})\right)^2}{\left(\Gamma\left(\frac{n-1}{2}\right)\right)^2}\left\{\frac{2}{(n-1)s^2}\right\}\right)$$

$$=\frac{1}{9n}+\frac{2(\bar{Y}-\ell_0)^2}{9(n-1)s^2}\left\{\frac{(n-1)}{2}-\frac{\Gamma^2(\frac{n}{2})}{\Gamma^2\left(\frac{n-1}{2}\right)}\right\}$$

which proves the result in equation (4.3.6).

$$m_3=m'_3-3m_2m'_1-(m'_1)^3$$

$$
=\frac{(\bar{Y}-\ell_0)}{9}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{1}{2}}\left(\frac{\Gamma(\frac{n}{2})}{n\Gamma\left(\frac{n-1}{2}\right)}+\frac{(\bar{Y}-\ell_0)^2}{3}\left\{\frac{2}{(n-1)s^2}\right\}\frac{\Gamma(\frac{n+2}{2})}{\Gamma\left(\frac{n-1}{2}\right)}\right)
$$

$$
-3\frac{(\bar{Y}-\ell_0)}{3}\frac{\Gamma(\frac{n}{2})}{\Gamma\left(\frac{n-1}{2}\right)}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{1}{2}}\left(\frac{1}{9n}+\frac{2(\bar{Y}-\ell_0)^2}{9(n-1)s^2}\left\{\frac{(n-1)}{2}-\frac{\Gamma^2(\frac{n}{2})}{\Gamma^2\left(\frac{n-1}{2}\right)}\right\}\right)
$$

$$
-\left(\frac{(\bar{Y}-\ell_0)^3}{27}\frac{\Gamma^3(\frac{n}{2})}{\Gamma^3\left(\frac{n-1}{2}\right)}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{3}{2}}\right)
$$

$$
=\left(\frac{(\bar{Y}-\ell_0)}{9n}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{1}{2}}\frac{\Gamma(\frac{n}{2})}{\Gamma\left(\frac{n-1}{2}\right)}+\frac{(\bar{Y}-\ell_0)^3}{27}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{3}{2}}\frac{\Gamma(\frac{n+2}{2})}{\Gamma\left(\frac{n-1}{2}\right)}\right)
$$

$$
-\left(\frac{(\bar{Y}-\ell_0)}{9n}\frac{\Gamma(\frac{n}{2})}{\Gamma\left(\frac{n-1}{2}\right)}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{1}{2}}+\frac{(\bar{Y}-\ell_0)^3}{9}\frac{\Gamma(\frac{n}{2})}{\Gamma\left(\frac{n-1}{2}\right)}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{3}{2}}\left\{\frac{(n-1)}{2}-\frac{\Gamma^2(\frac{n}{2})}{\Gamma^2\left(\frac{n-1}{2}\right)}\right\}\right)
$$

$$
-\left(\frac{(\bar{Y}-\ell_0)^3}{27}\frac{\Gamma^3(\frac{n}{2})}{\Gamma^3\left(\frac{n-1}{2}\right)}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{3}{2}}\right)
$$

$$
=\left(\frac{(\bar{Y}-\ell_0)^3}{27}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{3}{2}}\frac{\Gamma(\frac{n+2}{2})}{\Gamma\left(\frac{n-1}{2}\right)}\right)-
$$

$$
\left(\frac{(\bar{Y}-\ell_0)^3}{9}\frac{\Gamma(\frac{n}{2})}{\Gamma\left(\frac{n-1}{2}\right)}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{3}{2}}\left\{\frac{(n-1)}{2}-\frac{\Gamma^2(\frac{n}{2})}{\Gamma^2\left(\frac{n-1}{2}\right)}\right\}\right)
$$

$$
-\left(\frac{(\bar{Y}-\ell_0)^3}{27}\frac{\Gamma^3(\frac{n}{2})}{\Gamma^3\left(\frac{n-1}{2}\right)}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{3}{2}}\right)
$$

$$
=\frac{(\bar{Y}-\ell_0)^3}{27}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{3}{2}}\left(\frac{\Gamma(\frac{n+2}{2})}{\Gamma\left(\frac{n-1}{2}\right)}-\frac{3\Gamma(\frac{n}{2})}{\Gamma\left(\frac{n-1}{2}\right)}\left\{\frac{(n-1)}{2}-\frac{\Gamma^2(\frac{n}{2})}{\Gamma^2\left(\frac{n-1}{2}\right)}\right\}-\frac{\Gamma^3(\frac{n}{2})}{\Gamma^3\left(\frac{n-1}{2}\right)}\right)
$$

$$
=\frac{(\bar{Y}-\ell_0)^3}{27}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{3}{2}}\left(\frac{\Gamma(\frac{n+2}{2})}{\Gamma\left(\frac{n-1}{2}\right)}-\frac{(n-1)}{2}\frac{3\Gamma(\frac{n}{2})}{\Gamma\left(\frac{n-1}{2}\right)}+\frac{3\Gamma^3(\frac{n}{2})}{\Gamma^3\left(\frac{n-1}{2}\right)}-\frac{\Gamma^3(\frac{n}{2})}{\Gamma^3\left(\frac{n-1}{2}\right)}\right)
$$

$$
=\frac{(\bar{Y}-\ell_0)^3}{27}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{3}{2}}\left(\frac{\Gamma(\frac{n+2}{2})}{\Gamma\left(\frac{n-1}{2}\right)}-\frac{3(n-1)}{2}\frac{\Gamma(\frac{n}{2})}{\Gamma\left(\frac{n-1}{2}\right)}+\frac{2\Gamma^3(\frac{n}{2})}{\Gamma^3\left(\frac{n-1}{2}\right)}\right)
$$

$$
=\frac{(\bar{Y}-\ell_0)^3}{27}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{3}{2}}\left(\frac{2\Gamma^3(\frac{n}{2})}{\Gamma^3\left(\frac{n-1}{2}\right)}+\frac{\Gamma(\frac{n+2}{2})}{\Gamma\left(\frac{n-1}{2}\right)}-\frac{3(n-1)}{2}\frac{\Gamma(\frac{n}{2})}{\Gamma\left(\frac{n-1}{2}\right)}\right)
$$

$$=\frac{(\bar{Y}-\ell_0)^3}{27}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{3}{2}}\left(\frac{2\Gamma^3(\frac{n}{2})}{\Gamma^3\left(\frac{n-1}{2}\right)}+\frac{n\Gamma(\frac{n}{2})}{2\Gamma\left(\frac{n-1}{2}\right)}-\frac{3(n-1)}{2}\frac{\Gamma(\frac{n}{2})}{\Gamma\left(\frac{n-1}{2}\right)}\right)$$

$$=\frac{(\bar{Y}-\ell_0)^3}{27}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{3}{2}}\frac{\Gamma(\frac{n}{2})}{\Gamma\left(\frac{n-1}{2}\right)}\left(\frac{2\Gamma^2(\frac{n}{2})}{\Gamma^2\left(\frac{n-1}{2}\right)}+\frac{n}{2}-\frac{3(n-1)}{2}\right)$$

$$=\frac{(\bar{Y}-\ell_0)^3}{27}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{3}{2}}\frac{\Gamma(\frac{n}{2})}{\Gamma\left(\frac{n-1}{2}\right)}\left(\frac{2\Gamma^2(\frac{n}{2})}{\Gamma^2\left(\frac{n-1}{2}\right)}-\frac{(2n-3)}{2}\right)$$

which proves the result in equation (4.3.7).

$$m_4 = m'_4 - 4m_3m'_1 - 6m_2(m'_1)^2 - (m'_1)^4$$

$$=\left\langle\frac{1}{27}\left(\frac{1}{n^2}+(\bar{Y}-\ell_0)^2\left\{\frac{2}{(n-1)s^2}\right\}\left\{\frac{2}{n}\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}+\frac{(\bar{Y}-\ell_0)^2}{3}\left\{\frac{2}{(n-1)s^2}\right\}\frac{\Gamma\left(\frac{n+3}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\right\}\right)\right\rangle$$

$$-4\left\langle\frac{(\bar{Y}-\ell_0)^3}{27}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{3}{2}}\frac{\Gamma(\frac{n}{2})}{\Gamma\left(\frac{n-1}{2}\right)}\left(\frac{2\Gamma^2(\frac{n}{2})}{\Gamma^2\left(\frac{n-1}{2}\right)}-\frac{(2n-3)}{2}\right)\right\rangle\left(\frac{(\bar{Y}-\ell_0)}{3}\frac{\Gamma(\frac{n}{2})}{\Gamma\left(\frac{n-1}{2}\right)}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{1}{2}}\right)$$

$$-6\left(\frac{1}{9n}+\frac{2(\bar{Y}-\ell_0)^2}{9(n-1)s^2}\left\{\frac{(n-1)}{2}-\frac{\Gamma^2(\frac{n}{2})}{\Gamma^2\left(\frac{n-1}{2}\right)}\right\}\right)\left(\frac{(\bar{Y}-\ell_0)}{3}\frac{\Gamma(\frac{n}{2})}{\Gamma\left(\frac{n-1}{2}\right)}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{1}{2}}\right)^2$$

$$-\left(\frac{(\bar{Y}-\ell_0)}{3}\frac{\Gamma(\frac{n}{2})}{\Gamma\left(\frac{n-1}{2}\right)}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{1}{2}}\right)^4$$

$$=\left(\frac{1}{27n^2}\right)+\frac{2(\bar{Y}-\ell_0)^2}{27n}\left\{\frac{2}{(n-1)s^2}\right\}\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}+\frac{(\bar{Y}-\ell_0)^4}{81}\left\{\frac{2}{(n-1)s^2}\right\}^2\frac{\Gamma\left(\frac{n+3}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}$$

$$-4\left\langle\frac{(\bar{Y}-\ell_0)^3}{27}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{3}{2}}\frac{\Gamma(\frac{n}{2})}{\Gamma\left(\frac{n-1}{2}\right)}\left(\frac{2\Gamma^2(\frac{n}{2})}{\Gamma^2\left(\frac{n-1}{2}\right)}-\frac{(2n-3)}{2}\right)\right\rangle\left(\frac{(\bar{Y}-\ell_0)}{3}\frac{\Gamma(\frac{n}{2})}{\Gamma\left(\frac{n-1}{2}\right)}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{1}{2}}\right)$$

$$-6\left(\frac{1}{9n}+\frac{2(\bar{Y}-\ell_0)^2}{9(n-1)s^2}\left\{\frac{(n-1)}{2}-\frac{\Gamma^2(\frac{n}{2})}{\Gamma^2\left(\frac{n-1}{2}\right)}\right\}\right)\left(\frac{(\bar{Y}-\ell_0)}{3}\frac{\Gamma(\frac{n}{2})}{\Gamma\left(\frac{n-1}{2}\right)}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{1}{2}}\right)^2$$

$$-\left(\frac{(\bar{Y}-\ell_0)}{3}\frac{\Gamma(\frac{n}{2})}{\Gamma\left(\frac{n-1}{2}\right)}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{1}{2}}\right)^4$$

$$=\left(\frac{1}{27n^2}\right)+\frac{2(\bar{Y}-\ell_0)^2}{27n}\left\{\frac{2}{(n-1)s^2}\right\}\frac{(n-1)}{2}+\frac{1}{3}\frac{(\bar{Y}-\ell_0)^4}{27}\left\{\frac{2}{(n-1)s^2}\right\}^2\frac{\left(\frac{n+1}{2}\right)\left(\frac{n-1}{2}\right)\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}$$

$$-4\left\langle\frac{(\bar{Y}-\ell_0)^3}{27}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{3}{2}}\frac{\Gamma(\frac{n}{2})}{\Gamma\left(\frac{n-1}{2}\right)}\left(\frac{2\Gamma^2(\frac{n}{2})}{\Gamma^2\left(\frac{n-1}{2}\right)}-\frac{(2n-3)}{2}\right)\right\rangle\left(\frac{(\bar{Y}-\ell_0)}{3}\frac{\Gamma(\frac{n}{2})}{\Gamma\left(\frac{n-1}{2}\right)}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{1}{2}}\right)$$

$$-6\left(\frac{1}{9n}+\frac{2(\bar{Y}-\ell_0)^2}{9(n-1)s^2}\left\{\frac{(n-1)}{2}-\frac{\Gamma^2(\frac{n}{2})}{\Gamma^2\left(\frac{n-1}{2}\right)}\right\}\right)\left(\frac{(\bar{Y}-\ell_0)}{3}\frac{\Gamma(\frac{n}{2})}{\Gamma\left(\frac{n-1}{2}\right)}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{1}{2}}\right)^2$$

$$-\left(\frac{(\bar{Y}-\ell_0)}{3}\frac{\Gamma(\frac{n}{2})}{\Gamma\left(\frac{n-1}{2}\right)}\left\{\frac{2}{(n-1)s^2}\right\}^{\frac{1}{2}}\right)^4$$

$$=\frac{1}{27n^2}+\frac{(n-1)(\bar{Y}-\ell_0)^2}{27n}\left\{\frac{2}{(n-1)s^2}\right\}+\frac{1}{3}\frac{(\bar{Y}-\ell_0)^4}{27}\left\{\frac{2}{(n-1)s^2}\right\}^2\left(\frac{n+1}{2}\right)\left(\frac{n-1}{2}\right)$$

$$-4\left\langle\frac{(\bar{Y}-\ell_0)^4}{27}\left\{\frac{2}{(n-1)s^2}\right\}^2\frac{\Gamma^2(\frac{n}{2})}{3\Gamma^2\left(\frac{n-1}{2}\right)}\left(\frac{2\Gamma^2(\frac{n}{2})}{\Gamma^2\left(\frac{n-1}{2}\right)}-\frac{(2n-3)}{2}\right)\right\rangle$$

$$-2\left(\frac{(\bar{Y}-\ell_0)^2}{27n}\left\{\frac{2}{(n-1)s^2}\right\}\frac{\Gamma^2(\frac{n}{2})}{\Gamma^2\left(\frac{n-1}{2}\right)}+\frac{(\bar{Y}-\ell_0)^4}{27}\left\{\frac{2}{(n-1)s^2}\right\}^2\frac{\Gamma^2(\frac{n}{2})}{\Gamma^2\left(\frac{n-1}{2}\right)}\left\{\frac{(n-1)}{2}-\frac{\Gamma^2(\frac{n}{2})}{\Gamma^2\left(\frac{n-1}{2}\right)}\right\}\right)$$

$$-\frac{1}{3}\left(\frac{(\bar{Y}-\ell_0)^4}{27}\frac{\Gamma^4(\frac{n}{2})}{\Gamma^4\left(\frac{n-1}{2}\right)}\left\{\frac{2}{(n-1)s^2}\right\}^2\right)$$

$$=\frac{1}{27n^2}+\frac{(\bar{Y}-\ell_0)^2}{27n}\left\{\frac{2}{(n-1)s^2}\right\}\left((n-1)-2\frac{\Gamma^2(\frac{n}{2})}{\Gamma^2\left(\frac{n-1}{2}\right)}\right)+\frac{1}{3}\frac{(\bar{Y}-\ell_0)^4}{27}\left\{\frac{2}{(n-1)s^2}\right\}^2\left(\frac{n+1}{2}\right)\left(\frac{n-1}{2}\right)$$

$$-4\frac{(\bar{Y}-\ell_0)^4}{27}\left\{\frac{2}{(n-1)s^2}\right\}^2\frac{2\Gamma^4(\frac{n}{2})}{3\Gamma^4\left(\frac{n-1}{2}\right)}+4\frac{(\bar{Y}-\ell_0)^4}{27}\left\{\frac{2}{(n-1)s^2}\right\}^2\frac{\Gamma^2(\frac{n}{2})}{3\Gamma^2\left(\frac{n-1}{2}\right)}\frac{(2n-3)}{2}$$

$$-2\frac{(\bar{Y}-\ell_0)^4}{27}\left\{\frac{2}{(n-1)s^2}\right\}^2\frac{\Gamma^2(\frac{n}{2})}{\Gamma^2\left(\frac{n-1}{2}\right)}\left\{\frac{(n-1)}{2}-\frac{\Gamma^2(\frac{n}{2})}{\Gamma^2\left(\frac{n-1}{2}\right)}\right\}$$

$$-\frac{1}{3}\left(\frac{(\bar{Y}-\ell_0)^4}{27}\frac{\Gamma^4(\frac{n}{2})}{\Gamma^4\left(\frac{n-1}{2}\right)}\left\{\frac{2}{(n-1)s^2}\right\}^2\right)$$

$$=\frac{1}{27n^2}+\frac{(\bar{Y}-\ell_0)^2}{27n}\left\{\frac{2}{(n-1)s^2}\right\}\left((n-1)-2\frac{\Gamma^2(\frac{n}{2})}{\Gamma^2\left(\frac{n-1}{2}\right)}\right)$$

$$+\frac{(\bar{Y}-\ell_0)^4}{27}\left\{\frac{2}{(n-1)s^2}\right\}^2\left(\frac{1}{3}\left(\frac{n+1}{2}\right)\left(\frac{n-1}{2}\right)-\frac{2(n-1)}{2}\frac{\Gamma^2(\frac{n}{2})}{\Gamma^2\left(\frac{n-1}{2}\right)}+\frac{2\Gamma^4(\frac{n}{2})}{\Gamma^4\left(\frac{n-1}{2}\right)}-\frac{8\Gamma^4(\frac{n}{2})}{3\Gamma^4\left(\frac{n-1}{2}\right)}+\frac{4\Gamma^2(\frac{n}{2})}{3\Gamma^2\left(\frac{n-1}{2}\right)}\frac{(2n-3)}{2}-\frac{1}{3}\frac{\Gamma^4(\frac{n}{2})}{\Gamma^4\left(\frac{n-1}{2}\right)}\right)$$

$$=\frac{1}{27n^2}+\frac{(\overline{Y}-\ell_0)^2}{27n}\left\{\frac{2}{(n-1)s^2}\right\}\left(\left(n-1\right)-2\frac{\Gamma^2\left(\frac{n}{2}\right)}{\Gamma^2\left(\frac{n-1}{2}\right)}\right)$$

$$+\frac{(\overline{Y}-\ell_0)^4}{27}\left\{\frac{2}{(n-1)s^2}\right\}^2\left(\frac{1}{3}\left(\tfrac{n+1}{2}\right)\left(\tfrac{n-1}{2}\right)+\frac{2\Gamma^4\left(\frac{n}{2}\right)}{\Gamma^4\left(\frac{n-1}{2}\right)}-\frac{8\Gamma^4\left(\frac{n}{2}\right)}{3\Gamma^4\left(\frac{n-1}{2}\right)}-\frac{2(n-1)}{2}\frac{\Gamma^2\left(\frac{n}{2}\right)}{\Gamma^2\left(\frac{n-1}{2}\right)}+\frac{4\Gamma^2\left(\frac{n}{2}\right)}{3\Gamma^2\left(\frac{n-1}{2}\right)}\frac{(2n-3)}{2}-\frac{1}{3}\frac{\Gamma^4\left(\frac{n}{2}\right)}{\Gamma^4\left(\frac{n-1}{2}\right)}\right)$$

$$=\frac{1}{27n^2}+\frac{(\overline{Y}-\ell_0)^2}{27n}\left\{\frac{2}{(n-1)s^2}\right\}\left(\left(n-1\right)-2\frac{\Gamma^2\left(\frac{n}{2}\right)}{\Gamma^2\left(\frac{n-1}{2}\right)}\right)$$

$$+\frac{(\overline{Y}-\ell_0)^4}{27}\left\{\frac{2}{(n-1)s^2}\right\}^2\left(\frac{1}{3}\left(\tfrac{n+1}{2}\right)\left(\tfrac{n-1}{2}\right)-\frac{\Gamma^2\left(\frac{n}{2}\right)}{\Gamma^2\left(\frac{n-1}{2}\right)}(n-1)+\frac{2\Gamma^2\left(\frac{n}{2}\right)}{3\Gamma^2\left(\frac{n-1}{2}\right)}(2n-3)-\frac{\Gamma^4\left(\frac{n}{2}\right)}{\Gamma^4\left(\frac{n-1}{2}\right)}\right)$$

$$=\frac{1}{27n^2}+\frac{(\overline{Y}-\ell_0)^2}{27n}\left\{\frac{2}{(n-1)s^2}\right\}\left(\left(n-1\right)-2\frac{\Gamma^2\left(\frac{n}{2}\right)}{\Gamma^2\left(\frac{n-1}{2}\right)}\right)$$

$$+\frac{(\overline{Y}-\ell_0)^4}{27}\left\{\frac{2}{(n-1)s^2}\right\}^2\left(\frac{1}{3}\left(\tfrac{n+1}{2}\right)\left(\tfrac{n-1}{2}\right)+\frac{\Gamma^2\left(\frac{n}{2}\right)}{\Gamma^2\left(\frac{n-1}{2}\right)}\left\{\frac{(n-3)}{3}-\frac{\Gamma^2\left(\frac{n}{2}\right)}{\Gamma^2\left(\frac{n-1}{2}\right)}\right\}\right)$$

$$=\frac{1}{27n^2}+\frac{(\overline{Y}-\ell_0)^2}{27n}\left\{\frac{2}{(n-1)s^2}\right\}\left(\left(n-1\right)-2\frac{\Gamma^2\left(\frac{n}{2}\right)}{\Gamma^2\left(\frac{n-1}{2}\right)}\right)$$

$$+\frac{(\overline{Y}-\ell_0)^4}{27(n-1)^2\left(s^2\right)^2}\left(\frac{1}{3}\left(n+1\right)\left(n-1\right)+\frac{4\Gamma^2\left(\frac{n}{2}\right)}{\Gamma^2\left(\frac{n-1}{2}\right)}\left\{\frac{(n-3)}{3}-\frac{\Gamma^2\left(\frac{n}{2}\right)}{\Gamma^2\left(\frac{n-1}{2}\right)}\right\}\right)$$

which proves the result in equation (4.3.8).

Proof of theorem 4.6.1

Using theorem 4.3.3 and $r=\frac{1}{2}$

$$E(d_{12}\mid\chi^2_{\nu_1},\chi^2_{\nu_2},\underline{Y_i})=E_{\chi^2_{\nu_1}}\left((\overline{Y}_1-\ell_0)\frac{1}{3}\sqrt{\frac{\chi^2_{\nu_1}}{\nu_1 {s_1}^2}}\right)-E_{\chi^2_{\nu_2}}\left((\overline{Y}_2-\ell_0)\frac{1}{3}\sqrt{\frac{\chi^2_{\nu_2}}{\nu_2 {s_2}^2}}\right)$$

$$=\frac{1}{3}\left((\overline{Y}_1-\ell_0)\frac{\sqrt{2}\Gamma\left(\frac{\nu_1+1}{2}\right)}{\sqrt{\nu_1 {s_1}^2}\Gamma\left(\frac{\nu_1}{2}\right)}-(\overline{Y}_2-\ell_0)\frac{\sqrt{2}\Gamma(\left(\frac{\nu_2+1}{2}\right)}{\sqrt{\nu_2 {s_2}^2}\Gamma\left(\frac{\nu_2}{2}\right)}\right)$$

$$=\frac{\sqrt{2}}{3}\left(\frac{(\overline{Y}_1-\ell_0)}{\sqrt{\nu_1 {s_1}^2}}\frac{\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)}-\frac{(\overline{Y}_2-\ell_0)}{\sqrt{\nu_2 {s_2}^2}}\frac{\Gamma(\left(\frac{\nu_2+1}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)}\right)$$

$$Var(d_{12}\mid\chi^2_{\nu_1},\chi^2_{\nu_2},\underline{Y_i})=E_{\chi^2_{\nu_1}\chi^2_{\nu_2}}(Var(d_{12})\mid\chi^2_{\nu_1},\chi^2_{\nu_2},\underline{Y_i})+Var_{\chi^2_{\nu_1}\chi^2_{\nu_2}}(E(d_{12})\mid\chi^2_{\nu_1},\chi^2_{\nu_2}$$

$$=\frac{1}{9}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)+Var_{\chi^2_{\nu_1}}\left(\frac{(\overline{Y}_1-\ell_0)}{3\sqrt{\nu_1 s_1{}^2}}(\sqrt{\chi^2_{\nu_1}})+\frac{(\overline{Y}_2-\ell_0)}{3\sqrt{\nu_2 s_2{}^2}}(\sqrt{\chi^2_{\nu_2}})\right)$$

$$=\frac{1}{9}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)+\left(\frac{(\overline{Y}_1-\ell_0)^2}{9\nu_1 s_1{}^2}Var(\sqrt{\chi^2_{\nu_1}})\right)+\left(\frac{(\overline{Y}_2-\ell_0)^2}{9\nu_2 s_2{}^2}Var(\sqrt{\chi^2_{\nu_2}})\right)$$

$$=\frac{1}{9}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)+\frac{1}{9}\left(\frac{(\overline{Y}_1-\ell_0)^2}{\nu_1 s_1{}^2}\left\{\nu_1-\frac{2\Gamma^2\left(\frac{\nu_1+1}{2}\right)}{\Gamma^2\left(\frac{\nu_1}{2}\right)}\right\}+\frac{(\overline{Y}_2-\ell_0)^2}{\nu_2 s_2{}^2}\left\{\nu_2-\frac{2\Gamma^2\left(\frac{\nu_2+1}{2}\right)}{\Gamma^2\left(\frac{\nu_2}{2}\right)}\right\}\right)$$

Proof of Theorem 4.6.2

From theorem 4.6.1

$$\mu'_1=(\overline{Y}_1-\ell_0)\frac{1}{3}\sqrt{\frac{\chi^2_{\nu_1}}{\nu_1 s_1{}^2}}-(\overline{Y}_2-\ell_0)\frac{1}{3}\sqrt{\frac{\chi^2_{\nu_2}}{\nu_2 s_2{}^2}}.$$

$$\mu'_2=\frac{1}{9}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)$$

and now

$$\mu'_3=\mu_3+3\mu_2\mu'_1+(\mu'_1)^3$$

$$=0+3\frac{1}{9}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\left(\frac{(\overline{Y}_1-\ell_0)}{3}\sqrt{\frac{\chi^2_{\nu_1}}{\nu_1 s_1{}^2}}-\frac{(\overline{Y}_2-\ell_0)}{3}\sqrt{\frac{\chi^2_{\nu_2}}{\nu_2 s_2{}^2}}\right)+\left(\frac{(\overline{Y}_1-\ell_0)}{3}\sqrt{\frac{\chi^2_{\nu_1}}{\nu_1 s_1{}^2}}-\frac{(\overline{Y}_2-\ell_0)}{3}\sqrt{\frac{\chi^2_{\nu_2}}{\nu_2 s_2{}^2}}\right)^3$$

$$=\frac{3}{9}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\left(\frac{(\overline{Y}_1-\ell_0)}{3}\sqrt{\frac{\chi^2_{\nu_1}}{\nu_1 s_1{}^2}}-\frac{(\overline{Y}_2-\ell_0)}{3}\sqrt{\frac{\chi^2_{\nu_2}}{\nu_2 s_2{}^2}}\right)+\frac{(\overline{Y}_1-\ell_0)^3(\chi^2_{\nu_1})^{\frac{3}{2}}}{27(\nu_1 s_1{}^2)^{\frac{3}{2}}}$$

$$-\frac{3(\overline{Y}_1-\ell_0)^2(\chi^2_{\nu_1})}{9(\nu_1 s_1{}^2)}\frac{(\overline{Y}_2-\ell_0)(\chi^2_{\nu_2})^{\frac{1}{2}}}{3(\nu_2 s_2{}^2)^{\frac{1}{2}}}+\frac{3(\overline{Y}_1-\ell_0)(\chi^2_{\nu_1})^{\frac{1}{2}}}{3(\nu_1 s_1{}^2)^{\frac{1}{2}}}\frac{(\overline{Y}_2-\ell_0)^2(\chi^2_{\nu_2})}{9(\nu_2 s_2{}^2)}-\frac{(\overline{Y}_2-\ell_0)^3(\chi^2_{\nu_2})^{\frac{3}{2}}}{27(\nu_2 s_2{}^2)^{\frac{3}{2}}}$$

$$\mu'_4=\mu_4+4\mu'_1\mu_3+6(\mu'_1)^2\mu_2+(\mu'_1)^4$$

$$=3\left\{\frac{1}{9}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\right\}^2+0+6\left((\overline{Y}_1-\ell_0)\frac{1}{3}\sqrt{\frac{\chi^2_{\nu_1}}{\nu_1 s_1{}^2}}-(\overline{Y}_2-\ell_0)\frac{1}{3}\sqrt{\frac{\chi^2_{\nu_2}}{\nu_2 s_2{}^2}}\right)^2\frac{1}{9}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)$$

$$+\left((\overline{Y}_1-\ell_0)\frac{1}{3}\sqrt{\frac{\chi^2_{\nu_1}}{\nu_1 s_1{}^2}}-(\overline{Y}_2-\ell_0)\frac{1}{3}\sqrt{\frac{\chi^2_{\nu_2}}{\nu_2 s_2{}^2}}\right)^4$$

$$=3\left\{\frac{1}{9}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\right\}^2+0+6\left(\frac{(\overline{Y_1}-\ell_0)^2(\chi^2_{\nu_1})}{9(\nu_1 s_1{}^2)}-\frac{2}{9}(\overline{Y_1}-\ell_0)(\overline{Y_2}-\ell_0)\sqrt{\frac{\chi^2_{\nu_1}\chi^2_{\nu_2}}{\nu_1 s_1{}^2\nu_2 s_2{}^2}}+\frac{(\overline{Y_2}-\ell_0)^2(\chi^2_{\nu_2})}{9(\nu_2 s_2{}^2)}\right)\frac{1}{9}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)$$

$$+\frac{(\overline{Y_1}-\ell_0)^4(\chi^2_{\nu_1})^2}{81(\nu_1 s_1{}^2)^2}+4\frac{(\overline{Y_1}-\ell_0)^3(\chi^2_{\nu_1})^{\frac{3}{2}}}{27(\nu_1 s_1{}^2)^{\frac{3}{2}}}(\overline{Y_2}-\ell_0)\frac{1}{3}\sqrt{\frac{\chi^2_{\nu_2}}{\nu_2 s_2{}^2}}+6\frac{(\overline{Y_1}-\ell_0)^2(\chi^2_{\nu_1})}{9(\nu_1 s_1{}^2)}\frac{(\overline{Y_2}-\ell_0)^2(\chi^2_{\nu_2})}{9(\nu_2 s_2{}^2)}$$

$$+4(\overline{Y_1}-\ell_0)\frac{1}{3}\sqrt{\frac{\chi^2_{\nu_1}}{\nu_1 s_1{}^2}}\frac{(\overline{Y_2}-\ell_0)^3(\chi^2_{\nu_2})^{\frac{3}{2}}}{27(\nu_2 s_2{}^2)^{\frac{3}{2}}}+\frac{(\overline{Y_2}-\ell_0)^4(\chi^2_{\nu_2})^2}{81(\nu_2 s_2{}^2)^2}$$

$$=\frac{1}{27}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)^2+\frac{2}{27}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\left(\frac{(\overline{Y_1}-\ell_0)^2(\chi^2_{\nu_1})}{(\nu_1 s_1{}^2)}-2(\overline{Y_1}-\ell_0)(\overline{Y_2}-\ell_0)\sqrt{\frac{\chi^2_{\nu_1}\chi^2_{\nu_2}}{\nu_1 s_1{}^2\nu_2 s_2{}^2}}+\frac{(\overline{Y_2}-\ell_0)^2(\chi^2_{\nu_2})}{(\nu_2 s_2{}^2)}\right)$$

$$+\frac{(\overline{Y_1}-\ell_0)^4(\chi^2_{\nu_1})^2}{81(\nu_1 s_1{}^2)^2}+4\frac{(\overline{Y_1}-\ell_0)^3(\chi^2_{\nu_1})^{\frac{3}{2}}}{81(\nu_1 s_1{}^2)^{\frac{3}{2}}}(\overline{Y_2}-\ell_0)\sqrt{\frac{\chi^2_{\nu_2}}{\nu_2 s_2{}^2}}+6\frac{(\overline{Y_1}-\ell_0)^2(\overline{Y_2}-\ell_0)^2(\chi^2_{\nu_1})(\chi^2_{\nu_2})}{81(\nu_1 s_1{}^2)(\nu_2 s_2{}^2)}$$

$$+4(\overline{Y_1}-\ell_0)(\overline{Y_2}-\ell_0)^3\frac{(\chi^2_{\nu_1})^{\frac{1}{2}}(\chi^2_{\nu_2})^{\frac{3}{2}}}{81(\nu_1 s_1{}^2)^{\frac{1}{2}}(\nu_2 s_2{}^2)^{\frac{3}{2}}}+\frac{(\overline{Y_2}-\ell_0)^4(\chi^2_{\nu_2})^2}{81(\nu_2 s_2{}^2)^2}.$$

Proof of Theorem 4.6.3

From theorem 4.6.1

$$m'_1=\frac{\sqrt{2}}{3}\left(\frac{(\overline{Y_1}-\ell_0)}{\sqrt{\nu_1 s_1{}^2}}\frac{\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)}-\frac{(\overline{Y_2}-\ell_0)}{\sqrt{\nu_2 s_2{}^2}}\frac{\Gamma(\left(\frac{\nu_2+1}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)}\right)$$

$$m_2=\frac{1}{9}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)+\frac{1}{9}\left(\frac{(\overline{Y_1}-\ell_0)^2}{\nu_1 s_1{}^2}\left\{\nu_1-\frac{2\Gamma^2\left(\frac{\nu_1+1}{2}\right)}{\Gamma^2\left(\frac{\nu_1}{2}\right)}\right\}+\frac{(\overline{Y_2}-\ell_0)^2}{\nu_2 s_2{}^2}\left\{\nu_2-\frac{2\Gamma^2\left(\frac{\nu_2+1}{2}\right)}{\Gamma^2\left(\frac{\nu_2}{2}\right)}\right\}\right)$$

and now

$$m'_3=\frac{3}{9}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\left(\frac{(\overline{Y_1}-\ell_0)}{3}\sqrt{\frac{E(\chi^2_{\nu_1})}{\nu_1 s_1{}^2}}-\frac{(\overline{Y_2}-\ell_0)}{3}\sqrt{\frac{E(\chi^2_{\nu_2})}{\nu_2 s_2{}^2}}\right)+\frac{(\overline{Y_1}-\ell_0)^3E(\chi^2_{\nu_1})^{\frac{3}{2}}}{27(\nu_1 s_1{}^2)^{\frac{3}{2}}}$$

$$-\frac{3(\overline{Y_1}-\ell_0)^2 E(\chi^2_{\nu_1})}{9(\nu_1 {s_1}^2)}\frac{(\overline{Y_2}-\ell_0)E(\chi^2_{\nu_2})^{\frac{1}{2}}}{3(\nu_2 {s_2}^2)^{\frac{1}{2}}}+\frac{3(\overline{Y_1}-\ell_0)E(\chi^2_{\nu_1})^{\frac{1}{2}}}{3(\nu_1 {s_1}^2)^{\frac{1}{2}}}\frac{(\overline{Y_2}-\ell_0)^2 E(\chi^2_{\nu_2})}{9(\nu_2 {s_2}^2)}-\frac{(\overline{Y_2}-\ell_0)^3 E(\chi^2_{\nu_2})^{\frac{3}{2}}}{27(\nu_2 {s_2}^2)^{\frac{3}{2}}}$$

$$m'_3=\frac{3}{9}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\left(\frac{(\overline{Y_1}-\ell_0)}{3\sqrt{\nu_1 {s_1}^2}}\frac{\sqrt{2}\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)}-\frac{(\overline{Y_2}-\ell_0)}{3}\frac{\sqrt{2}\Gamma\left(\frac{\nu_2+1}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)}\right)+\frac{(\overline{Y_1}-\ell_0)^3 2^{\frac{3}{2}}}{27(\nu_1 {s_1}^2)^{\frac{3}{2}}}\frac{\Gamma\left(\frac{\nu_1+3}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)}$$

$$-\frac{3(\overline{Y_1}-\ell_0)^2\nu_1}{9(\nu_1 {s_1}^2)}\frac{(\overline{Y_2}-\ell_0)}{3(\nu_2 {s_2}^2)^{\frac{1}{2}}}\frac{\sqrt{2}\Gamma\left(\frac{\nu_2+1}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)}+\frac{3(\overline{Y_1}-\ell_0)}{3(\nu_1 {s_1}^2)^{\frac{1}{2}}}\frac{\sqrt{2}\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)}\frac{(\overline{Y_2}-\ell_0)^2\nu_2}{9(\nu_2 {s_2}^2)}-\frac{(\overline{Y_2}-\ell_0)^3}{27(\nu_2 {s_2}^2)^{\frac{3}{2}}}\frac{2^{\frac{3}{2}}\Gamma\left(\frac{\nu_2+3}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)}$$

$m_3 = m'_3 - 3m_2m'_1 - (m'_1)^2$

$$=\frac{3}{9}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\left(\frac{(\overline{Y_1}-\ell_0)}{3\sqrt{\nu_1 {s_1}^2}}\frac{\sqrt{2}\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)}-\frac{(\overline{Y_2}-\ell_0)}{3}\frac{\sqrt{2}\Gamma\left(\frac{\nu_2+1}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)}\right)+\frac{(\overline{Y_1}-\ell_0)^3 2^{\frac{3}{2}}}{27(\nu_1 {s_1}^2)^{\frac{3}{2}}}\frac{\Gamma\left(\frac{\nu_1+3}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)}$$

$$-\frac{3(\overline{Y_1}-\ell_0)^2\nu_1}{9(\nu_1 {s_1}^2)}\frac{(\overline{Y_2}-\ell_0)}{3(\nu_2 {s_2}^2)^{\frac{1}{2}}}\frac{\sqrt{2}\Gamma\left(\frac{\nu_2+1}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)}+\frac{3(\overline{Y_1}-\ell_0)}{3(\nu_1 {s_1}^2)^{\frac{1}{2}}}\frac{\sqrt{2}\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)}\frac{(\overline{Y_2}-\ell_0)^2\nu_2}{9(\nu_2 {s_2}^2)}-\frac{(\overline{Y_2}-\ell_0)^3}{27(\nu_2 {s_2}^2)^{\frac{3}{2}}}\frac{2^{\frac{3}{2}}\Gamma\left(\frac{\nu_2+3}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)}$$

$$-\left\{\frac{3}{9}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)+\frac{3}{9}\left(\frac{(\overline{Y_1}-\ell_0)^2}{\nu_1 {s_1}^2}\left\{\nu_1-\frac{2\Gamma^2\left(\frac{\nu_1+1}{2}\right)}{\Gamma^2\left(\frac{\nu_1}{2}\right)}\right\}+\frac{(\overline{Y_2}-\ell_0)^2}{\nu_2 {s_2}^2}\left\{\nu_2-\frac{2\Gamma^2\left(\frac{\nu_2+1}{2}\right)}{\Gamma^2\left(\frac{\nu_2}{2}\right)}\right\}\right)\right\}\times$$

$$\frac{\sqrt{2}}{3}\left(\frac{(\overline{Y_1}-\ell_0)}{\sqrt{\nu_1 {s_1}^2}}\frac{\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)}-\frac{(\overline{Y_2}-\ell_0)}{\sqrt{\nu_2 {s_2}^2}}\frac{\Gamma\left(\frac{\nu_2+1}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)}\right)-\frac{2^{\frac{3}{2}}}{27}\left(\frac{(\overline{Y_1}-\ell_0)}{\sqrt{\nu_1 {s_1}^2}}\frac{\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)}-\frac{(\overline{Y_2}-\ell_0)}{\sqrt{\nu_2 {s_2}^2}}\frac{\Gamma\left(\frac{\nu_2+1}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)}\right)^3$$

$$\therefore m_3=\frac{3}{9}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\left(\frac{(\overline{Y_1}-\ell_0)}{3\sqrt{\nu_1 {s_1}^2}}\frac{\sqrt{2}\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)}-\frac{(\overline{Y_2}-\ell_0)}{3}\frac{\sqrt{2}\Gamma\left(\frac{\nu_2+1}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)}\right)+\frac{(\overline{Y_1}-\ell_0)^3 2^{\frac{3}{2}}}{27(\nu_1 {s_1}^2)^{\frac{3}{2}}}\frac{\Gamma\left(\frac{\nu_1+3}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)}$$

$$-\frac{3(\overline{Y_1}-\ell_0)^2\nu_1}{9(\nu_1 {s_1}^2)}\frac{(\overline{Y_2}-\ell_0)}{3(\nu_2 {s_2}^2)^{\frac{1}{2}}}\frac{\sqrt{2}\Gamma\left(\frac{\nu_2+1}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)}+\frac{3(\overline{Y_1}-\ell_0)}{3(\nu_1 {s_1}^2)^{\frac{1}{2}}}\frac{\sqrt{2}\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)}\frac{(\overline{Y_2}-\ell_0)^2\nu_2}{9(\nu_2 {s_2}^2)}-\frac{(\overline{Y_2}-\ell_0)^3}{27(\nu_2 {s_2}^2)^{\frac{3}{2}}}\frac{2^{\frac{3}{2}}\Gamma\left(\frac{\nu_2+3}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)}$$

$$-\frac{3}{9}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\frac{\sqrt{2}}{3}\frac{(\overline{Y_1}-\ell_0)}{\sqrt{\nu_1 {s_1}^2}}\frac{\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)}+\frac{3}{9}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\frac{\sqrt{2}}{3}\frac{(\overline{Y_2}-\ell_0)}{\sqrt{\nu_2 {s_2}^2}}\frac{\Gamma\left(\frac{\nu_2+1}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)}$$

$$-\frac{3}{9}\frac{(\overline{Y_1}-\ell_0)^2}{\nu_1 {s_1}^2}\left\{\nu_1-\frac{2\Gamma^2\left(\frac{\nu_1+1}{2}\right)}{\Gamma^2\left(\frac{\nu_1}{2}\right)}\right\}\frac{\sqrt{2}}{3}\frac{(\overline{Y_1}-\ell_0)}{\sqrt{\nu_1 {s_1}^2}}\frac{\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)}+\frac{3}{9}\frac{(\overline{Y_1}-\ell_0)^2}{\nu_1 {s_1}^2}\left\{\nu_1-\frac{2\Gamma^2\left(\frac{\nu_1+1}{2}\right)}{\Gamma^2\left(\frac{\nu_1}{2}\right)}\right\}\frac{\sqrt{2}}{3}\frac{(\overline{Y_2}-\ell_0)}{\sqrt{\nu_2 {s_2}^2}}\frac{\Gamma\left(\frac{\nu_2+1}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)}$$

$$-\frac{3}{9}\frac{(\overline{Y}_2-\ell_0)^2}{\nu_2 s_2{}^2}\left\{\nu_2-\frac{2\Gamma^2\left(\frac{\nu_2+1}{2}\right)}{\Gamma^2\left(\frac{\nu_2}{2}\right)}\right\}\frac{\sqrt{2}}{3}\frac{(\overline{Y}_1-\ell_0)}{\sqrt{\nu_1 s_1{}^2}}\frac{\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)}+\frac{3}{9}\frac{(\overline{Y}_2-\ell_0)^2}{\nu_2 s_2{}^2}\left\{\nu_2-\frac{2\Gamma^2\left(\frac{\nu_2+1}{2}\right)}{\Gamma^2\left(\frac{\nu_2}{2}\right)}\right\}\frac{\sqrt{2}}{3}\frac{(\overline{Y}_2-\ell_0)}{\sqrt{\nu_2 s_2{}^2}}\frac{\Gamma\left(\frac{\nu_2+1}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)}$$

$$-\frac{2^{\frac{3}{2}}}{27}\frac{\left(\overline{Y}_1-\ell_0\right)^3}{\left(\nu_1 s_1{}^2\right)^{\frac{3}{2}}}\frac{\Gamma^3\left(\frac{\nu_1+1}{2}\right)}{\Gamma^3\left(\frac{\nu_1}{2}\right)}+\frac{2^{\frac{3}{2}}}{9}\frac{\left(\overline{Y}_1-\ell_0\right)^2}{\left(\nu_1 s_1{}^2\right)}\frac{\Gamma^2\left(\frac{\nu_1+1}{2}\right)}{\Gamma^2\left(\frac{\nu_1}{2}\right)}\frac{(\overline{Y}_2-\ell_0)}{\sqrt{\nu_2 s_2{}^2}}\frac{\Gamma\left(\frac{\nu_2+1}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)}$$

$$-\frac{2^{\frac{3}{2}}}{9}\frac{(\overline{Y}_1-\ell_0)}{\sqrt{\nu_1 s_1{}^2}}\frac{\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)}\frac{\left(\overline{Y}_2-\ell_0\right)^2}{\left(\nu_2 s_2{}^2\right)}\frac{\Gamma^2\left(\frac{\nu_2+1}{2}\right)}{\Gamma^2\left(\frac{\nu_2}{2}\right)}+\frac{2^{\frac{3}{2}}}{27}\frac{\left(\overline{Y}_2-\ell_0\right)^3}{\left(\nu_2 s_2{}^2\right)^{\frac{3}{2}}}\frac{\Gamma^3\left(\frac{\nu_2+1}{2}\right)}{\Gamma^3\left(\frac{\nu_2}{2}\right)}.$$

Let us look at $\left(\overline{Y}_1-\ell_0\right)\left(\overline{Y}_2-\ell_0\right)^2$ term

$$-\frac{\sqrt{2}}{9}(\overline{Y}_1-\ell_0)\left(\overline{Y}_2-\ell_0\right)^2\frac{\Gamma\left(\frac{\nu_1+1}{2}\right)}{\left(\nu_1 s_1{}^2\right)}\frac{\nu_2}{\Gamma\left(\frac{\nu_1}{2}\right)\left(\nu_2 s_2{}^2\right)}-\frac{3}{9}\frac{\sqrt{2}}{3}(\overline{Y}_1-\ell_0)\left(\overline{Y}_2-\ell_0\right)^2\frac{\Gamma\left(\frac{\nu_1+1}{2}\right)}{\left(\nu_1 s_1{}^2\right)}\frac{\nu_2}{\Gamma\left(\frac{\nu_1}{2}\right)\left(\nu_2 s_2{}^2\right)}$$

$$+\frac{6}{9}\frac{\sqrt{2}}{3}(\overline{Y}_1-\ell_0)\left(\overline{Y}_2-\ell_0\right)^2\frac{\Gamma\left(\frac{\nu_1+1}{2}\right)\Gamma^2\left(\frac{\nu_2+1}{2}\right)}{\sqrt{\nu_1 s_1{}^2}\left(\nu_2 s_2{}^2\right)\Gamma\left(\frac{\nu_1}{2}\right)\Gamma^2\left(\frac{\nu_2}{2}\right)}-\frac{2^{\frac{3}{2}}}{9}(\overline{Y}_1-\ell_0)\left(\overline{Y}_2-\ell_0\right)^2\frac{\Gamma\left(\frac{\nu_1+1}{2}\right)\Gamma^2\left(\frac{\nu_2+1}{2}\right)}{\sqrt{\nu_1 s_1{}^2}\left(\nu_2 s_2{}^2\right)\Gamma\left(\frac{\nu_1}{2}\right)\Gamma^2\left(\frac{\nu_2}{2}\right)}$$

$$=0$$

Also the $\left(\overline{Y}_1-\ell_0\right)^2\left(\overline{Y}_2-\ell_0\right)$ term is zero

The $\left(\overline{Y}_1-\ell_0\right)^3$ term is

$$\frac{(\overline{Y}_1-\ell_0)^3}{27}\left(\frac{2}{\nu_1 s_1{}^2}\right)^{\frac{3}{2}}\frac{\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)}\left\{\frac{2\Gamma^2\left(\frac{\nu_1+1}{2}\right)}{\Gamma^2\left(\frac{\nu_1}{2}\right)}-\frac{(2\nu_1-1)}{2}\right\}$$

and the $\left(\overline{Y}_2-\ell_0\right)^3$ term is

$$-\frac{(\overline{Y}_2-\ell_0)^3}{27}\left(\frac{2}{\nu_2 s_2{}^2}\right)^{\frac{3}{2}}\frac{\Gamma\left(\frac{\nu_2+1}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)}\left\{\frac{2\Gamma^2\left(\frac{\nu_2+1}{2}\right)}{\Gamma^2\left(\frac{\nu_2}{2}\right)}-\frac{(2\nu_2-1)}{2}\right\}$$

Therefore

$$m_3=\frac{(\overline{Y}_1-\ell_0)^3}{27}\left(\frac{2}{\nu_1 s_1{}^2}\right)^{\frac{3}{2}}\frac{\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)}\left\{\frac{2\Gamma^2\left(\frac{\nu_1+1}{2}\right)}{\Gamma^2\left(\frac{\nu_1}{2}\right)}-\frac{(2\nu_1-1)}{2}\right\}-\frac{(\overline{Y}_2-\ell_0)^3}{27}\left(\frac{2}{\nu_2 s_2{}^2}\right)^{\frac{3}{2}}\frac{\Gamma\left(\frac{\nu_2+1}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)}\left\{\frac{2\Gamma^2\left(\frac{\nu_2+1}{2}\right)}{\Gamma^2\left(\frac{\nu_2}{2}\right)}-\frac{(2\nu_2-1)}{2}\right\}$$

The fourth moment

$$\mu_4 = 3(\mu_2)^2 = 3\left(\frac{1}{9}\right)^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)^2 = \frac{1}{27}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)^2$$

$$\mu'_4 = \mu_4 + 4\mu'_1\mu_3 + 6(\mu'_1)^2\mu_2 + (\mu'_1)^4$$

$$\mu'_4 = \frac{1}{27}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)^2 + 0 + 6\left((\overline{Y}_1-\ell_0)\frac{1}{3}\sqrt{\frac{\chi^2_{\nu_1}}{\nu_1 s_1{}^2}} - (\overline{Y}_2-\ell_0)\frac{1}{3}\sqrt{\frac{\chi^2_{\nu_2}}{\nu_2 s_2{}^2}}\right)^2 \frac{1}{9}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)$$

$$+\left((\overline{Y}_1-\ell_0)\frac{1}{3}\sqrt{\frac{\chi^2_{\nu_1}}{\nu_1 s_1{}^2}} - (\overline{Y}_2-\ell_0)\frac{1}{3}\sqrt{\frac{\chi^2_{\nu_2}}{\nu_2 s_2{}^2}}\right)^4$$

Consider

$$E_{\chi^2_{\nu_1}\chi^2_{\nu_2}}\left((\overline{Y}_1-\ell_0)\frac{1}{3}\sqrt{\frac{\chi^2_{\nu_1}}{\nu_1 s_1{}^2}} - (\overline{Y}_2-\ell_0)\frac{1}{3}\sqrt{\frac{\chi^2_{\nu_2}}{\nu_2 s_2{}^2}}\right)^2$$

$$= E_{\chi^2_{\nu_1}\chi^2_{\nu_2}}\left((\overline{Y}_1-\ell_0)^2\frac{1}{9}\frac{\chi^2_{\nu_1}}{\nu_1 s_1{}^2} - 2(\overline{Y}_1-\ell_0)\frac{1}{3}\sqrt{\frac{\chi^2_{\nu_1}}{\nu_1 s_1{}^2}}(\overline{Y}_2-\ell_0)\frac{1}{3}\sqrt{\frac{\chi^2_{\nu_2}}{\nu_2 s_2{}^2}} + (\overline{Y}_2-\ell_0)^2\frac{1}{9}\frac{\chi^2_{\nu_2}}{\nu_2 s_2{}^2}\right)$$

$$= \left((\overline{Y}_1-\ell_0)^2\frac{1}{9}\frac{\nu_1}{\nu_1 s_1{}^2} - \frac{2(\overline{Y}_1-\ell_0)(\overline{Y}_2-\ell_0)}{9\sqrt{\nu_1 s_1{}^2}\sqrt{\nu_2 s_2{}^2}}\frac{\sqrt{2}\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)}\frac{\sqrt{2}\Gamma(\left(\frac{\nu_2+1}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)} + (\overline{Y}_2-\ell_0)^2\frac{1}{9}\frac{\nu_2}{\nu_2 s_2{}^2}\right)$$

$$= (\overline{Y}_1-\ell_0)^2\frac{1}{9}\frac{\nu_1}{\nu_1 s_1{}^2} + (\overline{Y}_2-\ell_0)^2\frac{1}{9}\frac{\nu_2}{\nu_2 s_2{}^2} - \frac{4(\overline{Y}_1-\ell_0)(\overline{Y}_2-\ell_0)}{9\sqrt{\nu_1 s_1{}^2}\sqrt{\nu_2 s_2{}^2}}\frac{\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)}\frac{\Gamma\left(\frac{\nu_2+1}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)}$$

$$E_{\chi^2_{\nu_1}\chi^2_{\nu_2}}\left((\overline{Y}_1-\ell_0)\frac{1}{3}\sqrt{\frac{\chi^2_{\nu_1}}{\nu_1 s_1{}^2}} - (\overline{Y}_2-\ell_0)\frac{1}{3}\sqrt{\frac{\chi^2_{\nu_2}}{\nu_2 s_2{}^2}}\right)^4$$

$$E_{\chi^2_{\nu_1}\chi^2_{\nu_2}}\{(\overline{Y}_1-\ell_0)^4\frac{1}{81}\frac{(\chi^2_{\nu_1})^2}{(\nu_1 s_1{}^2)^2} - \frac{4}{81}\frac{(\overline{Y}_1-\ell_0)^3(\overline{Y}_2-\ell_0)(\chi^2_{\nu_1})^{\frac{3}{2}}(\chi^2_{\nu_2})^{\frac{1}{2}}}{(\nu_1 s_1{}^2)^{\frac{3}{2}}(\nu_2 s_2{}^2)^{\frac{1}{2}}} + \frac{6(\overline{Y}_1-\ell_0)^2(\overline{Y}_2-\ell_0)^2\chi^2_{\nu_1}\chi^2_{\nu_2}}{81(\nu_1 s_1{}^2)(\nu_2 s_2{}^2)}$$

$$-\frac{4}{81}\frac{(\overline{Y}_1-\ell_0)(\overline{Y}_2-\ell_0)^3(\chi^2_{\nu_1})^{\frac{1}{2}}(\chi^2_{\nu_2})^{\frac{3}{2}}}{(\nu_1 s_1{}^2)^{\frac{1}{2}}(\nu_2 s_2{}^2)^{\frac{3}{2}}} + (\overline{Y}_2-\ell_0)^4\frac{1}{81}\frac{(\chi^2_{\nu_2})^2}{(\nu_2 s_2{}^2)^2}\}$$

$$= \{(\overline{Y}_1-\ell_0)^4\frac{1}{81}\frac{\nu_1(\nu_1+2)}{(\nu_1 s_1{}^2)^2} - \frac{8}{81}\frac{(\overline{Y}_1-\ell_0)^3(\overline{Y}_2-\ell_0)}{(\nu_1 s_1{}^2)^{\frac{3}{2}}(\nu_2 s_2{}^2)^{\frac{1}{2}}}\frac{(\nu_1+2)\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)}\frac{\Gamma\left(\frac{\nu_2+1}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)} + \frac{6(\overline{Y}_1-\ell_0)^2(\overline{Y}_2-\ell_0)^2\nu_1\nu_2}{81(\nu_1 s_1{}^2)(\nu_2 s_2{}^2)}$$

$$-\frac{8}{81}\frac{(\overline{Y_1}-\ell_0)(\overline{Y_2}-\ell_0)^3}{(\nu_1 s_1{}^2)^{\frac{1}{2}}(\nu_2 s_2{}^2)^{\frac{3}{2}}}\frac{\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)}\frac{(\nu_2+2)\Gamma\left(\frac{\nu_2+1}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)}+(\overline{Y_2}-\ell_0)^4\frac{1}{81}\frac{\nu_2(\nu_2+2)}{(\nu_2 s_2{}^2)^2}\}.$$

Fourth moment about zero unconditional

$$m'_4=\frac{1}{27}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)^2+\frac{6(\overline{Y_1}-\ell_0)^2\nu_1}{81(\nu_1 s_1{}^2)}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)+\frac{6(\overline{Y_2}-\ell_0)^2\nu_2}{81(\nu_2 s_2{}^2)}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)$$

$$-\frac{24(\overline{Y_1}-\ell_0)(\overline{Y_2}-\ell_0)}{81(\nu_1 s_1{}^2)^{\frac{1}{2}}(\nu_2 s_2{}^2)^{\frac{1}{2}}}\frac{\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)}\frac{\Gamma\left(\frac{\nu_2+1}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)+\frac{1}{81}\frac{(\overline{Y_1}-\ell_0)^4\nu_1(\nu_1+2)}{(\nu_1 s_1{}^2)^2}$$

$$-\frac{8}{81}\frac{(\overline{Y_1}-\ell_0)^3(\overline{Y_2}-\ell_0)}{(\nu_1 s_1{}^2)^{\frac{3}{2}}(\nu_2 s_2{}^2)^{\frac{1}{2}}}\frac{(\nu_1+2)\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)}\frac{\Gamma\left(\frac{\nu_2+1}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)}+\frac{6(\overline{Y_1}-\ell_0)^2(\overline{Y_2}-\ell_0)^2\nu_1\nu_2}{81(\nu_1 s_1{}^2)(\nu_2 s_2{}^2)}$$

$$-\frac{8}{81}\frac{(\overline{Y_1}-\ell_0)(\overline{Y_2}-\ell_0)^3}{(\nu_1 s_1{}^2)^{\frac{1}{2}}(\nu_2 s_2{}^2)^{\frac{3}{2}}}\frac{\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)}\frac{(\nu_2+2)\Gamma\left(\frac{\nu_2+1}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)}+(\overline{Y_2}-\ell_0)^4\frac{1}{81}\frac{\nu_2(\nu_2+2)}{(\nu_2 s_2{}^2)^2}\}.$$

Fourth moment about the mean unconditional

$$m_4 \quad = m'_4-4m'_1 m_3-6\left(m'_1\right)^2 m_2-\left(m'_1\right)^4$$

$$\therefore -4m'_1 m_3=\frac{-4\sqrt{2}}{3}\left(\frac{(\overline{Y_1}-\ell_0)}{\sqrt{\nu_1 s_1{}^2}}\frac{\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)}-\frac{(\overline{Y_2}-\ell_0)}{\sqrt{\nu_2 s_2{}^2}}\frac{\Gamma\left(\frac{\nu_2+1}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)}\right)\times$$

$$\left(\frac{(\overline{Y_1}-\ell_0)^3}{27}\left(\frac{2}{\nu_1 s_1{}^2}\right)^{\frac{3}{2}}\frac{\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)}\left\{\frac{2\Gamma^2\left(\frac{\nu_1+1}{2}\right)}{\Gamma^2\left(\frac{\nu_1}{2}\right)}-\frac{(2\nu_1-1)}{2}\right\}-\frac{(\overline{Y_2}-\ell_0)^3}{27}\left(\frac{2}{\nu_2 s_2{}^2}\right)^{\frac{3}{2}}\frac{\Gamma\left(\frac{\nu_2+1}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)}\left\{\frac{2\Gamma^2\left(\frac{\nu_2+1}{2}\right)}{\Gamma^2\left(\frac{\nu_2}{2}\right)}-\frac{(2\nu_2-1)}{2}\right\}\right)$$

and

$$-6\left(m'_1\right)^2 m_2 \quad = -\frac{12}{9}\left(\frac{(\overline{Y_1}-\ell_0)}{\sqrt{\nu_1 s_1{}^2}}\frac{\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)}-\frac{(\overline{Y_2}-\ell_0)}{\sqrt{\nu_2 s_2{}^2}}\frac{\Gamma\left(\frac{\nu_2+1}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)}\right)^2\times$$

$$\left[\frac{1}{9}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)+\frac{1}{9}\left(\frac{(\overline{Y_1}-\ell_0)^2}{\nu_1 s_1{}^2}\left\{\nu_1-\frac{2\Gamma^2\left(\frac{\nu_1+1}{2}\right)}{\Gamma^2\left(\frac{\nu_1}{2}\right)}\right\}+\frac{(\overline{Y_2}-\ell_0)^2}{\nu_2 s_2{}^2}\left\{\nu_2-\frac{2\Gamma^2\left(\frac{\nu_2+1}{2}\right)}{\Gamma^2\left(\frac{\nu_2}{2}\right)}\right\}\right)\right]$$

$$=-\frac{12}{9}\left\{\frac{(\overline{Y_1}-\ell_0)^2}{\nu_1 s_1{}^2}\frac{\Gamma^2\left(\frac{\nu_1+1}{2}\right)}{\Gamma^2\left(\frac{\nu_1}{2}\right)}+\frac{(\overline{Y_2}-\ell_0)^2}{\nu_2 s_2{}^2}\frac{\Gamma^2\left(\frac{\nu_2+1}{2}\right)}{\Gamma^2\left(\frac{\nu_2}{2}\right)}-\frac{2(\overline{Y_1}-\ell_0)}{\sqrt{\nu_1 s_1{}^2}}\frac{(\overline{Y_2}-\ell_0)}{\sqrt{\nu_2 s_2{}^2}}\frac{\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)}\frac{\Gamma\left(\frac{\nu_2+1}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)}\right\}\times$$

$$\left[\frac{1}{9}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)+\frac{1}{9}\left(\frac{(\overline{Y_1}-\ell_0)^2}{\nu_1 s_1{}^2}\left\{\nu_1-\frac{2\Gamma^2\left(\frac{\nu_1+1}{2}\right)}{\Gamma^2\left(\frac{\nu_1}{2}\right)}\right\}+\frac{(\overline{Y_2}-\ell_0)^2}{\nu_2 s_2{}^2}\left\{\nu_2-\frac{2\Gamma^2\left(\frac{\nu_2+1}{2}\right)}{\Gamma^2\left(\frac{\nu_2}{2}\right)}\right\}\right)\right]$$

$$-\left(m_1^{'}\right)^4 = -\frac{4}{81}\left(\frac{(\overline{Y_1}-\ell_0)}{\sqrt{\nu_1 s_1^{\,2}}}\frac{\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)} - \frac{(\overline{Y_2}-\ell_0)}{\sqrt{\nu_2 s_2^{\,2}}}\frac{\Gamma\left(\frac{\nu_2+1}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)}\right)^4$$

$$= -\frac{4}{81}\left\{\frac{(\overline{Y_1}-\ell_0)^4}{\left(\nu_1 s_1^{\,2}\right)^2}\frac{\Gamma^4\left(\frac{\nu_1+1}{2}\right)}{\Gamma^4\left(\frac{\nu_1}{2}\right)} - 4\frac{(\overline{Y_1}-\ell_0)^3(\overline{Y_2}-\ell_0)}{(\nu_1 s_1^{\,2})^{\frac{3}{2}}(\nu_2 s_2^{\,2})^{\frac{1}{2}}}\frac{\Gamma^3\left(\frac{\nu_1+1}{2}\right)}{\Gamma^3\left(\frac{\nu_1}{2}\right)}\frac{\Gamma\left(\frac{\nu_2+1}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)}\right.$$

$$\left. + 6\frac{(\overline{Y_1}-\ell_0)^2(\overline{Y_2}-\ell_0)^2}{(\nu_1 s_1^{\,2})(\nu_2 s_2^{\,2})}\frac{\Gamma^2\left(\frac{\nu_1+1}{2}\right)}{\Gamma^2\left(\frac{\nu_1}{2}\right)}\frac{\Gamma^2\left(\frac{\nu_2+1}{2}\right)}{\Gamma^2\left(\frac{\nu_2}{2}\right)} - 4\frac{(\overline{Y_1}-\ell_0)(\overline{Y_2}-\ell_0)^3}{(\nu_1 s_1^{\,2})^{\frac{1}{2}}(\nu_2 s_2^{\,2})^{\frac{3}{2}}}\frac{\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)}\frac{\Gamma^3\left(\frac{\nu_2+1}{2}\right)}{\Gamma^3\left(\frac{\nu_2}{2}\right)} + \frac{(\overline{Y_2}-\ell_0)^4}{\left(\nu_2 s_2^{\,2}\right)^2}\frac{\Gamma^4\left(\frac{\nu_2+1}{2}\right)}{\Gamma^4\left(\frac{\nu_2}{2}\right)}\right.$$

$$m_4 = \frac{1}{27}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)^2 + \frac{2}{27}(\overline{Y_1}-\ell_0)^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\frac{1}{(\nu_1 s_1^{\,2})}\left\{\nu_1 - \frac{2\Gamma^2\left(\frac{\nu_1+1}{2}\right)}{\Gamma^2\left(\frac{\nu_1}{2}\right)}\right\}$$

$$+\frac{2}{27}(\overline{Y_2}-\ell_0)^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\frac{1}{(\nu_2 s_2^{\,2})}\left\{\nu_2 - \frac{2\Gamma^2\left(\frac{\nu_2+1}{2}\right)}{\Gamma^2\left(\frac{\nu_2}{2}\right)}\right\}$$

$$+\frac{(\overline{Y_1}-\ell_0)^4}{27(\nu_1 s_1^{\,2})^2}\left\{\frac{\nu_1(\nu_1+2)}{3} + 4\frac{\Gamma^2\left(\frac{\nu_1+1}{2}\right)}{\Gamma^2\left(\frac{\nu_1}{2}\right)}\left[\frac{1}{3}(\nu_1-2) - \frac{\Gamma^2\left(\frac{\nu_1+1}{2}\right)}{\Gamma^2\left(\frac{\nu_1}{2}\right)}\right]\right\}$$

$$+\frac{(\overline{Y_2}-\ell_0)^4}{27(\nu_2 s_2^{\,2})^2}\left\{\frac{\nu_2(\nu_2+2)}{3} + 4\frac{\Gamma^2\left(\frac{\nu_2+1}{2}\right)}{\Gamma^2\left(\frac{\nu_2}{2}\right)}\left[\frac{1}{3}(\nu_2-2) - \frac{\Gamma^2\left(\frac{\nu_2+1}{2}\right)}{\Gamma^2\left(\frac{\nu_2}{2}\right)}\right]\right\}$$

$$+\frac{2(\overline{Y_1}-\ell_0)^2(\overline{Y_2}-\ell_0)^2}{27(\nu_1 s_1^{\,2})(\nu_2 s_2^{\,2})}\left\{\nu_1\nu_2 + 4\frac{\Gamma^2\left(\frac{\nu_1+1}{2}\right)}{\Gamma^2\left(\frac{\nu_1}{2}\right)}\frac{\Gamma^2\left(\frac{\nu_2+1}{2}\right)}{\Gamma^2\left(\frac{\nu_2}{2}\right)} - \frac{2\Gamma^2\left(\frac{\nu_1+1}{2}\right)\nu_2}{\Gamma^2\left(\frac{\nu_1}{2}\right)} - \frac{2\Gamma^2\left(\frac{\nu_2+1}{2}\right)\nu_1}{\Gamma^2\left(\frac{\nu_2}{2}\right)}\right\}$$

where

$\nu_1 = (n_1 - 1)$ and $\nu_2 = (n_2 - 1)$.

Proof of Theorem 4.7.1

The Fisher Information matrix of $\underline{\theta} = [\mu, \sigma]'$ per unit observation can be derived as follows:

The likelihood function is defined as

$$L(\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2\sigma^2}(Y-\mu)^2\right)$$

and

$$\ell n L = \text{Const} + \ell n\left(\frac{1}{\sigma}\right) - (Y-\mu)^2 / 2\sigma^2.$$

Also

$$\frac{\partial \ell n L}{\partial \mu} = 2(Y-\mu)/2\sigma^2 = \frac{(Y-\mu)}{\sigma^2}$$

and

$$\frac{\partial^2 \ell n L}{(\partial\mu)^2} = \frac{-1}{\sigma^2}. \text{ Therefore } -E\left(\frac{\partial^2 \ell n L}{(\partial\mu)^2}\right) = \frac{1}{\sigma^2}.$$

$$\frac{\partial^2 \ell n L}{\partial\mu\partial\sigma} = \frac{(Y-\mu)\sigma}{(\sigma^2)^2} = \frac{(Y-\mu)}{\sigma^3} \text{ and}$$

$$-E\left(\frac{\partial^2 \ell n L}{\partial\mu\partial\sigma}\right) = E\left\{\frac{(Y-\mu)\sigma}{(\sigma^2)^2}\right\} = \left\{\frac{(E(Y)-\mu)\sigma}{(\sigma^2)^2}\right\} = 0.$$

Further

$$\frac{\partial \ell n L}{\partial \sigma} = -\frac{1}{\sigma} + \left\{\frac{(Y-\mu)^2}{(\sigma^3)}\right\}$$

$$\frac{\partial^2 \ell n L}{(\partial\sigma)^2} = \frac{1}{\sigma^2} + (Y-\mu)^2\left\{\frac{-3\sigma^2}{(\sigma^3)^2}\right\}$$

$$= \frac{1}{\sigma^2} - \frac{3(Y-u)^2}{\sigma^4}$$

and

$$-E\left(\frac{\partial^2 \ell n L}{(\partial\sigma)^2}\right) = -\frac{1}{\sigma^2} + \frac{3E(Y-u)^2}{\sigma^4} = -\frac{1}{\sigma^2} + \frac{3\sigma^2}{\sigma^4} = -\frac{1}{\sigma^2} + \frac{3}{\sigma^2} = \frac{2}{\sigma^2}.$$

The Fisher Information matrix is then given by

$$F(\mu,\sigma) = F(\theta) = \begin{pmatrix} \sigma^{-2} & 0 \\ 0 & 2\sigma^{-2} \end{pmatrix} = \text{diag}[\sigma^{-2}, 2\sigma^{-2}] \text{ and}$$

$$F^{-1}(\mu,\sigma) = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \frac{1}{2}\sigma^2 \end{pmatrix} = \text{diag}\left[\sigma^2, \frac{1}{2}\sigma^2\right].$$

Consider

$$C_{pl} = t(\underline{\theta}) = \frac{\mu - \ell_0}{3\sigma},$$

then

$$\frac{\partial t(\underline{\theta})}{\partial \mu} = \frac{1}{3\sigma} \quad \text{and} \quad \frac{\partial t(\underline{\theta})}{\partial \sigma} = \frac{-(\mu - \ell_0)}{3} \frac{1}{\sigma^2}$$

and therefore

$$\nabla'_t(\underline{\theta}) = \begin{bmatrix} \frac{\partial t(\underline{\theta})}{\partial \mu} & \frac{\partial t(\underline{\theta})}{\partial \sigma} \end{bmatrix} = \begin{bmatrix} \frac{1}{3\sigma} & \frac{\ell_0 - \mu}{3\sigma^2} \end{bmatrix} = \frac{1}{3\sigma}\begin{bmatrix} 1 & \frac{\ell_0 - \mu}{\sigma} \end{bmatrix}.$$

Now

$$\nabla'_t(\underline{\theta}) F^{-1}(\underline{\theta}) = \frac{\sigma^2}{3\sigma}\begin{bmatrix} 1 & \frac{\ell_0 - \mu}{\sigma} \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$$

$$= \frac{\sigma}{3}\begin{bmatrix} 1 & \frac{\ell_0 - \mu}{2\sigma} \end{bmatrix}$$

and

$$\nabla'_t(\underline{\theta}) F^{-1}(\underline{\theta}) \nabla_t(\underline{\theta}) = \tfrac{1}{9}\left(1 + \frac{(\ell_0 - \mu)^2}{2\sigma^2}\right).$$

Also

$$\sqrt{\nabla'_t(\underline{\theta}) F^{-1}(\underline{\theta}) \nabla_t(\underline{\theta})} = \frac{1}{3}\sqrt{1 + \frac{(\ell_0 - \mu)^2}{2\sigma^2}}$$

which means that

$$\eta(\underline{\theta}) = \frac{\nabla'_t(\underline{\theta}) F^{-1}(\underline{\theta})}{\sqrt{\nabla'_t(\underline{\theta}) F^{-1}(\underline{\theta}) \nabla_t(\underline{\theta})}} = [\eta_1(\underline{\theta}) \quad \eta_2(\underline{\theta})]$$

where

$$\eta_1(\underline{\theta}) = \frac{\sigma}{\sqrt{1 + \frac{(\ell_0 - \mu)^2}{2\sigma^2}}}$$

and

$$\eta_2(\underline{\theta}) = \frac{\ell_0 - \mu}{2\sqrt{1 + \frac{(\ell_0 - \mu)^2}{2\sigma^2}}}$$

$$\eta(\underline{\theta}) = \frac{\nabla'_t(\underline{\theta}) F^{-1}(\underline{\theta})}{\sqrt{\nabla'_t(\underline{\theta}) F^{-1}(\underline{\theta}) \nabla_t(\underline{\theta})}} = [\eta_1(\underline{\theta}) \quad \eta_2(\underline{\theta})] \ .$$

where for a prior $\pi(\underline{\theta})$ to be a probability–matching prior, the differential equation

$$\sum_{\alpha=1}^{m} \frac{\partial}{\partial \theta_\alpha}\{\eta_\alpha(\underline{\theta})\pi(\underline{\theta})\} = 0$$

must be satisfied.

In what follows, it will be clear that if we set

$$\pi(\underline{\theta}) = \left\{1 + \frac{(\ell_0 - \mu)^2}{2\sigma^2}\right\}^{-\frac{1}{2}} \sigma^{-2}$$

then this will be the case, because

$$\pi(\underline{\theta})\eta_1(\underline{\theta}) = \left\{1 + \frac{(\ell_0 - \mu)^2}{2\sigma^2}\right\}^{-\frac{1}{2}} \sigma^{-2} \frac{\sigma}{\sqrt{1 + \frac{(\ell_0 - \mu)^2}{2\sigma^2}}} = \frac{\sigma^{-1}}{\left(1 + \frac{(\ell_0 - \mu)^2}{2\sigma^2}\right)} = \sigma^{-1}\left(1 + \frac{(\ell_0 - \mu)^2}{2\sigma^2}\right)^{-1}$$

and

$$\frac{\partial}{\partial\mu}\{\pi(\underline{\theta})\eta_1(\underline{\theta})\} = \frac{\sigma^{-1}\frac{(\ell_0 - \mu)}{\sigma^2}}{\left(1 + \frac{(\ell_0 - \mu)^2}{2\sigma^2}\right)^2}$$

$$= \frac{(\ell_0 - \mu)}{\sigma^3\left(1 + \frac{(\ell_0 - \mu)^2}{2\sigma^2}\right)^2}.$$

Also

$$\pi(\underline{\theta})\eta_2(\underline{\theta}) = \left\{1 + \frac{(\ell_0 - \mu)^2}{2\sigma^2}\right\}^{-\frac{1}{2}} \sigma^{-2} \frac{\ell_0 - \mu}{2\sqrt{1 + \frac{(\ell_0 - \mu)^2}{2\sigma^2}}} = \frac{(\ell_0 - \mu)\sigma^{-2}}{2\left(1 + \frac{(\ell_0 - \mu)^2}{2\sigma^2}\right)} = \frac{(\ell_0 - \mu)}{2\left(\sigma^2 + \frac{(\ell_0 - \mu)^2}{2}\right)}$$

$$\frac{\partial}{\partial\sigma}\{\pi(\theta)\eta_2(\theta)\} = \frac{(\ell_0 - \mu)}{2}\frac{\partial}{\partial\sigma}\left\{\frac{1}{\left(\sigma^2 + \frac{(\ell_0 - \mu)^2}{2}\right)}\right\}$$

$$= \frac{-(\ell_0 - \mu)}{2}\frac{2\sigma}{\left(\sigma^2 + \frac{(\ell_0 - \mu)^2}{2}\right)^2} = \frac{-(\ell_0 - \mu)}{\sigma^3\left(1 + \frac{(\ell_0 - \mu)^2}{2\sigma^2}\right)^2}.$$

Therefore

$$\sum_{\alpha=1}^{m}\frac{\partial}{\partial\theta_\alpha}\{\eta(\underline{\theta})\pi(\underline{\theta})\} = \frac{\partial}{\partial\mu}\{\pi(\underline{\theta})\eta_1(\underline{\theta})\} + \frac{\partial}{\partial\sigma}\{\pi(\underline{\theta})\eta_2(\underline{\theta})\}$$

$$= \frac{(\ell_0 - \mu)}{\sigma^3\left(1 + \frac{(\ell_0 - \mu)^2}{2\sigma^2}\right)^2} - \frac{(\ell_0 - \mu)}{\sigma^3\left(1 + \frac{(\ell_0 - \mu)^2}{2\sigma^2}\right)^2} = 0$$

which proves the theorem

$$\pi^M(\mu,\sigma)=\left\{1+\frac{(\ell_0-\mu)^2}{2\sigma^2}\right\}^{-\frac{1}{2}}\sigma^{-2}.$$

Proof of Corollary 4.7.1.1

From theorem (4.7.1)

$$\pi^M(\mu,\sigma)=\left\{1+\frac{(\ell_0-\mu)^2}{2\sigma^2}\right\}^{-\frac{1}{2}}\sigma^{-2}$$

let

$x=\sigma \quad y=\sigma^2$

then

$y=x^2$

$$\pi^M(\mu,y)\propto\left\{1+\frac{(\ell_0-\mu)^2}{2\sigma^2}\right\}^{-\frac{1}{2}}y^{-1}\left|\frac{dx}{dy}\right|$$

$$x=y^{\frac{1}{2}} \quad \frac{dx}{dy}=\frac{1}{2}y^{-\frac{1}{2}}$$

$$\pi^M(\mu,y)\propto\left\{1+\frac{(\ell_0-\mu)^2}{2\sigma^2}\right\}^{-\frac{1}{2}}y^{-1}\frac{1}{2}y^{-\frac{1}{2}}$$

$$\pi^M(\mu,y)\propto\frac{1}{2}\left\{1+\frac{(\ell_0-\mu)^2}{2\sigma^2}\right\}^{-\frac{1}{2}}y^{-\frac{3}{2}}$$

$$\pi^M(\mu,\sigma^2)\propto\left\{1+\frac{(\ell_0-\mu)^2}{2\sigma^2}\right\}^{-\frac{1}{2}}\left(\sigma^2\right)^{-\frac{3}{2}}$$

$$\pi^M(\mu,\sigma^2)=\left\{1+\frac{(\ell_0-\mu)^2}{2\sigma^2}\right\}^{-\frac{1}{2}}\sigma^{-3}$$

is therefore a probability matching prior.

Proof of Corollary 4.7.1.2

From theorem (4.7.1)

$$\eta(\underline{\theta}) = \frac{\nabla'_t(\underline{\theta})F^{-1}(\underline{\theta})}{\sqrt{\nabla'_t(\underline{\theta})F^{-1}(\underline{\theta})\nabla_t(\underline{\theta})}} = [\eta_1(\underline{\theta}) \quad \eta_2(\underline{\theta})]$$

where

$$\eta_1(\underline{\theta}) = \frac{\sigma}{\sqrt{1+\frac{(\ell_0-\mu)^2}{2\sigma^2}}}$$

and

$$\eta_2(\underline{\theta}) = \frac{\ell_0-\mu}{2\sqrt{1+\frac{(\ell_0-\mu)^2}{2\sigma^2}}}$$

For a prior $\pi(\theta)$ to be a probability–matching prior, the differential equation

$$\sum_{\alpha=1}^{m} \frac{\partial}{\partial\theta_\alpha}\{\eta_\alpha(\underline{\theta})\pi(\underline{\theta})\} = 0$$

must be satisfied.

In what follows it will be clear that if we set

$$\pi(\underline{\theta}) = \left\{1+\frac{(\ell_0-\mu)^2}{2\sigma^2}\right\}^{\frac{1}{2}}$$

then this will be the case, because

$$\pi(\underline{\theta})\eta_1(\underline{\theta}) = \left\{1+\frac{(\ell_0-\mu)^2}{2\sigma^2}\right\}\frac{\sigma}{\sqrt{1+\frac{(\ell_0-\mu)^2}{2\sigma^2}}} = \sigma$$

and

$$\frac{\partial}{\partial\mu}\{\pi(\underline{\theta})\eta_1(\underline{\theta})\} = 0$$

Also

$$\pi(\underline{\theta})\eta_2(\underline{\theta}) = \left\{1+\frac{(\ell_0-\mu)^2}{2\sigma^2}\right\}\frac{\ell_0-\mu}{2\sqrt{1+\frac{(\ell_0-\mu)^2}{2\sigma^2}}} = \frac{(\ell_0-\mu)}{2}$$

$$\frac{\partial}{\partial\sigma}\{\pi(\underline{\theta})\eta_2(\underline{\theta})\} = \frac{\partial}{\partial\sigma}\left\{\frac{(\ell_0-\mu)}{2}\right\}$$

$= 0$.

Therefore

$$\sum_{\alpha=1}^{m} \frac{\partial}{\partial \theta_\alpha} \{\eta(\underline{\theta})\pi(\underline{\theta})\} = \frac{\partial}{\partial \mu}\{\pi(\underline{\theta})\eta_1(\underline{\theta})\} + \frac{\partial}{\partial \sigma}\{\pi(\underline{\theta})\eta_2(\underline{\theta})\}$$

$$= 0 + 0 = 0 \ .$$

$\pi^M(\mu,\sigma) = \left\{1 + \frac{(\ell_0 - \mu)^2}{2\sigma^2}\right\}^{\frac{1}{2}}$ is therefore a probability matching prior.

Proof of Theorem 4.7.2

$C_{pl} = \frac{\mu - \ell_0}{3\sigma} = t(\underline{\theta}) \quad \therefore \mu = 3\sigma\, t(\underline{\theta}) + \ell_0$ and $\sigma = \frac{\mu - \ell_0}{3t(\underline{\theta})}$.

Define $\quad A = \frac{\partial(\mu,\sigma)}{\partial(t(\underline{\theta}),\sigma)} = \begin{bmatrix} 3\sigma & 3t(\underline{\theta}) \\ 0 & 1 \end{bmatrix}$.

Hence the Fisher information matrix under the re-parametrisation $(t(\underline{\theta}),\sigma)$ is given by

$$\begin{aligned} F(t(\underline{\theta}),\sigma) = A'F(\mu,\sigma)A \quad &= \begin{bmatrix} 3\sigma & 0 \\ 3t(\theta) & 1 \end{bmatrix} \begin{bmatrix} \sigma^{-2} & 0 \\ 0 & 2\sigma^{-2} \end{bmatrix} \begin{bmatrix} 3\sigma & 3t(\theta) \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 3\sigma^{-1} & 0 \\ 3t(\theta)\sigma^{-2} & 2\sigma^{-2} \end{bmatrix} \begin{bmatrix} 3\sigma & 3t(\theta) \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 9 & 9\sigma^{-1}t(\theta) \\ 9\sigma^{-1}t(\theta) & 9t^2(\theta)\sigma^{-2} + 2\sigma^{-2} \end{bmatrix}. \end{aligned}$$

Following the notations in Berger and Bernardo (1992) and Bernardo (1998), the functions $h_j\,(j = 1,2)$ which are needed to calculate the reference prior for the group ordering $(t(\theta),\sigma)$ can be obtained from $F(t(\theta),\sigma)$ as follows:

From the section above, let

$$\tilde{F} = A'F(\mu,\sigma)A = \begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix}$$

therefore

$$\underset{\sim}{F}^{-1} = \begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix}^{-1} = \begin{pmatrix} f_{11.2} & f_{12} \\ f_{21} & f_{22} \end{pmatrix}.$$

Now

$$h_1 = f_{11.2} = f_{11} - \frac{1}{f_{22}} f_{12} f_{21} = \left| 9 - \frac{9\sigma^{-1}t(\underline{\theta})9\sigma^{-1}t(\underline{\theta})}{9t^2(\underline{\theta})\sigma^{-2} + 2\sigma^{-2}} \right| = \left| \frac{81t^2(\underline{\theta})\sigma^{-2} + 18\sigma^{-2} - 81\sigma^{-2}t^2(\underline{\theta})}{9t^2(\underline{\theta})\sigma^{-2} + 2\sigma^{-2}} \right|$$

$$h_1 = \left| \frac{18\sigma^{-2}}{9t^2(\underline{\theta})\sigma^{-2} + 2\sigma^{-2}} \right| = \left| \frac{18}{9t^2(\underline{\theta}) + 2} \right| = (18)\{9t^2(\underline{\theta}) + 2\}^{-1}$$

$$h_1^{\frac{1}{2}} = (18)^{\frac{1}{2}}\{9t^2(\underline{\theta}) + 2\}^{-\frac{1}{2}}$$

and

$$h_2 = f_{22} = \left(9t^2(\underline{\theta})\sigma^{-2} + 2\sigma^{-2}\right) = \sigma^{-2}\left(9t^2(\underline{\theta}) + 2\right)$$

$$h_2^{\frac{1}{2}} = \sigma^{-1}\{9t^2(\underline{\theta}) + 2\}^{\frac{1}{2}}.$$

It follows that

$$p(t(\underline{\theta})) \propto h_1^{\frac{1}{2}} = (18)^{\frac{1}{2}}\{9t^2(\underline{\theta}) + 2\}^{-\frac{1}{2}} \propto \{1 + \frac{9}{2}t^2(\underline{\theta})\}$$

$$p(\sigma | t(\underline{\theta})) \propto h_2^{\frac{1}{2}} = \sigma^{-1}\{9t^2(\underline{\theta}) + 2\}^{\frac{1}{2}} \propto \sigma^{-1}.$$

Therefore the reference prior relative to the ordered parameterisation $(t(\underline{\theta}), \sigma)$ is given by

$$p(t(\underline{\theta}), \sigma) = p(t(\underline{\theta}))\, p(\sigma | t(\underline{\theta})$$

$p(t(\underline{\theta}), \sigma) = \{1 + \frac{9}{2}t^2(\underline{\theta})\}\sigma^{-1}$ which is the same as before.

Therefore the reference prior relative to the ordered parameterisation $(t(\underline{\theta}), \sigma)$ is given by

$$\pi^R(t(\underline{\theta}), \sigma) \propto \left\{1 + \frac{9}{2}t^2(\underline{\theta})\right\}^{-\frac{1}{2}} \sigma^{-1}.$$

In the (μ, σ) parameterisation this corresponds to

let

$x = t(\underline{\theta}) \quad y = \mu$

then

$$y = 3\sigma x + \ell_0 \quad x = \frac{y-\ell_0}{3\sigma} \quad \frac{dx}{dy} = \frac{1}{3\sigma}$$

$$\pi^R(y,\sigma) \propto \left\{1+\frac{9}{2}x^2\right\}^{-\frac{1}{2}} \sigma^{-1} \left|\frac{dx}{dy}\right|$$

$$= \frac{1}{3}\left\{1+\frac{9}{2}x^2\right\}^{-\frac{1}{2}} \sigma^{-1}\sigma^{-1}$$

$$\pi^R(\mu,\sigma) \propto \left\{1+\frac{1}{2}\frac{(\mu-\ell_0)^2}{\sigma^2}\right\}^{-\frac{1}{2}} \sigma^{-2}.$$

which proves the theorem. See also Bernardo (1998) pages 60 – 61 for a similar proof).

Proof of Theorem 4.9.1

$$\delta = \Phi\left(\frac{\mu-\ell_0}{\sigma}\right) = \int_{-\infty}^{\frac{\mu-\ell_0}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz$$

Let

$\dfrac{\mu-\ell_0}{\sigma} = \nu$, then

$$\frac{\partial\Phi(\nu)}{\partial\mu} = \frac{\partial\Phi(\nu)}{\partial\nu} \times \frac{\partial\nu}{\partial\mu}$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\nu^2}\left(\frac{1}{\sigma}\right)$$

and

$$\frac{\partial\Phi(\nu)}{\partial\sigma} = \frac{\partial\Phi(\nu)}{\partial\nu} \times \frac{\partial\nu}{\partial\sigma}$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\nu^2}\left(\frac{\ell_0-\mu}{\sigma^2}\right) .$$

Define

$$f = \frac{\partial\Phi(\nu)}{\partial\nu} = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\nu^2} ,$$

then

$$\nabla'_t(\theta) = \left[\frac{\partial\Phi(\nu)}{\partial\mu} \quad \frac{\partial\Phi(\nu)}{\partial\sigma}\right] = f\left[1 \quad \frac{\ell_0-\mu}{\sigma}\right].$$

Since

$$F^{-1}(\underline{\theta}) = F^{-1}(\mu,\sigma) = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \frac{1}{2}\sigma^2 \end{pmatrix} = \text{diag}\left[\sigma^2, \tfrac{1}{2}\sigma^2\right]$$

(See proof of Theorem 4.7.1) it follows that

$$\nabla'_t(\underline{\theta})F^{-1}(\underline{\theta}) = f\begin{bmatrix} 1 & \frac{\ell_0-\mu}{\sigma} \end{bmatrix}\begin{pmatrix} \sigma^2 & 0 \\ 0 & \frac{1}{2}\sigma^2 \end{pmatrix}$$

$$\nabla'_t(\underline{\theta})F^{-1}(\underline{\theta}) = \sigma^2 f\begin{bmatrix} 1 & \frac{\ell_0-\mu}{\sigma} \end{bmatrix}\begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} = \sigma^2 f\begin{bmatrix} 1 & \frac{\ell_0-\mu}{2\sigma} \end{bmatrix}$$

and

$$\nabla'_t(\underline{\theta})F^{-1}(\underline{\theta})\nabla_t(\underline{\theta}) = \sigma^2 f^2\left[1+\tfrac{1}{2}\left(\frac{\mu-\ell_0}{\sigma}\right)^2\right].$$

Therefore

$$\sqrt{\nabla'_t(\underline{\theta})F^{-1}(\underline{\theta})\nabla_t(\underline{\theta})} = \sigma f\left[1+\tfrac{1}{2}\left(\frac{\mu-\ell_0}{\sigma}\right)^2\right]^{\frac{1}{2}}$$

which means that

$$\eta_1(\underline{\theta}) = \frac{\sigma}{\sqrt{1+\frac{1}{2}\left(\frac{\mu-\ell_0}{\sigma}\right)^2}} \text{ and } \eta_2(\underline{\theta}) = \frac{-\frac{1}{2}(\mu-\ell_0)}{\sqrt{1+\frac{1}{2}\left(\frac{\mu-\ell_0}{\sigma}\right)^2}} \ .$$

Since $\eta_1(\underline{\theta})$ and $\eta_2(\underline{\theta})$ are exactly the same as the corresponding functions in Theorem 4.7.1, the probability matching priors for $C_{pl} = \frac{\mu-\ell_0}{3\sigma}$ and $\delta = \Phi\left(\frac{\mu-\ell_0}{\sigma}\right)$ are similar and is given by

$$\pi^M(\mu,\sigma) = \left\{1+\frac{(\ell_0-\mu)^2}{2\sigma^2}\right\}^{-\frac{1}{2}}\sigma^{-2}$$

or from corollary 4.7.1.1

$$\pi^M(\mu,\sigma^2) = \left\{1+\frac{(\ell_0-\mu)^2}{2\sigma^2}\right\}^{-\frac{1}{2}}\sigma^{-3}.$$

CHAPTER 5

A PROCESS CAPABILITY INDEX FOR AVERAGES OF OBSERVATIONS FROM NEW BATCHES IN THE CASE OF THE BALANCED RANDOM EFFECTS MODEL WITH TWO VARIANCE COMPONENTS

5.1 INTRODUCTION

Capability analysis is used in many facets of industrial processes. In this chapter, a process capability index $\left(P_{pl}^{\ 1}\right)$ is developed for the average of observations from new or unknown batches in the case of a balanced random effects model. Using a Bayesian approach theoretical and simulation results are derived for the index under two different but related prior distributions. For the conventional prior (obtained from the Jeffreys' rule), the exact posterior moments of the index are calculated. These moments are used to obtain a Pearson curve approximation to the posterior distribution for a real problem. The posterior distribution can also be obtained by using Monte Carlo simulation. A weighted Monte Carlo method is used for the probability matching (reference) prior to obtain the unconditional posterior distribution of $P_{pl}^{\ 1}$.

Data arising from multiple sources of variability is very common in practice. Virtually all industrial processes exhibit between-batch, as well as within-batch, components of variation. In some cases the between-batch (or between subgroup) component is viewed as part of the common-cause-system for the process. It seems worthwhile to develop a process capability index in more general settings. To do so, it is necessary to employ a statistical model which adequately handles multiple sources of variability. The variance component model is suitable for this task.

This chapter investigates a version of the process capability or performance index for the balanced random effects model from a Bayesian framework. The index is denoted by P_{pl}^{1} and can be used for the average of observations from new or unknown batches. The term capability index will be used as a generic term.

The next section briefly introduces and defines the lower process performance index. In section 3, a Bayesian analysis of the random effects model is considered. The conventional prior is discussed and used to derive posterior distributions of some parameters. Section 4 formally defines the lower performance index (P_{pl}^{1}) which is based on the random effects model. The exact posterior moments of the index are calculated. In section 5, the probability matching and reference priors for P_{pl}^{1}, are derived. In section 6, estimation procedures including posterior moments which are used to obtain a Pearson curve approximation to the posterior distribution of P_{pl}^{1}. Monte Carlo simulation procedures are explained for estimating the posterior distribution of P_{pl}^{1} and a related index for the conventional prior. A weighted Monte Carlo method is used for the probability matching (reference) prior to obtain the unconditional posterior distribution of P_{pl}^{1}. An illustrative example is provided in section 7.

5.2 DEFINITIONS AND NOTATIONS

The lower process performance index is defined as:

$$P_{pl}^{1} = \frac{\mu - l_0}{3\left(\frac{\sigma_1^2 + J\sigma_2^2}{J}\right)^{\frac{1}{2}}} \qquad (5.2.1)$$

where

$\mu =$ mean of future observation for a new or unknown batch

$\sigma_1^2 =$ within group variance

$\sigma_2^2 =$ between groups variance

$\sigma_3^2 =$ between groups variance

$l_0 =$ lower specification limit

$J =$ batch (package) size

P_{pl}^1 follows from the fact that the expected value of the sample mean of a new or unknown batch is μ and its variance is $\frac{\sigma_1^2}{J} + \sigma_2^2$.

Approximations of the exact posterior distribution of P_{pl}^1 can be obtained. Current knowledge indicates that a posterior analysis for this form of P_{pl}^1 does not exist. For notational convenience, the letters I and J will now be used to indicate number of groups (batch) and subgroup size respectively.

The next sections will, therefore, look at the variance component model and derive the exact posterior moments (mean, variance, third and fourth central moments) for the lower process performance index.

5.3 THE VARIANCE COMPONENT MODEL

The variance component model with two variance components is of the form:

$Y_{ij} = \mu + r_i + \varepsilon_{ij}$ for $i = 1, 2, \cdots, I$ and $j = 1, 2, \cdots, J$.

(5.3.1)

Equation (5.3.1) can be written as:

$Y_{ij} = \mu + \varepsilon_{ij}$ for $i = 1, 2, \cdots, I$ and $j = 1, 2, \cdots, J$.

(5.3.2)

where $\mu_i = \mu + r_i$

and Y_{ij} is the j^{th} response from the i^{th} batch

(5.3.3) $\varepsilon_{ij} \sim N(0, \sigma_1^2)$ independently of $r_i \sim N(0, \sigma_2^2)$

The single fixed effect μ denotes the overall mean and the random effect r_i denotes the deviation from this mean, specific to batch i. ε_{ij} represents the within group/batch variation. Y_{ij} is known and denotes the j^{th} response value in the i^{th} batch.

The model (5.3.2) explains the sampling variation or within group (lot) variation, while model (5.3.3) explains process parameter or between group variation and (5.3.1) explains both sources of variation, sampling variation and process parameter (between groups) variation.

Some important results of the of the variance component model are summarised in the theorem 5.3.1

Theorem 5.3.1

I. $Y_{ij} \mid \mu, \sigma_1^2, \sigma_2^2 \sim N(\mu, \sigma_1^2 + \sigma_2^2)$

II. $\overline{Y}_{i.} \mid \mu, \sigma_1^2, \sigma_2^2 \sim N(\mu, \frac{\sigma_1^2 + J\sigma_2^2}{J}) = N(\mu, \frac{\sigma_1^2}{J} + \sigma_2^2)$

III. $\overline{\overline{Y}}_{..} \mid \mu, \sigma_1^2, \sigma_2^2 \sim N(\mu, \frac{\sigma_1^2 + J\sigma_2^2}{IJ})$

where $\overline{Y}_{i.} = \frac{1}{J}\sum_{j=1}^{J} Y_{ij}$, the i$^{\text{th}}$ batch (group) mean and $\overline{\overline{Y}}_{..} = \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J} Y_{ij}$ is the overall sample mean

<u>Proof</u>

The proof is given in Appendix A5.

The variance component model is used when the variation of group means is more than we expect when using a simple within group variance model. Process parameters are known to vary slightly, even when the process is in statistical control, introducing extra variation σ_2^2; e.g. Kiln (brick oven) temperatures may vary process parameters (quality measures) in brick manufacture.

5.3.1 POSTERIOR DISTRIBUTION OF THE MEAN AND VARIANCE COMPONENTS

Consider model (5.3.1) again:

$Y_{ij} = \mu + r_i + \varepsilon_{ij}$ for $i = 1, 2, \cdots, I$ and $j = 1, 2, \cdots, J$.

where $\varepsilon_{ij} \sim N(0, \sigma_1^2)$ independently of $r_i \sim N(0, \sigma_2^2)$

The non-informative joint prior for the variance component model as defined by Box and Tiao (1973), page 251, viz.:

$$p(\mu, \sigma_1^2, \sigma_2^2) \propto \frac{1}{\sigma_1^2(\sigma_1^2 + J\sigma_2^2)} \qquad (5.3.4)$$

The prior may easily be obtained by applying Jeffreys' rule. Jeffreys' rule states that the prior distribution for a set of parameters is taken to be proportional to the square root of the determinant of the Fisher information matrix. By combining the prior with the likelihood, the joint posterior distribution of μ, σ_1^2 and σ_2^2 can be obtained.

Theorem 5.3.2

$$p(\mu | \underline{Y}, \sigma_1^2, \sigma_2^2) = \frac{1}{\sqrt{2\pi \frac{(\sigma_1^2 + J\sigma_2^2)}{IJ}}} \exp\left(-\frac{1}{2} \frac{IJ(\mu - \overline{\overline{Y}}_{..})^2}{(\sigma_1^2 + J\sigma_2^2)}\right) \qquad (5.3.5)$$

i.e. $\mu | \underline{Y}, \sigma_1^2, \sigma_2^2 \sim N\left(\overline{\overline{Y}}_{..}, \frac{(\sigma_1^2 + J\sigma_2^2)}{IJ}\right)$

and the joint posterior distribution of the variance components σ_1^2, σ_2^2 is

$$p\left(\sigma_1^2, \sigma_2^2 | \underline{Y}\right) \propto \left((\sigma_1^2)^{-\frac{1}{2}(\nu_1+2)} (\sigma_1^2 + J\sigma_2^2)^{-\frac{1}{2}(\nu_2+2)} \exp\left(-\frac{1}{2}\left\{\frac{\nu_1 m_1}{\sigma_1^2} + \frac{\nu_2 m_2}{(\sigma_1^2 + J\sigma_2^2)}\right\}\right) \right) \qquad (5.3.6)$$

where $\sigma_1^2 > 0 \quad \sigma_2^2 > 0 \quad \sigma_{12}^2 > \sigma_1^2$ and $\sigma_{12}^2 = \sigma_1^2 + J\sigma_2^2$

I is the number of groups/batches

J is the number of observations within each group

$\nu_1 = I(J-1) \quad ; \nu_2 = I-1$

$\nu_1 m_1 = \sum_{i=1}^{I}\sum_{j=1}^{J}(Y_{ij} - \overline{Y_{i.}})^2$ is residual sum of squares

$\nu_2 m_2 = J\sum_{i=1}^{I}(\overline{Y_{i.}} - \overline{\overline{Y}}_{..})^2$ is between groups sum of squares.

<u>Proof</u>

The proofs are given in Appendix A5.

Theorem 5.3.3

$E(m_1) = \sigma_1^2$ and $E(m_2) = (\sigma_1^2 + J\sigma_2^2)$.

<u>Proof</u>

The proofs are given in Appendix A5.

The posterior distribution for σ_1^2 and σ_{12}^2, each one proportional to an inverse gamma distribution, would be independent of each other if the restriction $\sigma_{12}^2 > \sigma_1^2$ did not apply. The joint posterior distribution for σ_1^2 and σ_{12}^2 would be the product of these two distributions.

$$p\left(\sigma_1^2, \sigma_{12}^2 \mid \underline{Y}\right) \propto \left((\sigma_1^2)^{-\frac{1}{2}(\nu_1+2)} \exp(-\frac{1}{2}\{\frac{\nu_1 m_1}{\sigma_1^2}\}) \times (\sigma_{12}^2)^{-\frac{1}{2}(\nu_2+2)} \exp(-\frac{1}{2}\{\frac{\nu_2 m_2}{(\sigma_{12}^2)}\}) \right)$$

$$\sigma_1^2, \sigma_{12}^2 \mid \underline{Y} \sim IG(\sigma_1^2 \mid \frac{\nu_1}{2}; \nu_1 m_1) \times IG(\sigma_{12}^2 \mid \frac{\nu_2}{2}; \nu_2 m_2)$$

$$IG(x \mid \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-\alpha-1} \exp(-\beta/x)$$

i.e. the inverted gamma density with positive parameters α and β.

i.e. the inverted gamma density with positive parameters α and β.

However, the above mentioned restrictions do apply. Nevertheless, using a two-step rejection sampling procedure (as discussed in section 5.6) it is straight forward to generate samples from the joint distribution.

By using the Jeffreys' prior (5.3.4) the exact posterior moments (mean, variance, third and fourth central moments) for the lower process performance index are defined and derived in section 5.4. A probability matching prior as well as the reference prior for the index is also derived.

5.4 POSTERIOR DISTRIBUTION OF THE LOWER PROCESS PERFORMANCE INDEX P^1_{pl}

The lower process performance index (P_{pl}^{1}) is defined in (5.2.1) section 5.2.

To illustrate how and when this index (5.2.1) is used consider a factory that manufactures medical tablets in very small batches. In this instance a small batch is likely to be a weekly or monthly intake of tablets for an individual patient. The interest is in whether the patient gets, on average, the required dosage of the drug from the batch in the specified time, given that each patient must get an average dosage of at least l_0. The question is, therefore, whether the process is capable of producing to this specification.

The above mentioned index is contrasted with the following index:

$$P_{pl} = \left(\frac{\mu - l_0}{3\sigma_{total}} \right) = \left(\frac{\mu - l_0}{3\left(\sigma_1^{\,2} + \sigma_2^{\,2}\right)^{\frac{1}{2}}} \right)$$

which is not dependent on J, the subgroup size. The index assesses whether the process is capable of producing each tablet (measurement) to specification as opposed to the mean of a group tablets from the batch. The following theorem can easily be proved.

Theorem 5.4.1

The lower process performance index (5.2.1) given the variance components, is normally distributed, i.e.

$$P_{pl}^{1} \mid \underline{Y}, \sigma_1^2, \sigma_2^2 \sim N\left(\frac{\overline{\overline{Y}}_{..} - l_0}{3\left(\frac{(\sigma_1^2 + J\sigma_2^2)}{J}\right)^{\frac{1}{2}}}, \frac{1}{9I} \right) \text{for all } i, j \qquad (5.4.1)$$

<u>Proof:</u>

The proof is given in Appendix A5.

5.4.1 EXACT POSTERIOR MOMENTS OF THE LOWER THE LOWER PERFORMANCE INDEX P^1_{pl}

Denote the first four posterior moments about the origin for P_{pl}^{1} (conditionally on σ_1^2, σ_2^2) by μ'_1, μ'_2, μ'_3 and μ'_4 the central moments by μ_2, μ_3 and μ_4.

Theorem 5.4.2

$$\mu'_1 = \frac{(\overline{\overline{Y}}_{..} - l_0) J^{\frac{1}{2}}}{3\left(\sigma_{12}^2\right)^{\frac{1}{2}}}, \mu'_2 = \frac{1}{9I} + \frac{(\overline{\overline{Y}}_{..} - l_0)^2 J}{9\left(\sigma_{12}^2\right)},$$

$$\mu'_3 = \frac{(\overline{\overline{Y}}_{..} - l_0) J^{\frac{1}{2}}}{27}\left(\frac{3}{I\left(\sigma_{12}^2\right)^{\frac{1}{2}}} + \frac{(\overline{\overline{Y}}_{..} - l_0)^2 J}{\left(\sigma_{12}^2\right)^{\frac{3}{2}}} \right),$$

$$\mu'_4 = \frac{1}{27}\left\{ \frac{1}{I^2} + (\overline{\overline{Y}}_{..} - l_0)^2 J \left(\frac{2}{I\sigma_{12}^2} + \frac{(\overline{\overline{Y}}_{..} - l_0)^2 J}{3\left(\sigma_{12}^2\right)^2} \right) \right\},$$

and $\sigma_{12}^2 = \sigma_1^2 + J\sigma_2^2$.

<u>Proof:</u>

The proof is given in Appendix A5.

The r^{th} posterior moment of $\frac{1}{\sigma_{12}^{\;2}}$ (in the case of Jeffreys' prior) is given in Box and Tiao (1973) as

$$E\left(\frac{1}{\sigma_{12}^{\;2}}\right)^r = \left(\frac{2}{v_2 m_2}\right)^r \frac{\Gamma\left(\frac{v_2}{2}+r\right)\Pr\left\{F_{v_2+2r;v_1} < \frac{v_2}{v_2+2r}\frac{m_2}{m_1}\right\}}{\Gamma\left(\frac{v_2}{2}\right)\Pr\left\{F_{v_2;v_1} < \frac{m_2}{m_1}\right\}}$$

where $F_{v_2;v_1}$ is an F-distribution with v_2 and v_1 degrees of freedom and $\sigma_{12}^{\;2} = \sigma_1^{\;2} + J\sigma_2^{\;2}$.

Using the results above, the unconditional moments about the origin $M_1^{'}, M_2^{'}, M_3^{'}, M_4^{'}$ and moments about the mean M_1, M_2, M_3, M_4 can be calculated. The probabilities from the F-distribution can be found using the Matlab package (see section 5.7).

5.5 THE PROBABILITY MATCHING AND REFERENCE PRIORS FOR THE LOWER PROCESS CAPABILITY INDEX P^1_{pl}

The Bayesian paradigm emerges as an attractive approach in many types of statistical problems – especially in capability index problems but the choice of an appropriate non-informative prior distribution has been controversial. Common non-informative priors in multi-parameter problems, such as Jeffreys' prior, can have features that have an unexpectedly dramatic effect on the posterior.

5.5.1 THE PROBABILITY MATCHING PRIOR FOR THE LOWER PROCESS CAPABILITY INDEX P^1_{pl}

As mentioned in section 4.7, Datta and Ghosh (1995) derived the differential equation that a prior must satisfy if the posterior probability of a one-sided credibility interval for a parametric function and its frequentist probability agree up to $0(n^{-1})$.

Theorem 5.5.1

The probability-matching priors for the P_{pl}^{1} index in the case of the balanced random effects model defined in (5.3.1) are:

i. $\pi_a(\underline{\theta}) = \pi_a(\mu, \sigma_2^{\,2}, \sigma_1^{\,2}) \propto \sigma_1^{-2}\left(\sigma_1^{\,2} + J\sigma_2^{\,2}\right)^{-\frac{3}{2}}\left(1 + \dfrac{J(\mu - l_0)^2}{2\left(\sigma_1^{\,2} + J\sigma_2^{\,2}\right)}\right)^{-\frac{1}{2}}$

ii. $\pi_b(\underline{\theta}) = \pi_b(\mu, \sigma_2^{\,2}, \sigma_1^{\,2}) \propto \sigma_1^{-2}\left(\sigma_1^{\,2} + J\sigma_2^{\,2}\right)^{-\frac{1}{2}}\left(1 + \dfrac{J(\mu - l_0)^2}{2\left(\sigma_1^{\,2} + J\sigma_2^{\,2}\right)}\right)^{\frac{1}{2}}$

iii. $\pi_c(\underline{\theta}) = \pi_c(\mu, \sigma_2^{\,2}, \sigma_1^{\,2}) \propto \left(\sigma_1^{\,2} + J\sigma_2^{\,2}\right)^{-\frac{1}{2}}\left(1 + \dfrac{J(\mu - l_0)^2}{2\left(\sigma_1^{\,2} + J\sigma_2^{\,2}\right)}\right)^{\frac{1}{2}}$.

Proof

The proof is given in Appendix A5.

$\pi_a(\mu, \sigma_2^{\,2}, \sigma_1^{\,2})$ is of the same form as the reference prior (see section 5.5.2).

Theorem 5.5.2

The probability-matching prior

$$\pi_a(\mu, \sigma_2^{\,2}, \sigma_1^{\,2}) \propto \sigma_1^{-2}\left(\sigma_1^{\,2} + J\sigma_2^{\,2}\right)^{-\frac{3}{2}}\left(1 + \frac{J(\mu - l_0)^2}{2\left(\sigma_1^{\,2} + J\sigma_2^{\,2}\right)}\right)^{-\frac{1}{2}}$$

leads to a proper posterior.

Proof

The proof is given in Appendix A5.

5.5.2 THE REFERENCE PRIOR FOR THE LOWER PROCESS PERFORMANCE INDEX P^1_{pl}

As mentioned in section 4.7, the Jeffreys' and probability matching priors are but two methods to obtain useful information regarding non-informative priors. The Jeffreys' prior is not always suitable for multi-parameter problems. In recognition of this problem Berger and Bernado (1992) proposes the reference prior approach to the development of non-informative priors, the key feature of which is a possible dependence of the reference prior on specification of parameters of interest and nuisance parameters. In this section the reference prior of Berger and Bernado (1992) is derived for the process capability index (P_{pl}^{1}).

As is the case of the Jeffreys' prior, the reference prior method is derived from the Fisher information matrix. Note that the reference priors depend on the group ordering of the parameters. Berger and Bernado (1992) suggests that multiple groups, ordered in terms of inferential importance, are allowed, with the reference prior being determined through a succession of analyses for the implied conditional problems. They particularly recommend the reference prior based on having each parameter in its own group, i.e., having each conditional reference prior be one dimensional. Notations such as $\left\{\mu, \sigma_2^{\,2}, \sigma_1^{\,2}\right\}$ will be used to specify the groups and importance of parameters; $\left\{\mu, \sigma_2^{\,2}, \sigma_1^{\,2}\right\}$ means that there are three groups, with μ being the most important and σ_1^{2} the least important.

As in section 4.7, there will also be an examination to see whether the reference priors satisfy the probability-matching criterion (4.7.1).

The following theorem can now be stated.

Theorem 5.5.3

For the lower process performance index, P_{pl}^{1}, the reference prior relative to the ordered parameterisation $\left\{\mu, \sigma_2^{\,2}, \sigma_1^{\,2}\right\}$ is given by:

$$p_R(\mu, \sigma_2^{\,2}, \sigma_1^{\,2}) \propto \sigma_1^{-2}\left(\sigma_1^{\,2} + J\sigma_2^{\,2}\right)^{-\frac{3}{2}}\left(1 + \frac{J(\mu - l_0)^2}{2\left(\sigma_1^{\,2} + J\sigma_2^{\,2}\right)}\right)^{-\frac{1}{2}}.$$

which is the same as theorem 5.5.1(i).

<u>Proof</u>

The proof is given in Appendix A5.

Corollary 5.5.3.1

The reference prior relative to the ordered parameterisation $\{\mu, \sigma_1{}^2, \sigma_2{}^2\}$ is the same as the reference prior relative to the ordered parameterisation $\{\mu, \sigma_2{}^2, \sigma_1{}^2\}$.

Proof

The proof is given in Appendix A5.

5.6 PROCEDURES FOR ESTIMATING THE PROCESS PERFORMANCE INDEX P^1_{pl}

5.6.1 PEARSON CURVE APPROXIMATION

As mentioned in the case of Jeffreys' prior, $p(\mu, \sigma_2{}^2, \sigma_1{}^2) \propto \sigma_1^{-2}\left(\sigma_1{}^2 + J\sigma_2{}^2\right)^{-1}$ the exact posterior moments can be calculated. A Pearson curve approximation of the posterior distribution of $P_{pl}{}^1$ can then be drawn. A type I Pearson curve can be used to approximate the posterior distribution of $P_{pl}{}^1$ for our data set. Monte Carlo simulation methods can also be used to graph the index.

5.6.2 MONTE CARLO SIMULATION

In the case of Jeffreys' prior, simulation of the posterior of σ_1^2 and $\sigma_2{}^2$ can be achieved through the following standard Monte Carlo simulation routines. By using the Matlab package, simulation of σ_1^2 and $\sigma_2{}^2$ can be obtained in the following way:

i. Draw τ from a $\chi^2_{v_1}$ distribution. $\tau = \sum_{i=1}^{v_1} Z_i^{\,2}$ where $Z_i \sim N(0,1)$ $(i = 1, 2, \cdots, v_1)$ and independently of each other.

ii. Let $\dfrac{1}{\tau} = \dfrac{\sigma_1^2}{v_1 m_1}$

iii. $\sigma_1^{2*} = \frac{v_1 m_1}{\tau}$ where the * indicates a simulated value.

iv. Draw u from a $\chi^2_{v_2}$ distribution. $u = \sum_{i=1}^{v_2} Z_i^2$

v. Let $\frac{1}{u} = \frac{\sigma_{12}{}^2}{v_2 m_2}$ *where* $\sigma_{12}{}^2 = \sigma_1^2 + J\sigma_2^2$

vi. $\sigma_{12}{}^{2*} = \frac{v_2 m_2}{u}$

vii. If $\sigma_{12}{}^{2*} > \sigma_1^{2*}$ (σ_1^{2*} was simulated in step (iii))

calculate $\sigma_2^{2*} = \frac{1}{J}(\sigma_{12}{}^{2*} - \sigma_1^{2*})$ from the expression $\sigma_{12}{}^2 = \sigma_1^2 + J\sigma_2^2$.

where the following expressions still hold

$v_1 m_1 = \sum_{i=1}^{I}\sum_{j=1}^{J}(Y_{ij} - \overline{Y_{i.}})^2$ is the residual sum of squares,

$v_2 m_2 = J\sum_{i=1}^{I}(\overline{Y_{i.}} - \overline{\overline{Y}}_{..})^2$ is between groups sums of squares.

It must be mentioned that even though σ_1^{2*} and $\sigma_{12}{}^{2*}$ will always be positive, the value for σ_2^{2*} can be negative. If a negative value for σ_2^{2*} is obtained, disregard this value as well as the corresponding value for σ_1^{2*} and start again.

viii. The simulation of μ involves substituting each of the values σ_1^{2*} and $\sigma_{12}{}^{2*}$ into the normal distribution, $\mu \mid \underline{Y}, \sigma_1^2, \sigma_2^2 \sim N(\overline{\overline{Y}}_{..}, \frac{(\sigma_1^{2*} + J\sigma_2^{2*})}{IJ})$ and then drawing a value for μ. The resulting set of values $\left(\sigma_2^{2*}, \sigma_1^{2*}, \mu^*\right)$ is then substituted into P_{pl}^1 to simulate

$$P_{pl}^{1*} = \frac{\mu^* - l_0}{3\left(\frac{\sigma_1^{2*} + J\sigma_2^{2*}}{J}\right)^{\frac{1}{2}}}.$$

Repeat steps i-viii until $\tilde{\ell}$ permissible values are obtained. For our example, $\tilde{\ell}$ was taken as 10 000. The resulting set of values P_{pl}^{1*} is plotted in a histogram, and the fitted Pearson curve represents the unconditional posterior distribution $p(P_{pl}^{1} \mid \underline{Y})$.

A second method, the Rao-Blackwell method, can also be used to obtain the unconditional posterior distribution $p(P_{pl}^{1} \mid \underline{Y})$.

I. Simulate a pair of variance components $\left(\sigma_1^{2*}, \sigma_2^{2*}\right)$ as discussed above and substitute them in the conditional posterior distribution $p(P_{pl}^{1} \mid \underline{Y}, \sigma_1^{2*}, \sigma_2^{2*})$ which is normally distributed with mean

$$\mathrm{E}(P_{pl}^{1} \mid \underline{Y}, \sigma_1^{2*}, \sigma_2^{2*}) = \frac{\overline{\overline{Y}} - l_0}{3\left(\dfrac{\sigma_1^{2*} + J\sigma_2^{2*}}{J}\right)^{\frac{1}{2}}}$$

and variance

$$\mathrm{Var}(P_{pl}^{1} \mid \underline{Y}, \sigma_1^{2*}, \sigma_2^{2*}) = \frac{1}{9I}$$

II. Draw $p(P_{pl}^{1} \mid \underline{Y}, \sigma_1^{2*}, \sigma_2^{2*})$

III. Repeat I and II $\tilde{\ell}$ times, Using a Rao–Blackwell argument (see Gelfand and Smith 1991), a density estimate of the unconditional posterior distribution of P_{pl}^{1} can be obtained by averaging $p(P_{pl}^{1} \mid \underline{Y}, \sigma_1^{2}, \sigma_2^{2})$ over the $\tilde{\ell}$ repetitions.

Since the two methods are for all practical purposes identical, only the first method is illustrated in section 5.7.

5.6.3 THE WEIGHTED MONTE CARLO METHOD-SAMPLING-IMPORTANCE RE-SAMPLING

This section describes how to apply a weighted Monte Carlo (WMC) method to simulate P_{pl}^{1} using the probability matching prior. This method is especially suitable for computing credibility intervals.

Let

$$\pi_a(\underline{\theta}) \propto \sigma_1^{-2}\left(\sigma_1^2 + J\sigma_2^2\right)^{-\frac{3}{2}}\left(1+ \quad \frac{J(\mu - l_0)^2}{2\left(\sigma_1^2 + J\sigma_2^2\right)}\right)^{-\frac{1}{2}} \tag{5.6.1}$$

and

$$q(\underline{\theta}) \propto \sigma_1^{-2}\left(\sigma_1^2 + J\sigma_2^2\right)^{-1}. \tag{5.6.2}$$

According to Smith and Gelfand (1992), Guttman and Menzefricke (2003) and Skare *et al*. (2003), the $\tilde{\ell}$ independent draws of $\underline{\theta}^{*(\ell)} = (\mu^{*(\ell)}, \sigma_1^{2*(\ell)}, \sigma_2^{2*(\ell)})$, as discussed in section 5.6.2, for $\ell = 1$ to $\tilde{\ell}$; are a weighted sample from the posterior distribution based on π_a, where the weights are

$$w_\ell = \frac{\pi(\underline{\theta}^{*(\ell)}) / q(\underline{\theta}^{*(\ell)})}{\sum_{\ell=1}^{\tilde{\ell}} \pi(\underline{\theta}^{*(\ell)}) / q(\underline{\theta}^{*(\ell)})}$$

$\pi_a(\underline{\theta}^{*(\ell)})$ and $q(\underline{\theta}^{*(\ell)})$ denote the realistic prior density and implied prior density defined in (5.6.1) and (5.6.2). For the algorithm to be efficient, it is important that q is a good approximation to π_a. This means that q should not have too light tails when compared to π_a. For further details see Skare *et al* (2003) and Li (2007).

To simulate using the results from the probability matching prior we associate with each lower performance index value $P_{pl}^{1*(\ell)}$

$$w_\ell = \frac{\left(\sigma_1^{2*(\ell)} + J\sigma_2^{2*(\ell)}\right)^{-\frac{1}{2}}\left\{1+\frac{J(\mu^{*(\ell)} - l_0)^2}{2\left(\sigma_1^{2*(\ell)} + J\sigma_2^{2*(\ell)}\right)}\right\}^{-\frac{1}{2}}}{\sum_{\ell=1}^{\tilde{\ell}}\left(\sigma_1^{2*(\ell)} + J\sigma_2^{2*(\ell)}\right)^{-\frac{1}{2}}\left\{1+\frac{J(\mu^{*(\ell)} - l_0)^2}{2\left(\sigma_1^{2*(\ell)} + J\sigma_2^{2*(\ell)}\right)}\right\}^{-\frac{1}{2}}}$$

a. Sort the $P_{pl}^{1*(\ell)}$ values calculated in section 5.6.2 in ascending order so that

$$P_{pl}^{1*(1)} \leq P_{pl}^{1*(2)} \leq \cdots \leq P_{pl}^{1*(10000)}$$

b. Compute the weighted function w_ℓ associated with the ℓth ordered $P_{pl}^{1*(\ell)}$.

c. Add the weights from left to right (from the first on) until you get $\sum_{\ell=1}^{k_1} w_\ell = 0.025$. Write down the corresponding ordered value $P_{pl}^{1*(k_1)}$ and denote it as $P_{pl}^{1*}{}_{(0.025)}$. Add the weights from right to left (from the last back) until you get $\sum_{\ell=k_2}^{\tilde{\ell}} w_\ell = 0.025$. Write down the corresponding ordered value $P_{pl}^{1*(k_2)}$ and denote it as $P_{pl}^{1*}{}_{(0.975)}$. The 95% interval is $P_{pl}^{1*}{}_{(0.025)} - P_{pl}^{1*}{}_{(0.975)}$.

Simulation results for the variance component process performance index P_{pl}^{1} as discussed in this chapter will now be compared with the following index simulated as

$$P_{pl}^{*} = \left(\frac{\mu^{*} - LSL}{3\sigma*_{total}} \right) = \frac{\mu^{*} - l_0}{3\left(\sigma_1^{2*} + \sigma_2^{2*}\right)^{\frac{1}{2}}}$$

which is independent of J, the subgroup size.

Simulation results using the probability matching (reference) prior will be considered.

5.7 AN APPLICATION

The data in Table 5.1 represents the amount of a certain drug in a tablet. Packages (5 in all) with tablets are sampled to determine whether the patient gets on average the required dosage of the drug from the batches in a specified time, given that each patient must get an average dosage of at least $l_0 = 350$.

Table 5.1: Amount of a certain drug per tablet

Batch	Measurements				
1	379	357	390	376	376
2	363	367	382	381	359
3	401	402	407	402	396
4	402	387	392	395	394
5	415	405	396	390	395

The lower specification limit is l_0=350. The above limit is selected solely for illustrative purposes. In practice, fixed in advance limits are often determined from medical or regulatory considerations. See also Wolfinger (1998).

In table 5.2 posterior moments of σ_{12}^{-2} are given for the data in table 5.1 and where $\sigma_{12}^{2} = \sigma_1^{2} + J\sigma_2^{2}$.

Table 5 2: Expected values of $\frac{1}{\sigma_{12}^{2}}$ *, moments about the origin*

r	$\frac{1}{2}$	1	$\frac{3}{2}$	2
$E\left(\frac{1}{\sigma_{12}^{2}}\right)^r$	0.0291352208	0.0009606939	0.0000349856	0.0000013842

Table 5.3: Moments about the origin and about the mean

r	1	2	3	4
M'_r	0.8330	0.8076	0.8732	1.0360
M_r		0.1136	0.0111	0.0392

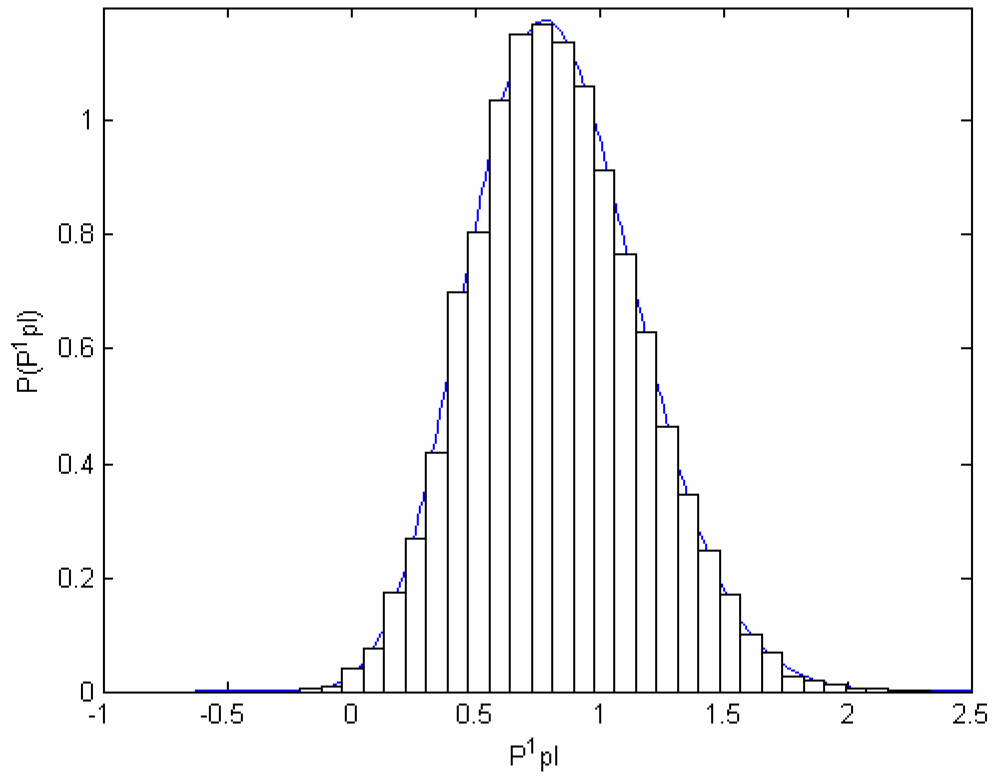


Figure 5.1: Process performance index Pearson's type 1 curve and histogram of simulated P_{pl}^{1} *values*

A density estimate of the unconditional posterior distribution of P_{pl}^{1} using Pearson's type I curve for the non-informative Jeffreys' prior is given in Figure 5.1. Further discussions on the Pearson's Type I curve are found in Van der Merwe and Chikobvu (2004). The simulated P_{pl}^{1} values for the same prior are plotted as a histogram in the same figure.

Figure 5.2 shows a histogram of the simulated P_{pl} values.

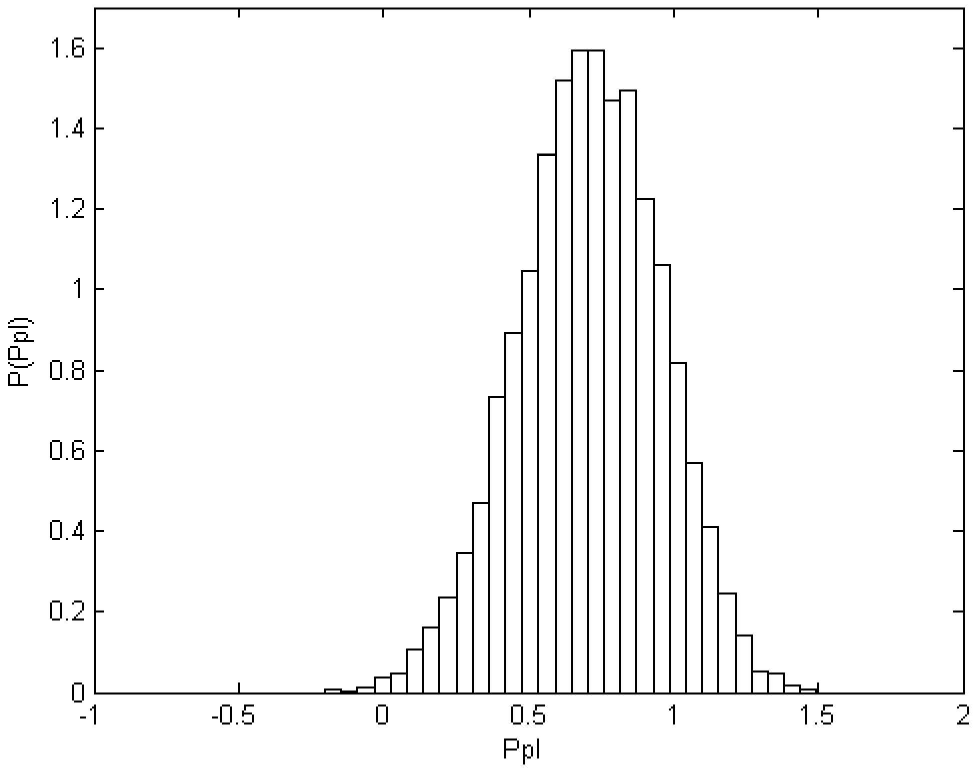


Figure 5.2: Histogram of simulated P_{pl} values

Table 5.4: Posterior mean, variance and 95% interval for P_{pl} and P_{pl}^{1} using the non-informative Jeffery's prior

Index	Mean	Variance	95% Interval
P_{pl}^{1}	0.8341	0.1139	(0.2161; 1.5396)
P_{pl}	0.7107	0.0596	(0.2082; 1.1653)

The corresponding 95% interval in the case of the probability-matching prior for P_{pl}^{1} is (0.2129; 1.5270).

Some features of these results worth noting are:

- The indices are all less than one suggesting that the process is not capable.
- P_{pl}^{1} gives mean values that are higher when compared to P_{pl}
- The variance of P_{pl}^{1} is larger than the variance P_{pl}

Appendix A5

Proof of theorem 5.3.1

$$I.\quad E[Y_{ij}\,|\,\mu,\sigma_1^2,\sigma_2^2] = E[\mu + r_i + \varepsilon_{ij}\,|\,\mu,\sigma_1^2,\sigma_2^2]$$

$$= E[\mu\,|\,\mu,\sigma_1^2,\sigma_2^2] + E[r_i\,|\,\mu,\sigma_1^2,\sigma_2^2] + E[\varepsilon_{ij}\,|\,\mu,\sigma_1^2,\sigma_2^2]$$

$$= \mu + 0 + 0$$

$$= \mu\,.$$

$$Var[Y_{ij}\,|\,\mu,\sigma_1^2,\sigma_2^2] = Var[\mu + r_i + \varepsilon_{ij}\,|\,\mu,\sigma_1^2,\sigma_2^2]$$

$$= Var[r_i\,|\,\mu,\sigma_1^2,\sigma_2^2] + Var[\varepsilon_{ij}\,|\,\mu,\sigma_1^2,\sigma_2^2]$$

$$= \sigma_2^2 + \sigma_1^2$$

$$hence\quad Y_{ij}\,|\,\mu,\sigma_1^2,\sigma_2^2 \sim N(\mu,\sigma_1^2+\sigma_2^2)\,.$$

$$II.\quad If\ \ we\ define\quad \overline{Y_{i.}} = \frac{\sum_{i=1}^{J} Y_{ij}}{J}$$

$$E[\overline{Y_{i.}}\,|\,\mu,\sigma_1^2,\sigma_2^2] = E[\frac{\sum_{j=1}^{J} Y_{ij}}{J}\,|\,\mu,\sigma_1^2,\sigma_2^2]$$

$$= E[\frac{\sum_{j=1}^{J}(\mu + r_i + \varepsilon_{ij})}{J}\,|\,\mu,\sigma_1^2,\sigma_2^2]$$

$$= E[\frac{J\mu}{J}\,|\,\mu,\sigma_1^2,\sigma_2^2] + E[\frac{Jr_i}{J}\,|\,\mu,\sigma_1^2,\sigma_2^2] + E[\frac{\sum_{j=1}^{J}\varepsilon_{ij}}{J}\,|\,\mu,\sigma_1^2,\sigma_2^2]$$

$$= E[\mu\,|\,\mu,\sigma_1^2,\sigma_2^2] + E[r_i\,|\,\mu,\sigma_1^2,\sigma_2^2] + [\frac{\sum_{j=1}^{J} E(\varepsilon_{ij}\,|\,\mu,\sigma_1^2,\sigma_2^2)}{J}]$$

$$= \mu\,.$$

$$Var[\overline{Y_{i.}}\,|\,\mu,\sigma_1^2,\sigma_2^2] = Var[\frac{\sum_{j=1}^{J} Y_{ij}}{J}\,|\,\mu,\sigma_1^2,\sigma_2^2]$$

$$= Var[\frac{\sum_{j=1}^{J}(\mu + r_i + \varepsilon_{ij})}{J} \mid \mu, \sigma_1^2, \sigma_2^2]$$

$$= Var[\frac{J\mu}{J} \mid \mu, \sigma_1^2, \sigma_2^2] + Var[\frac{Jr_i}{J} \mid \mu, \sigma_1^2, \sigma_2^2] + Var[\frac{\sum_{j=1}^{J}\varepsilon_{ij}}{J} \mid \mu, \sigma_1^2, \sigma_2^2]$$

$$= Var[\mu \mid \mu, \sigma_1^2, \sigma_2^2] + Var[r_i \mid \mu, \sigma_1^2, \sigma_2^2] + [\frac{\sum_{j=1}^{J} Var(\varepsilon_{ij} \mid \mu, \sigma_1^2, \sigma_2^2)}{J^2}]$$

$$= 0 + \sigma_2^2 + \frac{J\sigma_1^2}{J^2}$$

$$= \sigma_2^2 + \frac{\sigma_1^2}{J}.$$

III. $E[\overline{\overline{Y}}_{..} \mid \mu, \sigma_1^2, \sigma_2^2] = E[\frac{\sum_{i=1}^{I}\sum_{j=1}^{J} Y_{ij}}{IJ} \mid \mu, \sigma_1^2, \sigma_2^2]$

$$= E[\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}(\mu + r_i + \varepsilon_{ij})}{IJ} \mid \mu, \sigma_1^2, \sigma_2^2]$$

$$= E[\frac{IJ\mu}{IJ} \mid \mu, \sigma_1^2, \sigma_2^2] + E[\frac{J\sum_{i=1}^{I} r_i}{IJ} \mid \mu, \sigma_1^2, \sigma_2^2] + E[\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}\varepsilon_{ij}}{IJ} \mid \mu, \sigma_1^2, \sigma_2^2]$$

$$= E[\mu \mid \mu, \sigma_1^2, \sigma_2^2] + [\frac{\sum_{i=1}^{I} E(r_i \mid \mu, \sigma_1^2, \sigma_2^2)}{I}] + [\frac{\sum_{i=1}^{I}\sum_{j=1}^{J} E(\varepsilon_{ij} \mid \mu, \sigma_1^2, \sigma_2^2)}{IJ}]$$

$$= \mu + 0 + 0$$

$$= \mu.$$

$$Var[\overline{\overline{Y}}_{..} \mid \mu,\sigma_1^2,\sigma_2^2] = Var[\frac{\sum_{i=1}^{I}\sum_{j=1}^{J} Y_{ij}}{IJ} \mid \mu,\sigma_1^2,\sigma_2^2]$$

$$= Var[\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}(\mu + r_i + \varepsilon_{ij})}{IJ} \mid \mu,\sigma_1^2,\sigma_2^2]$$

$$= \text{var}[\frac{IJ\mu}{IJ} \mid \mu,\sigma_1^2,\sigma_2^2] + Var[\frac{J\sum_{i=1}^{I} r_i}{IJ} \mid \mu,\sigma_1^2,\sigma_2^2] + Var[\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}\varepsilon_{ij}}{IJ} \mid \mu,\sigma_1^2,\sigma_2^2]$$

$$= Var[\mu \mid \mu,\sigma_1^2,\sigma_2^2] + [\frac{\sum_{i=1}^{I} Var(r_i \mid \mu,\sigma_1^2,\sigma_2^2)}{I^2}] + [\frac{\sum_{i=1}^{I}\sum_{j=1}^{J} Var(\varepsilon_{ij} \mid \mu,\sigma_1^2,\sigma_2^2)}{I^2 J^2}]$$

$$= 0 + \frac{I\sigma_2^2}{I^2} + \frac{IJ\sigma_1^2}{I^2 J^2}$$

$$= \frac{\sigma_2^2}{I} + \frac{\sigma_1^2}{IJ}$$

$$= \frac{\sigma_1^2 + J\sigma_2^2}{IJ}.$$

Proof of theorem 5.3.2

Consider model (5.3.1) again:

$Y_{ij} = \mu + r_i + \varepsilon_{ij}$ for $i = 1, 2, \cdots, I$ and $j = 1, 2, \cdots, J$.

where $\varepsilon_{ij} \sim N(0,\sigma_1^2)$ independently of $r_i \sim N(0,\sigma_2^2)$

The non-informative joint prior for the variance component model as defined by Box and Tiao (1973), page 251, viz.:

$$p(\mu,\sigma_1^2,\sigma_2^2) \propto \frac{1}{\sigma_1^2(\sigma_1^2 + J\sigma_2^2)}.$$

The posterior joint distribution of μ and σ_1^2, σ_2^2 can be worked out, but first observe the result below.

$$\sum_{i=1}^{I}\sum_{j=1}^{J}(Y_{ij} - \mu)^2 = \sum_{i=1}^{I}\sum_{j=1}^{J}(Y_{ij} - \overline{Y}_{i.})^2 + J\sum_{i=1}^{I}(\overline{Y}_{i.} - \overline{\overline{Y}}_{..})^2 + IJ(\overline{\overline{Y}}_{..} - \mu)^2 \qquad (A5.3.1)$$

Therefore

$$\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}(Y_{ij}-\mu)^2}{(\sigma_1^2+\sigma_2^2)}=\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}(Y_{ij}-\overline{Y_{i.}})^2}{\sigma_1^2}+\frac{J\sum_{i=1}^{I}(\overline{Y_{i.}}-\overline{\overline{Y_{..}}})^2}{(\sigma_1^2+J\sigma_2^2)}+\frac{IJ(\overline{\overline{Y_{..}}}-\mu)^2}{(\sigma_1^2+J\sigma_2^2)} \qquad (A5.3.2)$$

Let

$\underline{Y}=[Y_{11,}Y_{12,}\cdots Y_{1J,}Y_{21,}\cdots Y_{2J,}\cdots\cdots Y_{I1,}\cdots Y_{IJ}]$

Posterior $\propto$ *likelihood x prior*

$$p\left(\mu,\sigma_1^2,\sigma_2^2\,|\,\underline{Y}\right)\propto\prod_{i=1}^{I}\prod_{j=1}^{J}f(Y_{ij}\,|\,\mu,\sigma_1^2,\sigma_2^2)\,p\left(\mu,\sigma_1^2,\sigma_2^2\right)$$

$$=\left[(\sigma_1^2+\sigma_2^2)^{\frac{-IJ}{2}}\exp\left(-\frac{1}{2}\left\{\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}(Y_{ij}-\mu)^2}{\sigma_1^2+\sigma_2^2}\right\}\right)\right]\frac{1}{\sigma_1^2(\sigma_1^2+J\sigma_2^2)}$$

Using result (A5.3.2), the expression in square brackets can be split into three

$$\propto\left[(\sigma_1^2)^{-\frac{1}{2}I(J-1)+1}(\sigma_1^2+J\sigma_2^2)^{-\frac{1}{2}(I)+1}\exp\left(-\frac{1}{2}\left\{\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}(Y_{ij}-\overline{Y_{i.}})^2}{\sigma_1^2}+\frac{J\sum_{i=1}^{I}(\overline{Y_{i.}}-\overline{\overline{Y_{..}}})^2}{(\sigma_1^2+J\sigma_2^2)}+\frac{IJ(\overline{\overline{Y_{..}}}-\mu)^2}{(\sigma_1^2+J\sigma_2^2)}\right\}\right)\right]$$

$$=\left[(\sigma_1^2)^{-(\frac{1}{2}I(J-1)+1)}(\sigma_1^2+J\sigma_2^2)^{-(\frac{1}{2}(I-1)+\frac{1}{2}+1)}\exp\left(-\frac{1}{2}\left\{\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}(Y_{ij}-\overline{Y_{i.}})^2}{\sigma_1^2}+\frac{J\sum_{i=1}^{I}(\overline{Y_{i.}}-\overline{\overline{Y_{..}}})^2}{(\sigma_1^2+J\sigma_2^2)}+\frac{IJ(\overline{\overline{Y_{..}}}-\mu)^2}{(\sigma_1^2+J\sigma_2^2)}\right\}\right)\right]$$

$$=\left[(\sigma_1^2)^{-\frac{1}{2}(\nu_1+2)}(\sigma_1^2+J\sigma_2^2)^{-\frac{1}{2}(\nu_2+3)}\exp\left(-\frac{1}{2}\left\{\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}(Y_{ij}-\overline{Y_{i.}})^2}{\sigma_1^2}+\frac{J\sum_{i=1}^{I}(\overline{Y_{i.}}-\overline{\overline{Y_{..}}})^2}{(\sigma_1^2+J\sigma_2^2)}+\frac{IJ(\overline{\overline{Y_{..}}}-\mu)^2}{(\sigma_1^2+J\sigma_2^2)}\right\}\right)\right]$$

where $\nu_1=I(J-1)\quad\nu_2=I-1$

$$=\left[(\sigma_1^2+J\sigma_2^2)^{\frac{-1}{2}}\exp\left(-\frac{1}{2}\frac{IJ(\overline{\overline{Y_{..}}}-\mu)^2}{(\sigma_1^2+J\sigma_2^2)}\right)\right]$$

$$\left((\sigma_1^2)^{-\frac{1}{2}(\nu_1+2)}(\sigma_1^2+J\sigma_2^2)^{-\frac{1}{2}(\nu_2+2)}\exp\left(-\frac{1}{2}\left\{\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}(Y_{ij}-\overline{Y_{i.}})^2}{\sigma_1^2}+\frac{J\sum_{i=1}^{I}(\overline{Y_{i.}}-\overline{\overline{Y_{..}}})^2}{(\sigma_1^2+J\sigma_2^2)}\right\}\right)\right).$$

The function inside the square brackets just above, when regarded as a function of μ, is proportional to the conditional distribution of a normal distribution for which the mean is $\overline{\overline{Y}}_{..}$ and variance $\frac{(\sigma_1^2 + J\sigma_2^2)}{IJ}$.

$$p(\mu \mid \underline{Y}, \sigma_1^2, \sigma_2^2) \propto \left[(\sigma_1^2 + J\sigma_2^2)^{\frac{-1}{2}} \exp\left(-\frac{1}{2}\frac{IJ(\overline{\overline{Y}}_{..} - \mu)^2}{(\sigma_1^2 + J\sigma_2^2)}\right) \right]$$

$$i.e \quad \mu \mid \underline{Y}, \sigma_1^2, \sigma_2^2 \sim N\left(\overline{\overline{Y..}}, \frac{(\sigma_1^2 + J\sigma_2^2)}{IJ}\right)$$

The function inside the curled brackets on the right hand side is therefore proportional to the marginal joint distribution of the variance components σ_1^2, σ_2^2.

$$\left((\sigma_1^2)^{-\frac{1}{2}(\nu_1+2)} (\sigma_1^2 + J\sigma_2^2)^{-\frac{1}{2}(\nu_2+2)} \exp\left(-\frac{1}{2}\left\{\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}(Y_{ij} - \overline{Y}_{i.})^2}{\sigma_1^2} + \frac{J\sum_{i=1}^{I}(\overline{Y}_{i.} - \overline{\overline{Y}}_{..})^2}{(\sigma_1^2 + J\sigma_2^2)}\right\}\right) \right)$$

$$p\left(\sigma_1^2, \sigma_2^2 \mid \underline{Y}\right) = \left((\sigma_1^2)^{-\frac{1}{2}(\nu_1+2)} (\sigma_1^2 + J\sigma_2^2)^{-\frac{1}{2}(\nu_2+2)} \exp\left(-\frac{1}{2}\left\{\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}(Y_{ij} - \overline{Y}_{i.})^2}{\sigma_1^2} + \frac{J\sum_{i=1}^{I}(\overline{Y}_{i.} - \overline{\overline{Y}}_{..})^2}{(\sigma_1^2 + J\sigma_2^2)}\right\}\right) \right)$$

$$p\left(\sigma_1^2, \sigma_2^2 \mid \underline{Y}\right) = \left((\sigma_1^2)^{-\frac{1}{2}(\nu_1+2)} (\sigma_1^2 + J\sigma_2^2)^{-\frac{1}{2}(\nu_2+2)} \exp\left(-\frac{1}{2}\left\{\frac{\nu_1 m_1}{\sigma_1^2} + \frac{\nu_2 m_2}{(\sigma_1^2 + J\sigma_2^2)}\right\}\right) \right)$$

$$p\left(\sigma_1^2, \sigma_2^2 \mid \underline{Y}\right) = \left((\sigma_1^2)^{-\frac{1}{2}(\nu_1+2)} \exp\left(-\frac{1}{2}\left\{\frac{\nu_1 m_1}{\sigma_1^2}\right\}\right) \times (\sigma_1^2 + J\sigma_2^2)^{-\frac{1}{2}(\nu_2+2)} \exp\left(-\frac{1}{2}\left\{\frac{\nu_2 m_2}{(\sigma_1^2 + J\sigma_2^2)}\right\}\right) \right)$$

$$p\left(\sigma_1^2, \sigma_2^2 \mid \underline{Y}\right) = \left((\sigma_1^2)^{-\frac{1}{2}\nu_1 - 1} \exp\left(-\frac{1}{2}\left\{\frac{\nu_1 m_1}{\sigma_1^2}\right\}\right) \times (\sigma_1^2 + J\sigma_2^2)^{-\frac{1}{2}\nu_2 - 1} \exp\left(-\frac{1}{2}\left\{\frac{\nu_2 m_2}{(\sigma_1^2 + J\sigma_2^2)}\right\}\right) \right)$$

$$p\left(\sigma_1^2, \sigma_2^2 \mid \underline{Y}\right) = IG\left(\sigma_1^2 \mid \frac{\nu_1}{2}; \nu_1 m_1\right) \times IG\left(\sigma_1^2 + J\sigma_2^2 \mid \frac{\nu_2}{2}; \nu_2 m_2\right).$$

Proof of theorem 5.3.3

We now prove that $E(m_1) = \sigma_1^2$ and $E(m_2) = (\sigma_1^2 + J\sigma_2^2)$.

First

$$E\left(m_1\right)=\sigma_1^{\,2}$$

$$\nu_1 m_1=\sum_{i=1}^{I}\sum_{j=1}^{J}(Y_{ij}-\overline{Y_{i.}})^2$$

$$E\left(\nu_1 m_1\right)=\sum_{i=1}^{I}\sum_{j=1}^{J}E(Y_{ij}-\overline{Y_{i.}})^2$$

$$=\sum_{i=1}^{I}\sum_{j=1}^{J}E\Big((Y_{ij}-\mu)-(\overline{Y_{i.}}-\mu)\ \Big)^2$$

$$=\sum_{i=1}^{I}\sum_{j=1}^{J}E\Big((Y_{ij}-\mu)^2+(\overline{Y_{i.}}-\mu)^2-2(Y_{ij}-\mu)(\overline{Y_{i.}}-\mu)\ \Big)$$

$$=\sum_{i=1}^{I}\sum_{j=1}^{J}\Big(E(Y_{ij}-\mu)^2+E(\overline{Y_{i.}}-\mu)^2-2E(Y_{ij}-\mu)(\overline{Y_{i.}}-\mu)\ \Big)$$

$$=\sum_{i=1}^{I}\sum_{j=1}^{J}\Big(Var(Y_{ij}\mid\mu,r_i,\sigma_1^{\,2})+Var(\overline{Y_{i.}}\mid\mu,r_i,\sigma_1^{\,2})-2Cov(Y_{ij}\mid\mu,r_i,\sigma_1^{\,2})(\overline{Y_{i.}}\mid\mu,r_i,\sigma_1^{\,2})\ \Big)$$

$$=\sum_{i=1}^{I}\sum_{j=1}^{J}\left(\sigma_1^{\,2}+\frac{\sigma_1^{\,2}}{J}-2\frac{\sigma_1^{\,2}}{J}\right)$$

$$=\sum_{i=}^{I}\left(J\sigma_1^{\,2}+\frac{J\sigma_1^{\,2}}{J}-2\frac{J\sigma_1^{\,2}}{J}\right)$$

$$=\left(IJ\sigma_1^{\,2}+\frac{IJ\sigma_1^{\,2}}{J}-2\frac{IJ\sigma_1^{\,2}}{J}\right)$$

$$=\left(IJ\sigma_1^{\,2}+I\sigma_1^{\,2}-2I\sigma_1^{\,2}\ \right)$$

$$=\left(IJ\sigma_1^{\,2}-I\sigma_1^{\,2}\ \ \right)$$

$$=(\sigma_1^{\,2}I(J-1))$$

$$=\sigma_1^{\,2}\nu_1\ \text{ since }\ \nu_1=I(J-1)$$

$$E\left(m_1\right)=\frac{\sigma_1^{\,2}\nu_1}{\nu_1}=\sigma_1^{\,2}.$$

$$E\left(m_1\right)=\frac{\sigma_1^{\,2}\nu_1}{\nu_1}=\sigma_1^{\,2}.$$

We now prove that

$$E\left(m_2\right)=(\sigma_1^{\,2}+J\sigma_2^{\,2})$$

$$\nu_2 m_2 = J\sum_{i=1}^{I}(\overline{Y_{i.}} - \overline{\overline{Y_{..}}})^2$$

$$E(\nu_2 m_2) = J\sum_{i=1}^{I}E(\overline{Y_{i.}} - \overline{\overline{Y_{..}}})^2$$

$$= J\sum_{i=1}^{I}E\left((\overline{Y_{i.}} - \mu) - (\overline{\overline{Y_{..}}} - \mu)\right)^2$$

$$= J\sum_{i=1}^{I}E\left((\overline{Y_{i.}} - \theta)^2 + (\overline{\overline{Y_{..}}} - \theta)^2 - 2(\overline{Y_{i.}} - \theta)(\overline{\overline{Y_{..}}} - \theta)\right)$$

$$= J\sum_{i=1}^{I}\left(E(\overline{Y_{i.}} - \mu)^2 + E(\overline{\overline{Y_{..}}} - \mu)^2 - 2E(\overline{Y_{i.}} - \mu)(\overline{\overline{Y_{..}}} - \mu)\right)$$

$$= J\sum_{i=1}^{I}\left(Var(\overline{Y_{i.}} \mid \mu, \sigma_1^2, \sigma_2^2) + Var(\overline{\overline{Y_{..}}} \mid \mu, \sigma_1^2, \sigma_2^2) - 2Cov(\overline{Y_{i.}} \mid \mu, \sigma_1^2, \sigma_2^2, \overline{\overline{Y_{..}}} \mid \mu, \sigma_1^2, \sigma_2^2)\right)$$

$$= J\sum_{i=1}^{I}\left(\frac{(\sigma_1^2 + J\sigma_2^2)}{J} + \frac{(\sigma_1^2 + J\sigma_2^2)}{IJ} - 2\frac{(\sigma_1^2 + J\sigma_2^2)}{IJ}\right)$$

$$= J\left(\frac{I(\sigma_1^2 + J\sigma_2^2)}{J} + \frac{I(\sigma_1^2 + J\sigma_2^2)}{IJ} - 2\frac{I(\sigma_1^2 + J\sigma_2^2)}{IJ}\right)$$

$$= \left(I(\sigma_1^2 + J\sigma_2^2) + (\sigma_1^2 + J\sigma_2^2) - 2(\sigma_1^2 + J\sigma_2^2)\right)$$

$$= \left(I(\sigma_1^2 + J\sigma_2^2) - (\sigma_1^2 + J\sigma_2^2)\right)$$

$$= (\sigma_1^2 + J\sigma_2^2)(I - 1)$$

$$= (\sigma_1^2 + J\sigma_2^2)\nu_2 \text{ since } \nu_2 = (I - 1)$$

$$E(m_2) = \frac{(\sigma_1^2 + J\sigma_2^2)\nu_2}{\nu_2} = (\sigma_1^2 + J\sigma_2^2).$$

Proof of theorem 5.4.1

Since $$\mu \mid \underline{Y}, \sigma_1^2, \sigma_2^2 \sim N(\overline{\overline{Y..}}, \frac{(\sigma_1^2 + J\sigma_2^2)}{IJ})$$

and

$$\mu - l_0 \mid \underline{Y}, \sigma_1^2, \sigma_2^2 \sim N(\overline{\overline{Y..}} - l_0, \frac{(\sigma_1^2 + J\sigma_2^2)}{IJ})$$

it follows that

$$\frac{\mu - l_0}{3\left(\frac{(\sigma_{12}{}^2)}{J}\right)^{\frac{1}{2}}} \mid \underline{Y}, \sigma_1{}^2, \sigma_2{}^2 \sim N\left(\frac{\overline{\overline{Y..}} - l_0}{\left(\frac{\sigma_{12}{}^2}{J}\right)^{\frac{1}{2}}},\quad \frac{1}{9I}\right) \text{ where } \sigma_1{}^2 + J\sigma_2{}^2 = \sigma_{12}{}^2 .$$

Therefore

$$P_{pl}{}^1 \mid \underline{Y}, \sigma_1{}^2, \sigma_2{}^2 \sim N\left(\frac{\overline{\overline{Y..}} - l_0}{\left(\frac{\sigma_{12}{}^2}{J}\right)^{\frac{1}{2}}},\quad \frac{1}{9I}\right).$$

Proof of theorem 5.4.2

Since

$$P_{pl}{}^1 \mid \underline{Y}, \sigma_1{}^2, \sigma_2{}^2 \sim N\left(\frac{\overline{\overline{Y..}} - l_0}{\left(\frac{\sigma_{12}{}^2}{J}\right)^{\frac{1}{2}}},\quad \frac{1}{9I}\right)$$

it follows that

$$\mu'_1 = \frac{(\overline{\overline{Y}}_{..} - l_0)}{3\left(\frac{\sigma_{12}{}^2}{J}\right)^{\frac{1}{2}}} = \frac{(\overline{\overline{Y}}_{..} - l_0)J^{\frac{1}{2}}}{3\left(\sigma_{12}{}^2\right)^{\frac{1}{2}}},$$

$$\mu'_2 = \mu_2 + (\mu'_1)^2 = \frac{1}{9I} + \frac{(\overline{\overline{Y}}_{..} - l_0)^2 J}{9\left(\sigma_{12}{}^2\right)},$$

$$\mu'_3 = \mu_3 + 3\mu_2\mu'_1 + (\mu'_1)^3$$

$$= 0 + 3\left((\frac{1}{9I}).\frac{(\overline{\overline{Y}}_{..} - l_0)J^{\frac{1}{2}}}{3\left(\sigma_{12}{}^2\right)^{\frac{1}{2}}}\right) + \frac{(\overline{\overline{Y}}_{..} - l_0)^3 J^{\frac{3}{2}}}{27\left(\sigma_{12}{}^2\right)^{\frac{3}{2}}}$$

$$= \frac{(\overline{\overline{Y}}_{..} - l_0) J^{\frac{1}{2}}}{27} \left(\frac{3}{I\left(\sigma_{12}{}^2\right)^{\frac{1}{2}}} + \frac{(\overline{\overline{Y}}_{..} - l_0)^2 J}{\left(\sigma_{12}{}^2\right)^{\frac{3}{2}}} \right)$$

and

$$\mu'_4 = \mu_4 + 4\mu'_1\mu_3 + 6(\mu'_1)^2 \mu_2 + (\mu'_1)^4$$

$$= 3(\frac{1}{9I})^2 + 0 + 6\frac{(\overline{\overline{Y}}_{..} - l_0)^2 J}{9\left(\sigma_{12}{}^2\right)^{\frac{1}{2}}} . \frac{1}{9I} + \frac{(\overline{\overline{Y}}_{..} - l_0)^4 J^2}{81\left(\sigma_{12}{}^2\right)^2}$$

$$= \frac{1}{27} \left\{ \frac{1}{I^2} + (\overline{\overline{Y}}_{..} - l_0)^2 J \left(\frac{2}{I\sigma_{12}{}^2} + \frac{(\overline{\overline{Y}}_{..} - l_0)^2 J}{3\left(\sigma_{12}{}^2\right)^2} \right) \right\}.$$

Proof of theorem 5.5.1

The probability matching prior is derived from the Fisher information matrix. To obtain the Fisher information matrix, the negative of the expected values of the second derivatives (with respect to the parameters of the log–likelihood) must be calculated.

The Fisher information matrix of $\underline{\theta} = \left[\mu, \sigma_1{}^2, \sigma_2{}^2\right]$ per unit observation can be derived as follows:

The likelihood function is defined as

$$L\left(\mu,\sigma_1{}^2,\sigma_2{}^2|\underline{Y}\right) = \left((\sigma_1{}^2)^{-\frac{1}{2}\nu_1} (\sigma_1{}^2 + J\sigma_2{}^2)^{-\frac{1}{2}(\nu_2+1)} \exp(-\frac{1}{2}\{\frac{IJ(\overline{\overline{Y}}_{..} - \mu)^2}{(\sigma_1{}^2 + J\sigma_2{}^2)} + \frac{\nu_2 m_2}{(\sigma_1{}^2 + J\sigma_2{}^2)} + \frac{\nu_1 m_1}{\sigma_1{}^2}\}) \right)$$

and

$$\ell n L\left(\mu,\sigma_1{}^2,\sigma_2{}^2|\underline{Y}\right) = \ell = -\frac{\nu_1}{2}\ell n\left(\sigma_1{}^2\right) - \frac{1}{2}(\nu_2+1)\ell n(\sigma_1{}^2 + J\sigma_2{}^2) - \frac{1}{2}\frac{IJ(\overline{\overline{Y}}_{..} - \mu)^2}{(\sigma_1{}^2 + J\sigma_2{}^2)} - \frac{1}{2}\frac{\nu_2 m_2}{(\sigma_1{}^2 + J\sigma_2{}^2)} - \frac{1}{2}\frac{\nu_1 m_1}{\sigma_1{}^2}$$

Now

$$\frac{\partial \ell}{\partial \sigma_1{}^2} = -\frac{\nu_1}{2}\frac{1}{\sigma_1{}^2} - \frac{1}{2}\frac{(\nu_2+1)}{(\sigma_1{}^2 + J\sigma_2{}^2)} + \frac{1}{2}\frac{IJ(\overline{\overline{Y}}_{..} - \mu)^2}{(\sigma_1{}^2 + J\sigma_2{}^2)^2} + \frac{1}{2}\frac{\nu_2 m_2}{(\sigma_1{}^2 + J\sigma_2{}^2)^2} + \frac{1}{2}\frac{\nu_1 m_1}{(\sigma_1{}^2)^2}$$

and

$$\frac{\partial^2\ell}{(\partial\sigma_1^2)^2}=\frac{\nu_1}{2}\frac{1}{(\sigma_1^2)^2}+\frac{1}{2}\frac{(\nu_2+1)}{(\sigma_1^2+J\sigma_2^2)^2}-\frac{IJ(\overline{\overline{Y}}_{..}-\mu)^2}{(\sigma_1^2+J\sigma_2^2)^3}-\frac{\nu_2 m_2}{(\sigma_1^2+J\sigma_2^2)^3}-\frac{\nu_1 m_1}{(\sigma_1^2)^3}.$$

Where

$$\overline{\overline{Y}}_{..}\mid\mu,\sigma_1^2,\sigma_2^2\sim N\left(\mu,\frac{\sigma_1^2+J\sigma_2^2}{IJ}\right),\quad E(m_1)=\sigma_1^2\quad and\quad E(m_2)=\sigma_1^2+J\sigma_2^2$$

Therefore

$$E\left(\frac{\partial^2\ell}{(\partial\sigma_1^2)^2}\right)=\frac{\nu_1}{2}\frac{1}{(\sigma_1^2)^2}+\frac{1}{2}\frac{(\nu_2+1)}{(\sigma_1^2+J\sigma_2^2)^2}-\frac{IJ.E(\overline{\overline{Y}}_{..}-\mu)^2}{(\sigma_1^2+J\sigma_2^2)^3}-\frac{\nu_2 E(m_2)}{(\sigma_1^2+J\sigma_2^2)^3}-\frac{\nu_1 E(m_1)}{(\sigma_1^2)^3}$$

$$-E\left(\frac{\partial^2\ell}{(\partial\sigma_1^2)^2}\right)=-\frac{\nu_1}{2}\frac{1}{(\sigma_1^2)^2}-\frac{1}{2}\frac{(\nu_2+1)}{(\sigma_1^2+J\sigma_2^2)^2}+\frac{IJ\,\frac{(\sigma_1^2+J\sigma_2^2)}{IJ}}{(\sigma_1^2+J\sigma_2^2)^3}+\frac{\nu_2(\sigma_1^2+J\sigma_2^2)}{(\sigma_1^2+J\sigma_2^2)^3}+\frac{\nu_1\sigma_1^2}{(\sigma_1^2)^3}$$

$$-E\left(\frac{\partial^2\ell}{(\partial\sigma_1^2)^2}\right)=-\frac{\nu_1}{2}\frac{1}{(\sigma_1^2)^2}-\frac{1}{2}\frac{(\nu_2+1)}{(\sigma_1^2+J\sigma_2^2)^2}+\frac{1}{(\sigma_1^2+J\sigma_2^2)^2}+\frac{\nu_2}{(\sigma_1^2+J\sigma_2^2)^2}+\frac{\nu_1}{(\sigma_1^2)^2}$$

$$-E\left(\frac{\partial^2\ell}{(\partial\sigma_1^2)^2}\right)=\frac{\nu_1}{(\sigma_1^2)^2}-\frac{\nu_1}{2}\frac{1}{(\sigma_1^2)^2}+\frac{\nu_2}{(\sigma_1^2+J\sigma_2^2)^2}+\frac{1}{(\sigma_1^2+J\sigma_2^2)^2}+-\frac{1}{2}\frac{(\nu_2+1)}{(\sigma_1^2+J\sigma_2^2)^2}$$

$$-E\left(\frac{\partial^2\ell}{(\partial\sigma_1^2)^2}\right)=\frac{\nu_1}{(\sigma_1^2)^2}-\frac{\nu_1}{2}\frac{1}{(\sigma_1^2)^2}+\frac{(\nu_2+1)}{(\sigma_1^2+J\sigma_2^2)^2}-\frac{1}{2}\frac{(\nu_2+1)}{(\sigma_1^2+J\sigma_2^2)^2}$$

$$-E\left(\frac{\partial^2\ell}{(\partial\sigma_1^2)^2}\right)=\frac{\nu_1}{2}\frac{1}{(\sigma_1^2)^2}+\frac{1}{2}\frac{(\nu_2+1)}{(\sigma_1^2+J\sigma_2^2)^2}.$$

Further

$$\frac{\partial\ell}{\partial\sigma_2^2}=-\frac{J}{2}(\nu_2+1)\frac{1}{(\sigma_1^2+J\sigma_2^2)}+\frac{1}{2}\frac{IJ^2(\overline{\overline{Y}}_{..}-\mu)^2}{(\sigma_1^2+J\sigma_2^2)^2}+\frac{1}{2}\frac{\nu_2 m_2 J}{(\sigma_1^2+J\sigma_2^2)^2}$$

$$\frac{\partial^2\ell}{(\partial\sigma_2^2)^2}=\frac{J^2}{2}\frac{(\nu_2+1)}{(\sigma_1^2+J\sigma_2^2)^2}-\frac{IJ^3(\overline{\overline{Y}}_{..}-\mu)^2}{(\sigma_1^2+J\sigma_2^2)^3}-\frac{\nu_2 m_2 J^2}{(\sigma_1^2+J\sigma_2^2)^3}$$

$$E\left(\frac{\partial^2\ell}{(\partial\sigma_2^2)^2}\right)=\frac{J^2}{2}\frac{(\nu_2+1)}{(\sigma_1^2+J\sigma_2^2)^2}-\frac{IJ^3E(\overline{\overline{Y}}_{..}-\mu)^2}{(\sigma_1^2+J\sigma_2^2)^3}-\frac{\nu_2 E(m_2)J^2}{(\sigma_1^2+J\sigma_2^2)^3}$$

$$E\left(\frac{\partial^2\ell}{(\partial\sigma_2^2)^2}\right)=\frac{J^2}{2}\frac{(\nu_2+1)}{(\sigma_1^2+J\sigma_2^2)^2}-\frac{IJ^3\,\frac{(\sigma_1^2+J\sigma_2^2)}{IJ}}{(\sigma_1^2+J\sigma_2^2)^3}-\frac{\nu_2(\sigma_1^2+J\sigma_2^2)J^2}{(\sigma_1^2+J\sigma_2^2)^3}$$

$$-E\left(\frac{\partial^2\ell}{(\partial\sigma_2{}^2)^2}\right)=-\frac{J^2}{2}\frac{(\nu_2+1)}{(\sigma_1{}^2+J\sigma_2{}^2)^2}+\frac{J^2}{(\sigma_1{}^2+J\sigma_2{}^2)^2}+\frac{\nu_2 J^2}{(\sigma_1{}^2+J\sigma_2{}^2)^2}$$

$$-E\left(\frac{\partial^2\ell}{(\partial\sigma_2{}^2)^2}\right)=\frac{J^2(\nu_2+1)}{(\sigma_1{}^2+J\sigma_2{}^2)^2}-\frac{J^2}{2}\frac{(\nu_2+1)}{(\sigma_1{}^2+J\sigma_2{}^2)^2}$$

$$-E\left(\frac{\partial^2\ell}{(\partial\sigma_2{}^2)^2}\right)=\frac{J^2(\nu_2+1)}{2(\sigma_1{}^2+J\sigma_2{}^2)^2}.$$

Also

$$\frac{\partial\ell}{\partial\mu}=\frac{IJ(\overline{\overline{Y}}_{..}-\mu)}{(\sigma_1{}^2+J\sigma_2{}^2)}.$$

And

$$\frac{\partial^2\ell}{\partial\mu^2}=\frac{-IJ}{(\sigma_1{}^2+J\sigma_2{}^2)}.$$

Therefore

$$E\left(\frac{\partial^2\ell}{\partial\mu^2}\right)=E\left(\frac{-IJ}{(\sigma_1{}^2+J\sigma_2{}^2)}\right)\ \text{therefore}\ -E\left(\frac{\partial^2\ell}{\partial\mu^2}\right)=\left(\frac{IJ}{(\sigma_1{}^2+J\sigma_2{}^2)}\right)$$

Also

$$\frac{\partial^2\ell}{\partial\mu\partial\sigma_1{}^2}=\frac{-IJ(\overline{\overline{Y}}_{..}-\mu)}{(\sigma_1{}^2+J\sigma_2{}^2)^2}\ \text{and}\ E\left(\frac{\partial^2\ell}{\partial\mu\partial\sigma_1{}^2}\right)=\frac{-IJ.E(\overline{\overline{Y}}_{..}-\mu)}{(\sigma_1{}^2+J\sigma_2{}^2)^2}=0.$$

Therefore

$$-E\left(\frac{\partial^2\ell}{\partial\mu\partial\sigma_1{}^2}\right)=0.$$

Similarly

$$\frac{\partial^2\ell}{\partial\mu\partial\sigma_2{}^2}=\frac{-IJ^2(\overline{\overline{Y}}_{..}-\mu)}{(\sigma_1{}^2+J\sigma_2{}^2)^2}\ \text{and}\ E\left(\frac{\partial^2\ell}{\partial\mu\partial\sigma_2{}^2}\right)=\frac{-IJ^2.E(\overline{\overline{Y}}_{..}-\mu)}{(\sigma_1{}^2+J\sigma_2{}^2)^2}=0.$$

$$-E\left(\frac{\partial^2\ell}{\partial\mu\partial{\sigma_2}^2}\right)=0 .$$

Further

$$\frac{\partial\ell}{\partial{\sigma_2}^2\partial{\sigma_1}^2}=\frac{J(\nu_2+1)}{2({\sigma_1}^2+J{\sigma_2}^2)^2}-\frac{IJ^2(\overline{\overline{Y}}_{..}-\theta)^2}{({\sigma_1}^2+J{\sigma_2}^2)^3}-\frac{\nu_2 m_2 J}{({\sigma_1}^2+J{\sigma_2}^2)^2}$$

$$E\left(\frac{\partial\ell}{\partial{\sigma_2}^2\partial{\sigma_1}^2}\right)=\frac{J(\nu_2+1)}{2({\sigma_1}^2+J{\sigma_2}^2)^2}-\frac{IJ^2E(\overline{\overline{Y}}_{..}-\theta)^2}{({\sigma_1}^2+J{\sigma_2}^2)^3}-\frac{\nu_2 E(m_2)J}{({\sigma_1}^2+J{\sigma_2}^2)^2}$$

$$E\left(\frac{\partial\ell}{\partial{\sigma_2}^2\partial{\sigma_1}^2}\right)=\frac{J(\nu_2+1)}{2({\sigma_1}^2+J{\sigma_2}^2)^2}-\frac{IJ^2\dfrac{({\sigma_1}^2+J{\sigma_2}^2)}{IJ}}{({\sigma_1}^2+J{\sigma_2}^2)^3}-\frac{\nu_2({\sigma_1}^2+J{\sigma_2}^2)J}{({\sigma_1}^2+J{\sigma_2}^2)^2}$$

$$E\left(\frac{\partial\ell}{\partial{\sigma_2}^2\partial{\sigma_1}^2}\right)=\frac{J(\nu_2+1)}{2({\sigma_1}^2+J{\sigma_2}^2)^2}-\frac{J}{({\sigma_1}^2+J{\sigma_2}^2)^2}-\frac{\nu_2 J}{({\sigma_1}^2+J{\sigma_2}^2)}$$

$$-E\left(\frac{\partial\ell}{\partial{\sigma_2}^2\partial{\sigma_1}^2}\right)=-\frac{J(\nu_2+1)}{2({\sigma_1}^2+J{\sigma_2}^2)^2}+\frac{J(\nu_2+1)}{({\sigma_1}^2+J{\sigma_2}^2)^2}$$

$$-E\left(\frac{\partial\ell}{\partial{\sigma_2}^2\partial{\sigma_1}^2}\right)=\frac{J(\nu_2+1)}{2({\sigma_1}^2+J{\sigma_2}^2)^2} .$$

To summarise the results derived so far and substituting $\nu_1=I(J-1), \quad \nu_2=I-1$:

$$-E\left(\frac{\partial^2\ell}{\partial\mu^2}\right)=\left(\frac{IJ}{({\sigma_1}^2+J{\sigma_2}^2)}\right),-E\left(\frac{\partial^2\ell}{(\partial{\sigma_2}^2)^2}\right)=\frac{J^2(\nu_2+1)}{2({\sigma_1}^2+J{\sigma_2}^2)^2}=\frac{J^2I}{2({\sigma_1}^2+J{\sigma_2}^2)^2}$$

$$-E\left(\frac{\partial^2\ell}{(\partial{\sigma_1}^2)^2}\right)=\frac{\nu_1}{2}\frac{1}{({\sigma_1}^2)^2}+\frac{1}{2}\frac{(\nu_2+1)}{({\sigma_1}^2+J{\sigma_2}^2)^2}=\frac{I(J-1)}{2({\sigma_1}^2)^2}+\frac{I}{2({\sigma_1}^2+J{\sigma_2}^2)^2}$$

$$-E\left(\frac{\partial^2\ell}{\partial\mu\partial{\sigma_2}^2}\right)=0, \qquad -E\left(\frac{\partial^2\ell}{\partial\mu\partial{\sigma_1}^2}\right)=0,$$

$$-E\left(\frac{\partial\ell}{\partial{\sigma_2}^2\partial{\sigma_1}^2}\right)=\frac{J(\nu_2+1)}{2({\sigma_1}^2+J{\sigma_2}^2)^2}=\frac{JI}{2({\sigma_1}^2+J{\sigma_2}^2)^2} .$$

The Fisher information matrix is then given by

$$F(\mu,\sigma_1^2,\sigma_2^2)=F(\underline{\theta})=\begin{pmatrix} -E\left(\dfrac{\partial^2\ell}{\partial\mu^2}\right) & -E\left(\dfrac{\partial^2\ell}{\partial\mu\partial\sigma_2^2}\right) & -E\left(\dfrac{\partial^2\ell}{\partial\mu\partial\sigma_1^2}\right) \\ -E\left(\dfrac{\partial^2\ell}{\partial\mu\partial\sigma_2^2}\right) & -E\left(\dfrac{\partial^2\ell}{(\partial\sigma_2^2)^2}\right) & -E\left(\dfrac{\partial\ell}{\partial\sigma_2^2\partial\sigma_1^2}\right) \\ -E\left(\dfrac{\partial^2\ell}{\partial\mu\partial\sigma_1^2}\right) & -E\left(\dfrac{\partial\ell}{\partial\sigma_2^2\partial\sigma_1^2}\right) & -E\left(\dfrac{\partial^2\ell}{(\partial\sigma_1^2)^2}\right) \end{pmatrix}$$

$$F(\mu,\sigma_1^2,\sigma_2^2)=F(\underline{\theta})=\begin{pmatrix} \dfrac{IJ}{(\sigma_1^2+J\sigma_2^2)} & 0 & 0 \\ 0 & \dfrac{IJ^2}{2(\sigma_1^2+J\sigma_2^2)^2} & \dfrac{IJ}{2(\sigma_1^2+J\sigma_2^2)^2} \\ 0 & \dfrac{IJ}{2(\sigma_1^2+J\sigma_2^2)^2} & \dfrac{I(J-1)}{2(\sigma_1^2)^2}+\dfrac{I}{2(\sigma_1^2+J\sigma_2^2)^2} \end{pmatrix}.$$

And the inverse of the Fisher information matrix is given by

$$F^{-1}(\mu,\sigma_1^2,\sigma_2^2)=F^{-1}(\underline{\theta})=\begin{pmatrix} \dfrac{(\sigma_1^2+J\sigma_2^2)}{IJ} & 0 & 0 \\ 0 & \dfrac{\left\{\dfrac{I(J-1)}{2(\sigma_1^2)^2}+\dfrac{I}{2(\sigma_1^2+J\sigma_2^2)^2}\right\}}{|H|} & \dfrac{-\dfrac{IJ}{2(\sigma_1^2+J\sigma_2^2)^2}}{|H|} \\ 0 & \dfrac{-\dfrac{IJ}{2(\sigma_1^2+J\sigma_2^2)^2}}{|H|} & \dfrac{\dfrac{IJ^2}{2(\sigma_1^2+J\sigma_2^2)^2}}{|H|} \end{pmatrix}$$

The determinant $|H|$ is equal to

$$|H| = \frac{IJ^2}{2(\sigma_1^2 + J\sigma_2^2)^2}\left\{\frac{I(J-1)}{2(\sigma_1^2)^2} + \frac{I}{2(\sigma_1^2 + J\sigma_2^2)^2}\right\} - \frac{I^2J^2}{4(\sigma_1^2 + J\sigma_2^2)^4}$$

$$= \frac{I^2J^2}{4(\sigma_1^2 + J\sigma_2^2)^2}\left\{\frac{(J-1)}{(\sigma_1^2)^2} + \frac{I}{(\sigma_1^2 + J\sigma_2^2)^2} - \frac{1}{(\sigma_1^2 + J\sigma_2^2)^2}\right\}$$

$$= \frac{I^2J^2}{4(\sigma_1^2 + J\sigma_2^2)^2}\left\{\frac{(J-1)}{(\sigma_1^2)^2}\right\}.$$

$$F^{-1}(\underline{\theta}) = F^{-1}(\mu, \sigma_2^2, \sigma_1^2) = \begin{pmatrix} \frac{(\sigma_1^2 + J\sigma_2^2)}{IJ} & 0 & 0 \\ 0 & \frac{2\left\{(J-1)(\sigma_1^2 + J\sigma_2^2)^2 + (\sigma_1^2)^2\right\}}{IJ^2(J-1)} & \frac{-2(\sigma_1^2)^2}{IJ(J-1)} \\ 0 & \frac{-2(\sigma_1^2)^2}{IJ(J-1)} & \frac{2(\sigma_1^2)^2}{I(J-1)} \end{pmatrix}$$

The interest of this study lies in the probability matching prior for (P_{pl}^{1}), the lower process performance index.

Let $\underline{\theta} = [\mu, \sigma_2^2, \sigma_1^2]'$. The capability index is

$$P_{pl}^{1} = t(\underline{\theta}) = \frac{\mu - l_0}{3\left(\frac{\sigma_1^2 + J\sigma_2^2}{J}\right)^{\frac{1}{2}}} = \frac{(\mu - l_0)J^{\frac{1}{2}}}{3\left(\sigma_1^2 + J\sigma_2^2\right)^{\frac{1}{2}}}.$$

Therefore

$$\frac{\partial t(\underline{\theta})}{\partial \mu} = \frac{J^{\frac{1}{2}}}{3\left(\sigma_1^2 + J\sigma_2^2\right)^{\frac{1}{2}}}, \quad \frac{\partial t(\underline{\theta})}{\partial \sigma_2^2} = \frac{-(\mu - l_0)J^{\frac{3}{2}}}{6\left(\sigma_1^2 + J\sigma_2^2\right)^{\frac{3}{2}}} \text{ and } \frac{\partial t(\underline{\theta})}{\partial \sigma_1^2} = \frac{-(\mu - l_0)J^{\frac{1}{2}}}{6\left(\sigma_1^2 + J\sigma_2^2\right)^{\frac{3}{2}}}$$

As mentioned

$$\nabla'_t(\tilde{\theta}) = \begin{bmatrix} \frac{\partial t(\underline{\theta})}{\partial \mu} & \frac{\partial t(\underline{\theta})}{\partial \sigma_2^2} & \frac{\partial t(\underline{\theta})}{\partial \sigma_1^2} \end{bmatrix} = \begin{bmatrix} \frac{J^{\frac{1}{2}}}{3\left(\sigma_1^2 + J\sigma_2^2\right)^{\frac{1}{2}}} & \frac{-(\theta - l_0)J^{\frac{3}{2}}}{6\left(\sigma_1^2 + J\sigma_2^2\right)^{\frac{3}{2}}} & \frac{-(\theta - l_0)J^{\frac{1}{2}}}{6\left(\sigma_1^2 + J\sigma_2^2\right)^{\frac{3}{2}}} \end{bmatrix}$$

$$\nabla'_t(\underline{\theta}) = \frac{J^{\frac{1}{2}}}{3\left(\sigma_1^2 + J\sigma_2^2\right)^{\frac{1}{2}}}\begin{bmatrix} 1 & \frac{-(\mu - l_0)J}{2\left(\sigma_1^2 + J\sigma_2^2\right)} & \frac{-(\mu - l_0)}{2\left(\sigma_1^2 + J\sigma_2^2\right)} \end{bmatrix}.$$

Further

$$\nabla'_t(\underline{\theta})F^{-1}(\underline{\theta}) = \frac{J^{\frac{1}{2}}}{3\left(\sigma_1^2+J\sigma_2^2\right)^{\frac{1}{2}}}\begin{bmatrix}1 & \frac{-(\mu-l_0)J}{2\left(\sigma_1^2+J\sigma_2^2\right)} & \frac{-(\mu-l_0)}{2\left(\sigma_1^2+J\sigma_2^2\right)}\end{bmatrix}\times$$

$$\begin{pmatrix}\frac{(\sigma_1^2+J\sigma_2^2)}{IJ} & 0 & 0 \\ 0 & \frac{\left\{\frac{I(J-1)}{2(\sigma_1^2)^2}+\frac{I}{2(\sigma_1^2+J\sigma_2^2)^2}\right\}}{|H|} & \frac{-\frac{IJ}{2(\sigma_1^2+J\sigma_2^2)^2}}{|H|} \\ 0 & \frac{-\frac{IJ}{2(\sigma_1^2+J\sigma_2^2)^2}}{|H|} & \frac{\frac{IJ^2}{2(\sigma_1^2+J\sigma_2^2)^2}}{|H|}\end{pmatrix}$$

$$\nabla'_t(\underline{\theta})F^{-1}(\underline{\theta})=\frac{J^{\frac{1}{2}}}{3\left(\sigma_1^2+J\sigma_2^2\right)^{\frac{1}{2}}}\begin{bmatrix}\frac{(\sigma_1^2+J\sigma_2^2)}{IJ} & \frac{-(\mu-l_0)\left(\sigma_1^2+J\sigma_2^2\right)}{IJ} & 0\end{bmatrix}$$

$$\nabla'_t(\underline{\theta})F^{-1}(\underline{\theta})\nabla_t(\underline{\theta})=\frac{J}{9\left(\sigma_1^2+J\sigma_2^2\right)}\begin{bmatrix}\frac{(\sigma_1^2+J\sigma_2^2)}{IJ} & \frac{-(\mu-l_0)\left(\sigma_1^2+J\sigma_2^2\right)}{IJ} & 0\end{bmatrix}\begin{bmatrix}1 \\ \frac{-(\mu-l_0)J}{2\left(\sigma_1^2+J\sigma_2^2\right)} \\ \frac{-(\mu-l_0)}{6\left(\sigma_1^2+J\sigma_2^2\right)}\end{bmatrix}$$

$$\nabla'_t(\underline{\theta})F^{-1}(\underline{\theta})\nabla_t(\underline{\theta}) = \frac{J}{9\left(\sigma_1^2+J\sigma_2^2\right)}\left(\frac{(\sigma_1^2+J\sigma_2^2)}{IJ}+\frac{(\mu-l_0)^2}{2I}\right)$$

$$=\left(\frac{1}{9I}+\frac{J(\mu-l_0)^2}{18I\left(\sigma_1^2+J\sigma_2^2\right)}\right)$$

$$\left\{\nabla'_t(\underline{\theta})F^{-1}(\underline{\theta})\nabla_t(\underline{\theta})\right\}^{\frac{1}{2}} = \left(\frac{1}{9I}+\frac{J(\mu-l_0)^2}{18I\left(\sigma_1^2+J\sigma_2^2\right)}\right)^{\frac{1}{2}}$$

$$\left\{\nabla'_t(\underline{\theta})F^{-1}(\underline{\theta})\nabla_t(\underline{\theta})\right\}^{\frac{1}{2}} = \frac{1}{3I^{\frac{1}{2}}}\left(1+\frac{J(\mu-l_0)^2}{2\left(\sigma_1^2+J\sigma_2^2\right)}\right)^{\frac{1}{2}}.$$

Define as before

$$\eta(\underline{\theta}) = \frac{\nabla'_t(\underline{\theta})(\underline{\theta})F^{-1}(\underline{\theta})}{\sqrt{\nabla'_t(\underline{\theta})F^{-1}(\underline{\theta})\nabla_t(\underline{\theta})}} = [\eta_1(\underline{\theta}) \quad \eta_2(\underline{\theta}) \quad \eta_3(\underline{\theta})]$$

where

$$\nabla'_t(\underline{\theta})F^{-1}(\underline{\theta}) = \frac{J^{\frac{1}{2}}}{3\left(\sigma_1^2 + J\sigma_2^2\right)^{\frac{1}{2}}}\left[\frac{(\sigma_1^2 + J\sigma_2^2)}{IJ} \quad \frac{-(\mu - l_0)\left(\sigma_1^2 + J\sigma_2^2\right)}{IJ} \quad 0\right]$$

and

$$\left\{\nabla'_t(\underline{\theta})F^{-1}(\underline{\theta})\nabla_t(\underline{\theta})\right\}^{\frac{1}{2}} = \frac{1}{3I^{\frac{1}{2}}}\left(1 + \frac{J(\mu - l_0)^2}{2\left(\sigma_1^2 + J\sigma_2^2\right)}\right)^{\frac{1}{2}}$$

giving

$$\eta_1(\underline{\theta}) = \frac{\left(\sigma_1^2 + J\sigma_2^2\right)^{\frac{1}{2}}}{3IJ^{\frac{1}{2}}} 3I^{\frac{1}{2}}\left(1 + \frac{J(\mu - l_0)^2}{2\left(\sigma_1^2 + J\sigma_2^2\right)}\right)^{-\frac{1}{2}}$$

$$\eta_2(\underline{\theta}) = \frac{-(\theta - l_0)\left(\sigma_1^2 + J\sigma_2^2\right)^{\frac{1}{2}}}{3IJ^{\frac{1}{2}}} 3I^{\frac{1}{2}}\left(1 + \frac{J(\mu - l_0)^2}{2\left(\sigma_1^2 + J\sigma_2^2\right)}\right)^{-\frac{1}{2}} \text{ and } \eta_3(\underline{\theta}) = 0 \; .$$

For a prior $\pi(\underline{\theta})$ to be a probability-matching prior, the differential equation

$$\sum_{\alpha=1}^{m} \frac{\partial}{\partial \theta_\alpha}\{\eta_\alpha(\underline{\theta})\pi(\underline{\theta})\} = 0$$

$$\frac{\partial}{\partial \mu}\{\eta_1(\underline{\theta})\pi(\underline{\theta})\} + \frac{\partial}{\partial \sigma_2^2}\{\eta_2(\underline{\theta})\pi(\underline{\theta})\} + \frac{\partial}{\partial \sigma_1^2}\{\eta_3(\underline{\theta})\pi(\underline{\theta})\} = 0 \; .$$

must be satisfied.

The priors

i. $\pi_a(\underline{\theta}) = \pi_a(\mu, \sigma_2^2, \sigma_1^2) \propto \sigma_1^{-2}\left(\sigma_1^2 + J\sigma_2^2\right)^{-\frac{3}{2}}\left(1 + \frac{J(\mu - l_0)^2}{2\left(\sigma_1^2 + J\sigma_2^2\right)}\right)^{-\frac{1}{2}}$

will also be a probability matching prior since

In this case

$$\frac{\partial}{\partial \mu}\{\eta_1(\underline{\theta})\pi_a(\underline{\theta})\}$$

$$=\frac{\partial}{\partial\mu}\left\{\frac{\left(\sigma_1{}^2+J\sigma_2{}^2\right)^{\frac{1}{2}}}{3IJ^{\frac{1}{2}}}3I^{\frac{1}{2}}\left(1+\frac{J(\mu-l_0)^2}{2\left(\sigma_1{}^2+J\sigma_2{}^2\right)}\right)^{-\frac{1}{2}}.\sigma_1{}^{-2}\left(\sigma_1{}^2+J\sigma_2{}^2\right)^{-\frac{3}{2}}\left(1+\frac{J(\mu-l_0)^2}{2\left(\sigma_1{}^2+J\sigma_2{}^2\right)}\right)^{-\frac{1}{2}}\right\}$$

$$=\frac{\partial}{\partial\mu}\left\{\frac{\sigma_1^{-2}\left(\sigma_1^{2}+J\sigma_2^{2}\right)^{-1}}{I^{\frac{1}{2}}J^{\frac{1}{2}}}\left(1+\frac{J(\mu-l_0)^2}{2\left(\sigma_1^{2}+J\sigma_2^{2}\right)}\right)^{-1}\right\}$$

$$=\frac{\partial}{\partial\mu}\left\{\frac{\sigma_1^{-2}\left(\sigma_1^{2}+J\sigma_2^{2}\right)^{-1}}{I^{\frac{1}{2}}J^{\frac{1}{2}}}\left(\frac{2\left(\sigma_1^{2}+J\sigma_2^{2}\right)+J(\mu-l_0)^2}{2\left(\sigma_1^{2}+J\sigma_2^{2}\right)}\right)^{-1}\right\}$$

$$=\frac{\partial}{\partial\mu}\left\{\frac{\sigma_1^{-2}\left(\sigma_1^{2}+J\sigma_2^{2}\right)^{-1}}{I^{\frac{1}{2}}J^{\frac{1}{2}}}\left(\frac{2\left(\sigma_1^{2}+J\sigma_2^{2}\right)}{2\left(\sigma_1^{2}+J\sigma_2^{2}\right)+J(\mu-l_0)^2}\right)\right\}$$

$$=\frac{\partial}{\partial\mu}\left\{\frac{\sigma_1^{-2}}{I^{\frac{1}{2}}J^{\frac{1}{2}}}\left(\frac{2}{2\left(\sigma_1^{2}+J\sigma_2^{2}\right)+J(\mu-l_0)^2}\right)\right\}$$

$$=\left\{\frac{\sigma_1^{-2}}{I^{\frac{1}{2}}J^{\frac{1}{2}}}\frac{\partial}{\partial\theta}\left(2\left[2\left(\sigma_1^{2}+J\sigma_2^{2}\right)+J(\mu-l_0)^2\right]^{-1}\right)\right\}$$

$$=\left\{\frac{\sigma_1^{-2}}{I^{\frac{1}{2}}J^{\frac{1}{2}}}\left((-1)2\left[2\left(\sigma_1^{2}+J\sigma_2^{2}\right)+J(\mu-l_0)^2\right]^{-2}2J(\mu-l_0)\right)\right\}$$

$$=\left\{\frac{-2J(\mu-l_0)\sigma_1^{-2}}{I^{\frac{1}{2}}J^{\frac{1}{2}}}\left(2\left[2\left(\sigma_1^{2}+J\sigma_2^{2}\right)+J(\mu-l_0)^2\right]^{-2}\right)\right\}$$

$$=\left\{\frac{-2J(\mu-l_0)\sigma_1^{-2}}{I^{\frac{1}{2}}J^{\frac{1}{2}}}\left(\frac{2}{\left(2\left(\sigma_1^{2}+J\sigma_2^{2}\right)+J(\mu-l_0)^2\right)^2}\right)\right\}$$

$$=\left\{\frac{-J(\mu-l_0)\sigma_1^{-2}}{I^{\frac{1}{2}}J^{\frac{1}{2}}}\left(\sigma_1^{2}+J\sigma_2^{2}\right)^{-2}\left(\frac{1}{\left(1+\frac{J(\mu-l_0)^2}{2\left(\sigma_1^{2}+J\sigma_2^{2}\right)}\right)^2}\right)\right\}$$

$$=\left\{\frac{-J(\mu-l_0)\sigma_1^{-2}}{I^{\frac{1}{2}}J^{\frac{1}{2}}}\left(\sigma_1^{2}+J\sigma_2^{2}\right)^{-2}\left(1+\frac{J(\mu-l_0)^2}{2\left(\sigma_1^{2}+J\sigma_2^{2}\right)}\right)^{-2}\right\}.$$

Further

$$\frac{\partial}{\partial {\sigma_2}^2}\{\eta_2(\underline{\theta})\pi_a(\underline{\theta})\}$$

$$=\frac{\partial}{\partial {\sigma_2}^2}\left\{\frac{-(\mu-l_0)\left({\sigma_1}^2+J{\sigma_2}^2\right)^{\frac{1}{2}}}{3IJ^{\frac{1}{2}}}3I^{\frac{1}{2}}\left(1+\frac{J(\mu-l_0)^2}{2\left({\sigma_1}^2+J{\sigma_2}^2\right)}\right)^{-\frac{1}{2}}.{\sigma_1}^{-2}\left({\sigma_1}^2+J{\sigma_2}^2\right)^{-\frac{3}{2}}\left(1+\frac{J(\mu-l_0)^2}{2\left({\sigma_1}^2+J{\sigma_2}^2\right)}\right)^{-\frac{1}{2}}\right\}$$

$$=\frac{\partial}{\partial {\sigma_2}^2}\left\{\frac{-(\mu-l_0){\sigma_1}^{-2}\left({\sigma_1}^2+J{\sigma_2}^2\right)^{-1}}{I^{\frac{1}{2}}J^{\frac{1}{2}}}\left(1+\frac{J(\mu-l_0)^2}{2\left({\sigma_1}^2+J{\sigma_2}^2\right)}\right)^{-1}\right\}$$

$$=\frac{\partial}{\partial {\sigma_2}^2}\left\{\frac{-(\mu-l_0){\sigma_1}^{-2}\left({\sigma_1}^2+J{\sigma_2}^2\right)^{-1}}{I^{\frac{1}{2}}J^{\frac{1}{2}}}\left(\frac{2\left({\sigma_1}^2+J{\sigma_2}^2\right)+J(\mu-l_0)^2}{2\left({\sigma_1}^2+J{\sigma_2}^2\right)}\right)^{-1}\right\}$$

$$=\frac{\partial}{\partial {\sigma_2}^2}\left\{\frac{-(\mu-l_0){\sigma_1}^{-2}\left({\sigma_1}^2+J{\sigma_2}^2\right)^{-1}}{I^{\frac{1}{2}}J^{\frac{1}{2}}}\left(\frac{2\left({\sigma_1}^2+J{\sigma_2}^2\right)}{2\left({\sigma_1}^2+J{\sigma_2}^2\right)+J(\mu-l_0)^2}\right)\right\}$$

$$=\frac{\partial}{\partial {\sigma_2}^2}\left\{\frac{-(\mu-l_0){\sigma_1}^{-2}}{I^{\frac{1}{2}}J^{\frac{1}{2}}}\left(\frac{1}{\left({\sigma_1}^2+J{\sigma_2}^2\right)+\frac{J(\mu-l_0)^2}{2}}\right)\right\}$$

$$=\left\{\frac{-(\mu-l_0){\sigma_1}^{-2}}{I^{\frac{1}{2}}J^{\frac{1}{2}}}\frac{\partial}{\partial {\sigma_2}^2}\left[\left({\sigma_1}^2+J{\sigma_2}^2\right)+\frac{J(\mu-l_0)^2}{2}\right]^{-1}\right\}$$

$$=\left\{\frac{-(\mu-l_0){\sigma_1}^{-2}}{I^{\frac{1}{2}}J^{\frac{1}{2}}}(-1)\left[\left({\sigma_1}^2+J{\sigma_2}^2\right)+\frac{J(\mu-l_0)^2}{2}\right]^{-2}J\right\}$$

$$=\left\{\frac{J(\mu-l_0){\sigma_1}^{-2}\left({\sigma_1}^2+J{\sigma_2}^2\right)^{-2}}{I^{\frac{1}{2}}J^{\frac{1}{2}}}\left[1+\frac{J(\mu-l_0)^2}{2\left({\sigma_1}^2+J{\sigma_2}^2\right)}\right]^{-2}\right\}.$$

Also

$$\frac{\partial}{\partial {\sigma_1}^2}\{\eta_3(\underline{\theta})\pi_a(\underline{\theta})\}$$

$$=\frac{\partial}{\partial\sigma_1^2}\left\{0\,.\,\sigma_1^{-2}\left(\sigma_1^2+J\sigma_2^2\right)^{-\frac{3}{2}}\left(1+\quad\frac{J(\mu-l_0)^2}{2\left(\sigma_1^2+J\sigma_2^2\right)}\right)^{-\frac{1}{2}}\right\}\quad=\frac{\partial}{\partial\sigma_1^2}\{0\}=0$$
.

Therefore

$$\sum_{\alpha=1}^{m}\frac{\partial}{\partial\theta_\alpha}\{\eta_\alpha(\underline{\theta})\pi_a(\underline{\theta})\}=\frac{\partial}{\partial\mu}\{\eta_1(\underline{\theta})\pi_a(\underline{\theta})\}+\frac{\partial}{\partial\sigma_2^2}\{\eta_2(\underline{\theta})\pi_a(\underline{\theta})\}+\frac{\partial}{\partial\sigma_2^2}\{\eta_3(\underline{\theta})\pi_a(\underline{\theta})\}$$

$$=\left\{\frac{-J(\mu-l_0)\sigma_1^{-2}\left(\sigma_1^2+J\sigma_2^2\right)^{-2}}{I^{\frac{1}{2}}J^{\frac{1}{2}}}\left[1+\frac{J(\mu-l_0)^2}{2\left(\sigma_1^2+J\sigma_2^2\right)}\right]^{-2}\right\}+$$

$$\left\{\frac{J(\mu-l_0)\sigma_1^{-2}\left(\sigma_1^2+J\sigma_2^2\right)^{-2}}{I^{\frac{1}{2}}J^{\frac{1}{2}}}\left[1+\frac{J(\mu-l_0)^2}{2\left(\sigma_1^2+J\sigma_2^2\right)}\right]^{-2}\right\}+0=0$$

ii. $\pi_b(\underline{\theta})=\pi_b(\theta,\sigma_2^2,\sigma_1^2)\propto\sigma_1^{-2}\left(\sigma_1^2+J\sigma_2^2\right)^{-\frac{1}{2}}\left(1+\quad\frac{J(\theta-l_0)^2}{2\left(\sigma_1^2+J\sigma_2^2\right)}\right)^{\frac{1}{2}}$

will also be a probability matching prior since

$$\frac{\partial}{\partial\mu}\{\eta_1(\underline{\theta})\pi_b(\underline{\theta})\}$$

$$=\frac{\partial}{\partial\mu}\left\{\frac{\left(\sigma_1^2+J\sigma_2^2\right)^{\frac{1}{2}}}{3IJ^{\frac{1}{2}}}3I^{\frac{1}{2}}\left(1+\quad\frac{J(\mu-l_0)^2}{2\left(\sigma_1^2+J\sigma_2^2\right)}\right)^{-\frac{1}{2}}.\sigma_1^{-2}\left(\sigma_1^2+J\sigma_2^2\right)^{-\frac{1}{2}}\left(1+\quad\frac{J(\mu-l_0)^2}{2\left(\sigma_1^2+J\sigma_2^2\right)}\right)^{\frac{1}{2}}\right\}$$

$$=\frac{\partial}{\partial\mu}\left\{\frac{3I^{\frac{1}{2}}\sigma_1^{-2}}{3IJ^{\frac{1}{2}}}\right\}=0$$
.

Further

$$\frac{\partial}{\partial\sigma_2^2}\{\eta_2(\underline{\theta})\pi_b(\underline{\theta})\}$$

$$=\frac{\partial}{\partial\sigma_2^2}\left\{\frac{-(\mu-l_0)\left(\sigma_1^2+J\sigma_2^2\right)^{\frac{1}{2}}}{3IJ^{\frac{1}{2}}}3I^{\frac{1}{2}}\left(1+\quad\frac{J(\mu-l_0)^2}{2\left(\sigma_1^2+J\sigma_2^2\right)}\right)^{-\frac{1}{2}}.\sigma_1^{-2}\left(\sigma_1^2+J\sigma_2^2\right)^{-\frac{1}{2}}\left(1+\quad\frac{J(\mu-l_0)^2}{2\left(\sigma_1^2+J\sigma_2^2\right)}\right)^{\frac{1}{2}}\right\}$$

$$=\frac{\partial}{\partial\sigma_2^2}\left\{\frac{-(\mu-l_0)3I^{\frac{1}{2}}\sigma_1^{-2}}{3IJ^{\frac{1}{2}}}\right\}=0$$
.

Also

$$\frac{\partial}{\partial \sigma_1^{\,2}}\{\eta_3(\underline{\theta})\pi_b(\underline{\theta})\}$$

$$=\frac{\partial}{\partial \sigma_1^2}\left\{0\,.\,\sigma_1^{-2}\left(\sigma_1^{\,2}+J\sigma_2^{\,2}\right)^{-\frac{1}{2}}\left(1+\quad\frac{J(\mu-l_0)^2}{2\left(\sigma_1^{\,2}+J\sigma_2^{\,2}\right)}\right)^{\frac{1}{2}}\right\}\quad=\frac{\partial}{\partial \sigma_1^2}\{0\}=0$$

.

iii. . $\pi_c(\underline{\theta})=\pi_c(\mu,\sigma_2^{\,2},\sigma_1^{\,2})\propto\left(\sigma_1^{\,2}+J\sigma_2^{\,2}\right)^{-\frac{1}{2}}\left(1+\quad\frac{J(\mu-l_0)^2}{2\left(\sigma_1^{\,2}+J\sigma_2^{\,2}\right)}\right)^{\frac{1}{2}}$

will also be such a prior because

$$\frac{\partial}{\partial \mu}\{\eta_1(\underline{\theta})\pi_c(\underline{\theta})\}$$

$$=\frac{\partial}{\partial \mu}\left\{\frac{\left(\sigma_1^{\,2}+J\sigma_2^{\,2}\right)^{\frac{1}{2}}}{3IJ^{\frac{1}{2}}}3I^{\frac{1}{2}}\left(1+\quad\frac{J(\mu-l_0)^2}{2\left(\sigma_1^{\,2}+J\sigma_2^{\,2}\right)}\right)^{-\frac{1}{2}}.\left(\sigma_1^{\,2}+J\sigma_2^{\,2}\right)^{-\frac{1}{2}}\left(1+\quad\frac{J(\mu-l_0)^2}{2\left(\sigma_1^{\,2}+J\sigma_2^{\,2}\right)}\right)^{\frac{1}{2}}\right\}$$

$$=\frac{\partial}{\partial \mu}\left\{\frac{3I^{\frac{1}{2}}}{3IJ^{\frac{1}{2}}}\right\}=0$$

.

Further

$$\frac{\partial}{\partial \sigma_2^{\,2}}\{\eta_2(\underline{\theta})\pi_c(\underline{\theta})\}$$

$$=\frac{\partial}{\partial \sigma_2^{\,2}}\left\{\frac{-(\mu-l_0)\left(\sigma_1^{\,2}+J\sigma_2^{\,2}\right)^{\frac{1}{2}}}{3IJ^{\frac{1}{2}}}3I^{\frac{1}{2}}\left(1+\quad\frac{J(\mu-l_0)^2}{2\left(\sigma_1^{\,2}+J\sigma_2^{\,2}\right)}\right)^{-\frac{1}{2}}.\left(\sigma_1^{\,2}+J\sigma_2^{\,2}\right)^{-\frac{1}{2}}\left(1+\quad\frac{J(\mu-l_0)^2}{2\left(\sigma_1^{\,2}+J\sigma_2^{\,2}\right)}\right)^{\frac{1}{2}}\right\}$$

$$=\frac{\partial}{\partial \sigma_2^{\,2}}\left\{\frac{-(\mu-l_0)3I^{\frac{1}{2}}}{3IJ^{\frac{1}{2}}}\right\}=0$$

.

Also

$$\frac{\partial}{\partial \sigma_1^{\,2}}\{\eta_3(\underline{\theta})\pi_c(\underline{\theta})\}$$

$$=\frac{\partial}{\partial \sigma_1^2}\left\{0\,.\left(\sigma_1^{\,2}+J\sigma_2^{\,2}\right)^{-\frac{1}{2}}\left(1+\quad\frac{J(\mu-l_0)^2}{2\left(\sigma_1^{\,2}+J\sigma_2^{\,2}\right)}\right)^{\frac{1}{2}}\right\}\quad=\frac{\partial}{\partial \sigma_1^2}\{0\}=0$$

.

Proof of theorem 5.5.2

Posterior $\propto$ likelihood$\times$prior.

For the probability matching prior $\pi_a\left(\mu,\sigma_2^{\,2},\sigma_1^{\,2}\mid \underline{Y}\right)$, the posterior distribution is given by

$$p_a\left(\mu,\sigma_2^{\,2},\sigma_1^{\,2}\mid \underline{Y}\right)=C(\sigma_1^{\,2})^{-\frac{1}{2}\nu_1}(\sigma_1^{\,2}+J\sigma_2^{\,2})^{-\frac{1}{2}(\nu_2+1)}\exp(-\frac{1}{2}\{\frac{IJ(\mu-\overline{\overline{Y}}_{..})^2}{(\sigma_1^{\,2}+J\sigma_2^{\,2})}+\frac{\nu_1 m_1}{\sigma_1^{\,2}}+\frac{\nu_2 m_2}{(\sigma_1^{\,2}+J\sigma_2^{\,2})}\})\times$$

$$\sigma_1^{-2}\left(\sigma_1^{\,2}+J\sigma_2^{\,2}\right)^{-\frac{3}{2}}\left(1+\quad\frac{J(\mu-l_0)^2}{2\left(\sigma_1^{\,2}+J\sigma_2^{\,2}\right)}\right)^{-\frac{1}{2}}.$$

C is a normalizing constant .

Therefore

$$p_a\left(\mu,\sigma_2^{\,2},\sigma_1^{\,2}\mid \underline{Y}\right)\le C(\sigma_1^{\,2})^{-\frac{1}{2}(\nu_1+2)}(\sigma_1^{\,2}+J\sigma_2^{\,2})^{-\frac{1}{2}(\nu_2+4)}\exp(-\frac{1}{2}\{\frac{\nu_1 m_1}{\sigma_1^{\,2}}+\frac{\nu_2 m_2}{(\sigma_1^{\,2}+J\sigma_2^{\,2})}\}).$$

Further

$$C\int_{-\infty}^{\infty}(\sigma_1^{\,2})^{-\frac{1}{2}(\nu_1+2)}(\sigma_1^{\,2}+J\sigma_2^{\,2})^{-\frac{1}{2}(\nu_2+4)}\exp(-\frac{1}{2}\{\frac{IJ(\mu-\overline{\overline{Y}}_{..})^2}{(\sigma_1^{\,2}+J\sigma_2^{\,2})}+\frac{\nu_1 m_1}{\sigma_1^{\,2}}+\frac{\nu_2 m_2}{(\sigma_1^{\,2}+J\sigma_2^{\,2})}\})d\mu$$

$$=C(2\pi)^{\frac{1}{2}}\left\{\frac{1}{IJ}(\sigma_1^{\,2}+J\sigma_2^{\,2})\right\}^{\frac{1}{2}}(\sigma_1^{\,2})^{-\frac{1}{2}(\nu_1+2)}(\sigma_1^{\,2}+J\sigma_2^{\,2})^{-\frac{1}{2}(\nu_2+4)}\exp(-\frac{1}{2}\{\frac{\nu_1 m_1}{\sigma_1^{\,2}}+\frac{\nu_2 m_2}{(\sigma_1^{\,2}+J\sigma_2^{\,2})}\})$$

and

$$C\left(\frac{2\pi}{IJ}\right)^{\frac{1}{2}}\int_0^{\infty}\int_0^{\infty}(\sigma_1^{\,2})^{-\frac{1}{2}(\nu_1+2)}(\sigma_1^{\,2}+J\sigma_2^{\,2})^{-\frac{1}{2}(\nu_2+3)}\exp(-\frac{1}{2}\{\frac{\nu_1 m_1}{\sigma_1^{\,2}}+\frac{\nu_2 m_2}{(\sigma_1^{\,2}+J\sigma_2^{\,2})}\})d\sigma_1^{\,2}d\sigma_2^{\,2}$$

$$=C\left(\frac{2\pi}{IJ}\right)^{\frac{1}{2}}\frac{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2+1}{2}\right)\Pr\left\{F_{\nu_2+1;\nu_1}<\frac{m_2}{m_1}\right\}\quad 2^{\frac{1}{2}(\nu_1+\nu_2+1)}}{J(\nu_1 m_1)^{\frac{1}{2}\nu_1}(\nu_2 m_2)^{\frac{1}{2}(\nu_2+1)}}$$

where $F_{\nu_2+1;\nu_1}$ is the F-distribution with ν_2+1 and ν_1 degrees of freedom.

Proof of Theorem 5.5.3

The Fisher information matrix is then given by

$$F(\mu,\sigma_2^2,\sigma_1^2)=\begin{pmatrix} \frac{IJ}{(\sigma_1^2+J\sigma_2^2)} & 0 & 0 \\ 0 & \frac{IJ^2}{2(\sigma_1^2+J\sigma_2^2)^2} & \frac{IJ}{2(\sigma_1^2+J\sigma_2^2)^2} \\ 0 & \frac{IJ}{2(\sigma_1^2+J\sigma_2^2)^2} & \frac{I(J-1)}{2(\sigma_1^2)^2}+\frac{I}{2(\sigma_1^2+J\sigma_2^2)^2} \end{pmatrix}$$

The interest lies in the Fisher information matrix for $t(\underline{\theta}),\nu$ and σ_1^2. This will be obtained in two stages. Substituting $\nu=\frac{\sigma_2^2}{\sigma_1^2}$, the Fisher information matrix is given by

$$F(\mu,\sigma_2^2,\sigma_1^2)=\begin{pmatrix} \frac{IJ}{\sigma_1^2(1+J\nu)} & 0 & 0 \\ 0 & \frac{IJ^2}{2(\sigma_1^2)^2(1+J\nu)^2} & \frac{IJ}{2(\sigma_1^2)^2(1+J\nu)^2} \\ 0 & \frac{IJ}{2(\sigma_1^2)^2(1+J\nu)^2} & \frac{I(J-1)}{2(\sigma_1^2)^2}+\frac{I}{2(\sigma_1^2)^2(1+J\nu)^2} \end{pmatrix}$$

We are interested in the Fisher Information matrix for $t(\underline{\theta}),\quad \nu,\quad \sigma_1^2$.

We will first derive the Fisher information matrix for μ,ν,σ_1^2

Let $A=\frac{\partial(\mu,\sigma_2^2,\sigma_1^2)}{\partial(\mu,\nu,\sigma_1^2)}=\begin{pmatrix} \frac{\partial\mu}{\partial\mu} & \frac{\partial\mu}{\partial\nu} & \frac{\partial\mu}{\partial\sigma_1^2} \\ \frac{\partial\sigma_2^2}{\partial\mu} & \frac{\partial\sigma_2^2}{\partial\nu} & \frac{\partial\sigma_2^2}{\partial\sigma_1^2} \\ \frac{\partial\sigma_1^2}{\partial\mu} & \frac{\partial\sigma_1^2}{\partial\nu} & \frac{\partial\sigma_1^2}{\partial\sigma_1^2} \end{pmatrix}=\begin{pmatrix} 1 & 0 & 0 \\ 0 & \sigma_1^2 & \nu \\ 0 & 0 & 1 \end{pmatrix}$

$$\nu=\frac{\sigma_2^2}{\sigma_1^2} \qquad\qquad \therefore \sigma_2^2=\nu\sigma_1^2$$

$$A' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sigma_1^2 & 0 \\ 0 & \nu & 1 \end{pmatrix}$$

Now $F(\mu,\nu,\sigma_1^2) = A'F(\mu,\sigma_2^2,\sigma_1^2)A$

$A'F(\mu,\sigma_1^2,\sigma_2^2)A =$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \sigma_1^2 & 0 \\ 0 & \nu & 1 \end{pmatrix} \begin{pmatrix} \frac{IJ}{\sigma_1^2(1+J\nu)} & 0 & 0 \\ 0 & \frac{IJ^2}{2(\sigma_1^2)^2(1+J\nu)^2} & \frac{IJ}{2(\sigma_1^2)^2(1+J\nu)^2} \\ 0 & \frac{IJ}{2(\sigma_1^2)^2(1+J\nu)^2} & \frac{I(J-1)}{2(\sigma_1^2)^2} + \frac{I}{2(\sigma_1^2)^2(1+J\nu)^2} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sigma_1^2 & \nu \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{IJ}{\sigma_1^2(1+J\nu)} & 0 & 0 \\ 0 & \frac{IJ^2}{2(\sigma_1^2)(1+J\nu)^2} & \frac{IJ}{2(\sigma_1^2)(1+J\nu)^2} \\ 0 & \frac{\nu IJ^2 + IJ}{2(\sigma_1^2)^2(1+J\nu)^2} & \frac{I(J-1)}{2(\sigma_1^2)^2} + \frac{I+\nu IJ}{2(\sigma_1^2)^2(1+J\nu)^2} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sigma_1^2 & \nu \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{IJ}{\sigma_1^2(1+J\nu)} & 0 & 0 \\ 0 & \frac{IJ^2}{2(\sigma_1^2)(1+J\nu)^2} & \frac{IJ}{2(\sigma_1^2)(1+J\nu)^2} \\ 0 & \frac{IJ(1+J\nu)}{2(\sigma_1^2)^2(1+J\nu)^2} & \frac{I(J-1)}{2(\sigma_1^2)^2} + \frac{I(1+J\nu)}{2(\sigma_1^2)^2(1+J\nu)^2} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sigma_1^2 & \nu \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{IJ}{\sigma_1^2(1+J\nu)} & 0 & 0 \\ 0 & \frac{IJ^2}{2(\sigma_1^2)(1+J\nu)^2} & \frac{IJ}{2(\sigma_1^2)(1+J\nu)^2} \\ 0 & \frac{IJ}{2(\sigma_1^2)^2(1+J\nu)} & \frac{I+I(J-1)(1+J\nu)}{2(\sigma_1^2)^2(1+J\nu)} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sigma_1^2 & \nu \\ 0 & 0 & 1 \end{pmatrix}$$

$$
=\begin{pmatrix} \frac{IJ}{\sigma_1^2(1+J\nu)} & 0 & 0 \\ 0 & \frac{IJ^2}{2(\sigma_1^2)(1+J\nu)^2} & \frac{IJ}{2(\sigma_1^2)(1+J\nu)^2} \\ 0 & \frac{IJ}{2(\sigma_1^2)^2(1+J\nu)} & \frac{I\left[1+(J-1)(1+J\nu)\right]}{2(\sigma_1^2)^2(1+J\nu)} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sigma_1^2 & \nu \\ 0 & 0 & 1 \end{pmatrix}
$$

Therefore $F(\mu,\nu,\sigma_1^2) =$

$$
=\begin{pmatrix} \frac{IJ}{\sigma_1^2(1+J\nu)} & 0 & 0 \\ 0 & \frac{(\sigma_1^2)IJ^2}{2(\sigma_1^2)(1+J\nu)^2} & \frac{\nu IJ^2+IJ}{2(\sigma_1^2)(1+J\nu)^2} \\ 0 & \frac{(\sigma_1^2)IJ}{2(\sigma_1^2)^2(1+J\nu)} & \frac{\nu IJ}{2(\sigma_1^2)^2(1+J\nu)}+\frac{I\left[1+(J-1)(1+J\nu)\right]}{2(\sigma_1^2)^2(1+J\nu)} \end{pmatrix}
$$

$$
=\begin{pmatrix} \frac{IJ}{\sigma_1^2(1+J\nu)} & 0 & 0 \\ 0 & \frac{IJ^2}{2(1+J\nu)^2} & \frac{IJ}{2(\sigma_1^2)(1+J\nu)} \\ 0 & \frac{IJ}{2(\sigma_1^2)(1+J\nu)} & \frac{\nu IJ}{2(\sigma_1^2)^2(1+J\nu)}+\frac{I\left[1+(J+J^2\nu-1-J\nu)\right]}{2(\sigma_1^2)^2(1+J\nu)} \end{pmatrix}
$$

$$
=\begin{pmatrix} \frac{IJ}{\sigma_1^2(1+J\nu)} & 0 & 0 \\ 0 & \frac{IJ^2}{2(1+J\nu)^2} & \frac{IJ}{2(\sigma_1^2)(1+J\nu)} \\ 0 & \frac{IJ}{2(\sigma_1^2)(1+J\nu)} & \frac{IJ\left[1+J\nu\right]}{2(\sigma_1^2)^2(1+J\nu)} \end{pmatrix}
$$

$$=\begin{pmatrix} \frac{IJ}{\sigma_1^2(1+J\nu)} & 0 & 0 \\ 0 & \frac{IJ^2}{2(1+J\nu)^2} & \frac{IJ}{2(\sigma_1^2)(1+J\nu)} \\ 0 & \frac{IJ}{2(\sigma_1^2)(1+J\nu)} & \frac{IJ}{2(\sigma_1^2)^2} \end{pmatrix}.$$

At this second stage the Fisher information matrix for μ, ν and $t(\underline{\theta})$ will be derived.

Now

$$t(\underline{\theta}) = \frac{\mu - l_0}{3\left(\frac{\sigma_1^2 + J\sigma_2^2}{J}\right)^{\frac{1}{2}}} = \frac{\mu - l_0}{3\left(\frac{\sigma_1^2(1+J\nu)}{J}\right)^{\frac{1}{2}}}$$

$$\therefore \mu = 3\left(\frac{\sigma_1^2(1+J\nu)}{J}\right)^{\frac{1}{2}} t(\underline{\theta}) + l_0$$

$$\therefore \mu = 3\frac{(\sigma_1^2)^{\frac{1}{2}}}{J^{\frac{1}{2}}}(1+J\nu)^{\frac{1}{2}} t(\underline{\theta}) + l_0$$

$$\frac{\partial\mu}{\partial t(\underline{\theta})} = 3\frac{(\sigma_1^2)^{\frac{1}{2}}}{J^{\frac{1}{2}}}(1+J\nu)^{\frac{1}{2}},$$

$$\frac{\partial\mu}{\partial\nu} = \frac{3}{2}\frac{J(\sigma_1^2)^{\frac{1}{2}}}{J^{\frac{1}{2}}}(1+J\nu)^{-\frac{1}{2}} t(\underline{\theta})$$

$$= \frac{3}{2}J^{\frac{1}{2}}(\sigma_1^2)^{\frac{1}{2}}(1+J\nu)^{-\frac{1}{2}} t(\underline{\theta}),$$

$$\frac{\partial\mu}{\partial\sigma_1^2} = \frac{3}{2}\frac{(\sigma_1^2)^{-\frac{1}{2}}}{J^{\frac{1}{2}}}(1+J\nu)^{\frac{1}{2}} t(\underline{\theta}).$$

$$\tilde{A} = \begin{pmatrix} \frac{\partial\mu}{\partial t(\underline{\theta})} & \frac{\partial\mu}{\partial\nu} & \frac{\partial\mu}{\partial\sigma_1^2} \\ \frac{\partial\nu}{\partial t(\underline{\theta})} & \frac{\partial\nu}{\partial\nu} & \frac{\partial\nu}{\partial\sigma_1^2} \\ \frac{\partial\sigma_1^2}{\partial t(\underline{\theta})} & \frac{\partial\sigma_1^2}{\partial\nu} & \frac{\partial\sigma_1^2}{\partial\sigma_1^2} \end{pmatrix} = \begin{pmatrix} 3\frac{(\sigma_1^2)^{\frac{1}{2}}}{J^{\frac{1}{2}}}(1+J\nu)^{\frac{1}{2}} & \frac{3}{2}J^{\frac{1}{2}}(\sigma_1^2)^{\frac{1}{2}}(1+J\nu)^{-\frac{1}{2}} t(\underline{\theta}) & \frac{3}{2}\frac{(\sigma_1^2)^{-\frac{1}{2}}}{J^{\frac{1}{2}}}(1+J\nu)^{\frac{1}{2}} t(\underline{\theta}) \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The Fisher information matrix for $t(\underline{\theta}), \nu$ and σ_1^2 is given by

$F(t(\underline{\theta}),\nu,\sigma_1^{\,2}) = \widetilde{A}' F(\mu,\nu,\sigma_1^{\,2})\widetilde{A}$ where

$\widetilde{A}' F(\mu,\nu,\sigma_1^{\,2}) =$

$$= \begin{pmatrix} 3\dfrac{(\sigma_1^2)^{\frac{1}{2}}}{J^{\frac{1}{2}}}(1+J\nu)^{\frac{1}{2}} & 0 & 0 \\ \dfrac{3}{2}J^{\frac{1}{2}}(\sigma_1^2)^{\frac{1}{2}}(1+J\nu)^{-\frac{1}{2}}t(\underline{\theta}) & 1 & 0 \\ \dfrac{3}{2}\dfrac{(\sigma_1^2)^{-\frac{1}{2}}}{J^{\frac{1}{2}}}(1+J\nu)^{\frac{1}{2}}t(\underline{\theta}) & 0 & 1 \end{pmatrix} \begin{pmatrix} \dfrac{IJ}{\sigma_1^2(1+J\nu)} & 0 & 0 \\ 0 & \dfrac{IJ^2}{2(1+J\nu)^2} & \dfrac{IJ}{2(\sigma_1^2)(1+J\nu)} \\ 0 & \dfrac{IJ}{2(\sigma_1^2)(1+J\nu)} & \dfrac{IJ}{2(\sigma_1^2)^2} \end{pmatrix}$$

$$= \begin{pmatrix} 3\dfrac{IJ^{\frac{1}{2}}}{(\sigma_1^2)^{\frac{1}{2}}(1+J\nu)^{\frac{1}{2}}} & 0 & 0 \\ \dfrac{3}{2}\dfrac{IJ^{\frac{3}{2}}}{(\sigma_1^2)^{\frac{1}{2}}(1+J\nu)^{\frac{3}{2}}}t(\underline{\theta}) & \dfrac{IJ^2}{2(1+J\nu)^2} & \dfrac{IJ}{2(\sigma_1^2)(1+J\nu)} \\ \dfrac{3}{2}\dfrac{IJ^{\frac{1}{2}}}{(\sigma_1^2)^{\frac{3}{2}}(1+J\nu)^{\frac{1}{2}}}t(\underline{\theta}) & \dfrac{IJ}{2(\sigma_1^2)(1+J\nu)} & \dfrac{IJ}{2(\sigma_1^2)^2} \end{pmatrix}.$$

$\widetilde{A}' F(\mu,\nu,\sigma_1^{\,2})\widetilde{A} =$

$$\begin{pmatrix} 3\dfrac{IJ^{\frac{1}{2}}}{(\sigma_1^2)^{\frac{1}{2}}(1+J\nu)^{\frac{1}{2}}} & 0 & 0 \\ \dfrac{3}{2}\dfrac{IJ^{\frac{3}{2}}}{(\sigma_1^2)^{\frac{1}{2}}(1+J\nu)^{\frac{3}{2}}}t(\underline{\theta}) & \dfrac{IJ^2}{2(1+J\nu)^2} & \dfrac{IJ}{2(\sigma_1^2)(1+J\nu)} \\ \dfrac{3}{2}\dfrac{IJ^{\frac{1}{2}}}{(\sigma_1^2)^{\frac{3}{2}}(1+J\nu)^{\frac{1}{2}}}t(\underline{\theta}) & \dfrac{IJ}{2(\sigma_1^2)(1+J\nu)} & \dfrac{IJ}{2(\sigma_1^2)^2} \end{pmatrix} \begin{pmatrix} 3\dfrac{(\sigma_1^2)^{\frac{1}{2}}}{J^{\frac{1}{2}}}(1+J\nu)^{\frac{1}{2}} & \dfrac{3}{2}J^{\frac{1}{2}}(\sigma_1^2)^{\frac{1}{2}}(1+J\nu)^{-\frac{1}{2}}t(\underline{\theta}) & \dfrac{3}{2}\dfrac{(\sigma_1^2)^{-\frac{1}{2}}}{J^{\frac{1}{2}}}(1+J\nu)^{\frac{1}{2}}t(\underline{\theta}) \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 9I & \dfrac{9}{2}\dfrac{IJ}{(1+J\nu)}t(\underline{\theta}) & \dfrac{9}{2}\dfrac{I}{(\sigma_1^2)}t(\underline{\theta}) \\ \dfrac{9}{2}\dfrac{IJ}{(1+J\nu)}t(\underline{\theta}) & \dfrac{9}{4}\dfrac{IJ^2}{(1+J\nu)^2}t^2(\underline{\theta})+\dfrac{IJ^2}{2(1+J\nu)^2} & \dfrac{9}{4}\dfrac{IJ}{(\sigma_1^2)(1+J\nu)}t^2(\underline{\theta})+\dfrac{IJ}{2(\sigma_1^2)(1+J\nu)} \\ \dfrac{9}{2}\dfrac{I}{(\sigma_1^2)}t(\underline{\theta}) & \dfrac{9}{4}\dfrac{IJ}{(\sigma_1^2)(1+J\nu)}t^2(\underline{\theta})+\dfrac{IJ}{2(\sigma_1^2)(1+J\nu)} & \dfrac{9}{4}\dfrac{I}{(\sigma_1^2)^2}t^2(\underline{\theta})+\dfrac{IJ}{2(\sigma_1^2)^2} \end{pmatrix}$$

$$
=\begin{pmatrix} 9I & \frac{9}{2}\frac{IJ}{(1+J\nu)}t(\underline{\theta}) & \frac{9}{2}\frac{I}{(\sigma_1^2)}t(\underline{\theta}) \\ \frac{9}{2}\frac{IJ}{(1+J\nu)}t(\underline{\theta}) & \frac{IJ^2\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(1+J\nu)^2} & \frac{IJ\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^2)(1+J\nu)} \\ \frac{9}{2}\frac{I}{(\sigma_1^2)}t(\underline{\theta}) & \frac{IJ\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^2)(1+J\nu)} & \frac{I\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^2)^2}+\frac{I(J-1)}{2(\sigma_1^2)^2} \end{pmatrix}
$$

$$
\widetilde{A}'F(\mu,\nu,\sigma_1^2)\widetilde{A}=\begin{pmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{pmatrix}=\begin{pmatrix} 9I & \frac{9}{2}\frac{IJ}{(1+J\nu)}t(\underline{\theta}) & \frac{9}{2}\frac{I}{(\sigma_1^2)}t(\underline{\theta}) \\ \frac{9}{2}\frac{IJ}{(1+J\nu)}t(\underline{\theta}) & \frac{IJ^2\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(1+J\nu)^2} & \frac{IJ\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^2)(1+J\nu)} \\ \frac{9}{2}\frac{I}{(\sigma_1^2)}t(\underline{\theta}) & \frac{IJ\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^2)(1+J\nu)} & \frac{I\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^2)^2}+\frac{I(J-1)}{2(\sigma_1^2)^2} \end{pmatrix}.
$$

As mentioned earlier the reference prior depends on the group ordering of the parameters and it is determined through a succession of analysis for the implied conditional problems. Berger and Bernado (1992) particularly recommended the reference prior based on having each parameter in its own group, i.e. having each conditional reference prior being one dimensional

Let

$$
F(t(\underline{\theta}),\nu,\sigma_1^2)=\widetilde{A}'F(\mu,\nu,\sigma_1^2)\widetilde{A}=\begin{pmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{pmatrix}
$$

Therefore consider the sub-matrix $\widetilde{F}=\begin{pmatrix} f_{22} & f_{23} \\ f_{32} & f_{33} \end{pmatrix}$ and its inverse

Therefore $\widetilde{F}=\begin{pmatrix} \frac{IJ^2\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(1+J\nu)^2} & \frac{IJ\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^2)(1+J\nu)} \\ \frac{IJ\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^2)(1+J\nu)} & \frac{I\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^2)^2}+\frac{I(J-1)}{2(\sigma_1^2)^2} \end{pmatrix}$

and

$$|\widetilde{F}| = \left[\frac{IJ^2\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(1+J\nu)^2}\left(\frac{I\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^{\,2})^2}+\frac{I(J-1)}{2(\sigma_1^{\,2})^2}\right)-\frac{IJ\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^{\,2})(1+J\nu)}\frac{IJ\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^{\,2})(1+J\nu)}\right]$$

$$|\widetilde{F}| = \left(\frac{I^2J^2\left(\frac{9}{2}t^2(\underline{\theta})+1\right)^2}{4(\sigma_1^{\,2})^2(1+J\nu)^2}+\frac{I^2J^2\left(\frac{9}{2}t^2(\underline{\theta})+1\right)(J-1)}{4(\sigma_1^{\,2})^2(1+J\nu)^2}\right)-\frac{I^2J^2\left(\frac{9}{2}t^2(\underline{\theta})+1\right)^2}{4(\sigma_1^{\,2})^2(1+J\nu)^2}$$

$$|\widetilde{F}| = \frac{I^2J^2\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{4(\sigma_1^{\,2})^2(1+J\nu)^2}\left\{\left(\frac{\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{1}+\frac{(J-1)}{1}\right)-\frac{\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{1}\right\}$$

$$|\widetilde{F}| = \frac{I^2J^2\left(\frac{9}{2}t^2(\underline{\theta})+1\right)(J-1)}{4(\sigma_1^{\,2})^2(1+J\nu)^2}.$$

$$\widetilde{F}^{-1} = \frac{1}{|\widetilde{F}|}\begin{pmatrix} \frac{I\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^{\,2})^2}+\frac{I(J-1)}{2(\sigma_1^{\,2})^2} & -\frac{IJ\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^{\,2})(1+J\nu)} \\ -\frac{IJ\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^{\,2})(1+J\nu)} & \frac{IJ^2\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(1+J\nu)^2} \end{pmatrix}$$

$$\widetilde{F}^{-1} = \frac{4(\sigma_1^{\,2})^2(1+J\nu)^2}{I^2J^2\left(\frac{9}{2}t^2(\underline{\theta})+1\right)(J-1)}\begin{pmatrix} \frac{I\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^{\,2})^2}+\frac{I(J-1)}{2(\sigma_1^{\,2})^2} & -\frac{IJ\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^{\,2})(1+J\nu)} \\ -\frac{IJ\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^{\,2})(1+J\nu)} & \frac{IJ^2\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(1+J\nu)^2} \end{pmatrix}$$

$$\tilde{F}^{-1} = \begin{pmatrix} \frac{2(1+J\nu)^2}{IJ^2(J-1)} + \frac{2(1+J\nu)^2}{IJ^2\left(\frac{9}{2}t^2(\underline{\theta})+1\right)} & -\frac{2(\sigma_1^2)(1+J\nu)}{IJ(J-1)} \\ -\frac{2(\sigma_1^2)(1+J\nu)}{IJ(J-1)} & \frac{2(\sigma_1^2)^2}{I(J-1)} \end{pmatrix}.$$

The reference prior in the $(t(\underline{\theta}), \nu, \sigma_1^2)$ parameterisation

Let $\tilde{F}^{-1} = \begin{pmatrix} f_{22} & f_{23} \\ f_{32} & f_{33} \end{pmatrix}^{-1}$ from the section above

$$h_1 = f_{11.2} = f_{11} - \begin{bmatrix} f_{12} & f_{13} \end{bmatrix} \tilde{F}^{-1} \begin{bmatrix} f_{21} \\ f_{31} \end{bmatrix}$$

$$h_1 = 9I - \begin{bmatrix} \frac{9}{2}\frac{IJ}{(1+J\nu)}t(\underline{\theta}) & \frac{9}{2}\frac{I}{(\sigma_1^2)}t(\underline{\theta}) \end{bmatrix} \begin{pmatrix} \frac{2(1+J\nu)^2}{IJ^2(J-1)} + \frac{2(1+J\nu)^2}{IJ^2\left(\frac{9}{2}t^2(\underline{\theta})+1\right)} & -\frac{2(\sigma_1^2)(1+J\nu)}{IJ(J-1)} \\ -\frac{2(\sigma_1^2)(1+J\nu)}{IJ(J-1)} & \frac{2(\sigma_1^2)^2}{I(J-1)} \end{pmatrix}^{-1} \begin{bmatrix} \frac{9}{2}\frac{IJ}{(1+J\nu)}t(\underline{\theta}) \\ \frac{9}{2}\frac{I}{(\sigma_1^2)}t(\underline{\theta}) \end{bmatrix}$$

$$h_1 = 9I - \left(\frac{9}{2}\right)^2 I^2 t^2(\underline{\theta}) \begin{bmatrix} \frac{J}{(1+J\nu)} & \frac{1}{(\sigma_1^2)} \end{bmatrix} \begin{pmatrix} \frac{2(1+J\nu)^2}{IJ^2(J-1)} + \frac{2(1+J\nu)^2}{IJ^2\left(\frac{9}{2}t^2(\underline{\theta})+1\right)} & -\frac{2(\sigma_1^2)(1+J\nu)}{IJ(J-1)} \\ -\frac{2(\sigma_1^2)(1+J\nu)}{IJ(J-1)} & \frac{2(\sigma_1^2)^2}{I(J-1)} \end{pmatrix} \begin{bmatrix} \frac{J}{(1+J\nu)} \\ \frac{1}{(\sigma_1^2)} \end{bmatrix}$$

$$h_1 = 9I - \left(\frac{9}{2}\right)^2 I^2 t^2(\underline{\theta}) \begin{bmatrix} \frac{2(1+J\nu)}{IJ(J-1)} + \frac{2(1+J\nu)}{IJ\left(\frac{9}{2}t^2(\underline{\theta})+1\right)} - \frac{2(1+J\nu)}{IJ(J-1)} & -\frac{2(\sigma_1^2)}{I(J-1)} + \frac{2(\sigma_1^2)}{I(J-1)} \end{bmatrix} \begin{bmatrix} \frac{J}{(1+J\nu)} \\ \frac{1}{(\sigma_1^2)} \end{bmatrix}$$

$$h_1 = 9I - \left(\frac{9}{2}\right)^2 I^2 t^2(\underline{\theta}) \begin{bmatrix} \frac{2(1+J\nu)}{IJ\left(\frac{9}{2}t^2(\underline{\theta})+1\right)} & 0 \end{bmatrix} \begin{bmatrix} \frac{J}{(1+J\nu)} \\ \frac{1}{(\sigma_1^2)} \end{bmatrix}$$

$$h_1 = 9I - \frac{2\left(\frac{9}{2}\right)^2 It^2(\underline{\theta})}{\left(\frac{9}{2}t^2(\underline{\theta})+1\right)} = \frac{9I\left(\frac{9}{2}t^2(\underline{\theta})+1\right) - 2\left(\frac{9}{2}\right)^2 It^2(\underline{\theta})}{\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}$$

$$h_1 = \frac{\left(\frac{81}{2} It^2(\underline{\theta}) + 9I\right) - \frac{81}{2} It^2(\underline{\theta})}{\left(\frac{9t^2(\underline{\theta}) + 2}{2}\right)} = 18I\left(9t^2(\underline{\theta}) + 2\right)^{-1}$$
.

Further

$$h_2 = f_{22} - \frac{1}{f_{33}} f_{23} f_{32}$$

$$h_2 = \frac{IJ^2\left(\frac{9}{2}t^2(\underline{\theta}) + 1\right)}{2(1+J\nu)^2} - \frac{2(\sigma_1^2)^2}{I\left(\frac{9}{2}t^2(\underline{\theta}) + 1\right) + I(J-1)} \left(\frac{IJ\left(\frac{9}{2}t^2(\underline{\theta}) + 1\right)}{2(\sigma_1^2)(1+J\nu)}\right)^2$$

$$h_2 = \frac{IJ^2\left(\frac{9}{2}t^2(\underline{\theta}) + 1\right)}{2(1+J\nu)^2} - \frac{2}{\left(\frac{9}{2}It^2(\underline{\theta}) + I + IJ - I\right)} \left(\frac{IJ\left(\frac{9}{2}t^2(\underline{\theta}) + 1\right)}{2(1+J\nu)}\right)^2$$

$$h_2 = \frac{IJ^2\left(\frac{9}{2}t^2(\underline{\theta}) + 1\right)}{2(1+J\nu)^2} - \frac{4}{\left(9It^2(\underline{\theta}) + 2IJ\right)} \left(\frac{IJ\left(\frac{9}{2}t^2(\underline{\theta}) + 1\right)}{2(1+J\nu)}\right)^2$$

$$h_2 = \frac{1}{(1+J\nu)^2} \left\{ \frac{IJ^2\left(\frac{9}{2}t^2(\underline{\theta}) + 1\right)}{2} - \frac{I^2J^2\left(\frac{9}{2}t^2(\underline{\theta}) + 1\right)^2}{\left(9It^2(\underline{\theta}) + 2IJ\right)} \right\}$$

$$h_2 = (1+J\nu)^{-2} \left\{ \frac{IJ^2\left(\frac{9}{2}t^2(\underline{\theta}) + 1\right)}{2} - \frac{IJ^2\left(\frac{9}{2}t^2(\underline{\theta}) + 1\right)^2}{\left(9t^2(\underline{\theta}) + 2J\right)} \right\}$$

$$h_2 = \frac{IJ^2\left(\frac{9}{2}t^2(\underline{\theta}) + 1\right)}{2(1+J\nu)^2} \left\{ 1 - \frac{\frac{9}{2}t^2(\underline{\theta}) + 1}{\frac{9}{2}t^2(\underline{\theta}) + J} \right\} \propto (1+J\nu)^{-2}$$

$$h_2 \propto (1+J\nu)^{-2}$$

$$h_3 = f_{33} = \frac{I\left(\frac{9}{2}t^2(\underline{\theta}) + 1\right)}{2(\sigma_1^2)^2} + \frac{I(J-1)}{2(\sigma_1^2)^2} = \frac{\left(\frac{9}{2}It^2(\underline{\theta}) + I + IJ - I\right)}{2(\sigma_1^2)^2} = \frac{\left(\frac{9}{2}It^2(\underline{\theta}) + IJ\right)}{2(\sigma_1^2)^2} \propto (\sigma_1^2)^{-2}$$
.

From the above it follows that

$$p(t(\underline{\theta})) \propto h_1^{\frac{1}{2}} = \left(9t^2(\underline{\theta})+2\right)^{-\frac{1}{2}}$$

$$p(\nu \,|\, t(\underline{\theta})) \propto h_2^{\frac{1}{2}} = \left(1+J\nu\right)^{-1}$$

$$p(\sigma_1^{\,2} \,|\, t(\underline{\theta}),\nu) \propto h_3^{\frac{1}{2}} = \sigma_1^{-2}.$$

Therefore the reference prior relative to the ordered parameterisation $(t(\underline{\theta}),\nu,\sigma_1^{\,2})$ is given by

$$p(t(\underline{\theta}),\nu,\sigma_1^{\,2}) = p(t(\underline{\theta}))\,p(\nu \,|\, t(\underline{\theta})\,p(\sigma_1^{\,2} \,|\, t(\underline{\theta}),\nu).$$

$$p(t(\underline{\theta}),\nu,\sigma_1^{\,2}) \propto \left(9t^2(\underline{\theta})+2\right)^{-\frac{1}{2}}\left(1+J\nu\right)^{-1}\sigma_1^{-2}.$$

The reference prior in the $(\mu,\sigma_2^{\,2},\sigma_1^{\,2})$ parameterisation

As defined

$$t\left(\underline{\theta}\right) = \frac{\mu - l_0}{3\left(\dfrac{\sigma_1^{\,2}+J\sigma_2^{\,2}}{J}\right)^{\frac{1}{2}}},$$

let

$$x_1 = t\left(\underline{\theta}\right) \quad y_1 = \mu$$

then

$$y_1 = 3x_1\left(\frac{\sigma_1^{\,2}+J\sigma_2^{\,2}}{J}\right)^{\frac{1}{2}} + l_0 \text{ and } x_1 = \frac{y_1 - l_0}{3\left(\dfrac{\sigma_1^{\,2}+J\sigma_2^{\,2}}{J}\right)^{\frac{1}{2}}},$$

$$\frac{\partial x_1}{\partial y_1} = \frac{1}{3}\left(\frac{\sigma_1^{\,2}+J\sigma_2^{\,2}}{J}\right)^{-\frac{1}{2}}.$$

For

$$\nu = \frac{\sigma_2^{\,2}}{\sigma_1^{\,2}},$$

let $x_2 = \nu \quad y_2 = \sigma_2{}^2$

$y_2 = x_2\sigma_1{}^2$ and $x_2 = \dfrac{y_2}{\sigma_1{}^2}$

$$\frac{\partial x_2}{\partial y_2} = \frac{1}{\sigma_1{}^2}$$

i.e. $\dfrac{\partial t(\underline{\theta})}{\partial \mu} = \dfrac{1}{3}\left(\dfrac{\sigma_1{}^2 + J\sigma_2{}^2}{J}\right)^{-\frac{1}{2}}$, $\nu = \dfrac{\sigma_2{}^2}{\sigma_1{}^2}, \dfrac{\partial \nu}{\partial \sigma_2{}^2} = \dfrac{1}{\sigma_1{}^2}$

$$p_{R_1}(y_1, y_2, \sigma_1{}^2) \propto p(x_1, x_2, \sigma_1{}^2)\left|\frac{dx_1}{dy_1}\right|\left|\frac{dx_2}{dy_2}\right|$$

$$p_{R_1}(\mu, \sigma_2{}^2, \sigma_1{}^2) \propto \left(\frac{J(\mu - l_0)^2}{2\left(\sigma_1{}^2 + J\sigma_2{}^2\right)} + 1\right)^{-\frac{1}{2}}\left(1 + \frac{J\sigma_2{}^2}{\sigma_1{}^2}\right)^{-1}\sigma_1{}^{-2}\,\frac{J^{\frac{1}{2}}}{3}\left(\sigma_1{}^2 + J\sigma_2{}^2\right)^{-\frac{1}{2}}\sigma_1{}^{-2}$$

$$p_{R_1}(\mu, \sigma_2{}^2, \sigma_1{}^2) \propto \left(\frac{J(\mu - l_0)^2}{2\left(\sigma_1{}^2 + J\sigma_2{}^2\right)} + 1\right)^{-\frac{1}{2}}\left(\sigma_1{}^2 + J\sigma_2{}^2\right)^{-\frac{3}{2}}\sigma_1{}^2\sigma_1{}^{-2}\sigma_1{}^{-2}$$

$$p_{R_1}(\mu, \sigma_2{}^2, \sigma_1{}^2) \propto \left(\frac{J(\mu - l_0)^2}{2\left(\sigma_1{}^2 + J\sigma_2{}^2\right)} + 1\right)^{-\frac{1}{2}}\left(\sigma_1{}^2 + J\sigma_2{}^2\right)^{-\frac{3}{2}}\sigma_1{}^{-2}$$

which corresponds to the Probability Matching Prior $\pi_a(\underline{\theta})$.

Proof of corollary 5.5.3.1

The reference prior in the group ordering $(t(\underline{\theta}), \sigma_1{}^2, \nu)$

The Fisher information matrix $F(t(\underline{\theta}), \sigma_1{}^2, \nu) =$

$$
\begin{pmatrix}
9I & \frac{9}{2}\frac{I}{(\sigma_1^2)}t(\underline{\theta}) & \frac{9}{2}\frac{IJ}{(1+J\nu)}t(\underline{\theta}) \\
\frac{9}{2}\frac{I}{(\sigma_1^2)}t(\underline{\theta}) & \frac{I\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^2)^2}+\frac{I(J-1)}{2(\sigma_1^2)^2} & \frac{IJ\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^2)(1+J\nu)} \\
\frac{9}{2}\frac{IJ}{(1+J\nu)}t(\underline{\theta}) & \frac{IJ\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^2)(1+J\nu)} & \frac{IJ^2\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(1+J\nu)^2}
\end{pmatrix}
=
\begin{pmatrix}
f_{11} & f_{12} & f_{13} \\
f_{21} & f_{22} & f_{23} \\
f_{31} & f_{32} & f_{33}
\end{pmatrix}
$$

Now consider the sub-matrix

$$
\widetilde{F}=\begin{pmatrix} f_{22} & f_{23} \\ f_{32} & f_{33} \end{pmatrix}=
\begin{pmatrix}
\frac{I\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^2)^2}+\frac{I(J-1)}{2(\sigma_1^2)^2} & \frac{IJ\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^2)(1+J\nu)} \\
\frac{IJ\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^2)(1+J\nu)} & \frac{IJ^2\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(1+J\nu)^2}
\end{pmatrix}
$$

and

$$
|\widetilde{F}|=\left[\frac{IJ^2\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(1+J\nu)^2}\left(\frac{I\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^2)^2}+\frac{I(J-1)}{2(\sigma_1^2)^2}\right)-\frac{IJ\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^2)(1+J\nu)}\frac{IJ\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^2)(1+J\nu)}\right]
$$

$$
|\widetilde{F}|=\left(\frac{I^2J^2\left(\frac{9}{2}t^2(\underline{\theta})+1\right)^2}{4(\sigma_1^2)^2(1+J\nu)^2}+\frac{I^2J^2\left(\frac{9}{2}t^2(\underline{\theta})+1\right)(J-1)}{4(\sigma_1^2)^2(1+J\nu)^2}\right)-\frac{I^2J^2\left(\frac{9}{2}t^2(\underline{\theta})+1\right)^2}{4(\sigma_1^2)^2(1+J\nu)^2}
$$

$$
|\widetilde{F}|=\frac{I^2J^2\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{4(\sigma_1^2)^2(1+J\nu)^2}\left\{\left(\frac{\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{1}+\frac{(J-1)}{1}\right)-\frac{\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{1}\right\}
$$

$$
|\widetilde{F}|=\frac{I^2J^2\left(\frac{9}{2}t^2(\underline{\theta})+1\right)(J-1)}{4(\sigma_1^2)^2(1+J\nu)^2}
$$

$$\widetilde{F}^{-1} = \frac{1}{\left|\widetilde{F}\right|}\begin{pmatrix} \dfrac{I\left(\dfrac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^{\,2})^2} + \dfrac{I(J-1)}{2(\sigma_1^{\,2})^2} & -\dfrac{IJ\left(\dfrac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^{\,2})(1+J\nu)} \\ -\dfrac{IJ\left(\dfrac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^{\,2})(1+J\nu)} & \dfrac{IJ^2\left(\dfrac{9}{2}t^2(\underline{\theta})+1\right)}{2(1+J\nu)^2} \end{pmatrix}$$

$$\widetilde{F}^{-1} = \frac{4(\sigma_1^{\,2})^2(1+J\nu)^2}{I^2J^2\left(\frac{9}{2}t^2(\underline{\theta})+1\right)(J-1)}\begin{pmatrix} \dfrac{IJ^2\left(\dfrac{9}{2}t^2(\underline{\theta})+1\right)}{2(1+J\nu)^2} & -\dfrac{IJ\left(\dfrac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^{\,2})(1+J\nu)} \\ -\dfrac{IJ\left(\dfrac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^{\,2})(1+J\nu)} & \dfrac{I\left(\dfrac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^{\,2})^2} + \dfrac{I(J-1)}{2(\sigma_1^{\,2})^2} \end{pmatrix}$$

$$\widetilde{F}^{-1} = \begin{pmatrix} \dfrac{2(\sigma_1^{\,2})^2}{I(J-1)} & -\dfrac{2(\sigma_1^{\,2})(1+J\nu)}{IJ(J-1)} \\ -\dfrac{2(\sigma_1^{\,2})(1+J\nu)}{IJ(J-1)} & \dfrac{2(1+J\nu)^2}{IJ^2(J-1)} + \dfrac{2(1+J\nu)^2}{IJ^2\left(\dfrac{9}{2}t^2(\underline{\theta})+1\right)} \end{pmatrix}$$

$$\widetilde{F}^{-1} = \begin{pmatrix} f_{22} & f_{23} \\ f_{32} & f_{33} \end{pmatrix}^{-1}.$$

$$h_1 = f_{11.2} = f_{11} - \begin{bmatrix} f_{12} & f_{13} \end{bmatrix} \widetilde{F}^{-1} \begin{bmatrix} f_{21} \\ f_{31} \end{bmatrix}$$

$$h_1 = f_{11.2} = f_{11} - \begin{bmatrix} f_{12} & f_{13} \end{bmatrix} \begin{pmatrix} f_{22} & f_{23} \\ f_{32} & f_{33} \end{pmatrix}^{-1} \begin{bmatrix} f_{21} \\ f_{31} \end{bmatrix}$$

$$h_1 = 9I - \begin{bmatrix} \frac{9}{2}\frac{I}{(\sigma_1^2)}t(\underline{\theta}) & \frac{9}{2}\frac{IJ}{(1+J\nu)}t(\underline{\theta}) \end{bmatrix} \begin{pmatrix} \frac{2(\sigma_1^2)^2}{I(J-1)} & -\frac{2(\sigma_1^2)(1+J\nu)}{IJ(J-1)} \\ -\frac{2(\sigma_1^2)(1+J\nu)}{IJ(J-1)} & \frac{2(1+J\nu)^2}{IJ^2(J-1)} + \frac{2(1+J\nu)^2}{IJ^2\left(\frac{9}{2}t^2(\underline{\theta})+1\right)} \end{pmatrix}^{-1} \begin{bmatrix} \frac{9}{2}\frac{I}{(\sigma_1^2)}t(\underline{\theta}) \\ \frac{9}{2}\frac{IJ}{(1+J\nu)}t(\underline{\theta}) \end{bmatrix}$$

$$h_1 = 9I - \left(\frac{9}{2}\right)^2 I^2 t^2(\underline{\theta}) \begin{bmatrix} \frac{1}{(\sigma_1^2)} & \frac{J}{(1+J\nu)} \end{bmatrix} \begin{pmatrix} \frac{2(\sigma_1^2)^2}{I(J-1)} & -\frac{2(\sigma_1^2)(1+J\nu)}{IJ(J-1)} \\ -\frac{2(\sigma_1^2)(1+J\nu)}{IJ(J-1)} & \frac{2(1+J\nu)^2}{IJ^2(J-1)} + \frac{2(1+J\nu)^2}{IJ^2\left(\frac{9}{2}t^2(\underline{\theta})+1\right)} \end{pmatrix} \begin{bmatrix} \frac{1}{(\sigma_1^2)} \\ \frac{J}{(1+J\nu)} \end{bmatrix}$$

$$h_1 = 9I - \left(\frac{9}{2}\right)^2 I^2 t^2(\underline{\theta}) \begin{bmatrix} \frac{2(\sigma_1^2)}{I(J-1)} - \frac{2(\sigma_1^2)}{I(J-1)} & -\frac{2(1+J\nu)}{IJ(J-1)} + \frac{2(1+J\nu)}{IJ(J-1)} + \frac{2(1+J\nu)}{IJ\left(\frac{9}{2}t^2(\underline{\theta})+1\right)} \end{bmatrix} \begin{bmatrix} \frac{1}{(\sigma_1^2)} \\ \frac{J}{(1+J\nu)} \end{bmatrix}$$

$$h_1 = 9I - \left(\frac{9}{2}\right)^2 I^2 t^2(\underline{\theta}) \begin{bmatrix} 0 & \frac{2(1+J\nu)}{IJ\left(\frac{9}{2}t^2(\underline{\theta})+1\right)} \end{bmatrix} \begin{bmatrix} \frac{1}{(\sigma_1^2)} \\ \frac{J}{(1+J\nu)} \end{bmatrix}$$

$$h_1 = 9I - \frac{2\left(\frac{9}{2}\right)^2 It^2(\underline{\theta})}{\left(\frac{9}{2}t^2(\underline{\theta})+1\right)} = \frac{9I\left(\frac{9}{2}t^2(\underline{\theta})+1\right) - 2\left(\frac{9}{2}\right)^2 It^2(\underline{\theta})}{\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}$$

$$h_1 = \frac{\left(\frac{81}{2}It^2(\underline{\theta})+9I\right) - \frac{81}{2}It^2(\underline{\theta})}{\left(\frac{9t^2(\underline{\theta})+2}{2}\right)} = 18I\left(9t^2(\underline{\theta})+2\right)^{-1} \text{ (as before)}$$

$$\therefore h_1^{\frac{1}{2}} \propto \left(9t^2(\underline{\theta})+2\right)^{-\frac{1}{2}}.$$

Let

$$h_2 = f_{22} - \frac{1}{f_{33}} f_{23} f_{32}$$

$$h_2 = \frac{I\left(\frac{9}{2}t^2(\underline{\theta})+1\right)+I(J-1)}{2(\sigma_1^{\ 2})^2} - \frac{2(1+J\nu)^2}{IJ^2\left(\frac{9}{2}t^2(\underline{\theta})+1\right)} \frac{IJ\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^{\ 2})(1+J\nu)} \frac{IJ\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^{\ 2})(1+J\nu)}$$

$$h_2 = \frac{I\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^{\ 2})^2} - \frac{I\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(\sigma_1^{\ 2})^2} = \frac{I(J-1)}{2(\sigma_1^{\ 2})^2} \propto (\sigma_1^{\ 2})^{-2}$$

$$h_2 \propto (\sigma_1^{\ 2})^{-2}$$

$$h_3 = f_{33} = \frac{IJ^2\left(\frac{9}{2}t^2(\underline{\theta})+1\right)}{2(1+J\nu)^2} \propto (1+J\nu)^{-2}.$$

It follows that

$$p(t(\underline{\theta})) \propto h_1^{\frac{1}{2}} = \left(9t^2(\underline{\theta})+2\right)^{-\frac{1}{2}}$$

$$p(\sigma_1^{\ 2} \mid t(\underline{\theta})) \propto h_2^{\frac{1}{2}} = \sigma_1^{-2}$$

$$p(\nu \mid t(\underline{\theta}), \sigma_1^{\ 2}) \propto h_3^{\frac{1}{2}} = \left(1+J\nu\right)^{-1}.$$

Therefore the reference prior relative to the ordered parameterisation $(t(\underline{\theta}), \sigma_1^{\ 2}, \nu)$ is given by

$$p(t(\underline{\theta}), \nu, \sigma_1^{\ 2}) = p(t(\underline{\theta}))\, p(\sigma_1^{\ 2} \mid t(\underline{\theta})\, p(\nu \mid t(\underline{\theta}), \sigma_1^{\ 2})$$

$$p(t(\underline{\theta}), \nu, \sigma_1^{\ 2}) = \left(9t^2(\underline{\theta})+2\right)^{-\frac{1}{2}} \sigma_1^{-2} \left(1+J\nu\right)^{-1}.$$

The reference prior in the $(\mu, \sigma_1^{\ 2}, \sigma_2^{\ 2})$ parameterisation

The reference prior relative to the ordered parameterisation $(t(\underline{\theta}), \sigma_1^{\ 2}, \nu)$ is given by

$$p(t(\underline{\theta}), \sigma_1^{\ 2}, \nu) \propto \left(9t^2(\underline{\theta})+2\right)^{-\frac{1}{2}} \left(1+J\nu\right)^{-1} \sigma_1^{-2}$$

which is the same as before and

$$t(\underline{\theta}) = \frac{\mu - l_0}{3\left(\frac{\sigma_1^{\,2} + J\sigma_2^{\,2}}{J}\right)^{\frac{1}{2}}}$$

let

$x_1 = t(\underline{\theta}) \quad y_1 = \mu$

then

$$y_1 = 3x_1\left(\frac{\sigma_1^{\,2} + J\sigma_2^{\,2}}{J}\right)^{\frac{1}{2}} + l_0 \text{ and } x_1 = \frac{y_1 - l_0}{3\left(\frac{\sigma_1^{\,2} + J\sigma_2^{\,2}}{J}\right)^{\frac{1}{2}}},$$

$$\frac{\partial x_1}{\partial y_1} = \frac{1}{3}\left(\frac{\sigma_1^{\,2} + J\sigma_2^{\,2}}{J}\right)^{-\frac{1}{2}}.$$

For

$\nu = \frac{\sigma_2^{\,2}}{\sigma_1^{\,2}}$,

Let $x_2 = \nu \quad y_2 = \sigma_2^{\,2}$

$y_2 = x_2\sigma_1^{\,2}$ and $x_2 = \frac{y_2}{\sigma_1^{\,2}}$

$$\frac{\partial x_2}{\partial y_2} = \frac{1}{\sigma_1^{\,2}}$$

i.e. $\frac{\partial t(\underline{\theta})}{\partial \mu} = \frac{1}{3}\left(\frac{\sigma_1^{\,2} + J\sigma_2^{\,2}}{J}\right)^{-\frac{1}{2}}$, $\nu = \frac{\sigma_2^{\,2}}{\sigma_1^{\,2}}$, $\frac{\partial \nu}{\partial \sigma_2^{\,2}} = \frac{1}{\sigma_1^{\,2}}$

$$p_{R_2}(y_1, \sigma_1^{\,2}, y_2) \propto p(x_1, \sigma_1^{\,2}, x_2)\left|\frac{dx_1}{dy_1}\right|\left|\frac{dx_2}{dy_2}\right|$$

$$p_{R_2}(\mu,\sigma_1{}^2,\sigma_2{}^2) \propto \left(\frac{J(\mu-l_0)^2}{2\left(\sigma_1{}^2+J\sigma_2{}^2\right)}+1\right)^{-\frac{1}{2}}\left(1+\frac{J\sigma_2{}^2}{\sigma_1{}^2}\right)^{-1}\sigma_1{}^{-2}\,\frac{J^{\frac{1}{2}}}{3}\left(\sigma_1{}^2+J\sigma_2{}^2\right)^{-\frac{1}{2}}\sigma_1{}^{-2}$$

$$p_{R_2}(\mu,\sigma_1{}^2,\sigma_2{}^2) \propto \left(\frac{J(\mu-l_0)^2}{2\left(\sigma_1{}^2+J\sigma_2{}^2\right)}1+\right)^{-\frac{1}{2}}\left(\sigma_1{}^2+J\sigma_2{}^2\right)^{-\frac{3}{2}}\sigma_1{}^2\sigma_1{}^{-2}\sigma_1{}^{-2}$$

$$p_{R_2}(\mu,\sigma_1{}^2,\sigma_2{}^2) \propto \left(\frac{J(\mu-l_0)^2}{2\left(\sigma_1{}^2+J\sigma_2{}^2\right)}1+\right)^{-\frac{1}{2}}\left(\sigma_1{}^2+J\sigma_2{}^2\right)^{-\frac{3}{2}}\sigma_1{}^{-2}$$

which corresponds to the probability matching prior $\pi_a(\underline{\theta})$.

CHAPTER 6

A PROCESS CAPABILITY INDEX FOR AVERAGES OF OBSERVATIONS FROM NEW BATCHES IN THE CASE OF THE BALANCED RANDOM EFFECTS MODEL WITH THREE VARIANCE COMPONENTS

6.1 INTRODUCTION

The results from the previous chapter may be generalised. In this chapter, an extension from the two variance component to the three variance component model will be discussed. The following development of the three variance component model is broadly similar to that for two variance component model. For notational convenience, the letters I, J and K will now be used to indicate number of time-periods, number of subgroups and subgroup size respectively.

Data arising from multiple sources of variability are very common in practice. Virtually all industrial processes exhibit between-batch, as well as within-batch components of variation. In some cases the between-batch (or between subgroup) component is viewed as part of the common-cause-system for the process. It seems worthwhile to develop a process capability index in even more general settings. In this chapter we add a third variance component. To do so, it is necessary to employ a statistical model which adequately handles multiple sources of variability.

A version of the process capability or performance index for the balanced random effects model with three variance components from a Bayesian framework is scrutinised again. The index is denoted by ${}_3P_{pl}{}^1$ and can be used for average of observations for a given time-period.

6.2 DEFINITIONS AND NOTATIONS

The lower process performance index for the three variance component model is defined as:

$$ {}_3P_{pl}^{1} = \frac{\mu - l_0}{3\left(\frac{\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2}{JK}\right)^{\frac{1}{2}}} \tag{6.2.1}$$

where

$\mu =$ mean of future observation for a new or unknown batch

$\sigma_1^{\,2} =$ residual variance

$\sigma_2^{\,2} =$ within groups variance

$\sigma_3^{\,2} =$ between groups variance

$l_0 =$ lower specification limit

$J =$ batch (package) size

$K =$ number of sub-samples per batch (package)

To illustrate how and when the index (6.2.1) will be used, consider a factory which manufactures chronic medication in the form of medical tablets in very small batches. A small batch in this instance is likely to be a weekly intake of tablets for an individual patient. Monthly samples of $J = 8$ packages say, of the tablet are sampled and various properties of the tablet are then replicated in the laboratory by analysing $K = 5$ tablets say, per package. Packages with tablets are sampled for a period of $I = 15$ months. The interest is in whether the patient gets on average the required dosage of the drug from the batches in the specified time, given that each patient must get an average dosage of at least l_0. The question therefore is whether the process is capable of producing to this specification over a one month future period (see section 6.9).

As introduced above, the variation observed could possibly be explained by several components such as a "*between*" months (month-to-month) component, a "*within*" months (package- to-package) component and a *residual* component. These sources of variation should be incorporated in a suitable model.

Approximations of the exact posterior distribution of ${}_3P_{pl}{}^{1}$ can be obtained. Current knowledge indicates that a posterior analysis for this form of capability index does not exist.

The above mentioned index will be contrasted with the following indices:

i. $${}_3P_{pl}^{11} = \left(\frac{\mu - l_0}{3\sigma_{total}}\right) = \left(\frac{\mu - l_0}{3\left(\sigma_1{}^2 + \sigma_2{}^2 + \sigma_3{}^2\right)^{\frac{1}{2}}}\right)$$

which is independent of I, J and K. This index assesses whether the process is capable of producing each future tablet to specification.

ii. $${}_3P_{pl}{}^{111} = \left(\frac{\mu - l_0}{3\left(\dfrac{\sigma_1{}^2 + K\sigma_2{}^2 + K\sigma_3{}^2}{K}\right)^{\frac{1}{2}}}\right)$$

which assesses whether the process is capable of producing a future batch to specification**.**

iii. $${}_3P_{pl}^{1V} = \frac{\mu - l_0}{3\left(\dfrac{\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2}{IJK}\right)^{\frac{1}{2}}}$$

which assesses whether the process is capable of producing to specification over a future period of I months.

iv. $${}_3P_{pl}^{V} = \left(\frac{\mu_i - l_0}{3\left(\dfrac{\sigma_1{}^2}{JK} + \dfrac{\sigma_2{}^2}{J}\right)^{\frac{1}{2}}}\right) = \left(\frac{\mu_i - l_0}{3\left(\dfrac{\sigma_1{}^2 + K\sigma_2{}^2}{JK}\right)^{\frac{1}{2}}}\right)$$ with $\mu_i = \mu + r_i$.

which is independent $\sigma_3^{\,2}$ and assesses whether the process is capable of producing to specification in a specific month, say the 10th month or a month similar to the 10th month.

In the next sections we will therefore look at the three variance component model. The posterior distributions of the mean and variance components are derived in section 6.4. In section 6.5, the posterior distribution of ${}_3P_{pl}^{\;1}$ conditional on the variance components is derived. The probability–matching prior for ${}_3P_{pl}^{\;1}$ will be derived in section 6.6. Sections 6.7 and 6.8 deal with the estimation of the indices. We conclude with an application in section 6.9.

6.3 THE VARIANCE COMPONENT MODEL

The variance component model with three variance components is of the form:

$$Y_{ijk} = \mu + r_i + c_{ij} + \varepsilon_{ijk} \quad i = 1,\ldots,I, \quad j = 1,\ldots,J \quad \text{and} \quad k = 1,\ldots,K, \tag{6.3.1}$$

where Y_{ijk} are the observations, μ is a common location parameter, r_i, c_{ij} and ε_{ijk} are three different kinds of random effects. We further assume that the random effects $(r_i, c_{ij}, \varepsilon_{ijk})$ are all independent and that

$$r_i \sim N(0,\sigma_3^2), \quad c_{ij} \sim N(0,\sigma_2^2) \quad \text{and} \quad \varepsilon_{ijk} \sim N(0,\sigma_1^2).$$

and the parameters $(\sigma_1^2, \sigma_2^2, \sigma_3^2)$ are the variance components. This model explains each data point as being additively influenced by three random effects (*residual, batches (packages)* and *groups).* The variance of each data point therefore consists of three components corresponding to each of these random effects.

The effects for our earlier example consist of *residual, packages* and *months* effects and are then denoted by $(\varepsilon_{ijk}, c_{ij}, r_i)$. Some important results of the variance component model are summarised in theorem 6.3.1.

Theorem 6.3.1

I. $Y_{ijk} \mid \mu, \sigma_1^{\,2}, \sigma_2^{\,2}, \sigma_3^{\,2} \sim N(\mu, \sigma_1^{\,2} + \sigma_2^{\,2} + \sigma_3^{\,2})$

II. $\overline{Y_{ij.}} \mid \mu, \sigma_1^{\ 2}, \sigma_2^{\ 2}, \sigma_3^{\ 2} \sim N(\mu, \sigma_3^{\ 2} + \sigma_2^{\ 2} + \frac{\sigma_1^{\ 2}}{K}) = N(\mu, \frac{K\sigma_3^{\ 2} + K\sigma_2^{\ 2} + \sigma_1^{\ 2}}{K})$

III. $\overline{\overline{Y_{i..}}} \mid \mu, \sigma_1^{\ 2}, \sigma_2^{\ 2}, \sigma_3^{\ 2} \sim N(\mu, \sigma_3^{\ 2} + \frac{\sigma_2^{\ 2}}{J} + \frac{\sigma_1^{\ 2}}{KJ}) = N(\mu, \frac{JK\sigma_3^{\ 2} + K\sigma_2^{\ 2} + \sigma_1^{\ 2}}{JK})$

IV. $\overline{\overline{\overline{Y_{...}}}} \mid \mu, \sigma_1^{\ 2}, \sigma_2^{\ 2}, \sigma_3^{\ 2} \sim N(\mu, \frac{\sigma_3^{\ 2}}{I} + \frac{\sigma_2^{\ 2}}{IJ} + \frac{\sigma_1^{\ 2}}{IJK}) = N(\mu, \frac{JK\sigma_3^{\ 2} + K\sigma_2^{\ 2} + \sigma_1^{\ 2}}{IJK})$

V. $\overline{\overline{Y_{i..}}} \mid \theta, r_i, \sigma_1^{\ 2}, \sigma_2^{\ 2} \sim N(\theta + r_i, \frac{\sigma_2^{\ 2}}{J} + \frac{\sigma_1^{\ 2}}{KJ}) = N(\theta + r_i, \frac{K\sigma_2^{\ 2} + \sigma_1^{\ 2}}{JK})$.

<u>Proof</u>

The proofs are very simple and similar to those in chapter 5 and only proofs to III and V are given in Appendix A6.

6.4 POSTERIOR DISTRIBUTION OF THE MEAN AND THE VARIANCE COMPONENTS

Consider the model in (6.2.1) again. The following theorems can now be stated.

Theorem 6.4.1

For the non-informative joint prior (see Box and Tiao (1973)):

$$p(\mu, \sigma_1^{\ 2}, \sigma_{12}^{\ 2}, \sigma_{123}^{\ 2}) \propto \sigma_1^{-2}(\sigma_1^2 + K\sigma_2^2)^{-1}(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)^{-1}.$$

The joint posterior distribution of $\sigma_1^{\ 2}, \sigma_2^{\ 2}$ and σ_3^2 can be worked out as

$$p\left(\sigma_1^{\ 2}, \sigma_{12}^{\ 2}, \sigma_{123}^{\ 2} \mid \underline{Y}\right) \propto (\sigma_1^{\ 2})^{\frac{-(\frac{1}{2}\nu_1+2)}{2}} \exp(-\frac{1}{2}\{\frac{\nu_1 m_1}{\sigma_1^{\ 2}}\}) \times (\sigma_1^{\ 2} + K\sigma_2^{\ 2})^{\frac{-\frac{1}{2}(\nu_2+2)}{2}} \exp(-\frac{1}{2}\{\frac{\nu_2 m_2}{(\sigma_1^{\ 2} + K\sigma_2^{\ 2})}\} \times$$
$$(\sigma_1^{\ 2} + K\sigma_2^{\ 2} + JK\sigma_3^{\ 2})^{\frac{-\frac{1}{2}(\nu_3+2)}{2}} \exp(-\frac{1}{2}\{\frac{\nu_3 m_3}{(\sigma_1^{\ 2} + K\sigma_2^{\ 2} + JK\sigma_3^{\ 2})}\})$$

and

$$p\left(\mu \mid \underline{Y}, \sigma_1^{\ 2}, \sigma_2^{\ 2}, \sigma_3^{\ 2}\right) \propto \left[(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)^{\frac{-1}{2}} \exp(-\frac{1}{2}\frac{IJK(\overline{\overline{\overline{Y}}}_{...} - \mu)^2}{(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)})\right]$$

i.e. $\mu \mid \underline{Y} \sim N(\overline{\overline{\overline{Y}}}_{...}, \frac{\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2}{IJK})$

where

$$v_1 m_1 = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(Y_{ijk} - \overline{Y_{ij.}})^2 \,, v_2 m_2 = K\sum_{i=1}^{I}\sum_{j=1}^{J}(\overline{Y_{ij.}} - \overline{\overline{Y_{i..}}})^2 \,,\ v_3 m_3 = JK\sum_{i=1}^{I}(\overline{\overline{Y_{i..}}} - \overline{\overline{\overline{Y_{...}}}})^2$$

$$\underline{Y} = \left[Y_{111,}\ Y_{112,} \cdots\cdots Y_{IJK}\right]'$$

$$v_1 = IJ(K-1), \quad v_2 = I(J-1), \quad v_3 = I-1.$$

Proof

The proof is given in Appendix A6.

The posterior distribution for $\sigma_1^{\ 2}, \sigma_{12}^{\ 2}$ and $\sigma_{123}^{\ 2}$ would be independent, each proportional to an inverted gamma distribution, if the restriction $\sigma_{123}^{\ 2} > \sigma_{12}^{\ 2} > \sigma_1^{\ 2}$ did not apply. The joint posterior distribution for $\sigma_1^{\ 2}, \sigma_{12}^{\ 2}$ and $\sigma_{123}^{\ 2}$ would be the product of the three distributions.

$$\sigma_1^{\ 2}, \sigma_{12}^{\ 2}, \sigma_{123}^{\ 2} \mid \underline{Y} \sim IG(\sigma_1^{\ 2} \mid \frac{\nu_1}{2}; \nu_1 m_1) \times IG(\sigma_1^{\ 2} + K\sigma_2^{\ 2} \mid \frac{\nu_2}{2}; \nu_2 m_2) \times IG(\sigma_1^{\ 2} + K\sigma_2^{\ 2} + JK\sigma_3^{\ 2} \mid \frac{\nu_3}{2}; \nu_3 m_3)$$

However the restrictions do apply. Nevertheless, using a three-step rejection sampling procedure (as will be discussed) it is straight-forward to generate samples from the joint distribution.

Theorem 6.4.2

I. $E(m_1) = \sigma_1^{\ 2}$ **II.** $E(m_2) = \sigma_{12}^{\ 2} = (\sigma_1^{\ 2} + K\sigma_2^{\ 2})$

III. $E(m_3) = \sigma_{123}^{\ 2} = (\sigma_1^{\ 2} + K\sigma_2^{\ 2} + JK\sigma_3^{\ 2})$.

Proof

The proofs are very simple and similar to those in chapter 5 and only III is given in Appendix A6.

The following theorem gives the posterior distribution of $\mu_i = \mu + r_i \quad (i = 1, \ldots, I)$ given $\underline{Y}$ and the variance components σ_1^2, σ_2^2 and σ_3^2.

Theorem 6.4.3

μ_i given σ_1^2, σ_2^2 and σ_3^2 is normally distributed with mean

$$E(\mu_i \mid \sigma_1^2, \sigma_2^2, \sigma_3{}^2, \underline{Y}) = \frac{JK\sigma_3{}^2 \overline{\overline{Y}}_{i..}}{(\sigma_1{}^2 + K\sigma_2{}^2) + JK\sigma_3{}^2} + \frac{(\sigma_1{}^2 + K\sigma_2{}^2)}{(\sigma_1{}^2 + K\sigma_2{}^2) + JK\sigma_3{}^2} \overline{\overline{\overline{Y}}}_{...}$$

and variance

$$\text{Var}(\mu_i \mid \underline{Y}, \sigma_1^2, \sigma_2^2, \sigma_3^2) = \frac{(\sigma_1{}^2 + K\sigma_2{}^2)}{(\sigma_1{}^2 + K\sigma_2{}^2 + JK\sigma^2{}_3)} \left\{ \frac{IJK\sigma_3^2 + \sigma_1{}^2 + K\sigma_2{}^2}{IJK} \right\}$$

Proof

The proof is given in Appendix A6 (See van der Merwe and Hugo (2007) for a similar proof).

6.5 POSTERIOR DISTRIBUTION OF THE LOWER PROCESS PERFORMANCE INDEX ${}_3P^1{}_{pl}$ WITH THREE VARIANCE COMPONENTS

Theorem 6.5.1

The posterior distribution of ${}_3P_{pl}^1$ given the variance components is

$${}_3P_{pl}^1 \mid \underline{Y}, \sigma_1{}^2, \sigma_2{}^2, \sigma_3^2 \sim N\left(\frac{\overline{\overline{\overline{Y}}}_{...} - l_0}{3\left(\dfrac{\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2}{JK}\right)^{\frac{1}{2}}}, \frac{1}{9I} \right) \textit{for all } i, j, k \qquad (6.5.1)$$

Proof

The proof is given in Appendix A6.

The unconditional posterior distribution of ${}_3P_{pl}^1$ can be obtained by using Monte Carlo simulation.

6.6 THE PROBABILITY MATCHING PRIOR FOR THE LOWER PROCESS CAPABILITY INDEX $_3P^1_{pl}$

Theorem 6.6.1

The probability-matching prior for the $_3P^1_{pl}$ Index for the balanced random effects model defined in (6.2.1) is:

$$\pi(\underline{\theta}) = \pi(\mu, \sigma_1^2, \sigma_2^2, \sigma_3^2) \propto \sigma_1^{-2}\left(\sigma_1^2 + K\sigma_2^2\right)^{-1}\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^{-\frac{1}{2}}\left(1 + \frac{(\mu - l_0)^2 (JK)}{2\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)}\right)^{\frac{1}{2}}$$

Proof

The proof is given in Appendix A6.

6.7 MONTE CARLO SIMULATION PROCEDURE FOR ESTIMATING THE POSTERIOR DISTRIBUTION OF $_3P^1_{pl}$

Simulation of the posterior of σ_1^2, σ_2^2 and σ_3^2 can be achieved through the following standard simulation routines. By using the Matlab package, simulation of σ_1^2, σ_2^2 and σ_3^2 can be obtained in the following way:

ix. Draw τ from a χ_{v_1}-distribution

x. Let $\frac{1}{\tau} = \frac{\sigma_1^2}{v_1 m_1}$

xi. $\sigma_1^{2*} = \frac{v_1 m_1}{\tau}$ where the * indicates a simulated value.

xii. Draw u from a χ_{v_2}-distribution

xiii. Let $\frac{1}{u} = \frac{\sigma_{12}^2}{v_2 m_2}$ where $\sigma_{12}^2 = \sigma_1^2 + J\sigma_2^2$

xiv. $\sigma_{12}{}^{2*} = \dfrac{v_2 m_2}{u}$

xv. If $\sigma_{12}{}^{2*} > \sigma_1^{2*}$ (σ_1^{2*} was simulated in step (iii))

Calculate $\sigma_2{}^{2*} = \dfrac{1}{J}(\sigma_{12}{}^{2*} - \sigma_1{}^{2*})$ from the expression $\sigma_{12}{}^{2*} = \sigma_1{}^{2*} + J\sigma_2{}^{2*}$.

If $\sigma_{12}{}^{2*} < \sigma_1{}^{2*}$ ignore the values of $\sigma_1{}^{2*}, \sigma_2{}^{2*}$ and start again.

xvi. Draw $\aleph$ from a χ_{v_3}-distribution

xvii. Let $\dfrac{1}{\aleph} = \dfrac{\sigma_{123}{}^2}{v_3 m_3}$ where $\sigma_{123}{}^2 = \sigma_1{}^2 + K\sigma_2{}^2 + JK\sigma_3{}^2$

xviii. $\sigma_{123}{}^{2*} = \dfrac{v_3 m_3}{\aleph}$

xix. If $\sigma_{123}{}^{2*} > \sigma_{12}{}^{2*} > \sigma_1{}^{2*}$ (σ_1^{2*} was simulated in step (iii) and σ_2^{2*} was simulated in step (iv)))

Calculate $\sigma_3{}^{2*} = \dfrac{\sigma_{123}{}^{2*} - [\sigma_1{}^{2*} + K\sigma_2{}^{2*}]}{JK}$ from the expression

$\sigma_{123}{}^{2*} = \sigma_1{}^{2*} + K\sigma_2{}^{2*} + JK\sigma_3{}^{2*}$.

If $\sigma_{123}{}^{2*} < \sigma_{12}{}^{2*}$ ignore the values of $\sigma_1{}^{2*}, \sigma_2{}^{2*}, \sigma_3{}^{2*}$ and start again.

By making use of the fact that $\mu \mid \underline{Y}, \sigma_1^2, \sigma_2^2, \sigma_3^2 \sim N\left\{\overline{\overline{\overline{Y}}}_{\dots}, \dfrac{\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2}{IJK}\right\}$ and from the definition of the performance index (6.2.1), it follows that ${}_3P_{pl}^1$ can be simulated.

Given $\sigma_1^2, \sigma_2^2, \sigma_3^2$, the conditional posterior density function $p({}_3P_{pl}^1 \mid \underline{Y}, \sigma_1{}^2, \sigma_2{}^2, \sigma_3{}^2)$ is calculated which is defined in (6.5.1). Repeat steps (i-xi) $\tilde{\ell}$ times. For our example $\tilde{\ell}$ was taken as 10000. Using a Rao–Blackwell argument (Gelfand and Smith, 1991), a density estimate of the unconditional posterior distribution of (6.2.1), can be obtained by averaging $p({}_3P_{pl}^1 \mid \underline{Y}, \sigma_1{}^2, \sigma_2{}^2, \sigma_3{}^2)$ over the $\tilde{\ell}$ repetitions.

Simulation results for the variance component process performance index ${}_3P_{pl}^1$ as discussed in this article will now be compared with the following indices simulated as ${}_3P_{pl}{}^{11*}, {}_3P_{pl}{}^{111*}, {}_3P_{pl}^{1V}$.

By making use of the fact that

$$\mu_i \mid \underline{Y}, \sigma_1^2, \sigma_2^2, \sigma_3^2 \sim N\left\{E(\mu_i \mid \sigma_1^2, \sigma_2^2, \sigma_3{}^2, \underline{Y}), \mathrm{Var}(\mu_i \mid \underline{Y}, \sigma_1^2, \sigma_2^2, \sigma_3^2)\right\}$$

where

$$E(\mu_i \mid \sigma_1^2, \sigma_2^2, \sigma_3{}^2, \underline{Y}) = \frac{JK\sigma_3{}^2 \overline{\overline{Y_{i..}}}}{(\sigma_1{}^2 + K\sigma_2{}^2) + JK\sigma_3{}^2} + \frac{(\sigma_1{}^2 + K\sigma_2{}^2)}{(\sigma_1{}^2 + K\sigma_2{}^2) + JK\sigma_3{}^2}\overline{\overline{\overline{Y}}}_{...} \quad \text{and}$$

$$\mathrm{Var}(\mu_i \mid \underline{Y}, \sigma_1^2, \sigma_2^2, \sigma_3^2) = \frac{(\sigma_1{}^2 + K\sigma_2{}^2)}{(\sigma_1{}^2 + K\sigma_2{}^2 + JK\sigma^2{}_3)}\left\{\frac{IJK\sigma_3^2 + \sigma_1{}^2 + K\sigma_2{}^2}{IJK}\right\}$$

and from the definition of the index, ${}_3P_{pl}^V$ can also be simulated.

$${}_3P_{pl}^{V*} = \left(\frac{\mu_i^* - l_0}{3\left(\frac{\sigma_1{}^{2*} + K\sigma_2{}^{2*}}{JK}\right)^{\frac{1}{2}}}\right).$$

Simulation results using the probability matching prior will be considered next.

6.8 THE WEIGHTED MONTE CARLO METHOD -SAMPLING-IMPORTANCE RE-SAMPLING

This section describes how to apply a weighted Monte Carlo (WMC) method to simulate ${}_3P_{pl}^1$ using the probability matching prior. This method is especially suitable for computing credibility intervals.

Let

$$\pi(\underline{\theta}) \propto \sigma_1^{-2}\left(\sigma_1^2 + K\sigma_2^2\right)^{-1}\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^{-\frac{1}{2}}\left(1 + \frac{(\mu - l_0)^2 (JK)}{2\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)}\right)^{\frac{1}{2}} \qquad (6.8.1)$$

and

$$q(\underline{\theta}) \propto \sigma_1^{-2}(\sigma_1^2 + K\sigma_2^2)^{-1}(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)^{-1} \tag{6.8.2}$$

According to Smith and Gelfand (1992), Guttman and Menzefricke (2003) and Skare *et al.* (2003), the $\tilde{\ell}$ independent draws of $\underline{\theta}^{*(\ell)} = (\mu^{*(\ell)}, \sigma_1^{2*(\ell)}, \sigma_2^{2*(\ell)}, \sigma_3^{2*(\ell)})$, as discussed in section 6.7, for $\ell = 1$ to $\tilde{\ell}$; is a weighted sample from the posterior distribution based on π, where the weights are

$$w_\ell = \frac{\pi(\underline{\theta}^{*(\ell)}) / q(\underline{\theta}^{*(\ell)})}{\sum_{\ell=1}^{\tilde{\ell}} \pi(\underline{\theta}^{*(\ell)}) / q(\underline{\theta}^{*(\ell)})}$$

$\pi(\underline{\theta}^{*(\ell)})$ and $q(\underline{\theta}^{*(\ell)})$ denote the realistic prior density and implied prior density defined in (6.8.1) and (6.8.2). For the algorithm to be efficient, it is important that the posterior resulting from q is a good approximation to the one obtained from π. This means that the posterior obtained from q should not have too light tails when compared with the posterior from π. For further details see Skare *et al* (2003) and Li (2007).

To simulate using the results from the probability matching prior we associate with each lower performance index value ${}_3P_{pl}^{1*(\ell)}$

$$w_\ell = \frac{\sigma_1^{-2*(\ell)}\left(\sigma_1^{2*(\ell)} + K\sigma_2^{2*(\ell)}\right)^{-1}\left(\sigma_1^{2*(\ell)} + K\sigma_2^{2*(\ell)} + JK\sigma_3^{2*(\ell)}\right)^{-\frac{1}{2}}\left(1 + \frac{(\mu^{*(\ell)} - l_0)^2 (JK)}{2\left(\sigma_1^{2*(\ell)} + K\sigma_2^{2*(\ell)} + JK\sigma_3^{2*(\ell)}\right)}\right)^{\frac{1}{2}}}{\sum_{\ell=1}^{\tilde{\ell}} \sigma_1^{-2*(\ell)}\left(\sigma_1^{2*(\ell)} + K\sigma_2^{2*(\ell)}\right)^{-1}\left(\sigma_1^{2*(\ell)} + K\sigma_2^{2*(\ell)} + JK\sigma_3^{2*(\ell)}\right)^{-\frac{1}{2}}\left(1 + \frac{(\mu^{*(\ell)} - l_0)^2 (JK)}{2\left(\sigma_1^{2*(\ell)} + K\sigma_2^{2*(\ell)} + JK\sigma_3^{2*(\ell)}\right)}\right)^{\frac{1}{2}}}$$

d. Sort the ${}_3P_{pl}^{1*(\ell)}$ values calculated in ascending order so that

$${}_3P_{pl}^{1\,*(1)} \leq {}_3P_{pl}^{1\,*(2)} \leq \cdots \leq {}_3P_{pl}^{1\,*(10000)}$$

e. Compute the weighted function w_{ℓ} associated with the ℓth ordered $P_{pl}^{1*(\ell)}$.

f. Add the weights from left to right (from the first on) until you get $\sum_{\ell=1}^{k_1} w_{\ell} = 0.025$. Write down the corresponding ordered value ${}_3P_{pl}^{1*(k_1)}$ and denote it as $P_{pl}{}^{1*}{}_{(0.025)}$. Add the weights from right to left (from the last back) until you get $\sum_{\ell=k_2}^{\tilde{\ell}} w_{\ell} = 0.025$. Write down the corresponding ordered value ${}_3P_{pl}^{1\,*(k_2)}$ and denote it as ${}_3P_{pl\,(0.975)}^{1*}$. The 95% interval is ${}_3P_{pl\,(0.025)}^{1*} - {}_3P_{pl\,(0.975)}^{1*}$.

6.9 APPLICATION

In a process to manufacture chronic medication, various properties of the manufactured tablet have to be monitored. Monthly samples of J=8 packages of the tablet are sampled and various physical properties of the tablet are replicated in the laboratory by analysing K=5 tablets per package. The data in Table 6.1 represents the amount of drug in a tablet (the percentage of the drug per tablet). Packages with tablets sampled for the first I=15 months starting January of a particular year are selected as review data to determine whether the patient gets on average the required dosage of the drug from the batches in a specified time, given that each patient must get an average dosage of at least $l_0 = 20$.

Table 6.1: Drug data arising from multiple sources of variability

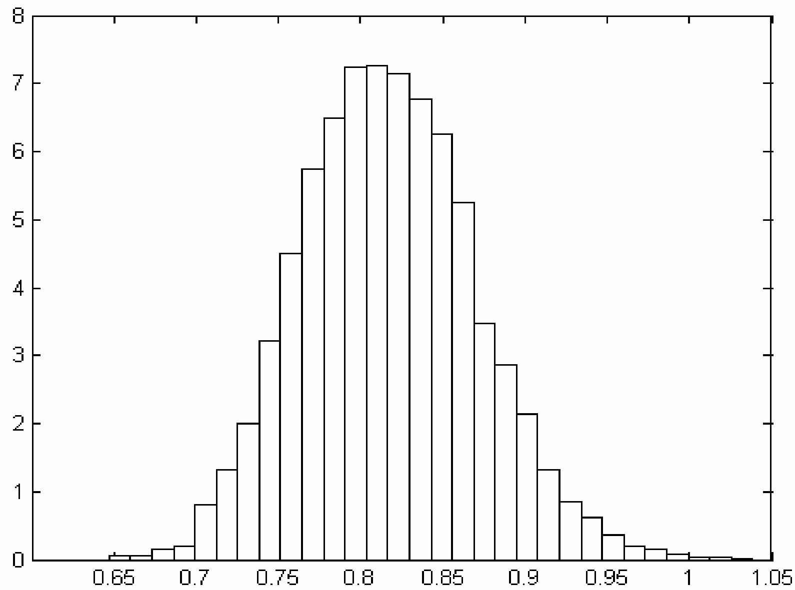
Day	Package	Average Package Dosage	Day	Package	Average Package Dosage	Day	Package	Average Package Dosage
1	1	23.3000	6	1	24.7900	11	1	24.4700
1	2	23.0600	6	2	24.6500	11	2	23.4200
1	3	23.4800	6	3	24.7500	11	3	23.6100
1	4	22.2200	6	4	24.9400	11	4	22.9100
1	5	22.1600	6	5	24.7700	11	5	22.8600

1	6	22.9400	6	6	25.1700	11	6	23.5700
1	7	23.6900	6	7	24.6700	11	7	23.6100
1	8	22.0200	6	8	24.5100	11	8	22.3700
2	1	24.9100	7	1	24.5200	12	1	22.2200
2	2	24.6800	7	2	23.4400	12	2	23.5900
2	3	25.0900	7	3	23.8800	12	3	23.6800
2	4	25.3900	7	4	23.1300	12	4	23.2300
2	5	24.5600	7	5	23.6400	12	5	23.9300
2	6	25.3800	7	6	23.5800	12	6	23.9800
2	7	24.8200	7	7	23.5000	12	7	23.6900
2	8	24.9000	7	8	23.7100	12	8	23.4600
3	1	25.2900	8	1	25.0900	13	1	24.5800
3	2	24.1900	8	2	23.7100	13	2	25.2700
3	3	24.9000	8	3	23.6800	13	3	24.5200
3	4	24.7000	8	4	23.6700	13	4	25.6200
3	5	24.6900	8	5	23.6700	13	5	24.7300
3	6	24.6500	8	6	23.9800	13	6	24.8000
3	7	24.8500	8	7	23.9400	13	7	25.0000
3	8	24.5000	8	8	24.1300	13	8	24.5300
4	1	23.8600	9	1	22.1900	14	1	25.3100
4	2	24.1200	9	2	22.8400	14	2	24.8400
4	3	23.2000	9	3	22.9500	14	3	25.2900
4	4	23.3600	9	4	23.2500	14	4	25.3800
4	5	23.6400	9	5	24.6000	14	5	25.4000
4	6	22.8500	9	6	22.6500	14	6	24.6000
4	7	22.7200	9	7	23.7700	14	7	25.1400
4	8	23.3300	9	8	23.3700	14	8	23.9400
5	1	23.4800	10	1	23.9900	15	1	25.4800
5	2	23.4300	10	2	23.5300	15	2	24.6500
5	3	23.0700	10	3	22.3700	15	3	25.0000
5	4	22.6900	10	4	22.9300	15	4	25.7100
5	5	23.6000	10	5	22.3100	15	5	25.2800
5	6	23.4900	10	6	22.8600	15	6	24.3800
5	7	23.4600	10	7	22.0000	15	7	24.3500
5	8	22.7400	10	8	22.8100	15	8	25.4300

In addition, $v_1 m_1 = 390.6720$. The above data and limit is selected solely for illustrative purposes. In practice, fixed in advance limits are often determined from medical or regulatory considerations.

As introduced above, the variation observed could possibly be explained by several components such as a “between” months (month-to-month) component, a “within” month (package-to-package) component and a residual component.

$p(\sigma_1^2 \mid \underline{Y})$

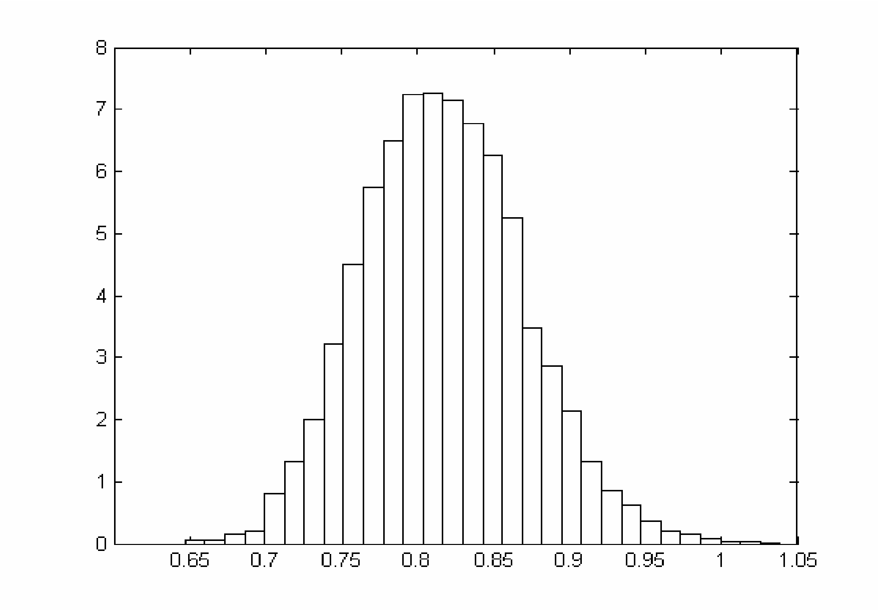


$\sigma_1^2 \mid \underline{Y}$

Figure 6.1: Histogram of simulated variance component

A plot of the posterior distribution of $\sigma_1^2 \mid \underline{Y}$ is symmetrical or fairly symmetrical. The reason for this is the large number of degrees of freedom $v_1 = IJ(K-1) = 480$ associated with the residual variance.

$p(\sigma_2^{\,2} \mid \underline{Y})$

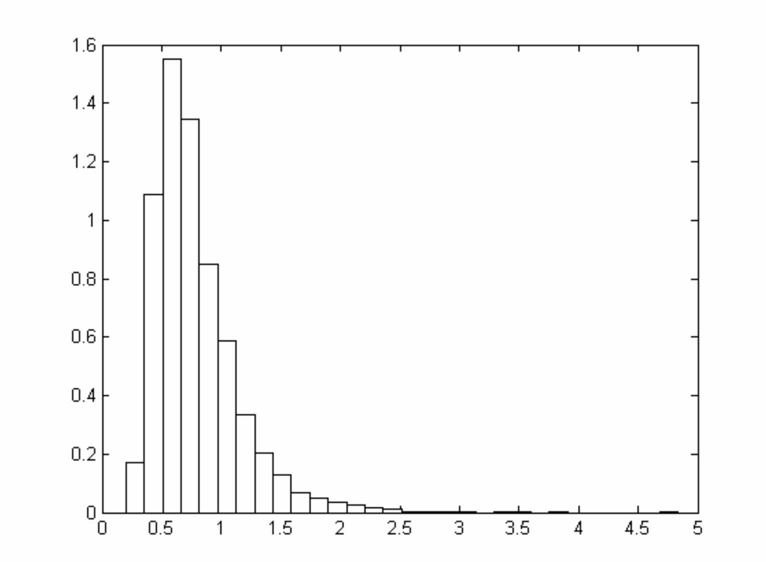


$\sigma_2^{\,2} \mid \underline{Y}$

Figure 6.2: Histogram of simulated variance component

A plot of the posterior distribution of $\sigma_2^{\,2} \mid \underline{Y}$ is also fairly symmetrical. The reason for this is again the large number of degrees of freedom $v_2 = I(J-1) = 105$ associated with the residual variance.

$p(\sigma_3^{\,2} \mid \underline{Y})$

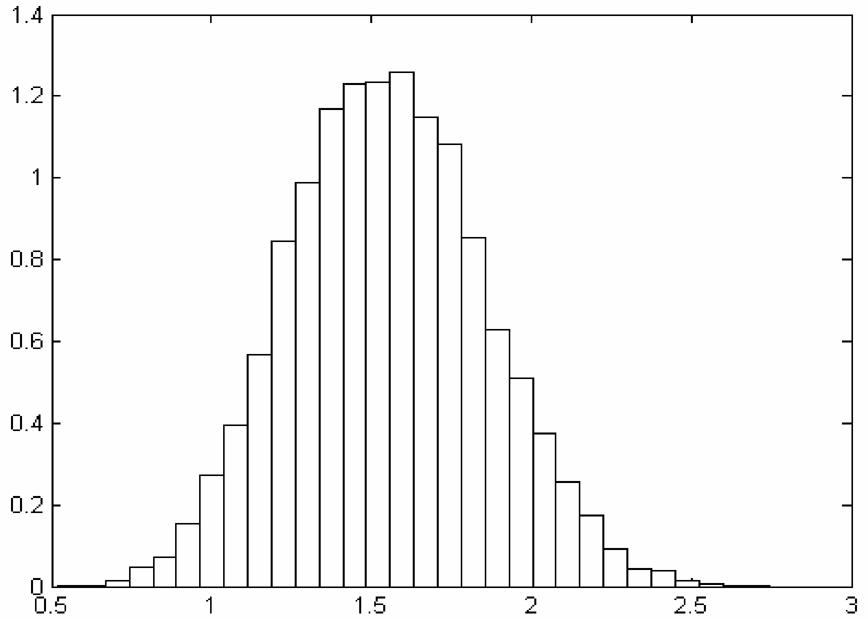


$\sigma_3^{\,2} \mid \underline{Y}$

Figure 6.3: Histogram of simulated variance component

The posterior distribution of $\sigma_3^2 \mid \underline{Y}$ on the other hand is quite skewed. The between months variation is much larger than the within months variation.

$p({}_3P_{pl}^1)$

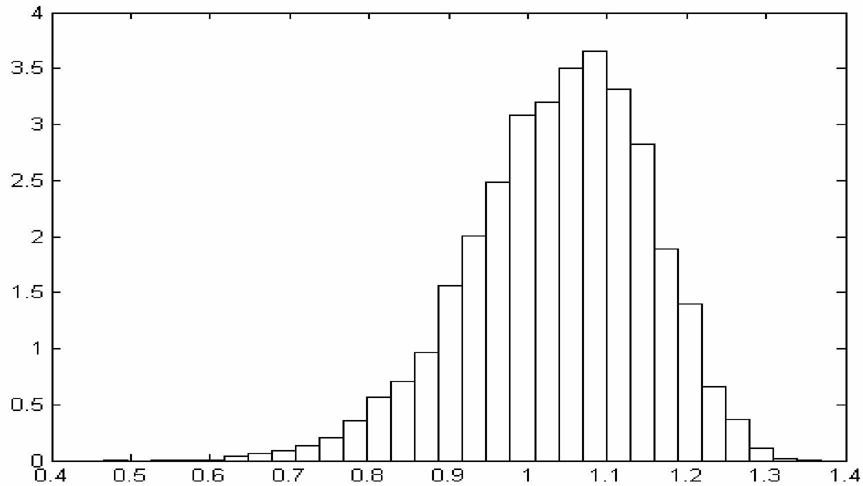


${}_3P_{pl}^1$

Figure 6.4: Histogram of simulated index

The mean of the index ${}_3P_{pl}^1$ is 1.5499 showing that the process is capable.

$p({}_3P_{pl}^{11})$



${}_3P_{pl}^{11}$

Figure 6.5: Histogram of simulated index

The mean of the index ${}_3P_{pl}^{11}$ is 1.0385 showing that the process is capable.

The results of the indices are summarised in the table below.

Table 6.2: Posterior mean, variance and 95% interval for the indices using the non-informative Jeffery's prior

Index	Mean	Variance	95% Interval
${}_3P_{pl}^{1}$	1.5499	0.0925	(0.9861;2.1634)
${}_3P_{pl}^{11}$	1.0385	0.0126	(0.7936;1.2307)
${}_3P_{pl}^{111}$	1.3642	0.0471	(0.9304;1.7672)
${}_3P_{pl}^{1V}$	6.0029	1.3873	(3.8192;8.3787)
${}_3P_{pl}^{V}$	5.8081	0.1059	(5.1653;6.4356)

The corresponding 95% interval in the case of the probability-matching prior for ${}_3P_{pl}^{1}$ is (0.9800;2.1654).

Some features of these results worth noting are:

- The mean of the indices are greater than 1 suggesting that the process is capable.
- The mean of the indices ${}_3P_{pl}^{1V}$ and ${}_3P_{pl}^{V}$ are much greater than 1 suggesting that the process is very capable when a patient takes the tablets for longer periods. In fact the data suggest that the process is super.
- The variance of the index ${}_3P_{pl}^{11}$ is small when compared to the variances of the other indices.

It has been suggested that some advanced patients may require a slightly higher dosage of the drug. The question therefore is whether the process is capable of producing to specification under these same machine settings and for $l_0 = 21,\ 22$ and 23.

$l_0 = 21.$

Table 6.3: Posterior mean, variance and 95% interval for the Indices using the non-informative Jeffery's prior

Index	Mean	Variance	95% Interval
${}_3P_{pl}^{1}$	1.1584	0.0548	(0.7208;1.6302)
${}_3P_{pl}^{11}$	0.7762	0.0086	(0.5744;0.9359)
${}_3P_{pl}^{111}$	1.0196	0.0288	(0.6777;1.3368)
${}_3P_{pl}^{1V}$	4.4864	0.8218	(2.7918;6.3136)
${}_3P_{pl}^{V}$	3.8008	0.0583	(3.3201;4.2691)

The corresponding 95% interval in the case of the probability-matching prior for ${}_3P_{pl}{}^{1}$ is (0.7199;1.6318).

Some features of these results worth noting are:

- The mean of the indices ${}_3P_{pl}{}^{11}$ is less than 1 suggesting that the process is not capable of producing the individual tablets to specification but the other indices are greater than 1 suggesting that the process is capable once the tablets are analysed at least as a batch.
- The variance of the indices ${}_3P_{pl}{}^{11}$ and ${}_3P_{pl}{}^{111}$ are smaller when compared to the variances of the other indices.

$l_0 = 22.$

Table 6.4: Posterior mean, variance and 95% interval for the Indices using the non-informative Jeffery's prior

Index	Mean	Variance	95% Interval
${}_3P_{pl}^{1}$	0.7669	0.0280	(0.4470;1.004)
${}_3P_{pl}^{11}$	0.5138	0.0057	(0.3533;0.6469)
${}_3P_{pl}^{111}$	0.6750	0.0158	(0.4196;0.9129)

${}_3P_{pl}^{1V}$	2.9700	0.4203	(1.7314;4.2619)
${}_3P_{pl}^{V}$	1.7935	0.0315	(1.4390;2.1314)

The corresponding 95% interval in the case of the probability-matching prior for ${}_3P_{pl}{}^1$ is (0.4465;1.1015).

Some features of these results worth noting are:

- The mean of the indices ${}_3P_{pl}{}^1$, ${}_3P_{pl}{}^{11}$ and ${}_3P_{pl}{}^{111}$ are less than 1 suggesting that the process is not capable but the other indices are greater than 1 suggesting that the process is capable. The process is capable once the tablets are taken over longer periods.
- The variance of the indices ${}_3P_{pl}{}^{11}$ and ${}_3P_{pl}{}^{111}$ are smaller when compared to the variances of the other indices.

$l_0 = 23$.

Table 6.5: Posterior mean, variance and 95% interval for the indices using the non-informative Jeffery's prior

Index	Mean	Variance	95% Interval
${}_3P_{pl}^{1}$	0.3753	0.0122	(0.1636;0.5922)
${}_3P_{pl}^{11}$	0.2515	0.0039	(0.1243;0.3668)
${}_3P_{pl}^{111}$	0.3304	0.0081	(0.1525;0.5000)
${}_3P_{pl}^{1V}$	1.4536	0.1827	(0.6338;2.2937)
${}_3P_{pl}^{V}$	-0.2138	0.0254	(-0.5347; 0.0851)

The corresponding 95% interval in the case of the probability-matching prior for ${}_3P_{pl}{}^1$ is (0.1631; 0.5931).

Some features of these results worth noting are:

- The mean of all the indices except ${}_3P_{pl}^{1V}$ are less than 1 suggesting that the process is not capable. The process is capable once the tablets are taken over longer periods of I months
- The variance of the indices ${}_3P_{pl}^{11}$ and ${}_3P_{pl}^{111}$ are smaller when compared to the variances of the other indices.
- The index ${}_3P_{pl}^{V}$ which was been giving the second highest figure all along now suddenly becomes the lowest and even negative. This is because once we know the month, the third variance component falls away and the index is computed using a smaller variance and becomes very sensitive to departures from the mean.

Appendix A6

Proof of theorem 6.3.1

III.

$$\textit{If we define} \quad \overline{\overline{Y_{i..}}} = \frac{\sum_{j=1}^{J}\sum_{k=1}^{K} Y_{ijk}}{KJ}$$

$$E[\overline{\overline{Y_{i..}}} \mid \mu,\sigma_1^2,\sigma_2^2,\sigma_3^2] = E[\frac{\sum_{j=1}^{J}\sum_{k=1}^{K} Y_{ijk}}{KJ} \mid \mu,\sigma_1^2,\sigma_2^2,\sigma_3^2]$$

$$= E[\frac{\sum_{j=1}^{J}\sum_{k=1}^{K} (\mu + r_i + c_{ij} + \varepsilon_{ijk})}{KJ} \mid \mu,\sigma_1^2,\sigma_2^2,\sigma_3^2]$$

$$= E[\frac{KJ\mu}{KJ} \mid \mu,\sigma_1^2,\sigma_2^2,\sigma_3^2] + E[\frac{KJr_i}{KJ} \mid \mu,\sigma_1^2,\sigma_2^2,\sigma_3^2] +$$

$$E[\frac{K\sum_{j=1}^{J}(c_{ij} \mid \mu,\sigma_1^2,\sigma_2^2,\sigma_3^2)}{KJ}] + E[\frac{\sum_{j=1}^{J}\sum_{k=1}^{K}(\varepsilon_{ijk} \mid \mu,\sigma_1^2,\sigma_2^2,\sigma_3^2)}{KJ}]$$

$$= E[\mu \mid \mu,\sigma_1^2,\sigma_2^2,\sigma_3^2] + E[r_i \mid \mu,\sigma_1^2,\sigma_2^2,\sigma_3^2] + [\frac{\sum_{j=1}^{J} E(c_{ij} \mid \mu,\sigma_1^2,\sigma_2^2,\sigma_3^2)}{J}]$$

$$+[\frac{\sum_{j=1}^{J}\sum_{k=1}^{K} E(\varepsilon_{ijk} \mid \mu,\sigma_1^2,\sigma_2^2,\sigma_3^2)}{KJ}]$$

$$= \mu + 0 + 0 + 0 = \mu \, .$$

$$V[\overline{\overline{Y_{i..}}} \mid \mu,\sigma_1^2,\sigma_2^2,\sigma_3^2] = V[\frac{\sum_{j=1}^{J}\sum_{k=1}^{K} Y_{ijk}}{JK} \mid \mu,\sigma_1^2,\sigma_2^2,\sigma_3^2]$$

$$=V[\frac{\sum_{j=1}^{J}\sum_{k=1}^{K}(\mu+r_i+c_{ij}+\varepsilon_{ijk})}{JK}\,|\,\mu,\sigma_1^2,\sigma_2^2,\sigma_3^2]$$

$$=V[\frac{JK\mu}{JK}\,|\,\mu,\sigma_1^2,\sigma_2^2,\sigma_3^2]+V[\frac{JKr_i}{JK}\,|\,\mu,\sigma_1^2,\sigma_2^2,\sigma_3^2]+$$

$$V[\frac{K\sum_{j=1}^{J}(c_{ij}\,|\,\mu,\sigma_1^2,\sigma_2^2,\sigma_3^2)}{JK}]+V[\frac{\sum_{j=1}^{J}\sum_{k=1}^{K}(\varepsilon_{ijk}\,|\,\mu,\sigma_1^2,\sigma_2^2,\sigma_3^2)}{JK}]$$

$$=V[\mu\,|\,\mu,\sigma_1^2,\sigma_2^2,\sigma_3^2]+V[r_i\,|\,\mu,\sigma_1^2,\sigma_2^2,\sigma_3^2]+$$

$$[\frac{\sum_{j=1}^{J}V(c_{ij}\,|\,\mu,\sigma_1^2,\sigma_2^2,\sigma_3^2)}{J^2}]+[\frac{\sum_{j=1}^{J}\sum_{k=1}^{K}V(\varepsilon_{ijk}\,|\,\mu,\sigma_1^2,\sigma_2^2,\sigma_3^2)}{J^2K^2}]$$

$$=0+\sigma_3^2+\frac{J\sigma_2^2}{J^2}+\frac{JK\sigma_1^2}{J^2K^2}=\sigma_3^2+\frac{\sigma_2^2}{J}+\frac{\sigma_1^2}{JK}=\frac{KJ\sigma_3^2+K\sigma_2^2+\sigma_1^2}{JK}.$$

V.

$$Define \quad \overline{\overline{Y}}_{i..}=\frac{\sum_{j=1}^{J}\sum_{k=1}^{K}Y_{ijk}}{KJ}$$

$$E[\overline{\overline{Y}}_{i..}\,|\,\theta,r_i,\sigma_1^2,\sigma_2^2]=E[\frac{\sum_{j=1}^{J}\sum_{k=1}^{K}Y_{ijk}}{KJ}\,|\,\theta,r_i,\sigma_1^2,\sigma_2^2]$$

$$=E[\frac{\sum_{j=1}^{J}\sum_{k=1}^{K}(\theta+r_i+c_{ij}+\varepsilon_{ijk})}{KJ}\,|\,\theta,r_i,\sigma_1^2,\sigma_2^2]$$

$$= E[\frac{KJ\theta}{KJ} \mid \theta, r_i, \sigma_1{}^2, \sigma_2{}^2] + E[\frac{KJr_i}{KJ} \mid \theta, r_i, \sigma_1{}^2, \sigma_2{}^2] +$$

$$E[\frac{K\sum_{j=1}^{J}(c_{ij} \mid \theta, r_i, \sigma_1{}^2, \sigma_2{}^2)}{KJ}] + E[\frac{\sum_{j=1}^{J}\sum_{k=1}^{K}(\varepsilon_{ijk} \mid \theta, r_i, \sigma_1{}^2, \sigma_2{}^2)}{KJ}]$$

$$= E[\theta \mid \theta, r_i, \sigma_1{}^2, \sigma_2{}^2] + E[r_i \mid \theta, r_i, \sigma_1{}^2, \sigma_2{}^2] + [\frac{\sum_{j=1}^{J} E(c_{ij} \mid \theta, r_i, \sigma_1{}^2, \sigma_2{}^2)}{J}]$$

$$+ [\frac{\sum_{j=1}^{J}\sum_{k=1}^{K} E(\varepsilon_{ijk} \mid \theta, r_i, \sigma_1{}^2, \sigma_2{}^2)}{KJ}]$$

$$= \theta + r_i + 0 + 0$$

$$= \theta + r_i \, .$$

$$V[\overline{\overline{Y}}_{i..} \mid \theta, r_i, \sigma_1{}^2, \sigma_2{}^2] = V[\frac{\sum_{j=1}^{J}\sum_{k=1}^{K} Y_{ijk}}{JK} \mid \theta, r_i, \sigma_1{}^2, \sigma_2{}^2]$$

$$= V[\frac{\sum_{j=1}^{J}\sum_{k=1}^{K}(\theta + r_i + c_{ij} + \varepsilon_{ijk})}{JK} \mid \theta, r_i, \sigma_1{}^2, \sigma_2{}^2]$$

$$= V[\frac{JK\theta}{JK} \mid \theta, r_i, \sigma_1{}^2, \sigma_2{}^2] + V[\frac{JKr_i}{JK} \mid \theta, r_i, \sigma_1{}^2, \sigma_2{}^2] +$$

$$V[\frac{K\sum_{j=1}^{J}(c_{ij} \mid \theta, r_i, \sigma_1{}^2, \sigma_2{}^2)}{JK}] + V[\frac{\sum_{j=1}^{J}\sum_{k=1}^{K}(\varepsilon_{ijk} \mid \theta, r_i, \sigma_1{}^2, \sigma_2{}^2)}{JK}]$$

$$= V[\theta \mid \theta, r_i, \sigma_1{}^2, \sigma_2{}^2] + V[r_i \mid \theta, r_i, \sigma_1{}^2, \sigma_2{}^2] +$$

$$[\frac{\sum_{j=1}^{J} V(c_{ij} \mid \theta, r_i, \sigma_1{}^2, \sigma_2{}^2)}{J^2}] + [\frac{\sum_{j=1}^{J}\sum_{k=1}^{K} V(\varepsilon_{ijk} \mid \theta, r_i, \sigma_1{}^2, \sigma_2{}^2)}{J^2K^2}]$$

$$= 0 + 0 + \frac{J\sigma_2{}^2}{J^2} + \frac{JK\sigma_1{}^2}{J^2K^2}$$

$$= \frac{\sigma_2{}^2}{J} + \frac{\sigma_1{}^2}{JK}$$

$$= \frac{\sigma_2{}^2}{J} + \frac{\sigma_1{}^2}{JK}$$

$$= \frac{K\sigma_2{}^2 + \sigma_1^2}{JK}.$$

Proof of theorem 6.4.1

The non-informative joint prior:

$$p(\mu,\sigma_1{}^2,\sigma_{12}{}^2,\sigma_{123}{}^2) \propto p(\mu)\,p(\sigma_1{}^2,\sigma_{12}{}^2,\sigma_{123})$$

$$= c \times p(\sigma_1{}^2,\sigma_{12}{}^2,\sigma_{123})$$

$$\propto \sigma_1^{-2}(\sigma_1^2 + K\sigma_2^2)^{-1}(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)^{-1},$$

The joint posterior distribution of μ and $\sigma_1{}^2,\sigma_2{}^2$ and σ_3^2 can be worked out.

The posterior is computed as follows:

Let $\underline{Y} = [Y_{111,}Y_{112,}Y_{113,}\cdots Y_{IJK}]$

Posterior $\propto$ *likelihood x prior*

$$p\left(\mu,\sigma_1{}^2,\sigma_{12}{}^2,\sigma_{123} \mid \underline{Y}\right) \propto \prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{k=1}^{K} f(Y_{ijk} \mid \mu,\sigma_1{}^2,\sigma_2{}^2,\sigma_3{}^2)\,p\left(\mu,\sigma_1{}^2,\sigma_{12}{}^2,\sigma_{123}\right)$$

$$= \left[(\sigma_1{}^2+\sigma_2{}^2+\sigma_3{}^2)^{\frac{-IJK}{2}}\exp(-\frac{1}{2}\{\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(Y_{ijk}-\mu)^2}{(\sigma_1{}^2+\sigma_2{}^2+\sigma_3{}^2)}\}\right]\frac{1}{\sigma_1^2(\sigma_1^2+K\sigma_2^2)(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2)}$$

$$\propto \sigma_1^{-\frac{1}{2}v_1}(\sigma_1{}^2+K\sigma_2{}^2)^{-\frac{1}{2}v_2}(\sigma_1{}^2+K\sigma_2{}^2+JK\sigma_3{}^2)^{-\frac{1}{2}(v_3+1)}\times$$

$$\left[\exp(-\frac{1}{2}\{\frac{v_1m_1}{\sigma_1{}^2}+\frac{v_2m_2}{(\sigma_1{}^2+K\sigma_2{}^2)}+\frac{v_3m_3}{(\sigma_1{}^2+K\sigma_2{}^2+JK\sigma_3{}^2)}+\frac{IJK(\overline{\overline{\overline{Y}}}_{\ldots}-\mu)^2}{(\sigma_1{}^2+K\sigma_2{}^2+JK\sigma_3{}^2)}\}\right]\times$$

$$\frac{1}{\sigma_1^2(\sigma_1^2+K\sigma_2^2)(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2)}$$

$$\propto (\sigma_1^2)^{-\frac{1}{2}(v_1+2)}(\sigma_1^2+K\sigma_2^2)^{-\frac{1}{2}(v_2+2)}(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2)^{-\frac{1}{2}(v_3+3)}\times$$

$$\exp\left\{-\frac{1}{2}\left[\frac{IJK(\overline{\overline{\overline{Y}}}_{\ldots}-\theta)^2}{\sigma_1^2+K\sigma_2^2+JK\sigma_3^2}+\frac{v_3m_3}{\sigma_1^2+K\sigma_2^2+JK\sigma_3^2}+\frac{v_2m_2}{\sigma_1^2+K\sigma_2^2}+\frac{v_1m_1}{\sigma_1^2}\right]\right\}.$$

This is the joint posterior of all the parameters.

$$p\left(\mu,\sigma_1{}^2,\sigma_{12}{}^2,\sigma_{123}{}^2 \mid \underline{Y}\right) \propto \left[(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2)^{\frac{-1}{2}} \exp(-\frac{1}{2}\frac{IJK(\overline{\overline{\overline{Y}}}_{...}-\mu)^2}{(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2)})\right]$$

$$\left((\sigma_1^2)^{-\frac{1}{2}(\nu_1+2)}(\sigma_1^2+K\sigma_2^2)^{-\frac{1}{2}(\nu_2+2)}(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2)^{-\frac{1}{2}(\nu_3+2)}\times \exp(-\frac{1}{2}\{\frac{\nu_3 m_3}{\sigma_1^2+K\sigma_2^2+JK\sigma_3^2}+\frac{\nu_2 m_2}{\sigma_1^2+K\sigma_2^2}+\frac{\nu_1 m_1}{\sigma_1^2}\})\right)$$

The function inside the curled brackets on the right hand side is therefore proportional to the marginal joint distribution of the variance components $\sigma_1{}^2,\sigma_2{}^2,\sigma_3{}^2$.

The joint posterior distribution of the variance components is given by

$$p(\sigma_1{}^2,\sigma_{12}{}^2,\sigma_{123}{}^2 \mid \underline{Y}) \propto \quad (\sigma_1^2)^{-\frac{1}{2}(\nu_1+2)}(\sigma_1^2+K\sigma_2^2)^{-\frac{1}{2}(\nu_2+2)}(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2)^{-\frac{1}{2}(\nu_3+2)}\times \exp\left\{-\frac{1}{2}\left[\frac{\nu_3 m_3}{\sigma_1^2+K\sigma_2^2+JK\sigma_1^2}+\frac{\nu_2 m_2}{\sigma_1^2+K\sigma_2^2}+\frac{\nu_1 m_1}{\sigma_1^2}\right]\right\}$$

$$p\left(\sigma_1{}^2,\sigma_{12}{}^2,\sigma_{123}{}^2 \mid \underline{Y}\right)=\left((\sigma_1{}^2)^{\frac{-(\frac{1}{2}\nu_1+2)}{2}}\exp(-\frac{1}{2}\{\frac{\nu_1 m_1}{\sigma_1{}^2}\})\times(\sigma_1{}^2+K\sigma_2{}^2)^{\frac{-\frac{1}{2}(\nu_2+2)}{2}}\exp(-\frac{1}{2}\{\frac{\nu_2 m_2}{(\sigma_1{}^2+K\sigma_2{}^2)}\}\times (\sigma_1{}^2+K\sigma_2{}^2+JK\sigma_3{}^2)^{\frac{-\frac{1}{2}(\nu_3+2)}{2}}\exp(-\frac{1}{2}\{\frac{\nu_3 m_3}{(\sigma_1{}^2+K\sigma_2{}^2+JK\sigma_3{}^2)}\})\right)$$

i.e.

$$p\left(\sigma_1{}^2,\sigma_{12}{}^2,\sigma_{123}{}^2 \mid \underline{Y}\right)=IG(\sigma_1{}^2 \mid \frac{\nu_1}{2};\nu_1 m_1)\times IG(\sigma_1{}^2+K\sigma_2{}^2 \mid \frac{\nu_2}{2};\nu_2 m_2)\times IG(\sigma_1{}^2+K\sigma_2{}^2+JK\sigma_3{}^2 \mid \frac{\nu_3}{2};\nu_3 m_3)$$

where

$$IG(x \mid \alpha,\beta)=\frac{\beta^\alpha}{\Gamma(\alpha)}x^{-\alpha-1}\exp(-\beta/x)$$

i.e. the inverted gamma density with positive parameters α and β.

The function inside the square brackets just above, when regarded as a function of μ, is proportional to the conditional distribution of a normal distribution for which the mean is:

$$E(\mu \mid \underline{Y}, \sigma_1^2, \sigma_2^2, \sigma_3^2) = \overline{\overline{\overline{Y}}}_{\ldots}$$

and variance:

$$\mathrm{Var}(\mu \mid \underline{Y}, \sigma_1^2, \sigma_2^2, \sigma_3^2) = \frac{\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2}{IJK}$$

i.e. $\mu \mid \underline{Y} \sim N(\overline{\overline{\overline{Y}}}_{\ldots}, \dfrac{\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2}{IJK})$.

Proof of theorem 6.4.2

III. We show that $E\left(m_3\right) = (\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)$

$$v_3 m_3 = JK\sum_{i=1}^{I}(\overline{\overline{Y}}_{i..} - \overline{\overline{\overline{Y}}}_{\ldots})^2$$

$$E(v_3 m_3) = JK\sum_{i=1}^{I}E(\overline{\overline{Y}}_{i..} - \overline{\overline{\overline{Y}}}_{\ldots})^2$$

$$E\left(v_3 m_3\right) = JK\sum_{i=1}^{I}E\left((\overline{\overline{Y}}_{i..} - \mu) - (\overline{\overline{\overline{Y}}}_{\ldots} - \mu) \right)^2$$

$$E\left(v_3 m_3\right) = JK\sum_{i=1}^{I}E\left((\overline{\overline{Y}}_{i..} - \mu)^2 + (\overline{\overline{\overline{Y}}}_{\ldots} - \mu)^2 - 2(\overline{\overline{Y}}_{i..} - \mu)(\overline{\overline{\overline{Y}}}_{\ldots} - \mu) \right)$$

$$E\left(v_3 m_3\right) = JK\sum_{i=1}^{I}\left(E(\overline{\overline{Y}}_{i..} - \mu)^2 + E((\overline{\overline{\overline{Y}}}_{\ldots} - \mu)^2)^2 - 2E(\overline{\overline{Y}}_{i..} - \mu)((\overline{\overline{\overline{Y}}}_{\ldots} - \mu)^2) \right)$$

$$E\left(v_3 m_3\right) = JK\sum_{i=1}^{I}\left(\begin{array}{l} Var(\overline{\overline{Y}}_{i..} \mid \mu, \sigma_1^2, \sigma_2^2, \sigma_3^2) + Var(\overline{\overline{\overline{Y}}}_{\ldots} \mid \mu, \sigma_1^2, \sigma_2^2, \sigma_3^2) \\ \qquad - 2Cov(\overline{\overline{Y}}_{i..} \mid \mu, \sigma_1^2, \sigma_2^2, \sigma_3^2, \overline{\overline{\overline{Y}}}_{\ldots} \mid \mu, \sigma_1^2, \sigma_2^2, \sigma_3^2) \end{array} \right)$$

$$E\left(v_3 m_3\right) = JK\sum_{i=1}^{I}\left(\frac{(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)}{JK} + \frac{(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)}{IJK} - 2\frac{(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)}{IJK} \right)$$

$$E\left(v_3 m_3\right) = JK\left(\frac{I(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)}{JK} + \frac{I(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)}{IJK} - 2\frac{I(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)}{IJK} \right)$$

$$E\left(v_3 m_3\right) = \left(I(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2) + (\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2) - 2(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2) \right)$$

$$E\left(v_3 m_3\right) = \left(I(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2) - (\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2) \right)$$

$$E\left(v_3 m_3\right) = (\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)\left(I - 1\right)$$

$$E\left(v_3 m_3\right) = (\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)v_3 \quad \sin ce \quad v_3 = I - 1$$

$$E\left(m_3\right)=\frac{(\sigma_1{}^2+K\sigma_2{}^2+JK\sigma_3{}^2)\nu_3}{\nu_3}=(\sigma_1{}^2+K\sigma_2{}^2+JK\sigma_3{}^2).$$

Proof of theorem 6.4.3

We want to derive the posterior distribution of $\mu_i=\mu+r_i \quad (i=1,\ldots,I)$ given $\underline{Y}$ and the variance components σ_1^2,σ_2^2 and σ_3^2. To do this we derive the posterior distribution of μ_i given $\mu,\sigma_1^2,\ \sigma_2^2,\sigma_3^2$ and $\underline{Y}$ which is normally distributed.

$$Y_{ijk}=\mu+r_i+c_{ij}+\varepsilon_{ijk} \quad i=1,\ldots,I, \quad j=1,\ldots,J \quad \text{and} \quad k=1,\ldots,K.$$

where

$$r_i\sim N(0,\sigma_3^2), \quad c_{ij}\sim N(0,\sigma_2^2) \quad \text{and} \quad \varepsilon_{ijk}\sim N(0,\sigma_1^2).$$

Therefore

$$\mu+r_i\sim N(\mu,\sigma_3{}^2) \quad or \quad simply \quad \mu_i\sim N(\mu,\sigma_3{}^2) \quad \sin ce \quad \mu_i=\mu+r_i\,.$$

The posterior distribution of $\mu_i\,|\,\underline{Y},\mu,\sigma_1{}^2,\sigma_2{}^2$ is now calculated as follows

$$p(\mu_i\,|\,\underline{Y},\mu,\sigma_1{}^2,\sigma_2{}^2)\propto\prod_{j=1}^{J}\prod_{k=1}^{K}f(Y_{ijk}\,|\,\mu_i)p\left(\mu_i\right)$$

$$\propto\prod_{j=1}^{J}\prod_{k=1}^{K}\exp\Big(-\frac{1}{2}\frac{(Y_{ijk}-\mu_i)^2}{(\sigma_1{}^2+\sigma_2{}^2)}\Big)\exp\Big(-\frac{1}{2}\frac{(\mu_i-\mu)^2}{\sigma_3{}^2}\Big)$$

$$=\exp\Big(-\frac{1}{2}\frac{\sum_{j=1}^{J}\sum_{k=1}^{K}(Y_{ijk}-\mu_i)^2}{(\sigma_1{}^2+\sigma_2{}^2)}\Big)\exp\Big(-\frac{1}{2}\frac{(\mu_i-\mu)^2}{\sigma_3{}^2}\Big)$$

$$=\exp\left(-\frac{1}{2}\left\{\frac{\sum_{j=1}^{J}\sum_{k=1}^{K}(Y_{ijk}-\mu_i)^2}{(\sigma_1{}^2+\sigma_2{}^2)}+\frac{(\mu_i-\mu)^2}{\sigma_3{}^2}\right\}\right)$$

but

$$\sum_{j=1}^{J}\sum_{k=1}^{K}(Y_{ijk}-\mu_i)^2=\sum_{j=1}^{J}\sum_{k=1}^{K}\left((Y_{ijk}-\overline{\overline{Y}}_{i..})+(\overline{\overline{Y}}_{i..}-\mu_i)\right)^2$$

$$=\sum_{j=1}^{J}\sum_{k=1}^{K}\left((Y_{ijk}-\overline{\overline{Y}}_{i..})^2+(\overline{\overline{Y}}_{i..}-\mu_i)^2+2(Y_{ijk}-\overline{\overline{Y}}_{i..})(\overline{\overline{Y}}_{i..}-\mu_i)\right)$$

$$= \sum_{j=1}^{J}\sum_{k=1}^{K}(Y_{ijk} - \overline{\overline{Y}}_{i..})^2 + \sum_{j=1}^{J}\sum_{k=1}^{K}(\overline{\overline{Y}}_{i..} - \mu_i)^2 + 2\sum_{j=1}^{J}\sum_{k=1}^{K}(Y_{ijk} - \overline{\overline{Y}}_{i..})(\overline{\overline{Y}}_{i..} - \mu_i)$$

$$= \sum_{j=1}^{J}\sum_{k=1}^{K}(Y_{ijk} - \overline{\overline{Y}}_{i..})^2 + JK(\overline{\overline{Y}}_{i..} - \mu_i)^2 + 0$$

$$= \sum_{j=1}^{J}\sum_{k=1}^{K}(Y_{ijk} - \overline{\overline{Y}}_{i..})^2 + JK(\overline{\overline{Y}}_{i..} - \mu_i)^2 .$$

Therefore

$$p(\mu_i \mid \underline{Y}, \mu, \sigma_1^2, \sigma_2^2, \sigma_3^2) \propto \exp\left(-\frac{1}{2}\left\{ \frac{\sum_{j=1}^{J}\sum_{k=1}^{K}(Y_{ijk} - \overline{\overline{Y}}_{i..})^2}{(\sigma_1^2 + \sigma_2^2)} + \frac{JK(\overline{\overline{Y}}_{i..} - \mu_i)^2}{(\sigma_1^2 + K\sigma_2^2)} + \frac{(\mu_i - \mu)^2}{\sigma_3^2} \right\} \right)$$

We omit all factors that involve Y_{ij} but do not depend on μ_i.

Therefore:

$$p(\mu_i \mid \underline{Y}, \mu, \sigma_1^2, \sigma_2^2) \propto \exp\left(-\frac{1}{2}\left\{ \frac{JK(\overline{\overline{Y}}_{i..} - \mu_i)^2}{(\sigma_1^2 + K\sigma_2^2)} + \frac{(\mu_i - \mu)^2}{\sigma_3^2} \right\} \right)$$

$$\frac{JK(\overline{\overline{Y}}_{i..} - \mu_i)^2}{(\sigma_1^2 + K\sigma_2^2)} + \frac{(\mu_i - \mu)^2}{\sigma_3^2} = \frac{JK(\overline{\overline{Y}}_{i..}^{\,2} + \mu_i^2 - 2\overline{\overline{Y}}_{i..}\mu_i)}{(\sigma_1^2 + K\sigma_2^2)} + \frac{(\mu_i^2 + \mu^2 - 2\mu_i\mu)}{\sigma_3^2}$$

$$= \mu_i^2\left(\frac{JK}{(\sigma_1^2 + K\sigma_2^2)} + \frac{1}{\sigma_3^2}\right) - 2\mu_i\left(\frac{J\overline{\overline{Y}}_{i..}}{(\sigma_1^2 + K\sigma_2^2)} + \frac{\mu}{\sigma_3^2}\right) + \frac{J\overline{\overline{Y}}_{i..}^{\,2}}{(\sigma_1^2 + K\sigma_2^2)} + \frac{\mu^2}{\sigma_3^2}$$

$$= \mu_i^2\left(\frac{(\sigma_1^2 + K\sigma_2^2) + JK\sigma_3^2}{(\sigma_1^2 + K\sigma_2^2)\sigma_3^2}\right) - 2\mu_i\left(\frac{(\sigma_1^2 + K\sigma_2^2)\mu + J\overline{\overline{Y}}_{i..}\sigma_3^2}{(\sigma_1^2 + K\sigma_2^2)\sigma_3^2}\right) + \frac{J\overline{\overline{Y}}_{i..}^{\,2}}{(\sigma_1^2 + K\sigma_2^2)} + \frac{\mu^2}{\sigma_3^2}$$

$$=\frac{1}{\frac{(\sigma_1^2+K\sigma_2^2)\sigma_3^2}{(\sigma_1^2+K\sigma_2^2+JK\sigma^2{}_3)}}\left[\mu_i^2-2\mu_i(\frac{(\sigma_1^2+K\sigma_2^2)\mu+JK\overline{\overline{Y}}_{i..}\sigma_3^2}{(\sigma_1^2+K\sigma_2^2)+JK\sigma_3^2})\right]+\frac{JK\overline{\overline{Y}}_{i..}^2}{(\sigma_1^2+K\sigma_2^2)}+\frac{\mu^2}{\sigma_3^2}$$

$$=\frac{1}{\frac{(\sigma_1^2+K\sigma_2^2)\sigma_3^2}{(\sigma_1^2+K\sigma_2^2+JK\sigma^2{}_3)}}\left[\begin{array}{l}\mu_i^2-2\mu_i(\frac{(\sigma_1^2+K\sigma_2^2)\mu+JK\overline{\overline{Y}}_{i..}\sigma_3^2}{(\sigma_1^2+K\sigma_2^2)+JK\sigma_3^2})+\\ \qquad\qquad\left(\frac{(\sigma_1^2+K\sigma_2^2)\mu+JK\overline{\overline{Y}}_{i..}\sigma_3^2}{(\sigma_1^2+K\sigma_2^2)+JK\sigma_3^2}\right)^2\end{array}\right]$$

$$+\frac{J\overline{Y}_{i.}^2}{\sigma_1^2}+\frac{\mu^2}{\sigma_2^2}-\left(\frac{(\sigma_1^2+K\sigma_2^2)\mu+JK\overline{\overline{Y}}_{i..}\sigma_3^2}{(\sigma_1^2+K\sigma_2^2)+JK\sigma_3^2}\right)^2$$

$$=\frac{1}{\frac{(\sigma_1^2+K\sigma_2^2)\sigma_3^2}{(\sigma_1^2+K\sigma_2^2+JK\sigma^2{}_3)}}\left[\mu_i-(\frac{(\sigma_1^2+K\sigma_2^2)\mu+JK\overline{\overline{Y}}_{i..}\sigma_3^2}{(\sigma_1^2+K\sigma_2^2)+JK\sigma_3^2})\right]^2+\textit{other terms}$$

not involving μ_i.

If we drop all terms which do not involve μ_i in the expression for $p(\mu_i\,|\,\underline{Y},\mu,\sigma_1^2,\sigma_2^2,\sigma_3^2)$ we get

$$p(\mu_i\,|\,\underline{Y},\mu,\sigma_1^2,\sigma_2^2,\sigma_3^2)\propto\exp\left(-\frac{1}{2}\left\{\frac{\left[\mu_i-(\frac{(\sigma_1^2+K\sigma_2^2)\mu+JK\overline{\overline{Y}}_{i..}\sigma_3^2}{(\sigma_1^2+K\sigma_2^2)+JK\sigma_3^2})\right]^2}{\frac{(\sigma_1^2+K\sigma_2^2)\sigma_3^2}{(\sigma_1^2+K\sigma_2^2+JK\sigma^2{}_3)}}\right\}\right)$$

This is proportional to a normal distribution with mean $(\frac{(\sigma_1^2+K\sigma_2^2)\mu+JK\sigma_3^2\overline{\overline{Y}}_{i..}}{(\sigma_1^2+K\sigma_2^2)+JK\sigma_3^2})$

and variance $\frac{(\sigma_1^2+K\sigma_2^2)\sigma_3^2}{(\sigma_1^2+K\sigma_2^2+JK\sigma^2{}_3)}$

$$\textit{i.e.}\quad \mu_i\,|\,\underline{Y},\mu,\sigma_1^2,\sigma_2^2,\sigma_3^2\sim N\left((\frac{(\sigma_1^2+K\sigma_2^2)\mu+JK\sigma_3^2\overline{\overline{Y}}_{i..}}{(\sigma_1^2+K\sigma_2^2)+JK\sigma_3^2}),\frac{(\sigma_1^2+K\sigma_2^2)\sigma_3^2}{(\sigma_1^2+K\sigma_2^2+JK\sigma^2{}_3)}\right)$$

We will now derive the posterior distribution of $\mu_i = \mu + r_i \quad (i = 1,\ldots, I)$ given $\underline{Y}$ and the variance components σ_1^2, σ_2^2 and σ_3^2. To do this we firstly appeal to the derived posterior distribution of μ_i given μ, σ_1^2, σ_2^2, σ_3^2 and $\underline{Y}$ which is normal with mean

$$E(\mu_i \mid \mu, \sigma_1^2, \sigma_2^2, {\sigma_3}^2, \underline{Y}) = \frac{JK{\sigma_3}^2 \overline{\overline{Y_{i..}}}}{({\sigma_1}^2 + K{\sigma_2}^2) + JK{\sigma_3}^2} + \frac{({\sigma_1}^2 + K{\sigma_2}^2)\mu}{({\sigma_1}^2 + K{\sigma_2}^2) + JK{\sigma_3}^2}$$

and variance

$$\mathrm{Var}(\mu_i \mid \mu, \sigma_1^2, \sigma_2^2, {\sigma^2}_3, \underline{Y}) = \frac{({\sigma_1}^2 + K{\sigma_2}^2){\sigma_3}^2}{({\sigma_1}^2 + K{\sigma_2}^2 + JK{\sigma^2}_3)}$$

Since $\mu \mid Y, \sigma_1^2, \sigma_2^2, \sigma_3^2 \sim N\left\{\overline{\overline{\overline{Y}}}_{...}, \frac{\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2}{IJK}\right\}$, it follows that μ_i given σ_1^2, σ_2^2 and σ_3^2 will be normal with mean

$$E(\mu_i \mid \sigma_1^2, \sigma_2^2, {\sigma_3}^2, \underline{Y}) = \frac{JK{\sigma_3}^2 \overline{\overline{Y_{i..}}}}{({\sigma_1}^2 + K{\sigma_2}^2) + JK{\sigma_3}^2} + \frac{({\sigma_1}^2 + K{\sigma_2}^2)}{({\sigma_1}^2 + K{\sigma_2}^2) + JK{\sigma_3}^2} \overline{\overline{\overline{Y}}}_{...}$$

and variance

$$\mathrm{Var}(\mu_i \mid \underline{Y}, \sigma_1^2, \sigma_2^2, \sigma_3^2) = E_\mu\left\{\mathrm{Var}(\mu_i \mid \underline{Y}, \mu, \sigma_1^2, \sigma_2^2, \sigma_3^2)\right\} + \mathrm{Var}_\mu\left\{E(\mu_i \mid \underline{Y}, \mu, \sigma_1^2, \sigma_2^2, \sigma_3^2)\right\}$$

$$\mathrm{Var}(\mu_i \mid \underline{Y}, \sigma_1^2, \sigma_2^2, \sigma_3^2) = E_\mu\left\{\frac{({\sigma_1}^2 + K{\sigma_2}^2){\sigma_3}^2}{({\sigma_1}^2 + K{\sigma_2}^2 + JK{\sigma^2}_3)}\right\} + \mathrm{Var}_\mu\left\{\frac{JK{\sigma_3}^2 \overline{\overline{Y_{i..}}}}{({\sigma_1}^2 + K{\sigma_2}^2) + JK{\sigma_3}^2} + \frac{({\sigma_1}^2 + K{\sigma_2}^2)\mu}{({\sigma_1}^2 + K{\sigma_2}^2) + JK{\sigma_3}^2}\right\}$$

$$= \frac{({\sigma_1}^2 + K{\sigma_2}^2){\sigma_3}^2}{({\sigma_1}^2 + K{\sigma_2}^2 + JK{\sigma^2}_3)} + \left(\frac{({\sigma_1}^2 + K{\sigma_2}^2)}{({\sigma_1}^2 + K{\sigma_2}^2 + JK{\sigma^2}_3)}\right)^2 \frac{\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2}{IJK}$$

$$= \frac{({\sigma_1}^2 + K{\sigma_2}^2){\sigma_3}^2}{({\sigma_1}^2 + K{\sigma_2}^2 + JK{\sigma^2}_3)} + \left(\frac{({\sigma_1}^2 + K{\sigma_2}^2)^2}{IJK({\sigma_1}^2 + K{\sigma_2}^2 + JK{\sigma^2}_3)}\right)$$

$$= \frac{({\sigma_1}^2 + K{\sigma_2}^2)}{({\sigma_1}^2 + K{\sigma_2}^2 + JK{\sigma^2}_3)}\left\{\sigma_3^2 + \frac{({\sigma_1}^2 + K{\sigma_2}^2)}{IJK}\right\}$$

$$= \frac{({\sigma_1}^2 + K{\sigma_2}^2)}{({\sigma_1}^2 + K{\sigma_2}^2 + JK{\sigma^2}_3)}\left\{\frac{IJK\sigma_3^2 + {\sigma_1}^2 + K{\sigma_2}^2}{IJK}\right\}.$$

Proof of theorem 6.5.1

$\sin ce$ $$\mu \mid \underline{Y}, \sigma_1^2, \sigma_2^2, \sigma_3^2 \sim N\left\{\overline{\overline{\overline{Y}}}_{\ldots}, \frac{\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2}{IJK}\right\}$$

$$\frac{\mu - l_0}{\left(\frac{\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2}{JK}\right)^{\frac{1}{2}}} \mid \underline{Y}, \sigma_1^2, \sigma_2^2, \sigma_3^2 \sim N\left(\frac{\overline{\overline{\overline{Y}}}_{\ldots} - l_0}{\left(\frac{\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2}{JK}\right)^{\frac{1}{2}}}, \quad \frac{1}{I}\right)$$

$$\frac{\mu - l_0}{3\left(\frac{\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2}{JK}\right)^{\frac{1}{2}}} \mid \underline{Y}, \sigma_1^2, \sigma_2^2, \sigma_3^2 \sim N\left(\frac{\overline{\overline{\overline{Y}}}_{\ldots} - l_0}{\left(\frac{\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2}{JK}\right)^{\frac{1}{2}}}, \quad \frac{1}{9I}\right)$$

$$\frac{\mu - l_0}{3\left(\frac{(\sigma_{123}{}^2)}{JK}\right)^{\frac{1}{2}}} \mid \underline{Y}, \sigma_1^2, \sigma_2^2, \sigma_3^2 \sim N\left(\frac{\overline{\overline{\overline{Y}}}_{\ldots} - l_0}{\left(\frac{\sigma_{123}{}^2}{JK}\right)^{\frac{1}{2}}}, \quad \frac{1}{9I}\right) \sin ce\ \sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2 = \sigma_{123}{}^2$$

$$_3P_{pl}^1 \mid \underline{Y}, \sigma_1^2, \sigma_2^2, \sigma_3^2 \sim N\left(\frac{\overline{\overline{\overline{Y}}}_{\ldots} - l_0}{\left(\frac{\sigma_{123}{}^2}{JK}\right)^{\frac{1}{2}}}, \quad \frac{1}{9I}\right).$$

Proof of theorem 6.6.1

The integrated likelihood function is given by

$$L(\mu, \sigma_1^2, \sigma_2^2, \sigma_3^2 \mid \underline{Y}) \propto (\sigma_1^2)^{-\frac{1}{2}\nu_1} (\sigma_1^2 + K\sigma_2^2)^{-\frac{1}{2}\nu_2} (\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)^{-\frac{1}{2}(\nu_3+1)} \times$$

$$\exp\left\{-\frac{1}{2}\left[\frac{JK\sum_{i=1}^{I}(\overline{\overline{Y}}_{i..} - \mu)^2}{\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2} + \frac{\nu_2 m_2}{\sigma_1^2 + K\sigma_2^2} + \frac{\nu_1 m_1}{\sigma_1^2}\right]\right\}$$

$$\propto (\sigma_1^2)^{-\frac{1}{2}\nu_1} (\sigma_1^2 + K\sigma_2^2)^{-\frac{1}{2}\nu_2} (\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)^{-\frac{1}{2}(\nu_3+1)} \times$$

$$\exp\left\{-\frac{1}{2}\left[\frac{IJK(\overline{\overline{\overline{Y}}}_{...} - \mu)^2}{\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2} + \frac{\nu_3 m_3}{\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2} + \frac{\nu_2 m_2}{\sigma_1^2 + K\sigma_2^2} + \frac{\nu_1 m_1}{\sigma_1^2}\right]\right\}.$$

The Fisher information matrix is obtained by differentiating $\log L(\mu, \sigma_1^2, \sigma_2^2, \sigma_3^2 \mid \underline{Y})$ twice with respect to the unknown parameters and taking minus the expected values.

The Fisher information matrix is given as:

$$F(\mu, \sigma_1^2, \sigma_2^2, \sigma_3^2) = \begin{bmatrix} F_{11} & F_{12} & F_{13} & F_{14} \\ F_{21} & F_{22} & F_{23} & F_{24} \\ F_{31} & F_{32} & F_{33} & F_{34} \\ F_{41} & F_{42} & F_{43} & F_{44} \end{bmatrix}$$

where

$$L(\mu, \sigma_1^2, \sigma_2^2, \sigma_3^2 \mid \underline{Y}) \propto (\sigma_1^2)^{-\frac{1}{2}\nu_1} (\sigma_1^2 + K\sigma_2^2)^{-\frac{1}{2}\nu_2} (\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)^{-\frac{1}{2}(\nu_3+1)} \times$$
$$\exp\left\{-\frac{1}{2}\left[\frac{IJK(\overline{\overline{\overline{Y}}}_{...} - \mu)^2}{\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2} + \frac{\nu_3 m_3}{\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2} + \frac{\nu_2 m_2}{\sigma_1^2 + K\sigma_2^2} + \frac{\nu_1 m_1}{\sigma_1^2}\right]\right\}$$

$$\ln L(\mu, \sigma_1^2, \sigma_2^2, \sigma_3^2 \mid \underline{Y}) = \ell \propto \ln[(\sigma_1^2)^{-\frac{1}{2}\nu_1} (\sigma_1^2 + K\sigma_2^2)^{-\frac{1}{2}\nu_2} (\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)^{-\frac{1}{2}(\nu_3+1)} \times$$
$$\exp\left\{-\frac{1}{2}\left[\frac{IJK(\overline{\overline{\overline{Y}}}_{...} - \mu)^2}{\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2} + \frac{\nu_3 m_3}{\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2} + \frac{\nu_2 m_2}{\sigma_1^2 + K\sigma_2^2} + \frac{\nu_1 m_1}{\sigma_1^2}\right]\right\}]$$

$$\ell = -\frac{\nu_1}{2}\ln(\sigma_1^2) - \frac{\nu_2}{2}\ln(\sigma_1^2 + K\sigma_2^2) - \frac{(\nu_3+1)}{2}\ln(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)$$
$$-\frac{1}{2}\left[\frac{IJK(\overline{\overline{\overline{Y}}}_{...} - \mu)^2}{\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2} + \frac{\nu_3 m_3}{\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2} + \frac{\nu_2 m_2}{\sigma_1^2 + K\sigma_2^2} + \frac{\nu_1 m_1}{\sigma_1^2}\right]$$

$$\frac{\partial \ell}{\partial \mu} = \frac{IJK(\overline{\overline{\overline{Y}}}_{...} - \mu)}{\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2}$$

$$\frac{\partial \ell}{\partial \mu^2} = -\frac{IJK(\overline{\overline{\overline{Y}}}_{...} - \mu)^0}{\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2} = -\frac{IJK}{\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2}$$

since $\nu_3 = I - 1$

$$\frac{\partial \ell}{\partial \mu^2} = -\frac{(\nu_3+1)JK}{\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2} \text{ and } -E\left(\frac{\partial \ell}{\partial \mu^2}\right) = F_{11} = \frac{(\nu_3+1)JK}{\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2}.$$

Further

$$\frac{\partial \ell}{\partial \sigma_1^2} = -\frac{\nu_1}{2\sigma_1^2} - \frac{\nu_2}{2(\sigma_1^2 + K\sigma_2^2)} - \frac{(\nu_3+1)}{2(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)} + \frac{1}{2}\frac{IJK(\overline{\overline{\overline{Y}}}_{\ldots} - \mu)^2}{\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^2}$$
$$+\frac{1}{2}\frac{\nu_3 m_3}{\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^2} + \frac{1}{2}\frac{\nu_2 m_2}{\left(\sigma_1^2 + K\sigma_2^2\right)^2} + \frac{1}{2}\frac{\nu_1 m_1}{\left(\sigma_1^2\right)^2}$$

$$\frac{\partial^2 \ell}{(\partial \sigma_1^2)^2} = \frac{\nu_1}{2\left(\sigma_1^2\right)^2} + \frac{\nu_2}{2(\sigma_1^2 + K\sigma_2^2)^2} + \frac{(\nu_3+1)}{2(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)^2} - \frac{1}{2}\frac{IJK(\overline{\overline{\overline{Y}}}_{\ldots} - \mu)^2}{\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^3}$$
$$-\frac{2}{2}\frac{\nu_3 m_3}{\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^3} - \frac{2}{2}\frac{\nu_2 m_2}{\left(\sigma_1^2 + K\sigma_2^2\right)^3} - \frac{2}{2}\frac{\nu_1 m_1}{\left(\sigma_1^2\right)^3}$$

and we know $E\left(m_1\right) = \sigma_1^{\,2}$, $E\left(m_2\right) = (\sigma_1^{\,2} + K\sigma_2^{\,2})$ and $E\left(m_3\right) = (\sigma_1^{\,2} + K\sigma_2^{\,2} + JK\sigma_3^{\,2})$.

Therefore

$$-E\left(\frac{\partial^2 \ell}{(\partial \sigma_1^2)^2}\right) = -\frac{\nu_1}{2\left(\sigma_1^2\right)^2} - \frac{\nu_2}{2(\sigma_1^2 + K\sigma_2^2)^2} - \frac{(\nu_3+1)}{2(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)^2} + \frac{1}{2}\frac{IJK.E(\overline{\overline{\overline{Y}}}_{\ldots} - \mu)^2}{\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^3}$$
$$+\frac{\nu_3 E(m_3)}{\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^3} + \frac{\nu_2 E(m_2)}{\left(\sigma_1^2 + K\sigma_2^2\right)^3} + \frac{\nu_1 E(m_1)}{\left(\sigma_1^2\right)^3}$$

$$-E\left(\frac{\partial^2 \ell}{(\partial \sigma_1^2)^2}\right) = -\frac{\nu_1}{2\left(\sigma_1^2\right)^2} - \frac{\nu_2}{2(\sigma_1^2 + K\sigma_2^2)^2} - \frac{(\nu_3+1)}{2(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)^2} + \frac{1}{2}\frac{IJK\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)}{IJK\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^3}$$
$$+\frac{\nu_3(\sigma_1^{\,2} + K\sigma_2^{\,2} + JK\sigma_3^{\,2})}{\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^3} + \frac{\nu_2(\sigma_1^{\,2} + K\sigma_2^{\,2})}{\left(\sigma_1^2 + K\sigma_2^2\right)^3} + \frac{\nu_1 \sigma_1^{\,2}}{\left(\sigma_1^2\right)^3}$$

$$-E\left(\frac{\partial^2 \ell}{(\partial \sigma_1^2)^2}\right) = -\frac{\nu_1}{2\left(\sigma_1^2\right)^2} + \frac{\nu_1}{\left(\sigma_1^2\right)^2} - \frac{\nu_2}{2(\sigma_1^2 + K\sigma_2^2)^2} + \frac{\nu_2}{\left(\sigma_1^2 + K\sigma_2^2\right)^2} - \frac{(\nu_3+1)}{2(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)^2}$$
$$+\frac{(\nu_3+1)}{2\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^2}$$

$$-E\left(\frac{\partial^2\ell}{(\partial\sigma_1{}^2)^2}\right)=\frac{\nu_1}{2\left(\sigma_1^2\right)^2}+\frac{\nu_2}{2\left(\sigma_1^2+K\sigma_2^2\right)^2}+\frac{(\nu_3+1)}{2\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^2}$$

$$-E\left(\frac{\partial^2\ell}{(\partial\sigma_1{}^2)^2}\right)=\frac{\nu_1}{2\left(\sigma_1^2\right)^2}+\frac{\nu_2}{2\left(\sigma_1^2+K\sigma_2^2\right)^2}+\frac{(\nu_3+1)}{2\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^2}$$

$$-E\left(\frac{\partial^2\ell}{(\partial\sigma_1{}^2)^2}\right)=F_{22}=\frac{1}{2}\left\{\frac{\nu_1}{\left(\sigma_1^2\right)^2}+\frac{\nu_2}{\left(\sigma_1^2+K\sigma_2^2\right)^2}+\frac{(\nu_3+1)}{\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^2}\right\}.$$

Further

$$\frac{\partial\ell}{\partial\sigma_2{}^2}=0-\frac{K\nu_2}{2(\sigma_1^2+K\sigma_2^2)}-\frac{K(\nu_3+1)}{2(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2)}+\frac{IJK^2(\overline{\overline{\overline{Y}}}_{\dots}-\mu)^2}{2\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^2}+\frac{K\nu_3m_3}{2\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^2}$$
$$+\frac{K\nu_2m_2}{2\left(\sigma_1^2+K\sigma_2^2\right)^2}$$

$$\frac{\partial^2\ell}{(\partial\sigma_2^2)^2}=0+\frac{K^2\nu_2}{2(\sigma_1^2+K\sigma_2^2)^2}+\frac{K^2(\nu_3+1)}{2(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2)^2}-\frac{IJK^3(\overline{\overline{\overline{Y}}}_{\dots}-\mu)^2}{\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^3}-\frac{K^2\nu_3m_3}{\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^3}$$
$$-\frac{K^2\nu_2m_2}{\left(\sigma_1^2+K\sigma_2^2\right)^3}$$

$$-E\left(\frac{\partial^2\ell}{(\partial\sigma_2{}^2)^2}\right)=0-\frac{K^2\nu_2}{2(\sigma_1^2+K\sigma_2^2)^2}-\frac{K^2(\nu_3+1)}{2(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2)^2}+\frac{IJK^3.E(\overline{\overline{\overline{Y}}}_{\dots}-\mu)^2}{\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^3}+\frac{K^2\nu_3E(m_3)}{\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^3}$$
$$+\frac{K^2\nu_2E(m_2)}{\left(\sigma_1^2+K\sigma_2^2\right)^3}$$

$$-E\left(\frac{\partial^2\ell}{(\partial\sigma_2{}^2)^2}\right)=-\frac{K^2\nu_2}{2(\sigma_1^2+K\sigma_2^2)^2}-\frac{K^2(\nu_3+1)}{2(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2)^2}+\frac{IJK^3.\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)}{IJK\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^3}+\frac{K^2\nu_3(\sigma_1{}^2+K\sigma_2{}^2+JK\sigma_3{}^2)}{\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^3}$$
$$+\frac{K^2\nu_2\left(\sigma_1^2+K\sigma_2^2\right)}{\left(\sigma_1^2+K\sigma_2^2\right)^3}$$

$$-E\left(\frac{\partial^2\ell}{(\partial\sigma_2{}^2)^2}\right)=-\frac{K^2\nu_2}{2(\sigma_1^2+K\sigma_2^2)^2}-\frac{K^2(\nu_3+1)}{2(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2)^2}+\frac{K^2}{\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^2}+\frac{K^2\nu_3}{\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^2}+\frac{K^2\nu_2}{\left(\sigma_1^2+K\sigma_2^2\right)^2}$$

$$-E\left(\frac{\partial^2\ell}{(\partial\sigma_2{}^2)^2}\right)=-\frac{K^2\nu_2}{2(\sigma_1^2+K\sigma_2^2)^2}+\frac{K^2\nu_2}{\left(\sigma_1^2+K\sigma_2^2\right)^2}-\frac{K^2(\nu_3+1)}{2(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2)^2}+\frac{K^2(\nu_3+1)}{\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^2}$$

$$-E\left(\frac{\partial^2\ell}{(\partial\sigma_2{}^2)^2}\right)=\frac{K^2\nu_2}{2(\sigma_1^2+K\sigma_2^2)^2}+\frac{K^2(\nu_3+1)}{2(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2)^2}$$

$$-E\left(\frac{\partial^2 \ell}{(\partial {\sigma_2}^2)^2}\right)=F_{33}=\frac{K^2}{2}\left\{\frac{\nu_2}{(\sigma_1^2+K\sigma_2^2)^2}+\frac{(\nu_3+1)}{(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2)^2}\right\}.$$

Further

$$\frac{\partial \ell}{\partial {\sigma_3}^2}=0+0-\frac{JK(\nu_3+1)}{2(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2)}+\frac{I\left(JK\right)^2(\overline{\overline{\overline{Y}}}_{\dots}-\mu)^2}{2\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^2}+\frac{JK\nu_3 m_3}{2\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^2}$$

$$\frac{\partial^2 \ell}{\left(\partial {\sigma_3}^2\right)^2}=\frac{\left(JK\right)^2(\nu_3+1)}{2(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2)^2}-\frac{I\left(JK\right)^3(\overline{\overline{\overline{Y}}}_{\dots}-\mu)^2}{\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^3}-\frac{\left(JK\right)^2\nu_3 m_3}{\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^3}$$

$$-E\left(\frac{\partial^2 \ell}{\left(\partial {\sigma_3}^2\right)^2}\right)=-\frac{\left(JK\right)^2(\nu_3+1)}{2(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2)^2}+\frac{I\left(JK\right)^3E(\overline{\overline{\overline{Y}}}_{\dots}-\mu)^2}{\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^3}+\frac{\left(JK\right)^2\nu_3 E(m_3)}{\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^3}$$

$$-E\left(\frac{\partial^2 \ell}{\left(\partial {\sigma_3}^2\right)^2}\right)=-\frac{\left(JK\right)^2(\nu_3+1)}{2(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2)^2}+\frac{I\left(JK\right)^3\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)}{IJK\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^3}+\frac{\left(JK\right)^2\nu_3\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)}{\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^3}$$

$$-E\left(\frac{\partial^2 \ell}{\left(\partial {\sigma_3}^2\right)^2}\right)=-\frac{\left(JK\right)^2(\nu_3+1)}{2(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2)^2}+\frac{\left(JK\right)^2}{\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^2}+\frac{\left(JK\right)^2\nu_3}{\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^2}$$

$$-E\left(\frac{\partial^2 \ell}{\left(\partial {\sigma_3}^2\right)^2}\right)=F_{44}=-\frac{\left(JK\right)^2(\nu_3+1)}{2(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2)^2}+\frac{\left(JK\right)^2(\nu_3+1)}{\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^2}=\frac{\left(JK\right)^2(\nu_3+1)}{2(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2)^2}$$

Also

$$\frac{\partial \ell}{\partial\mu\partial\sigma_1^2}=-\frac{IJK(\overline{\overline{\overline{Y}}}_{\dots}-\mu)}{\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^2}\qquad \therefore -E\left(\frac{\partial \ell}{\partial\mu\partial\sigma_1^2}\right)=F_{12}=F_{21}=\frac{IJKE(\overline{\overline{\overline{Y}}}_{\dots}-\mu)}{\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^2}=0$$

$$\frac{\partial \ell}{\partial\mu\partial\sigma_2^2}=-\frac{IJK^2(\overline{\overline{\overline{Y}}}_{\dots}-\mu)}{\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^2}\qquad \therefore -E\left(\frac{\partial \ell}{\partial\mu\partial\sigma_2^2}\right)=F_{13}=F_{31}=\frac{IJK^2E(\overline{\overline{\overline{Y}}}_{\dots}-\mu)}{\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^2}=0$$

$$\frac{\partial \ell}{\partial\mu\partial\sigma_3^2}=-\frac{IJ^2K^2(\overline{\overline{\overline{Y}}}_{\dots}-\mu)}{\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^2}\qquad \therefore -E\left(\frac{\partial \ell}{\partial\mu\partial\sigma_3^2}\right)=F_{14}=F_{41}=\frac{IJ^2K^2E(\overline{\overline{\overline{Y}}}_{\dots}-\mu)}{\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^2}=0$$

Further

$$\frac{\partial^2 \ell}{\partial\sigma_1^2\partial\sigma_2^2} = -0 + \frac{K\nu_2}{2(\sigma_1^2 + K\sigma_2^2)^2} + \frac{K(\nu_3+1)}{2(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)^2} - \frac{IJK^2(\overline{\overline{\overline{Y}}}_{\dots} - \mu)^2}{\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^3}$$

$$- \frac{K\nu_3 m_3}{\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^3} - \frac{K\nu_2 m_2}{\left(\sigma_1^2 + K\sigma_2^2\right)^3}$$

$$-E\left(\frac{\partial^2 \ell}{\partial\sigma_1^2\partial\sigma_2^2}\right) = -\frac{K\nu_2}{2(\sigma_1^2 + K\sigma_2^2)^2} - \frac{K(\nu_3+1)}{2(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)^2} + \frac{IJK^2.E(\overline{\overline{\overline{Y}}}_{\dots} - \mu)^2}{\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^3}$$

$$+ \frac{K\nu_3 E(m_3)}{\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^3} + \frac{K\nu_2 E(m_2)}{\left(\sigma_1^2 + K\sigma_2^2\right)^3}$$

$$-E\left(\frac{\partial^2 \ell}{\partial\sigma_1^2\partial\sigma_2^2}\right) = -\frac{K\nu_2}{2(\sigma_1^2 + K\sigma_2^2)^2} - \frac{K(\nu_3+1)}{2(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)^2} + \frac{IJK^2.\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)}{IJK\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^3}$$

$$+ \frac{K\nu_3\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)}{\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^3} + \frac{K\nu_2\left(\sigma_1^2 + K\sigma_2^2\right)}{\left(\sigma_1^2 + K\sigma_2^2\right)^3}$$

$$-E\left(\frac{\partial^2 \ell}{\partial\sigma_1^2\partial\sigma_2^2}\right) = -\frac{K\nu_2}{2(\sigma_1^2 + K\sigma_2^2)^2} - \frac{K(\nu_3+1)}{2(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)^2} + \frac{K}{\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^2}$$

$$+ \frac{K\nu_3}{\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^2} + \frac{K\nu_2}{\left(\sigma_1^2 + K\sigma_2^2\right)^2}$$

$$-E\left(\frac{\partial^2 \ell}{\partial\sigma_1^2\partial\sigma_2^2}\right) = -\frac{K\nu_2}{2(\sigma_1^2 + K\sigma_2^2)} + \frac{K\nu_2}{\left(\sigma_1^2 + K\sigma_2^2\right)^2} - \frac{K(\nu_3+1)}{2(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)} + \frac{K(\nu_3+1)}{\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^2}$$

$$-E\left(\frac{\partial^2 \ell}{\partial\sigma_1^2\partial\sigma_2^2}\right) = \frac{K\nu_2}{2(\sigma_1^2 + K\sigma_2^2)} + \frac{K(\nu_3+1)}{2(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)}$$

$$-E\left(\frac{\partial^2 \ell}{\partial\sigma_1^2\partial\sigma_2^2}\right) = F_{23} = F_{32} = \frac{K}{2}\left\{\frac{\nu_2}{(\sigma_1^2 + K\sigma_2^2)} + \frac{(\nu_3+1)}{(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)}\right\}$$

.

Further

$$\frac{\partial \ell}{\partial \sigma_1{}^2 \partial \sigma_3{}^2} = 0 + 0 + \frac{JK(\nu_3 + 1)}{2(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)^2} - \frac{I\left(JK\right)^2 (\overline{\overline{\overline{Y}}}_{\dots} - \mu)^2}{\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^3}$$

$$- \frac{JK\nu_3 m_3}{\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^3}$$

$$-E\left(\frac{\partial \ell}{\partial \sigma_1{}^2 \partial \sigma_3{}^2}\right) = -\frac{JK(\nu_3 + 1)}{2(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)^2} + \frac{I\left(JK\right)^2 E(\overline{\overline{\overline{Y}}}_{\dots} - \mu)^2}{\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^3}$$

$$+ \frac{JK\nu_3 E\left(m_3\right)}{\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^3}$$

$$-E\left(\frac{\partial \ell}{\partial \sigma_1{}^2 \partial \sigma_3{}^2}\right) = -\frac{JK(\nu_3 + 1)}{2(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)^2} + \frac{I\left(JK\right)^2 \left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)}{IJK\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^3}$$

$$+ \frac{JK\nu_3\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)}{\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^3}$$

$$-E\left(\frac{\partial \ell}{\partial \sigma_1{}^2 \partial \sigma_3{}^2}\right) = F_{24} = F_{42} = -\frac{JK(\nu_3 + 1)}{2(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)^2} + \frac{JK(\nu_3 + 1)}{\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^2} = \frac{JK(\nu_3 + 1)}{2(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)^2}$$

Also

$$\frac{\partial^2 \ell}{\partial \sigma_2{}^2 \partial \sigma_3{}^2} = \frac{JK^2(\nu_3 + 1)}{2(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)^2} - \frac{IJ^2K^3(\overline{\overline{\overline{Y}}}_{\dots} - \mu)^2}{\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^3} - \frac{JK^2\nu_3 m_3}{\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^3}$$

$$-E\left(\frac{\partial^2 \ell}{\partial \sigma_2{}^2 \partial \sigma_3{}^2}\right) = -\frac{JK^2(\nu_3 + 1)}{2(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)^2} + \frac{IJ^2K^3E(\overline{\overline{\overline{Y}}}_{\dots} - \mu)^2}{\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^3} + \frac{JK^2\nu_3 E(m_3)}{\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^3}$$

$$-E\left(\frac{\partial^2 \ell}{\partial \sigma_2{}^2 \partial \sigma_3{}^2}\right) = -\frac{JK^2(\nu_3 + 1)}{2(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)^2} + \frac{IJ^2K^3\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)}{IJK\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^3} + \frac{JK^2\nu_3\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)}{\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^3}$$

$$-E\left(\frac{\partial^2 \ell}{\partial \sigma_2{}^2 \partial \sigma_3{}^2}\right) = F_{34} = F_{43} = -\frac{JK^2(\nu_3 + 1)}{2(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)^2} + \frac{JK^2(\nu_3 + 1)}{\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^2} = \frac{JK^2(\nu_3 + 1)}{2(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)^2}$$

The inverse of the Fisher information matrix is given by

$$F^{-1}(\mu, \sigma_1^2, \sigma_2^2, \sigma_3^2) = F^{-1}(\underline{\theta}) = \begin{bmatrix} F^{11} & F^{12} & F^{13} & F^{14} \\ F^{21} & F^{22} & F^{23} & F^{24} \\ F^{31} & F^{32} & F^{33} & F^{34} \\ F^{41} & F^{42} & F^{43} & F^{44} \end{bmatrix}$$

where $F^{11} = \dfrac{\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2}{(\nu_3+1)JK}$

$$F^{12} \quad = \quad F^{21} = 0; F^{13} = F^{31} = 0; F^{14} = F^{41} = 0;$$

$$F^{22} \quad = \quad \frac{1}{|H|}\frac{\nu_2(\nu_3+1)J^2K^4}{4(\sigma_1^2+K\sigma_2^2)^2(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2)^2};$$

$$F^{23} = \frac{-1}{|H|}\frac{J^2K^3\nu_2(\nu_3+1)}{4(\sigma_1^2+K\sigma_2^2)^2(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2)^2} = F^{32};$$

$$F^{24} = 0 = F^{42};$$

$$F^{33} = \frac{J^2K^2(\nu_3+1)}{4|H|(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2)^2}\left\{\frac{\nu_1}{(\sigma_1^2)^2}+\frac{\nu_2}{(\sigma_1^2+K\sigma_2^2)^2}\right\};$$

$$F^{34} = \frac{-\nu_1(\nu_3+1)JK^2}{4|H|(\sigma_1^2)(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2)^2} = F^{43};$$

$$F^{44} = \frac{\nu_1K^2}{4|H|(\sigma_1^2)^2}\left\{\frac{\nu_2}{(\sigma_1^2+K\sigma_2^2)^2}+\frac{(\nu_3+1)}{(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2)^2}\right\}$$

and

$$|H| = \frac{\nu_1\nu_2(\nu_3+1)J^2K^4}{8(\sigma_1^2)^2(\sigma_1^2+K\sigma_2^2)^2(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2)^2}.$$

We are interested in the probability matching prior for (${}_3P^1_{pl}$), the lower process performance index.

Let $\underline{\theta} = [\mu, \sigma_3{}^2, \sigma_2{}^2, \sigma_1{}^2]'$. The capability index is

$$ {}_3P^1_{pl} = \frac{\mu - l_0}{3\left(\dfrac{\sigma_1^2+K\sigma_2^2+JK\sigma_3^2}{JK}\right)^{\frac{1}{2}}} = \frac{(\mu-l_0)(JK)^{\frac{1}{2}}}{3\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^{\frac{1}{2}}}.$$

Therefore

$$\frac{\partial t(\underline{\theta})}{\partial\mu} = \frac{(JK)^{\frac{1}{2}}}{3\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^{\frac{1}{2}}}, \quad \frac{\partial t(\underline{\theta})}{\partial\sigma_1{}^2} = \frac{-(\mu-l_0)(JK)^{\frac{1}{2}}}{6\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^{\frac{3}{2}}}$$

$$\frac{\partial t(\underline{\theta})}{\partial \sigma_2{}^2} = \frac{-(\mu - l_0) J^{\frac{1}{2}} K^{\frac{3}{2}}}{6\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^{\frac{3}{2}}} \text{ and } \frac{\partial t(\underline{\theta})}{\partial \sigma_3{}^2} = \frac{-(\mu - l_0)\left(JK\right)^{\frac{3}{2}}}{6\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^{\frac{3}{2}}}.$$

As mentioned

$$\nabla'_t(\tilde{\theta}) = \left[\frac{\partial t(\underline{\theta})}{\partial \mu} \quad \frac{\partial t(\underline{\theta})}{\partial \sigma_1{}^2} \quad \frac{\partial t(\underline{\theta})}{\partial \sigma_2{}^2} \quad \frac{\partial t(\underline{\theta})}{\partial \sigma_3{}^2}\right] =$$

$$\left[\frac{\left(JK\right)^{\frac{1}{2}}}{3\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^{\frac{1}{2}}} \quad \frac{-(\mu - l_0)\left(JK\right)^{\frac{1}{2}}}{6\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^{\frac{3}{2}}} \quad \frac{-(\mu - l_0) J^{\frac{1}{2}} K^{\frac{3}{2}}}{6\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^{\frac{3}{2}}} \quad \frac{-(\mu - l_0)\left(JK\right)^{\frac{3}{2}}}{6\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^{\frac{3}{2}}}\right]$$

$$\nabla'_t(\underline{\theta}) = \frac{\left(JK\right)^{\frac{1}{2}}}{3\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^{\frac{1}{2}}}\left[1 \; \frac{-(\mu - l_0)}{2\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)} \; \frac{-(\mu - l_0)K}{2\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)} \; \frac{-(\mu - l_0)\left(JK\right)}{2\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)}\right]$$

Further

$$\nabla'_t(\underline{\theta})F^{-1}(\underline{\theta}) = \frac{\left(JK\right)^{\frac{1}{2}}}{3\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^{\frac{1}{2}}}\left[\frac{\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2}{(\nu_3+1)JK} \quad 0 \quad 0 \; \frac{-(\mu-l_0)(JK)}{2\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)} \times \frac{\nu_1 K^2}{4|H|(\sigma_1^2)^2} \frac{\nu_2}{(\sigma_1^2 + K\sigma_2^2)^2}\right]$$

$$\text{with} \left|H\right| = \frac{\nu_1 \nu_2 (\nu_3 + 1) J^2 K^4}{8(\sigma_1^2)^2 (\sigma_1^2 + K\sigma_2^2)^2 (\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)^2}$$

$$\nabla'_t(\underline{\theta})F^{-1}(\underline{\theta}) = \frac{\left(JK\right)^{\frac{1}{2}}}{3\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^{\frac{1}{2}}}\left[\frac{\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)}{(\nu_3+1)JK} \quad 0 \quad 0 \; \frac{-(\mu-l_0)(JK)}{2\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)} \times \frac{2(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)^2}{(\nu_3+1)J^2K^2}\right]$$

$$\nabla'_t(\underline{\theta})F^{-1}(\underline{\theta}) = \frac{\left(JK\right)^{\frac{1}{2}}}{3\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^{\frac{1}{2}}}\left[\frac{\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)}{(\nu_3+1)JK} \quad 0 \quad 0 \; \frac{-(\mu-l_0)(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2)}{(\nu_3+1)\left(JK\right)}\right]$$

$$\nabla'_t(\underline{\theta})F^{-1}(\underline{\theta})\nabla_t(\underline{\theta}) = \frac{(JK)}{9\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)}\left[\frac{\sigma_1^2+K\sigma_2^2+JK\sigma_3^2}{(\nu_3+1)JK} \quad 0 \quad 0 \quad \frac{-(\mu-l_0)(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2)}{(\nu_3+1)(JK)}\right]\begin{bmatrix} 1 \\ \frac{-(\mu-l_0)}{2\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)} \\ \frac{-(\mu-l_0)K}{2\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)} \\ \frac{-(\mu-l_0)(JK)}{2\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)} \end{bmatrix}$$

$$= \frac{(JK)}{9\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)}\left(\frac{\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)}{(\nu_3+1)JK} + \frac{-(\mu-l_0)(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2)}{(\nu_3+1)(JK)} \times \frac{-(\mu-l_0)(JK)}{2\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)}\right)$$

$$= \frac{(JK)}{9\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)}\left(\frac{\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)}{(\nu_3+1)JK} + \frac{(\mu-l_0)^2(JK)}{2(\nu_3+1)(JK)}\right)$$

$$= \frac{1}{9}\left(\frac{1}{(\nu_3+1)} + \frac{(\mu-l_0)^2(JK)}{2(\nu_3+1)\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)}\right)$$

$$= \left(\frac{1}{9(\nu_3+1)} + \frac{(\mu-l_0)^2(JK)}{18(\nu_3+1)\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)}\right)$$

$$= \left(\frac{1}{9I} + \frac{(\mu-l_0)^2(JK)}{18I\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)}\right) \text{ since } I = \nu_3 + 1$$

$$\nabla'_t(\underline{\theta})F^{-1}(\underline{\theta})\nabla_t(\underline{\theta}) = \frac{1}{3I^{\frac{1}{2}}}\left(1+\frac{(\mu-l_0)^2(JK)}{2\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)}\right)^{\frac{1}{2}}.$$

Define as before

$$\eta(\underline{\theta}) = \frac{\nabla'_t(\underline{\theta})(\underline{\theta})F^{-1}(\underline{\theta})}{\sqrt{\nabla'_t(\underline{\theta})F^{-1}(\underline{\theta})\nabla_t(\underline{\theta})}} = [\eta_1(\underline{\theta}) \quad \eta_2(\underline{\theta}) \quad \eta_3(\underline{\theta}) \quad \eta_4(\underline{\theta})]$$

where

$$\nabla'_t(\underline{\theta})F^{-1}(\underline{\theta}) = \frac{(JK)^{\frac{1}{2}}}{3\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^{\frac{1}{2}}}\left[\frac{\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)}{(\nu_3+1)JK} \quad 0 \quad 0 \quad \frac{-(\mu-l_0)(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2)}{(\nu_3+1)(JK)}\right]$$

$$\nabla' t(\underline{\theta}) F^{-1}(\underline{\theta}) = \frac{(JK)^{\frac{1}{2}}}{3\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^{\frac{1}{2}}} \begin{bmatrix} \frac{\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)}{(IJK)} & 0 & 0 & \frac{-(\mu-l_0)(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2)}{(IJK)} \end{bmatrix}$$

since $I = \nu_3 + 1$

and

$$\nabla'_t(\underline{\theta}) F^{-1}(\underline{\theta}) \nabla_t(\underline{\theta}) = \frac{1}{3I^{\frac{1}{2}}} \left(1+\frac{(\mu-l_0)^2 (JK)}{2\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)}\right)^{\frac{1}{2}}$$

giving

$$\eta_1(\underline{\theta}) = \frac{(JK)^{\frac{1}{2}}}{3\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^{\frac{1}{2}}} \frac{\frac{\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)}{(IJK)}}{\frac{1}{3I^{\frac{1}{2}}}\left(1+\frac{(\mu-l_0)^2 (JK)}{2\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)}\right)^{\frac{1}{2}}}$$

$$\eta_1(\underline{\theta}) = \frac{(JK)^{\frac{1}{2}}}{3\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^{\frac{1}{2}}} 3I^{\frac{1}{2}} \frac{\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)}{(IJK)} \left(1+\frac{(\mu-l_0)^2 (JK)}{2\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)}\right)^{-\frac{1}{2}}$$

$$\eta_1(\underline{\theta}) = \frac{\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^{\frac{1}{2}}}{(IJK)^{\frac{1}{2}}} \left(1+\frac{(\mu-l_0)^2 (JK)}{2\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)}\right)^{-\frac{1}{2}}$$

$\eta_2(\underline{\theta}) = 0$ and $\eta_3(\underline{\theta}) = 0$

$$\eta_4(\underline{\theta}) = \frac{(JK)^{\frac{1}{2}}}{3\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^{\frac{1}{2}}} \frac{\frac{-(\mu-l_0)(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2)}{(IJK)}}{\frac{1}{3I^{\frac{1}{2}}}\left(1+\frac{(\mu-l_0)^2 (JK)}{2\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)}\right)^{\frac{1}{2}}}$$

$$\eta_4(\underline{\theta}) = \frac{(JK)^{\frac{1}{2}}}{3\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^{\frac{1}{2}}} 3I^{\frac{1}{2}} \frac{-(\mu-l_0)(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2)}{(IJK)} \left(1+\frac{(\mu-l_0)^2 (JK)}{2\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)}\right)^{-\frac{1}{2}}$$

$$\eta_4(\underline{\theta}) = \frac{-(\mu - l_0)\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^{\frac{1}{2}}}{\left(IJK\right)^{\frac{1}{2}}}\left(1+\frac{(\mu-l_0)^2(JK)}{2\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)}\right)^{-\frac{1}{2}}.$$

For a prior $\pi(\underline{\theta})$ to be a probability-matching prior, the differential equation

$$\sum_{\alpha=1}^{m}\frac{\partial}{\partial\theta_\alpha}\{\eta_\alpha(\underline{\theta})\pi(\underline{\theta})\} = 0$$

$$\frac{\partial}{\partial\mu}\{\eta_1(\underline{\theta})\pi(\underline{\theta})\} + \frac{\partial}{\partial\sigma_1^{\ 2}}\{\eta_2(\underline{\theta})\pi(\underline{\theta})\} + \frac{\partial}{\partial\sigma_2^{\ 2}}\{\eta_3(\underline{\theta})\pi(\underline{\theta})\} + \frac{\partial}{\partial\sigma_3^{\ 2}}\{\eta_4(\underline{\theta})\pi(\underline{\theta})\} = 0$$

must be satisfied.

The probability matching prior is

$$\pi(\underline{\theta}) = \pi(\mu, \sigma_1^{\ 2}, \sigma_2^{\ 2}, \sigma_3^2) \propto \sigma_1^{-2}\left(\sigma_1^2 + K\sigma_2^2\right)^{-1}\left(\sigma_1^2 + K\sigma_2^2 + JK\sigma_3^2\right)^{-\frac{1}{2}}\left(1+\frac{(\mu-l_0)^2(JK)}{2\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)}\right)^{\frac{1}{2}}$$

since

$$\frac{\partial}{\partial\mu}\left\{\frac{\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^{\frac{1}{2}}}{(IJK)^{\frac{1}{2}}}\left(1+\frac{(\mu-l_0)^2(JK)}{2\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)}\right)^{-\frac{1}{2}}\times\sigma_1^{-2}\left(\sigma_1^2+K\sigma_2^2\right)^{-1}\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^{-\frac{1}{2}}\left(1+\frac{(\mu-l_0)^2(JK)}{2\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)}\right)^{\frac{1}{2}}\right\} + 0 + 0 +$$

$$\frac{\partial}{\partial\sigma_3^{2}}\left\{\frac{-(\mu-l_0)\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^{\frac{1}{2}}}{(IJK)^{\frac{1}{2}}}\left(1+\frac{(\mu-l_0)^2(JK)}{2\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)}\right)^{-\frac{1}{2}}\times\sigma_1^{-2}\left(\sigma_1^2+K\sigma_2^2\right)^{-1}\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)^{-\frac{1}{2}}\left(1+\frac{(\mu-l_0)^2(JK)}{2\left(\sigma_1^2+K\sigma_2^2+JK\sigma_3^2\right)}\right)^{\frac{1}{2}}\right.$$

$$\frac{\partial}{\partial\mu}\left\{\frac{1}{(IJK)^{\frac{1}{2}}}\times\sigma_1^{-2}\left(\sigma_1^2+K\sigma_2^2\right)^{-1}\right\} + 0 + 0 + \frac{\partial}{\partial\sigma_3^{2}}\left\{\frac{-(\mu-l_0)}{(IJK)^{\frac{1}{2}}}\times\sigma_1^{-2}\left(\sigma_1^2+K\sigma_2^2\right)^{-1}\right\}$$

$$0 + 0 + 0 + 0 = 0.$$

CHAPTER 7

SUMMARY, CONCLUSIONS AND FURTHER RESEARCH

7.1 SUMMARY

Process capability indices have been extensively studied by frequentists. This thesis illustrated the flexibility and unique features of the Bayesian approach for the analysis of process capability indices. Posterior distributions and credibility intervals for the indices are obtained through Monte Carlo simulations where independent samples are obtained from joint posterior distributions. Non-informative and probability-matching priors were derived for single and multiple variance components. Some specific processes may be well described by a single variance component but for others, more complicated structures may be required. It is the task of the statistician to find an appropriate model describing the process data at hand and to base conclusions on that model. In this thesis situations involving one, two and three variance components were used to describe production processes.

7.2 CONCLUSIONS

Bayesian inference has a number of advantages. The variance component model provides a natural way of taking into account all sources of uncertainty in the estimation of the process capability indices. Uncertainty about the true values of parameters for the process capability indices are incorporated into the analysis through the choice of a non-informative prior distribution. The probability-matching and reference priors were recommended because they are designed to produce posterior credibility intervals which are asymptotically identical to their frequentist counterparts.

The non-informative priors yield near exact frequentist inferences for all the process capability indices used. The means of indices using Bayesian simulation techniques were compared to their more commonly used frequentist estimates. From the simulation results it is clear that the probability matching and reference priors work quite well. It seems, therefore, that the frequentist properties of Bayesian inferences of capability indices based on the probability matching and reference priors work well. Berger and Sun (2008) came to similar conclusions using many statistical functions including the reciprocal of the coefficient of variation.

The added advantage of the Bayesian approach is that, from the posterior distributions (represented by histograms) of the capability indices, one is in a position to obtain quantiles, credible regions and can also perform other inferential tasks. In some of the cases, the exact distributions of the indices were derived.

For the conventional prior the exact posterior moments of the indices can be calculated. By using these moments, Pearson curve and Cornish-Fisher approximations of the posterior distribution are obtained. Gibbs sampling can also be used to obtain the unconditional posterior distribution of indices.

Bayesian simulation techniques were used to solve the supplier selection problem using process capability indices and information, which would not otherwise be available to a frequentist, was made available in the form of posterior distributions of the indices and probabilities of suppliers being ranked first to last were computed.

7.3 BAYESIAN SIMULATION OF OTHER INDICES AND FURTHER RESEARCH

In chapters 4, 5 and 6 the simple case of the lower process capability index only was investigated. In chapter 4, for instance, the index, $C_{pl} = \frac{\mu - LSL}{3\sigma}$ is used in defense of the Bayesian approach. This form of the index is convenient and easy to work with from a Bayesian point of view. Some applications require an upper limit rather than a lower limit. In a similar way to C_{pl}, the reference and probability-matching priors for the upper capability index $C_{pu} = \frac{USL-\mu}{3\sigma}$ could be derived without much extra effort. The Bayesian results derived, will still hold in the case of C_{pu} since $C_{pu} = -\left(\frac{\mu - USL}{3\sigma}\right)$ which is now of the same form as C_{pl} except for the negative sign.

The index C_{pl} is extended to two variance components in chapter 5 and to three variance components in chapter 6. The two and three variance component upper process capability indices could be defined similarly. For example, the upper process capability indices for the two variance component case are specified as:

$$P_{pu} = \frac{USL-\mu}{3\left(\sigma_1^2+\sigma_2^2\right)^{\frac{1}{2}}} \quad \text{and} \quad P_{pu}^{\;1} = \frac{USL-\mu}{3\left(\frac{\sigma_1^2+J\sigma_2^2}{J}\right)^{\frac{1}{2}}}$$

Notice also that for the symmetric case

$$\begin{aligned} C_p &= \frac{USL-LSL}{6\sigma} \\ &= \frac{1}{2}\left(\frac{USL-\mu}{3\sigma}+\frac{\mu-LSL}{3\sigma}\right) \end{aligned}$$

Therefore

$$C_p = \frac{1}{2}(C_{pu} + C_{pl})$$

if the process is centred within the specification range. Therefore the Bayesian results derived thus far would also apply to an index like C_p. The result could also be extended to the two and three variance component indices of chapters 5 and 6 without much extra effort.

Again notice also that C_{pk} is made up of the two indices namely C_{pl} and C_{pu}.

$C_{pk} = \min\left(\frac{USL-\mu}{3\sigma}, \frac{\mu-LSL}{3\sigma}\right) = \min\left(C_{pu}, C_{pl}\right)$. The posterior distribution of this form of index is easy to simulate from a Bayesian point but is not so easy to derive algebraically. The simulation results can also be extended to the two and three variance component indices of chapters 5 and 6 without much extra effort.

The lower process capability index can be viewed as a brick that can be used to build the other indices.

Although the analysis ended up only on three random effects (three variance components), the methods can be extended and applied to more complex designs and unbalanced data sets.

However, the reference and probability-matching priors derived for C_{pl} will not be easily derived for C_{pk} and other indices. Work still needs to be done from a Bayesian point of view to try and develop probability-matching and/or reference priors for indices where the denominator of the index is in a form such as in the following index:

$$C_{pm} = \frac{USL - LSL}{6\sqrt{\sigma^2 + (\mu - T)^2}}$$

In this thesis an important assumption of normality was made about the data. More work still needs to be done when this assumption is violated. It would be interesting to work on the indices from a Bayesian point of view when the data is assumed to be non-symmetrical, or follows a distribution which is not normal such as the t-distribution or double exponential.

REFERENCES

AAIG (1995). *Statistical Process Control Reference Manual*. Automotive Industry Action Group, Southfield.

Berger, J.O. and Bernardo, J.M. (1992). *On the Development of Reference Priors in Bayesian Statistics IV*, Eds. J.M. Bernardo, J.O. Berger, A.P. David and A.F.M. Smith, Oxford University Press, 35 – 70.

Berger, J.O. and Dongchu, S. (2008). Objective priors for the Bivariate normal model. *The Annals of Statistics* 36(2), 963-982.

Bernardo, J.M. and Irony, T.Z. (1996). A general multivariate Bayesian process capability index, *The Statistician*, 45, 487–502.

Bothe, D.R. (2002). Discussion, *Journal of Quality Technology*, 34(1): 32-37.

Boyles, R.A. (1991). The Taguchi capability Index. *Journal of Quality Technology*, 23: 17-26.

Box, G.E.P. and Tiao, G.C. (1973). *Bayesian Inference in Statistical Analysis*. Addison –Wesley, Reading, MA.

Carlin, B.P., Gelfand, A.E. and Smith, A.F.M. (1992). Hierarchical Bayes analysis of change point problems. *Applied Statistics*, 41(2), 389–405.

Chan, L.K., Cheng, S.W. and Spiring, F.A. (1988). A new measure of process capability: Cpm, *Journal of Quality Technology*, 20, 162–175.

Chen, K.S. and Pearn, W.L. (1997). An application of the non–normal process capability indices. *Quality & Reliability Engineering International*, 13, 355–360.

Cheng, S.W. and Siring, F.A. (1989). Assessing process capability: A Bayesian approach, *IEE Transactions*, 21, 97–98.

Chikobvu, D. and van der Merwe A.J. (2007). A process Capability Index for Averages of Observations from New Batches in the case of the Balanced Random Effects Model. *Technical Report, UFS.*

Choi, B.C. and Owen, D.B. (1990). A study of new process capability index. *Communications in Statistics-Theory and Methods*, 19, 1231-1245.

Chou, Y.M. (1994). Selecting a better supplier by testing process capability indices. *Quality Engineering*, 6(3): 427-438.

Chou, Y.M. and Owen, D.B. (1989). On the distribution of the estimated process capability indices. *Communs. Statist. Theory. Meth*., 18, 4549–4560 .

Cornish, E.A. and Fisher, R.A. (1937). Moments and cumulants in the specification of distributions. *Revue de l'Institut International de Statisque,* 5, 307–320.

Datta, G.S. and Gosh, J.K. (1995). On Priors providing frequentist validity for Bayesian inference. *Biometrika*, 82: 37–45.

Elderton, W.P. (1953). *Frequency Curves and Correlation*. Cambridge University Press.

Elderton, W.P. and Johnson, N.L. (1969). *Systems of Frequency Curves*. Cambridge University Press.

Fisher, R.A. and Cornish, E.A. (1960). The Percentile Points of distributions having known cumulants. *Technometrics*, 2, 209–226.

Gelfand, A.E., Hills, S.E., Racine–Poon, A. and Smith, A.F.M. (1990). Illustration of Bayesian inference in normal data models using Gibbs sampling. *J. Amer. Statist. Assoc.* 85, 972–985.

Gelfand, A.E. and Smith, A.F.M. (1991). Gibbs sampling for marginal posterior expectations, *Comm. Statist. Theory Methods*, 20(5, 6) 1747–1766.

Gelfand, A.E. , Smith, A.F.M. and Lee, T.M. (1992). Bayesian Analysis of constrained parameters and truncated data problems using Gibbs sampling. *J. Amer. Statist. Assoc.* 87, 523–532.

Gelman, A. and Rubin, D.R. (1992). A single series from the Gibbs sampler provides a false sense of security, *Bayesian Statistics* 4 (ed. J.M. Bernardo, J. Berger, A.P. Dawid and A. F.M. Smith), 627–635, Oxford University Press, Oxford.

Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. Pat Anal. and Mach. Intel*. 6, 721–741.

Greenwich, M. and Jahr-Shaffroth, B. (1995). A Process Incapability index. *International Journal of Quality, Reliability and Management* 12, 58-71.

Guttman, I. and Menzefriecke, U. (2003). Posterior Distributions for Functions of Variance Components. *Sociedad de Estadistica e Investigacio'n Operativa Test.*, 12, 115-123.

Hastings, W.K. (1970). Monte Carlo sample methods using Markov chain and their applications. *Biometrika*, 57, 97–109.

Herman, J.T. (1989). Capability Index-Enough for Process Industries? *Transactions of the ASQC Quality Congress*, Toronto 670-675.

Hsiang, T.C., Taguchi, G. (1985). Tutorial on quality control and assurance-The Taguch methods. *Joint meetings of the American Statistical Associations*, Las Vegas, Nevada, pp. 188.

Hoskins, J., Stuart, B. and Taylor, J. (1988). A Motorola Commitment: A six sigma mandate. The Motorola guide to statistical process control for continuous improvement towards six sigma quality.

Hsiang, T.C., Taguchi, G. (1985). Tutorial on quality control and assurance-The Taguch methods. Joint meetings of the American Statistical Associations, Las Vegas, Nevada, pp. 188.

Hubele, N.F., Berrado, A. and Gel, E.S. (2005). A Wald test for comparing multiple capability Indices. *Journal of Quality Technology*, 37(4), 304–307.

Hubele, N.F., Montgomery, D.C. and Chin, W.H. (1991). An application of statistical process control in jet–turbine engine component manufacturing. *Quality Engineering*, 4(2), 197–210.

Jeffreys, H. (1961). *Theory of Probabaility.* Oxford University Press, New York, N.Y.

Juran J.M. (1974). *Juran's Quality Control Handbook* (3rd Edition). McGraw-Hill, New York, N.Y.

Kane, V.E. (1986). Process capability indices. *Journal of Quality Technology*, 18(1), 41–52.

Kotz S., Johnson, N.L. (2002). Process capability Indices-A review, 1992-2000. *Journal of Quality Technology*, 34(1): 2-53.

Li, K. (2007). Pool Size Selection for the Sampling/Importance Resampling Algorithm. *Statistica Sinica* 17, 895-907.

Lu, M.W. and Rudy, R.J. (2002). Discussion, *Journal of Quality Technology*, 34(1): 38-39.

Lynch, D.P. (2004). Setting the Record Straight with Capability Indices. *SAE Technical paper series*. 2004 SAE World Congress Detroit, Michigan March 8-11, 2004.

Lyth, D.M. and Rabiej, R.J. (1995). Critical variables in wood manufacturing's process capability: species, structure, and moisture content. *Quality Engineering*, 8(2), 275–281.

Mukerjee, R. and Dey, D.K. (1993). Frequentist validity of posterior quantiles in the presence of a nuisance parameter: higher order asymptotics. *Biometrika*, 80, 3, pp. 499–505.

Niverthi, M and Dey, D.K. (2000). Multivariate process capability: A Bayesian Perspective, *Commun. Statist. – Theory Meth*, 667–687.

Parlar, M., Wesolowsky, G.O. (1999). Specification limits, capability Indices and centering in assembly manufacture. *Journal of Quality Technology*, 31: 317-325.

Pearn, W.L., Kotz, S., and Johnson, N.L. (1992). Distributional and inferential properties of Process Capability Indices, *Journal of Quality Technology*, 24, 216–231.

Polansky, A.M., (2006). Permutation methods for comparing process capabilities. *Journal of Quality Technology*, 38(3): 254-266.

Rodriguez, R.N. (2002). Discussion, *Journal of Quality Technology*, 34(1): 28-31.

Singpurwala, N.D., (1998). The Stochastic control of Process capability Indices. Sociedoad de Estyadistica e Investigacion Operativa Test, 7(1) 1-74.

Skare,O., Bolviken, E. and Holden, L. (2003). Improved Sampling-Importance Resampling and Reduced Bias Importance Sampling. *J. Scand. Statist,.* 30, 719-737.

Smith, A. and Gelfand, A. (1992). Bayesian Analysis Statistics without Tears: A Sampling-Resampling Perspective. *Amer. Statist.*, 46(2), 84-88.

Spiring, F., Cheng, S., Yeung, A. and Leung B. (2002). Discussion, *Journal of Quality Technology*, 34(1): 23-27.

Tong, L.I., Chen, C.L., Hsu, H.H. (1998). Construction of the confidence interval using bootstrap simulation to distinguish between two process capability indices.

Proceedings of the 3^rd Annual International Conference on Industrial Engineering Theories, Application and Practice, Hong Kong, pp. 1035-1042.

Van der Merwe, A.J. and Chikobvu D. (2004). Bayesian Estimation of Process Capability Index C_{pk}. *South African Statistical Journal*, 38 (2), 139-158.

Vangel, M.G. (1992). New Methods for One sided Tolerance Limits for a One-Way Balanced Random-Effects ANOVA Model. *Technometrics* 34, 176-185.

Vannman K. (1995). Unified Approach to Capability Indices. *Statistica Sinica,* 5: 805-820.

Wolfinger, R.D. (1998). Tolerance Intervals for Variance Component Models Using Bayesian Simulation. *Journal of Quality Technology*. 30, 18-32.

Yeh, A.B. and Bhattacharya, S. (1998). A Robust Capability Index. Communications in Statistics-Simulation and Computation 27, 565-589.

Zellner, A. (1971). *An Introduction to Bayesian Inference in Econometrics.* New York. John Wiley & Sons.

Zhang, N.F., Stenback, G.A. and Wardrop, D.M. (1990). Interval estimation of the process capability index C_{pk}. *Communs. Statist. Theory Meth*, 19, 4455–4470.