BAYESIAN TOLERANCE INTERVALS WITH PROBABILITY MATCHING PRIORS

DHARINI A/P PATHMANATHAN

THESIS SUBMITTED IN FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

INSTITUTE OF MATHEMATICAL SCIENCES
FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2014

UNIVERSITI MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: **DHARINI A/P PATHMANATHAN**

I/C/Passport No: **840510-14-5480**

Regisration/Matric No.: **SHB100001**

Name of Degree: **DOCTOR OF PHILOSOPHY**

Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"):

"BAYESIAN TOLERANCE INTERVALS WITH PROBABILITY MATCHING PRIORS"

Field of Study: **BAYESIAN STATISTICS**

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work,
(2) This Work is original,
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work,
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work,
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained,
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

(Candidate Signature) Date: 28 JANUARY 2014

Subscribed and solemnly declared before,

Witness's Signature Date: 28 JANUARY 2014

Name: **PROFESSOR DR ONG SENG HUAT**

Designation: Prof. Dr. Ong Seng Huat
Institut Sains Matematik
Universiti Malaya

ABSTRACT

A review on statistical tolerance intervals shows that the derivation of two-sided tolerance intervals is far more challenging than that of their one-sided counterparts. Much of the existing construction of two-sided tolerance intervals are through a numerical approach. This study addresses the problems of constructing two-sided tolerance intervals in balanced one-way random effects models and for a general family of distributions. The Bayesian tolerance interval developed by Ong and Mukerjee (2011) using probability matching priors (PMP) is compared via Monte Carlo simulation with the modified large sample (MLS) tolerance interval of Krishnamoorthy and Mathew (2009) for normal and non-normal experimental errors with respect to coverage probabilities and expected widths. Data generated from normal and non-normal experimental errors were studied to see the effects on the tolerance intervals since real data may not necessarily follow the normal distribution. Results show that the PMP tolerance interval appears to be less conservative for data with moderate and large number of classes while the MLS tolerance interval is preferable for smaller sample sizes. For the second part of the study, the PMP as well as frequentist two-sided tolerance intervals are constructed for a general family of parametric models. Simulation studies show that the asymptotic results are well-reflected in finite sample sizes. The findings are then applied to real data. The results obtained in this research are a contribution to the area of statistical tolerance regions.

ABSTRAK

Kajian tentang selang toleransi statistik menunjukkan bahawa penerbitan selang toleransi dua bahagian adalah jauh lebih mencabar berbanding penerbitan selang toleransi satu bahagian. Kebanyakan kaedah penerbitan selang toleransi dua bahagian yang sedia ada menggunakan pendekatan berangka. Penyelidikan ini memberi perhatian kepada masalah pembinaan selang toleransi dua bahagian bagi model kesan rawak sehala berimbang dan famili umum taburan. Kaedah simulasi Monte Carlo digunakan untuk membandingkan selang toleransi Bayesian yang dibina oleh Ong dan Mukerjee (2011) yang menggunakan prior berpadanan kebarangkalian (PBK) dengan selang toleransi hampir berbentuk tertutup melalui kaedah sampel besar terubahsuai (SBT) oleh Krishnamoorthy dan Mathew (2009). Ini melibatkan ralat eksperimen bertaburan normal dan tidak normal berdasarkan kebarangkalian liputan serta jangkaan lebar. Data yang dijana daripada ralat eksperimen bertaburan normal dan tidak normal dikaji bagi melihat kesan terhadap selang-selang toleransi ini kerana data sebenar tidak semestinya bertaburan normal. Hasil kajian menunjukkan bahawa selang toleransi PBK kelihatan kurang konservatif bagi data dengan bilangan kelas yang sederhana dan besar manakala selang toleransi SBT disyorkan bagi sampel bersaiz kecil. Dalam bahagian kedua penyelidikan ini, selang toleransi dua bahagian PBK serta frekuentis dibina bagi famili umum model-model berparameter. Kajian simulasi menunjukkan bahawa hasil-hasil asimptotik yang diperoleh dicerminkan dengan baik oleh sampel terhingga. Hasil-hasil yang diperoleh daripada kajian ini turut diaplikasi dalam data sebenar. Hasil-hasil penyelidikan ini merupakan satu sumbangan kepada bidang kajian yang melibatkan rantau toleransi statistik.

ACKNOWLEDGEMENTS

My journey throughout the years of doing my PhD definitely had its ups and downs. I am deeply touched by the kindness of many to stand by me in making this dream come true.

First and foremost, I would like to thank my mother Sorna Letchumy and brother Darmar whose moral support and fervent prayers shielded me during the stormy phases of this journey. Their words of encouragement boosted my confidence to achieve my goals.

I would like to dedicate my deepest appreciation to my supervisor Professor Ong Seng Huat who guided me and generously supported me by giving me all the facilities needed in doing my research. He was ever willing and never hesitant in providing me with all the materials which enabled my work to flow smoothly without any hitches. I am very fortunate because he is understanding and gave me the opportunities and freedom to correspond with some of the greatest scholars in my field of study.

I am ever grateful to Professor Rahul Mukerjee from the Indian Institute of Management Kolkata who has the substance of a genius for sparing his precious time in helping me with the completion of this thesis. I treasure his ideas which played a great role in my research. One must really be blessed to receive the guidance of such a sincere and dedicated mentor.

Furthermore, I am indebted to the University of Malaya and the Ministry of Higher Education, Malaysia for the generous financial support for my studies and research via the Skim Latihan Akademik IPTA (SLAI) scholarship and Postgraduate Research Grant (Project no. PS334-2010B). I also feel greatly honoured to be given the opportunity to study in the University of Malaya.

Sincere thanks to the examiners, Prof. Kalimuthu Krishnamoorthy from University of Louisiana at Lafayette, U.S.A., Prof. Thomas Mathew from University of Maryland, Baltimore County, U.S.A. and Dr. Ng Choung Min from University of Malaya for their constructive suggestions which led to the improvement of this thesis.

Special thanks to Professor Angelina Chin Yan Mui, the administrative staff and the computer laboratory staff of the Institute of Mathematical Sciences, University of Malaya for their help throughout my candidature.

In addition, thanks to my cousin, Rupani and friends especially Safwati Ibrahim, Huda Zuhrah, Nurul Najihah and Lily Wong for their constant prayers, kindness and support.

Last but not least, I offer my most humble obeisance to my Heavenly Father, the Almighty for His divine intervention which made the impossible possible and gave me the strength to surpass all difficulties.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

β	-	content in (β,γ) tolerance interval
γ	-	confidence level in (β,γ) tolerance interval
π	-	prior
$\phi(\cdot)$	-	standard normal density function
$\Phi(\cdot)$	-	standard normal distribution function
$O_p(n^{-1})$	-	at most of order n^{-1} in probability
$O(n^{-1})$	-	at most of order n^{-1}
$o_p(n^{-1})$	-	smaller order in probability than n^{-1}
$o(n^{-1})$	-	smaller order than n^{-1}
a	-	observed information element
Y_{ij}	-	ij^{th} observation in balanced one-way random effects models

e_{ij} - experimental error associated with Y_{ij} in balanced one-way random effects model

ν_i - random effects parameter for i^{th} class in balanced one-way random effects model

f - probability density function

g - additional term in tolerance interval

l - log-likelihood

C - observed information matrix

D - derivative operator

E - expectation

F - cumulative distribution function

I - Fisher information

N - normal distribution

P - probability

Q	-	percentile function
R	-	content
cdf	-	cumulative distribution function
pdf	-	probability density function
GLD	-	generalized lambda distribution
MLE	-	maximum likelihood estimator
MLS	-	modified large sample
MSB	-	mean square between classes
MSW	-	mean square within classes
PMP	-	probability matching priors

CHAPTER 1

INTRODUCTION

1.1 Background of study

The prior distribution plays a vital role in Bayesian analysis. It represents information regarding the uncertainty about a parameter, say θ, which is combined with the probability distribution of a new data resulting the posterior distribution. We are able to obtain the posterior distribution by multiplying the prior distribution and likelihood distribution. The posterior distribution is used in making future inferences related to θ. The prior distribution is an important asset since we will never be able to perform any Bayesian inference without it. However, the choice of prior distributions is the most crucial and criticized point in Bayesian analysis. Undeniably, selecting the prior distribution which is the key to Bayesian inference is a challenging task. According to Ghosh *et al.* (2008), with sufficient information from past experience, expert opinion or previously collected data, subjective priors are ideal, and indeed should be used for inferential purposes. However, we can use Bayesian techniques efficiently even without adequate prior information with some default or objective priors. A specific objectivity criterion for such priors known as the probability matching criterion has found appeal to both frequentists and Bayesians. Based on Datta and Sweeting (2005), a probability matching prior (PMP) in the context of credible regions is a prior distribution under which the posterior probabilities of certain regions coincide with their coverage probabilities, either exactly or approximately.

According to Ong and Mukerjee (2011), probability matching priors play an important role with regard to the Bayesian versus frequentist inference in statistical inference which is gaining significant attention in recent years. Ong and Mukerjee (2011) also mentioned that the problem of ensuring approximate frequentist validity of Bayesian credible sets namely those based on posterior quantiles for a parameter or parameteric function of interest is a substantial body of work concerning the probability matching priors. Hence, this appears to be appealing to Bayesians as noninformative or objective priors with an external validation and frequentists as means to obtain precise frequentist confidence sets with a Bayesian interpretation.

On the other hand, the computation of statistical intervals based on random samples has wide applicability. The choice of interval to be constructed depends on the underlying problem and application (Krishnamoorthy and Mathew, 2009). The commonly used statistical intervals consist of the confidence intervals, tolerance intervals and prediction intervals. Contrary to the confidence interval which provides an estimated range of values concerning an unknown population parameter such as the population mean and population variance, the tolerance interval gives information on the entire population where it captures at least a certain proportion of the population with a given confidence level. A prediction interval based on a random sample provides bounds for future observations (Krishnamoorthy and Mathew, 2009). We illustrate the distinction among a confidence interval, tolerance interval and prediction interval to get a clearer picture of their applications. For example, the 95% confidence interval estimated for a population mean indicates that 95% of the intervals in repeated sampling include the population mean. The prediction interval is interpreted similarly to the confidence interval where it provides information concerning a single value (Krishnamoorthy and Mathew, 2009). The tolerance interval is implemented in situations where we intend to use the sample to conclude at least 90% of the population is within the bounds with a certain confidence level, say 95%.

Tolerance intervals have a wide range of applications in diverse fields such as engineering, quality control, pharmaceutical studies, manufacturing, environmental monitoring and so on. The theory of statistical tolerance intervals has undergone vast development since the pioneering works by Wilks in 1941 and 1942. Various methods were implemented in constructing tolerance intervals within the framework of random effects models, regression models, multivariate normal populations, continuous distributions as well as discrete distributions.

The two-sided (β, γ) frequentist tolerance interval, say (*L, U*), contains at least a specified proportion β of the sampled population with a specified confidence γ, where *L* and *U* are called respectively the lower and upper tolerance limits. On the other hand, an equal-tailed tolerance interval (L_e, U_e) is constructed so that it includes the interval $\left(q_{\frac{1-\beta}{2}}, q_{\frac{1+\beta}{2}}\right)$, where q_p is the *p* quantile of the sampled population. Note that the equal-tailed tolerance interval (L_e, U_e) is constructed so that no more than $100\left(\frac{1-\beta}{2}\right)\%$ of the population is less than L_e, and no more than $100\left(\frac{1-\beta}{2}\right)\%$ is greater than U_e. Because of this constraint, an equal-tailed tolerance interval is wider than the corresponding two-sided tolerance intervals. Furthermore, the problem of finding an equal-tailed tolerance interval simplifies to simultaneous estimation of quantiles, and so Bonferroni method could be used to find and approximate equal-tailed tolerance interval in cases where the exact ones are difficult to obtain.

The formal definitions of one- and two-sided tolerance intervals are as follows: Let X be a continuous random variable with cumulative distribution function (c.d.f) $F_X(\cdot;\theta)$ where θ is a possibly vector valued unknown parameter. Let L and U be respectively the lower and upper bounds of a tolerance interval such that $L < U$. Let $P[.]$ denote the probability set function.

i. The one-sided (β,γ) tolerance interval associated with the lower tolerance limit, L of the form $[L,+\infty)$ is required to satisfy the condition

$$P[1 - F_X(L;\theta) \geq \beta] = \gamma$$

ii. The one-sided (β,γ) tolerance interval associated with the upper tolerance limit, U of the form $(-\infty,U]$ is required to satisfy the condition

$$P[F_X(U;\theta) \geq \beta] = \gamma$$

iii. The two-sided (β,γ) tolerance interval $[L, U]$ satisfies

$$P[F_X(U;\theta) - F_X(L;\theta) \geq \beta] = \gamma$$

The construction of two-sided tolerance intervals is more challenging than that of its one-sided counterpart.

Ong and Mukerjee (2011) explored matching priors in the context of tolerance intervals in balanced one-way and two-way nested random effects models. They have derived matching conditions for both models which characterize priors under which a $\beta-$content two-sided Bayesian tolerance interval with posterior credibility level $\gamma + O_p(n^{-1})$ which also has frequentist confidence level $\gamma + O(n^{-1})$ for every β and γ, where n is the sample size. Wolfinger (1998) and Van der Merwe and Hugo (2007) studied the models mentioned using non-informative priors where the analysis was done on matching prior for the posterior quantiles of the error variance. It is yet unknown if these priors enjoy the matching property specifically for the tolerance intervals. Besides using the Bayesian approach, Krishnamoorthy and Mathew (2009) applied the modified large sample (MLS) method in constructing tolerance interval based on the procedure by Graybill and Wang (1980) for the aforementioned models. Krishnamoorthy and Lian (2012) studied the merits of these intervals in their work.

Previous studies have shown that most of the tolerance intervals from various distributions are one-sided since the computation of two-sided intervals is rather a daunting task. So far, there are no general formulae that can be readily invoked to obtain a two-sided Bayesian or frequentist tolerance interval in a general framework of parametric models. In an attempt to address this problem, we first explore via higher order asymptotic considerations, two-sided Bayesian tolerance intervals under a fairly general framework of parametric models.

1.2 Objective of research

Ong and Mukerjee (2011) developed two-sided Bayesian tolerance intervals, with approximate frequentist validity, in balanced one-way and two-way nested random effects models using probability matching priors (PMP). On the other hand, Krishnamoorthy and Lian (2012) examined closed-form approximate tolerance intervals by the modified large sample (MLS) approach which was proposed by Krishnamoorthy and Mathew (2009). The objective in the first part of this work is to evaluate and perform a comparative study via Monte Carlo simulation between the PMP and MLS tolerance intervals for both normal and non-normal error distributions when the balanced one-way random effects models are of concern. The non-normal error distributions which are applied include the t-distribution, skew-normal (see Azzalini,

1985) and the generalized lambda distribution (see Karian and Dudewicz, 2000). Both t- and skew-normal distributions have heavier tails than the normal distribution while the generalized lambda distribution is a versatile four-parameter distribution which is able to produce distributions with various shapes and skewness.

The second part of the research aims at developing two-sided tolerance intervals in a fairly general framework of parametric models. Higher order asymptotics are developed to obtain explicit analytical formulae for these intervals in both Bayesian and frequentist setups. This, in turn, leads to a characterization of probability matching priors for the two-sided tolerance intervals and paves the way for the development of the corresponding frequentist results. For instances where the probability matching priors are difficult to be obtained, we develop purely frequentist tolerance intervals which cater to situations of this kind. The results are then applied to real life examples. The software MATLAB was used for the simulations and data analysis.

1.3 Outline of research

Chapter 2 of this thesis involves the literature review of this study where all the significant academic literature related to the study is discussed. In the first part, we will examine the development of probability matching priors. This is followed by the background and progress of statistical tolerance intervals in the balanced one-way random effects models. We shall also look into the development of tolerance intervals using the modified large sample (MLS) method as well as the tolerance intervals involving probability matching priors (PMP). The final part of this chapter focuses on the development of tolerance intervals in a fairly general framework of parametric models.

In Chapter 3, we study the two-sided tolerance intervals for the balanced one-way random effects models which include the Bayesian tolerance intervals with approximate frequentist validity (Ong and Mukerjee, 2011). We also look into the probability matching prior and examine if other priors satisfy the probability matching condition for the two-sided tolerance interval.

In Chapter 4, we study the modified large sample (MLS) tolerance intervals (Krishnamoorthy and Mathew, 2009). We also conduct a comparative study between the probability matching priors (PMP) discussed in Chapter 3 and the modified large sample (MLS) tolerance intervals by varying the error distributions for the balanced one-way random effects models. The distributions of interest include the normal, t-, skew-normal and the generalized lambda distributions. We examine the merits such as the expected widths, expected contents and the coverage probabilities of the tolerance intervals computed using these distributions.

Chapter 5 of this thesis emphasizes on the construction of two-sided tolerance intervals in a general framework of parametric models. We derive asymptotic results leading to explicit formulae for two-sided Bayesian and frequentist tolerance intervals. This process characterizes the probability matching priors for such intervals and indicates their roles in finding frequentist tolerance intervals via the Bayesian approach. We also develop purely frequentist tolerance intervals in situations where the matching priors are difficult to be obtained. These intervals are applied to real data. Simulation studies are conducted to provide backing to the asymptotic results in finite samples.

Chapter 6 provides the concluding remarks as well as some significant contributions of this research. Suggestions on extending research works related to this research are also included in this chapter.

1.4 Contributions of research

Research results obtained in Chapter 4 and Chapter 5 for this thesis have respectively led to the acceptance of the following research papers for publication:

Pathmanathan, D., & Ong, S. H. (2013). A Monte Carlo simulation study of two-sided tolerance intervals in balanced one-way random effects model for non-normal errors. *Journal of Statistical Computation and Simulation*, (ahead-of-print), 1-16. doi: 10.1080/00949655.2013.792820

Pathmanathan, D., Mukerjee, R., & Ong, S. H. (2013). Two-sided Bayesian and frequentist tolerance intervals: general asymptotic results with applications. *Statistics*, (ahead-of-print), 1-15. doi: 10.1080/02331888.2012.748774

CHAPTER 2

LITERATURE REVIEW

2.1 Probability matching priors

The selection of priors is the most critical and controversial task in Bayesian analysis. In order to form the joint posterior distribution of the parameters given the data, the information provided in the prior distribution which should represent what is known about the unknown parameters before the data is available is combined with the information given by the data via the likelihood function (Box and Tiao, 1973). Continuous research has contributed in reducing the controversies due to this topic. (Van Boekel et al., 2004). According to Robert (2007), rarely the available prior information is accurate enough to lead to determining the exact prior distribution in practice. This is due to the sense that many probability distributions maybe compatible with this information. Some of the common techniques in determining prior distributions include the conjugate prior approach which requires a limited amount of information and the non-informative approach which can directly be derived from the sampling distribution (Robert, 2007). The use of these priors as well as the probability matching priors has greatly contributed in overcoming some of the issues surrounding the choice of prior distribution in Bayesian analysis (Hugo, 2012). Jeffreys' (1946) work on non-informative priors was a gift to the Bayesians because it shows a method to derive the prior distribution from the sampling distribution (Robert, 2007). However, some Bayesians were not in favour of such automated methods. Based on Scricciolo (1999), the Jeffreys' prior is given by $\pi_j(\theta) \propto [\det \mathrm{I}(\theta)]$ where $\det \mathrm{I}(\theta)$ is the determinant of the per observation $(n \times n)$ expected Fisher information matrix. This was designed primarily as a remedy for the lack of invariance to reparameterization of uniform priors. Therefore, we may consider the uniform prior as the distribution corresponding to that parameterization making $\mathrm{I}(\theta)$ independent of θ.

The probability matching criterion has found appeal to both Bayesians and frequentists (Ghosh et al., 2008). This criterion amounts to the requirement that the coverage probability of a Bayesian credible region is asymptotically equivalent to the coverage probability of the frequentist confidence region up to a certain order (Ghosh et al., 2008). These priors are attractive to frequentists as they are able to produce accurate frequentist confidence intervals with Bayesian interpretation while to the Bayesians, these priors can be considered as objective priors (Ong and Mukerjee, 2010). Datta and Sweeting (2005) defined a probability matching prior as a prior distribution under which the posterior probabilities coincide either exactly or approximately with their coverage probabilities. Situations where probability matching priors exist are very limited. Most of the literature on this topic focuses on approximate probability matching priors, usually for large *n*, based on the asymptotic theory of the maximum likelihood estimator. (Datta and Sweeting, 2005)

An example which illustrates the probability matching priors is as follows (Datta and Sweeting, 2005):

We consider an observation X from a $N(\theta,1)$ distribution where the parameter θ is unknown. When we take an improper uniform prior π over the real line of θ, the posterior distribution of $Z=\theta-X$ becomes exactly the same as its sampling distribution. Thus, $P_{\pi}\{\theta \leq \theta_{\alpha}(X) \mid X\} = P_{\theta}\{\theta \leq \theta_{\alpha}(X)\} = \alpha$,

where $\theta_{\alpha}(X) = X + z_{\alpha}$ and z_{α} represents the $\alpha-$quantile of a standard normal distribution. Hence, every credible interval based on the pivotal quantity Z with posterior probability α, is also a confidence interval with confidence level α. Therefore, the uniform distribution represents a probability matching prior.

It is mentioned in Datta and Sweeting (2005), that Lindley (1958) was one of the pioneers to review the probability matching problem in a different setup. He attempted to provide a Bayesian interpretation of Fisher's (1956) fiducial distribution for a scalar parameter. Under the assumption of a single sufficient statistic, Lindley (1958) showed that if a suitable transformation results in a location model with a location parameter $\tau = g(\theta)$, then exact matching holds by using a uniform prior on the location parameter τ (Datta and Sweeting, 2005).

The construction of probability matching priors has been actively studied for the past two decades. Scricciolo (1999) mentioned that Welch and Peers (1963) were among the first to study frequentist coverage properties of Bayesian intervals in cases involving scalar and vector parameters. Welch and Peers (1963) extended the study by Lindley (1958) to any location family model and developed the corresponding asymptotic theory. An explicit proof of these results was provided by Datta, Ghosh and Mukerjee (2000) and Datta and Mukerjee (2004, p.22). Datta and Mukerjee (2004) provided an excellent monograph on probability matching priors which stated that these priors are appealing to Bayesians as objective priors with an external validation, and to frequentists as a means of getting accurate intervals with a Bayesian interpretation. Among others who studied this topic are Mukerjee and Dey (1993) and Mukerjee and Ghosh (1997) who investigated higher order matching conditions. Along the same lines, Ghosh and Mukerjee (1998) examined the latest developments on probability matching priors. Some of the other review papers include Kass and Wasserman (1996) and Mukerjee and Reid (1999).

A prior is known as first- or second-order matching if it ensures approximate frequentist validity of posterior quantiles with margin of error $o(n^{-1/2})$ or $o(n^{-1})$ respectively (Ong and Mukerjee, 2010), where n is the sample size. Following Ong and Mukerjee (2010), the prior $\pi(\theta)$ is called first- or second-order probability matching if the relationship

$$P_\theta\left\{\theta \le \theta^{(1-\alpha)}(\pi, X)\right\} = 1 - \alpha + o(n^{-r/2}) \tag{2.1}$$

holds for $r = 1$ or 2 and for each α $(0 < \alpha < 1)$,

where $X_i, 1 \le i \le n$, are independent and identically distributed possibly vector-valued absolutely continuous random variables with common density $f(x;\theta)$, indexed by a scalar parameter θ. Given $X = (X_1, ..., X_n)'$, let $\theta^{(1-\alpha)}(\pi, X)$ be the $(1-\alpha)$-th posterior quantile of θ under a prior $\pi(\theta)$. Let P_θ denote the frequentist probability measure with respect to θ. The Jeffreys' prior was characterized as first-order probability matching by Welch and Peers (1963). They also studied model conditions under which it is second-order matching (Mukerjee and Reid, 1999).

According to Ghosh et al. (2008), there are several probability matching criteria which are achieved through:

a) posterior quantiles
b) distribution functions
c) highest posterior density (HPD) regions
d) inversion of certain test statistics

However priors based on (a)-(d) need not always be identical. A phenomenon where any prior satisfying all four criteria does not exist may occur (Ghosh et al., 2008). Mukerjee and Reid (2001) applied probability matching priors in computing Bayesian tolerance limits.

Ong and Mukerjee (2011) derived probability matching conditions in relation to tolerance intervals for both balanced one-way and two-way nested random effects models. These conditions enable us to evaluate if the priors satisfy the matching property for tolerance intervals.

2.2 The shrinkage argument

The derivation of a frequentist property from a Bayesian property usually proceeds by the introduction of an auxiliary prior distribution which is allowed to shrink in the true parameter value and thus producing the required frequentist probability. According to Datta and Mukerjee (2004), the shrinkage argument which plays a significant part in the development of matching priors was the brainchild of J.K. Ghosh who suggested to Ghosh and Mukerjee (1991) and Ghosh (1994, Ch. 9). Early applications of the shrinkage argument are foreshadowed in Bickel and Ghosh (1990) and Dawid (1991). The argument is presented in detail in Mukerjee and Reid (2000), the unpublished thesis of Li (1998) and Datta and Mukerjee (2004).

We shall closely follow this argument based on Mukerjee and Reid (2001) and Datta and Mukerjee (2004). Let X be a possibly vector-valued random variable with a probability density function, pdf $f(\cdot;\theta)$. The parameter θ belongs to the $p-$dimensional Euclidean space $\Re^p$ or some open subset. Suppose we intend to find an expression of the expectation $E_\theta\{h(X,\theta)\}$ where h is a measurable function. This expectation is known to exist and is continuous for all θ. In the present context, h is an indicator function in the case where $E_\theta\{h(X,\theta)\}$ represents a frequentist probability. As given in Datta and Mukerjee (2004), the following steps show a Bayesian approach for evaluating $E_\theta\{h(X,\theta)\}$.

Step 1: Consider a proper prior density $\bar{\pi}(\cdot)$ for θ where the support of $\bar{\pi}(\cdot)$ is a compact rectangle in the parameter space. $\bar{\pi}(\cdot)$ vanishes on the boundary of support while remaining positive in the interior. The support of $\bar{\pi}(\cdot)$ is a closure of the set on which it is positive. Thus, obtain $E^{\bar{\pi}}\{h(X,\theta) \mid X\}$ which is the expectation of $h(X,\theta)$ in the posterior setup. (Datta and Mukerjee, 2004)

Step 2: Find $E_\theta E^{\bar{\pi}}\{h(X,\theta) \mid X\}$ $(= \lambda(\theta)$ say), for θ in the interior of the support of $\bar{\pi}(\cdot)$. (Datta and Mukerjee, 2004)

Step 3: Integrate $\lambda(\cdot)$ with respect to $\bar{\pi}(\cdot)$ and then allow $\bar{\pi}(\cdot)$ to converge weakly to the degenerate prior at θ. This yields $E_\theta\{h(X,\theta)\}$. (Datta and Mukerjee, 2004)

We justify the above steps as follows. As shown in Datta and Mukerjee (2004), we note that the posterior density of θ under the prior $\bar{\pi}(\cdot)$ is given by $f(X;\theta)\ \bar{\pi}(\cdot) / N(X)$ where

$$N(X) = \int f(X;\theta)\ \bar{\pi}(\cdot)\ d\theta \tag{2.2}$$

Therefore, **Step 1** yields

$$E^{\bar{\pi}}\{h(X,\theta) \mid X\} = K(X) / N(X), \tag{2.3}$$

where

$$K(X) = \int h(X,\theta) f(X;\theta)\ \bar{\pi}(\cdot)\ d\theta \tag{2.4}$$

In view of **(2.3)**, **Step 2** yields

$$\lambda(\theta) = \int \{K(X) / N(X)\} f(x;\theta) dx$$

Thus in **Step 3**, integrating $\lambda(\cdot)$ with respect to $\bar{\pi}(\cdot)$, we get

$$\int \lambda(\theta)\ \bar{\pi}(\cdot)\ d\theta = \int\int \{K(X)/N(X)\} f(x;\theta)\ \bar{\pi}(\theta)\ dx\ d\theta$$

$$= \int \{K(X)/N(X)\}\ \{\int f(x;\theta)\bar{\pi}(\theta) d\theta\} dx$$

$$= \int K(X) dx$$

$$= \iint h(x,\theta) f(x;\theta)\ \bar{\pi}(\theta)\ d\theta\ dx$$

$$= \int \{\int h(x,\theta) f(x;\theta) dx\}\ \bar{\pi}(\theta)\ d\theta$$

$$= \int [E_\theta\{h(X,\theta)\}]\ \bar{\pi}(\theta)\ d\theta \tag{2.5}$$

using (2.2) and (2.4). By the assumed continuity of $E_\theta\{h(X,\theta)\}$ for all θ, as well as the compactness of the support of $\bar{\pi}(\cdot)$, the validity of the claim made in **Step 3** is proven in the last line of (2.5).

We observe in **Step 3** that $\bar{\pi}(\cdot)$ is allowed to converge weakly to a degenerate prior. Due to this, the present Bayesian approach is said to be based on the shrinkage argument (Datta and Mukerjee, 2004). The shrinkage argument is extensively used in Datta and Mukerjee (2004). It simplifies the derivation of matching priors in various contexts and also plays an important role in purely frequentist problems. Thus, it is applied in constructing Bayesian and frequentist tolerance intervals.

2.3 Statistical tolerance intervals: Introduction

The computation of tolerance intervals for continuous distributions was extensively studied since the pioneering work of Wilks (1941, 1942). Early works contributed by Wald and Wolfowitz (1946) demonstrated the construction of tolerance limits for normal distribution. Burrows (1963) gave a general introduction to tolerance intervals which played the role as a starting point to enhance the understanding of the utility of tolerance intervals. Apart from that, Patel (1986) provided a fairly comprehensive review at that time of publication which discussed tolerance intervals for various univariate distributions. The problem with the work by Patel (1986) is that there are many inconsistencies with the notations used. Hence, it is advised to refer to the primary sources for a clearer picture of when studying the formulae. Easterling and Weeks (1970) proposed and illustrated an accuracy criterion for a Bayesian approach for the exponential and normal densities. According to Krishnamoorthy and Mathew (2009), the last three decades have shown a vast development in the theory of statistical tolerance intervals and tolerance regions. The derivation of tolerance intervals in the framework of random effects models and simultaneous tolerance intervals for regression was only implemented during the 1980s and 1990s while satisfactory tolerance regions for multivariate normal populations and multivariate regression models were only accomplished in the last decade (Krishnamoorthy and Mathew, 2009).

Guttman (1970) and Hahn and Meeker (1971) provided informative reviews up to various stages while Krishnamoorthy and Mathew (2009) did an excellent and up-to-date study on tolerance intervals. Jilek (1981) compiled a bibliography which lists about 270 articles related to this topic and Jilek and Ackerman (1989) listed an additional 130 articles. Since then, the literature on this topic has shown a significant increase. The computation of tolerance intervals associated with continuous distributions has been studied extensively. Some examples of the literature on discrete cases include Zacks (1970), Hahn and Chandra (1981) and Cai and Wang (2005).

2.3.1 Tolerance intervals for variance component models

Several authors explored tolerance intervals for the one-way random effects model for both balanced and unbalanced cases as well as the two-way nested random effects model. Sahai and Ojeda (2004) gave a comprehensive and detailed study on fixed, random and mixed analysis of variance (ANOVA) models.

The work by Fertig and Mann (1974) who discussed the point estimations of the percentiles of the observations in the balanced one-way random effects model was a motivation to the derivation of one-sided tolerance intervals for the balanced one-way random effects model (Krishnamoorthy and Mathew, 2009). Lemon (1977) made the first attempt in formally deriving a lower tolerance limit for the distribution $N(\theta_1, \theta_2 + \theta_3)$ which turned out to be quite conservative; see Krishnamoorthy and Mathew (2009). The construction of one-sided tolerance limits has been well addressed by Mee and Owen (1983), Mee (1984), Vangel (1992), Bhaumik and Kulkarni (1996), Krishnamoorthy and Mathew (2004) and Liao *et al.* (2005). The methods available in order to obtain one-sided tolerance intervals are approximate; see Krishnamoorthy and Mathew (2004) for a comparative study of some approximate methods. Chen and Harris (2006) discussed numerical approach by conditioning on an estimator of the unknown expected mean squared ratio. Undoubtedly, constructing the one-sided tolerance intervals is much easier than that of the two-sided case.

Our concern in this research is with the two-sided tolerance intervals. Mee (1984) extended the procedures in Mee and Owen (1983) to find two-sided tolerance intervals; see Beckman and Tietjen (1989) for further results in this direction. Some methods are required to make these methods less conservative. Hoffman and Kringle

(2005) constructed two-sided tolerance intervals for general random-effects model for both balanced and unbalanced cases. Rebafka, Clémencon and Feinberg (2007) derived the new nonparametric bootstrap approach for two-sided mean coverage and guaranteed coverage tolerance limits for a balanced one-way random effects model. A solution to the tolerance interval problem (from the frequentist perspective) is given in the recent work of Sharma and Mathew (2012), under a very general mixed or random effects model.

Wolfinger (1998) presented the Bayesian simulation approach which handles different types of Bayesian tolerance intervals. The three kinds of commonly used tolerance intervals proposed by Wolfinger (1998) are as follows:

1. The (β, γ) tolerance interval, where β represents the content or the proportion of the population to be included in the interval and γ is the confidence level (reliability of the interval). Both β and γ lie between 0 and 1 and are typically assigned values of 0.90, 0.95 or 0.99 (Wolfinger, 1998).
2. The $\beta-$expectation tolerance interval, where β represents the expected coverage of the interval. β is again measured on a probability scale and is typically set to a value close to 1. This interval focuses on prediction of one or a few future observations from the process and consequently tends to be narrower than the corresponding (β, γ) intervals (Wolfinger, 1998).
3. The fixed-in-advance tolerance interval is the one that is specified in advance, and the intent is to estimate the actual proportion of the population that is included in the interval. (Wolfinger, 1998).

The intervals (1)-(3) can take forms of a lower limit (L, ∞), an upper limit $(-\infty, U)$ or a two-sided limit (L, U) (Wolfinger, 1998).

Recently, Ong and Mukerjee (2011) studied two-sided Bayesian tolerance intervals with approximate frequentist validity, in balanced one-way and two-way nested random effects models using probability matching priors (PMP). Ong and Mukerjee (2011) derived probability matching conditions specific to the aforementioned problem via a technique involving inversion of approximate posterior characteristic functions. These conditions are beneficial in the evaluation of some other priors which have been applied. It was unknown whether the priors employed by Wolfinger (1998) (balanced one-way random effects model) and Van der Merwe and Hugo (2007) (two-way nested random effects model) enjoy matching properties specifically for tolerance intervals until Ong and Mukerjee (2011) showed that these priors did not meet the requirements of the matching criterion.

Krishnamoorthy and Lian (2012) studied closed-form approximate tolerance intervals by the modified large sample (MLS) approach which was introduced by Krishnamoorthy and Mathew (2009). The MLS approach is based on the procedure by Graybill and Wang (1980) for finding upper confidence limits for a linear combination of variance components. Krishnamoorthy and Lian (2012) also compared the MLS tolerance intervals with the tolerance intervals constructed using the generalized variable approach which was introduced by Liao *et al.* (2005). The MLS method in computing tolerance intervals in various models was illustrated by Krishnamoorthy and Mathew (2009). They found that the MLS approach produced results similar to the generalized variable case. Moreover, the MLS tolerance intervals are easier to be computed as they are in closed-form (Krishnamoorthy and Lian, 2012).

In the first part of this research, both PMP and MLS intervals were applied for non-normal errors and the distributions of interest are the t-distribution, skew-normal

(Azzalini, 1985) and generalized lambda distributions. Karian and Dudewicz (2000) extensively studied the generalized lambda distribution.

2.3.2 Two-sided Bayesian and frequentist tolerance intervals for a general framework of parametric models

In the second part of the study, we develop two-sided Bayesian and frequentist tolerance intervals for a general framework of parametric models. Probability matching priors for one-sided tolerance intervals were characterized in Mukerjee and Reid (2001). The tolerance intervals which will be studied involve the normal, Weibull and inverse Gaussian distributions.

As mentioned earlier, Wolfinger (1998) came up with an approach based on Bayesian simulation whereas in our work we give analytical formulae applicable to wide ranging parametric models, based on the foundation of higher order asymptotic theory. His approach cannot be easily adapted for frequentist tolerance intervals. The development of general results on such frequentist tolerance intervals is a main thrust of our research. We explicitly try to ensure posterior credibility level $\gamma + O_p(n^{-1})$ for a $\beta -$content two-sided tolerance interval, where n is the sample size.

Young (2010) gave a useful R package for obtaining tolerance intervals involving discrete and continuous cases as well as regression tolerance intervals. Krishnamoorthy and Mathew (2009) discussed non-normal tolerance intervals such as log-normal, gamma, two-parameter exponential, Weibull and other related distributions. There is no general method available for constructing a two-sided tolerance interval. However, Krishnamoorthy and Xie (2011) provided a general framework for a symmetric location-scale family which can be readily applied to find tolerance intervals and equal tailed tolerance intervals. These authors illustrated the approach to find tolerance intervals for normal, Laplace and logistic distributions with censored data. For the Weibull distribution, tolerance limits were constructed using the generalized variable method. Statistical problems concerning the Weibull distribution are not simple due to the MLEs not being in closed form. Thus, they are computed numerically. Monte Carlo procedures were applied by Thoman, Bain and Antle (1969) based on the distributions of certain pivotal quantities involving the maximum likelihood estimators, MLEs (see Krishnamoorthy and Mathew, 2009). The results obtained for the MLEs enable the empirical finding of the distributions of some pivotal quantities. This is done according to the inferential procedures for Weibull parameters (Krishnamoorthy and Mathew, 2009). This approach contributed to the development of methods for confidence limits for reliability and one-sided tolerance limits based on the MLEs of the Weibull distribution; see Thoman et al. (1970). Approximate methods were proposed in constructing one-sided tolerance intervals for the Weibull case and these do not require simulation. Some of the works include Mann and Fertig (1975, 1977), Mann (1978), Engelhardt and Bain (1977) and Bain and Engelhardt (1981). Krishnamoorthy and Mathew (2009) discussed Monte Carlo procedures for the computation of one-sided tolerance limits, estimating a survival probability and for constructing lower limits for the stress-strength reliability involving the Weibull distribution. Tang and Doug (1994) proposed one-sided tolerance limits for the inverse Gaussian model and carried out Monte Carlo simulations to evaluate these limits in terms of coverage probability and average values.

The tolerance intervals for the Weibull, inverse Gaussian and other models in the literature are mainly one-sided since it is difficult to construct two-sided tolerance intervals. Guenther (1972) and Hahn and Meeker (1991) mentioned that one-sided tolerance limits can be used to obtain approximate equal-tailed tolerance intervals via the Bonferroni's inequality. The Bonferroni's approximation is used to control the central $100 \times \beta\%$ of the sampled population while controlling both tails to achieve at least $100 \times \gamma\%$ confidence (Young, 2010). Studies have shown that no procedure to compute two-sided tolerance intervals for the parametric models is available in the literature.

We apply the two-sided tolerance intervals to real data. For the Weibull tolerance interval, we consider the shelf life data in Gacula and Kubala (1973). As for the inverse Gaussian case, it is mentioned in Chhikara and Folks (1989) that the inverse Gaussian model fits the failure of ball bearings data in Lieblin and Zelen (1956).

CHAPTER 3

PROBABILITY MATCHING TOLERANCE INTERVAL FOR BALANCED ONE-WAY RANDOM EFFECTS MODEL

3.1 Introduction

Recently, Ong and Mukerjee (2011) studied the $\beta-$content tolerance interval with posterior credibility level $\gamma + O_p(n^{-1})$ which also has frequentist confidence level $\gamma + O(n^{-1})$ for balanced one-way and two-way nested random effects models. These were computed via probability matching priors (PMP). Wolfinger (1998) presented the two-sided tolerance intervals obtained via Bayesian simulation. In this chapter, we discuss the PMP two-sided tolerance intervals for balanced one-way random effects model. We shall also study if the prior used by Wolfinger (1998) satisfies the probability matching criteria for tolerance intervals given by Ong and Mukerjee (2011).

3.2 Balanced one-way random effects model

The balanced one-way random effects model is defined as follows:

$$Y_{ij} = \theta_1 + v_i + e_{ij} \tag{3.1}$$

for $i = 1,2,...,n$, where n represents the number of classes and $j = 1,2,...,t$, where t represents the number of observations per class. Here Y_{ij} denotes the ij^{th} observation and θ_1 is the population mean. v_i is a random effect for the i^{th} class and e_{ij} is the experimental error associated with Y_{ij}. v_i and e_{ij} are independent with $v_i \sim N(0,\theta_2)$ and $e_{ij} \sim N(0,\theta_3)$.

ρ is the intra-class correlation coefficient and

$$\rho = \frac{\theta_2}{\theta_2 + \theta_3} . \tag{3.2}$$

By fixing the value of ρ, the relationship between the variance of v_i, θ_2 and the variance of e_{ij}, θ_3 is given by

$$\theta_2 = \frac{\rho}{1-\rho}\theta_3 . \tag{3.3}$$

3.3 Preliminaries

Let $Y_i = (Y_{i1},...,Y_{it})'$ where $i = 1, 2, ..., n$. Under the model assumption, Y_i's are independent with the same t-variate normal distribution.

The maximum likelihood estimator (MLE) of $\theta = (\theta_1, \theta_2, \theta_3)$ is given by $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3)$ where

$$\hat{\theta}_1 = \overline{Y},\ \hat{\theta}_2 = \frac{(\text{MSB} - \text{MSW})}{t},\ \hat{\theta}_3 = \text{MSW}, \tag{3.4}$$

Note: The MLE is calculated without imposing the non-negativity constraint on θ_2.

$\overline{Y}$ is the grand mean of Y_{ij} while MSW and MSB are the usual mean squares within and between classes, that is,

$$\text{MSW} = \frac{1}{n(t-1)}\sum_{i=1}^{n}\sum_{j=1}^{t}\left(Y_{ij} - \overline{Y}_i\right)^2 \quad \text{and} \quad \text{MSB} = \frac{t}{n}\sum_{i=1}^{n}\left(\overline{Y}_i - \overline{Y}\right)^2 . \tag{3.5}$$

$\overline{Y}_i$ denotes the mean of the i^{th} class. In the following sections we consider asymptotics as $n \to \infty$ so as to ensure the consistency of these MLEs.

$Y_1, Y_2, ..., Y_n$ are independent and identically distributed (i.i.d) $N_t(\theta_1 L_t, \theta_2 J_t + \theta_3 I_t)$

$$L_t = \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \qquad I_t = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & & \\ \vdots & \vdots & \vdots & \vdots & \\ 0 & 0 & 0 & \ldots & 1 \end{bmatrix}, \qquad J_t = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 1 & 1 & \cdots & 1 \\ 1 & 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots & 1 \\ 1 & 1 & 1 & \ldots & 1 \end{bmatrix}$$

L_t is $n \times 1$ matrix with each element equals to $1. I_t$ is an identity matrix of order n and J_t is a $n \times n$ matrix with every element equals to 1. The variance of Y_i is

$$V = \theta_2 J_t + \theta_3 I_t . \tag{3.6}$$

Following Sahai and Ojeda (2004) and applying Result A.1 and Result A.2 from Appendix A, we get

$$|V| = \theta_3^{t-1}(\theta_3 + t\theta_2), \tag{3.7}$$

and

$$V^{-1} = \frac{1}{\theta_3} I_t - \frac{\theta_2}{\theta_3(\theta_3 + t\theta_2)} J_t . \tag{3.8}$$

Applying the definition of the multivariate normal distribution,

Recall that the probability density function of Y_i is given by

$$f(Y_i;\theta) = \frac{1}{(2\pi)^{t/2} |V|^{1/2}} \exp\left\{ -\frac{1}{2}(Y_i - \theta_1 \ell_t)' V^{-1} (Y_i - \theta_1 \ell_t) \right\}. \tag{3.9}$$

Therefore, the likelihood function of Y is given by

$$f(Y;\theta) = \frac{1}{(2\pi)^{nt/2} |V|^{n/2}} \exp\left\{ -\frac{1}{2}\sum_{i=1}^{n}(Y_i - \theta_1 \ell_t)' V^{-1} (Y_i - \theta_1 \ell_t) \right\}. \tag{3.10}$$

For the exponent of the function in (3.9), we have

$$(Y_i - \theta_1 \ell_t)' V^{-1} (Y_i - \theta_1 \ell_t) = (Y_i - \theta_1 \ell_t)' \left\{ \frac{1}{\theta_3} I_t - \frac{\theta_2}{\theta_3(\theta_3 + t\theta_2)} J_t \right\} (Y_i - \theta_1 \ell_t)$$

$$=\frac{1}{\theta_3}\sum_{j=1}^{t}(Y_{ij}-\theta_1)^2-\frac{\theta_2}{\theta_3(\theta_3+t\theta_2)}\left[\sum_{j=1}^{t}(Y_{ij}-\theta_1)\right]^2$$

$$(Y_i-\theta_1\ell_t)'V^{-1}(Y_i-\theta_1\ell_t)\ =\frac{1}{\theta_3}\sum_{j=1}^{t}(Y_{ij}-\theta_1)^2-\frac{t^2\theta_2}{\theta_3(\theta_3+t\theta_2)}(\overline{Y}_i-\theta_1)^2\,. \tag{3.11}$$

Furthermore,

$$\sum_{j=1}^{t}(Y_{ij}-\theta_1)^2\ =\sum_{j=1}^{t}(Y_{ij}-\overline{Y}_i)^2+t(\overline{Y}_i-\theta_1)^2\,. \tag{3.12}$$

We substitute (3.12) into (3.11) and hence the exponent term reduces to

$$\frac{1}{\theta_3}\sum_{j=1}^{t}(Y_{ij}-\overline{Y}_i)^2+\frac{t}{(\theta_3+t\theta_2)}(\overline{Y}_i-\theta_1)^2\ . \tag{3.13}$$

Therefore,

$f(Y_i;\theta)=$

$$\frac{1}{(2\pi)^{t/2}\theta_3^{(t-1)/2}(\theta_3+t\theta_2)^{1/2}}\exp\left\{-\left[\frac{1}{2\theta_3}\sum_{j=1}^{t}(Y_{ij}-\overline{Y}_i)^2+\frac{t}{(\theta_3+t\theta_2)}(\overline{Y}_i-\theta_1)^2\right]\right\}. \tag{3.14}$$

<u>Solutions to the likelihood equation</u>

Let $l(\theta)=n^{-1}\sum_{i=1}^{n}\log f(Y_i;\theta)$. Then, writing $\overline{Y}=\frac{1}{n}\sum_{i=1}^{n}\overline{Y}_i=\frac{1}{nt}\sum_{i=1}^{n}\sum_{j=1}^{t}Y_{ij}$,

$$l(\theta)=-\frac{t}{2}\log(2\pi)-\left(\frac{t-1}{2}\right)\log\theta_3-\frac{1}{2}\log(\theta_3+t\theta_2)-\frac{1}{2n\theta_3}\sum_{i=1}^{n}\sum_{j=1}^{t}(Y_{ij}-\overline{Y}_i)^2$$
$$-\frac{t}{2n(\theta_3+t\theta_2)}\left[\sum_{i=1}^{n}(\overline{Y}_i-\theta_1)^2\right]$$
$$=-\frac{t}{2}\log(2\pi)-\left(\frac{t-1}{2}\right)\log\theta_3-\frac{1}{2}\log(\theta_3+t\theta_2)-\frac{1}{2n\theta_3}\sum_{i=1}^{n}\sum_{j=1}^{t}(Y_{ij}-\overline{Y}_i)^2$$
$$-\frac{t}{2n(\theta_3+t\theta_2)}\left\{\sum_{i=1}^{n}(\overline{Y}_i-\overline{Y})^2+n(\overline{Y}-\theta_1)^2\right\}$$

$$\therefore l(\theta) = -\frac{t}{2}\log(2\pi) - \left(\frac{t-1}{2}\right)\log\theta_3 - \frac{1}{2}\log(\theta_3 + t\theta_2) - \frac{U_2}{2n\theta_3} - \frac{U_1}{2n(\theta_3 + t\theta_2)}$$
$$- \frac{t}{2(\theta_3 + t\theta_2)}(\bar{Y} - \theta_1)^2 \tag{3.15}$$

where $U_1 = t\sum_{i=1}^{n}(\bar{Y}_i - \bar{Y})^2$ and $U_2 = \sum_{i=1}^{n}\sum_{j=1}^{t}(Y_{ij} - \bar{Y}_i)^2$.

The first partial derivatives are

$$D_1 l(\theta) = \frac{\delta}{\delta\theta_1} l(\theta) = \frac{t}{\theta_3 + t\theta_2}(\bar{Y} - \theta_1)$$

$$D_2 l(\theta) = \frac{\delta}{\delta\theta_2} l(\theta) = -\frac{t}{2(\theta_3 + t\theta_2)} + \frac{tU_1}{2n(\theta_3 + t\theta_2)^2} + \frac{t^2}{2(\theta_3 + t\theta_2)^2}(\bar{Y} - \theta_1)^2$$

$$D_3 l(\theta) = \frac{\delta}{\delta\theta_3} l(\theta)$$
$$= -\frac{(t-1)}{2\theta_3} - \frac{1}{2(\theta_3 + t\theta_2)} + \frac{U_2}{2n\theta_3^2} + \frac{U_1}{2n(\theta_3 + t\theta_2)^2} + \frac{t}{2(\theta_3 + t\theta_2)^2}(\bar{Y} - \theta_1)^2.$$

Equating these first partial derivatives with zero, we get the maximum likelihood estimators (MLE)

$$\hat{\theta}_1 = \bar{Y},\ \hat{\theta}_2 = \frac{1}{t}\left(\frac{U_1}{n} - \frac{U_2}{n(t-1)}\right),\ \hat{\theta}_3 = \frac{U_2}{n(t-1)}$$

i.e. $\hat{\theta}_1 = \bar{Y}$, $\hat{\theta}_2 = \dfrac{(\mathrm{MSB} - \mathrm{MSW})}{t}$, $\hat{\theta}_3 = \mathrm{MSW}$ as given in Equation (3.4).

3.4 Bayesian tolerance interval with approximate frequentist validity

3.4.1 The setup

Following Ong and Mukerjee (2011), let $f(.;\theta)$ be the common t-variate normal density of $Y_1, Y_2, ..., Y_n$ where $Y_i = (Y_{i1}, ..., Y_{it})'$, $1 \le i \le n$. For $1 \le s, u, w \le 3$, we define

$$l(\theta)=n^{-1}\sum_{i=1}^{n}\log f(Y_i;\theta),\qquad c_{su}=-\{D_sD_ul(\theta)\}_{\theta=\hat{\theta}},\qquad a_{suw}=\{D_sD_uD_wl(\theta)\}_{\theta=\hat{\theta}}\tag{3.16}$$

where $D_s=\partial/\partial\theta_s$ and the matrix $C=(c_{su})$ is positive definite. The derivation of c_{su} and a_{suw} is available in Appendix A.

$C^{-1}=(c^{su})$ can be found almost instantly by using MATLAB symbolic computation. Hence,

$$c^{11}=\frac{\hat{\theta}_3+t\hat{\theta}_2}{t},\qquad c^{12}=c^{13}=c^{21}=c^{31}=0$$

$$c^{22}=2\left(\frac{\hat{\theta}_3^2}{t(t-1)}+\frac{1}{t}\hat{\theta}_2\hat{\theta}_3+\hat{\theta}_2^2\right)$$

$$c^{23}=c^{32}=-\frac{2\hat{\theta}_3^2}{t(t-1)}\qquad c^{33}=-\frac{2\hat{\theta}_3^2}{t-1}\tag{3.17}$$

From Appendix A,

$$a_{111}=a_{122}=a_{212}=a_{221}=a_{133}=a_{313}=a_{331}=a_{123}=a_{132}=a_{213}=a_{231}=a_{312}=a_{321}=0$$

$$a_{112}=a_{121}=a_{211}=\frac{t^2}{(\hat{\theta}_3+t\hat{\theta}_2)^2};\qquad a_{113}=a_{131}=a_{311}=\frac{t}{(\hat{\theta}_3+t\hat{\theta}_2)^2};$$

$$a_{222}=\frac{2t^3}{(\hat{\theta}_3+t\hat{\theta}_2)^3};\qquad a_{223}=a_{232}=a_{322}=\frac{2t^2}{(\hat{\theta}_3+t\hat{\theta}_2)^3};$$

$$a_{332}=a_{323}=a_{233}=\frac{2t}{(\hat{\theta}_3+t\hat{\theta}_2)^3};\qquad a_{333}=\frac{2(t-1)}{\hat{\theta}_3^3}+\frac{2}{(\hat{\theta}_3+t\hat{\theta}_2)^3}\tag{3.18}$$

3.4.2 Joint posterior density

Consider

$$\pi_s(\theta) = \frac{\delta}{\delta\theta_s}\pi(\theta) = D_s\pi(\theta), \quad s = 1, 2, 3$$

where $\pi(\theta)$ is a smooth prior. Let $\pi(\hat{\theta}) = \hat{\pi}$, $\pi_s(\hat{\theta}) = \frac{\delta}{\delta\theta_s}\pi(\theta)\bigg|_{\theta=\hat{\theta}} = \hat{\pi}_s$.

Let $h_s = \sqrt{n}(\theta_s - \hat{\theta}_s)$ $s = 1, 2, 3$. $Y = (Y_{ij})$ $i = 1, 2, ..., n$, $j = 1, 2, ..., t$ (collection of all the observations). According to Datta and Mukerjee (2004, Ch. 2), the posterior density of $h = (h_1, h_2, h_3)'$ is given by

$$\pi_{post}(h \mid Y) = \psi_3(h; C^{-1})\left[1 + n^{-1/2}\left\{\left(\frac{\hat{\pi}_s}{\hat{\pi}}\right)h_s + \frac{1}{6}a_{suw}h_s h_u h_w\right\}\right] + O_p(n^{-1}) \qquad (3.19)$$

where $\psi_3(\cdot; C^{-1}) \sim N_3(0, C^{-1})$ (trivariate normal with null mean vector and covariance matrix C^{-1}.

Remark 3.1: *In* (3.19) *and the rest of Section 3.4, the summation convention is followed with implicit sums over repeated sub- or superscripts ranging over* $\{1, 2, 3\}$ *(Ong and Mukerjee, 2011) i.e.* $a_{suw}h_s h_u h_w$ *in (3.19) represents* $\sum_{s=1}^{3}\sum_{u=1}^{3}\sum_{w=1}^{3}a_{suw}h_s h_u h_w$.

Under the balanced one-way random effects model in Equation (3.1), each $Y_{ij} \sim N(\theta_1, \theta_2 + \theta_3)$. Ong and Mukerjee (2011) considered a Bayesian tolerance interval, under a prior $\pi(\cdot)$ for the $N(\theta_1, \theta_2 + \theta_3)$. The limits, $\theta_1 \mp z(\theta_2 + \theta_3)^{1/2}$ cover a proportion β of this distribution where $z = \Phi^{-1}\left(\frac{1}{2}(1+\beta)\right)$ and $\Phi(\cdot)$ is a $N(0,1)$ distribution function. This motivates a tolerance interval with limits of the form

$\hat{\theta}_1 \mp z(b+n^{-1/2}g_1+n^{-1}g_2)^{1/2}$, where $b=\sqrt{\hat{\theta}_2+\hat{\theta}_3}$. g_1, g_2 are $O_p(1)$ functions of $Y=(Y_1,...,Y_n)$ which are determined so that the interval has $\beta-$content with posterior credibility level $\gamma+O_p(n^{-1})$. Thus, the two-sided tolerance interval in Ong and Mukerjee (2011), is given by

$$\left[\hat{\theta}_1+\frac{d}{n}-z\left(b+\frac{g_1}{\sqrt{n}}+\frac{g_2}{n}\right),\ \hat{\theta}_1+\frac{d}{n}+z\left(b+\frac{g_1}{\sqrt{n}}+\frac{g_2}{n}\right)\right] \tag{3.20}$$

Here d, g_1 and g_2 are functions of the observations $\{Y_{ij}\}$ which can potentially involve the prior $\pi(\cdot)$ and are of order $O_p(1)$. The tolerance interval in (3.20) is centered around $\hat{\theta}_1+\frac{d}{n}$. The presence of d in (3.20) induces flexibility in centering the interval; see Ong and Mukerjee (2011). We can center it at $\hat{\theta}_1$ by choosing $d=0$ or around the posterior mean or the posterior mode of θ_1 which are both of the form $\hat{\theta}_1+\frac{d}{n}$ with d being of order $O_p(1)$. Interestingly, the probability matching condition to be obtained in what follows, does not depend on the choice of d.

Let R be the content of the tolerance interval in (3.20) following the $N(\theta_1,\theta_2+\theta_3)$ distribution. Therefore, R is given by

$$R=\Phi(W_2)-\Phi(W_1), \tag{3.21}$$

where

$$W_2=\frac{\hat{\theta}_1+n^{-1}d+z(b+n^{-1/2}g_1+n^{-1}g_2)-\theta_1}{(\theta_2+\theta_3)^{1/2}},$$

$$W_1=\frac{\hat{\theta}_1+n^{-1}d-z(b+n^{-1/2}g_1+n^{-1}g_2)-\theta_1}{(\theta_2+\theta_3)^{1/2}}$$

g_1 and g_2 are determined such that

$$P^{\pi}(R \geq \beta \mid Y) = \gamma + O_p(n^{-1}), \tag{3.22}$$

where $P(\cdot \mid Y)$ is the posterior probability measure under the prior $\pi(\cdot)$.

$h_s = \sqrt{n}(\theta_s - \hat{\theta}_s)$, $s = 1, 2, 3$

$$\frac{h_2 + h_3}{\sqrt{n}} = \theta_2 + \theta_3 - (\hat{\theta}_2 + \hat{\theta}_3) \tag{3.23}$$

Now,

$$(\theta_2 + \theta_3)^{-\frac{1}{2}} = \left(\hat{\theta}_2 + \hat{\theta}_3 + \frac{h_2 + h_3}{\sqrt{n}}\right)^{-\frac{1}{2}} \qquad \text{[using (3.23)]}$$

$$= \left(b^2 + \frac{h_2 + h_3}{\sqrt{n}}\right)^{-\frac{1}{2}} \qquad \left[\text{since } b = \sqrt{\hat{\theta}_2 + \hat{\theta}_3}\ \right]$$

$$= \frac{1}{b}\left(1 + \frac{h_2 + h_3}{b^2\sqrt{n}}\right)^{-\frac{1}{2}}$$

$$= \frac{1}{b}\left[1 - \frac{1}{2}\left(\frac{h_2 + h_3}{b^2\sqrt{n}}\right) + \frac{3}{8}\frac{(h_2 + h_3)^2}{b^4 n}\right] + O_p(n^{-3/2}) \qquad \text{[Taylor expansion]}$$

$$= b^{-1}\left[1 - n^{-1/2}H + \frac{3}{2}n^{-1}H^2\right] + O_p(n^{-3/2}). \qquad \left[H = \frac{h_2 + h_3}{2b^2}\right]$$

Using the expansion above, we express

$W_2 = \dfrac{\hat{\theta}_1 + n^{-1}d + z(b + n^{-1/2}g_1 + n^{-1}g_2) - \theta_1}{(\theta_2 + \theta_3)^{1/2}}$ as

$$W_2 = [\hat{\theta}_1 + n^{-1}d + z(b + n^{-1/2}g_1 + n^{-1}g_2) - \theta_1]\ \frac{1}{b}\left[1 - n^{-1/2}H + \frac{3}{2}n^{-1}H^2\right] + O_p(n^{-3/2})$$

$$= [zb + n^{-1/2}(zg_1 - h_1) + n^{-1}(zg_2 + d)]\ b^{-1}\left[1 - n^{-1/2}H + \frac{3}{2}n^{-1}H^2\right] + O_p(n^{-3/2}).$$

[since $\theta_1 - \hat{\theta}_1 = n^{-1/2}h_1$]

By multiplying out and rearranging the terms above,

$$W_2 = z + n^{-1/2}A_1 + n^{-1}A_2 + O_p(n^{-3/2})$$

where $A_1 = \dfrac{zg_1 - h_1}{b} - \dfrac{z(h_2 + h_3)}{2b^2} = b^{-1}(zg_1 - h_1) - zH$,

and $$A_2 = \frac{zg_2 + d}{b} - \frac{(zg_1 - h_1)(h_2 + h_3)}{2b^3} + \frac{3}{8}\frac{z(h_2 + h_3)^2}{b^4}$$

$$= b^{-1}(zg_2 + d) - b^{-1}(zg_1 - h_1)H + \frac{3}{2}zH^2 .$$

This leads to

$$W_2 = z + b^{-1}\{n^{-1/2}[z(g_1 - bH) - h_1]$$

$$+ n^{-1}[z(g_2 - g_1H + \tfrac{3}{2}bH^2) + d + h_1H]\} + O_p(n^{-3/2})$$

$W_2 = z + n^{-1/2}A_1^* + n^{-1}A_2^* + O_p(n^{-3/2})$ where

$A_1^* = b^{-1}[z(g_1 - bH) - h_1]$ and $A_2^* = b^{-1}[z(g_2 - g_1H + \tfrac{3}{2}bH^2) + d + h_1H]$

For W_1, replace z by $-z$

$$W_1 = -z + b^{-1}\{n^{-1/2}[-z(g_1 - bH) - h_1]$$

$$+ n^{-1}[-z(g_2 - g_1H + \tfrac{3}{2}bH^2) + d + h_1H]\} + O_p(n^{-3/2})$$

$W_1 = -z + n^{-1/2}B_1^* + n^{-1}B_2^* + O_p(n^{-3/2})$ where

$B_1^* = b^{-1}[-z(g_1 - bH) - h_1]$ and $B_2^* = b^{-1}[-z(g_2 - g_1H + \tfrac{3}{2}bH^2) + d + h_1H]$

From Equation (3.21),

$$R = \Phi(W_2) - \Phi(W_1)$$

$$= \Phi(z + n^{-1/2}A_1^* + n^{-1}A_2^*) - \Phi(-z + n^{-1/2}B_1^* + n^{-1}B_2^*) + O_p(n^{-3/2})$$

$$= \Phi(z + n^{-1/2}A_1^* + n^{-1}A_2^*) + \Phi(z - n^{-1/2}B_1^* - n^{-1}B_2^*) - 1 + O_p(n^{-3/2})$$

By using the Taylor expansion and applying the fact $-\frac{\delta}{\delta z}\phi(z) = z\phi(z)$, we obtain

$$\Phi(z + n^{-1/2}A_1^* + n^{-1}A_2^*) = \Phi(z) + (n^{-1/2}A_1^* + n^{-1}A_2^*)\phi(z) - \tfrac{1}{2}n^{-1}A_1^{*2}z\phi(z) + \ldots$$

$$\Phi(z - n^{-1/2}B_1^* - n^{-1}B_2^*) = \Phi(z) - (n^{-1/2}B_1^* + n^{-1}B_2^*)\phi(z) - \tfrac{1}{2}n^{-1}B_1^{*2}z\phi(z) + \ldots$$

We retain the terms up to n^{-1} for the Taylor expansion above.

Thus,

$$R = 2\Phi(z) - 1 + n^{-1/2}(A_1^* - B_1^*)\phi(z) + n^{-1}[A_2^* - B_2^* - \tfrac{1}{2}z(A_1^{*2} + B_1^{*2})]\phi(z) + O_p(n^{-3/2})$$

$$= \beta + n^{-1/2}(A_1^* - B_1^*)\phi(z) + n^{-1}[A_2^* - B_2^* - \tfrac{1}{2}z(A_1^{*2} + B_1^{*2})]\phi(z) + O_p(n^{-3/2}) \qquad (3.24)$$

since $\Phi(z) = \frac{1+\beta}{2}$.

After some simplification,

$$\left.\begin{aligned} A_1^* - B_1^* &= 2z[b^{-1}g_1 - H] \\ A_2^* - B_2^* &= 2z[b^{-1}g_2 - b^{-1}g_1H + \tfrac{3}{2}H^2] \\ A_1^{*2} + B_1^{*2} &= 2z^2(b^{-1}g_1 - H)^2 + 2b^{-2}h_1^2 \end{aligned}\right\} \qquad (3.25)$$

Therefore, from (3.24) and (3.25),

$$R - \beta = \frac{2z\phi(z)}{\sqrt{n}}\left(\frac{g_1}{b} - H\right) + \frac{2z\phi(z)}{n}\left(\frac{g_2}{b} - \frac{g_1}{b}H + \frac{3}{2}H^2 - \frac{z^2}{2}\left(\frac{g_1}{b} - H\right)^2 - \frac{h_1^2}{2b^2}\right)$$

$$+ O_p(n^{-3/2})$$

$$= \frac{2z\phi(z)}{b\sqrt{n}}\left\{g_1 - bH + \frac{1}{\sqrt{n}}\left[g_2 - g_1H + \frac{3}{2}H^2b - \frac{z^2}{2}\frac{(g_1 - bH)^2}{b} - \frac{h_1^2}{2b}\right]\right\}$$

$$+ O_p(n^{-3/2})$$

$$= \frac{2z\phi(z)}{b\sqrt{n}}(g_1 - kX)$$

$$\therefore\ R = \beta + \frac{2z\phi(z)}{b\sqrt{n}}(g_1 - kX)$$

with $k = (\mu' C^{-1}\mu)^{1/2}$ and $\mu = \frac{1}{2b}(0, 1, 1)'$.

$X = X_0 + n^{-1/2}X_1 + O_p(n^{-1})$ where

$$X_0 = k^{-1}bH \text{ and } X_1 = k^{-1}\{g_1H - g_2 - \tfrac{3}{2}bH^2 + \tfrac{1}{2}b^{-1}[z^2(g_1 - bH)^2 + h_1^2]\} \tag{3.26}$$

$\therefore\ R - \beta \geq 0$ is equivalent to $g_1 - kX \geq 0$ since z and $\phi(z) > 0$ i.e. $R \geq \beta \Leftrightarrow X \leq \frac{g_1}{k}$.

Hence, $P_\pi(R \geq \beta \mid Y) = \gamma + O_p(n^{-1})$ is equivalent to $P_\pi\left(X \leq \frac{g_1}{k} \mid Y\right) = \gamma + O_p(n^{-1})$. To that effect, we consider the approximate posterior characteristic function and hence an expansion for the posterior density of X in the next section. This will facilitate the fact that the leading term in the posterior density of X is the standard univariate normal density. This is how the above representation of R in terms of X helps.

Posterior density of X

Let $\xi = (-1)^{1/2}\tau$ where τ is an auxiliary variable. Then,

$\exp(\xi X) = \exp(\xi X_0)[1 + n^{-1/2}X_1] + O_p(n^{-1})$.

Recalling (3.19), we get

$$\pi_{post}(h \mid Y)\exp(\xi X) = \phi_3(h; C^{-1})\exp(\xi X_0)\left\{1 + n^{-1/2}\left[\left(\frac{\hat{\pi}_s}{\hat{\pi}}\right)h_s + \frac{1}{6}a_{suw}h_sh_uh_w + \xi X_1\right]\right\}$$
$$+ O_p(n^{-1}) \tag{3.27}$$

Let $\lambda = (\lambda_1, \lambda_2, \lambda_3)' = \frac{1}{k}C^{-1}\mu$ to get

$$(h - \xi\lambda)'C(h - \xi\lambda) = h'Ch - 2\xi h'C\lambda + \xi^2\lambda'C\lambda$$

$$= h'Ch - \frac{2\xi}{k} h'\mu + \frac{\xi^2}{k^2} \mu' C^{-1} \mu$$

$$= h'Ch - \frac{\xi}{bk}(h_2 + h_3) + \xi^2 \qquad (k^2 = \mu' C^{-1} \mu)$$

$$\therefore\ -\frac{1}{2} h'Ch + \frac{\xi}{2bk}(h_2 + h_3) = \frac{1}{2}\xi^2 - \frac{1}{2}(h - \xi\lambda)'C(h - \xi\lambda)$$

Hence, $\phi_3(h; C^{-1}) \exp\left(\frac{\xi}{2bk}(h_2 + h_3)\right) = \phi_3(h; \xi\lambda, C^{-1}) \exp\left(\frac{1}{2}\xi^2\right)$ (3.28)

where $\phi_3(\cdot; \xi\lambda, C^{-1})$ is the trivariate normal density with mean vector $\xi\lambda$ and covariance matrix C^{-1}.

If $h = (h_1, ..., h_p)'$ has density $\phi_p(h; \xi\lambda, C^{-1})$, then

$$\left.\begin{aligned} &E[h_s] = \xi\lambda_s \\ &E[h_s h_u] = c^{su} + \xi^2 \lambda_s \lambda_u \\ &E[h_s h_u h_w] = \xi^3 \lambda_s \lambda_u \lambda_w + \xi(\lambda_s c^{uw} + \lambda_u c^{sw} + \lambda_w c^{uw}) \end{aligned}\right\} \qquad (3.29)$$

Then by applying (3.29), we integrate Equation (3.27) with respect to h to obtain the approximate posterior characteristic function of X under $\pi(\cdot)$ that is (note that $\lambda_1 = 0$)

$$\begin{aligned} &\exp\left(\frac{1}{2}\xi^2\right)\Bigg\{1 + n^{-1/2}\Bigg[\left(\frac{\hat{\pi}_s}{\hat{\pi}}\right)\xi\lambda_s + \frac{1}{6} a_{suw}\left(\xi^3 \lambda_s \lambda_u \lambda_w + \xi\lambda_s c^{uw} + \xi\lambda_u c^{sw} + \xi\lambda_w c^{su}\right) \\ &\qquad + \frac{\xi^2 g_1}{2kb^2}(\lambda_2 + \lambda_3) - \frac{\xi g_2}{k} - \frac{3}{8}\frac{\xi}{kb^3}\left(c^{22} + 2c^{23} + c^{33} + \xi^2(\lambda_2 + \lambda_3)^2\right) \\ &\qquad + \frac{\xi c^{11}}{2kb} + \frac{\xi z^2}{2kb}\left(g_1^2 - \frac{\xi g_1}{b}(\lambda_2 + \lambda_3) + \frac{1}{4b^2}(c^{22} + 2c^{23} + c^{33} + \xi^2(\lambda_2 + \lambda_3)^2\right)\Bigg]\Bigg\} \\ &\qquad + O_p(n^{-1}) \end{aligned} \qquad (3.30)$$

We write (3.30) in the form

$$\exp\left(\tfrac{1}{2}\xi^2\right)\left\{1 + n^{-1/2}\left[L_1\xi + L_2\xi^2 + L_3\xi^3\right]\right\} + O_p(n^{-1}) \qquad (3.31)$$

where $L_1 = L_{11} + L_{12}$, $L_2 = \frac{g_1}{k} L_{21}$ and

$$L_{11} = \frac{\hat{\pi}_s}{\hat{\pi}} \lambda_s + \frac{1}{2} a_{suw} \lambda_s c^{uw} + \frac{(z^2 - 3)}{8kb^3} (c^{22} + 2c^{23} + c^{33}) + \frac{c^{11}}{2kb}$$

$$L_{12} = \frac{z^2 g_1^2}{2kb} - \frac{g_2}{k}$$

$$L_{21} = \frac{1}{2b^2} (\lambda_2 + \lambda_3) - \frac{z^2}{2b^2} (\lambda_2 + \lambda_3) = -\frac{(\lambda_2 + \lambda_3)(z^2 - 1)}{2b^2}$$

$$L_3 = \frac{1}{6} a_{suw} \lambda_s \lambda_u \lambda_w + \frac{(z^2 - 3)(\lambda_2 + \lambda_3)^2}{8kb^3}$$

$$\varepsilon = (0, 1, 1)', \ \mu = \frac{1}{2b} \varepsilon, \ k = \frac{1}{2} b^{-1} (\varepsilon' C^{-1} \varepsilon)^{1/2}, \quad \lambda = (\lambda_1, \lambda_2, \lambda_3)' = \frac{1}{2} (kb)^{-1} C^{-1} \varepsilon \tag{3.32}$$

After some simplification (see Appendix A), we get

$$L_{11} = \frac{\hat{\pi}_s}{\hat{\pi}} \lambda_s + \frac{1}{2} a_{suw} \lambda_s c^{uw} + \frac{1}{2} kb^{-1} (z^2 - 3) + \frac{1}{2} (kb)^{-1} c^{11} \tag{3.33}$$

$$L_{12} = \frac{1}{2} (kb)^{-1} z^2 g_1^2 - k^{-1} g_2 \tag{3.34}$$

$$L_{21} = -kb^{-1} (z^2 - 1) \tag{3.35}$$

$$L_3 = \frac{1}{6} a_{suw} \lambda_s \lambda_u \lambda_w + \frac{1}{2} kb^{-1} (z^2 - 3) \tag{3.36}$$

Note that, L_{11}, L_{21} and L_3 do not involve g_1 and g_2 while L_{12} does. Retaining this distinction helps in simplifying the notations later. The prior only appears in L_{11} through the term $\frac{\hat{\pi}_s}{\hat{\pi}} \lambda_s$.

Since $\int_{-\infty}^{\infty} e^{\xi x} H_j(x)\phi(x)dx = \xi^j e^{\frac{1}{2}\xi^2}$, where $\phi(\cdot)$ is the standard univariate normal density and $H_j(\cdot)$ is the Hermite polynomial of degree *j*, inverting (3.31), we now get the posterior density of *X*, under $\pi(\cdot)$ as:

$$\left\{1 + n^{-1/2}[L_1 x + L_2(x^2-1) + L_3(x^3-3x)]\right\}\phi(x) + O_p(n^{-1}) \tag{3.37}$$

Integrating (3.37) over $(-\infty, g_1/k]$,

$$\begin{aligned}
P^{\pi}(R \geq \beta \mid Y) &= P^{\pi}\left(X \leq \frac{g_1}{k} \mid Y \right) \\
&= \Phi(x) - n^{-1/2}\left\{L_1 + L_2 x + L_3(x^2-1)\right\}\phi(x)\Big|_{-\infty}^{g_1/k} + O_p(n^{-1}) \\
&= \Phi\left(\frac{g_1}{k}\right) - \frac{1}{\sqrt{n}}\left[L_1 + L_2\frac{g_1}{k} + L_3\left(\frac{g_1^2}{k^2} - 1\right)\right]\phi\left(\frac{g_1}{k}\right) + O_p(n^{-1})
\end{aligned}$$

$$\therefore\ P^{\pi}(R \geq \beta \mid Y) = \Phi\left(\frac{g_1}{k}\right) - n^{-1/2}\left\{L_{11} + L_{12} - L_3 + (L_{21} + L_3)\left(\frac{g_1}{k}\right)^2\right\}\phi\left(\frac{g_1}{k}\right) + O_p(n^{-1}) \tag{3.38}$$

Let $\Phi\left(\frac{g_1}{k}\right) = \Phi(q) = \gamma$, Then, $P^{\pi}(R \geq \beta \mid Y) = \gamma + O_p(n^{-1})$ provided $g_1 = kq$ and $L_1 + L_2 q + L_3(q^2 - 1) = 0$ i.e.

$$L_{11} + \frac{z^2 k}{2b}q^2 - \frac{g_2}{k} + L_{21}q^2 + L_3(q^2-1) = 0 \text{ i.e. } g_2 = k\left[L_{11} + \left(\frac{z^2 k}{2b} + L_{21}\right)q^2 + L_3(q^2-1)\right]$$

$$\therefore\ g_1 = kq \text{ and } g_2 = k\left[L_{11} + \left(\frac{z^2 k}{2b} + L_{21}\right)q^2 + L_3(q^2-1)\right] \tag{3.39}$$

3.4.3 Matching condition

Applying g_1 and g_2 in (3.39), the tolerance interval in (3.20) has $\beta-$content with posterior credibility level $\gamma+O_p(n^{-1})$. We now characterize priors for which it has $\beta-$content also with frequentist confidence level $\gamma+O(n^{-1})$. Such priors will be probability matching in the present context of two-sided tolerance limits.

With the above objective, we now study the frequentist coverage $P_\theta(R\geq\beta)=P_\theta(X\leq g_1/k)$, with g_1 and g_2 in (3.39). Ong and Mukerjee (2011) employed the shrinkage argument as shown in Datta and Mukerjee (2004) which involves the following steps:

Step 1: Consider an auxiliary prior $\pi^*(\theta)$ which vanishes on the boundaries of a rectangle containing the true θ and obtain $P^{\pi^*}(R\geq\beta\,|\,Y)$ with margin of error $O_p(n^{-1})$. As in the previous section, the posterior density of *X*, under $\pi^*(\theta)$, turns out to be

$$\left\{1+n^{-1/2}[L_1^*x+L_2(x^2-1)+L_3(x^3-3x)]\right\}\phi(x)+O_p(n^{-1}) \tag{3.40}$$

where $L_1^*=L_{11}^*+L_{12}$ and $L_2=qL_{21}$,

L_{11}^* is similar to L_{11} in (3.33) with the only change that is the term $\frac{\hat{\pi}_s}{\hat{\pi}}\lambda_s$ is replaced by $\frac{\hat{\pi}_s^*}{\hat{\pi}^*}\lambda_s$, i.e.

$$L_{11}^*=L_{11}-\left\{\left(\frac{\hat{\pi}_s}{\hat{\pi}}\right)-\left(\frac{\hat{\pi}_s^*}{\hat{\pi}^*}\right)\right\}\lambda_s \tag{3.41}$$

and L_{12}, L_{21}, L_3 are as in equations (3.34)-(3.36), with g_1 and g_2 as shown in (3.39).

As usual $\pi^*(\hat{\theta}) = \hat{\pi}^*$, $\pi_s^*(\hat{\theta}) = \hat{\pi}_s^*$, with $\pi_s^*(\theta) = D_s \pi^*(\theta)$.

Integrating the posterior density of X under $\pi^*(\cdot)$,

$$P^{\pi^*}(R \geq \beta \mid Y) = P^{\pi^*}\left(X \leq \frac{g_1}{k} \mid Y \right)$$

$$= \Phi(q) - n^{-1/2}\{L_1^* + L_2 q + L_3(q^2-1)\}\phi(q) + O_p(n^{-1}) \qquad [\text{as } g_1 = kq\,]$$

$$= \gamma - n^{-1/2}\left\{ L_{11} + \left[\left(\frac{\hat{\pi}_s^*}{\hat{\pi}^*}\right) - \left(\frac{\hat{\pi}_s}{\hat{\pi}}\right)\right]\lambda_s + \frac{z^2 k}{2b} q^2 - \frac{g_2}{k} + L_{21} q^2 + L_3(q^2-1) \right\}\phi(q)$$
$$+ O_p(n^{-1})$$

$$= \gamma - n^{-1/2}\left\{ L_{11} + \left[\left(\frac{\hat{\pi}_s^*}{\hat{\pi}^*}\right) - \left(\frac{\hat{\pi}_s}{\hat{\pi}}\right)\right]\lambda_s + \frac{z^2 k}{2b} q^2 - L_{11} - \left(\frac{z^2 k}{2b} + L_{21}\right) q^2 \right.$$
$$\left. - L_3(q^2-1) + L_{21} q^2 + L_3(q^2-1) \right\}\phi(q) + O_p(n^{-1})$$

$$= \gamma - n^{-1/2}\left\{ \left[\left(\frac{\hat{\pi}_s^*}{\hat{\pi}^*}\right) - \left(\frac{\hat{\pi}_s}{\hat{\pi}}\right)\right]\lambda_s \right\}\phi(q) + O_p(n^{-1}). \qquad (3.42)$$

Before proceeding further, we are required to obtain λ_1, λ_2 and λ_3 explicitly. From (3.32), $\lambda = \frac{1}{k} C^{-1}\mu$ where $k = (\mu' C^{-1} \mu)^{1/2}$ and $\mu = \frac{1}{2b}(0,1,1)'$.

From Equation (3.17),

$c^{12} = c^{13} = c^{21} = c^{31} = 0$. Thus,

$$C^{-1}\mu = \frac{1}{2b}\begin{pmatrix} c^{11} & 0 & 0 \\ 0 & c^{22} & c^{23} \\ 0 & c^{32} & c^{33} \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{2b}\begin{pmatrix} 0 \\ c^{22} + c^{23} \\ c^{32} + c^{33} \end{pmatrix}$$

$c^{22} + c^{23} = 2\psi_2(\hat{\theta})$, where $\psi_2(\theta) = \frac{2}{t}\theta_2\theta_3 + \theta_2^2$

$c^{32} + c^{33} = 2\psi_3(\hat{\theta})$, where $\psi_3(\theta) = \dfrac{\theta_3^2}{t}$,

Therefore, $k^2 = \mu' C^{-1} \mu$

$$= \frac{1}{(2b)^2}(c^{22} + 2c^{23} + c^{33})$$

$$= \frac{1}{2b^2}\left(\psi_2(\hat{\theta}) + \psi_3(\hat{\theta})\right) = \frac{1}{b^2}\left(\psi(\hat{\theta})\right)^2$$

where $\psi(\theta) = \left\{\dfrac{1}{2}\left[\psi_2(\theta) + \psi_3(\theta)\right]\right\}^{\frac{1}{2}} = \left\{\dfrac{1}{2}t^{-1}\left[t\theta_2^2 + 2\theta_2\theta_3 + \theta_3^2\right]\right\}^{\frac{1}{2}}$

Hence, $\lambda = \dfrac{1}{k}C^{-1}\mu = \dfrac{1}{2bk}\begin{pmatrix} 0 \\ 2\psi_2(\hat{\theta}) \\ 2\psi_3(\hat{\theta}) \end{pmatrix} = \dfrac{1}{bk}\begin{pmatrix} 0 \\ \psi_2(\hat{\theta}) \\ \psi_3(\hat{\theta}) \end{pmatrix}$.

Since $k^2 = \dfrac{1}{b^2}\left(\psi(\hat{\theta})\right)^2$, we get $\lambda_1 = 0$, $\lambda_2 = \dfrac{\psi_2(\hat{\theta})}{\psi(\hat{\theta})}$, $\lambda_3 = \dfrac{\psi_3(\hat{\theta})}{\psi(\hat{\theta})}$.

The required notations are summarized as follows

$$\lambda_s = \frac{\psi_s(\hat{\theta})}{\psi(\hat{\theta})}$$

where

$$\psi_1(\theta) = 0,\ \psi_2(\theta) = \frac{2}{t}\theta_2\theta_3 + \theta_2^2,\ \psi_3(\theta) = \frac{\theta_3^2}{t}$$

$$\psi(\theta) = \left\{\frac{1}{2}t^{-1}\left[t\theta_2^2 + 2\theta_2\theta_3 + \theta_3^2\right]\right\}^{\frac{1}{2}} \tag{3.43}$$

Hence via substituting (3.43), Equation (3.42) becomes

$$P^{\pi^*}(R \ge \beta \mid Y) = \gamma - n^{-1/2}\left\{\left[\frac{\hat{\pi}_2^*}{\hat{\pi}^*} - \frac{\hat{\pi}_2}{\hat{\pi}}\right]\frac{\psi_2(\hat{\theta})}{\psi(\hat{\theta})} + \left[\frac{\hat{\pi}_3^*}{\hat{\pi}^*} - \frac{\hat{\pi}_3}{\hat{\pi}}\right]\frac{\psi_3(\hat{\theta})}{\psi(\hat{\theta})}\right\}\phi(q) + O_p(n^{-1}).$$

We resume to the next step of the shrinkage argument.

Step 2: Find $E_\theta[P^{\pi^*}(R \geq \beta \mid Y)]$ with margin of error $O(n^{-1})$.

$$E_\theta[P^{\pi^*}(R \geq \beta \mid Y)] = \gamma - n^{-1/2}\left\{\left[\frac{\pi_2^*(\theta)}{\pi^*(\theta)} - \frac{\pi_2(\theta)}{\pi(\theta)}\right]\frac{\psi_2(\theta)}{\psi(\theta)} + \left[\frac{\pi_3^*(\theta)}{\pi^*(\theta)} - \frac{\pi_3(\theta)}{\pi(\theta)}\right]\frac{\psi_3(\theta)}{\psi(\theta)}\right\}\phi(q) + O(n^{-1}).$$

Step 3: Integrate $E_\theta[P^{\pi^*}(R \geq \beta \mid Y)]$ by parts with respect to $\pi^*(\cdot)$ and finally allow $\pi^*(\cdot)$ to converge weakly to the degenerate prior at the true θ; see Datta and Mukerjee, 2004, Ch. 1. These steps yield the frequentist confidence level of our tolerance interval to have $\beta-$content as

$$P_\theta(R \geq \beta) = \gamma - n^{-1/2}\left\{-D_2\left(\frac{\psi_2(\theta)}{\psi(\theta)}\right) - \frac{\pi_2(\theta)}{\pi(\theta)}\frac{\psi_2(\theta)}{\psi(\theta)} - D_3\left(\frac{\psi_3(\theta)}{\psi(\theta)}\right) - \frac{\pi_3(\theta)}{\pi(\theta)}\frac{\psi_3(\theta)}{\psi(\theta)}\right\}\phi(q) + O(n^{-1})$$

$$= \gamma + n^{-1/2}\left\{\frac{1}{\pi(\theta)}D_2\left(\frac{\psi_2(\theta)}{\psi(\theta)}\pi(\theta)\right) + \frac{1}{\pi(\theta)}D_3\left(\frac{\psi_3(\theta)}{\psi(\theta)}\pi(\theta)\right)\right\}\phi(q) + O(n^{-1})$$

$$P_\theta(R \geq \beta) = \gamma + n^{-1/2}[\pi(\theta)]^{-1}\left\{D_s\left(\frac{\psi_s(\theta)}{\psi(\theta)}\pi(\theta)\right)\right\}\phi(q) + O(n^{-1}) \qquad (3.44)$$

Equation (3.44) becomes $\gamma + O(n^{-1})$ provided

$$D_s\left\{\left[\frac{\psi_s(\theta)}{\psi(\theta)}\right]\pi(\theta)\right\} = 0 \qquad (3.45)$$

where $s = 1, 2, 3$.

Therefore, the two-sided tolerance interval in (3.20) has $\beta-$content with posterior credibility level $\gamma + O_p(n^{-1})$ as well as frequentist confidence level $\gamma + O(n^{-1})$. This is only possible if and only if the prior $\pi(\cdot)$ satisfies (3.45). It is interesting to see that the matching condition given in (3.45) is free from β and γ. Moreover, we cognize that

(3.45) does not depend on the quantity d i.e. the matching condition works irrespective of the centering. This is anticipated because neither g_1 nor g_2 in (3.39) involves d.

We proceed to solving Equation (3.45). Using (3.43), it is not difficult to see that

$$\pi(\theta) = \frac{\psi(\theta)}{[\theta_3(\theta_3 + t\theta_2)]^2} \tag{3.46}$$

solves (3.45). In fact, as mentioned in Ong and Mukerjee (2011), the solution in (3.46) belongs to a more general class of solutions as given by

$$\pi(\theta) = \frac{\psi(\theta)}{\theta_2^2}\,\omega\left(\frac{\theta_2}{\theta_3(\theta_3 + t\theta_2)}\right) \tag{3.47}$$

where $\omega(\cdot)\ (>0)$ is any smooth function.

Since

$$\psi(\theta) = \left\{\frac{1}{2}t^{-1}\left[t\theta_2^2 + 2\theta_2\theta_3 + \theta_3^2\right]\right\}^{\frac{1}{2}} < \left\{\frac{1}{2}t^{-1}\left[t^2\theta_2^2 + 2t\theta_2\theta_3 + \theta_3^2\right]\right\}^{\frac{1}{2}} = \frac{\theta_3 + t\theta_2}{\sqrt{2t}},$$

proceeding along the lines in Datta and Mukerjee (2004, pp. 51-52), it can be shown that the solution in (3.46) ensures the propriety of the posterior for every $n \geq 2$. However, not all solutions of the form (3.47) do so. For instance, the solution

$$\pi(\theta) = \frac{\psi(\theta)}{\theta_2^2}, \tag{3.48}$$

which corresponds to $\omega(\cdot) = 1$ fails to ensure the propriety of the posterior for any $n \geq 2$.

Next, we verify that any prior of the form (3.47) satisfies the condition (3.45). Thus recalling (3.43) we obtain,

$$D_2\psi_2(\theta) = \frac{2\theta_3}{t} + 2\theta_2 \text{ and } D_3\psi_3(\theta) = \frac{2\theta_3}{t}$$

$$D_2\psi(\theta) = \frac{1}{2\psi(\theta)}\left(\frac{1}{t}\theta_3 + \theta_2\right) \text{ and } D_3\psi(\theta) = \frac{\theta_2 + \theta_3}{2t\psi(\theta)} \tag{3.49}$$

For any such prior in (3.47), the left-hand side, LHS of Equation (3.45) is:

$$D_2\left\{\frac{\psi_2(\theta)}{\theta_2^2}\omega\left(\frac{\theta_2}{\theta_3(\theta_3+t\theta_2)}\right)\right\}+D_3\left\{\frac{\psi_3(\theta)}{\theta_2^2}\omega\left(\frac{\theta_2}{\theta_3(\theta_3+t\theta_2)}\right)\right\}$$

$$=D_2\left(\frac{\psi_2(\theta)}{\theta_2^2}\right)\omega\left(\frac{\theta_2}{\theta_3(\theta_3+t\theta_2)}\right)+\frac{\psi_2(\theta)}{\theta_2^2}\omega'\left(\frac{\theta_2}{\theta_3(\theta_3+t\theta_2)}\right)D_2\left(\frac{\theta_2}{\theta_3(\theta_3+t\theta_2)}\right)$$
$$+D_3\left(\frac{\psi_3(\theta)}{\theta_2^2}\right)\omega\left(\frac{\theta_2}{\theta_3(\theta_3+t\theta_2)}\right)+\frac{\psi_3(\theta)}{\theta_2^2}\omega'\left(\frac{\theta_2}{\theta_3(\theta_3+t\theta_2)}\right)D_3\left(\frac{\theta_2}{\theta_3(\theta_3+t\theta_2)}\right)$$

Now, $\frac{\psi_2(\theta)}{\theta_2^2}=\frac{2\theta_3}{t\theta_2}+1$, $D_2\left(\frac{\psi_2(\theta)}{\theta_2^2}\right)=-\frac{2\theta_3}{t\theta_2^2}$ and $\frac{\psi_3(\theta)}{\theta_2^2}=\frac{\theta_3^2}{t\theta_2^2}$, $D_3\left(\frac{\psi_3(\theta)}{\theta_2^2}\right)=\frac{2\theta_3}{t\theta_2^2}$

i.e. $D_2\left(\frac{\psi_2(\theta)}{\theta_2^2}\right)+D_3\left(\frac{\psi_3(\theta)}{\theta_2^2}\right)=0$, so that under the prior in (3.47),

$$\omega'\left(\frac{\theta_2}{\theta_3(\theta_3+t\theta_2)}\right)\left[\frac{\psi_2(\theta)}{\theta_2^2}D_2\left(\frac{\theta_2}{\theta_3(\theta_3+t\theta_2)}\right)+\frac{\psi_3(\theta)}{\theta_2^2}D_3\left(\frac{\theta_2}{\theta_3(\theta_3+t\theta_2)}\right)\right] \tag{3.50}$$

$D_2\left(\frac{\theta_2}{\theta_3(\theta_3+t\theta_2)}\right)=\frac{1}{(\theta_3+t\theta_2)^2}$ and $D_3\left(\frac{\theta_2}{\theta_3(\theta_3+t\theta_2)}\right)=-\frac{\theta_2(2\theta_3+t\theta_2)}{\theta_3^2(\theta_3+t\theta_2)^2}$.

After some simplification, from (3.50), we can see that,

$\frac{\psi_2(\theta)}{\theta_2^2}D_2\left(\frac{\theta_2}{\theta_3(\theta_3+t\theta_2)}\right)+\frac{\psi_3(\theta)}{\theta_2^2}D_3\left(\frac{\theta_2}{\theta_3(\theta_3+t\theta_2)}\right)=0$ i.e. any prior of the form (3.47) satisfies the matching condition in (3.45).

We proceed to show that the prior $\frac{\psi(\theta)}{\theta_2^2}$ in (3.48) does not ensure the propriety of the posterior for any *n*. For this we show that $\int \exp(nl(\theta))\frac{\psi(\theta)}{\theta_2^2}d\theta=\infty$, i.e.

$$I=\int\frac{1}{\theta_3^{n(t-1)/2}(\theta_3+t\theta_2)^{n/2}}\exp\left[-\frac{U_2}{2\theta_3}-\frac{U_1}{2(\theta_3+t\theta_2)}-\frac{tn(\theta_1-\bar{Y})^2}{2(\theta_3+t\theta_2)}\right]\frac{\psi(\theta)}{\theta_2^2}d\theta=\infty$$

where $U_1=t\sum_{i=1}^{n}(\bar{Y}_i-\bar{Y})^2$ and $U_2=\sum_{i=1}^{n}\sum_{j=1}^{t}(Y_{ij}-\bar{Y}_i)^2$; see (3.15).

Since $\psi(\theta)$ does not involve θ_1 and $\int_{-\infty}^{\infty} \exp\left[-\frac{tn(\theta_1 - \bar{Y})^2}{2(\theta_3 + t\theta_2)}\right] d\theta_1 = k_1\sqrt{\theta_3 + t\theta_2}$ where k_1 is a constant free from θ, we eventually get

$$I = \int_0^\infty \int_0^\infty \frac{k_1}{\theta_3^{n(t-1)/2}(\theta_3 + t\theta_2)^{(n-1)/2}} \exp\left[-\frac{U_2}{2\theta_3} - \frac{U_1}{2(\theta_3 + t\theta_2)}\right] \frac{\psi(\theta)}{\theta_2^2} d\theta_2 d\theta_3$$

We now observe that for every $x > 0$, $(t-1)x + \frac{1}{x} \geq 2\sqrt{t-1}$ (by simple differentiation).

Taking $x = \frac{\theta_3}{\theta_3 + t\theta_2}$ we get

$$2\sqrt{t-1} \leq \frac{(t-1)\theta_3}{\theta_3 + t\theta_2} + \frac{\theta_3 + t\theta_2}{\theta_3} = \frac{2t^2\{\psi(\theta)\}^2}{\theta_3(\theta_3 + t\theta_2)}, \text{ i.e.}$$

$$\psi(\theta) \geq \left(\frac{\sqrt{t-1}}{t^2}\right)^{1/2} \sqrt{\theta_3(\theta_3 + t\theta_2)}, \text{ so that}$$

$$I \geq k_2 \int_0^\infty \int_0^\infty \frac{1}{\theta_3^{\frac{n(t-1)-1}{2}} (\theta_3 + t\theta_2)^{\frac{n}{2}-1}} \exp\left[-\frac{U_2}{2\theta_3} - \frac{U_1}{2(\theta_3 + t\theta_2)}\right] \frac{1}{\theta_2^2} d\theta_2 d\theta_3$$

[$k_2(>0)$ is free from θ]

$$I \geq k_2 \int_1^2 \int_0^\infty \frac{1}{\theta_3^{\frac{n(t-1)-1}{2}} (\theta_3 + t\theta_2)^{\frac{n}{2}-1}} \exp\left[-\frac{U_2}{2\theta_3} - \frac{U_1}{2(\theta_3 + t\theta_2)}\right] \frac{1}{\theta_2^2} d\theta_2 d\theta_3$$

Now $n \geq 2$ implies that $\frac{n(t-1)-1}{2} \geq \frac{n-1}{2} > 0$ and $\frac{n}{2} - 1 \geq 0$.

Hence, for $n \geq 2$, if $1 \leq \theta_3 \leq 2$ then

$$\frac{1}{\theta_3^{\frac{n(t-1)-1}{2}}} \geq \frac{1}{2^{\frac{n(t-1)-1}{2}}}, \quad \frac{1}{(\theta_3 + t\theta_2)^{\frac{n}{2}-1}} \geq \frac{1}{(2 + t\theta_2)^{\frac{n}{2}-1}},$$

$$\frac{U_2}{2\theta_3} \leq \frac{U_2}{2}, \quad \frac{U_1}{2(\theta_3 + t\theta_2)} \leq \frac{U_1}{2(1 + t\theta_2)}$$

Therefore,

$$I \geq k_3 \int_1^2 \int_0^\infty \frac{1}{(2+t\theta_2)^{\frac{n}{2}-1}} \exp\left[-\frac{U_2}{2} - \frac{U_1}{2(1+t\theta_2)}\right] \frac{1}{\theta_2^2} \, d\theta_2 d\theta_3 \qquad [k_3(>0) \text{ is free from } \theta]$$

$$= k_4 \int_0^\infty \frac{1}{(2+t\theta_2)^{\frac{n}{2}-1}} \exp\left[-\frac{U_1}{2(1+t\theta_2)}\right] \frac{1}{\theta_2^2} \, d\theta_2 \qquad [k_4(>0) \text{ is free from } \theta]$$

$$= k_4 \int_0^1 \frac{\exp\left[-\frac{1}{2}(U_1)w\right]}{\left(1+\frac{1}{w}\right)^{\frac{n}{2}-1}} \frac{w^2 t^2}{(1-w)^2} \frac{1}{tw^2} \, dw$$

$$\left[\text{where } w = \frac{1}{1+t\theta_2} \text{ i.e. } \theta_2 = \frac{1-w}{tw},\ d\theta_2 = -\frac{1}{tw^2} dw\right]$$

$$= k_5 \int_0^1 \left(\frac{w}{1+w}\right)^{\frac{n}{2}-1} \frac{\exp\left[-\frac{1}{2}(U_1)w\right]}{(1-w)^2} \frac{1}{tw^2} \, dw \qquad [k_5(>0) \text{ is free from } \theta]$$

$$\geq k_5 \int_{1/2}^1 \left(\frac{w}{1+w}\right)^{\frac{n}{2}-1} \frac{\exp\left(-\frac{1}{2}(U_1)w\right)}{(1-w)^2} \, dw$$

Now, $\frac{1}{2} \leq w \leq 1 \Rightarrow \frac{w}{1+w} \geq \frac{1}{2}\big/\left(1+\frac{1}{2}\right) = \frac{1}{3}$ and $\frac{1}{2}(U_1)w \leq \frac{1}{2}U_1$. Thus, with $k_6(>0)$ free from θ,

$$I \geq k_6 \int_{1/2}^1 \frac{dw}{(1-w)^2} = k_6 \left[\frac{1}{1-w}\right]_{1/2}^{\infty} = \infty \text{, as we had claimed.}$$

We next show that the prior $\pi(\theta)=\dfrac{\psi(\theta)}{[\theta_3(\theta_3+t\theta_2)]^2}$ in (3.46) ensures the propriety of the posterior for every $n \geq 2$. As noted earlier, $\psi(\theta)<\dfrac{\theta_3+t\theta_2}{\sqrt{2t}}$ and hence it suffices to show that $\displaystyle\int \exp\big(nl(\theta)\big)\frac{d\theta}{\theta_3^2(\theta_3+t\theta_2)}<\infty$.

Proceeding as with the other prior $\pi(\theta)=\dfrac{\psi(\theta)}{\theta_2^2}$, we get

$$I=\int_0^\infty\int_0^\infty \frac{k_7}{\theta_3^{n(t-1)/2}(\theta_3+t\theta_2)^{n/2}}\exp\left[-\frac{U_2}{2\theta_3}-\frac{U_1}{2(\theta_3+t\theta_2)}\right]\frac{\sqrt{\theta_3+t\theta_2}}{\theta_3^2(\theta_3+t\theta_2)}d\theta_2 d\theta_3$$

[$k_7(>0)$ is free from θ]

$$=k_7\int_0^\infty\int_0^\infty \frac{1}{\theta_3^{\frac{n(t-1)}{2}+2}(\theta_3+t\theta_2)^{\frac{n+1}{2}}}\exp\left[-\frac{U_2}{2\theta_3}-\frac{U_1}{2(\theta_3+t\theta_2)}\right]d\theta_2 d\theta_3$$

Transform $u=\dfrac{1}{\theta_3+t\theta_2}$, $v=\dfrac{1}{\theta_3}$; then $u^{-1}>v^{-1}>0$ i.e. $v>u>0$; also $\theta_2=\dfrac{1}{t}\left(\dfrac{1}{u}-\dfrac{1}{v}\right)$, $\theta_3=\dfrac{1}{v}$, so that $|J|=\begin{vmatrix}-\dfrac{1}{tu^2} & \dfrac{1}{tv^2}\\ 0 & -\dfrac{1}{v^2}\end{vmatrix}=\dfrac{1}{tu^2v^2}$.

Therefore, $I=\dfrac{k_7}{t}\displaystyle\int_0^\infty\int_u^\infty v^{\frac{n(t-1)}{2}}u^{\frac{n-3}{2}}\exp\left[-\frac{1}{2}U_2v-\frac{1}{2}U_1u\right]dvdu$.

Since $\dfrac{n(t-1)}{2}>0$, we have $I<\infty$ whenever $\dfrac{n-3}{2}>-1$ i.e. $n\geq 2$.

3.4.4 On other priors

The elements of the Fisher information matrix for the balanced one-way random effects model are as follows:

$$I_{11} = \frac{t}{\theta_3 + t\theta_2}, \; I_{12} = I_{21} = I_{13} = I_{31} = 0, \; I_{22} = \frac{t^2}{2(\theta_3 + t\theta_2)^2}$$

$$I_{23} = I_{32} = \frac{t}{2(\theta_3 + t\theta_2)^2}, \; I_{33} = \frac{1}{2}\left[\frac{t-1}{\theta_3^2} + \frac{1}{(\theta_3 + t\theta_2)^2}\right]$$

Thus, $|I| \propto \dfrac{1}{\theta_3^2(\theta_3 + t\theta_2)^3}$, i.e. the Jeffreys' prior is given by

$$\pi_J(\theta) = \frac{1}{\theta_3(\theta_3 + t\theta_2)^{3/2}} \tag{3.51}$$

Similarly $\begin{vmatrix} I_{22} & I_{23} \\ I_{32} & I_{33} \end{vmatrix} \propto \dfrac{1}{\theta_3^2(\theta_3 + t\theta_2)^2}$ and the modified version of the Jeffreys' prior based on the principal sub-matrix of I associated with θ_2 and θ_3 is given by $\dfrac{1}{\theta_3(\theta_3 + t\theta_2)}$.

Wolfinger (1998) considered the non-informative prior

$$\pi_W(\theta) = \frac{1}{\theta_3(\theta_3 + t\theta_2)}, \tag{3.52}$$

based on the modified version of the Jeffreys' prior. $\pi_W(\theta)$ enjoys the matching property of both θ_3 and $\dfrac{\theta_3}{\theta_2}$ (Datta and Mukerjee, 2004, pp.39). However, the priors in (3.51) and (3.52) do not satisfy the matching conditions given in (3.45) and thus, are not probability matching for tolerance intervals. (Ong and Mukerjee, 2011).

We show that neither Jeffreys' prior nor the non-informative prior by Wolfinger (1998) satisfies (3.45). It would suffice to show that they do not satisfy (3.45) when $\theta_2 = \theta_3 = 1$. From (3.45),

$\pi(\theta)D_2\left(\frac{\psi_2(\theta)}{\psi(\theta)}\right)+\frac{\psi_2(\theta)}{\psi(\theta)}\pi_2(\theta)+\pi(\theta)D_3\left(\frac{\psi_3(\theta)}{\psi(\theta)}\right)+\frac{\psi_3(\theta)}{\psi(\theta)}\pi_3(\theta)$, so that (3.45) holds if and only if

$$D_2\left(\frac{\psi_2(\theta)}{\psi(\theta)}\right)+D_3\left(\frac{\psi_3(\theta)}{\psi(\theta)}\right)+\frac{\psi_2(\theta)}{\psi(\theta)}[D_2\log\pi(\theta)]+\frac{\psi_3(\theta)}{\psi(\theta)}D_3[\log\pi(\theta)]=0 \tag{3.53}$$

At $\theta_2=\theta_3=1$,

$\frac{\psi_2(\theta)}{\psi(\theta)}=\frac{\frac{2}{t}+1}{\left[\frac{1}{2}\left(\frac{3}{t}+1\right)\right]^{\frac{1}{2}}}$ and $\frac{\psi_3(\theta)}{\psi(\theta)}=\frac{\frac{1}{t}}{\left[\frac{1}{2}\left(\frac{3}{t}+1\right)\right]^{\frac{1}{2}}}$

Furthermore, $D_2\left(\frac{\psi_2(\theta)}{\psi(\theta)}\right)$ is $\frac{\left(\frac{1}{t}+1\right)\left(\frac{2}{t}+\frac{1}{2}\right)}{\left[\frac{1}{2}\left(\frac{3}{t}+1\right)\right]^{\frac{3}{2}}}$ and $D_3\left(\frac{\psi_3(\theta)}{\psi(\theta)}\right)$ is $\frac{\frac{1}{t}\left(\frac{2}{t}+1\right)}{\left[\frac{1}{2}\left(\frac{3}{t}+1\right)\right]^{\frac{3}{2}}}$.Then, writing $[D_2\log\pi(\theta)]_1$ and $[D_3\log\pi(\theta)]_1$ respectively for the values of these quantities at $\theta_2=\theta_3=1$, it is clear that (3.53) holds at $\theta_2=\theta_3=1$ if and only if

$$\left(\tfrac{1}{t}+1\right)\left(\tfrac{2}{t}+\tfrac{1}{2}\right)+\tfrac{1}{t}\left(\tfrac{2}{t}+1\right)+\tfrac{1}{2}\left(\tfrac{3}{t}+1\right)\left\{\left(\tfrac{2}{t}+1\right)[D_2\log\pi(\theta)]_1+\tfrac{1}{t}[D_3\log\pi(\theta)]_1\right\}=0$$

Multiplying both sides by $2t^2$, we get

$$(t+1)(t+4)+2(t+2)+(t+3)\{(t+2)[D_2\log\pi(\theta)]_1+[D_3\log\pi(\theta)]_1\}=0$$

$$t^2+7t+8+(t+3)\{(t+2)[D_2\log\pi(\theta)]_1+[D_3\log\pi(\theta)]_1\}=0 \tag{3.54}$$

For the Jeffreys' prior, $\pi_J(\theta)=\frac{1}{\theta_3(\theta_3+t\theta_2)^{3/2}}$,

$$\log\pi_J(\theta)=-\log\theta_3-\frac{3}{2}\log(\theta_3+t\theta_2),$$

$$D_2\log\pi_J(\theta)=-\frac{3t}{2(\theta_3+t\theta_2)},\quad D_3\log\pi_J(\theta)=-\frac{1}{\theta_3}-\frac{3}{2(\theta_3+t\theta_2)}$$

$$[D_2\log\pi_J(\theta)]_1=-\frac{3t}{2(t+1)}\text{ and }[D_3\log\pi_J(\theta)]_1=-\frac{2t+5}{2(t+1)}$$

Hence the LHS of (3.54) becomes

$$t^2+7t+8-\frac{(t+3)(3t+5)}{2}=t^2+7t+8-\frac{3}{2}t^2-7t-\frac{15}{2}=-\frac{1}{2}(t^2-1)\neq 0,\ \forall t\geq 2$$

For the prior, $\pi_W(\theta)=\dfrac{1}{\theta_3(\theta_3+t\theta_2)}$ in (3.52),

$$\log\pi_W(\theta)=-\log\theta_3-\log(\theta_3+t\theta_2)$$

$$D_2\log\pi_W(\theta)=-\frac{t}{\theta_3+t\theta_2},\ D_3\log\pi_W(\theta)=-\frac{1}{\theta_3}-\frac{1}{\theta_3+t\theta_2}$$

$$[D_2\log\pi_W(\theta)]_1=-\frac{t}{t+1},\ [D_3\log\pi_W(\theta)]_1=-\frac{t+2}{t+1}$$

Hence the LHS of (3.54) becomes

$$t^2+7t+8-(t+3)(t+2)=2t+2\neq 0,\ \forall t\geq 2.$$

The priors $\pi_J(\theta)$ and $\pi_W(\theta)$ do not satisfy the probability matching conditions given in (3.45) and thus are not probability matching for tolerance intervals.

According to the numerical study by Ong and Mukerjee (2011), the prior $\pi_W(\theta)$ comes close to being probability matching for tolerance intervals. We explore the expressions for the frequentist probability $P_\theta(R\geq\beta)$ as shown in (3.44) under the prior $\pi_W(\theta)$. Applying (3.43), Ong and Mukerjee (2011) gave the expression under $\pi_W(\theta)$ for the term of order $n^{-1/2}$ on the right-hand side of (3.44) i.e.

$$V_1=\left(\frac{2}{nt}\right)^{1/2}\frac{(1+\delta)(1+t\delta)}{(1+2\delta+t\delta^2)^{3/2}}\phi(q), \tag{3.55}$$

where $\delta=\dfrac{\theta_2}{\theta_3}$.

Derivation of (3.55)

For the balanced one-way random effects model, the term (3.44) is of the form

$$\frac{1}{\sqrt{n}}\left[D_s\left(\frac{\psi_s(\theta)}{\psi(\theta)}\right)+\frac{\pi_s(\theta)}{\pi(\theta)}\frac{\psi_s(\theta)}{\psi(\theta)}\right]\phi(q),$$

the implicit sum on s being over $s=1,2,3$ for the one-way model.

$\psi_1(\theta)=0$, $D_1\left(\frac{\psi_1(\theta)}{\psi(\theta)}\right)=0$,

After some simplification,

$D_2\left(\frac{\psi_2(\theta)}{\psi(\theta)}\right)=\frac{(\theta_3+t\theta_2)(2\theta_3^2+2\theta_2\theta_3+t\theta_2^2)}{2t^2[\psi(\theta)]^3}$ and $D_3\left(\frac{\psi_3(\theta)}{\psi(\theta)}\right)=\frac{\theta_3(\theta_3^2+3\theta_2\theta_3+2t\theta_2^2)}{2t^2[\psi(\theta)]^3}$

$$D_2\left(\frac{\psi_2(\theta)}{\psi(\theta)}\right)+D_3\left(\frac{\psi_3(\theta)}{\psi(\theta)}\right)=\frac{3\theta_3^3+(5+2t)\theta_2\theta_3^2+5t\theta_2^2\theta_3+t^2\theta_2^3}{2t^2[\psi(\theta)]^3}$$

Also, $\frac{\pi_1(\theta)}{\pi(\theta)}\frac{\psi_1(\theta)}{\psi(\theta)}=0$. Noting that $\pi_W(\theta)=\frac{1}{\theta_3(\theta_3+t\theta_2)}$, we have

$\log\pi_W(\theta)=-\log\theta_3-\log(\theta_3+t\theta_2)$, i.e.

$\frac{\pi_2(\theta)}{\pi(\theta)}=-\frac{t}{\theta_3+t\theta_2}$ and $\frac{\pi_3(\theta)}{\pi(\theta)}=-\frac{1}{\theta_3}-\frac{1}{\theta_3+t\theta_2}=-\frac{2\theta_3+t\theta_2}{\theta_3(\theta_3+t\theta_2)}$. Therefore,

$$\frac{\pi_2(\theta)}{\pi(\theta)}\frac{\psi_2(\theta)}{\psi(\theta)}=-\frac{\theta_2(2\theta_3+t\theta_2)}{(\theta_3+t\theta_2)\psi(\theta)},$$

$$\frac{\pi_3(\theta)}{\pi(\theta)}\frac{\psi_3(\theta)}{\psi(\theta)}=-\frac{\theta_3(2\theta_3+t\theta_2)}{t(\theta_3+t\theta_2)\psi(\theta)},$$

$$\frac{\pi_2(\theta)}{\pi(\theta)}\frac{\psi_2(\theta)}{\psi(\theta)}+\frac{\pi_3(\theta)}{\pi(\theta)}\frac{\psi_3(\theta)}{\psi(\theta)}=-\frac{2\theta_3+t\theta_2}{t\psi(\theta)}.$$

As a result,

$$D_s\left(\frac{\psi_s(\theta)}{\psi(\theta)}\right)+\frac{\pi_s(\theta)}{\pi(\theta)}\frac{\psi_s(\theta)}{\psi(\theta)}=\frac{3\theta_3^3+(5+2t)\theta_2\theta_3^2+5t\theta_2^2\theta_3+t^2\theta_2^3}{2t^2[\psi(\theta)]^3}-\frac{2\theta_3+t\theta_2}{t\psi(\theta)}$$

$$=\frac{[\theta_3^3+(1+t)\theta_2\theta_3^2+t\theta_2^2\theta_3]}{2t^2[\psi(\theta)]^3}$$

$$=\frac{[\theta_3^3+(1+t)\theta_2\theta_3^2+t\theta_2^2\theta_3]}{2t^2\left[\frac{1}{2}\left(\frac{\theta_3^2}{t}+\frac{2}{t}\theta_2\theta_3+\theta_2^2\right)\right]^{3/2}}$$

$$=\frac{(2t)^{3/2}}{2t^2}\frac{[\theta_3^3+(1+t)\theta_2\theta_3^2+t\theta_2^2\theta_3]}{(\theta_3^2+2\theta_2\theta_3+t\theta_2^2)^{3/2}}$$

$$=\sqrt{\frac{2}{t}}\frac{[1+(1+t)\delta+t\delta^2]}{(1+2\delta+t\delta^2)^{3/2}} \qquad \text{(Here } \delta=\frac{\theta_2}{\theta_3}\text{)}$$

$$=\sqrt{\frac{2}{t}}\frac{(1+\delta)(1+t\delta)}{(1+2\delta+t\delta^2)^{3/2}},$$

i.e. the term of order $n^{-1/2}$ in the frequentist coverage is

$V_1=\left(\frac{2}{nt}\right)^{1/2}\frac{(1+\delta)(1+t\delta)}{(1+2\delta+t\delta^2)^{3/2}}\phi(q)$, as given in (3.55).

The quantity V_1 is of order $O(t^{-1})$ for any fixed n, δ and q. Even for small values of t it turns out to be small for moderately large values of *n.* To illustrate this point, let $\gamma=0.95$ i.e. $q=1.6449$. Ong and Mukerjee (2011) presented a table of values (see Table 3.1) for fixed values of t with smallest n denoted by $n_0(t)$ such that $V_1\le 0.02$ for every δ where $\delta=\frac{\theta_2}{\theta_3}$ in the set $\Delta=\{0.5, 1, 1.5, 2, 2.5, 3\}$. Table 3.1 gives the values of $n_0(t)$ for $2\le t\le 10$ in order to achieve $V_1\le 0.02$ for every δ in Δ. Based on this table, if $t=2$, then $V_1\le 0.02$ for every δ in Δ when $n\ge 16$. Similarly, the same occurs if $t=3$, $n\ge 12$ and so on.

Table 3.1: Values of $n_0(t)$ corresponding to t.

t	2	3	4	5	6	7	8	9	10
$n_0(t)$	16	12	10	9	8	7	6	6	5

The numerical study by Ong and Mukerjee (2011) revealed that although the prior $\pi_W(\theta)$ is not probability matching for tolerance intervals, it comes close to being so. If one is not too particular about the probability matching criteria, the simplicity of this prior makes a strong case in their favour. Ong and Mukerjee (2011) suggested the use of the relatively more complex matching prior shown in (3.46) if the matching property is compulsory. They also remarked that the results for the tolerance intervals in (3.20) are heavily dependent on balance in classes. We can see that t appears in the matching prior given in (3.46) for the balanced one-way random effects model. Denoting the number of observations in n classes by $t_1, ..., t_n$, some version of the present higher order asymptotics should go through provided the lower order moments of $t_1, ..., t_n$ remains bounded as n tends to infinity.

CHAPTER 4

TOLERANCE INTERVALS IN BALANCED ONE-WAY RANDOM EFFECTS MODEL WITH NON-NORMAL ERRORS: A COMPARATIVE STUDY

4.1 Introduction

In Chapter 3, we discussed the two-sided Bayesian tolerance intervals with approximate frequentist validity, via the use of probability matching priors (PMP). As mentioned in Chapter 3, these intervals, constructed from higher order asymptotic considerations, have $\beta-$content with posterior credibility level $\gamma + O_p(n^{-1})$ and have a frequentist confidence level $\gamma + O(n^{-1})$, where n is the number of classes. Thus, the method by Ong and Mukerjee (2011) depends heavily on the balance in the classes and is meaningful when the number of classes is large.

Krishnamoorthy and Mathew (2004) introduced the modified large sample (MLS) approach in constructing two-sided tolerance intervals for balanced one-way random effects model based on the procedure by Graybill and Wang (1980) for finding an upper confidence limit for a linear combination of variance components. The MLS tolerance intervals are in closed-form and this makes them easy to be computed. The merits of this tolerance interval were evaluated by Krishnamoorthy and Lian (2012) based upon the expected widths as well as coverage probabilities.

In this chapter, we evaluate and compare the performance of the PMP and MLS tolerance intervals, in particular when the errors are non-normal. Non-normal error distributions are represented by the t-distribution, skew-normal distribution with various shape parameters (Azzalini, 1985) and the generalized lambda (GLD) distribution (Karian and Dudewicz, 2000). The t-distribution and skew-normal distribution have heavier tails than the normal case. The GLD family is considered because of its versatility to produce distributions with wide range of shapes and skewness. The effects of non-normal experimental errors are studied by considering the expected widths as well as their standard errors and coverage probabilities.

4.2 Tolerance intervals for balanced one-way random effects model

We recall the balanced one-way random effects given in Equation (3.1) of Chapter 3 i.e.

$$Y_{ij} = \theta_1 + v_i + e_{ij}$$

for $i = 1, 2, ..., n$, where n represents the number of classes and $j = 1, 2, ..., t$, where t represents the number of observations per class. Here Y_{ij} denotes the ij^{th} observation and θ_1 is the population mean. v_i and e_{ij} are independent with $v_i \sim N(0, \theta_2)$ and $e_{ij} \sim N(0, \theta_3)$. As mentioned in Equation (3.2) and Equation (3.3) of Chapter 3, ρ, the intra-class correlation coefficient, given by $\rho = \dfrac{\theta_2}{\theta_2 + \theta_3}$ and the relationship between θ_2 and θ_3 is $\theta_2 = \dfrac{\rho}{1-\rho}\theta_3$.

Using (3.4) from Chapter 3, the maximum likelihood estimator (MLE) of $\theta = (\theta_1, \theta_2, \theta_3)$ is given by $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3)$ where

$$\hat{\theta}_1 = \overline{Y}, \hat{\theta}_2 = (\mathrm{MSB} - \mathrm{MSW})/t, \hat{\theta}_3 = \mathrm{MSW}.$$

In the above, $\overline{Y}$ is the grand mean of the Y_{ij}'s, while MSW and MSB in (3.5) are the usual mean squares within and between classes, that is,

$$\mathrm{MSW} = \frac{1}{n(t-1)} \sum_{i=1}^{n} \sum_{j=1}^{t} \left(Y_{ij} - \overline{Y}_i\right)^2 \quad \text{and} \quad \mathrm{MSB} = \frac{t}{n} \sum_{i=1}^{n} \left(\overline{Y}_i - \overline{Y}\right)^2.$$

4.2.1 Bayesian tolerance interval with approximate frequentist validity

Summarizing the results in Chapter 3, under the balanced one-way random effects model, each $Y_{ij} \sim N(\theta_1, \theta_2 + \theta_3)$. From Equation (3.20) of Chapter 3, the Bayesian tolerance interval under a prior $\pi(.)$, having the same $N(\theta_1, \theta_2 + \theta_3)$ which has $\beta-$content with posterior credibility level $\gamma + O_p(n^{-1})$, is given by:

$$\left[\hat{\theta}_1 - \left(b + \frac{g_1}{\sqrt{n}} + \frac{g_2}{n}\right)z,\ \hat{\theta}_1 + \left(b + \frac{g_1}{\sqrt{n}} + \frac{g_2}{n}\right)z\right] \tag{4.1}$$

By choosing $d = 0$, the tolerance interval in (3.20) becomes (4.1) where

$b = \sqrt{\hat{\theta}_2 + \hat{\theta}_3}$, $z = \Phi^{-1}\left(\frac{1}{2}(1+\beta)\right)$ (Φ is a standard normal distribution function),

$g_1 = kq$ and $g_2 = k\left[L_{11} - L_3 + \left(\frac{1}{2}kb^{-1}z^2 + L_{21} + L_3\right)q^2\right]$

$\varepsilon = (0,1,1)'$, $\quad k = \frac{1}{2}b^{-1}(\varepsilon' C^{-1}\varepsilon)^{1/2}$, $\quad \lambda = (\lambda_1, \lambda_2, \lambda_3)' = \frac{1}{2}(kb)^{-1}C^{-1}\varepsilon$

$$L_{11} = \frac{\hat{\pi}_s}{\hat{\pi}}\lambda_s + \frac{1}{2}a_{suw}\lambda_s c^{uw} + \frac{1}{2}kb^{-1}(z^2-3) + \frac{1}{2}(kb)^{-1}c^{11}$$

$$L_{21} = -kb^{-1}(z^2-1)$$

$$L_3 = \frac{1}{6}a_{suw}\lambda_s\lambda_u\lambda_w + \frac{1}{2}kb^{-1}(z^2-3)$$

$$q = \Phi^{-1}(\gamma)$$

$C^{-1} = (c^{su})$ and a_{suw} are respectively given in (3.17) and (3.18) of Chapter 3. Summation convention is followed with implicit sums over repeated sub- or superscripts ranging over {1, 2, 3}.

The interval in (4.1) has approximate frequentist validity, i.e., it has $\beta-$content with frequentist confidence level $\gamma + O(n^{-1})$, when $\pi(\cdot)$ is taken as a probability matching prior (PMP). Based on Ong and Mukerjee (2011) and as shown in Chapter 3, such a prior is given by

$$\pi(\theta) = \frac{\psi(\theta)}{\{\theta_3(\theta_3 + t\theta_2)\}^2}$$

where

$$\psi(\theta) = \{\tfrac{1}{2}t^{-1}(t\theta_2^2 + 2\theta_2\theta_3 + \theta_3^2)\}^{1/2}$$

In our comparisons, we will consider the interval (4.1) based on the aforesaid PMP.

4.2.2 Modified large sample tolerance intervals

Following Krishnamoorthy and Lian (2012), the modified large sample (MLS) tolerance intervals constructed are functions of

$$\overline{Y} = \frac{1}{nt}\sum_{i=1}^{n}\sum_{j=1}^{t}Y_{ij}\,,\ \overline{Y}_i = \frac{1}{t}\sum_{j=1}^{t}Y_{ij}\,,\ \text{SSB} = t\sum_{i=1}^{n}\left(\overline{Y}_i - \overline{Y}\right)^2 \text{ and SSW} = \sum_{i=1}^{n}\sum_{j=1}^{t}\left(Y_{ij} - \overline{Y}_i\right)^2. \qquad (4.2)$$

It is also noted that $\bar{Y}$, SSB and SSW are independent with

$$Z = \sqrt{nt}\,\frac{(\bar{Y}-\theta_1)}{\sqrt{t\theta_2+\theta_3}} \sim N(0,1), \quad U_B^2 = \frac{SSB}{t\theta_2+\theta_3} \sim \chi_{n-1} \text{ and } U_W^2 = \frac{SSW}{\theta_3} \sim \chi_{n(t-1)}. \tag{4.3}$$

To construct the (β,γ) tolerance intervals for a $N(\theta_1, \theta_2+\theta_3)$ distribution, let $\sigma_1^2 = t\theta_2+\theta_3$ and $\sigma_2^2 = \theta_3$, so that $\theta_2 = \frac{(\sigma_1^2-\sigma_2^2)}{t}$. Let $\hat{\sigma}_1^2 = \frac{t}{n-1}\sum_{i=1}^{n}\left(\bar{Y}_i-\bar{Y}\right)^2 = \frac{SSB}{n-1}$ and from (3.5), $\hat{\sigma}_2^2 = MSW$.

$$Y_{ij} \sim N\left(\theta_1, \frac{\sigma_1^2}{t}+\left(1-\frac{1}{t}\right)\sigma_2^2\right), \; \bar{Y} \sim N\left(\theta_1, \frac{\sigma_1^2}{nt}\right), \; \frac{\hat{\sigma}_1^2}{\sigma_1^2} \sim \frac{\chi_{n-1}^2}{n-1} \text{ and } \frac{\hat{\sigma}_2^2}{\sigma_2^2} \sim \frac{\chi_{n(t-1)}^2}{n(t-1)} \tag{4.4}$$

$\bar{Y}$, $\hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$ are mutually independent.

Note that $r_1 = 1/t$, $r_2 = 1-1/t$, $s_1 = 1/(nt)$ and $s_2 = 0$.

The construction of the (β,γ) tolerance interval simplifies to the construction of an γ upper confidence limit for $n_1 = r_1+s_1 = (1+1/n)/t$ and $n_2 = r_2+s_2 = (1-1/t)$.

The MLS tolerance interval is given by:

$$\bar{Y} \pm z_{(1+\beta)/2}\sqrt{U_{mls;\gamma}^{(1)}}\;, \tag{4.5}$$

where the MLS upper confidence limit for $n_1\sigma_1^2+n_2\sigma_2^2$ is given by

$$U_{mls;\gamma}^{(1)} = \left(n_1\hat{\sigma}_1^2+n_2\hat{\sigma}_2^2\right)+\left(n_1^2\hat{\sigma}_1^4\left(\frac{n-1}{\chi_{n-1;1-\gamma}^2}-1\right)^2+n_2^2\hat{\sigma}_2^4\left(\frac{n(t-1)}{\chi_{n(t-1);1-\gamma}^2}-1\right)^2\right)^{1/2}. \tag{4.6}$$

4.3 Monte Carlo simulation study and discussion

In this section, we conduct a Monte Carlo simulation study to compare the performance of the two-sided Bayesian PMP tolerance interval by Ong and Mukerjee (2011) and the MLS tolerance interval by Krishnamoorthy and Mathew (2009) for the balanced one-way random effects model with the experimental error following the standard normal distribution and non-normal distributions such as the t-distribution, skew-normal distribution and generalized lambda distribution (GLD). We note in passing that if the error e_{ij} follows the Student's t-distribution, then Y_{ij} in (3.1) is the sum of a normal (v_i) and a Student's t (e_{ij}) random variables, and an explicit expression for the probability density function (pdf) is given by Nason (2006). By fixing the value of ρ, the relationship between the variance of v_i, θ_2 and the variance of e_{ij}, θ_3 is given by $\theta_2 = \frac{\rho}{1-\rho}\theta_3$; see (3.3) of Chapter 3.

The PMP and MLS tolerance intervals were used for all cases as if the assumptions where all underlying distributions are normal are justified even though the data comes from another distribution. Our purpose is to see the effect on the expected width as well as the coverage probability when the distribution generating the data deviates from the normal.

Since the small sample behavior of the Bayesian tolerance intervals using probability matching priors (PMP) was never studied, it is of interest to examine its performance. As mentioned earlier, this approach depends heavily on balance in the classes and is meaningful when the number of classes *n* is large. On the other hand, if the number of observations *t* per class is large but *n* is small, then this approach is not expected to behave well because it draws its strength from the consistency of $\hat{\theta}$, which holds as $n \to \infty$. From this perspective, various combinations of (*n, t*), were considered in the simulation and a comparative study was done between the PMP and MLS tolerance intervals. The cases of non-normal error distributions were of particular interest in order to see the behavior of both tolerance intervals when there is a departure from normality in the data.

The following distributions are used to represent the non-normal experimental errors in our study:

<u>Experimental error following the t-distribution</u>

The experimental error is taken to follow the t-distribution, with mean 0 and variance $\frac{\nu}{\nu-2}, \nu > 2$ where ν is the degrees of freedom. It is known that the t-distribution approaches the standard normal distribution when ν increases.

The pdf of the t-distribution is:

$$f(x;\nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}. \tag{4.7}$$

<u>Experimental error following the skew-normal distribution</u>

The pdf for the skew-normal distribution (Azzalini, 1985) is

$$\frac{1}{\omega\pi} e^{-\frac{(x-\xi)^2}{2\omega^2}} \int_{-\infty}^{\alpha\left(\frac{x-\xi}{\omega}\right)} e^{-\frac{t^2}{2}} dt \tag{4.8}$$

where ξ, ω and α are the location, scale and shape parameters respectively.

The data generated has e_{ij} following a skew normal distribution with $\xi = 0$, $\omega = 1$ and shape= α. The error distribution reduces to a standard normal when $\alpha = 0$.

The mean of the skew-normal distribution is given by

$$\xi + \omega\delta\sqrt{\frac{2}{\pi}} \text{ where } \delta = \frac{\alpha}{\sqrt{1+\alpha^2}} \tag{4.9}$$

The mean of the skew-normal distribution is no longer 0 when $\alpha \neq 0$.

Experimental error following the generalized lambda distribution

The generalized lambda distribution (GLD) family with parameters λ_1, λ_2, λ_3 and λ_4, denoted as GLD($\lambda_1, \lambda_2, \lambda_3, \lambda_4$), is most easily specified in terms of its percentile function (Karian and Dudewicz, 2000). The following percentile function uses the Ramberg and Schmeiser's parameterization

$$Q(y) = Q(y; \lambda_1, \lambda_2, \lambda_3, \lambda_4) = \lambda_1 + \frac{y^{\lambda_3} - (1-y)^{\lambda_4}}{\lambda_2} \tag{4.10}$$

where $0 \le y \le 1$. λ_1 and λ_2 are respectively the location and scale parameters while λ_3 and λ_4 jointly determine the shape (with λ_3 mostly affecting the left tail and λ_4 the right tail).

The pdf of the GLD is given by

$$\frac{\lambda_2}{\lambda_3 y^{\lambda_3 - 1} + \lambda_4 (1-y)^{\lambda_4 - 1}} \text{ at } x = Q(y) \tag{4.11}$$

We shall utilize the parameters estimated in Karian and Dudewicz (2000) to generate our data whose error distribution follows the GLD. We consider parameters fitted using the method of moments approach which approximately fit the standard normal distribution. The parameter estimates are given as GLD (0, 0.1975, 0.1349, 0.1349).

The two-sided PMP and MLS tolerance intervals were constructed for $\beta = 0.90$ and $\gamma = 0.95$ for data from both normal and non-normal experimental error distributions. For each simulated interval, the content was calculated as $\Phi(U) - \Phi(L)$ where U and L respectively represent the upper and lower bounds of the tolerance intervals. We found that 2500 simulation runs were sufficient for our study and hence this number of runs was used for various combinations of *(n, t)* and ρ, the intra-class correlation coefficient. The coverage probability or the proportion of times the content of the simulated intervals was at least β was computed. The coverage probability depends on parameters estimated via ρ. We will not vary the mean, θ_1 in the balanced one-way random effects model as it has no impact on the interval. The tables in Section 4.3.1 show the coverage probabilities and expected widths with their respective standard errors (bracketed) for the PMP and MLS tolerance intervals applying various error distributions.

4.3.1 Simulation results

Table 4.1: *Simulated coverage probabilities of the PMP and MLS tolerance intervals. Experimental error: standard normal distribution.*

		(*n*,*t*)									
ρ		(15,2)	(25,2)	(40,2)	(50,2)	(75,2)	(45,3)	(60,3)	(60,4)	(80,4)	(75,5)
0.100	PMP	0.949	0.950	0.958	0.951	0.963	0.962	0.970	0.972	0.977	0.974
	MLS	0.973	0.971	0.975	0.974	0.978	0.980	0.975	0.976	0.978	0.981
0.300	PMP	0.938	0.947	0.959	0.957	0.966	0.966	0.963	0.966	0.972	0.973
	MLS	0.970	0.975	0.974	0.963	0.976	0.977	0.976	0.970	0.976	0.976
0.500	PMP	0.932	0.948	0.946	0.954	0.966	0.952	0.958	0.965	0.966	0.967
	MLS	0.968	0.964	0.968	0.971	0.967	0.966	0.967	0.972	0.971	0.973
0.700	PMP	0.912	0.940	0.960	0.947	0.964	0.951	0.964	0.957	0.964	0.968
	MLS	0.964	0.964	0.964	0.964	0.969	0.968	0.966	0.962	0.968	0.965
0.900	PMP	0.924	0.920	0.948	0.954	0.956	0.952	0.955	0.958	0.957	0.962
	MLS	0.950	0.955	0.966	0.963	0.968	0.963	0.962	0.964	0.963	0.964
0.990	PMP	0.924	0.945	0.947	0.950	0.963	0.948	0.954	0.958	0.952	0.958
	MLS	0.958	0.959	0.963	0.958	0.965	0.959	0.963	0.965	0.963	0.960
0.999	PMP	0.924	0.946	0.952	0.953	0.960	0.952	0.951	0.949	0.954	0.954
	MLS	0.958	0.953	0.958	0.960	0.966	0.953	0.954	0.957	0.962	0.964

Table 4.2: *Expected widths of the PMP and MLS tolerance intervals. Experimental error: standard normal distribution.*

		(n,t)									
ρ		(15,2)	(25,2)	(40,2)	(50,2)	(75,2)	(45,3)	(60,3)	(60,4)	(80,4)	(75,5)
0.100	PMP	4.423	4.143	3.997	3.919	3.829	3.857	3.811	3.758	3.721	3.697
		(0.584)	(0.427)	(0.324)	(0.289)	(0.223)	(0.243)	(0.208)	(0.176)	(0.151)	(0.142)
	MLS	4.686	4.305	4.068	4.002	3.881	3.909	3.838	3.787	3.730	3.716
		(0.644)	(0.455)	(0.333)	(0.285)	(0.229)	(0.243)	(0.207)	(0.181)	(0.152)	(0.138)
0.300	PMP	5.062	4.759	4.565	4.488	4.363	4.435	4.361	4.311	4.260	4.248
		(0.716)	(0.531)	(0.395)	(0.349)	(0.264)	(0.306)	(0.264)	(0.240)	(0.199)	(0.193)
	MLS	5.414	4.931	4.649	4.545	4.412	4.489	4.394	4.343	4.278	4.266
		(0.786)	(0.542)	(0.389)	(0.350)	(0.272)	(0.309)	(0.263)	(0.244)	(0.203)	(0.194)
0.500	PMP	6.088	5.749	5.449	5.359	5.233	5.326	5.222	5.202	5.114	5.115
		(0.969)	(0.690)	(0.520)	(0.447)	(0.348)	(0.429)	(0.354)	(0.344)	(0.286)	(0.280)
	MLS	6.499	5.888	5.562	5.436	5.248	5.412	5.267	5.245	5.142	5.142
		(1.028)	(0.716)	(0.522)	(0.446)	(0.353)	(0.444)	(0.363)	(0.349)	(0.296)	(0.289)
0.700	PMP	8.109	7.530	7.168	7.007	6.821	7.017	6.876	6.840	6.735	6.731
		(1.441)	(1.005)	(0.723)	(0.638)	(0.491)	(0.636)	(0.531)	(0.524)	(0.445)	(0.456)
	MLS	8.573	7.765	7.252	7.097	6.850	7.106	6.926	6.877	6.769	6.758
		(1.447)	(0.995)	(0.729)	(0.626)	(0.497)	(0.630)	(0.550)	(0.531)	(0.451)	(0.458)
0.900	PMP	14.339	13.198	12.533	12.360	11.947	12.439	12.122	12.126	11.837	11.913
		(2.679)	(1.903)	(1.360)	(1.209)	(0.959)	(1.267)	(1.080)	(1.057)	(0.882)	(0.930)
	MLS	15.228	13.678	12.814	12.504	12.030	12.587	12.234	12.228	11.899	11.966
		(2.873)	(1.933)	(1.400)	(1.209)	(0.929)	(1.263)	(1.094)	(1.052)	(0.899)	(0.923)
0.990	PMP	45.776	42.596	40.212	39.220	37.928	39.647	38.620	38.772	37.795	38.006
		(8.700)	(6.076)	(4.549)	(3.975)	(3.091)	(4.228)	(3.605)	(3.556)	(3.003)	(3.105)
	MLS	48.573	43.674	40.847	39.745	38.225	40.105	38.954	38.898	37.964	38.182
		(9.173)	(6.303)	(4.542)	(4.063)	(3.135)	(4.364)	(3.611)	(3.597)	(3.008)	(3.182)
0.999	PMP	143.763	134.439	127.150	124.289	120.190	125.290	122.287	122.451	119.616	119.924
		(27.904)	(19.097)	(14.378)	(12.726)	(9.828)	(13.237)	(11.524)	(11.780)	(9.731)	(9.766)
	MLS	154.145	137.831	128.586	125.985	121.127	126.710	122.970	123.539	120.140	120.745
		(29.291)	(20.420)	(14.502)	(12.773)	(9.972)	(13.437)	(11.606)	(11.630)	(9.703)	(9.962)

Table 4.3: *Simulated coverage probabilities of the PMP and MLS tolerance intervals. Experimental error: t-distribution.*

(n,t)	ρ					
	0.900		0.990		0.999	
	PMP	MLS	PMP	MLS	PMP	MLS
degrees of freedom=3						
(15,2)	0.957	0.976	0.929	0.954	0.925	0.952
(25,2)	0.978	0.984	0.946	0.960	0.950	0.953
(40,2)	0.987	0.988	0.955	0.968	0.947	0.952
(45,3)	0.992	0.994	0.954	0.962	0.955	0.960
(60,3)	0.993	0.996	0.964	0.966	0.956	0.960
(60,4)	0.994	0.993	0.967	0.966	0.947	0.958
(80,4)	0.996	0.998	0.968	0.971	0.962	0.962
degrees of freedom=5						
(15,2)	0.941	0.970	0.926	0.952	0.920	0.952
(25,2)	0.958	0.973	0.946	0.952	0.938	0.962
(40,2)	0.965	0.978	0.948	0.965	0.945	0.953
(45,3)	0.975	0.976	0.951	0.970	0.959	0.955
(60,3)	0.981	0.981	0.948	0.962	0.957	0.960
(60,4)	0.981	0.979	0.958	0.965	0.955	0.953
(80,4)	0.986	0.989	0.964	0.971	0.958	0.962
degrees of freedom=10						
(15,2)	0.930	0.956	0.926	0.956	0.930	0.953
(25,2)	0.949	0.959	0.936	0.958	0.943	0.953
(40,2)	0.961	0.968	0.954	0.959	0.956	0.954
(45,3)	0.962	0.964	0.955	0.962	0.959	0.958
(60,3)	0.962	0.973	0.958	0.964	0.951	0.962
(60,4)	0.971	0.973	0.953	0.964	0.954	0.965
(80,4)	0.972	0.972	0.960	0.966	0.962	0.965
degrees of freedom=15						
(15,2)	0.934	0.959	0.923	0.953	0.926	0.950
(25,2)	0.940	0.963	0.936	0.959	0.945	0.961
(40,2)	0.955	0.961	0.956	0.957	0.945	0.960
(45,3)	0.950	0.967	0.946	0.958	0.956	0.954
(60,3)	0.963	0.962	0.951	0.962	0.956	0.960
(60,4)	0.961	0.968	0.960	0.955	0.948	0.963
(80,4)	0.963	0.966	0.957	0.964	0.959	0.969
degrees of freedom=25						
(15,2)	0.918	0.953	0.926	0.958	0.921	0.950
(25,2)	0.944	0.957	0.952	0.956	0.942	0.962
(40,2)	0.951	0.962	0.957	0.959	0.950	0.955
(45,3)	0.958	0.965	0.953	0.960	0.953	0.958
(60,3)	0.960	0.961	0.957	0.964	0.954	0.966
(60,4)	0.958	0.959	0.962	0.962	0.958	0.963
(80,4)	0.965	0.968	0.962	0.963	0.955	0.968

Table 4.4: *Expected widths of the PMP and MLS tolerance intervals. Experimental error: t-distribution.*

(n,t)	ρ					
	0.900		0.990		0.999	
	PMP	MLS	PMP	MLS	PMP	MLS
degrees of freedom=3						
(15,2)	15.247 (2.964)	16.220 (3.134)	45.737 (8.504)	48.971 (9.357)	145.031 (27.222)	153.337 (29.473)
(25,2)	14.316 (2.453)	14.730 (2.328)	42.744 (6.088)	43.935 (6.335)	134.556 (19.298)	138.342 (20.112)
(40,2)	13.612 (1.886)	13.780 (1.740)	40.308 (4.524)	41.167 (4.611)	127.419 (14.424)	128.946 (15.012)
(45,3)	13.384 (1.497)	13.584 (2.273)	39.876 (4.265)	40.494 (4.381)	125.807 (13.322)	126.879 (13.603)
(60,3)	13.089 (1.370)	13.198 (1.461)	38.949 (3.514)	39.243 (3.574)	122.520 (11.440)	123.205 (11.324)
(60,4)	13.009 (1.366)	13.096 (1.267)	39.135 (3.881)	39.127 (3.565)	122.460 (11.270)	122.932 (11.396)
(80,4)	12.812 (1.226)	12.832 (1.049)	38.129 (3.011)	38.220 (3.080)	120.102 (9.587)	120.061 (9.761)
degrees of freedom=5						
(15,2)	14.691 (2.696)	15.595 (2.782)	45.780 (8.730)	48.409 (9.170)	143.845 (27.686)	154.586 (29.299)
(25,2)	13.615 (1.895)	14.083 (1.943)	42.741 (6.248)	43.761 (6.462)	133.181 (19.606)	138.870 (19.915)
(40,2)	12.970 (1.418)	13.190 (1.406)	40.263 (4.631)	40.826 (4.532)	126.741 (14.415)	128.958 (14.952)
(45,3)	12.807 (1.270)	12.887 (1.290)	39.703 (4.260)	40.418 (4.227)	125.810 (13.049)	127.317 (13.955)
(60,3)	12.443 (1.064)	12.518 (1.054)	38.817 (3.635)	38.948 (3.616)	122.447 (11.066)	123.216 (11.331)
(60,4)	12.443 (1.067)	12.526 (1.066)	38.872 (3.564)	38.980 (3.527)	122.452 (11.294)	123.036 (11.436)
(80,4)	12.158 (0.895)	12.223 (0.890)	37.882 (2.931)	38.074 (3.007)	119.565 (9.634)	120.397 (9.533)
degrees of freedom=10						
(15,2)	14.402 (2.671)	15.307 (2.833)	45.578 (8.531)	48.868 (9.122)	144.004 (27.065)	154.151 (29.196)
(25,2)	13.477 (1.901)	13.811 (1.949)	42.159 (6.077)	43.713 (6.234)	134.008 (18.967)	138.231 (20.286)
(40,2)	12.711 (1.392)	12.915 (1.420)	40.170 (4.449)	40.741 (4.514)	127.442 (14.426)	128.933 (15.156)
(45,3)	12.544 (1.239)	12.731 (1.308)	39.770 (4.235)	38.988 (3.556)	125.847 (13.147)	127.049 (13.650)
(60,3)	12.239 (1.087)	12.365 (1.076)	38.764 (3.592)	39.083 (3.603)	121.814 (11.442)	123.147 (11.398)
(60,4)	12.241 (1.068)	12.350 (1.057)	38.635 (3.562)	39.034 (3.504)	122.245 (11.266)	123.355 (11.380)
(80,4)	11.982 (0.888)	12.052 (0.924)	37.708 (2.997)	38.001 (2.953)	119.772 (9.520)	120.466 (9.665)

(Table 4.4 continued)

degrees of freedom=15						
(15,2)	14.385 (2.626)	15.263 (2.884)	45.614 (8.596)	48.628 (9.285)	144.412 (27.135)	153.590 (29.337)
(25,2)	13.341 (1.863)	13.785 (1.934)	42.354 (6.097)	43.686 (6.289)	134.329 (19.441)	137.961 (19.585)
(40,2)	12.670 (1.365)	12.865 (1.436)	40.251 (4.471)	40.707 (4.577)	127.329 (14.514)	128.735 (14.508)
(45,3)	12.465 (1.293)	12.669 (1.285)	39.727 (4.282)	40.115 (4.299)	125.375 (13.440)	126.775 (13.621)
(60,3)	12.182 (1.040)	12.292 (1.095)	38.692 (3.562)	39.042 (3.555)	122.401 (11.373)	123.089 (11.249)
(60,4)	12.173 (1.080)	12.300 (1.105)	38.718 (3.510)	38.835 (3.583)	122.146 (11.403)	123.564 (11.475)
(80,4)	11.907 (0.880)	11.988 (0.904)	37.819 (2.959)	38.057 (3.019)	119.684 (9.622)	120.736 (9.528)
degrees of freedom=25						
(15,2)	14.282 (2.701)	15.185 (2.869)	45.544 (8.819)	48.789 (9.324)	144.915 (28.321)	153.884 (28.911)
(25,2)	13.336 (1.863)	13.762 (1.907)	42.599 (6.061)	43.805 (6.270)	133.964 (19.020)	138.966 (19.987)
(40,2)	12.667 (1.369)	12.875 (1.427)	40.228 (4.465)	40.669 (4.610)	127.017 (14.318)	128.531 (14.725)
(45,3)	12.502 (1.268)	12.587 (1.297)	39.710 (4.137)	40.274 (4.315)	125.359 (13.013)	127.103 (13.778)
(60,3)	12.167 (1.050)	12.251 (1.071)	38.611 (3.526)	38.857 (3.510)	122.074 (11.140)	123.494 (11.449)
(60,4)	12.143 (1.054)	12.224 (1.080)	38.746 (3.461)	39.036 (3.595)	119.915 (9.722)	123.238 (11.425)
(80,4)	11.902 (0.872)	11.972 (0.914)	37.850 (2.923)	38.009 (3.002)	119.748 (9.681)	120.145 (9.559)

Table 4.5: *Simulated coverage probabilities of the PMP and MLS tolerance intervals. Experimental error: skew-normal distribution.*

(n,t)	ρ					
	0.900		0.990		0.999	
	PMP	MLS	PMP	MLS	PMP	MLS
shape parameter=0.4						
(15,2)	0.913	0.947	0.916	0.947	0.932	0.947
(25,2)	0.934	0.946	0.946	0.959	0.938	0.959
(40,2)	0.940	0.962	0.952	0.966	0.942	0.966
(45,3)	0.944	0.952	0.952	0.957	0.957	0.960
(60,3)	0.949	0.951	0.953	0.959	0.955	0.964
(60,4)	0.949	0.952	0.956	0.966	0.952	0.963
(80,4)	0.951	0.955	0.964	0.968	0.960	0.965
shape parameter=1.0						
(15,2)	0.888	0.936	0.913	0.955	0.920	0.956
(25,2)	0.915	0.929	0.940	0.952	0.939	0.955
(40,2)	0.916	0.925	0.945	0.956	0.948	0.960
(45,3)	0.922	0.928	0.947	0.958	0.943	0.959
(60,3)	0.916	0.919	0.952	0.960	0.958	0.962
(60,4)	0.909	0.922	0.950	0.959	0.962	0.950
(80,4)	0.910	0.918	0.954	0.960	0.956	0.957
shape parameter=2.0						
(15,2)	0.884	0.927	0.923	0.953	0.921	0.954
(25,2)	0.895	0.921	0.934	0.955	0.932	0.947
(40,2)	0.887	0.895	0.944	0.954	0.948	0.958
(45,3)	0.886	0.912	0.943	0.956	0.949	0.959
(60,3)	0.883	0.898	0.953	0.958	0.962	0.960
(60,4)	0.880	0.902	0.956	0.954	0.949	0.966
(80,4)	0.873	0.888	0.953	0.960	0.959	0.962
shape parameter=5.0						
(15,2)	0.869	0.927	0.913	0.956	0.926	0.953
(25,2)	0.884	0.921	0.940	0.948	0.948	0.960
(40,2)	0.874	0.895	0.948	0.958	0.951	0.956
(45,3)	0.864	0.912	0.943	0.955	0.946	0.957
(60,3)	0.856	0.898	0.946	0.962	0.953	0.955
(60,4)	0.855	0.902	0.947	0.958	0.951	0.969
(80,4)	0.826	0.888	0.949	0.953	0.957	0.958
shape parameter=10.0						
(15,2)	0.879	0.920	0.922	0.953	0.923	0.959
(25,2)	0.882	0.910	0.948	0.950	0.940	0.954
(40,2)	0.872	0.910	0.943	0.953	0.946	0.959
(45,3)	0.856	0.883	0.949	0.954	0.952	0.954
(60,3)	0.853	0.874	0.950	0.957	0.956	0.962
(60,4)	0.850	0.884	0.946	0.964	0.954	0.956
(80,4)	0.841	0.858	0.948	0.959	0.956	0.963

Table 4.6: *Expected widths of the PMP and MLS tolerance intervals. Experimental error: skew-normal distribution.*

(n,t)	ρ					
	0.900		0.990		0.999	
	PMP	MLS	PMP	MLS	PMP	MLS
shape parameter=0.4						
(15,2)	14.249 (2.650)	15.071 (2.814)	45.362 (8.723)	48.383 (9.308)	145.011 (27.457)	154.328 (29.955)
(25,2)	13.198 (1.842)	13.661 (1.933)	42.474 (6.144)	43.537 (6.215)	134.006 (19.313)	138.478 (20.094)
(40,2)	12.546 (1.403)	12.809 (1.388)	40.178 (4.542)	40.764 (4.571)	127.199 (14.551)	129.144 (14.327)
(45,3)	12.404 (1.266)	12.534 (1.302)	39.673 (4.237)	40.114 (4.329)	125.830 (13.573)	127.000 (13.640)
(60,3)	12.086 (1.056)	12.139 (1.055)	38.600 (3.522)	38.904 (3.604)	122.370 (11.321)	123.418 (11.451)
(60,4)	12.031 (1.041)	12.154 (1.071)	38.621 (3.520)	38.985 (3.502)	122.352 (11.330)	123.144 (11.260)
(80,4)	11.843 (0.890)	11.873 (0.882)	37.737 (2.967)	38.014 (2.968)	119.903 (9.639)	120.566 (9.550)
shape parameter=1.0						
(15,2)	14.026 (2.625)	15.065 (2.833)	45.593 (8.739)	48.769 (9.413)	144.454 (27.706)	154.504 (29.901)
(25,2)	13.133 (1.844)	12.368 (1.295)	42.440 (6.100)	43.650 (6.175)	134.436 (19.747)	137.961 (20.015)
(40,2)	12.430 (1.386)	12.582 (1.401)	40.115 (4.536)	40.643 (4.610)	126.692 (14.397)	129.144 (14.806)
(45,3)	12.279 (1.225)	12.363 (1.289)	39.681 (4.282)	40.130 (4.256)	124.810 (13.752)	126.787 (13.411)
(60,3)	11.955 (1.060)	12.052 (1.081)	38.608 (3.547)	38.836 (3.476)	122.955 (11.158)	123.191 (11.391)
(60,4)	11.959 (1.055)	12.034 (1.047)	38.649 (3.620)	38.841 (3.584)	122.258 (11.222)	123.050 (11.612)
(80,4)	11.711 (0.904)	11.779 (0.901)	37.648 (2.945)	37.922 (3.073)	119.467 (9.600)	119.961 (9.549)
shape parameter=2.0						
(15,2)	14.013 (2.664)	14.859 (2.810)	45.711 (8.663)	48.489 (9.200)	144.196 (27.358)	153.727 (29.149)
(25,2)	13.017 (1.845)	13.402 (1.895)	42.368 (6.173)	43.520 (6.233)	134.558 (19.745)	137.891 (20.612)
(40,2)	12.327 (1.405)	12.446 (1.383)	40.104 (4.508)	40.651 (4.655)	127.209 (14.689)	128.878 (14.831)
(45,3)	12.164 (1.267)	12.337 (1.256)	39.604 (4.232)	40.116 (4.322)	125.220 (13.581)	127.210 (13.580)
(60,3)	11.871 (1.079)	11.950 (1.059)	38.584 (3.562)	39.047 (3.549)	122.454 (11.047)	123.168 (11.460)
(60,4)	11.865 (1.042)	11.948 (1.058)	38.501 (3.503)	38.795 (3.556)	122.081 (11.454)	123.504 (11.219)
(80,4)	11.631 (0.891)	11.681 (0.897)	37.679 (2.999)	37.930 (2.920)	119.526 (9.424)	120.319 (9.571)

(Table 4.6 continued)

shape parameter=5.0						
(15,2)	13.975 (2.701)	14.806 (2.788)	45.459 (8.574)	48.360 (8.887)	144.138 (26.991)	154.672 (29.416)
(25,2)	12.972 (1.800)	13.334 (1.916)	42.337 (6.129)	43.413 (6.350)	134.156 (19.101)	138.212 (19.483)
(40,2)	12.284 (1.375)	12.451 (1.327)	40.090 (4.581)	40.722 (4.583)	126.939 (14.155)	128.827 (14.632)
(45,3)	12.110 (1.275)	12.240 (1.302)	39.599 (4.300)	39.964 (4.271)	125.649 (13.843)	127.087 (13.772)
(60,3)	11.833 (1.087)	11.905 (1.094)	38.534 (3.522)	38.804 (3.525)	122.171 (11.312)	123.007 (11.438)
(60,4)	11.801 (1.064)	11.908 (1.069)	38.449 (3.579)	38.895 (3.571)	122.707 (11.344)	123.297 (11.020)
(80,4)	11.523 (0.905)	11.609 (0.903)	37.628 (3.062)	37.859 (3.052)	119.668 (9.590)	120.170 (9.655)
shape parameter=10.0						
(15,2)	13.958 (2.574)	14.864 (2.793)	45.282 (8.506)	48.539 (9.429)	144.751 (27.533)	154.434 (29.631)
(25,2)	12.986 (1.849)	13.216 (1.887)	42.633 (5.946)	43.523 (6.421)	134.051 (19.424)	138.230 (20.357)
(40,2)	12.274 (1.351)	12.434 (1.396)	40.118 (4.590)	40.657 (4.590)	126.779 (14.261)	128.785 (15.055)
(45,3)	12.031 (1.272)	12.237 (1.254)	39.554 (4.221)	39.921 (4.236)	125.605 (13.412)	126.890 (13.476)
(60,3)	11.807 (1.104)	11.867 (1.091)	38.535 (3.461)	38.825 (3.586)	122.537 (11.487)	123.701 (11.208)
(60,4)	11.786 (1.081)	11.871 (1.079)	38.483 (3.549)	38.866 (3.498)	122.064 (11.060)	123.143 (11.539)
(80,4)	11.537 (0.863)	11.612 (0.898)	37.730 (3.025)	37.812 (2.969)	119.455 (9.597)	119.968 (9.629)

Table 4.7: *Simulated coverage probabilities of the PMP and MLS tolerance intervals. Experimental error: GLD where $\lambda_1 = 0$, $\lambda_2 = 0.1975$.*

(n,t)	ρ					
	0.900		0.990		0.999	
	PMP	MLS	PMP	MLS	PMP	MLS
Normal approximation:						
$\lambda_3 = \lambda_4 = 0.1349$						
(15,2)	0.910	0.958	0.924	0.954	0.927	0.952
(25,2)	0.945	0.964	0.940	0.959	0.945	0.961
(40,2)	0.947	0.958	0.954	0.955	0.946	0.960
(45,3)	0.955	0.956	0.948	0.962	0.944	0.963
(60,3)	0.953	0.959	0.956	0.961	0.952	0.959
(60,4)	0.959	0.963	0.957	0.968	0.956	0.960
(80,4)	0.952	0.963	0.958	0.966	0.960	0.960
$\lambda_3 = \lambda_4 = 0.30$						
(15,2)	0.975	0.985	0.934	0.951	0.924	0.952
(25,2)	0.980	0.989	0.940	0.967	0.945	0.964
(40,2)	0.990	0.996	0.955	0.968	0.952	0.952
(45,3)	0.996	0.996	0.958	0.970	0.946	0.960
(60,3)	0.996	0.998	0.961	0.978	0.954	0.963
(60,4)	0.998	1.000	0.961	0.976	0.951	0.968
(80,4)	0.999	1.000	0.963	0.976	0.960	0.966
$\lambda_3 = 0.1349$, $\lambda_4 = 0.50$						
(15,2)	0.930	0.963	0.924	0.949	0.930	0.953
(25,2)	0.950	0.964	0.945	0.954	0.940	0.956
(40,2)	0.962	0.968	0.952	0.961	0.940	0.959
(45,3)	0.960	0.960	0.952	0.954	0.954	0.960
(60,3)	0.967	0.978	0.960	0.963	0.960	0.957
(60,4)	0.972	0.974	0.948	0.956	0.952	0.962
(80,4)	0.970	0.978	0.957	0.960	0.964	0.963
$\lambda_3 = 0.1349$, $\lambda_4 = 1.00$						
(15,2)	0.879	0.934	0.920	0.954	0.929	0.955
(25,2)	0.891	0.917	0.941	0.954	0.946	0.954
(40,2)	0.892	0.906	0.940	0.957	0.952	0.957
(45,3)	0.878	0.890	0.942	0.955	0.958	0.960
(60,3)	0.883	0.896	0.956	0.946	0.945	0.962
(60,4)	0.870	0.890	0.947	0.954	0.952	0.967
(80,4)	0.870	0.881	0.956	0.960	0.958	0.959

Table 4.8: *Expected widths of the PMP and MLS tolerance intervals. Experimental error: GLD where $\lambda_1 = 0$, $\lambda_2 = 0.1975$.*

(n,t)	ρ					
	0.900		0.990		0.999	
	PMP	MLS	PMP	MLS	PMP	MLS
Normal approximation:						
$\lambda_3 = \lambda_4 = 0.1349$						
(15,2)	14.216 (2.663)	14.688 (1.944)	45.510 (8.671)	48.397 (9.120)	144.831 (27.263)	154.111 (28.942)
(25,2)	13.243 (1.868)	13.677 (1.882)	42.397 (5.961)	43.584 (6.386)	134.257 (13.215)	138.684 (19.919)
(40,2)	12.598 (1.375)	12.740 (1.395)	40.124 (4.434)	40.858 (4.695)	127.376 (14.523)	128.754 (14.653)
(45,3)	12.468 (1.252)	12.584 (1.290)	39.583 (4.177)	40.240 (4.214)	125.186 (13.492)	127.187 (13.668)
(60,3)	12.126 (1.057)	12.213 (1.077)	38.720 (3.588)	38.883 (3.603)	122.259 (11.490)	122.874 (11.270)
(60,4)	12.127 (1.043)	12.205 (1.068)	38.572 (3.538)	39.093 (3.622)	122.297 (11.320)	123.091 (11.282)
(80,4)	11.833 (0.902)	11.893 (0.887)	37.753 (2.958)	37.926 (2.987)	119.824 (9.392)	120.182 (9.650)
$\lambda_3 = \lambda_4 = 0.30$						
(15,2)	15.455 (2.670)	16.428 (2.861)	46.173 (8.687)	49.004 (9.574)	145.554 (28.053)	153.888 (29.309)
(25,2)	14.458 (1.944)	14.846 (1.990)	42.621 (6.147)	44.106 (6.319)	134.666 (19.224)	138.361 (20.136)
(40,2)	13.682 (1.409)	13.980 (1.418)	40.450 (4.558)	41.125 (4.536)	127.383 (14.519)	129.061 (14.948)
(45,3)	13.487 (1.269)	13.638 (1.285)	39.897 (4.147)	40.476 (4.231)	125.453 (13.621)	126.925 (13.673)
(60,3)	13.163 (1.041)	13.296 (1.075)	39.003 (3.508)	39.407 (3.450)	122.526 (11.313)	123.172 (11.267)
(60,4)	13.112 (1.009)	13.251 (1.048)	39.024 (3.630)	39.369 (3.541)	122.008 (11.332)	123.385 (11.194)
(80,4)	12.850 (0.881)	12.958 (0.862)	37.989 (3.032)	38.394 (3.014)	119.694 (9.566)	120.075 (9.392)

(Table 4.8 continued)

$\lambda_3 = 0.1349$, $\lambda_4 = 0.50$						
(15,2)	15.237 (2.756)	16.229 (2.876)	46.037 (8.679)	48.764 (9.234)	145.226 (28.319)	155.172 (29.960)
(25,2)	14.272 (1.934)	14.681 (1.963)	42.610 (6.070)	43.893 (6.418)	134.152 (19.403)	137.564 (19.568)
(40,2)	13.524 (1.392)	13.735 (1.411)	40.570 (4.528)	41.060 (4.587)	127.275 (14.581)	129.036 (14.911)
(45,3)	13.301 (1.290)	13.400 (1.262)	40.054 (4.225)	40.251 (4.306)	125.848 (13.167)	127.358 (13.677)
(60,3)	12.998 (1.071)	13.112 (1.034)	38.944 (3.521)	39.235 (3.522)	122.755 (11.235)	123.334 (11.605)
(60,4)	12.029 (0.943)	13.079 (1.047)	38.881 (3.595)	39.206 (3.670)	122.427 (11.363)	123.139 (11.549)
(80,4)	12.727 (0.869)	12.776 (0.871)	38.005 (2.990)	38.200 (2.985)	119.875 (9.457)	120.046 (9.607)
$\lambda_3 = 0.1349$, $\lambda_4 = 1.00$						
(15,2)	15.739 (2.675)	16.894 (2.910)	46.157 (8.592)	49.103 (9.246)	145.380 (27.889)	153.122 (28.782)
(25,2)	14.782 (1.904)	15.228 (1.961)	43.052 (6.012)	44.198 (6.350)	134.836 (19.458)	137.380 (20.310)
(40,2)	14.102 (1.400)	14.248 (1.419)	40.655 (4.558)	41.257 (4.556)	127.099 (14.412)	128.452 (14.610)
(45,3)	13.787 (1.257)	13.943 (1.282)	40.240 (4.233)	40.628 (4.229)	126.280 (13.372)	127.423 (13.847)
(60,3)	13.205 (1.052)	13.645 (1.061)	39.238 (3.451)	39.396 (3.563)	121.962 (11.382)	123.247 (11.340)
(60,4)	13.416 (1.013)	13.554 (1.042)	39.061 (3.546)	39.419 (3.540)	122.588 (11.364)	123.200 (11.193)
(80,4)	13.205 (0.875)	13.258 (0.873)	38.122 (2.904)	38.406 (2.945)	120.029 (9.444)	120.379 (9.570)

Figure 4.1: *The density shapes produced by GLD (0, 0.1975, λ_3, λ_4)*

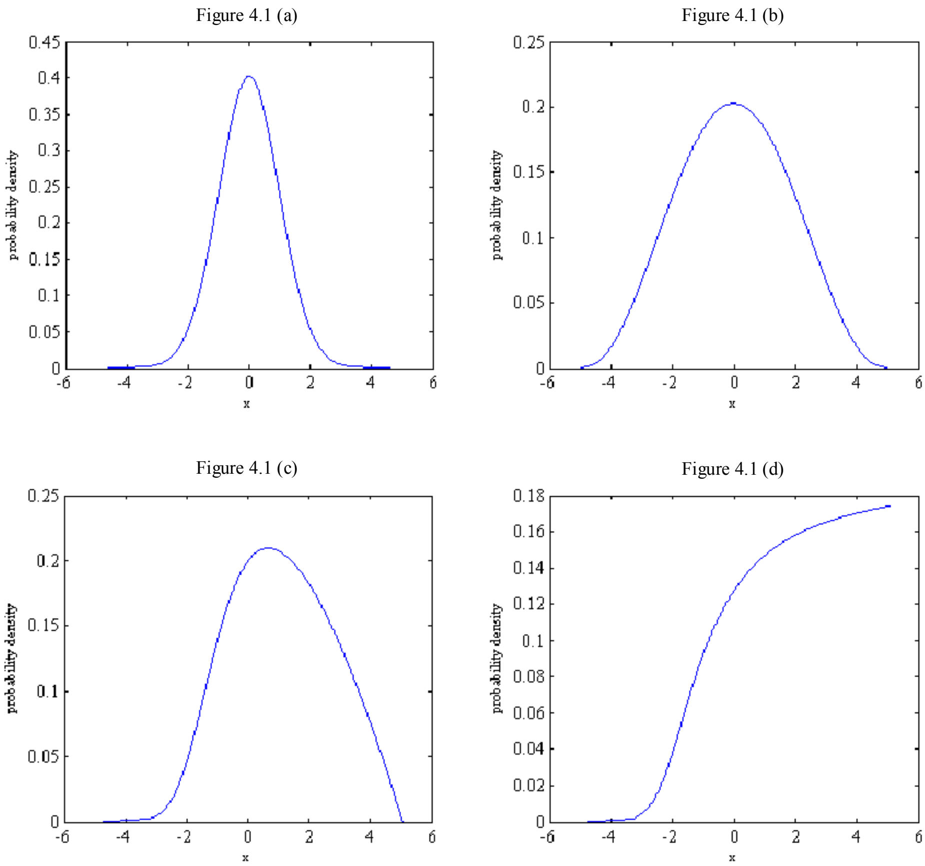

Figure 4.1 (a) $\lambda_3 = \lambda_4 = 0.1349$ *(standard normal fit)*

Figure 4.1 (b) $\lambda_3 = \lambda_4 = 0.30$

Figure 4.1 (c) $\lambda_3 = 0.1349, \lambda_4 = 0.50$

Figure 4.1 (d) $\lambda_3 = 0.1349, \lambda_4 = 1.00$

4.3.2 Discussion

Experimental error following the standard normal distribution

Tables 4.1 and 4.2 give the coverage probabilities and expected widths for various ρ and some combinations of n and t when the error distribution is standard normal. Both PMP and MLS tolerance intervals show conservatism in terms of coverage probabilities for small and moderate values of ρ but the PMP method is slightly less conservative for moderate ρ. The MLS tolerance interval seems to work well for smaller sample sizes and shows slight conservatism as the number of classes increases. The PMP tolerance interval appears to be more accurate for larger values of ρ and has coverage probability close to the nominal value 0.95 when the number of classes is around 25 to 50, when t remains as 2. It is necessary to maintain the balance between n and t to achieve coverage probability close to 0.95. The ratio n: t is approximately 12.5:1 to 25:1 to attain this for the PMP case. The expected widths for the MLS tolerance interval are wider than that of the PMP for sample sizes less than 50. The wider expected widths for the MLS case enable it to cover a proportion closer to 0.95 for smaller sample sizes.

Experimental error following the t- distribution

Tables 4.3 and 4.4 show the coverage probabilities and expected widths for data generated with experimental error following the t-distribution with degrees of freedom 3, 5, 10, 15 and 25. The results for small and moderate ρ were not reported as they are conservative and have coverage probabilities close to 1. Since $\theta_2 = \theta_3 \rho /(1-\rho)$, when ρ is small, the distribution of the experimental error following the t-distribution (variance θ_3) dominates the distribution of the normal random effects (variance θ_2). This will have an effect on both PMP and MLS tolerance intervals which have been derived under the assumption that the underlying distribution is normal (Gaussian). We noticed that the coverage probabilities tend to 1 for small and moderate ρ. This is due to the wider expected widths since the t-distribution has heavier tail than the normal distribution. When $\rho = 0.900$, the coverage probabilities for both PMP and MLS tolerance intervals get closer to 0.95 as the degrees of freedom increase from 15 onwards. It seems that the coverage probabilities happen to be close to the nominal level 0.95 for degrees of freedom as small as 3 for both cases as $\rho = 0.990$ and 0.999. The expected widths for both instances are comparable to the standard normal case.

Experimental error following the skew-normal distribution

We study the tolerance intervals for the skew-normal distribution whose tail is heavier than the normal distribution involving different shape parameters. Both PMP and MLS tolerance intervals seem to have coverage probabilities close to 0.95 when α is small i.e. 0.40 for $\rho = 0.900$. The results become conservative as α increases and are still acceptable for $\alpha = 1.00$. However, the results involving the expected widths and coverage probabilities tend to be comparable to that of the standard normal case when $\rho = 0.990$ and 0.999. The results for negative shape are very similar to the positive shape parameters.

The results for small and moderate ρ were not reported here since they are conservative and the coverage probabilities are very close to 0. The convergence of the coverage probability to 0 is more rapid for larger sample sizes as the number of classes increase and the skew-normal characteristics in the data become more dominant. The non-symmetrical behavior of the skew-normal distribution, where one tail is pulled in one direction, affects the coverage probability as the data becomes non-centred when ρ becomes smaller. Hence, both PMP and MLS tolerance intervals for symmetrical distributions such as the normal case become conservative for the skew-normal case for small and moderate ρ.

Experimental error following the generalized lambda distribution

As for the experimental error following the generalized lambda distribution, we refer to Equation (4.10) and use the normal approximation parameters suggested by Karian and Dudewicz (2000). Tables 4.7 and 4.8 clearly show that the results for these estimates are comparable with the standard normal case.

We examine the performance of the GLD error distribution by varying the parameters λ_3 and λ_4. Here $\lambda_1 > 0$ and $\lambda_2 > 0$ for all the cases studied. According to Karian and Dudewicz (2000), the GLD density has limited support given by $[\lambda_1 - 1/\lambda_2, \lambda_1 + 1/\lambda_2]$. In our study, the support of the GLD density is [-5.063, 5.063]. Figure 4.1 shows the shapes produced by the pdf plots by varying these parameters. For $\lambda_3 = \lambda_4 = 0.30$, the distribution is symmetrical with flatter and heavier tail than the normal distribution. For $\lambda_4 = 0.50$ and 1.00 where $\lambda_1 = 0, \lambda_2 = 0.1975, \lambda_3 = 0.1349$, the distribution is no longer symmetrical. The distribution for $\lambda_4 = 0.50$ is also flatter and has a heavier tail than that of the normal distribution. Generally, the standard normal fit in Figure 4.1(a) produces outputs close to the standard normal case. The results involving PMP and MLS tolerance intervals are conservative for $\rho \leq 0.900$ and more accurate for $\rho = 0.990$ and 0.999. We did not report the outputs for $\rho \leq 0.900$ in the tables as the distributions in Figure 4.1(b) and Figure 4.1(d) have coverage probabilities very close to 1. This is due to the aforementioned characteristics of these distributions which dominate the normal random effect and result in wider expected widths than the normal case. The coverage probabilities for Figure 4.1(d) which is *S*-shaped with limited support and no longer symmetrical are conservative for $\rho \le 0.900$ as they stay slightly lower than that reported for $\rho = 0.900$. The PMP and MLS tolerance intervals seem to be comparable with the normal case when $\rho = 0.990$ and 0.999.

The simulation results in Tables 4.1-4.8 show that both PMP and MLS tolerance intervals are comparable for large number of classes, n and small number of observations per class, t with coverage probability closer to the nominal value, 0.95. The MLS tolerance interval appears to be good for small sample sizes. For non-normal distributions whose tails are heavier than the normal distribution, both PMP and MLS tolerance intervals appear to be less conservative for large values of intra-class correlation coefficient, ρ.

CHAPTER 5

BAYESIAN AND FREQUENTIST TOLERANCE INTERVALS IN A GENERAL CASE

5.1 Introduction

It is an unquestionable fact that the study of two-sided tolerance intervals is more challenging than that of its one-sided counterpart. To appreciate the reason, consider a random sample from a univariate population characterized by a cumulative distribution function (c.d.f.) $F(x;\theta)$, where θ is a possibly vector valued unknown parameter. Then a one-sided $\beta-$content tolerance interval, associated with a lower tolerance limit, is of the form $[T,\infty)$, where T is a statistic so chosen that the relationship

$$1-F(T;\theta)\geq\beta, \tag{5.1}$$

holds with credibility or confidence level γ. On the other hand, a two-sided $\beta-$content tolerance interval is of the form $[T_1,T_2]$, the statistics T_1 and T_2 being such that the relationship

$$F(T_2;\theta)-F(T_1;\theta)\geq\beta. \tag{5.2}$$

holds with credibility or confidence level γ. If we write $q(\alpha;\theta)$ for the α th quantile of the population, then clearly (5.1) holds if and only if

$$T\leq q(1-\beta;\theta) \tag{5.3}$$

Therefore we can regard T as the $(1-\gamma)$ th posterior quantile of $q(1-\beta;\theta)$ in the Bayesian setup, or as a lower confidence limit for $q(1-\beta;\theta)$ with confidence coefficient γ in the frequentist setup, and this simplifies the study of one-sided tolerance intervals (Pathmanathan et al., 2013). However, no such reduction occurs for the inequality (5.2) arising in the two-sided case. This makes the construction of two-sided tolerance intervals intrinsically difficult. So far, no direct method for this purpose, which works under reasonable generality, is available.

Earlier, Mukerjee and Reid (2001) characterized the probability matching priors for one-sided tolerance intervals by taking note of the equivalence between (5.1) and (5.3). The corresponding results in the two-sided case were so far unknown. The results obtained in this chapter enable us to fill in this gap. In contrast, as indicated above, we give analytical formulae for such intervals, applicable to a wide range of parametric models and based on the foundation of higher order asymptotics. Moreover, the main idea of this chapter concerns the development of general results in the frequentist setup, where the Bayesian simulation approach in Wolfinger (1998) does not work. We aim at exploring two-sided tolerance intervals in a fairly general framework of parametric models. Explicit analytical formulae for these tolerance intervals in both Bayesian and frequentist setups were obtained by developing higher order asymptotics. The Bayesian results lead to a characterization for probability matching priors ensuring approximate frequentist validity of two-sided Bayesian tolerance intervals. We also examine such matching priors and their role in finding frequentist tolerance intervals via a Bayesian route. Based on our observation, we take cognizance of the fact that it is difficult to obtain matching priors in some situations. Hence, we formulate purely frequentist tolerance intervals that cater to situations of this kind. We address computational issues as well and note that it is straightforward to write programs for easy implementation of our explicit formulae. Finally, applications to real data from Gacula and Kubala (1975) for the Weibull case and Lieblin and Zelen (1956) for the inverse Gaussian tolerance intervals are presented. Simulation studies were conducted to investigate if the asymptotic results are well reflected in finite samples.

5.2 Two-sided Bayesian tolerance intervals

Let $X_1,\ldots,X_n$ be independent and identically distributed scalar-valued observation from a population specified by a density $f(x;\theta)$. Here $\theta=(\theta_1,\ldots,\theta_p)'$ is an unknown parameter that belongs to the *p*-dimensional Euclidean space or some open subset thereof. We work under the assumptions in Johnson (1970) for the Bayesian tolerance intervals. The Edgeworth assumptions in Bhattacharya and Ghosh (1978) will be applied for the frequentist calculation reported later. These two sets of assumptions hold under wider generality for models belonging to the exponential and curved exponential families and also for many other models such as Cauchy, Student's *t* and so on; see Datta and Mukerjee (2004) for more details. In what follows, for any $t \geq 0$, we write $O_p(n^{-t})$ to represent a quantity which, even when multiplied by n^t, remains bounded in probability as n tends to infinity; see Rao (1973).

Let $F(x;\theta)$ be the cumulative distribution function (cdf) corresponding to $f(x;\theta)$. $q(\alpha;\theta)$ is the α th quantile of the population represented by $F(x;\theta)$.
The interval $[q(\beta_2;\theta), q(1-\beta_1;\theta)]$, covers a proportion β of this population for a known θ where β_1, β_2 (>0) satisfy $1-\beta_1-\beta_2=\beta$. For notational simplicity, we write $b(\theta)=q(1-\beta_1;\theta)$ and $d(\theta)=q(\beta_2;\theta)$ which motivates us to consider a two-sided Bayesian tolerance interval of the following form:

$$[d(\hat{\theta})-g^{(n)},\, b(\hat{\theta})+g^{(n)}] \tag{5.4}$$

where $\hat{\theta} = (\hat{\theta}_1, ..., \hat{\theta}_p)'$ is the maximum likelihood estimator for $\theta = (\theta_1, ..., \theta_p)'$ for the data $X = (X_1, ..., X_n)$, and

$$g^{(n)} = n^{-1/2} g_1 + n^{-1} g_2 + O_p(n^{-3/2}) \tag{5.5}$$

where g_1, g_2 are $O_p(1)$ functions of data X (which may as well involve the prior). We have to choose g_1, g_2 so that interval (5.5) has $\beta-$content and posterior credibility level $\gamma + O_p(n^{-1})$, that is,

$$P^{\pi}\{F(b(\hat{\theta}) + g^{(n)}; \theta) - F(d(\hat{\theta}) - g^{(n)}; \theta) \geq \beta \mid X\} = \gamma + O_p(n^{-1}) \tag{5.6}$$

$P^{\pi}(. \mid X)$ is the posterior probability measure under the prior $\pi(\theta)$.

Further details on the precise form of $g^{(n)}$ in Equation (5.5) will be discussed in Remarks 5.1 and 5.5. We note that in most applications, especially with a symmetric density $f(x;\theta)$, taking $\beta_1 = \beta_2$ is fine since our results go through for arbitrary β_1, β_2 (>0) satisfying $1 - \beta_1 - \beta_2 = \beta$.

One may want to choose $\beta_1 > \beta_2$ if being at the upper extreme is considered more atypical than being at the lower extreme or $\beta_1 < \beta_2$ if it is the other way round.

Theorem 5.1 gives explicit formulae for g_1 and g_2 which ensures the attainment of Equation (5.6). Some of the notations used in presenting Theorem 5.1 as well as the rest of this chapter are summarized below.

For $1 \leq s, u, w \leq p$,

$$l(\theta) = n^{-1} \sum_{i=1}^{n} \log f(X_i; \theta), \tag{5.7}$$

$$D_s \equiv \partial / \partial\theta_s,\ D_{su} \equiv \partial^2 / \partial\theta_s \delta\theta_u,\ D_{suw} \equiv \partial^3 / \partial\theta_s \delta\theta_u \delta\theta_w \tag{5.8}$$

$$c_{su} = -\{D_s D_u l(\theta)\}_{\theta=\hat{\theta}},\ a_{suw} = \{D_s D_u D_w l(\theta)\}_{\theta=\hat{\theta}} \tag{5.9}$$

$C =$ (c_{su}) is the per observation observed information matrix at $\hat{\theta}$. We write

$C^{-1} = (c^{su})$, (5.10)

$\hat{\pi} = \pi(\hat{\theta})$, $\pi_s(\theta) = D_s\pi(\theta)$, $\hat{\pi}_s = \pi_s(\hat{\theta})$, (5.11)

$\widetilde{f}(x;\theta) = \partial f(x;\theta)/\partial x$, $f_s(x;\theta) = D_s f(x;\theta)$ [$= \partial F_s(x;\theta)/\partial x$],

$F_s(x;\theta) = D_s F(x;\theta)$, $F_{su}(x;\theta) = D_s D_u F(x;\theta)$,

$f^d = f(d(\hat{\theta});\hat{\theta})$, $\widetilde{f}^d = \widetilde{f}(d(\hat{\theta});\hat{\theta})$, $f_s^d = f_s(d(\hat{\theta});\hat{\theta})$,

$F_s^d = F_s(d(\hat{\theta});\hat{\theta})$, $F_{su}^d = F_{su}(d(\hat{\theta});\hat{\theta})$,

$f^b = f(b(\hat{\theta});\hat{\theta})$, $\widetilde{f}^b = \widetilde{f}(b(\hat{\theta});\hat{\theta})$, $f_s^b = f_s(b(\hat{\theta});\hat{\theta})$,

$F_s^b = F_s(b(\hat{\theta});\hat{\theta})$, $F_{su}^b = F_{su}(b(\hat{\theta});\hat{\theta})$,

$K_s(\theta) = F_s(d(\theta);\theta) - F_s(b(\theta);\theta)$, $\hat{K}_s = K_s(\hat{\theta}) = F_s^d - F_s^b$.

Let $K(\theta)$ and $\hat{K}$ be $p \times 1$ vectors with sth elements given by $K_S(\theta)$ and $\hat{K}_S$, respectively. We assume that $K(\theta)$ is non-null for every θ. This implies that $\hat{K}$ is also non-null and that, as a result, the quantity $M = (c^{su}\hat{K}_s\hat{K}_u)^{1/2}$ is positive. When defining M and also in the rest of this chapter, the summation convention is followed, with implicit sums on repeated sub- or superscripts in a product ranging over 1,…, p i.e.

$\sum_{s=1}^{p}\sum_{u=1}^{p} c^{su}\hat{K}_s\hat{K}_u$ is written as $c^{su}\hat{K}_s\hat{K}_u$.

For $1 \le s,u \le p$, let

$$A_s = \frac{(F_s^d - F_s^b)}{M} = \frac{\hat{K}_s}{M}, \qquad B_s = \frac{(f_s^d + f_s^b)}{M},$$

$$V_{su} = \frac{(F_{su}^d - F_{su}^b)}{M}, \qquad \lambda_s = c^{su}A_u \tag{5.12}$$

THEOREM 5.1 *The tolerance interval* $[\,d(\hat\theta)-g^{(n)}, b(\hat\theta)+g^{(n)}\,]$ *has* $\beta-$*content with posterior credibility level* $\gamma+O_p(n^{-1})$, *i.e. Equation* (5.6) *holds, provided* g_1 *and* g_2 *in the expression* (5.5) *for* $g^{(n)}$ *satisfy*

$$g_1 = \frac{M z_\gamma}{(f^d+f^b)} \quad \text{and} \quad g_2 = \left\{\frac{M}{(f^d+f^b)}\right\}\{L_1(\pi)+L_2+L_3(z_\gamma^2-1)\}+g_1^2 L_4,$$

where z_γ *is the* γ *th quantile of the standard univariate normal distribution, and*

$$L_1(\pi)=\frac{\hat\pi_s}{\hat\pi}\lambda_s, \qquad L_2=\frac{1}{2}(a_{suw}\lambda_w c^{su}+c^{su}V_{su}),$$

$$L_3=\frac{1}{6}a_{suw}\lambda_s\lambda_u\lambda_w+\frac{1}{2}\lambda_s\lambda_u V_{su}, \qquad L_4=\frac{1}{2}\left\{\frac{(\tilde f^d-\tilde f^b)}{(f^d+f^b)}\right\}-\lambda_s B_s \tag{5.13}$$

Proof of Theorem 5.1:

We define $h=(h_1,...,h_p)' = n^{1/2}(\theta-\hat\theta)$. Thus, $\theta=\hat\theta+\frac{h}{\sqrt n}$. Invoking Equation (5.5), we find that by Taylor's expansion,

$$F(b(\hat\theta)+g^{(n)};\theta)=F(b(\hat\theta)+g^{(n)};\hat\theta+\frac{h}{\sqrt n})$$

$$=F(b(\hat\theta)+g^{(n)};\hat\theta)+\frac{h_s}{\sqrt n}F_s(b(\hat\theta)+g^{(n)};\hat\theta)+\frac{1}{2n}h_s h_u F_{su}(b(\hat\theta)+g^{(n)};\hat\theta)$$
$$+O_p(n^{-3/2})$$

Next, we expand with respect to $b(\hat\theta)+g^{(n)}$ and ignore terms higher than $1/n$ for

$$F(b(\hat\theta)+g^{(n)};\hat\theta)=\left\{F(b(\hat\theta);\hat\theta)+\left(\frac{g_1}{\sqrt n}+\frac{g_2}{n}\right)f^b+\frac{1}{2n}g_1^2\tilde f^b+...\right\}$$

$$F_s(b(\hat\theta)+g^{(n)};\hat\theta)=\left[F_s(b(\hat\theta);\hat\theta)+\frac{1}{\sqrt n}g_1 f_s(b(\hat\theta);\hat\theta)+...\right]$$

to obtain

$$F(b(\hat\theta)+g^{(n)};\theta)=\left\{F(b(\hat\theta);\hat\theta)+\left(\frac{g_1}{\sqrt n}+\frac{g_2}{n}\right)f^b+\frac{1}{2n}g_1^2\tilde f^b\right\}+\frac{h_s}{\sqrt n}\left[F_s^b+\frac{1}{\sqrt n}g_1 f_s^b\right]$$

$$+\frac{1}{2n}h_s h_u F_{su}^b + O_p(n^{-3/2})$$

$$F(b(\hat{\theta})+g^{(n)};\theta) = F(b(\hat{\theta});\hat{\theta}) + n^{-1/2}(g_1 f^b + h_s F_s^b)$$
$$+n^{-1}\left(\frac{1}{2}g_1^2\tilde{f}^b + g_2 f^b + g_1 h_s f_s^b + \frac{1}{2}h_s h_u F_{su}^b\right) + O_p(n^{-3/2})$$

In a similar manner, we obtain

$$F(d(\hat{\theta})-g^{(n)};\theta) = F(d(\hat{\theta});\hat{\theta}) + n^{-1/2}(-g_1 f^d + h_s F_s^d)$$
$$+n^{-1}\left(\frac{1}{2}g_1^2\tilde{f}^d - g_2 f^d - g_1 h_s f_s^d + \frac{1}{2}h_s h_u F_{su}^d\right) + O_p(n^{-3/2})$$

Since $F(b(\hat{\theta});\hat{\theta}) - F(d(\hat{\theta});\hat{\theta}) = 1-\beta_1-\beta_2 = \beta$, by the definitions of $b(\theta)$ and $d(\theta)$, recalling the definition of R, we get

$$R = F(b(\hat{\theta})+g^{(n)};\theta) - F(d(\hat{\theta})-g^{(n)};\theta)$$
$$= \beta + n^{-1/2}\left\{g_1(f^b+f^d) + h_s(F_s^b - F_s^d)\right\}$$
$$+n^{-1}\left(\frac{1}{2}g_1^2(\tilde{f}^b-\tilde{f}^d) + g_2(f^b+f^d) + g_1 h_s(f_s^b+f_s^d) + \frac{1}{2}h_s h_u(F_{su}^b - F_{su}^d)\right)$$
$$+O_p(n^{-3/2}) \qquad (5.14)$$

From (5.14) above, on rearranging the terms,

$$\frac{\sqrt{n}(R-\beta)}{M} = \frac{g_1(f^b+f^d)}{M} - \frac{h_s(F_s^d - F_s^b)}{M} - n^{-1/2}\left[-\frac{g_1 h_s(f_s^d+f_s^b)}{M} + \frac{1}{2}h_s h_u\frac{(F_{su}^d - F_{su}^b)}{M}\right.$$
$$\left.-\left(\frac{\frac{1}{2}g_1^2(\tilde{f}^b-\tilde{f}^d) + g_2(f^d+f^b)}{M}\right)\right] + O_p(n^{-1})$$

Let $G_1 = \dfrac{g_1(f^b + f^d)}{M}$, $\quad G_2 = \left(\dfrac{\frac{1}{2} g_1^2(\tilde{f}^b - \tilde{f}^d) + g_2(f^d + f^b)}{M} \right)$ (5.15)

A_s, B_s and V_{su} are as given in (5.12)

Therefore, we write

$$\begin{aligned} \frac{\sqrt{n}(R-\beta)}{M} &= G_1 - \left\{ h_s A_s + n^{-1/2}\left(-g_1 h_s B_s + \frac{1}{2} h_s h_u V_{su} - G_2 \right) + O_p(n^{-1}) \right\} \\ &= G_1 - Y \end{aligned} \tag{5.16}$$

where

$$Y = h_s A_s + n^{-1/2}\left(-g_1 h_s B_s + \frac{1}{2} h_s h_u V_{su} - G_2 \right) + O_p(n^{-1}) \tag{5.17}$$

In view of Equation (5.16), we next consider the posterior density of *Y*. Following chapter 2 of Datta and Mukerjee (2004), note that, the posterior density of *h*, $\pi(\theta)$, can be expressed as

$$\pi_{\text{post}}(h \mid X) = \phi_p(h; C^{-1})\left[1 + n^{-1/2}\left\{ \left(\frac{\hat{\pi}_s}{\hat{\pi}} \right) h_s + \frac{1}{6} a_{suw} h_s h_u h_w \right\} \right] + O_p(n^{-1})$$

where $\phi_p(.;C^{-1})$ represents the *p*-variate normal density with null mean vector and covariance matrix C^{-1}.

Let $\xi = (-1)^{1/2}\tau$, with τ as an auxiliary variate. Then, by Equation (5.17) we get

$$\begin{aligned} \exp(\xi Y) &= \exp\left\{ \xi h_s A_s + n^{-1/2}\xi\left[-g_1 h_s B_s + \frac{1}{2} h_s h_u V_{su} - G_2 \right] + O_p(n^{-1}) \right\} \\ &= \exp(\xi h_s A_s) \exp\left\{ n^{-1/2}\xi\left[-g_1 h_s B_s + \frac{1}{2} h_s h_u V_{su} - G_2 \right] \right\} + O_p(n^{-1}) \\ &= \exp(\xi h_s A_s)\left\{ 1 + n^{-1/2}\xi\left[-g_1 h_s B_s + \frac{1}{2} h_s h_u V_{su} - G_2 \right] \right\} + O_p(n^{-1}) \end{aligned}$$

(Taylor series expansion, $e^x = 1 + x + \dfrac{x^2}{2!} + \ldots$)

$$\exp(\xi Y)\pi_{\text{post}}(h \mid X) = \exp(\xi h_s A_s)\phi_p(h; C^{-1})\left\{1 + n^{-1/2}\left[\xi\left(-g_1 h_s B_s + \frac{1}{2}h_s h_u V_{su} - G_2\right)\right.\right.$$

$$\left.\left. + \left(\frac{\hat{\pi}_s}{\hat{\pi}}\right)h_s + \frac{1}{6}a_{suw}h_s h_u h_w\right]\right\} + O_p(n^{-1})$$

Hence, recalling the definition of λ_s from Equation (5.12),

let $\lambda = (\lambda_1, \lambda_2, ..., \lambda_p)' = C^{-1}A$, $A = (A_1, A_2, ..., A_p)'$.

$$\begin{aligned}(h - \xi\lambda)'C(h - \xi\lambda) &= h'Ch - 2\xi\lambda'Ch + \xi^2\lambda'C\lambda \\ &= h'Ch - 2\xi h'A + \xi^2\end{aligned}$$

where $\lambda'C = (A'C^{-1})C = A'$ and

$\lambda'C\lambda = (A'C^{-1})C(C^{-1}A) = A'C^{-1}A = 1 \qquad \because c^{su}A_s A_u = 1$

$\therefore\ \exp(\xi h_s A_s)\phi_p(h; C^{-1}) = \exp(\tfrac{1}{2}\xi^2)\phi_p(h; \xi\lambda, C^{-1})$

where $\phi_p(h; \xi\lambda, C^{-1})$ is a p-variate normal density with mean vector $\xi\lambda$ and covariance matrix C^{-1}.

$$\exp(\xi Y)\pi_{\text{post}}(h \mid X) = \exp\left(\frac{1}{2}\xi^2\right)\phi_p(h; \xi\lambda, C^{-1})\left\{1 + n^{-1/2}\left[\xi\left(-g_1 h_s B_s + \frac{1}{2}h_s h_u V_{su} - G_2\right)\right.\right.$$

$$\left.\left. + \left(\frac{\hat{\pi}_s}{\hat{\pi}}\right)h_s + \frac{1}{6}a_{suw}h_s h_u h_w\right]\right\} + O_p(n^{-1}) \qquad (5.18)$$

If $h = (h_1, ..., h_p)'$ has density $\phi_p(h; \xi\lambda, C^{-1})$, then

$$\left.\begin{aligned}&E[h_s] = \xi\lambda_s \\ &E[h_s h_u] = c^{su} + \xi^2\lambda_s\lambda_u \\ &E[h_s h_u h_w] = \xi^3\lambda_s\lambda_u\lambda_w + \xi(\lambda_s c^{uw} + \lambda_u c^{sw} + \lambda_w c^{uw})\end{aligned}\right\} \qquad (5.19)$$

Then by applying (5.19), we integrate Equation (5.18) with respect to h to obtain the approximate posterior characteristic function of Y which is

$$\exp(\tfrac{1}{2}\xi^2)\Big\{1+n^{-1/2}\Big[-\xi^2 g_1\lambda_s B_s+\tfrac{1}{2}\xi(c^{su}+\xi^2\lambda_s\lambda_u)V_{su}-\xi G_2+(\hat{\pi}_s/\hat{\pi})\xi\lambda_s$$
$$+\tfrac{1}{6}a_{suw}\Big(\xi^3\lambda_s\lambda_u\lambda_w+\xi(\lambda_s c^{uw}+\lambda_u c^{sw}+\lambda_w c^{su})\Big)\Big]\Big\}+O_p(n^{-1})$$
$$=\exp(\tfrac{1}{2}\xi^2)\Big\{1+n^{-1/2}[\big(L_1(\pi)+L_2-G_2\big)\xi-g_1(\lambda_s B_s)\xi^2+L_3\xi^3]\Big\}+O_p(n^{-1})$$

Inverting the approximate posterior characteristic function of Y as noted above, we now get the posterior density of Y, under $\pi(\theta)$, as given by

$$\widetilde{\pi}(y\mid X)=\phi(y)[1+n^{-1/2}\{(L_1(\pi)+L_2-G_2)y-g_1(\lambda_s B_s)(y^2-1)+L_3(y^3-3y)\}]+O_p(n^{-1}),$$

After some simplification, $L_1(\pi)$, L_2 and L_3 are as shown in Equation (5.13).

From Equation (5.16), we observe that

$R\geq\beta$ if and only if $Y\leq G_1$ where $R=F(b(\hat{\theta})+g^{(n)};\theta)-F(d(\hat{\theta})-g^{(n)};\theta)$

Therefore, the integration of $\widetilde{\pi}(y\mid X)$ over $Y\leq G_1$ yields

$$P^{\pi}\{F(b(\hat{\theta})+g^{(n)};\theta)-F(d(\hat{\theta})-g^{(n)};\theta)\geq\beta\mid X\}$$
$$=P^{\pi}\{Y\leq G_1\mid X\}$$
$$=\Phi(G_1)-n^{-1/2}\{L_1(\pi)+L_2-G_2-g_1(\lambda_s B_s)G_1+L_3(G_1^2-1)\}\phi(G_1)+O_p(n^{-1}) \qquad (5.20)$$

where $\Phi(\cdot)$ is the standard univariate normal cdf.

The right hand side of (5.20) equals $\gamma+O_p(n^{-1})$ provided

$$G_1=z_\gamma \text{ and } G_2=L_1(\pi)+L_2-g_1(\lambda_s B_s)z_\gamma+L_3(z_\gamma^2-1) \qquad (5.21)$$

i.e. recalling Equation (5.15), provided g_1 and g_2 are as in the statement of Theorem 5.1.

Remark 5.1

Based on Theorem 5.1, g_1 is free from the prior while g_2 involves the prior only via the term $L_1(\pi)$. It is easy to find $g^{(n)}$ satisfying Equation (5.5) with g_1 and g_2 as in Theorem 5.1. For instance, we can simply take $g^{(n)} = n^{-1/2}g_1 + n^{-1}g_2$. However, other choices of $g^{(n)}$ are possible and may be helpful in certain situations. We will come back to this point in the next section.

Remark 5.2

It is noted that Theorem 5.1 is applicable even to models, such as the inverse Gaussian, which do not admit analytical expressions for $d(\theta)$ and $b(\theta)$. We only require the values of these functions at $\theta = \hat{\theta}$, i.e. $d(\hat{\theta})$ and $b(\hat{\theta})$ and these can be found computationally. The quantities in Equations (5.12) and (5.13) as well as the expressions for g_1 and g_2 in Theorem 5.1 involve the partial derivatives of $f(x;\theta)$ and $F(x;\theta)$, as evaluated at $\theta = \hat{\theta}$ and $x = d(\hat{\theta})$ or $b(\hat{\theta})$. It is easy to obtain these partial derivatives via symbolic computation via the software MATLAB which calculates g_1 and g_2 almost instantaneously, given the data X.

5.3 Frequentist tolerance intervals via probability matching prior

We consider the frequentist behaviour of the Bayesian tolerance interval in Theorem 5.1 with a view of characterizing priors under which it has $\beta-$content not only with posterior credibility level $\gamma + O_p(n^{-1})$ but also with frequentist confidence level $\gamma + O(n^{-1})$. Such a prior is referred to as probability matching prior for a two-sided tolerance interval. Therefore, the Bayesian tolerance interval in Theorem 5.1, when constructed using a prior of this kind, is also frequentist. A consideration of these priors provides a Bayesian route for obtaining two-sided frequentist tolerance intervals. To achieve this, we let $I = (I_{su})$ denote the per observation Fisher information matrix at θ, and write $M_0 = \{I_{su} K_s(\theta) K_u(\theta)\}^{1/2}$, where $I^{-1} = (I^{su})$. We note that $M_0 > 0$ because of our assumption that the vector $K(\theta)$ is non-null for every θ. Then the following results characterizing probability matching priors in the present context, holds. A shrinkage argument which is popular in Bayesian asymptotics will be employed in proving Theorem 5.2 and Theorem 5.3 (Datta and Mukerjee, 2004, Ch. 4).

The following will be useful in proving Theorem 5.2 and Theorem 5.3.

From Datta and Mukerjee, 2004, pp.5-7,

$h = (h_1, h_2, ..., h_p)' = n^{1/2}(\theta - \hat{\theta})$

$\hat{\theta} = \theta - n^{-1/2} h$

$$\hat{\theta} = \theta + O_p(n^{-1/2}) \qquad (5.22)$$

From (5.7), $l(\theta) = n^{-1} \sum_{i=1}^{n} \log f(X_i; \theta)$.

From (5.9) and (5.10), $c_{su} = -\{D_s D_u l(\theta)\}_{\theta=\hat{\theta}}$ and $C^{-1} = (c^{su})$.

$I_{su} \equiv I_{su}(\theta) = -E_\theta[\{D_s D_u \log f(X_i; \theta)\}]$ is the Fisher information matrix at θ. Define

$I^{-1} = (I^{su})$

$$E_\theta(\hat{\theta}) = \theta + O(n^{-1/2}) \qquad (5.23)$$

Comment: (5.23) follows from (5.22) noting that the expectations of both sides of (5.22) follow the same pattern as (5.22) itself under very general conditions such as those in Bhattacharya and Ghosh (1978).

Expanding $D_s D_u l(\hat{\theta})$ about θ, we get

$c_{su} = I_{su} + O(n^{-1/2})$

$c^{su} = I^{su} + O(n^{-1/2})$

$$\therefore\ E_\theta(c^{su}) = I^{su} + O(n^{-1/2}) \tag{5.24}$$

Comment: Since c_{su} and I_{su} are the observed and expected information elements and c^{su} and I^{su} are the (s, u)-th elements of the inverses of $C = (c_{su})$ and $I = (I_{su})$, we readily have $c_{su} = I_{su} + O(n^{-1/2})$ and $c^{su} = I^{su} + O(n^{-1/2})$. Again under very general conditions such as those in Bhattacharya and Ghosh (1978), the same pattern holds for the expectations of both sides of these equations.

THEOREM 5.2 *The Bayesian tolerance interval in Theorem 5.1 has $\beta-$content with frequentist confidence level $\gamma + O(n^{-1})$ if and only if the prior $\pi(\theta)$ satisfies the partial differential equation*

$$D_s\{M_0^{-1} I^{su} K_u(\theta)\pi(\theta)\} = 0 \tag{5.25}$$

Proof of Theorem 5.2:

Take an auxiliary prior $\pi^*(\cdot)$ which vanishes along the boundaries of a rectangle containing the true θ. Then, with $g^{(n)}$, g_1 and g_2 as in Theorem 5.1, analogously to (5.20),

$$P^{\pi *}\{F(b(\hat{\theta}) + g^{(n)};\theta) - F(d(\hat{\theta}) - g^{(n)};\theta) \geq \beta \mid X\}$$

$$= \Phi(G_1) - n^{-1/2}\{L_1(\pi^*) + L_2 - G_2 - g_1(\lambda_s B_s)G_1 + L_3(G_1^2 - 1)\}\phi(G_1) + O_p(n^{-1})$$

From (5.21),

$$\therefore P^{\pi *}\{F(b(\hat{\theta})+g^{(n)};\theta)-F(d(\hat{\theta})-g^{(n)};\theta)\geq \beta \mid X\}$$

$$= \gamma + n^{-1/2}\{L_1(\pi)-L_1(\pi^*)\}\phi(z_\gamma)+O_p(n^{-1}) \tag{5.26}$$

As in (5.13), here $L_1(\pi^*)=\left(\frac{\hat{\pi}_s^*}{\hat{\pi}^*}\right)\lambda_s$, with $\hat{\pi}^*=\pi^*(\hat{\theta})$, $\hat{\pi}_s^*=\pi_s^*(\hat{\theta})$ and $\pi_s^*(\theta)=D_s\pi^*(\theta)$.

Now,

$$L_1(\pi^*)=\frac{\hat{\pi}_s^*}{\hat{\pi}^*}\lambda_s$$

$$=\frac{\hat{\pi}_s^*}{\hat{\pi}^*}c^{su}A_u$$

$$=\frac{\hat{\pi}_s^*}{\hat{\pi}^*}c^{su}\frac{K_u(\hat{\theta})}{M} \tag{5.27}$$

Comment: Therefore,

$$E_\theta[L_1(\pi^*)] = \; = \frac{\pi_s^*}{\pi^*}M_0^{-1}I^{su}K_u(\theta)+O_p(n^{-1/2}), \tag{5.28}$$

where $M_0=\{I^{su}K_s(\theta)K_u(\theta)\}^{1/2}$. We are simply replacing each term in $L_1(\pi^*)$ by its population analogue. Hence, recalling the forms of λ_s and A_s as shown in (5.12), it follows from (5.26) and (5.28) that

$$E_\theta[P^{\pi *}\{F(b(\hat{\theta})+g^{(n)};\theta)-F(d(\hat{\theta})-g^{(n)};\theta)\geq \beta \mid X\}]$$

$$= \gamma + n^{-1/2}\left[\left(\frac{\pi_s(\theta)}{\pi(\theta)}-\frac{\pi_s^*(\theta)}{\pi^*(\theta)}\right)M_0^{-1}I^{su}K_u(\theta)\right]\phi(z_\gamma)+O(n^{-1}). \tag{5.29}$$

The last step of shrinkage argument involves integrating the $E_\theta -$expectation in (5.29) by parts with respect to $\pi^*(\theta)$ and then allowing $\pi^*(\theta)$ to converge weakly to the degenerate prior at the true θ. After some simplification, this shows that the tolerance interval in Theorem 5.1 has $\beta -$content with frequentist confidence level

$$P_\theta\{F(b(\hat{\theta})+g^{(n)};\theta)-F(d(\hat{\theta})-g^{(n)};\theta)\geq\beta\}$$
$$=\gamma+n^{-1/2}\{\pi(\theta)\}^{-1}D_s\{M_0^{-1}I^{su}K_u(\theta)\pi(\theta)\}\phi(z_\gamma)+O(n^{-1}).$$

The above equals $\gamma+O(n^{-1})$ if and only if $\pi(\theta)$ satisfies the partial differential equation (5.25).

Remark 5.3

The matching condition in (5.25) has a striking similarity with that for the posterior quantiles of a parametric function $\psi(\theta)$, with margin of error $O(n^{-1})$, as given by (Datta and Mukerjee, 2004, p.42),

$$D_s\{H_0^{-1}I^{su}\psi_u(\theta)\pi(\theta)\} = 0 \tag{5.30}$$

where $H_0 = \{I^{su}\psi_s(\theta)\psi_u(\theta)\}^{1/2}$ and $\psi_u(\theta) = D_u\psi(\theta)$, $1\leq u\leq p$. Although for $p\geq 2$, it is difficult to find an example of a $\psi(\theta)$ satisfying $\psi_u(\theta) = K_u(\theta)$, $1\leq u\leq p$, in many situations, the I^{su} are such that a solution to (5.25) also meets (5.30) for some $\psi(\theta)$. For instance, the solutions to (5.25) in the location-scale and Weibull models below satisfy (5.30) as well for $\psi(\theta)=\theta_2$. This is a bit surprising because posterior quantiles give one-sided credible sets while we are considering two-sided tolerance intervals here. Moreover, there is no obvious link between a two-sided tolerance interval and the posterior quantiles of a parametric function.

Remark 5.4

In particular, if $p = 1$, i.e., θ is a scalar, then both $K(\theta)$ and I are scalars. In this case, if the parameter space is an interval then the assumption that $K(\theta)$ is nonnull for every θ implies that $K(\theta)$ is either positive for all θ or negative for all θ. Hence, by the definition of M_0, the matching condition (5.25) reduces to $d\{I^{-1/2}\pi(\theta)\}/d\theta = 0$, with unique solution $\pi_0(\theta) \propto I^{1/2}$, the Jeffreys' prior. Thus, we obtain a probability matching property of Jeffreys' prior for two-sided tolerance intervals in the case of scalar θ.

Remark 5.5

Returning to the case of general *p*, if a matching prior, say $\pi_0(\theta)$, satisfying (5.25) is available, then as hinted earlier, Theorem 5.1 readily yields a two-sided tolerance interval which has $\beta-$content with frequentist confidence level $\gamma + O(n^{-1})$. For this purpose, one only needs to work with the prior $\pi_0(\theta)$ in Theorem 5.1, and this amounts to keeping g_1 as stated there while replacing $L_1(\pi)$ by $L_1(\pi_0)$ in the expression for g_2. With g_1 and g_2 so determined, there are numerous choices of $g^{(n)}$ satisfying (5.5). These include $g^{(n)} = g^{(1n)}$, $g^{(2n)}$ and $g^{(3n)}$, where

$$g^{(1n)} = n^{-1/2}g_1 + n^{-1}g_2, \qquad g^{(2n)} = n^{-1/2}g_1 \exp\left(\frac{n^{-1/2}g_2}{g_1}\right),$$

$$\begin{aligned} g^{(3n)} &= \frac{n^{-1/2}g_1}{(1 - n^{-1/2}g_2/g_1)}, \qquad \text{if } \frac{n^{-1/2}g_2}{g_1} < 1, \\ &= g^{(2n)}, \text{ otherwise.} \end{aligned} \tag{5.31}$$

Typically, $\gamma > 0.5$, so that by Theorem 5.1, $g_1 > 0$. Therefore, $g^{(1n)} \le g^{(2n)} \le g^{(3n)}$, as $e^x \ge 1 + x$ for every real *x*. As a result, tolerance intervals given by $g^{(n)} = g^{(1n)}$, $g^{(2n)}$

and $g^{(3n)}$ have $\beta-$content with progressively higher exact frequentist confidence levels, but this comes at the cost of progressively higher expected widths. Simulation studies enable us to get the suitable choice of $g^{(n)}$ for a given model. For j = 1, 2, 3 and sample size n, write $\gamma^{(jn)}$ for the simulated frequentist coverage probability for the tolerance interval given by $g^{(n)} = g^{(jn)}$. Take $g^{(n)} = g^{(1n)}$ if $\gamma^{(1n)}$ converges fast to the target γ. In this case, taking $g^{(n)} = g^{(2n)}$ or $g^{(3n)}$ will make the interval unnecessarily long. If, however, $\gamma^{(1n)}$ falls short of γ even for moderate n, then take $g^{(n)} = g^{(2n)}$ provided $\gamma^{(2n)}$ converges fast to γ. On the other hand, if $\gamma^{(2n)}$ too falls short of γ even for moderate n, then try $g^{(n)}=g^{(3n)}$. In our examples and also others not shown here, the above strategy works well and one of $g^{(1n)}$, $g^{(2n)}$ and $g^{(3n)}$ leads to a fast convergence to the target γ; e.g., as seen in Section 5.5, in the setups of the univariate normal model with both mean and variance equal to $\theta(>0)$ and the Weibull model below which happens with $g^{(n)} = g^{(1n)}$ and $g^{(n)} = g^{(2n)}$, respectively.

(a) Univariate normal model with mean=variance= θ

Let $f(x;\theta)$ represent the univariate normal model with mean and variance both equal to $\theta(>0)$. Here θ, and hence $K(\theta)$, are scalars. Suppose $\beta > \frac{1}{2}$, as in most practical situations. This implies that $\beta_j < \frac{1}{2}$ (j = 1, 2), as $\beta=1-\beta_1-\beta_2$. The standard univariate normal density is written as $\phi(\cdot)$. For notational simplicity, let $z^{(j)}$ be its $(1-\beta_j)$ th quantile (j = 1, 2). Then $b(\theta)= \theta+z^{(1)}\sqrt{\theta}$, $d(\theta)=\theta-z^{(2)}\sqrt{\theta}$.

$$f(x;\theta)=\frac{1}{\sqrt{2\pi\theta}}\exp\left(-\frac{1}{2\theta}(x-\theta)^2\right), \qquad -\infty<x<\infty \tag{5.32}$$

$$F(x;\theta) = \Phi\left(\frac{x-\theta}{\sqrt{\theta}}\right)$$

$$F_1(x;\theta) = \frac{dF(x;\theta)}{d\theta}$$

$$= \phi\left(\frac{x-\theta}{\sqrt{\theta}}\right)\left[-\frac{1}{2}x\theta^{-3/2} - \frac{1}{2}\theta^{-1/2}\right]$$

$$= -\frac{1}{2\sqrt{\theta}}\phi\left(\frac{x-\theta}{\sqrt{\theta}}\right)\left(\frac{x}{\theta}+1\right)$$

$$F_1(d(\theta);\theta) = -\frac{1}{2\sqrt{\theta}}\phi(z^{(2)})\left[-\frac{z^{(2)}}{\sqrt{\theta}}+2\right] \text{ and}$$

$$F_1(b(\theta);\theta) = -\frac{1}{2\sqrt{\theta}}\phi(z^{(1)})\left[\frac{z^{(1)}}{\sqrt{\theta}}+2\right]$$

$$K_1(\theta) = F_1(d(\theta);\theta) - F_1(b(\theta);\theta)\}$$

$$\therefore\ K_1(\theta) = (2\theta)^{-1}\{z^{(1)}\phi(z^{(1)}) + z^{(2)}\phi(z^{(2)})\} + \theta^{-1/2}\{\phi(z^{(1)}) - \phi(z^{(2)})\} \tag{5.33}$$

Note that $z^{(1)}$ and $z^{(2)}$ are both positive, because $\beta_j < \frac{1}{2}$ (j = 1, 2). Thus, $K_1(\theta)$ does not change sign over $\theta > 0$ if and only if $z^{(1)} \le z^{(2)}$, or equivalently, $\beta_1 \ge \beta_2$. The lack of symmetry in the condition just obtained is not totally unexpected even though $f(x;\theta)$ is symmetric about θ in this case while the quantity $F_1(x;\theta) = dF(x;\theta)/d\theta$ is not so. At any rate, following Remark 5.4 when $\beta_1 \ge \beta_2$, the Jeffreys' prior, $\pi_0(\theta) \propto \theta^{-1}\sqrt{2\theta+1}$ satisfies the matching condition in (5.25).

The Jeffreys' prior is easily obtained by noting that for $f(x;\theta)$ in (5.32),

$$\log f(x;\theta) = -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log\theta - \frac{x^2}{2\theta} + 2 - \frac{\theta}{2},$$

$$\frac{d\log f(x;\theta)}{d\theta} = -\frac{1}{2\theta} + \frac{x^2}{2\theta^2} - \frac{1}{2}$$

$$\frac{d^2 \log f(x;\theta)}{d\theta^2} = \frac{1}{2\theta^2} - \frac{x^2}{\theta^3},$$

$$I = \frac{E_\theta(X^2)}{\theta^3} - \frac{1}{2\theta^2} \qquad \text{where} \qquad E_\theta(X^2) = Var_\theta(X) + [E_\theta(X)]^2 = \theta + \theta^2$$

$$= \frac{\theta + \theta^2}{\theta^3} - \frac{1}{2\theta^2}$$

$$= \frac{1}{\theta} + \frac{1}{2\theta^2}$$

Hence, Jeffreys' prior is given by $\pi_0(\theta) = \sqrt{\frac{1}{\theta} + \frac{1}{2\theta^2}}$ i.e. $\pi_0(\theta) \propto \theta^{-1}\sqrt{2\theta + 1}$

(b) Location scale model

Consider the location-scale model

$$f(x;\theta) = \frac{1}{\theta_2}\rho\left(\frac{x-\theta_1}{\theta_2}\right), \qquad -\infty < x < \infty, \tag{5.34}$$

where $\theta = (\theta_1, \theta_2)'$, $-\infty < \theta_1 < \infty$, $\theta_2 > 0$, where $\rho(.)$ is a density on the real line with $\rho(t) > 0$, $-\infty < t < \infty$. Let $Q(u) = \int_{-\infty}^{u} \rho(t)dt$. Then, $F(x;\theta) = Q\left(\frac{x-\theta_1}{\theta_2}\right)$.

Define q_a as $Q(q_a) = \alpha$. Then the α th quantile of $F(x;\theta)$ is given by $\theta_1 + q_\alpha \theta_2$

Let q_1 and q_2 denote respectively the $(1-\beta_1)$ th and β_2 th quantiles of the distribution represented by $\rho(.)$. Then $b(\theta) = \theta_1 + q_1\theta_2$ and $d(\theta) = \theta_1 + q_2\theta_2$.

$F_1(x;\theta) = \rho\left(\frac{x-\theta_1}{\theta_2}\right)\left(-\frac{1}{\theta_2}\right)$ and $F_2(x;\theta) = \rho\left(\frac{x-\theta_1}{\theta_2}\right)\left(-\frac{(x-\theta_1)}{\theta_2^2}\right)$

$F_1(d(\theta);\theta) = -\theta_2^{-1}\rho(q_2)$ and $F_1(b(\theta);\theta) = -\theta_2^{-1}\rho(q_1)$

$F_2(d(\theta);\theta) = -q_2\theta_2^{-1}\rho(q_2)$ and $F_2(b(\theta);\theta) = -q_1\theta_2^{-1}\rho(q_1)$

$K_s(\theta) = F_s(d(\theta);\theta) - F_s(b(\theta);\theta)$, s=1, 2

$$K_1(\theta) = \theta_2^{-1}\{\rho(q_1) - \rho(q_2)\}, \qquad K_2(\theta)) = \theta_2^{-1}\{q_1\rho(q_1) - q_2\rho(q_2)\} \tag{5.35}$$

Therefore, the assumption that $K(\theta)$ is non-null for every θ holds. If $K(\theta)$ is null, $\rho(q_1) = \rho(q_2)$ and $q_1\rho(q_1) = q_2\rho(q_2)$. These imply that $q_1 = q_2$, i.e., $\beta = 1 - \beta_1 - \beta_2 =$ 0, which is impossible. For the location-scale model, $I^{su} \propto \theta_2^2$ for every s, u, and hence by Equation (5.35), M_0 is a constant free from θ. As a result, $\pi_0(\theta) \propto \theta_2^{-1}$ emerges as a solution to the matching condition (5.25). As mentioned previously in Mukerjee and Reid (2001), the same prior also enjoys the probability matching property for one-sided tolerance intervals in this case.

(c) Weibull Model

Consider the Weibull model given by

$$f(x;\theta) = \frac{\theta_2}{\theta_1}\left(\frac{x}{\theta_1}\right)^{\theta_2 - 1} \exp\left\{-\left(\frac{x}{\theta_1}\right)^{\theta_2}\right\}, \; x > 0, \tag{5.36}$$

where $\theta = (\theta_1, \theta_2)'$ and $\theta_1, \theta_2 > 0$.

Here $b(\theta) = \theta_1 \varepsilon_1^{1/\theta_2}$ and $d(\theta) = \theta_1 \varepsilon_2^{1/\theta_2}$, where $\varepsilon_1 = -\log \beta_1$ and $\varepsilon_2 = -\log(1 - \beta_2)$.

$$F(x;\theta) = 1 - \exp\left\{-\left(\frac{x}{\theta_1}\right)^{\theta_2}\right\}$$

$$F_1(x;\theta) = -\exp\left\{-\left(\frac{x}{\theta_1}\right)^{\theta_2}\right\}\frac{\theta_2}{\theta_1}\left(\frac{x}{\theta_1}\right)^{\theta_2}$$

$$F_2(x;\theta) = \exp\left\{-\left(\frac{x}{\theta_1}\right)^{\theta_2}\right\}\left(\frac{x}{\theta_1}\right)^{\theta_2}\log\left(\frac{x}{\theta_1}\right)$$

$$F_1(d(\theta);\theta) = -\frac{\theta_2}{\theta_1}(1 - \beta_2)\varepsilon_2 \text{ and } F_1(b(\theta);\theta) = -\frac{\theta_2}{\theta_1}(\beta_1)\varepsilon_1$$

$$F_2(d(\theta);\theta) = \theta_2^{-1}(1 - \beta_2)\varepsilon_2 \log \varepsilon_2 \text{ and } F_2(b(\theta);\theta) = \theta_2^{-1}\beta_1 \varepsilon_1 \log \varepsilon_1$$

Therefore,

$$K_1(\theta) = \frac{\theta_2}{\theta_1}\{\beta_1\varepsilon_1 - (1-\beta_2)\varepsilon_2\}\,,\;\; K_2(\theta) = \theta_2^{-1}\{(1-\beta_2)\varepsilon_2 \log\varepsilon_2 - \beta_1\varepsilon_1 \log\varepsilon_1\}\,, \tag{5.37}$$

$I^{11} \propto (\theta_1/\theta_2)^2$, $I^{12}(=I^{21}) \propto \theta_1$ and $I^{22} \propto \theta_2^2$ (see section 5.5 for details). Therefore by (5.37), M_0 is a constant free from θ. As a result, $\pi_0(\theta) \propto (\theta_1\theta_2)^{-1}$ appears as a solution to the matching condition (5.25).

Solutions to the matching condition (5.25) were available for the three models discussed in (a)-(c). However, there are instances where finding a solution to Equation (5.25) can be a daunting task and one example is the inverse Gaussian model. This is mainly because such models do not seem to admit analytical expressions for $b(\theta)$ and $d(\theta)$, and consequently do not allow us to explicitly write Equation (5.25). Hence, it is not always possible for us to obtain a two-sided frequentist tolerance interval using a matching prior in Theorem 5.1, and a direct method is employed. We will discuss this in the following section. Interestingly, although this is a purely frequentist problem, Bayesian arguments continue to be handy.

5.4 Purely frequentist two-sided tolerance intervals

As mentioned in Remark 5.1, the Bayesian tolerance interval obtained in Theorem 5.1 depends on the prior $\pi(\cdot)$ only via the term $L_1(\pi)$, of order $O_p(1)$, in the expression for g_2. This leads us to consider a purely frequentist tolerance interval of the same form, with $L_1(\pi)$ substituted appropriately by a term which is also of order $O_p(1)$ but does not involve any prior. Theorem 5.3 encapsulates the results so obtained.

We write $I^{su} \equiv I^{su}(\theta)$ to make explicit the dependence of I^{su} on θ, and define $\hat{I}^{su} = I^{su}(\hat{\theta})$ and $\hat{I}_w^{su} = I_w^{su}(\hat{\theta})$, where $I_w^{su}(\theta) = D_w I^{su}(\theta)$, $1 \le s,u,w \le p$.

Also, let $\hat{\Delta}_{su} = \Delta_{su}(\hat{\theta})$, where $\Delta_{su}(\theta) = \Delta_{su}^d(\theta) - \Delta_{su}^b(\theta)$, with,

$$\Delta_{su}^d(\theta) = F_{su}(d(\theta);\theta) - \frac{F_s(d(\theta);\theta) f_u(d(\theta);\theta)}{f(d(\theta);\theta)}, \tag{5.38}$$

$\Delta_{su}^b(\theta)$ is similarly defined by replacing $d(\theta)$ by $b(\theta)$ in (5.38).

Lemma 5.1 $L_{1f} = \mu(\theta) + O_p(n^{-1/2})$, where

$$\mu(\theta) = -D_s\{M_0^{-1} I^{su} K_u(\theta)\}. \tag{5.39}$$

Lemma 5.1 plays a crucial role in proving Theorem 5.3

THEOREM 5.3. *The tolerance interval* $[\,d(\hat{\theta}) - g_f^{(n)}, b(\hat{\theta}) + g_f^{(n)}\,]$, *where*

$$g_f^{(n)} = n^{-1/2} g_1 + n^{-1} g_{2f} + O_p(n^{-3/2}),$$

$$g_1 = \frac{M z_\gamma}{(f^d + f^b)} \quad \textit{and} \quad g_{2f} = \left\{\frac{M}{(f^d + f^b)}\right\}\left[L_{1f} + L_2 + L_3(z_\gamma^2 - 1)\right] + g_1^2 L_4,$$

with L_2, L_3, L_4 *as in* (5.13)*, and*

$$L_{1f} = \frac{1}{2} M^{-3} \hat{I}^{su} \hat{K}_u (\hat{I}_s^{vw} \hat{K}_v \hat{K}_w + 2\hat{I}^{vw} \hat{K}_v \hat{\Delta}_{sw}) - M^{-1}(\hat{I}_s^{su} \hat{K}_u + \hat{I}^{su} \hat{\Delta}_{su}), \tag{5.40}$$

has $\beta -$*content with frequentist confidence level* $\gamma + O(n^{-1})$.

Although the expression for L_{1f} in (5.40) seems a bit involved, it has a simple interpretation. Thus, in a sense L_{1f} can be considered as the sample analogue of $\mu(\theta)$. The advantage of the form in (5.40) is that it allows calculation of L_{1f} even when analytical expressions for $d(\theta)$ and $b(\theta)$ are not available, because we only require $d(\hat{\theta})$ and $b(\hat{\theta})$ for this purpose.

Proof of Lemma 5.1:

With a view to writing $\mu(\theta)$ explicitly, we first note that $F(d(\theta);\theta) = \beta_2$, by the definition of $d(\theta)$. Upon partial differentiation with respect to θ_s, we have

$$F_s(d(\theta);\theta) + f(d(\theta);\theta)\{D_s d(\theta)\} = 0$$

$$D_s d(\theta) = -\frac{F_s(d(\theta);\theta)}{f(d(\theta);\theta)}$$

Thus, by Equation (5.38),

$$\begin{aligned} D_s F_u(d(\theta);\theta) &= F_{su}(d(\theta);\theta) + f_u(d(\theta);\theta)\{D_s d(\theta)\} \\ &= F_{su}(d(\theta);\theta) - \frac{F_s(d(\theta);\theta)}{f(d(\theta);\theta)} f_u(d(\theta);\theta) \\ &= \Delta_{su}^d(\theta) \end{aligned} \qquad (5.41)$$

We obtain a similar expression for $D_s F_u(b(\theta);\theta)$ i.e. $\Delta_{su}^b(\theta)$

Since $K_u(\theta) = F_u(d(\theta);\theta) - F_u(b(\theta);\theta)$,

$$\begin{aligned} D_s K_u(\theta) &= \Delta_{su}^d(\theta) - \Delta_{su}^b(\theta) \\ &= \Delta_{su}(\theta) \end{aligned} \qquad (5.42)$$

Recalling that $M_0 = \{I^{su} K_s(\theta) K_u(\theta)\}^{1/2}$,

Here $M_0^{-1} = \{I^{vw} K_v(\theta) K_w(\theta)\}^{-1/2}$

$$\begin{aligned} D_s M_0^{-1} = -\frac{1}{2}\{I^{vw} K_v(\theta) K_w(\theta)\}^{-3/2}[(D_s I^{vw}) K_v(\theta) K_w(\theta) + I^{vw}[D_s K_v(\theta)] K_w(\theta) \\ + I^{vw} K_v(\theta)[D_s K_w(\theta)] \end{aligned}$$

$$= -\frac{1}{2} M_0^{-3} [I_s^{vw} K_v(\theta) K_w(\theta) + 2 I^{vw} K_v(\theta) \Delta_{sw}(\theta)]$$

$$\mu(\theta) = -D_s \{ M_0^{-1} I^{su} K_u(\theta) \}$$

$$= \frac{1}{2} M_0^{-3} I^{su} K_u(\theta) \{ I_s^{vw} K_v(\theta) K_w(\theta) + 2 I^{vw} K_v(\theta) \Delta_{sw}(\theta) \}$$

$$- M_0^{-1} \{ I_s^{su} K_u(\theta) + I^{su} \Delta_{su}(\theta) \} \tag{5.43}$$

From (5.40) and (5.43), the conclusion of the lemma is evident.

Proof of Theorem 5.3:

We use *Lemma 5.1* and the shrinkage argument to find an expression for $P_\theta \{ F(b(\hat{\theta}) + g_f^{(n)}; \theta) - F(d(\hat{\theta}) - g_f^{(n)}; \theta) \geq \beta \}$, the frequentist confidence interval considered.

We take an auxiliary prior $\pi^*(.)$ which vanishes on the boundaries of a rectangle containing the true θ. Then, we get as in the derivation of the Bayesian tolerance interval,

$$P^{\pi^*} \{ F(b(\hat{\theta}) + g_f^{(n)}; \theta) - F(d(\hat{\theta}) - g_f^{(n)}; \theta) \geq \beta \mid X \}$$

$$= \Phi(G_1) - n^{-1/2} \{ L_1(\pi^*) + L_2 - G_{2f} - g_1(\lambda_s B_s) G_1 + L_3(G_1^2 - 1) \} \phi(G_1) + O_p(n^{-1})$$

where G_1 and G_{2f} are given in (5.21).

By using the expressions for g_1 and g_{2f},

$$P^{\pi^*} \{ F(b(\hat{\theta}) + g_f^{(n)}; \theta) - F(d(\hat{\theta}) - g_f^{(n)}; \theta) \geq \beta \mid X \}$$

$$= \Phi(G_1) - n^{-1/2} \{ L_1(\pi^*) + L_2 - G_{2f} - g_1(\lambda_s B_s) G_1 + L_3(G_1^2 - 1) \} \phi(G_1) + O_p(n^{-1})$$

$$= \gamma - n^{-1/2} \{ L_1(\pi^*) + L_2 - [L_{1f} + L_2 - g_1(\lambda_s B_s) z_\gamma + L_3(z_\gamma^2 - 1)] - g_1(\lambda_s B_s) z_\gamma$$

$$+ L_3(z_\gamma^2 - 1) \} \phi(z_\gamma) + O_p(n^{-1})$$

$$=\gamma + n^{-1/2}\{L_{1f} - L_1(\pi^*)\}\phi(z_\gamma) + O_p(n^{-1}), \tag{5.44}$$

which is analogous to Equation (5.26).

From Equation (5.28), $E_\theta[L_1(\pi^*)] = \dfrac{\pi_s^*}{\pi^*} M_0^{-1} I^{su} K_u(\theta) + O_p(n^{-1/2})$.

Hence by Equation (5.44) and Lemma 5.1,

$$E_\theta[P^{\pi^*}\{F(b(\hat{\theta}) + g_f^{(n)};\theta) - F(d(\hat{\theta}) - g_f^{(n)};\theta) \geq \beta \mid X\}]$$

$$= \gamma - n^{-1/2}\left[D_s\{M_0^{-1} I^{su} K_u(\theta)\} + \left\{\frac{\pi_s^*(\theta)}{\pi^*(\theta)}\right\} M_0^{-1} I^{su} K_u(\theta) \right]\phi(z_\gamma) + O(n^{-1}) \tag{5.45}$$

We shall integrate the $E_\theta -$ expectation in the above by parts with respect to $\pi^*(\theta)$ and then allow $\pi^*(\theta)$ to converge weakly to the degenerate prior at the true θ. Then it is immediate that the tolerance interval in Theorem 5.3 has $\beta -$ content with frequentist confidence level,

$$P_\theta\{F(b(\hat{\theta}) + g_f^{(n)};\theta) - F(d(\hat{\theta}) - g_f^{(n)};\theta) \geq \beta\}$$

$$= \gamma - n^{-1/2}[D_s\{M_0^{-1} I^{su} K_u(\theta)\} - D_s\{M_0^{-1} I^{su} K_u(\theta)\}]\phi(z_\gamma) + O(n^{-1})$$

$$= \gamma + O(n^{-1}).$$

All the other points stated in Remark 5.2 in the context of Theorem 5.1 also hold for Theorem 5.3. Consequently, it is not difficult to write a program which almost instantaneously computes the values of g_1 and g_{2f} in Theorem 5.3 for a given model and a given data set.

Analogues to Equation (5.31), the choices for $g_f^{(n)}$ include $g_f^{(n)} = g_f^{(1n)}$, $g_f^{(2n)}$ and $g_f^{(3n)}$, where

$$g_f^{(n)} = g_f^{(1n)} = n^{-1/2} g_1 + n^{-1} g_{2f}, \qquad g_f^{(n)} = g_f^{(2n)} = n^{-1/2} g_1 \exp\left(\frac{n^{-1/2} g_{2f}}{g_1} \right)$$

$$g_f^{(n)} = g_f^{(3n)} = \frac{n^{-1/2} g_1}{(1 - n^{-1/2} g_{2f} / g_1)} \quad \text{if} \quad \frac{n^{-1/2} g_{2f}}{g_1} < 1$$

$$= g_f^{(2n)}, \text{ otherwise.} \tag{5.46}$$

Along the lines of Remark 5.5, consideration of simulated frequentist coverage probabilities can again throw light on a suitable choice of $g_f^{(n)}$ for a given model. For the inverse Gaussian model, the choice $g_f^{(n)} = g_f^{(3n)}$ works reasonably well.

While concluding this section, we note the strong similarity between the expression for $\mu(\theta)$ in (5.39) and the matching condition (5.25). This suggests that Bayesian arguments should be useful even in proving Theorem 5.3 which is a purely frequentist result. The proof above shows that this is, indeed, the case.

5.5 Simulation study and application to real data

In this section, the numerical studies relate to normal models whose mean and variance are equal, the Weibull model and the inverse Gaussian model. The derivation of expressions required while applying Theorem 5.1 and Theorem 5.3 are shown in (I)-(III).

(I) The normal model with equal mean and variance

For the normal model whose mean and variance are equal, it can be seen that maximum likelihood estimate MLE is $\hat{\theta} = \frac{1}{2}\{[1 + 4n^{-1}\sum_{i=1}^{n} X_i^2]^{1/2} - 1\}$.

Here p=1, and $c_{11} = \frac{(2\hat{\theta}+1)}{(2\hat{\theta}^2)}$, $a_{111} = \frac{(3\hat{\theta}+2)}{\hat{\theta}^3}$.

To obtain c_{11} and a_{111}

From (5.32), $f(x;\theta) = \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{1}{2\theta}(x-\theta)^2}$

$$\log f(x;\theta) = \text{constant} - \frac{1}{2}\log\theta - \frac{1}{2\theta}(x-\theta)^2$$

$$= \text{constant} - \frac{1}{2}\log\theta - \frac{x^2}{2\theta} - \frac{\theta}{2}$$

where the constant is free from θ.

$$l(\theta) = -\frac{1}{2}\log\theta - \frac{1}{2\theta n}\sum_{i=1}^{n} x_i^2 - \frac{\theta}{2} \qquad (5.47)$$

$$\frac{d}{d\theta} l(\theta) = -\frac{1}{2\theta} + \frac{1}{2\theta^2}\left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right) - \frac{1}{2} = 0$$

$$\theta^2 + \theta - \frac{1}{n}\sum_{i=1}^{n} x_i^2 = 0$$

Hence, $\hat{\theta} = \frac{1}{2}\{[1 + 4n^{-1}\sum_{i=1}^{n} X_i^2]^{1/2} - 1\}$.

$$\frac{d^2}{d\theta} l(\theta) = \frac{1}{2\theta^2} - \frac{1}{n\theta^3}\sum_{i=1}^{n} x_i^2$$

$$c_{11} = -\frac{d^2}{d\theta} l(\hat{\theta})$$

$$= -\frac{1}{2\hat{\theta}^2} + \frac{1}{n\hat{\theta}^3}(n\hat{\theta}^2 + n\hat{\theta}) \qquad \left(\because \sum_{i=1}^{n} x_i^2 = n\hat{\theta}^2 + n\hat{\theta}\right)$$

$$c_{11} = \frac{2\hat{\theta}+1}{2\hat{\theta}^2}$$

$$\frac{d^3}{d\theta} l(\theta) = -\frac{1}{\theta^3} + \frac{3}{n\theta^4}\sum_{i=1}^{n} x_i^2$$

$$a_{111} = -\frac{1}{\hat{\theta}^3} + \frac{3}{n\hat{\theta}^4}(n\hat{\theta}^2 + n\hat{\theta}) = \frac{3\hat{\theta}+2}{\hat{\theta}^3}$$

(II) The Weibull model

The pdf of the Weibull distribution is shown in (5.36). The closed-form expression for the MLE, $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2)'$ is not available but $\hat{\theta}$ can be readily calculated for a given data set using standard statistical software.

Here $p = 2$, and we can check that

$$c_{11} = \left(\frac{\hat{\theta}_2}{\hat{\theta}_1}\right)^2, \qquad c_{12} = -\frac{m_1}{\hat{\theta}_1}, \qquad c_{22} = \frac{(1+m_2)}{\hat{\theta}_2^2},$$

$$a_{111} = \frac{\hat{\theta}_2^2(\hat{\theta}_2+3)}{\hat{\theta}_1^3}, \qquad a_{112} = \frac{-\{2\hat{\theta}_2 + (1+\hat{\theta}_2)m_1\}}{\hat{\theta}_1^2},$$

$$a_{122} = \frac{2m_1+m_2}{\hat{\theta}_1\hat{\theta}_2}, \qquad a_{222} = \frac{2-m_3}{\hat{\theta}_2^3},$$

where $m_j = n^{-1}\sum_{i=1}^{n} z_i (\log z_i)^j$ (j = 1, 2, 3) and $z_i = \left(\frac{x_i}{\hat{\theta}_1}\right)^{\hat{\theta}_2}$.

Derivation of expressions for the Weibull distribution:

$$\log f(x;\theta) = \log\theta_2 + (\theta_2 - 1)\log x - \theta_2 \log\theta_1 - \left(\frac{x}{\theta_1}\right)^{\theta_2} \tag{5.48}$$

$$l(\theta) = \log\theta_2 - \theta_2\log\theta_1 + (\theta_2 - 1)\frac{1}{n}\sum_{i=1}^{n}\log x_i - \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i}{\theta_1}\right)^{\theta_2}$$

First derivatives:

$$D_1 l(\theta) = -\frac{\theta_2}{\theta_1} + \frac{\theta_2}{\theta_1}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i}{\theta_1}\right)^{\theta_2}$$

$$D_2 l(\theta) = \frac{1}{\theta_2} - \log\theta_1 + \frac{1}{n}\sum_{i=1}^{n}x_i - \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i}{\theta_1}\right)^{\theta_2}\log\left(\frac{x_i}{\theta_1}\right)$$

Second derivatives:

$$D_1^2 l(\theta) = \frac{\theta_2}{\theta_1^2} - \frac{\theta_2}{\theta_1^2}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i}{\theta_1}\right)^{\theta_2} - \left(\frac{\theta_2}{\theta_1}\right)^2\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i}{\theta_1}\right)^{\theta_2}$$

$$D_2^2 l(\theta) = -\frac{1}{\theta_2^2} - \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i}{\theta_1}\right)^{\theta_2}\left[\log\left(\frac{x_i}{\theta_1}\right)\right]^2$$

$$D_1 D_2 l(\theta) = -\frac{1}{\theta_1} + \frac{1}{\theta_1}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i}{\theta_1}\right)^{\theta_2} + \frac{\theta_2}{\theta_1}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i}{\theta_1}\right)^{\theta_2}\log\left(\frac{x_i}{\theta_1}\right)$$

Let $z_i = \left(\dfrac{x_i}{\hat{\theta}_1}\right)^{\hat{\theta}_2}$

From $D_1 l(\theta)\big|_{\theta=\hat{\theta}} = 0$, we get $\bar{z} = 1$ where $\bar{z} = \dfrac{1}{n}\left(z_1 + z_2 + ... + z_n\right)$ and

$$m_j = n^{-1}\sum_{i=1}^{n} z_i\,(\log z_i)^j \;(j = 1, 2, 3) \tag{5.49}$$

$$\therefore c_{11} = -\frac{\hat{\theta}_2}{\hat{\theta}_1^2}(1-\bar{z}) - \left(\frac{\hat{\theta}_2}{\hat{\theta}_1}\right)^2\bar{z} = \left(\frac{\hat{\theta}_2}{\hat{\theta}_1}\right)^2$$

$$c_{12} = \frac{1}{\hat{\theta}_1}(1-\bar{z}) - \frac{\hat{\theta}_2}{\hat{\theta}_1}\frac{1}{n}\sum_{i=1}^{n}\frac{z_i\log z_i}{\hat{\theta}_2} = -\frac{m_1}{\hat{\theta}_1}$$

$$c_{22} = \frac{1}{\hat{\theta}_2^2} + \frac{1}{n}\sum_{i=1}^{n} z_i \left[\frac{\log z_i}{\hat{\theta}_2}\right]^2 = \frac{1+m_2}{\hat{\theta}_2^2}$$

$$\left(c_{11}c_{22} - c_{12}^2 = \frac{1+m_2}{\hat{\theta}_1^2} - \frac{m_1^2}{\hat{\theta}_2^2} > 0 \right)$$

Third derivatives:

$$D_1^3 l(\theta) = -\frac{2\theta_2}{\theta_1^3} + \frac{2\theta_2}{\theta_1^3}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i}{\theta_1}\right)^{\theta_2} + \frac{\theta_2^2}{\theta_1^3}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i}{\theta_1}\right)^{\theta_2} + \frac{2\theta_2^2}{\theta_1^3}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i}{\theta_1}\right)^{\theta_2}$$
$$+\left(\frac{\theta_2}{\theta_1}\right)^3 \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i}{\theta_1}\right)^{\theta_2}$$

$$D_1^2 D_2 l(\theta) = \frac{1}{\theta_1^2} - \frac{1}{\theta_1^2}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i}{\theta_1}\right)^{\theta_2} - \frac{\theta_2}{\theta_1^2}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i}{\theta_1}\right)^{\theta_2}\log\left(\frac{x_i}{\theta_1}\right) - \frac{2\theta_2}{\theta_1^2}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i}{\theta_1}\right)^{\theta_2}$$
$$-\left(\frac{\theta_2}{\theta_1}\right)^2 \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i}{\theta_1}\right)^{\theta_2}\log\left(\frac{x_i}{\theta_1}\right)$$

$$D_1 D_2^2 l(\theta) = \frac{1}{\theta_1}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i}{\theta_1}\right)^{\theta_2}\log\left(\frac{x_i}{\theta_1}\right) + \frac{1}{\theta_1}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i}{\theta_1}\right)^{\theta_2}\log\left(\frac{x_i}{\theta_1}\right)$$
$$+\frac{\theta_2}{\theta_1}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i}{\theta_1}\right)^{\theta_2}\left[\log\left(\frac{x_i}{\theta_1}\right)\right]^2$$

$$D_2^3 l(\theta) = \frac{2}{\theta_2^3} - \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i}{\theta_1}\right)^{\theta_2}\left[\log\left(\frac{x_i}{\theta_1}\right)\right]^3$$

From (5.49),

$$a_{111} = -\frac{2\hat{\theta}_2}{\hat{\theta}_1^3} + \frac{2\hat{\theta}_2}{\hat{\theta}_1^3}\bar{z} + \frac{\hat{\theta}_2^2}{\hat{\theta}_1^3}\bar{z} + \frac{2\hat{\theta}_2^2}{\hat{\theta}_1^3}\bar{z} + \left(\frac{\hat{\theta}_2}{\hat{\theta}_1}\right)^3\bar{z}$$
$$= \frac{\hat{\theta}_2^2(\hat{\theta}_2+3)}{\hat{\theta}_1^3}$$

$$a_{112} = \frac{1}{\hat{\theta}_1^2}(1-\bar{z}) - \frac{\hat{\theta}_2}{\hat{\theta}_1^2}\frac{1}{n}\sum_{i=1}^{n}\frac{z_i \log z_i}{\hat{\theta}_2} - \frac{2\hat{\theta}_2}{\hat{\theta}_1^2}\bar{z} - \left(\frac{\hat{\theta}_2}{\hat{\theta}_1}\right)^2 \frac{1}{n}\sum_{i=1}^{n}\frac{z_i \log z_i}{\hat{\theta}_2}$$

$$= -\frac{m_1}{\hat{\theta}_1^2} - \frac{\hat{\theta}_2 m_1}{\hat{\theta}_1^2} - \frac{2\hat{\theta}_2}{\hat{\theta}_1^2}$$

$$= \frac{-\left\{2\hat{\theta}_2 + (1+\hat{\theta}_2)m_1\right\}}{\hat{\theta}_1^2}$$

$$a_{122} = \frac{2}{\hat{\theta}_1}\frac{1}{n}\sum_{i=1}^{n}\frac{z_i \log z_i}{\hat{\theta}_2} + \frac{\hat{\theta}_2}{\hat{\theta}_1}\frac{1}{n}\sum_{i=1}^{n} z_i\left[\frac{\log z_i}{\hat{\theta}_2}\right]^2$$

$$= \frac{2m_1 + m_2}{\hat{\theta}_1\hat{\theta}_2}$$

$$a_{222} = \frac{2}{\hat{\theta}_2^3} - \frac{1}{n}\sum_{i=1}^{n} z_i\left[\frac{\log z_i}{\hat{\theta}_2}\right]^3$$

$$= \frac{2 - m_3}{\hat{\theta}_2^3}$$

From (5.48),

$$D_1 \log f(x;\theta) = -\frac{\theta_2}{\theta_1} + \frac{\theta_2}{\theta_1}\left(\frac{x}{\theta_1}\right)^{\theta_2}$$

$$D_2 \log f(x;\theta) = \frac{1}{\theta_2} - \log\theta_1 + \log x - \left(\frac{x}{\theta_1}\right)^{\theta_2}\log\left(\frac{x}{\theta_1}\right)$$

$$D_1^2 \log f(x;\theta) = -\frac{\theta_2}{\theta_1^2}\left[\theta_2\left(\frac{x}{\theta_1}\right)^{\theta_2} + \left(\frac{x}{\theta_1}\right)^{\theta_2} - 1\right]$$

$$D_1 D_2 \log f(x;\theta) = \frac{1}{\theta_1}\left[\left(\frac{x}{\theta_1}\right)^{\theta_2} - 1 + \theta_2\left(\frac{x}{\theta_1}\right)^{\theta_2}\log\left(\frac{x}{\theta_1}\right)\right]$$

$$D_2^2 \log f(x;\theta) = -\log\left(\frac{x}{\theta_1}\right)^2\left(\frac{x}{\theta_1}\right)^{\theta_2} - \frac{1}{\theta_2^2}$$

Fisher information matrix, $(I(\theta))_{su} = I_{su}$

The Fisher information matrix, where $s,u = 1,2$

$$\left.\begin{aligned} I_{ss} &= -E[(D_s^2 \log f(x;\theta) \mid \theta] && \text{for } s = u \\ I_{su} &= -E[(D_s D_u \log f(x;\theta) \mid \theta] && \text{for } s \neq u \end{aligned}\right\} \tag{5.50}$$

$I = (I_{su})$ is the Fisher information matrix. We write $I^{-1} = (I^{su})$

As a result,

$I_{11} \propto \left(\dfrac{\theta_2}{\theta_1}\right)^2$, $I_{12} \propto \dfrac{1}{\theta_1}$, $I_{21} = I_{12}$ and $I_{22} \propto \dfrac{1}{\theta_2^2}$ so that

$I^{11} \propto \left(\dfrac{\theta_1}{\theta_2}\right)^2$, $I^{12} (= I^{21}) \propto \theta_1$ and $I^{22} \propto \theta_2^2$

(III) The inverse Gaussian model

$$f(x;\theta) = \left(\frac{\theta_2}{2\pi x^3}\right)^{1/2} \exp\left\{-\frac{\theta_2}{2x}\left(\frac{x}{\theta_1} - 1\right)^2\right\}, \qquad x > 0,$$

where $\phi(.)$ is as usual the standard univariate normal density, $\theta = (\theta_1, \theta_2)'$ and $\theta_1, \theta_2 > 0$. Here p = 2. $\hat{\theta}_1 = \overline{X}$, $\hat{\theta}_2 = (\overline{X}_{\text{har}} \overline{X}) / (\overline{X} - \overline{X}_{\text{har}})$, where $\overline{X}$ and $\overline{X}_{\text{har}}$ are the arithmetic and harmonic means, respectively, of $X_1, \ldots, X_n$.

It can be seen that:

$$c_{11} = \frac{\hat{\theta}_2}{\hat{\theta}_1^3}, \qquad c_{12} = 0, \qquad c_{22} = \frac{1}{2\hat{\theta}_2^2},$$

$$a_{111} = \frac{6\hat{\theta}_2}{\hat{\theta}_1^4}, \qquad a_{112} = -\frac{1}{\hat{\theta}_1^3}, \qquad a_{122} = 0, \qquad a_{222} = \frac{1}{\hat{\theta}_2^3},$$

$$\hat{I}^{11} = \frac{\hat{\theta}_1^3}{\hat{\theta}_2}, \qquad \hat{I}_1^{11} = \frac{3\hat{\theta}_1^2}{\hat{\theta}_2}, \qquad \hat{I}_2^{11} = -\frac{\hat{\theta}_1^3}{\hat{\theta}_2^2},$$

$$\hat{I}^{12} = \hat{I}_1^{12} = \hat{I}_2^{12} = 0$$

$$\hat{I}^{22} = 2\hat{\theta}_2^2, \qquad \hat{I}_1^{22} = 0, \qquad \hat{I}_2^{22} = 4\hat{\theta}_2$$

The expression (5.40) for L_{1f} under the inverse Gaussian model is simplified to some extent as $\hat{I}^{12} = \hat{I}_1^{12} = \hat{I}_2^{12} = 0$.

Derivation of expressions for c_{su} and a_{suw} are as follows:

$$\log f(x;\theta) = \text{constant} + \frac{1}{2}\log\theta_2 - \frac{\theta_2}{2}\left(\frac{x}{\theta_1^2} - \frac{2}{\theta_1} + \frac{1}{x}\right) \tag{5.51}$$

$$l(\theta) = \text{constant} + \frac{1}{2}\log\theta_2 - \frac{\theta_2}{2}\left(\frac{\overline{x}}{\theta_1^2} - \frac{2}{\theta_1} + \frac{1}{\overline{x}_I}\right)$$

where $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ (arithmetic mean), $\overline{x}_H = \frac{1}{\frac{1}{n}\sum_{i=1}^{n} x_i}$ (harmonic mean)

First derivatives:

$$D_1 l(\theta) = -\frac{\theta_2}{\theta_1^3}(\theta_1 - \overline{x})$$

When $D_1 l(\theta) = 0$, $\hat{\theta}_1 = \overline{x}$.

$$D_2 l(\theta) = \frac{1}{2\theta_2} - \frac{\theta_1^2}{2\theta_1^2 \overline{x}_H} - \frac{1}{\theta_1^2} + \overline{x}$$

When $D_2 l(\theta) = 0$, $\hat{\theta}_2 = \left(\frac{1}{\overline{x}_H} - \frac{1}{\overline{x}}\right)$.

Second derivatives:

$D_1^2 l(\theta) = \theta_2 \left(\frac{2}{\theta_1^3} - \frac{3\bar{x}}{\theta_1^4} \right)$ and thus, $c_{11} = \frac{\hat{\theta}_2}{\hat{\theta}_1^3} \quad (\because \hat{\theta}_1 = \bar{x})$.

$D_1 D_2 l(\theta) = -\frac{1}{2} \left(\frac{-2\bar{x}}{\theta_1^3} + \frac{2}{\theta_1^2} \right)$ and $c_{12} = 0$

Thus, $c_{21} = c_{12} = 0$.

$D_2^2 l(\theta) = -\frac{1}{2\theta_2^2}$. Therefore, $c_{22} = \frac{1}{2\hat{\theta}_2^2}$.

Third derivatives:

$D_1^3 l(\theta) = \theta_2 \left(-\frac{6}{\theta_1^4} + \frac{12\bar{x}}{\theta_1^5} \right)$ leads to $a_{111} = \frac{6\hat{\theta}_2}{\hat{\theta}_1^4}$.

$D_1^2 D_2 l(\theta) = \frac{2}{\theta_1^3} - \frac{3\bar{x}}{\theta_1^4}$ and thus, $a_{112} = -\frac{1}{\hat{\theta}_1^3}$

$\therefore a_{112} = a_{121} = a_{211}$

$D_1 D_2^2 l(\theta) = 0$

$a_{122} = 0$

$a_{122} = a_{212} = a_{221} = 0$

$D_2^3 l(\theta) = \frac{1}{\theta_2^3}$. This results in $a_{222} = \frac{1}{\hat{\theta}_2^3}$.

Information matrix

$$D_1 \log f(x;\theta) = -\frac{\theta_2}{\theta_1^2} + \frac{\theta_2 x}{\theta_1^3}$$

$$D_2 \log f(x;\theta) = \frac{1}{2\theta_2} - \frac{1}{2}\left(\frac{x}{\theta_1^2} - \frac{2}{\theta_1} + \frac{1}{x}\right)$$

$$D_1^2 \log f(x;\theta) = -\frac{3\theta_2 x}{\theta_1^4} + \frac{2\theta_2}{\theta_1^3}$$

$$D_1 D_2 \log f(x;\theta) = -\frac{1}{\theta_1^2} + \frac{x}{\theta_1^3}$$

$$D_2 D_1 \log f(x;\theta) = D_1 D_2 \log f(x;\theta)$$

$$D_2^2 \log f(x;\theta) = -\frac{1}{2\theta_2^2}$$

Using (5.50), the Fisher information matrix,

$$I_{11} = -E\left(-\frac{3\theta_2 x}{\theta_1^4} + \frac{2\theta_2}{\theta_1^3}\right)$$

$$= \frac{3\theta_2 E_\theta(x)}{\theta_1^4} - \frac{2\theta_2}{\theta_1^3}$$

Since $E_\theta(x) = \theta_1$,

$$I_{11} = \frac{\theta_2}{\theta_1^3}$$

$$\therefore \hat{I}_{11} = \frac{\hat{\theta}_2}{\hat{\theta}_1^3}$$

Similarly using (5.50), we obtain $\hat{I}_{22} = \dfrac{1}{2\hat{\theta}_2^2}$ and $\hat{I}_{12} = \hat{I}_{21} = 0$.

Therefore, as mentioned earlier, the fact that $\hat{I}^{12} = \hat{I}_1^{12} = \hat{I}_2^{12} = 0$ simplifies, to some extent expression (5.40) for L_{1f} under the inverse Gaussian model.

5.5.1 Simulation study

We carried out a simulation study to examine the finite sample implications of our results by taking $\beta_1 = 0.05$, $\beta_2 = 0.05$, i.e., $\beta = 0.90$, and $\gamma = 0.90$ and 0.95. The following tolerance intervals were studied in our simulation study:

(i) The Bayesian-cum-frequentist interval $[d(\hat{\theta}) - g^{(1n)}, b(\hat{\theta}) + g^{(1n)}]$ for the univariate normal model with both mean and variance θ, under the matching prior $\pi_0(\theta) \propto \theta^{-1}\sqrt{2\theta+1}$.

(ii) The Bayesian-cum-frequentist interval $[d(\hat{\theta}) - g^{(2n)}, b(\hat{\theta}) + g^{(2n)}]$ for the Weibull model under the matching prior $\pi_0(\theta) \propto (\theta_1\theta_2)^{-1}$.

(iii) The purely frequentist interval $[d(\hat{\theta}) - g_f^{(3n)}, b(\hat{\theta}) + g_f^{(3n)}]$ for the inverse Gaussian model.

(iv) The Bayesian tolerance interval $[d(\hat{\theta}) - g^{(3n)}, b(\hat{\theta}) + g^{(3n)}]$ for the inverse Gaussian model using the highest posterior density regions prior $\pi(\theta) \propto (\theta_1^2\theta_2)^{-1}$ (see pp. 72, Datta and Mukerjee, 2004).

Simulation studies showed that the best choice of $g^{(n)}$ to ensure the fastest convergence to the nominal value of the confidence levels for the normal model is $g^{(1n)}$ and Weibull model is $g^{(2n)}$. For the inverse Gaussian model, $g_f^{(n)} = g_f^{(3n)}$ (refer Tables B1 to B18 in the Appendix for results). We have also taken the expected width and expected content into consideration to obtain the tolerance interval. The Bayesian tolerance interval $[d(\hat{\theta}) - g^{(3n)}, b(\hat{\theta}) + g^{(3n)}]$ for the inverse Gaussian model was computed for the purpose of comparing its performance with the frequentist case (see Equation (5.31)). We shall use the prior $\pi(\theta) \propto (\theta_1^2\theta_2)^{-1}$ (see pp. 72, Datta and

Mukerjee, 2004). This prior does not enjoy the matching property in our context but it does so for highest posterior density regions for θ (see Table B19 and Table B20 in Appendix B).

The expected content, confidence levels and expected widths were computed based on 10000 simulation runs. For each simulated interval, we calculated the content as $F(U;\theta) - F(L;\theta)$ where U and L are respectively the upper and lower limits of the interval and $F(x;\theta)$ is the cumulative distribution function (cdf) where θ is a possibly vector valued unknown parameter. The confidence level is the proportion of time the content of the simulated tolerance intervals was at least β. As mentioned earlier, the intervals in (i)-(iv) resulting from higher order asymptotic considerations, have $\beta-$content with frequentist confidence level $\gamma + O(n^{-1})$. We also include the simulated coverage probabilities of the naive interval $[d(\hat{\theta}) - n^{-1/2}g_1, b(\hat{\theta}) + n^{-1/2}g_1]$, where g_1 = $Mz_\gamma /(f^b + f^d)$ as in Theorem 5.1 or 5.3 for comparative purposes. It is very obvious that this naive interval, based on simpler asymptotics, has $\beta-$content with frequentist confidence level $\gamma + O(n^{-1/2})$ rather than $\gamma + O(n^{-1})$.

Note: The top entry of Table 5.1 to Table 5.3 shows the naive tolerance interval while the bottom entry shows the higher order asymptotic tolerance interval.

Table 5.1: *Simulated coverage probabilities*: *univariate normal model with mean = variance = θ*

	$\gamma = 0.90$					$\gamma = 0.95$				
	Sample size					Sample size				
θ	15	20	25	30	50	15	20	25	30	50
4	0.795	0.806	0.817	0.822	0.834	0.839	0.853	0.860	0.867	0.882
	0.888	0.889	0.892	0.893	0.894	0.942	0.945	0.946	0.948	0.948
8	0.767	0.777	0.793	0.803	0.813	0.813	0.829	0.842	0.859	0.865
	0.895	0.895	0.890	0.901	0.896	0.944	0.947	0.945	0.947	0.949
12	0.749	0.760	0.775	0.790	0.805	0.791	0.812	0.829	0.834	0.850
	0.890	0.897	0.899	0.899	0.899	0.944	0.948	0.950	0.946	0.947
16	0.720	0.736	0.766	0.774	0.790	0.767	0.797	0.808	0.820	0.843
	0.887	0.890	0.891	0.891	0.896	0.940	0.950	0.944	0.946	0.945

Table 5.2: *Simulated coverage probabilities*: *Weibull model.*

	$\gamma = 0.90$					$\gamma = 0.95$				
	Sample size					Sample size				
(θ_1,θ_2)	15	20	25	30	50	15	20	25	30	50
(1,2)	0.717	0.749	0.770	0.777	0.816	0.791	0.827	0.845	0.855	0.882
	0.878	0.882	0.884	0.892	0.894	0.916	0.927	0.936	0.937	0.940
(5,5)	0.709	0.740	0.757	0.774	0.810	0.787	0.818	0.834	0.850	0.878
	0.889	0.890	0.893	0.895	0.899	0.931	0.939	0.941	0.940	0.942
(10,3)	0.726	0.749	0.764	0.787	0.813	0.797	0.816	0.834	0.853	0.878
	0.886	0.890	0.893	0.894	0.899	0.934	0.939	0.942	0.941	0.942
(15,6)	0.708	0.736	0.762	0.777	0.804	0.795	0.813	0.827	0.843	0.876
	0.887	0.892	0.890	0.895	0.896	0.934	0.933	0.941	0.943	0.947

Table 5.3: *Simulated coverage probabilities*: *inverse Gaussian model.*

	$\gamma = 0.90$					$\gamma = 0.95$				
	Sample size					Sample size				
(θ_1,θ_2)	15	20	25	30	50	15	20	25	30	50
(7,14)	0.768	0.796	0.812	0.823	0.864	0.693	0.726	0.759	0.765	0.805
	0.873	0.883	0.885	0.893	0.901	0.891	0.907	0.918	0.923	0.937
(8,12)	0.755	0.791	0.815	0.827	0.858	0.694	0.735	0.753	0.765	0.798
	0.857	0.871	0.883	0.886	0.896	0.869	0.889	0.909	0.915	0.926
(15,25)	0.755	0.789	0.812	0.828	0.866	0.703	0.725	0.750	0.768	0.780
	0.865	0.871	0.883	0.889	0.897	0.881	0.895	0.908	0.917	0.933
(20,50)	0.771	0.792	0.817	0.833	0.867	0.695	0.732	0.754	0.766	0.805
	0.895	0.892	0.895	0.895	0.899	0.930	0.920	0.925	0.930	0.940

Table 5.4: *Simulated coverage probabilities for the higher order asymptotic Bayesian tolerance interval: inverse Gaussian model.*

	$\gamma = 0.90$					$\gamma = 0.95$				
	Sample size					Sample size				
(θ_1,θ_2)	15	20	25	30	50	15	20	25	30	50
(7,14)	0.876	0.887	0.886	0.893	0.893	0.893	0.908	0.915	0.924	0.939
(8,12)	0.861	0.865	0.880	0.887	0.887	0.875	0.896	0.898	0.912	0.935
(15,25)	0.858	0.874	0.879	0.889	0.890	0.882	0.896	0.903	0.919	0.934
(20,50)	0.887	0.891	0.892	0.896	0.899	0.911	0.923	0.927	0.930	0.941

For the for the univariate normal model with both mean and variance θ, the convergence of the simulated frequentist coverage probability to the nominal value γ is quite rapid as shown in Table 5.1. It is also reasonably fast for both $\gamma =$ 0.90 and 0.95 in Table 5.2 and for $\gamma =$ 0.90 in Table 5.3, though slightly slow for $\gamma =$ 0.95 in Table 5.3. The outputs in our tables show that our higher order asymptotic results are well reflected in finite samples. It is interesting to see that, despite not working with a matching prior for the two-sided tolerance intervals, the Bayesian interval for the inverse Gaussian model comes quite close to the frequentist case for both $\gamma =$ 0.90 and 0.95 (see Table 5.4). Based on the results, the convergence to the target for the naïve interval is much slower. Hence, considering higher order asymptotics as shown here results in significant gains.

5.5.2 Application to real data

Data set 1: The data from (Gacula and Kubala, 1975) represent shelf life (in days) of a food product.

24	24	26	26	32	32	33	33	33	35	41	42	43
47	48	48	48	50	52	54	55	57	57	57	57	61

As mentioned by Gacula and Kubala (1975), the Weibull model fits the data well; see also Chhikara and Folks (1989). Our results are applied to this data set under the framework of the Weibull model. The prior $\pi_0(\theta) \propto (\theta_1\theta_2)^{-1}$ meets the matching condition in (5.25). As a result, the two-sided Bayesian tolerance interval in Theorem 5.1, obtained based on this prior, is also frequentist. We shall calculate this interval by choosing $g^{(n)} = g^{(2n)}$ (refer to Equation (5.31)). The Bayesian cum frequentist tolerance interval reported here is $[d(\hat{\theta}) - g^{(2n)}, b(\hat{\theta}) + g^{(2n)}]$.

Here, the sample size, $n = 26$. We take $\beta_1 =$0.05, $\beta_2 =$ 0.05, i.e., $\beta =$ 0.90, and $\gamma =$ 0.90 and 0.95. For the present data set under the Weibull model, we obtain , $\hat{\theta}_1 =$ 47.2816 and $\hat{\theta}_2 =$ 4.3329, so that $b(\hat{\theta}) =$ 60.9067 and $d(\hat{\theta}) = 23.8223$. Then, applying Equation (5.13) as well as the facts shown in (II) for the Weibull model, we get:

$M = 0.2425,\ L_1(\pi_0) = 0.7191,\ L_2 = 2.0224,\ L_3 = 0.8436,\ L_4 = -0.0219,$

upon symbolic computation of the required partial derivatives of $F(x;\theta)$ and $f(x;\theta)$. Thus, using Theorem 5.1, for $\gamma = 0.90$ and 0.95, the pair (g_1, g_2) and the corresponding Bayesian-cum-frequentist tolerance interval as specified above turns out to be as follows:

$\gamma =$ 0.90: $(g_1, g_2) =$ (15.9195, 35.2285), tolerance interval = [19.0037, 65.7253].

$\gamma =$ 0.95: $(g_1, g_2) =$ (20.4324, 42.7722), tolerance interval = [17.7811, 66.9480].

The two-sided Bayesian tolerance interval in Theorem 5.1, obtained based on this prior, is also frequentist. Based on the findings for $\gamma =$ 0.95, a frequentist may assert with about 0.95 confidence that at least 90% of the food product lasts between 17.7811 and 66.9480 days while a Bayesian may conclude with about 0.95 credibility that at least at least 90% of the food product lasts between 17.7811 and 66.9480 days. We interpret the results for $\gamma =$ 0.90 similarly.

Data set 2: The following data originally from Lieblin and Zelen (1956) represents the number of million revolutions before failure for each of 23 ball bearings.

17.88	28.92	33.00	41.52	42.12	45.60	48.48	51.84	51.96	54.12
55.56	67.80	68.64	68.64	68.88	84.12	93.12	98.64	105.12	105.84
127.92	128.04	173.40							

We apply our results to this data set under the framework of the inverse Gaussian model since it is mentioned in Chhikara and Folks (1989) that this model fits the data well. Since it is difficult to obtain a solution to the matching condition (5.25), we obtain a purely frequentist two-sided tolerance interval $[d(\hat{\theta})-g_f^{(n)}, b(\hat{\theta})+g_f^{(n)}]$ as given by Theorem 5.3 by choosing $g_f^{(n)}=g_f^{(3n)}$ (refer to Equation (5.46)) a choice which the simulation studies prove to work well for the inverse Gaussian model. We have also reported the Bayesian tolerance interval $[d(\hat{\theta})-g^{(3n)}, b(\hat{\theta})+g^{(3n)}]$ using the prior $\pi(\theta)\propto(\theta_1^2\theta_2)^{-1}$.

Here, the sample size, $n=23$. We take $\beta_1=0.05$, $\beta_2=$ 0.05, i.e., $\beta=$ 0.90, and $\gamma=$ 0.90 and 0.95. For the present data set under the inverse Gaussian model, $\hat{\theta}_1=$ 72.2243 and $\hat{\theta}_2=$ 231.6741, so that $b(\hat{\theta})=150.1856$ and $d(\hat{\theta})=$ 26.9034. Therefore, using (5.13), (5.40) and the facts noted in (III) above, we get:

$M=$ 0.2397, $L_{1f}=$ 1.0385, $L_1(\pi)=0.9493$,

$L_2=1.7643$, $L_3=$ 0.8377, $L_4=-$ 0.0098,

upon symbolic computation of the required partial derivatives of $F(x;\theta)$ and $f(x;\theta)$. Thus, applying Theorems 5.3 and 5.1, for $\gamma=0.90$ and 0.95, the pairs (g_1, g_{2f}) and (g_1, g_2), and the corresponding frequentist and Bayesian tolerance intervals as mentioned above are as follows:

$\gamma=0.90$: $(g_1, g_{2f})=$ (32.9318, 75.2455), $(g_1, g_2)=$ (32.9318, 72.9541),

Frequentist tolerance interval = [13.7880, 163.3009],

Bayesian tolerance interval = [14.1417, 162.9473].

$\gamma=$ 0.95: $(g_1, g_{2f})=$ (42.2675, 91.2664), $(g_1, g_2)=$(42.2675, 88.9750),

Frequentist tolerance interval = [10.8721, 166.2168],

Bayesian tolerance interval = [11.1951, 165.8938].

Hence, on the basis of the findings for $\gamma = 0.95$, a frequentist may claim with 0.95 confidence that at least 90% of the ball bearings went through between 10.8721 and 166.2168 million revolutions before failure. In the Bayesian viewpoint, with about 0.95 credibility, at least 90% of the balls bearings underwent between 11.1951 and 165.8938 million revolutions before failing. We shall interpret the results for $\gamma = 0.90$ in the similar manner. As shown in the simulation results earlier, the Bayesian interval here comes quite close to the frequentist case for both $\gamma = 0.90$ and 0.95.

CHAPTER 6

CONCLUDING REMARKS AND FUTURE RESEARCH

6.1 Concluding remarks

In Chapter 3 of this thesis, we studied the two-sided Bayesian tolerance interval with approximate frequentist validity by Ong and Mukerjee (2011), in balanced one-way random effects model. This tolerance interval uses probability matching priors (PMP) and thus we refer to it as the PMP tolerance interval here. Studies by Ong and Mukerjee (2011) reveal that the modified Jeffreys' prior by Wolfinger (1998) is not probability matching but it comes quite close to being so. The simplicity of this prior makes a strong case in its favour if one is not too particular about the probability matching criteria. It is recommended to use the relatively more complex matching prior given in Equation (3.46) if one considers the probability matching criteria a must. (Ong and Mukerjee, 2011)

As discussed in Chapter 4, we examined the results in Chapter 3 for two-sided Bayesian tolerance intervals in the balanced one-way random effects model derived via probability matching priors. The effect of non-normal experimental errors on these intervals is examined by simulation. We also did a comparative study with the MLS tolerance intervals proposed by Krishnamoorthy and Mathew (2009) and studied by Krishnamoorthy and Lian (2012). We applied these aforesaid tolerance intervals for all cases as if the assumptions where all underlying distributions are normal are justified even though the data comes from other distributions. We also note that the probability matching prior for the normal case is used for the PMP tolerance interval for all instances. For normal error distributions where moderate intra-class correlation coefficient, ρ is concerned, the PMP tolerance intervals are slightly less conservative compared to the MLS case. The outputs appear to be more accurate for larger intra-class correlation coefficients ρ. Both PMP and MLS tolerance intervals are comparable for non-normal error distributions when all the underlying distributions for the tolerance intervals are assumed to be normal. The MLS tolerance interval has confidence level close to the nominal value for smaller sample sizes and is comparable to the PMP case for large number of classes. We recommend the MLS tolerance interval for smaller sample sizes (less than 50). The PMP and MLS intervals can be used for sample sizes larger than 50 by noting that the results for the PMP tolerance intervals are heavily dependent on the balance in the classes. The PMP interval works well when the ratio $n:t$ is approximately 12.5:1 to 25:1. This criterion is important because it ensures the consistency of the maximum likelihood estimator. Expanding the margin of error of the frequentist confidence helps to achieve improvement in the PMP tolerance interval for smaller sample sizes. However, the underlying algebra will be extremely tedious and difficult.

In Chapter 5, we derived asymptotic results leading to explicit formulae for two-sided Bayesian and frequentist tolerance intervals in a general framework of parametric models. We also identified the probability matching priors for such intervals and studied their roles in determining frequentist tolerance intervals via the Bayesian route. For instances when the solution to the matching condition given in Equation (5.25) is difficult to be obtained such as for the inverse Gaussian model, the purely frequentist tolerance interval was employed and interestingly Bayesian arguments were very helpful. The convergence of the simulated frequentist confidence level to the nominal value is quite fast for the univariate normal model, reasonably fast for the Weibull model and slightly slower for the inverse Gaussian model. Interestingly, although we did not work with the matching prior for the inverse Gaussian case, the Bayesian tolerance interval here produces comparable results with its purely frequentist counterpart. Our higher order asymptotic results are well reflected in finite samples. Hence, consideration of higher order asymptotic as studied here entails significant gains.

6.2 Future research

We discussed the two-sided Bayesian tolerance interval with approximate frequentist validity for the balanced one-way random effects model. We hope to extend this study for models with smaller sample sizes. Currently, the only method of doing so is by extending the margin of error of the frequentist confidence level. This process can be excruciating. We intend to explore other priors which can help us to make the higher-order asymptotic tolerance intervals favourable for finite sample sizes. It is of interest to strengthen this work in terms of robustness study of the tolerance intervals since no such study has been done.

Ong and Mukerjee (2011) also discussed the two-way nested random effects model. We would also like to study the effects of non-normality on these probability matching tolerance intervals. We hope to extend the results in Chapter 3 to find the probability matching tolerance intervals for the one-way random model with unbalanced data in future works as it is a difficult task. Some of the satisfactory tolerance intervals for this case are reported in book by Krishnamoorthy and Mathew (2009).

Gupta and Kundu (1999) introduced the generalized exponential distribution which can be used quite effectively in analyzing lifetime data. This distribution is definitely a good alternative in replacing the gamma and Weibull distribution. According to Gupta and Kundu (1999), the gamma distribution has its drawback where the distribution function or the shape function is difficult to be computed if the shape parameter is not an integer. It is of great interest to compute the two-sided tolerance intervals for the generalized exponential model as it will be a good alternative to both gamma and Weibull distributions in life data analysis. We also intend to work with other distributions such as the Rayleigh, Pareto etc.

Currently, many resort to combining two one-sided tolerance intervals via Bonferonni's inequality to obtain approximate two-sided tolerance intervals. However, this approach is rather conservative and the tolerance intervals computed are unduly long. We hope that our work will significantly contribute towards computation of tolerance intervals since computing two-sided tolerance intervals is very challenging.

REFERENCES

Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics, 12*(2), 171-178.

Bain, L. J., & Engelhardt, M. (1981). Simple approximate distributional results for confidence and tolerance limits for the Weibull distribution based on maximum-likelihood estimators. *Technometrics, 23*(1), 15-20.

Beckman, R. J., & Tietjen, G. L. (1989). Two-sided tolerance limits for balanced random-effects ANOVA models. *Technometrics, 31*(2), 185-197. (Correction 1998, *40,* p. 269).

Bhattacharya, R. N., & Ghosh, J. K. (1978). On the validity of the formal Edgeworth expansion. *The Annals of Statistics, 6*(2), 434-451.

Bhaumik, D. K., & Kulkarni, P. M. (1996). A simple and exact method of constructing tolerance intervals for the one-way ANOVA with random effects. *American Statistician, 50*(4), 319-323.

Bickel, P. J., & Ghosh, J. K. (1990). A decomposition for the likelihood ratio statistic and the Bartlett correction - a Bayesian argument. *Annals of Statistics, 18*(3), 1070-1090.

Box, G. E., & Tiao, G. C. (1973). *Bayesian inference in statistical analysis*. Reading, MA: Addison-Wesley.

Burrows, G. L. (1963). Statistical tolerance limits-what are they? *Journal of the Royal Statistical Society: Series C (Applied Statistics), 12*(2), 133-144.

Cai, T. T., & Wang, H. (2009). Tolerance intervals for discrete distributions in exponential families. *Statistica Sinica, 19*(3), 905-923.

Chen, S. Y., & Harris, B. (2006). On lower tolerance limits with accurate coverage probabilities for the normal random effects model. *Journal of the American Statistical Association, 101*(475), 1039-1049.

Chhikara, R. S., & Folks, J. L. (1989). *The inverse Gaussian distribution: Theory, methodology, and applications*. New York, NY: Marcel Dekker, Inc.

Datta, G. S., Ghosh, M., & Mukerjee, R. (2000). Some new results on probability matching priors. *Calcutta Statistical Association Bulletin, 50*, 179-192.

Datta, G. S., & Mukerjee, R. (2004). *Probability matching priors*: *Higher order asymptotics*. New York, NY: Springer.

Datta, G. S., & Sweeting, T. J. (2005). Probability matching priors. *Handbook of statistics, 25*, 91–114.

Dawid, A. P. (1991). Fisherian inference in likelihood and prequential frames of reference. *Journal of the Royal Statistical Society: Series B (Methodological), 53*(1), 79-109.

Easterling, R. G., & Weeks, D. L. (1970). An accuracy criterion for Bayesian tolerance intervals. *Journal of the Royal Statistical Society: Series B (Methodological)*, *32*(2), 236-240.

Engelhardt, M., & Bain, L. J. (1977). Simplified statistical procedures for the Weibull or extreme-value distribution. *Technometrics, 19*(3), 323-331.

Fertig, K. W., & Mann, N. R. (1974). Population percentile estimation for the unbalanced random-effect model: Estimation of the strength of a material when sampling from various batches of the material. *Reliability and Biometry: Statistical Analysis of Lifelength,* 205-228.

Fisher, R. A. (1956). *Statistical methods and scientific inference.* Edinburgh, Scotland: Oliver & Boyd.

Gacula, M., & Kubala, J. (1975). Statistical models for shelf life failures. *Journal of Food Science, 40*(2), 404-409.

Ghosh, J.K., & Mukerjee, R. (1991). Characterization of priors under which Bayesian and frequentist Bartlett corrections are equivalent in the multiparameter case. *Journal of Multivariate Analysis, 38*(2), 385-393.

Ghosh, M., & Mukerjee, R. (1998). Recent developments on probability matching priors. *Applied Statistical Science III,* 227-252.

Ghosh, M., Santra, U., & Kim, D. (2008). Probability matching priors for some parameters of the bivariate normal distribution. *Pushing the Limits of Contemporary Statistics Contributions in Honor of Jayanta Kumar Ghosh, 3*, 71-81.

Graybill, F. A., & Wang, C. M. (1980). Confidence intervals on nonnegative linear combinations of variances. *Journal of the American Statistical Association, 75*(372), 869-873.

Guenther, W. C. (1972). Tolerance intervals for univariate distributions. *Naval Research Logistics Quarterly, 19*(2), 309-333.

Gupta, R. D., & Kundu, D. (1999). Generalized exponential distributions. *Australian & New Zealand Journal of Statistics*, *41*(2), 173-188.

Guttman, I. (1970). *Statistical tolerance regions: Classical and Bayesian* (Vol. 26): London, England: Charles Griffin and Company.

Hahn, G. J., & Chandra, R. (1981). Tolerance intervals for Poisson and binomial variables. *Journal of Quality Technology, 13*(2), 100-110.

Hoffman, D., & Kringle, R. (2005). Two-sided tolerance intervals for balanced and unbalanced random effects models. *Journal of Biopharmaceutical Statistics, 15*(2), 283-293.

Hugo, J. (2012). *Bayesian tolerance intervals for variance component models*. (Unpublished Doctoral Dissertation). University of the Free State, South Africa.

Jeffreys, H. (1946). An invariant form for the prior probability in estimation problems. *Proceedings of the Royal Society of London, Series A (Mathematical and Physical Sciences), 186*(1007), 453-461.

Jilek, M. (1981). A bibliography of statistical tolerance regions. *Statistics, 12*(3), 441-456.

Jílek, M., & Ackermann, H. (1989). A bibliography of statistical tolerance regions, II. *Statistics, 20*(1), 165-172.

Johnson, R. A. (1970). Asymptotic expansions associated with posterior distributions. *The Annals of Mathematical Statistics, 41*(3), 851-864.

Karian, Z. A., & Dudewicz, E. J. (2000). *Fitting statistical distributions: The generalized lambda distribution and generalized bootstrap methods*. Florida, FL: CRC Press.

Kass, R. E., & Wasserman, L. (1996). The selection of prior distributions by formal rules. *Journal of the American Statistical Association, 91*(435), 1343-1370.

Krishnamoorthy, K., & Lian, X. D. (2012). Closed-form approximate tolerance intervals for some general linear models and comparison studies. *Journal of Statistical Computation and Simulation, 82*(4), 547-563.

Krishnamoorthy, K., & Mathew, T. (2004). One-sided tolerance limits in balanced and unbalanced one-way random models based on generalized confidence intervals. *Technometrics, 46*(1), 44-52.

Krishnamoorthy, K., & Mathew, T. (2009). *Statistical Tolerance Regions: Theory, Applications and Computation.* Hoboken, NJ: Wiley.

Krishnamoorthy, K., & Xie, F. (2011). Tolerance intervals for symmetric location-scale families based on censored and uncensored data. *Journal of Statistical Planning and Inference, 141*(3), 1170-1182.

Lemon, G. H. (1977). Factors for one-sided tolerance limits for balanced one-way-ANOVA random-effects model. *Journal of the American Statistical Association, 72*(359), 676-680.

Li, J. (1998). *Some aspects of Bayesian and frequentist asymptotics. (Unpublished PhD thesis*). Ontario: University of Toronto.

Liao, C. T., Lin, T. Y., & Iyer, H. K. (2005). One- and two-sided tolerance intervals for general balanced mixed models and unbalanced one-way random models. *Technometrics, 47*(3), 323-335.

Lieblin, J., & Zelen, M. (1956). Statistical investigation of the fatigue life of deep-groove ball bearings. *Journal of the Research of the National Bureau of Standards, 57*(5), 273-316.

Lindley, D. V. (1958). Fiducial distributions and Bayes' theorem. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 20*(1), 102-107.

Mann, N. R. (1978). Weibull tolerance intervals associated with moderate to small survival proportions for use in a new formulation of lanchester combat theory. *Naval Research Logistics Quarterly, 25*(1), 121-128.

Mann, N. R., & Fertig, K. W. (1975). Simplified efficient point and interval estimators for Weibull parameters. *Technometrics, 17*(3), 361-368.

Mann, N. R., & Fertig, K. W. (1977). Efficient unbiased quantile estimators for moderate-size complete samples from extreme-value and Weibull distributions: Confidence bounds and tolerance and prediction intervals. *Technometrics, 19*(1), 87-93.

Mee, R. W. (1984). β-Expectation and β -Content tolerance limits for balanced one-way ANOVA random model. *Technometrics, 26*(3), 251-254.

Mee, R. W., & Owen, D. B. (1983). Improved factors for one-sided tolerance limits for balanced one-way ANOVA random model. *Journal of the American Statistical Association, 78*(384), 901-905.

Mukerjee, R., & Dey, D. K. (1993). Frequentist validity of posterior quantiles in the presence of a nuisance parameter - higher-order asymptotics. *Biometrika, 80*(3), 499-505.

Mukerjee, R., & Ghosh, M. (1997). Second-order probability matching priors. *Biometrika, 84*(4), 970-975.

Mukerjee, R., & Reid, N. (1999). On a property of probability matching priors: Matching the alternative coverage probabilities. *Biometrika, 86*(2), 333-340.

Mukerjee, R., & Reid, N. (2000). On the Bayesian approach for frequentist computations. *Brazilian Journal of Probability and Statistics, 14*(2), 159-166.

Mukerjee, R., & Reid, N. (2001). Second-order probability matching priors for a parametric function with application to Bayesian tolerance limits. *Biometrika, 88*(2), 587-592.

Nason, G. P. (2006). On the sum of t and Gaussian random variables. *Statistics & Probability Letters, 76*(12), 1280-1286.

Ong, S. H., & Mukerjee, R. (2010). Data-dependent probability matching priors of the second order. *Statistics, 44*(3), 291-302.

Ong, S. H., & Mukerjee, R. (2011). Probability matching priors for two-sided tolerance intervals in balanced one-way and two-way nested random effects models. *Statistics, 45*(4), 403-411.

Patel, J. K. (1986). Tolerance Limits - a Review. *Communications in Statistics-Theory and Methods, 15*(9), 2719-2762.

Pathmanathan, D., & Ong, S. H. (2013). A Monte Carlo simulation study of two-sided tolerance intervals in balanced one-way random effects model for non-normal errors. *Journal of Statistical Computation and Simulation*, (ahead-of-print), 1-16.

Pathmanathan, D., Mukerjee, R., & Ong, S. H. (2013). Two-sided Bayesian and frequentist tolerance intervals: general asymptotic results with applications. *Statistics*, (ahead-of-print), 1-15.

Rao, C. R. (1973). *Linear statistical inference and its applications* (2nd ed.). New York, NY: Wiley.

Rebafka, T., Clemencon, S., & Feinberg, M. (2007). Bootstrap-based tolerance intervals for application to method validation. *Chemometrics and Intelligent Laboratory Systems, 89*(2), 69-81.

Robert, C. (2007). *The Bayesian choice: From decision-theoretic foundations to computational implementation.* New York, NY: Springer.

Sahai, H., & Ojeda, M. (2004). *Analysis of variance for random models: Balanced data theory, methods, applications and data analysis.* (Vol. 1). Boston, MA: Birkhäuser.

Scricciolo, C. (1999). Probability matching priors: A review. *Journal of the Italian Statistical Society, 8*(1), 83-100.

Sharma, G., & Mathew, T. (2012). One-sided and two-sided tolerance intervals in general mixed and random effects models using small-sample asymptotics. J*ournal of the American Statistical Association*, *107*(497), 258-267.

Tang, L. C., & Chang, D. S. (1994). Tolerance limits for inverse Gaussian distribution. *Journal of Statistical Computation and Simulation, 51*(1), 21-29.

Thoman, D. R., Bain, L. J., & Antle, C. E. (1969). Inferences on the parameters of the Weibull distribution. *Technometrics, 11*(3), 445-460.

Thoman, D. R., Bain, L. J., & Antle, C. E. (1970). Maximum likelihood estimation, exact confidence intervals for reliability, and tolerance limits in the Weibull distribution. *Technometrics, 12*(2), 363-371.

Van Boekel, M. A. J. S., Stein, A., & Van Bruggen, A. H. C. (2004). *Bayesian statistics and quality modelling in the agro-food production chain.* (Vol. 3). Dordrecht, Netherlands: Kluwer Academic.

Van der Merwe, A. J., & Hugo, J. (2007). Bayesian tolerance intervals for the balanced two-factor nested random effects model. *Test, 16*(3), 598-612.

Vangel, M. G. (1992). New methods for one-sided tolerance limits for a one-way balanced random-effects ANOVA model. *Technometrics, 34*(2), 176-185

Wald, A., & Wolfowitz, J. (1946). Tolerance limits for a normal distribution. *The Annals of Mathematical Statistics, 17*(2), 208-215.

Welch, B. L., & Peers, H. W. (1963). On formulae for confidence points based on integrals of weighted likelihoods. *Journal of the Royal Statistical Society: Series B (Methodological), 25*(2), 318-329.

Wilks, S. S. (1941). Determination of sample sizes for setting tolerance limits. *The Annals of Mathematical Statistics, 12*(1), 91-96.

Wilks, S. S. (1942). Statistical prediction with special reference to the problem of tolerance limits. *The Annals of Mathematical Statistics, 13*(4), 400-409.

Wolfinger, R. D. (1998). Tolerance intervals for variance component models using Bayesian simulation. *Journal of Quality Technology, 30*(1), 18-32.

Young, D. S. (2010). Tolerance: An R package for estimating tolerance intervals. *Journal of Statistical Software, 36*(5), 1-39.

Zacks, S. (1970). Uniformly most accurate upper tolerance limits for monotone likelihood ratio families of discrete distributions. *Journal of the American Statistical Association, 65*(329), 307-316.

APPENDIX A

We present two results on the determinant of a matrix A and its inverse A^{-1} by Sahai and Ojeda (2004). These frequently occur in many linear model problems.

Let the matrix A be

$$A = \begin{bmatrix} a+b & a & a & \cdots & a \\ a & a+b & a & \ldots & a \\ a & a & a+b & \ldots & a \\ \vdots & \vdots & \vdots & \vdots & \\ a & a & a & \cdots & a+b \end{bmatrix}, \tag{A.3.1}$$

where a and b are either scalars or square matrices of the same order. If a and b are scalars, the matrix A is written as

$$A = bI_n + aJ_n \tag{A.3.2}$$

where I_n is an $n \times n$ identity matrix while J_n is an $n \times n$ matrix with every element equal to unity.

Result A.1 (Sahai and Ojeda, 2004)

For matrix A defined by (A.1),

$$|A| = (|b + na|)(|b|^{n-1}) \text{ and } |A^{-1}| = (|b + na|^{-1})(|b|^{1-n}) \tag{A.3.3}$$

where $|A|$ and $|A^{-1}|$ denote the determinants of the matrix A and A^{-1} respectively.

Result A.2 (Sahai and Ojeda, 2004)

For the matrix A shown in (A.2.2),

$$A^{-1} = \theta_1 I_n + \theta_2 J_n, \tag{A.3.4}$$

where $\theta_1 = \dfrac{1}{b}$ and $\theta_1 = 1/b$ for $b \neq 0$, $b \neq -na$.

Derivation of c_{su} :

From (3.16), we note that, $c_{su} = -\{D_s D_u l(\theta)\}_{\theta=\hat{\theta}}$

$$c_{11} = - D_1^2 l(\theta)\Big|_{\theta=\hat{\theta}}$$

$$D_1^2 l(\theta) = -\frac{t}{\theta_3 + t\theta_2}$$

$$\therefore c_{11} = \frac{t}{\hat{\theta}_3 + t\hat{\theta}_2}$$

$$D_1 D_2 l(\theta) = -\frac{t^2}{(\theta_3 + t\theta_2)^2}(\overline{Y} - \theta_1)$$

$$\therefore c_{12} = 0$$

Similarly, $c_{13} = 0$

$$c_{12} = c_{21} = c_{13} = c_{31} = 0 .$$

$$D_2^2 l(\theta) = \frac{t^2}{2(\theta_3 + t\theta_2)^2} - \frac{U_1 t^2}{n(\theta_3 + t\theta_2)^3} - \frac{t^3}{(\theta_3 + t\theta_2)^3}(\overline{Y} - \theta_1)^2$$

Note that $\hat{\theta}_3 + t\hat{\theta}_2 = \dfrac{U_1}{n}$.

$$\therefore c_{22} = \frac{t^2}{2(\hat{\theta}_3 + t\hat{\theta}_2)^2}$$

$$D_2 D_3 l(\theta) = \frac{t}{2(\theta_3 + t\theta_2)^2} - \frac{tU_1}{n(\theta_3 + t\theta_2)^3} - \frac{t^2}{(\theta_3 + t\theta_2)^3}(\overline{Y} - \theta_1)^2$$

$$\therefore c_{23} = \frac{t}{2(\hat{\theta}_3 + t\hat{\theta}_2)^2} = c_{32}$$

$$D_3^2 l(\theta) = \frac{(t-1)}{2\theta_3^2} + \frac{1}{2(\theta_3 + t\theta_2)^2} - \frac{U_2}{n\theta_3^3} - \frac{U_1}{n(\theta_3 + t\theta_2)^3}$$

$$\therefore c_{33} = \frac{1}{2}\left[\frac{t-1}{\hat{\theta}_3^2} + \frac{1}{(\hat{\theta}_3 + t\hat{\theta}_2)^2}\right]$$

Therefore, we have in particular,

$$c_{11} = \frac{t}{\hat{\theta}_3 + t\hat{\theta}_2}, \; c_{12} = c_{21} = c_{13} = c_{31} = 0$$

$$c_{22} = \frac{t^2}{2(\hat{\theta}_3 + t\hat{\theta}_2)^2}, \; c_{23} = \frac{t}{2(\hat{\theta}_3 + t\hat{\theta}_2)^2} = c_{32}$$

$$c_{33} = \frac{1}{2}\left[\frac{t-1}{\hat{\theta}_3^2} + \frac{1}{(\hat{\theta}_3 + t\hat{\theta}_2)^2}\right]$$

__Derivation of__ a_{suw} :

From equation (3.16), $a_{suw} = \{D_s D_u D_w l(\theta)\}_{\theta=\hat{\theta}}$.

Derivation of a_{suw} *:*

$D_1^3 l(\theta) = 0$

$\therefore a_{111} = 0$

Similarly, $a_{122} = a_{212} = a_{221} = a_{133} = a_{313} = a_{331} = 0$ and

$a_{123} = a_{132} = a_{213} = a_{231} = a_{312} = a_{321} = 0$.

$$D_1^2 D_2 l(\theta) = \frac{t^2}{(\theta_3 + t\theta_2)^2}$$

$$\therefore a_{112} = \frac{t^2}{(\hat{\theta}_3 + t\hat{\theta}_2)^2}$$

i.e. $a_{112} = a_{121} = a_{211} \dfrac{t^2}{(\hat{\theta}_3 + t\hat{\theta}_2)^2}$

$$D_1^2 D_3 l(\theta) = \frac{t}{(\theta_3 + t\theta_2)^2}$$

$$\therefore a_{113} = \frac{t}{(\hat{\theta}_3 + t\hat{\theta}_2)^2}$$

i.e. $a_{113} = a_{131} = a_{311} \dfrac{t}{(\hat{\theta}_3 + t\hat{\theta}_2)^2}$

$$D_2^3 l(\theta) = -\frac{t^3}{(\theta_3 + t\theta_2)^3} + \frac{3U_1 t^3}{n(\theta_3 + t\theta_2)^4} + \frac{3t^4}{(\theta_3 + t\theta_2)^4}(\overline{Y} - \theta_1)^2$$

$$\therefore a_{222} = \frac{2t^3}{(\hat{\theta}_3 + t\hat{\theta}_2)^3}$$

$$D_2^2 D_3 l(\theta) = -\frac{t^2}{(\theta_3 + t\theta_2)^3} + \frac{3U_1 t^2}{n(\theta_3 + t\theta_2)^4} + \frac{3t^3}{(\theta_3 + t\theta_2)^4}(\overline{Y} - \theta_1)^2$$

i.e. $a_{223} = a_{232} = a_{322} = \dfrac{2t^2}{(\hat{\theta}_3 + t\hat{\theta}_2)^3}$

$$D_3^2 D_2 l(\theta) = -\frac{t}{(\theta_3 + t\theta_2)^3} + \frac{3U_1 t}{n(\theta_3 + t\theta_2)^4}$$

$$\therefore a_{332} = a_{323} = a_{233} = \frac{2t}{(\hat{\theta}_3 + t\hat{\theta}_2)^3}$$

$$D_3^3 l(\theta) = -\frac{(t-1)}{\theta_3^3} - \frac{1}{(\theta_3 + t\theta_2)^3} + \frac{3U_2}{n\theta_3^4} + \frac{3U_1}{n(\theta_3 + t\theta_2)^4}$$

$$\therefore a_{333} = \frac{2(t-1)}{\hat{\theta}_3^3} + \frac{2}{(\hat{\theta}_3 + t\hat{\theta}_2)^3}$$

Hence, we can see that

$$a_{111} = a_{122} = a_{212} = a_{221} = a_{133} = a_{313} = a_{331} = a_{123} = a_{132} = a_{213} = a_{231} = a_{312} = a_{321} = 0$$

Derivation of simplified notation in equations (3.33)-(3.36)

From (3.32), let

$\varepsilon = (0, 1, 1)'$. Then $\mu = \frac{1}{2b}\varepsilon$, $k = \frac{1}{2}b^{-1}(\varepsilon' C^{-1}\varepsilon)^{1/2}$,

$$\lambda = (\lambda_1, \lambda_2, \lambda_3)' = \frac{1}{k}C^{-1}\mu = \frac{1}{2bk}C^{-1}\varepsilon = \frac{1}{\sqrt{\varepsilon^{-1}C^{-1}\varepsilon}}C^{-1}\varepsilon$$

In L_{11}, $\frac{c^{22} + 2c^{23} + c^{33}}{8kb^3} = \frac{\varepsilon' C^{-1}\varepsilon}{8kb^3} = \frac{(2bk)^2}{8kb^3} = \frac{k}{2b}$

$$\begin{aligned} L_{11} &= \frac{\hat{\pi}_s}{\hat{\pi}}\lambda_s + \frac{1}{2}a_{suw}\lambda_s c^{uw} + \frac{(z^2-3)}{8kb^3}(c^{22} + 2c^{23} + c^{33}) + \frac{c^{11}}{2kb} \\ &= \frac{\hat{\pi}_s}{\hat{\pi}}\lambda_s + \frac{1}{2}a_{suw}\lambda_s c^{uw} + (z^2-3)\left(\frac{k}{2b}\right) + \frac{c^{11}}{2kb} \\ &= \frac{\hat{\pi}_s}{\hat{\pi}}\lambda_s + \frac{1}{2}a_{suw}\lambda_s c^{uw} + \frac{1}{2}kb^{-1}(z^2-3) + \frac{1}{2}(kb)^{-1}c^{11} \end{aligned}$$

In L_{21}, $\frac{\lambda_2 + \lambda_3}{2b^2} = \frac{\varepsilon'\lambda}{2b^2} = \frac{\left(\frac{1}{2bk}C^{-1}\varepsilon\right)}{2b^2} = \frac{\varepsilon' C^{-1}\varepsilon}{4kb^3} = \frac{4k^2b^2}{4kb^3} = \frac{k}{b}$

$$\begin{aligned} L_{21} &= -\frac{(\lambda_2 + \lambda_3)(z^2-1)}{2b^2} \\ &= -\frac{k}{b}(z^2-1) \end{aligned}$$

In L_3, $\frac{(\lambda_2 + \lambda_3)^2}{8kb^3} = \frac{(\varepsilon'\lambda)^2}{8kb^3} = \frac{\left(\frac{1}{2bk}\varepsilon' C^{-1}\varepsilon\right)^2}{8kb^3} = \frac{\left(\frac{1}{2bk}4k^2b^2\right)^2}{8kb^3} = \frac{k}{2b}$

$$\begin{aligned} L_3 &= \frac{1}{6}a_{suw}\lambda_s\lambda_u\lambda_w + \frac{(z^2-3)(\lambda_2+\lambda_3)^2}{8kb^3} \\ &= \frac{1}{6}a_{suw}\lambda_s\lambda_u\lambda_w + (z^2-3)\left(\frac{k}{2b}\right) \end{aligned}$$

APPENDIX B

OTHER TABLES FOR CHAPTER 5

Tables B1 to B20 show the simulated expected widths, expected contents and their respective standard errors; and coverage probabilities for the higher order asymptotic tolerance intervals where $\beta_1 = \beta_2 = 0.05$.

Table B1: Higher order asymptotic tolerance interval, $[d(\hat{\theta}) - g^{(1n)}; b(\hat{\theta}) + g^{(1n)}]$ where $\gamma = 0.90$: *univariate normal model with mean= variance=* θ

θ	Sample size	Expected width	Standard error of width	Expected content	Standard error of content	Coverage probability
4	15	7.419	0.434	0.927	0.025	0.888
	20	7.271	0.369	0.923	0.020	0.889
	25	7.179	0.326	0.921	0.018	0.892
	30	7.105	0.293	0.919	0.016	0.893
	50	6.955	0.226	0.915	0.012	0.894
8	15	10.326	0.435	0.923	0.021	0.895
	20	10.128	0.374	0.919	0.017	0.894
	25	9.999	0.333	0.917	0.015	0.890
	30	9.927	0.299	0.916	0.012	0.901
	50	9.737	0.229	0.912	0.009	0.896
12	15	12.533	0.441	0.920	0.019	0.890
	20	12.317	0.374	0.917	0.015	0.897
	25	12.176	0.335	0.915	0.013	0.899
	30	12.072	0.303	0.913	0.011	0.897
	50	11.867	0.232	0.910	0.008	0.899
16	15	14.391	0.442	0.918	0.018	0.887
	20	14.147	0.377	0.916	0.015	0.890
	25	13.996	0.337	0.914	0.012	0.891
	30	13.891	0.307	0.912	0.011	0.890
	50	13.660	0.233	0.909	0.007	0.896

Table B2: Higher order asymptotic tolerance interval, $[d(\hat{\theta})-g^{(1n)};b(\hat{\theta})+g^{(1n)}]$ where $\gamma=0.95$: *univariate normal model with mean= variance=* θ.

θ	Sample size	Expected width	Standard error of width	Expected content	Standard error of content	Coverage probability
4	15	7.789	0.444	0.940	0.022	0.942
	20	7.562	0.373	0.935	0.018	0.945
	25	7.423	0.332	0.931	0.016	0.946
	30	7.332	0.297	0.928	0.014	0.948
	50	7.107	0.227	0.921	0.011	0.948
8	15	10.788	0.452	0.935	0.019	0.944
	20	10.492	0.379	0.930	0.015	0.947
	25	10.309	0.337	0.926	0.013	0.945
	30	10.188	0.307	0.923	0.012	0.947
	50	9.919	0.232	0.917	0.009	0.949
12	15	13.081	0.455	0.932	0.018	0.944
	20	12.740	0.383	0.927	0.014	0.948
	25	12.528	0.340	0.924	0.012	0.950
	30	12.380	0.307	0.921	0.011	0.946
	50	12.070	0.233	0.915	0.008	0.947
16	15	14.997	0.463	0.930	0.017	0.940
	20	14.627	0.385	0.926	0.013	0.950
	25	14.390	0.347	0.922	0.012	0.944
	30	14.231	0.311	0.920	0.010	0.946
	50	13.883	0.241	0.914	0.008	0.945

Table B3: Higher order asymptotic tolerance interval, $[d(\hat{\theta}) - g^{(2n)}; b(\hat{\theta}) + g^{(2n)}]$ where $\gamma = 0.90$: *univariate normal model with mean= variance=* θ.

$\boldsymbol{\theta}$	**Sample size**	**Expected width**	**Standard error of width**	**Expected content**	**Standard error of content**	**Coverage probability**
4	15	7.589	0.445	0.933	0.023	0.918
	20	7.372	0.378	0.927	0.020	0.911
	25	7.248	0.329	0.924	0.017	0.913
	30	7.163	0.298	0.922	0.012	0.910
	50	6.986	0.223	0.916	0.011	0.910
8	15	10.664	0.469	0.932	0.020	0.932
	20	10.346	0.393	0.925	0.017	0.927
	25	10.152	0.342	0.921	0.014	0.918
	30	10.033	0.307	0.919	0.012	0.918
	50	9.792	0.231	0.913	0.009	0.917
12	15	13.096	0.488	0.933	0.018	0.943
	20	12.657	0.396	0.925	0.015	0.940
	25	12.410	0.348	0.921	0.013	0.932
	30	12.253	0.312	0.918	0.011	0.928
	50	11.946	0.237	0.912	0.008	0.920
16	15	15.198	0.512	0.934	0.017	0.951
	20	14.639	0.409	0.926	0.014	0.946
	25	14.318	0.354	0.921	0.012	0.937
	30	14.132	0.316	0.918	0.011	0.937
	50	13.769	0.235	0.912	0.007	0.926

Table B4: Higher order asymptotic tolerance interval, $[d(\hat{\theta}) - g^{(2n)}; b(\hat{\theta}) + g^{(2n)}]$ where $\gamma = 0.95$: *univariate normal model with mean= variance=* θ.

θ	Sample size	Expected width	Standard error of width	Expected content	Standard error of content	Coverage probability
4	15	8.147	0.495	0.950	0.021	0.965
	20	7.790	0.407	0.942	0.018	0.963
	25	7.570	0.353	0.936	0.017	0.957
	30	7.436	0.314	0.932	0.015	0.956
	50	7.158	0.232	0.923	0.011	0.959
8	15	11.599	0.540	0.952	0.016	0.982
	20	10.988	0.427	0.942	0.015	0.976
	25	10.642	0.366	0.935	0.013	0.974
	30	10.427	0.327	0.930	0.012	0.967
	50	10.023	0.238	0.921	0.009	0.962
12	15	14.418	0.586	0.956	0.015	0.988
	20	13.546	0.443	0.944	0.013	0.985
	25	13.066	0.379	0.935	0.012	0.978
	30	12.774	0.332	0.930	0.011	0.974
	50	12.239	0.243	0.920	0.008	0.972
16	15	16.964	0.634	0.959	0.014	0.991
	20	15.771	0.468	0.945	0.012	0.989
	25	15.161	0.391	0.937	0.011	0.984
	30	14.784	0.344	0.931	0.010	0.979
	50	14.120	0.245	0.919	0.007	0.976

Table B5: Higher order asymptotic tolerance interval, $[d(\hat{\theta}) - g^{(3n)}; b(\hat{\theta}) + g^{(3n)}]$ where $\gamma = 0.90$: *univariate normal model with mean= variance=* θ.

θ	Sample size	Expected width	Standard error of width	Expected content	Standard error of content	Coverage probability
4	15	8.336	0.654	0.954	0.024	0.960
	20	7.713	0.447	0.939	0.020	0.948
	25	7.441	0.366	0.931	0.018	0.940
	30	7.293	0.316	0.926	0.015	0.934
	50	7.030	0.229	0.918	0.011	0.923
8	15	15.405	946.584	0.604	0.758	0.778
	20	12.242	0.964	0.963	0.018	0.989
	25	10.953	0.489	0.942	0.015	0.977
	30	10.484	0.375	0.931	0.013	0.966
	50	9.915	0.243	0.917	0.009	0.945
12	15	8.066	1.543	0.733	0.111	0.000
	20	4.069×10^{4}	4.068×10^{6}	0.011	0.795	0.260
	25	16.397	2.500	0.975	0.016	0.997
	30	13.639	0.577	0.947	0.012	0.990
	50	12.211	0.255	0.919	0.008	0.964
16	15	11.569	0.611	0.838	0.032	0.000
	20	10.261	0.986	0.787	0.056	0.000
	25	89.376	1.0874×10^{3}	0.206	0.549	0.041
	30	30.908	970.354	0.969	0.195	0.989
	50	7.033	0.230	0.918	0.012	0.923

Table B6: Higher order asymptotic tolerance interval, $[d(\hat{\theta})-g^{(3n)}; b(\hat{\theta})+g^{(3n)}]$ where $\gamma=0.95$: *univariate normal model with mean= variance=* θ.

θ	Sample size	Expected width	Standard error of width	Expected content	Standard error of content	Coverage probability
4	15	16.803	202.124	0.978	0.168	0.990
	20	9.078	0.796	0.971	0.018	0.992
	25	8.200	0.478	0.955	0.016	0.990
	30	7.820	0.376	0.945	0.014	0.984
	50	7.273	0.243	0.928	0.011	0.975
8	15	6.100	2.083	0.693	0.126	0.000
	20	-24.210	1886.100	-0.024	0.605	0.057
	25	13.859	1263.100	0.893	0.437	0.947
	30	13.086	1.120	0.975	0.013	0.999
	50	10.432	0.285	0.932	0.009	0.990
12	15	9.963	0.571	0.836	0.034	0
	20	9.315	0.663	0.809	0.040	0
	25	7.823	1.218	0.726	0.090	0
	30	-0.878	145.508	0.153	0.511	0.019
	50	13.458	0.411	0.945	0.009	0.998
16	15	12.176	0.496	0.859	0.026	0
	20	11.915	0.458	0.854	0.022	0
	25	11.519	0.485	0.842	0.023	0
	30	10.842	0.602	0.816	0.029	0
	50	20.162	3.028	0.984	0.011	1.000

Table B7: Higher order asymptotic tolerance interval, $[d(\hat{\theta})-g^{(1n)}; b(\hat{\theta})+g^{(1n)}]$ where $\gamma = 0.90$: *Weibull model.*

(θ_1, θ_2)	**Sample size**	**Expected width**	**Standard error of width**	**Expected content**	**Standard error of content**	**Coverage probability**
(1, 2)	15	1.799	0.305	0.945	0.049	0.850
	20	1.756	0.263	0.943	0.043	0.859
	25	1.731	0.230	0.943	0.038	0.876
	30	1.704	0.211	0.939	0.036	0.867
	50	1.656	0.157	0.934	0.028	0.884
(5,5)	15	4.406	0.766	0.946	0.049	0.857
	20	4.257	0.646	0.943	0.043	0.864
	25	4.151	0.554	0.941	0.038	0.869
	30	4.099	0.496	0.940	0.034	0.876
	50	3.934	0.363	0.934	0.028	0.887
(10,3)	15	13.348	2.224	0.946	0.050	0.849
	20	12.941	1.838	0.944	0.043	0.860
	25	12.689	1.607	0.942	0.038	0.869
	30	12.499	1.436	0.940	0.035	0.875
	50	12.071	1.080	0.934	0.028	0.891
(15,6)	15	11.231	2.037	0.945	0.049	0.854
	20	10.872	1.690	0.943	0.043	0.861
	25	10.619	1.474	0.941	0.039	0.865
	30	10.449	1.303	0.939	0.035	0.876
	50	10.073	0.945	0.934	0.027	0.887

Table B8: Higher order asymptotic tolerance interval, $[d(\hat{\theta}) - g^{(1n)}; b(\hat{\theta}) + g^{(1n)}]$ where $\gamma = 0.95$: *Weibull model.*

(θ_1, θ_2)	**Sample size**	**Expected width**	**Standard error of width**	**Expected content**	**Standard error of content**	**Coverage probability**
(1, 2)	15	1.886	0.314	0.955	0.044	0.899
	20	1.830	0.263	0.954	0.038	0.916
	25	1.794	0.233	0.953	0.035	0.919
	30	1.768	0.211	0.951	0.032	0.928
	50	1.701	0.159	0.943	0.026	0.935
(5,5)	15	4.679	0.810	0.959	0.042	0.910
	20	4.496	0.669	0.956	0.037	0.924
	25	4.379	0.584	0.953	0.034	0.927
	30	4.280	0.513	0.951	0.031	0.928
	50	4.077	0.376	0.943	0.025	0.940
(10,3)	15	14.152	2.307	0.959	0.043	0.907
	20	14.172	2.288	0.960	0.042	0.913
	25	13.298	1.656	0.954	0.034	0.926
	30	13.074	1.476	0.952	0.031	0.932
	50	12.479	1.103	0.944	0.025	0.938
(15,6)	15	11.996	2.154	0.958	0.042	0.910
	20	11.469	1.771	0.955	0.037	0.914
	25	11.190	1.516	0.953	0.033	0.930
	30	10.962	1.360	0.951	0.031	0.928
	50	10.415	0.994	0.943	0.026	0.933

Table B9: Higher order asymptotic tolerance interval, $[d(\hat{\theta})-g^{(2n)};b(\hat{\theta})+g^{(2n)}]$ where $\gamma=0.90$: *Weibull model.*

(θ_1, θ_2)	**Sample size**	**Expected width**	**Standard error of width**	**Expected content**	**Standard error of content**	**Coverage probability**
(1, 2)	15	1.840	0.310	0.951	0.046	0.878
	20	1.777	0.260	0.947	0.041	0.882
	25	1.746	0.234	0.945	0.038	0.884
	30	1.723	0.209	0.943	0.034	0.892
	50	1.662	0.158	0.935	0.028	0.894
(5,5)	15	4.552	0.795	0.953	0.045	0.889
	20	4.357	0.653	0.949	0.040	0.890
	25	4.232	0.568	0.946	0.036	0.893
	30	4.143	0.499	0.943	0.033	0.895
	50	3.958	0.363	0.936	0.027	0.899
(10,3)	15	13.739	2.253	0.953	0.046	0.886
	20	13.222	1.858	0.949	0.041	0.890
	25	12.882	1.619	0.946	0.037	0.893
	30	12.626	1.433	0.943	0.034	0.894
	50	12.108	1.062	0.935	0.027	0.899
(15,6)	15	11.636	2.078	0.953	0.045	0.887
	20	11.128	1.721	0.949	0.040	0.892
	25	10.815	1.497	0.945	0.036	0.890
	30	10.605	1.313	0.943	0.033	0.895
	50	10.147	0.963	0.936	0.027	0.896

Table B10: Higher order asymptotic tolerance interval, $[d(\hat{\theta})-g^{(2n)};b(\hat{\theta})+g^{(2n)}]$ where $\gamma=0.95$: *Weibull model.*

(θ_1, θ_2)	**Sample size**	**Expected width**	**Standard error of width**	**Expected content**	**Standard error of content**	**Coverage probability**
(1, 2)	15	1.927	0.316	0.959	0.041	0.916
	20	1.860	0.267	0.036	0.036	0.927
	25	1.818	0.236	0.956	0.033	0.936
	30	1.782	0.210	0.953	0.030	0.937
	50	1.708	0.161	0.944	0.026	0.940
(5,5)	15	4.826	0.839	0.964	0.039	0.931
	20	4.606	0.686	0.961	0.034	0.939
	25	4.459	0.592	0.957	0.032	0.941
	30	4.337	0.526	0.954	0.030	0.940
	50	4.098	0.379	0.945	0.025	0.942
(10,3)	15	14.610	2.331	0.965	0.038	0.934
	20	13.948	1.934	0.961	0.035	0.939
	25	13.519	1.678	0.958	0.032	0.942
	30	13.206	1.503	0.954	0.030	0.941
	50	12.528	1.103	0.945	0.025	0.942
(15,6)	15	12.408	2.225	0.964	0.038	0.934
	20	11.767	1.809	0.960	0.035	0.933
	25	11.383	1.547	0.956	0.032	0.941
	30	11.120	1.375	0.954	0.030	0.943
	50	10.511	1.005	0.945	0.025	0.947

Table B11: Higher order asymptotic tolerance interval, $[d(\hat{\theta}) - g^{(3n)}; b(\hat{\theta}) + g^{(3n)}]$ where $\gamma = 0.90$: *Weibull model.*

(θ_1, θ_2)	**Sample size**	**Expected width**	**Standard error of width**	**Expected content**	**Standard error of content**	**Coverage probability**
(1, 2)	15	1.946	0.309	0.962	0.039	0.931
	20	1.838	0.261	0.956	0.036	0.922
	25	1.779	0.232	0.951	0.036	0.909
	30	1.744	0.208	0.947	0.033	0.913
	50	1.672	0.159	0.937	0.027	0.904
(5,5)	15	4.997	0.847	0.970	0.035	0.952
	20	4.589	0.670	0.960	0.034	0.939
	25	4.375	0.576	0.954	0.033	0.929
	30	4.246	0.513	0.949	0.031	0.922
	50	3.997	0.367	0.938	0.026	0.918
(10,3)	15	14.881	2.312	0.969	0.035	0.948
	20	13.813	1.888	0.959	0.035	0.933
	25	13.260	1.624	0.953	0.033	0.926
	30	12.922	1.463	0.949	0.032	0.920
	50	12.246	1.076	0.939	0.026	0.914
(15,6)	15	12.847	2.242	0.970	0.034	0.954
	20	11.749	1.818	0.959	0.035	0.935
	25	11.197	1.523	0.953	0.033	0.929
	30	10.878	1.340	0.949	0.031	0.929
	50	10.218	0.962	0.938	0.026	0.915

Table B12: Higher order asymptotic tolerance interval, $[d(\hat{\theta}) - g^{(3n)}; b(\hat{\theta}) + g^{(3n)}]$ where $\gamma = 0.95$: *Weibull model.*

(θ_1, θ_2)	**Sample size**	**Expected width**	**Standard error of width**	**Expected content**	**Standard error of content**	**Coverage probability**
(1, 2)	15	2.035	0.318	0.968	0.035	0.950
	20	1.918	0.266	0.964	0.032	0.951
	25	1.849	0.232	0.960	0.031	0.950
	30	1.806	0.210	0.956	0.029	0.956
	50	1.721	0.160	0.947	0.025	0.950
(5,5)	15	5.323	0.897	0.979	0.028	0.975
	20	4.872	0.700	0.971	0.028	0.971
	25	4.606	0.603	0.964	0.029	0.960
	30	4.457	0.532	0.959	0.027	0.963
	50	4.143	0.379	0.947	0.024	0.953
(10,3)	15	15.758	2.390	0.978	0.030	0.971
	20	14.563	1.949	0.970	0.030	0.964
	25	13.881	1.693	0.963	0.029	0.957
	30	13.497	1.503	0.959	0.028	0.960
	50	12.649	1.108	0.948	0.024	0.953
(15,6)	15	13.750	2.381	0.979	0.027	0.977
	20	12.512	1.874	0.970	0.028	0.969
	25	11.860	1.604	0.964	0.028	0.965
	30	11.429	1.409	0.959	0.027	0.961
	50	10.613	0.988	0.947	0.024	0.957

Table B13: Purely frequentist higher order asymptotic tolerance interval, $[d(\hat{\theta}) - g_f^{(1n)}, b(\hat{\theta}) + g_f^{(1n)}]$ where $\gamma = 0.90$: *inverse Gaussian model.*

(θ_1, θ_2)	**Sample size**	**Expected width**	**Standard error of width**	**Expected content**	**Standard error of content**	**Coverage probability**
(7, 14)	15	16.476	4.599	0.935	0.052	0.803
	20	16.185	3.912	0.937	0.044	0.828
	25	16.024	3.510	0.937	0.039	0.842
	30	15.856	3.207	0.936	0.036	0.846
	50	15.602	2.426	0.933	0.028	0.876
(8, 12)	15	20.749	6.353	0.934	0.052	0.793
	20	20.430	5.386	0.936	0.044	0.819
	25	20.236	4.828	0.936	0.039	0.839
	30	20.086	4.408	0.936	0.037	0.850
	50	19.748	3.420	0.932	0.029	0.864
(15, 25)	15	37.490	11.170	0.933	0.053	0.795
	20	37.134	9.509	0.937	0.044	0.826
	25	36.570	8.481	0.936	0.040	0.843
	30	36.412	7.595	0.937	0.035	0.859
	50	35.774	5.946	0.963	0.028	0.875
(20, 50)	15	43.702	11.269	0.937	0.050	0.821
	20	42.742	9.729	0.938	0.043	0.832
	25	42.339	8.596	0.938	0.038	0.849
	30	41.761	7.840	0.937	0.036	0.856
	50	40.870	6.037	0.932	0.029	0.874

Table B14: Purely frequentist higher order asymptotic tolerance interval, $[d(\hat{\theta})-g_f^{(1n)}, b(\hat{\theta})+g_f^{(1n)}]$ where $\gamma = 0.95$: *inverse Gaussian model.*

(θ_1, θ_2)	**Sample size**	**Expected width**	**Standard error of width**	**Expected content**	**Standard error of content**	**Coverage probability**
(7, 14)	15	17.105	4.676	0.942	0.046	0.846
	20	16.802	3.988	0.944	0.040	0.874
	25	16.472	3.504	0.944	0.035	0.891
	30	16.280	3.185	0.944	0.033	0.900
	50	15.894	2.460	0.940	0.026	0.923
(8, 12)	15	21.365	6.365	0.939	0.048	0.826
	20	20.924	5.449	0.941	0.041	0.859
	25	20.728	4.885	0.942	0.036	0.881
	30	20.538	4.459	0.942	0.033	0.892
	50	20.076	3.386	0.940	0.026	0.922
(15, 25)	15	38.814	11.022	0.940	0.047	0.840
	20	38.087	9.491	0.943	0.040	0.866
	25	37.556	8.469	0.943	0.036	0.882
	30	37.221	7.826	0.943	0.033	0.896
	50	36.288	5.966	0.940	0.027	0.916
(20, 50)	15	45.486	11.537	0.944	0.046	0.857
	20	44.400	9.880	0.945	0.040	0.879
	25	43.610	8.726	0.945	0.035	0.892
	30	43.062	7.846	0.945	0.032	0.907
	50	41.842	6.027	0.941	0.026	0.926

Table B15: Purely frequentist higher order asymptotic tolerance interval, $[d(\hat{\theta}) - g_f^{(2n)}, b(\hat{\theta}) + g_f^{(2n)}]$ where $\gamma = 0.90$: *inverse Gaussian model.*

(θ_1, θ_2)	**Sample size**	**Expected width**	**Standard error of width**	**Expected content**	**Standard error of content**	**Coverage probability**
(7, 14)	15	16.810	4.593	0.939	0.050	0.830
	20	16.361	3.912	0.940	0.042	0.846
	25	16.173	3.563	0.939	0.039	0.857
	30	15.944	3.212	0.938	0.036	0.861
	50	15.609	2.449	0.934	0.028	0.884
(8, 12)	15	20.949	6.355	0.936	0.050	0.811
	20	20.640	5.452	0.938	0.043	0.832
	25	20.358	4.825	0.938	0.039	0.854
	30	20.167	4.451	0.937	0.035	0.856
	50	19.825	3.389	0.933	0.028	0.877
(15, 25)	15	38.230	11.092	0.938	0.049	0.823
	20	37.234	9.575	0.938	0.043	0.836
	25	36.861	8.585	0.938	0.040	0.849
	30	36.540	7.697	0.938	0.035	0.866
	50	35.718	5.988	0.933	0.028	0.876
(20, 50)	15	44.790	11.552	0.941	0.049	0.839
	20	43.489	9.858	0.941	0.042	0.856
	25	42.701	8.666	0.940	0.037	0.867
	30	42.170	7.920	0.939	0.035	0.870
	50	40.975	5.977	0.934	0.028	0.883

Table B16: Purely frequentist higher order asymptotic tolerance interval, $[d(\hat{\theta}) - g_f^{(2n)}, b(\hat{\theta}) + g_f^{(2n)}]$ where $\gamma = 0.95$: *inverse Gaussian model.*

(θ_1, θ_2)	**Sample size**	**Expected width**	**Standard error of width**	**Expected content**	**Standard error of content**	**Coverage probability**
(7, 14)	15	17.491	4.670	0.945	0.044	0.860
	20	16.976	3.968	0.946	0.038	0.885
	25	16.629	3.510	0.946	0.034	0.906
	30	16.404	3.218	0.945	0.032	0.907
	50	15.953	2.474	0.942	0.026	0.928
(8, 12)	15	21.644	6.336	0.941	0.046	0.842
	20	21.255	5.594	0.943	0.040	0.866
	25	20.797	4.917	0.943	0.036	0.884
	30	20.664	4.451	0.944	0.033	0.904
	50	20.122	3.436	0.940	0.027	0.920
(15, 25)	15	39.553	11.352	0.942	0.046	0.850
	20	38.491	9.562	0.944	0.039	0.877
	25	37.904	8.553	0.945	0.035	0.899
	30	37.447	7.761	0.945	0.032	0.908
	50	36.518	5.999	0.941	0.026	0.929
(20, 50)	15	46.713	11.648	0.948	0.043	0.884
	20	45.195	9.965	0.948	0.038	0.897
	25	43.999	8.709	0.947	0.034	0.908
	30	43.353	7.939	0.946	0.031	0.913
	50	41.968	6.081	0.942	0.026	0.930

Table B17: Purely frequentist higher order asymptotic tolerance interval, $[d(\hat{\theta})-g_f^{(3n)}, b(\hat{\theta})+g_f^{(3n)}]$ where $\gamma = 0.90$: *inverse Gaussian model.*

(θ_1, θ_2)	**Sample size**	**Expected width**	**Standard error of width**	**Expected content**	**Standard error of content**	**Coverage probability**
(7, 14)	15	17.784	4.730	0.947	0.043	0.873
	20	16.909	3.958	0.946	0.038	0.883
	25	16.430	3.532	0.944	0.036	0.885
	30	16.222	3.180	0.943	0.033	0.894
	50	15.715	2.426	0.936	0.027	0.901
(8, 12)	15	22.104	6.363	0.943	0.045	0.857
	20	21.195	5.521	0.943	0.040	0.871
	25	20.723	4.881	0.942	0.036	0.883
	30	20.423	4.439	0.941	0.034	0.886
	50	19.888	3.389	0.936	0.027	0.896
(15, 25)	15	40.086	11.101	0.944	0.044	0.865
	20	38.455	9.449	0.944	0.039	0.871
	25	37.504	8.559	0.943	0.036	0.883
	30	36.985	7.688	0.942	0.034	0.889
	50	35.990	5.983	0.936	0.027	0.897
(20, 50)	15	47.546	11.655	0.950	0.041	0.895
	20	45.011	9.974	0.948	0.038	0.892
	25	43.690	8.720	0.946	0.035	0.895
	30	42.754	7.938	0.943	0.033	0.895
	50	41.269	6.055	0.936	0.028	0.899

Table B18: Purely frequentist higher order asymptotic tolerance interval, $[d(\hat{\theta})-g_f^{(3n)}, b(\hat{\theta})+g_f^{(3n)}]$ where $\gamma = 0.95$: *inverse Gaussian model.*

(θ_1, θ_2)	**Sample size**	**Expected width**	**Standard error of width**	**Expected content**	**Standard error of content**	**Coverage probability**
(7, 14)	15	18.447	4.788	0.950	0.041	0.891
	20	17.434	3.972	0.949	0.036	0.907
	25	16.987	3.555	0.949	0.032	0.918
	30	16.643	3.209	0.948	0.030	0.923
	50	16.015	2.463	0.943	0.025	0.937
(8, 12)	15	22.703	6.422	0.946	0.042	0.869
	20	21.724	5.501	0.946	0.037	0.889
	25	21.233	4.871	0.947	0.033	0.909
	30	20.882	4.414	0.946	0.031	0.915
	50	20.233	3.479	0.942	0.026	0.926
(15, 25)	15	41.630	11.230	0.948	0.041	0.881
	20	39.506	9.577	0.947	0.037	0.895
	25	38.510	8.593	0.947	0.033	0.908
	30	37.877	7.764	0.947	0.031	0.917
	50	36.545	5.985	0.943	0.026	0.933
(20, 50)	15	39.490	6.931	0.942	0.026	0.930
	20	46.741	10.103	0.953	0.034	0.920
	25	45.059	8.895	0.951	0.032	0.925
	30	44.232	8.052	0.950	0.030	0.930
	50	42.278	6.101	0.944	0.025	0.940

Table B19: Bayesian higher order asymptotic tolerance interval, $[d(\hat{\theta}) - g^{(3n)}; b(\hat{\theta}) + g^{(3n)}]$ where $\gamma = 0.90$: *inverse Gaussian model.*

(θ_1, θ_2)	**Sample size**	**Expected width**	**Standard error of width**	**Expected content**	**Standard error of content**	**Coverage probability**
(7, 14)	15	17.874	4.738	0.948	0.042	0.876
	20	16.998	4.006	0.946	0.038	0.887
	25	16.472	3.535	0.944	0.036	0.886
	30	16.206	3.197	0.942	0.033	0.893
	50	15.684	2.453	0.936	0.028	0.893
(8, 12)	15	22.166	6.484	0.943	0.044	0.861
	20	21.229	5.548	0.943	0.040	0.865
	25	20.749	4.849	0.943	0.036	0.880
	30	20.473	4.432	0.942	0.034	0.887
	50	19.867	3.408	0.935	0.028	0.887
(15, 25)	15	40.202	11.412	0.944	0.045	0.858
	20	38.212	9.460	0.943	0.040	0.874
	25	37.534	8.508	0.943	0.036	0.879
	30	36.928	7.691	0.942	0.033	0.889
	50	35.893	5.927	0.935	0.028	0.890
(20, 50)	15	47.391	11.685	0.950	0.041	0.887
	20	44.767	9.772	0.947	0.038	0.891
	25	43.503	8.767	0.945	0.035	0.892
	30	42.765	7.870	0.943	0.032	0.896
	50	41.153	6.040	0.936	0.028	0.899

Table B20: Bayesian higher order asymptotic tolerance interval, $[d(\hat{\theta}) - g^{(3n)}; b(\hat{\theta}) + g^{(3n)}]$ where $\gamma = 0.95$: *inverse Gaussian model.*

(θ_1, θ_2)	**Sample size**	**Expected width**	**Standard error of width**	**Expected content**	**Standard error of content**	**Coverage probability**
(7, 14)	15	18.473	4.800	0.850	0.041	0.893
	20	17.516	4.049	0.950	0.036	0.908
	25	16.976	3.596	0.949	0.033	0.915
	30	16.615	3.209	0.948	0.030	0.924
	50	16.023	2.488	0.943	0.025	0.939
(8, 12)	15	22.827	6.488	0.946	0.042	0.875
	20	21.863	5.565	0.947	0.037	0.896
	25	21.208	4.961	0.946	0.035	0.898
	30	20.854	4.466	0.946	0.032	0.912
	50	20.295	3.425	0.943	0.026	0.935
(15, 25)	15	41.562	11.346	0.948	0.042	0.882
	20	39.597	9.830	0.947	0.037	0.896
	25	38.528	8.749	0.947	0.034	0.903
	30	37.818	7.767	0.947	0.031	0.919
	50	39.583	5.956	0.943	0.025	0.934
(20, 50)	15	49.455	11.878	0.954	0.039	0.911
	20	46.630	10.000	0.953	0.034	0.923
	25	45.129	8.902	0.951	0.032	0.927
	30	44.037	8.003	0.949	0.031	0.930
	50	42.129	6.028	0.944	0.025	0.941