BAYESIAN CONTROL CHARTS FOR TOLERANCE LIMITS IN THE CASE OF NORMAL POPULATIONS

R. van Zyl [1]
Quintiles International
e-mail: *ruaan.vanzyl@quintiles.com*

A.J. van der Merwe
University of the Free State

Key words: Control charts, Probability-matching prior, Reference prior, Tolerance limits.

Abstract: By using air-lead data analysed by Krishnamoorthy and Mathew (2009) a Bayesian procedure is applied to obtain control limits for the upper one-sided tolerance limit. Reference and probability matching priors are derived for the *p*th quantile of a normal distribution. By simulating the predictive density of a future upper one-sided tolerance limit, "run-lengths" and average "run-lengths" are derived. In the second part of this paper control limits are derived for one-sided tolerance limits for the distribution of the difference between two normal random variables. This article illustrates the flexibility and unique features of the Bayesian simulation method for obtaining the posterior predictive distribution of a future one-sided tolerance limit.

1. Introduction

Krishnamoorthy and Mathew (2009) and Hahn and Meeker (1991) defined a tolerance interval as an interval that is constructed in such a way that it will contain a specified proportion (or more) of the population with a certain degree of confidence. The proportion is also called the content of the tolerance interval. As opposed to confidence intervals that give information on unknown population parameters, a one-sided upper tolerance limit provides, for example, information about a quantile of the population. According to Hahn and Meeker (1991) tolerance intervals would be of importance in obtaining limits on the process capability of a product manufactured in large quantities. Further application examples of tolerance intervals include statistical process control, wood manufacturing, clinical and industrial applications, environmental monitoring and assessment and for exposure data analysis. For more applications see Krishnamoorthy and Mathew (2009) and Hugo (2012).

2. One-sided Tolerance Limits for a Normal Population

Suppose $X_1, X_2, \ldots, X_n$ is a random sample from a $N\left(\mu, \sigma^2\right)$ population. The maximum likelihood estimators of the unknown mean, μ, and unknown variance, σ^2, are the sample mean, $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$,

[1] Corresponding author.

AMS: 62F15, 62F25

and sample variance, $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$, respectively. Using the same notation as given in Krishnamoorthy and Mathew (2009), the pth quantile of a $N\left(\mu, \sigma^2\right)$ population is

$$q_p = \mu + z_p\sigma \tag{1}$$

where z_p denotes the pth quantile of a standard normal distribution.

A $1-\alpha$ upper confidence limit for q_p is a $(p, 1-\alpha)$ one-sided upper tolerance limit for the normal distribution. By using the posterior predictive distribution a Bayesian procedure will be developed to obtain control limits for a one-sided upper tolerance limit in the case of future samples.

Bayarri and García-Donato (2005) give the following reasons for recommending a Bayesian analysis:

- Control charts are based on future observations and Bayesian methods are very natural for prediction.
- Uncertainty in the estimation of the unknown parameters is adequately handled.
- Implementation with complicated models and in a sequential scenario poses no methodological difficulty, the numerical difficulties are easily handled via Monte Carlo methods.
- Objective Bayesian analysis is possible without introduction of external information other than the model, but any kind of prior information can be incorporated into the analysis, if desired.

There do not appear to be many papers on control charts for tolerance intervals from a Bayesian point of view. Hamada (2002) derived Bayesian tolerance interval control limits for np, p, c and u charts which control the probability content at a specified level with a given confidence while we are deriving posterior predictive intervals. It is therefore clear that our Bayesian method differs substantially from his.

3. Bayesian Procedure

By assigning a prior distribution to the unknown parameters the uncertainty in the estimation of the unknown parameters can adequately be handled. The information contained in the prior is combined with the likelihood function to obtain the posterior distribution of q_p. By using the posterior distribution the predictive distribution of a future sample one-sided upper tolerance limit can be obtained. The predictive distribution on the other hand can be used to obtain control limits and to determine the distribution of the "run length" and "expected run length". Determination of reasonable non-informative priors is however not an easy task. Therefore, in the next section, reference and probability matching priors will be derived for $q_p = \mu + z_p\sigma$, the pth quantile of a $N\left(\mu, \sigma^2\right)$ distribution.

4. Reference and Probability-Matching Priors for $q_p = \mu + z_p\sigma$

As mentioned the Bayesian paradigm emerges as attractive in many types of statistical problems, also in the case of q_p, the pth quantile of a $N\left(\mu, \sigma^2\right)$ population.

Prior distributions are needed to complete the Bayesian specification of the model. Determination of reasonable non-informative priors in multi-parameter problems is not easy; common non-informative priors, such as the Jeffreys' prior can have features that have an unexpectedly dramatic effect on the posterior.

Reference and probability-matching priors often lead to procedures with good frequency properties while returning the Bayesian flavour. The fact that the resulting Bayesian posterior intervals of the level $1-\alpha$ are also good frequentist intervals at the same level is a very desirable situation.

See also Bayarri and Berger (2004) and Severine, Mukerjee and Ghosh (2002) for a general discussion.

4.1. The Reference Prior

In this section the reference prior of Berger and Bernardo (1992) will be derived for $q_p = \mu + z_p\sigma$. In general, the derivation of the reference prior depends on the ordering of the parameters and how the parameter vector is divided into sub-vectors. As mentioned by Pearn and Wu (2005) the reference prior maximises the difference in information (entropy) about the parameter provided by the prior and posterior. In other words, the reference prior is derived in such a way that it provides as little information possible about the parameter of interest. The reference prior algorithm is relatively complicated and, as mentioned, the solution depends on the ordering of the parameters and how the parameter vector is partitioned into sub-vectors. In spite of these difficulties, there is growing evidence, mainly through examples that reference priors provide "sensible" answers from a Bayesian point of view and that frequentist properties of inference from reference posteriors are asymptotically "good". As in the case of the Jeffreys' prior, the reference prior is obtained from the Fisher information matrix. In the case of a scalar parameter, the reference prior is the Jeffreys' prior.

The following theorem can be proved:

Theorem 1 The reference prior for the ordering $\left\{q_p, \sigma^2\right\}$ is given by $p_R\left(q_p, \sigma^2\right) \propto \sigma^{-2}$.

In the (μ, σ) parametrisation this corresponds to $p_R(\mu, \sigma) \propto \sigma^{-2}$.

Proof. The proof is given in the Appendix. ■

Note: The ordering $\left\{q_p, \sigma^2\right\}$ means that the parameter $q_p = \mu + z_p\sigma$ is a more important parameter than σ^2.

4.2. Probability-Matching Priors

The reference prior algorithm is but one way to obtain a useful non-informative prior. Another type of non-informative prior is the probability-matching prior. This prior has good frequentist properties. Two reasons for using probability-matching priors are that they provide a method for constructing accurate frequentist intervals, and that they could be potentially useful for comparative purposes in a Bayesian analysis.

There are two methods for generating probability-matching priors due to Tibshirani (1989) and Datta and Ghosh (1995).

Tibshirani (1989) generated probability-matching priors by transforming the model parameters so that the parameter of interest is orthogonal to the other parameters. The prior distribution is then taken to be proportional to the square root of the upper left element of the information matrix in the new parametrisation.

Datta and Ghosh (1995) provided a different solution to the problem of finding probability-matching priors. They derived the differential equation that a prior must satisfy if the posterior probability of a one-sided credibility interval for a parametric function and its frequentist probability agree up to $O\left(n^{-1}\right)$ where n is the sample size. Using the method of Datta and Ghosh (1995) the following theorem will be proved.

Theorem 2 The probability-matching prior for q_p and σ^2 is $p_M\left(q_p,\sigma^2\right) \propto \sigma^{-2}$.

Proof. The proof is given in the Appendix. ■

4.3. The Posterior Distribution

As mentioned, by combining the information contained in the prior with the likelihood function the posterior distribution can be obtained. Since our non-informative prior for q_p in the $\left(\mu,\sigma^2\right)$ parametrisation is $p\left(\mu,\sigma^2\right) \propto \sigma^{-2}$, it follows that the posterior distribution of σ^2 has an inverse gamma distribution which means that $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$ and $\mu|\sigma^2, data \sim N\left(\bar{X},\frac{\sigma^2}{n}\right)$.

The posterior distribution of q_p is therefore equal to

$$\bar{X} + \frac{Z+z_p\sqrt{n}}{U}\frac{S}{\sqrt{n}} = \bar{X} + \frac{1}{\sqrt{n}}t_{n-1}\left(z_p\sqrt{n}\right)S,$$

where $Z \sim N(0,1)$ and independently distributed of $U^2 \sim \frac{\chi^2_{n-1}}{n-1}$. Thus a $(p,1-\alpha)$ upper tolerance limit is given by

$$\bar{X} + k_1 S = \bar{X} + t_{n-1;1-\alpha}\left(z_p\sqrt{n}\right)\frac{S}{\sqrt{n}},$$

where $t_{n-1;1-\alpha}\left(z_p\sqrt{n}\right)$ denotes the $1-\alpha$ quantile of a non-central t-distribution with $n-1$ degrees of freedom and non-centrality parameter $z_p\sqrt{n}$. $\bar{X}+k_1S$ is an exact tolerance limit (i.e., has the correct coverage probability) and as mentioned by Krishnamoorthy and Mathew (2009) is the same solution that is obtained by the frequentist approach. The tolerance factor k_1, which is derived from the non-central t-distribution, can be obtained from Table B1 in Krishnamoorthy and Mathew (2009).

In this paper we are firstly interested in the predictive distribution of a future sample one-sided upper tolerance limit. By using the predictive distribution a Bayesian procedure will be developed to obtain control limits for a future sample one-sided upper tolerance limit. Assuming that the process remains stable, the predictive distribution can be used to derive the distribution of the "run length" and "average run length".

5. A Future Sample One-sided Upper Tolerance Limit

Consider a future sample of m observations from the $N\left(\mu,\sigma^2\right)$ population: $X_{1f}, X_{2f}, \ldots, X_{mf}$. The future sample mean is defined as $\bar{X}_f = \frac{1}{m}\sum_{j=1}^{m} X_{jf}$ and a future sample variance by $S_f^2 = \frac{1}{m-1}\sum_{j=1}^{m}\left(X_{jf}-\bar{X}_f\right)^2$.

A $(p, 1-\alpha)$ upper tolerance limit for the future sample is defined as

$$\tilde{q} = \bar{X}_f + \tilde{k}_1 S_f,$$

where

$$\tilde{k}_1 = \frac{1}{\sqrt{m}} t_{m-1;1-\alpha}\left(z_p\sqrt{m}\right).$$

Although the posterior predictive distribution of $\tilde{q}$ can easily be obtained by simulation, the exact mean and variance can be derived analytically. The following theorem can now be proved.

Theorem 3 The exact mean and variance of $\tilde{q} = \bar{X}_f + \tilde{k}_1 S_f$ are given by

$$E(\tilde{q}|data) = \bar{X} + \tilde{k}_1 \frac{\Gamma\left(\frac{m}{2}\right)}{\Gamma\left(\frac{m-1}{2}\right)} \frac{\Gamma\left(\frac{n-2}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)} \frac{\sqrt{n-1}}{\sqrt{m-1}} S$$

and

$$Var(\tilde{q}|data) = \left(\frac{m+n}{nm}\right)\left(\frac{n-1}{n-3}\right) + \tilde{k}_1^2 \left\{\frac{n-1}{n-3} - \frac{\Gamma^2\left(\frac{m}{2}\right)\Gamma^2\left(\frac{n-2}{2}\right)(n-1)}{\Gamma^2\left(\frac{m-1}{2}\right)\Gamma^2\left(\frac{n-1}{2}\right)(m-1)}\right\} S^2.$$

Proof. The proof is given in the Appendix. ■

Corollary: If $m = n$, then

$$E(\tilde{q}|data) = \bar{X} + \tilde{k}_1 \frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n-2}{2}\right)}{\Gamma^2\left(\frac{n-1}{2}\right)} S$$

and

$$Var(\tilde{q}|data) = \frac{2}{n}\left(\frac{n-1}{n-3}\right) + \tilde{k}_1^2 \left(\frac{n-1}{n-3} - \frac{\Gamma^2\left(\frac{n}{2}\right)\Gamma^2\left(\frac{n-2}{2}\right)}{\Gamma^4\left(\frac{n-1}{2}\right)}\right) S^2.$$

6. The Predictive Distribution of $\tilde{q} = \bar{X}_f + \tilde{k}_1 S_f$

As mentioned previously, the posterior predictive distribution of $\tilde{q}$ can easily be simulated. This can be done in the following way:

$$\tilde{q}|\sigma^2, S_f^2, data \sim N\left(\bar{X} + \tilde{k}_1 S_f, \sigma^2\left(\frac{1}{m} + \frac{1}{n}\right)\right).$$

Therefore,

$$f\left(\tilde{q}|\sigma^2, S_f^2, data\right) = \left(\frac{mn}{\sigma^2(m+n)2\pi}\right)^{\frac{1}{2}} \exp\left\{-\frac{mn}{2\sigma^2(m+n)}\left[\tilde{q} - \left(\bar{X} + \tilde{k}_1 S_f\right)\right]^2\right\}. \quad (2)$$

The unconditional predictive distribution can be obtained by first simulating σ^2 and then S_f. From the posterior distribution it follows that $\sigma^2 \sim \frac{(n-1)S^2}{\chi^2_{n-1}}$ and given σ^2, $S_f \sim \left\{\frac{\sigma^2\chi^2_{m-1}}{m-1}\right\}^{\frac{1}{2}}$. Substitute the simulated σ^2 and S_f values in (2) and draw the normal density function. Repeat the procedure l times and average the l simulated normal density functions (Rao-Blackwell method) to obtain the unconditional predictive density function $f(\tilde{q}|data)$. In the example that follows $l = 100,000$.

7. Example

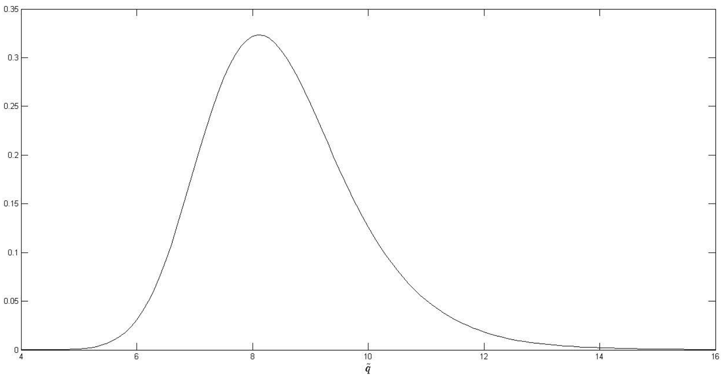
According to Krishnamoorthy and Mathew (2009) one-sided upper tolerance limits can commonly be used to assess the pollution level in a work place or in a region. The data in Table 1 represent air lead levels collected by the National Institute of Occupational Safety and Health at a laboratory, for health hazard evaluation. The air lead levels were collected from $n = 15$ different areas within the facility.

Table 1: Air lead levels $\left(\mu g/m^3\right)$.

200	120	15	7	8	6	48	61
380	80	29	1,000	350	1,400	110	

A normal distribution fitted the log-transformed lead levels quite well. The sample mean and standard deviation of the log-transformed data are calculated as $\bar{X} = 4.3329$ and $S = 1.7394$, respectively. For $n = 15$, $1-\alpha = 0.90$, $p = 0.95$ and using the non-central t-distribution in MATLAB, $k_1 = \frac{1}{\sqrt{n}} t_{n-1,1-\alpha}\left(z_p\sqrt{n}\right)$ =2.3290. A $(0.95, 0.90)$ upper tolerance limit for the air lead level is $\bar{X} + k_1 S = 8.3840$.

In this paper we are however interested in the predictive distribution of $\tilde{q} = \bar{X}_f + \tilde{k}_1 S_f$, the tolerance limit for a future sample of $m = n = 15$ observations. Using the simulated procedure described in Section 6, the predictive distribution is illustrated in Figure 1.



Mean = 8.5214, Mode = 8.2741

Figure 1: Predictive density function of a future tolerance limit $\tilde{q} = \bar{X}_f + \tilde{k}_1 S_f$.

The mean of the predictive distribution of $\tilde{q}$ is somewhat larger and the mode somewhat smaller than 8.384, the sample upper tolerance limit of the air lead level.

In Table 2 it is shown that the calculated means and variances from the simulation and formulae are, for all practical purposes, the same.

Table 2: Mean and variance of $\tilde{q}$.

	$E(\tilde{q}\|data)$	$Var(\tilde{q}\|data)$
From simulated $\tilde{q}$	8.5214	1.8981
Using Formulae	8.5427	1.8950

In Table 3 confidence limits for $\tilde{q}$ are given.

Table 3: Prediction limits for $\tilde{q}$.

	95% Left One-sided	95% Right One-sided	95% Two-sided
Left Limit	6.5683	-	6.2421
Right Limit	-	11.0320	11.6827

8. Control Chart for a Future One-sided Upper Tolerance Limit

Statistically, quality control is actually implemented in two phases. In Phase I the primary interest is to assess process stability. The practitioner must therefore be sure that the process is in statistical control before control limits can be determined for online monitoring in Phase II.

By using the predictive distribution a Bayesian procedure will be developed to obtain a control chart for a future one-sided upper tolerance limit. Assuming the process remains stable, the predictive distribution can be used to derive the distribution of the "run-length" and average "run-length". From Figure 1 it follows a 99.73% upper control limit for $\tilde{q} = \bar{X}_f + \tilde{k}_1 S_f$ is 13.7. Therefore the rejection region of size β $(\beta = 0.0027)$ for the predictive distribution is

$$\beta = \int_{R(\beta)} f(\tilde{q}|data)\, d\tilde{q},$$

where $R(\beta)$ represents those values of $\tilde{q}$ that are larger than 13.7.

The "run-length" is defined as the number of future $\tilde{q}$ values (r) until the control chart signals for the first time (Note that r does not include that $\tilde{q}$ value when the control chart signals). Given μ and σ^2 and a stable Phase I process, the distribution of the "run-length" r is geometric with parameter

$$\psi\left(\mu, \sigma^2\right) = \int_{R(\beta)} f\left(\tilde{q}|\mu, \sigma^2\right) d\tilde{q},$$

where $f\left(\tilde{q}|\mu, \sigma^2\right)$ is the distribution of a future $\tilde{q}$ given that μ and σ^2 are known. The values of μ and σ^2 are however unknown and the uncertainty of these parameter values is described by their joint posterior distribution $p\left(\mu, \sigma^2|data\right)$. By simulating μ and σ^2 from $p\left(\mu, \sigma^2|data\right)$, the probability density function $f\left(\tilde{q}|\mu, \sigma^2\right)$ (for the charting statistic $\tilde{q}$) can be obtained in the following way:

1. $\tilde{q}|\mu, \sigma^2, \chi^2_{m-1} \sim N\left(\mu + k_1 \sigma \frac{\sqrt{\chi^2_{m-1}}}{\sqrt{m-1}}, \frac{\sigma^2}{m}\right).$

2. The next step is to simulate $l = 100,000$ χ^2_{m-1} values to obtain l normal density functions for given μ and σ^2.

3. By averaging the l density functions (Rao-Blackwell method), $f\left(\tilde{q}|\mu,\sigma^2\right)$ can be obtained.

This must be done for each future sample. In other words, for each future sample μ and σ^2 must first be simulated from $p\left(\mu,\sigma^2|data\right)$ and then the procedure described in steps 1, 2 and 3.

As mentioned, the "run-length" r (given μ and σ^2) is geometrically distributed with mean

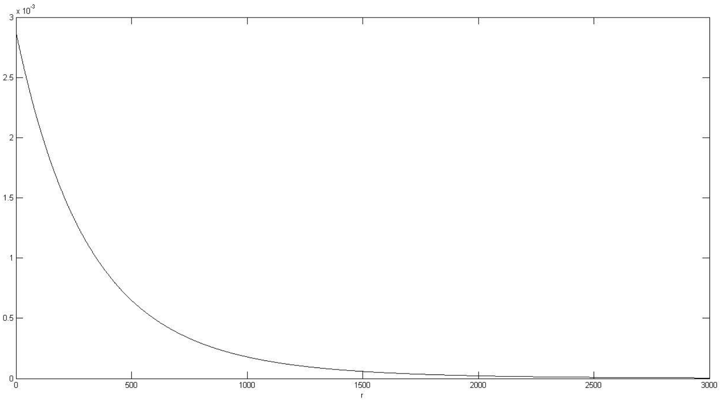
$$E\left(r|\mu,\sigma^2\right) = \frac{1-\psi\left(\mu,\sigma^2\right)}{\psi\left(\mu,\sigma^2\right)}$$

and variance

$$Var\left(r|\mu,\sigma^2\right) = \frac{1-\psi\left(\mu,\sigma^2\right)}{\psi^2\left(\mu,\sigma^2\right)}.$$

The unconditional moments $E\left(r|data\right)$, $E\left(r^2|data\right)$ and $Var\left(r|data\right)$ can therefore easily be obtained by simulation or numerical integration. For further details see Menzefricke (2002, 2007, 2010a, 2010b) .

In Figure 2 the predictive distribution of the "run-length" is displayed for the 99.73% upper control limit. As mentioned, for given μ and σ^2, the "run-length" r is geometric with parameter $\psi\left(\mu,\sigma^2\right)$. The unconditional "run-length" as given in Figure 2 is therefore obtained using the Rao-Blackwell method, i.e., the average of a large number of conditional "run-lengths".



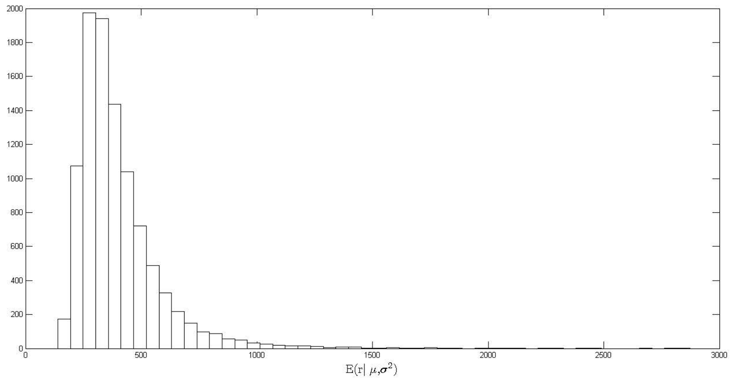
$E\left(r|data\right) = 396.27438, Median\left(r|data\right) = 251, Var\left(r|data\right) = 2.0945e^5$
95% Equal-tail Interval = $(8; 1644)$ Length = 1636
95% HPD Interval = $(3; 1266)$ Length = 1263

Figure 2: Predictive distribution of the "run-length" $f\left(r|data\right)$ for $n = m = 15$.

In Figure 3 the distribution of the average "run-length" is given.

For known μ and σ the expected "run-length" is $\frac{1}{0.0027} = 370$. If μ and σ^2 are unknown and estimated from the posterior distribution the expected "run-length" will usually be larger than 370 - especially if the sample size is small.

A control chart for an Upper Tolerance Limit can therefore be used to monitor the mean and standard deviation simultaneously.



$Mean = 400.1084, Median = 353.5753, Var = 3.5605e^4$

95% Equal-tail Interval = (202.1834; 874.6607) Length = 672.4773

95% HPD Interval = (163.4804; 566.0842) Length = 402.6038

Figure 3: Distribution of the average "run-length".

9. One-sided Tolerance Limits for the Difference Between Two Normal Populations

Let $X_1 \sim N\left(\mu_1, \sigma_1^2\right)$ and independently distributed from $X_2 \sim N\left(\mu_2, \sigma_2^2\right)$ where μ_1 and μ_2 are the unknown population means and σ_1^2 and σ_2^2 the unknown population variances. If z_p is the pth quantile of a standard normal distribution then

$$L_p = \mu_1 - \mu_2 - z_p\sqrt{\sigma_1^2 + \sigma_2^2} \tag{3}$$

is the $1-p$ quantile of the distribution of $X_1 - X_2$. As mentioned by Krishnamoorthy and Mathew (2009), a $(p, 1-\alpha)$ lower tolerance limit for the distribution of $X_1 - X_2$ is a $1-\alpha$ lower confidence limit of L_p.

Suppose $\bar{X}_i$ and S_i^2 are respectively the sample mean and variance of a random sample of n_i observations from $N\left(\mu_i, \sigma_i^2\right)$, $i = 1, 2$. By using exact distributional results Krishnamoorthy and Mathew (2009) constructed a $(1-\alpha)$ lower confidence limit for L_p as

$$\bar{X}_1 - \bar{X}_2 - t_{n_1+n_2-2, 1-\alpha}\left(z_p\sqrt{v_1}\right)\frac{S_d}{\sqrt{v_1}}, \tag{4}$$

where

$$v_1 = \frac{\sigma_1^2 + \sigma_2^2}{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

and

$$S_d^2 = \frac{\left(\sigma_1^2 + \sigma_2^2\right)}{n_1 + n_2 - 2}\left\{\frac{(n_1-1)S_1^2}{\sigma_1^2} + \frac{(n_2-1)S_2^2}{\sigma_2^2}\right\},$$

which is an exact $(p, 1-\alpha)$ lower tolerance limit for the distribution of $X_1 - X_2$ if the variance ratio is known. This is usually not the case, but the problem can be overcome by making use of the following "semi-Bayesian" approach.

The expression in (4) is derived from the fact that the pivotal quantity

$$\frac{\bar{X}_1 - \bar{X}_2 - L_p}{S_d} \sim \frac{Z + z_p\sqrt{v_1}}{\sqrt{v_1}\sqrt{\frac{\chi^2_{n_1+n_2-2}}{(n_1+n_2-2)}}} = N_2, \tag{5}$$

where $Z \sim N(0,1)$ and independent of $\chi^2_{n_1+n_2-2}$ (See equation [2.4.4] in Krishnamoorthy and Mathew (2009)).

From (5) it follows that

$$L_p|data, \chi^2_{n_1+n_2-2}, \sigma_1^2, \sigma_2^2 \sim N\left\{(\bar{X}_1 - \bar{X}_2) - \frac{z_p S_d}{\sqrt{\frac{\chi^2_{n_1+n_2-2}}{(n_1+n_2-2)}}}, \frac{S_d^2}{v_1 \frac{\chi^2_{n_1+n_2-2}}{(n_1+n_2-2)}}\right\}. \tag{6}$$

If the non-informative prior $p\left(\sigma_1^2, \sigma_2^2\right) \propto \sigma_1^{-2}\sigma_2^{-2}$ is used, then the posterior distribution of σ_i^2 $(i = 1,2)$ is an inverse-Gamma distribution, which means that

$$\sigma_i^2 \sim \frac{(n_i - 1)S_i^2}{\chi^2_{n_i-1}} \quad (i = 1,2),$$

and $p(L_p|data)$ — the unconditional posterior distribution of L_p — can be obtained by using the following simulation procedure:

1. Simulate σ_i^2 $(i = 1,2)$ and substitute the simulated variance components in S_d^2 and v_1.
2. Simulate $\chi^2_{n_1+n_2-2}$ and substitute the value in (6).
3. Draw the normal density function.
4. Repeat the procedure l times and average the l simulated normal density functions (Rao-Blackwell method) to obtain $p(L_p|data)$.

$l = 100,000$ is used in the example that follows.

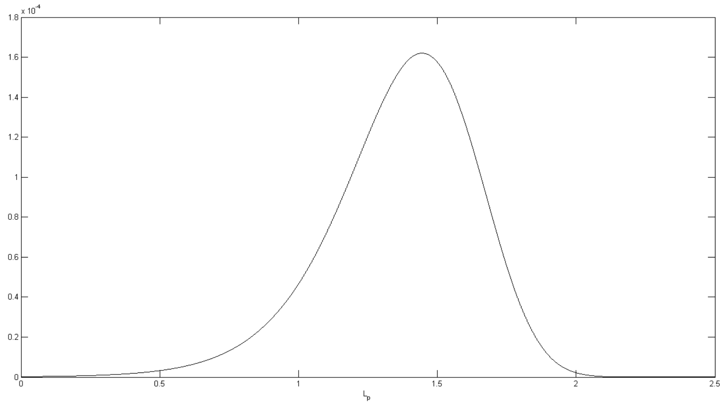
10. Example - Breakdown Voltage - Power Supply

This example is taken from Krishnamoorthy and Mathew (2009) and it is concerned with the proportion of times the breakdown voltage of X_1 of a capacitor exceeds the voltage output X_2 of a transverter (power supply). A sample of $n_1 = 50$ capacitors yielded $\bar{X}_1 = 6.75kV$ and $S_1^2 = 0.123$. The voltage output from $n_2 = 20$ transverters produced $\bar{X}_2 = 400kV$ and $S_2^2 = 0.53$.

Using the simulation procedure as described the unconditional posterior distribution $p(L_p|data)$ is obtained and illustrated in Figure 4 for $z_p = z_{0.95} = 1.6449$.

A $(0.95, 0.95)$ lower tolerance limit for the distribution of $X_1 - X_2$ is therefore $q_p = 0.8895$.

In what follows it will be shown that a more formal Bayesian procedure gives for all practical purposes the same estimate for the lower tolerance limit.



$p\left(L_p < 0.8895\right) = 0.05$

$Mean = 1.3671$, $Median = 1.3954$, $Mode = 1.4448$, $Var = 0.0714$

Figure 4: $p\left(L_p|data\right)$ – unconditional posterior distribution of L_p.

If the objective prior $p\left(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2\right) \propto \sigma_1^{-2}\sigma_2^{-2}$ is used then it is well known that the posterior distribution of L_p (see (3)) given the variance components is normal with mean

$$E\left(L_p^B|\sigma_1^2,\sigma_2^2,data\right) = \bar{X}_1 - \bar{X}_2 - z_p\sqrt{\sigma_1^2+\sigma_2^2}$$

and variance

$$Var\left(L_p^B|\sigma_1^2,\sigma_2^2,data\right) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}.$$

Also,

$$\sigma_i^2 \sim \frac{\left(n_i-1\right)S_i^2}{\chi^2_{n_1-1}}. \quad \left(i=1,2\right).$$

From this, the posterior distribution follows

$$L_p^B|data,U_1,U_2 \sim N\left\{\left(\bar{X}_1-\bar{X}_2\right) - z_p\sqrt{\frac{S_1^2}{U_1^2}+\frac{S_2^2}{U_2^2}}, \frac{S_1^2}{n_1U_1^2}+\frac{S_2^2}{n_2U_2^2}\right\}, \tag{7}$$

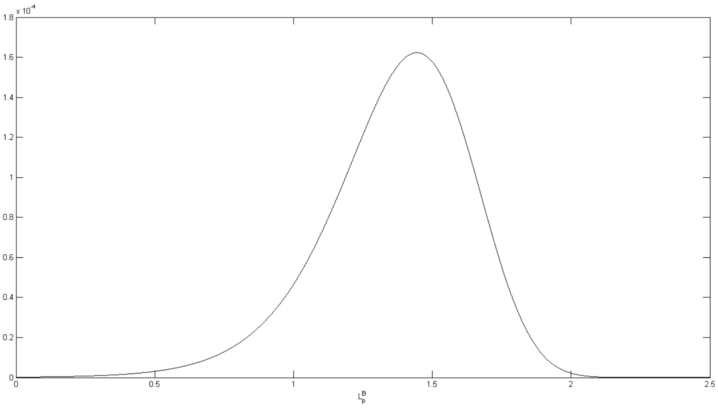
where

$$U_i^2 \sim \frac{\chi^2_{n_i-1}}{n_i-1}.$$

The expression in (7) is exactly the same as the generalised variable approach described by Krishnamoorthy and Mathew (2009).

By simulating 100,000 $U_i^2 \sim \frac{\chi^2_{n_i-1}}{n_1-1}$ $\left(i=1,2\right)$ values and using the Rao-Blackwell method the unconditional posterior distribution of L_p^B is obtained and given in Figure 5.

Since the two methods give for all practical purposes the same estimate for the lower tolerance limit, only the first method will be used to obtain the predictive distribution of a future one-sided tolerance limit.



$p\left(L_p^B < 0.88861\right) = 0.05$

$Mean = 1.3682$, $Median = 1.3964$, $Mode = 1.4462$, $Var = 0.0713$

Figure 5: $p\left(L_p^B | data\right)$.

11. The Predictive Distribution of a Future One-sided Lower Tolerance Limit for $X_1 - X_2$

By using the predictive distribution a Bayesian procedure will be developed to obtain a control chart for a future one-sided lower tolerance limit. Assuming the process remains stable, the predictive distribution can be used to derive the distribution of the "run-length" and average "run-length".

Consider a future sample of m_1 observations from the $N\left(\mu_1, \sigma_1^2\right)$ population $X_{1f}^{(1)}, X_{2f}^{(1)}, \ldots, X_{m_1 f}^{(1)}$. The future sample mean is defined as $\bar{X}_f^{(1)} = \frac{1}{m_1}\sum_{j=1}^{m_1} X_{jf}^{(1)}$ and a future sample variance by $S_f^{2\,(1)} = \frac{1}{m_1 - 1}\sum_{j=1}^{m_1}\left(X_{jf}^{(1)} - \bar{X}_f^{(1)}\right)^2$. Similar for a future sample of m_2 observations from $N\left(\mu_2, \sigma_2^2\right)$, $\bar{X}_f^{(2)} = \frac{1}{m_2}\sum_{j=1}^{m_2} X_{jf}^{(2)}$ and $S_f^{2\,(2)} = \frac{1}{m_2 - 1}\sum_{j=1}^{m_2}\left(X_{jf}^{(2)} - \bar{X}_f^{(2)}\right)$.

Thus using the exact distributional results a $(p, 1-\alpha)$ lower tolerance limit for the difference between two future samples is defined as

$$q_f = \bar{X}_f^{(1)} - \bar{X}_f^{(2)} - t_{m_1 + m_2 - 2; 1-\alpha}\left(z_p \sqrt{v_1^f}\right)\frac{S_d^f}{\sqrt{v_1^f}}. \tag{8}$$

If the sample sizes of the two future samples are the same, i.e., $m_1 = m_2 = m$, then (8) simplifies to

$$\begin{aligned} q_f &= \bar{X}_f^{(1)} - \bar{X}_f^{(2)} - t_{2(m-1);1-\alpha}\left(z_p\sqrt{m}\right)\frac{S_d^f}{\sqrt{m}} \\ &= \bar{X}_f^{(1)} - \bar{X}_f^{(2)} - \tilde{\tilde{k}} S_d^f, \end{aligned}$$

where

$$\tilde{\tilde{k}} = \frac{1}{\sqrt{m}} t_{2(m-1);1-\alpha}\left(z_p\sqrt{m}\right),$$

$$v_1^f = m,$$

and

$$S_d^{2\,f} = \frac{\left(\sigma_1^2 + \sigma_2^2\right)}{2}\left\{\frac{S_f^{2\,(1)}}{\sigma_1^2} + \frac{S_f^{2\,(2)}}{\sigma_2^2}\right\}.$$

We are interested in the predictive distribution of q_f. This cannot be derived analytically, but can easily be obtained by simulation.

Now,

$$q_f|\sigma_1^2, \sigma_2^2, S_f^{2\,(1)}, S_f^{2\,(2)}, \mu_1, \mu_2 \sim N\left\{(\mu_1 - \mu_2) - \tilde{\tilde{k}} S_d^f, \frac{1}{m}\left(\sigma_1^2 + \sigma_2^2\right)\right\},$$

since

$$\bar{X}_f^{(1)} - \bar{X}_f^{(2)} \sim N\left\{\mu_1 - \mu_2, \frac{1}{m}\left(\sigma_1^2 + \sigma_2^2\right)\right\}.$$

Also if the prior $p\left(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2\right) \propto \sigma_1^{-2}\sigma_2^{-2}$ is used,

$$(\mu_1 - \mu_2)\,|data, \sigma_1^2, \sigma_2^2 \sim N\left\{\bar{X}_1 - \bar{X}_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right\}.$$

Therefore

$$q_f|\sigma_1^2, \sigma_2^2, S_f^{2\,(1)}, S_f^{2\,(2)}, data \sim N\left\{(\bar{X}_1 - \bar{X}_2) - \tilde{\tilde{k}} S_d^f, \frac{1}{m}\left(\sigma_1^2 + \sigma_2^2\right) + \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right\}. \tag{9}$$

The unconditional predictive distribution of q_f can be obtained by first simulating σ_i^2 and then $S_f^{2\,(i)}$ $(i = 1, 2)$. As before it follows from the posterior distribution that

$$\sigma_i^2 \sim \frac{(n_i - 1)\,S_i^2}{\chi^2_{n_i - 1}} \quad (i = 1, 2)$$

and given σ_i^2

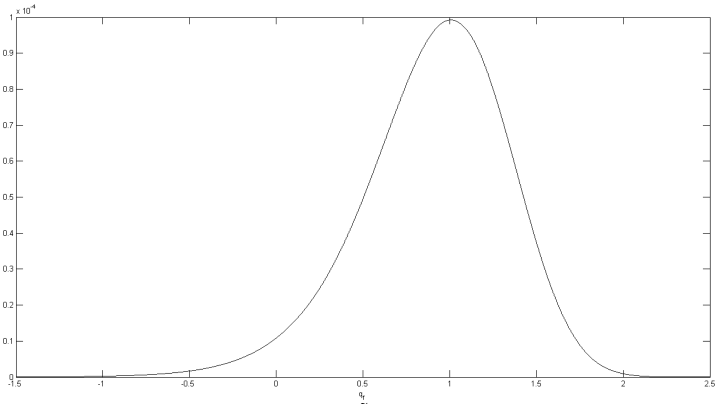
$$S_f^{2\,(i)} \sim \frac{\sigma_i^2 \chi_{m-1}^{2\,(i)}}{m - 1} \quad (i = 1, 2).$$

Substitute the simulated σ_i^2 and $S_f^{2\,(i)}$ values in (9) and draw the normal density function. Repeat the procedure l times and average the l simulated density functions (Rao-Blackwell method) to obtain the unconditional predictive density function $f\left(q_f|data\right)$. The predictive density function for the Breakdown Voltage example given in Section 10 for $m_1 = m_2 = 20$ is illustrated in Figure 6.

The mean of the predictive distribution of q_f is 0.9060 and is somewhat larger than 0.08895, the sample lower tolerance limit for the break down voltage data.

12. Control Chart for a Future One-sided Lower Tolerance Limit

As before, a Bayesian procedure will be developed to obtain a control chart for a future one-sided lower tolerance limit. Assuming the process remains stable, the predictive distribution can be used to derive the distribution of the "run-length" and average "run-length".



$\alpha = 0.05, \tilde{\tilde{k}} = 2.2096$
$Mean = 0.9060, Median = 0.9427, Mode = 1.0081, Var = 0.1834$
95% Equal-tail Interval $= (-0.0417; 1.6406)$
95% HPD Interval $= (0.0081; 1.6703)$
99.73% Equal-tail Interval $= (-0.8068; 1.9614)$
$p(q_f \leq -0.6258) = 0.0027$
$p(q_f \geq 1.8955) = 0.0027$

Figure 6: $f(q_f|data)$ – predictive density of the lower tolerance limit, q_f.

From Figure 6 it follows that for a 99.73% two-sided control chart the lower control limit is $LCL = -0.8068$ and the upper control limit is $UCL = 1.9614$. Therefore the rejection region of size β $(\beta = 0.0027)$ for the predictive distribution is

$$\beta = \int_{R(\beta)} f(q_f|data)\, dq_f$$

where $R(\beta)$ represents those values of q_f that are smaller than LCL and larger than UCL.

As mentioned, the "run-length" is defined as the number of future q_f values (r) until the control chart signals for the first time. Given μ_i, σ_i^2 $(i = 1,2)$, the distribution of the "run-length" r is geometric with parameter

$$\psi(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2) = \int_{R(\beta)} f(q_f|\mu_1, \mu_2, \sigma_1^2, \sigma_2^2)\, d_{qf}$$

where $f(q_f|\mu_1, \mu_2, \sigma_1^2, \sigma_2^2)$ is the distribution of q_f given that $\mu_{i,}\sigma_i^2$ $(i = 1,2)$ are known. These parameter values are however unknown but can be simulated from their joint posterior distribution $p(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2|data)$ as follows:

$$\sigma_i^2|data \sim \frac{(n_i - 1) S_i^2}{\chi_{n_i-1}^2} \quad (i = 1,2)$$

and

$$\mu_i|\sigma_i^2, data \sim N\left(\bar{X}_i, \frac{\sigma_i^2}{n_1}\right) \quad (i = 1,2).$$

1. From this it follows that

$$q_f|\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, S_f^{2\,(1)}, S_f^{2\,(2)} \sim N\left\{(\mu_1 - \mu_2) - \tilde{\tilde{k}} S_d^{(f)}, \frac{1}{m}\left(\sigma_1^2 + \sigma_2^2\right)\right\}.$$

2. The next step is to simulate $l = 100,000$ $S_f^{2\,(i)}$ values for given σ_i^2 $(i = 1,2)$. $S_f^{2\,(i)} \sim \frac{\sigma_i^2 \chi_{m-1}^{2\,(i)}}{m-1}$ $(i = 1,2)$.

3. By averaging the l density functions (Rao-Blackwell method), $f\left(q_f|\mu_1, \mu_2, \sigma_1^2, \sigma_2^2\right)$ can be obtained.

This must be done for each future sample. In other words for each future sample μ_i and σ_i^2 $(i = 1,2)$ must first be simulated from their joint posterior distribution $p\left(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2|data\right)$ and then the procedure described in steps 1, 2 and 3 follows.

The mean of the predictive distribution of the "run-length" for the 99.73% two-sided control limits is $E(r|data) = 201500$, much larger than the 370 that one would have expected if $\beta = 0.0027$. The reason for this large average "run-length" is mainly the small future sample sizes $(m_1 = m_2 = 20)$. The median "run-length" is 3770. Define $\bar{\psi}(\mu_1, \mu_2, \sigma_1, \sigma_2) = \frac{1}{\tilde{m}} \sum_{i=1}^{\tilde{m}} \psi\left(\mu_1^{(i)}, \mu_2^{(i)}, \sigma_1^{2(i)}, \sigma_2^{2\,(i)}\right)$. According to Menzefricke (2002), $\bar{\psi}\left(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2\right) \to \beta$ if $\tilde{m} \to \infty$. The harmonic mean of $r = \frac{1}{\beta}$ and if $\beta = 0.0027$, the harmonic mean = $\frac{1}{0.0027} = 370$.

In Table 4 the average "run-length" for different values of β is given.

Table 4: β Values and corresponding average "run-length".

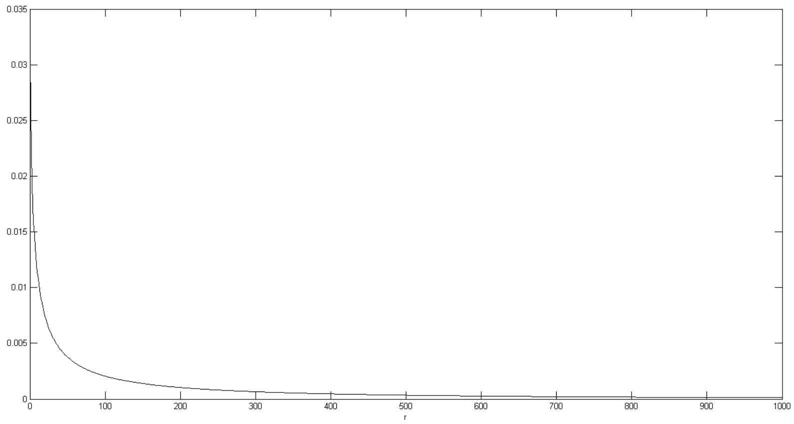
β	0.0027	0.012	0.016	0.02	0.024	0.028	0.03	0.032	0.036	0.04
$E(r\|data)$	201500	4044	1785	959	633	438	400	340	240	218

In Figure 7 the distribution of the "run-length" for $m_1 = m_2 = 20$ and $\beta = 0.03$ is illustrated and in Figure 8 the histogram of the expected "run-length" is given.

13. Conclusion

In the first part of this article a Bayesian control chart was developed for monitoring an upper one-sided tolerance limit across a range of sample values. In the Bayesian approach prior knowledge about the unknown parameters is formally incorporated into the process of inference by assigning a prior distribution to the parameters. The information contained in the prior is combined with the likelihood function to obtain the posterior distribution. By using the posterior distribution the predictive distribution of an upper one-sided tolerance limit can be obtained.

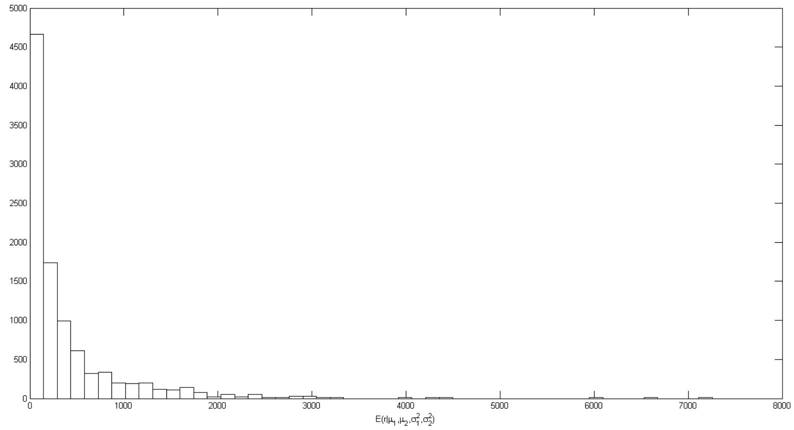
Determination of reasonable non-informative priors in multi-parameter problems is not an easy task. The Jeffreys' prior can, for example, have a bad effect on the posterior distribution. Reference and probability matching priors are therefore derived for the pth quantile of a normal distribution. The theory and results have been applied to air-lead level data analysed by Krishnamoorthy and



$E\left(r|data\right) = 400.41, Median\left(r|data\right) = 79.51, Var\left(r|data\right) = 1.8668e^{6}$

95% HPD Interval = $(0;1688)$

Figure 7: Predictive distribution of the "run-length" $f\left(r|data\right)$ for $m_1 = m_2 = m = 20$.



$Mean = 400.41, Median = 138.79, Var = 7.0775e^{5}$

95% HPD Interval = $(1.11;1552.1)$

Figure 8: Distribution of the average "run-length".

Mathew (2009) to illustrate the flexibility and unique features of the Bayesian simulation method for obtaining posterior distributions, prediction intervals and run lengths.

In the second part of this article the Bayesian procedure was extended to control charts of one-sided tolerance limits for a distribution of the difference between two independent normal variables.

Acknowledgements

We would like to thank the anonymous referee and the editor whose contributions improved and made this publication possible.

References

BAYARRI, M. J. AND BERGER, J. O. (2004). The interplay of Bayesian and frequentist analysis. *Statistical Science*, **19** (1), 58–80.

BAYARRI, M. J. AND GARCÍA-DONATO, G. (2005). A Bayesian sequential look at u-control charts. *Technometrics*, **47** (2), 142–151.

BERGER, J. O. AND BERNARDO, J. M. (1992). On the development of reference priors. *In* BERNARDO, J. M., BERGER, J. O., DAWID, A. P., AND SMITH, A. F. M. (Editors) *Bayesian Statistics 4*. Oxford University Press: New York, pp. 35–60.

DATTA, G. S. AND GHOSH, J. K. (1995). On priors providing frequentist validity for Bayesian inference. *Biometrika*, **82** (1), 37–45.

HAHN, G. AND MEEKER, W. (1991). *Statistical Intervals: A Guide for Practitioners*. John Wiley & Sons, Inc.: New Jersey. doi:10.1002/9780470316771.

HAMADA, M. (2002). Bayesian tolerance interval control limits for attributes. *Quality and Reliability Engineering International*, **18** (1), 45–52.

HUGO, J. (2012). *Bayesian Tolerance Intervals for Variance Component Models*. Ph.D. thesis, University of the Free State, South Africa.

KRISHNAMOORTHY, K. AND MATHEW, T. (2009). *Statistical Tolerance Regions: Theory, Applications and Computation*. John Wiley & Sons, Inc.: Hoboken, New Jersey. doi: 10.1002/9780470473900.ch11.

MENZEFRICKE, U. (2002). On the evaluation of control chart limits based on predictive distributions. *Communications in Statistics – Theory and Methods*, **31** (8), 1423–1440.

MENZEFRICKE, U. (2007). Control chart for the generalized variance based on its predictive distribution. *Communications in Statistics – Theory and Methods*, **36** (5), 1031–1038.

MENZEFRICKE, U. (2010a). Control chart for the variance and the coefficient of variation based on their predictive distribution. *Communications in Statistics – Theory and Methods*, **39** (16), 2930–2941.

MENZEFRICKE, U. (2010b). Multivariate exponentially weighted moving average chart for a mean based on its predictive distribution. *Communications in Statistics – Theory and Methods*, **39** (16), 2942–2960.

PEARN, W. L. AND WU, C. W. (2005). A Bayesian approach for assessing process precision based on multiple samples. *European Journal of Operational Research*, **165** (3), 685–695.

SEVERINE, T. A., MUKERJEE, R., AND GHOSH, M. (2002). On an exact probability matching property of right-invariant priors. *Biometrika*, **89** (4), 952–957.

TIBSHIRANI, R. (1989). Noninformative priors for one parameter of many. *Biometrika*, **76** (3), 604–608.

Appendix

Theorem 1

Assume $X_i\,(i=1,2,\ldots,n)$ are independently and identically normally distributed with mean μ and variance σ^2. The Fisher information matrix for the parameter vector $\underline{\theta}=\left[\mu,\sigma^2\right]'$ is given by

$$F\left(\mu,\sigma^2\right)=\begin{bmatrix}\frac{n}{\sigma^2} & 0\\ 0 & \frac{n}{2\left(\sigma^2\right)^2}\end{bmatrix}.$$

Let $q_p=\mu+z_p\sigma=t\left(\mu,\sigma^2\right)=t\left(\underline{\theta}\right)$.

To obtain the reference prior, the Fisher information matrix $F\left(t\left(\underline{\theta}\right),\sigma\right)$ must first be derived.

Let

$$A=\begin{bmatrix}\frac{\partial\mu}{\partial t(\underline{\theta})} & \frac{\partial\mu}{\partial\sigma^2}\\ \frac{\partial\sigma^2}{\partial t(\underline{\theta})} & \frac{\partial\sigma^2}{\partial\sigma^2}\end{bmatrix}=\begin{bmatrix}1 & -\frac{1}{2}\frac{z_p}{\sigma}\\ 0 & 1\end{bmatrix}.$$

Now,

$$F\left(t\left(\underline{\theta}\right),\sigma^2\right)=A'F\left(\mu,\sigma^2\right)A=\begin{bmatrix}\frac{n}{\sigma^2} & -\frac{nz_p}{2\sigma^3}\\ -\frac{nz_p}{2\sigma^3} & \frac{nz_p^2}{4\sigma^4}+\frac{n}{2\sigma^4}\end{bmatrix}=\begin{bmatrix}F_{11} & F_{12}\\ F_{21} & F_{22}\end{bmatrix}$$

and the inverse

$$F^{-1}\left(t\left(\underline{\theta}\right),\sigma^2\right)=\frac{2\sigma^6}{n^2}\begin{bmatrix}\frac{n}{2\sigma^4}\left(\frac{z_p^2}{2}+1\right) & \frac{nz_p}{2\sigma^3}\\ \frac{nz_p}{2\sigma^3} & \frac{n}{\sigma^2}\end{bmatrix}=\begin{bmatrix}F^{11} & F^{12}\\ F^{21} & F^{22}\end{bmatrix}.$$

Therefore,

$$F^{11}=\frac{\sigma^2}{n}\left(\frac{z_p^2}{2}+1\right),$$

$$\left(F^{11}\right)^{-1}=\frac{n}{\sigma^2}\left(\frac{z_p^2}{2}+1\right)^{-1}=h_1$$

and

$$p\left(t\left(\underline{\theta}\right)\right)\propto h_1^{\frac{1}{2}}\propto \text{ constant because it does not contain } t\left(\underline{\theta}\right).$$

Further,

$$h_2=F_{22}=\frac{n}{2\sigma^4}\left(\frac{z_p^2}{2}+1\right)$$

and

$$p\left(\sigma^2|t\left(\underline{\theta}\right)\right)\propto h_2^{\frac{1}{2}}\propto\sigma^{-2}.$$

Therefore the reference prior for the ordering $\left\{t\left(\underline{\theta}\right),\sigma^2\right\}=\left\{q_p,\sigma^2\right\}$ is $P_R\left(q_p,\sigma^2\right)\propto\sigma^{-2}$.

In the (μ, σ^2) parametrisation this corresponds to $P_R(\mu, \sigma^2) = p(t(\underline{\theta}), \sigma^2)\left|\frac{\partial t(\underline{\theta})}{\partial \mu}\right|$.

Since $\left|\frac{\partial t(\underline{\theta})}{\partial \mu}\right| = 1$, it follows that $P_R(\mu, \sigma^2) \propto \sigma^{-2}$.

Theorem 2

Let

$$t(\underline{\theta}) = t(\mu, \sigma^2) = q_p$$

and

$$\nabla_t'(\underline{\theta}) = \left[\frac{\partial}{\partial q_p} t(\underline{\theta}) \quad \frac{\partial}{\partial \sigma^2} t(\underline{\theta})\right] = \begin{bmatrix}1 & 0\end{bmatrix}.$$

Also,

$$\nabla_t'(\underline{\theta}) F^{-1}\left(t(\underline{\theta}), \sigma^2\right) = \begin{bmatrix}F^{11} & F^{12}\end{bmatrix}$$

and

$$\sqrt{\nabla_t'(\underline{\theta}) F^{-1}(t(\underline{\theta}), \sigma^2) \nabla_t(\underline{\theta})} = \left(F^{11}\right)^{\frac{1}{2}}.$$

Further,

$$\Upsilon'(\underline{\theta}) = \frac{\nabla_t'(\underline{\theta}) F^{-1}\left(t(\underline{\theta}), \sigma^2\right)}{\sqrt{\nabla_t'(\underline{\theta}) F^{-1}(t(\underline{\theta}), \sigma^2) \nabla_t(\underline{\theta})}} = \begin{bmatrix}\Upsilon_1(\underline{\theta}) & \Upsilon_2(\underline{\theta})\end{bmatrix},$$

where

$$\Upsilon_1(\underline{\theta}) = \left(F^{11}\right)^{\frac{1}{2}} = \frac{\sigma}{\sqrt{n}}\left(\frac{z_p^2}{2} + 1\right)^{\frac{1}{2}}$$

and

$$\Upsilon_2(\underline{\theta}) = \frac{F^{12}}{\sqrt{F^{11}}} = \frac{\sigma^2 z_p}{\sqrt{n}}\left(\frac{z_p^2}{2} + 1\right)^{-\frac{1}{2}}.$$

According to Datta and Ghosh (1995) a prior $P_M(\underline{\theta}) = P_M(q_p, \sigma^2)$ will be a probability matching prior if the following differential equation is satisfied

$$\frac{\partial}{\partial q_p}\left\{\Upsilon_1(\underline{\theta}) P_M(\underline{\theta})\right\} + \frac{\partial}{\partial \sigma^2}\left\{\Upsilon_2(\underline{\theta}) P_M(\underline{\theta})\right\} = 0.$$

It is therefore clear that if $P_M(\underline{\theta}) \propto \sigma^{-2}$ the differential equation is satisfied.

Theorem 3

It is well known that if $Y \sim \chi_u^2$, then

$$E(Y^r) = \frac{2^r \Gamma\left(\frac{u}{2} + r\right)}{\Gamma\left(\frac{u}{2}\right)}.$$

Also since $\bar{X}_f|\mu,\sigma^2 \sim N\left(\mu, \frac{\sigma^2}{m}\right)$ and $S_f \sim \left\{\frac{\sigma^2\chi^2_{m-1}}{m-1}\right\}^{\frac{1}{2}}$ for given σ^2, it follows that

$$E\left(\tilde{q}|\mu,\sigma^2\right) = \mu + \frac{\tilde{k}_1\sqrt{2}\sigma}{\sqrt{m-1}}\frac{\Gamma\left(\frac{m}{2}\right)}{\Gamma\left(\frac{m-1}{2}\right)}$$

and

$$E\left(\tilde{q}^2|\mu,\sigma^2\right) = \mu^2 + 2\tilde{k}_1\mu\frac{\sigma\sqrt{2}\Gamma\left(\frac{m}{2}\right)}{\sqrt{m-1}\Gamma\left(\frac{m-1}{2}\right)} + \sigma^2\left(\frac{1}{m} + \tilde{k}_1^2\right).$$

From the posterior distribution it follows that $\mu|\sigma^2, data \sim N\left(\bar{X}, \frac{\sigma^2}{n}\right)$ and $\sigma \sim \left\{\frac{(n-1)S^2}{\chi^2_{n-1}}\right\}$ given the data. Therefore,

$$E\left(\tilde{q}|data\right) = \bar{X} + \tilde{k}_1\sqrt{\frac{n-1}{m-1}}\frac{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n-2}{2}\right)}{\Gamma\left(\frac{m-1}{2}\right)\Gamma\left(\frac{n-1}{2}\right)}S \tag{10}$$

and

$$E\left(\tilde{q}^2|data\right) = \bar{X}^2 + 2\tilde{k}_1\bar{X}\sqrt{\frac{n-1}{m-1}}\frac{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n-2}{2}\right)}{\Gamma\left(\frac{m-1}{2}\right)\Gamma\left(\frac{n-1}{2}\right)}S + \left(\frac{1}{n} + \frac{1}{m} + \tilde{k}_1^2\right)\left(\frac{n-1}{n-3}\right)S \tag{11}$$

By making use of (10) and (11) and the fact that

$$Var\left(\tilde{q}|data\right) = E\left(\tilde{q}^2|data\right) - \left\{E\left(\tilde{q}|data\right)\right\}^2,$$

it follows that

$$Var\left(\tilde{q}|data\right) = \left(\frac{m+n}{nm}\right)\left(\frac{n-1}{n-3}\right) + \tilde{k}_1^2\left\{\frac{n-1}{n-3} - \frac{\Gamma^2\left(\frac{m}{2}\right)\Gamma^2\left(\frac{n-2}{2}\right)(n-1)}{\Gamma^2\left(\frac{m-1}{2}\right)\Gamma^2\left(\frac{n-1}{2}\right)(m-1)}\right\}S^2.$$

Manuscript received, 2015-09-12, revised, 2016-01-21, accepted, 2016-01-22.