South African Statistical Journal
Vol. 53, No. 2, 87–113

BAYESIAN TESTING FOR PROCESS CAPABILITY INDICES

A. J. van der Merwe
University of the Free State
e-mail: *vandermerweaj6@gmail.com*

M. R. Sjölander
University of the Free State
e-mail: *sjolanderm@ufs.ac.za*

R. van Zyl
University of the Free State
e-mail: *ruaanvz@gmail.com*

Process capability indices have been widely used in the manufacturing industry. They measure the ability of a manufacturing process to produce items that meet certain specifications. A capability index relates the voice of the customer (specification limits) to the voice of the process. There is a need to understand and interpret process capability indices. Most of the existing work in this area has been devoted to classical frequentist large sample theory. An alternative approach to the problem of making inference about capability indices is the Bayesian approach. In this paper a Bayesian version of Tukey's method is used for constructing simultaneous credibility intervals for all pairwise differences. A Bayesian procedure for testing all possible contrasts is also given. The problem of selecting the best supplier(s) has received considerable attention in the literature, but mainly from a classical frequentist point of view. A Bayesian simulation procedure is also illustrated to find the best supplier or group of suppliers. This method seems much easier to perform than the Monte Carlo integration method given in Wu, Shiau, Pearn and Hung (2016). In section 10, a sensitivity analysis regarding the prior choice is considered and in the last section, t-distributed data are analysed.

Key words: All possible contrasts, Bayesian procedure, Best supplier, Capability indices, t-distribution.

1. Introduction

Process capability indices have been widely used in the manufacturing industry. They measure the ability of a manufacturing process to produce items that meet certain specifications. A capability index relates the voice of the customer (specification limits) to the voice of the process. A large value of the index indicates that the current process is capable of producing items (parts, tablets) that will meet or exceed the customer's requirements. Capability indices are convenient because they reduce complex information about the process to a single number and measure relative variability similar to the coefficient of variation.

Application examples include the manufacturing of semiconductor products (Hoskins, Stuart and Taylor, 1988), jet-turbine engine components (Hubele, Shahriari and Cheng, 1991), wood products (Lyth and Rabiej, 1995), audio speaker drivers (Chen and Pearn, 1997), wavelength division

MSC2010 subject classifications. 62F15.

multiplexes (Wu and Pearn, 2005), measuring EEPROM process capability (Wu and Pearn, 2006) and many others.

There is a need to understand and interpret process capability indices. In the literature on statistical quality control there have been some attempts to study the inferential aspects of these indices. Most of the existing work in this area has been devoted to classical frequentist large sample theory.

As mentioned by Wu and Pearn (2005) a point estimate of the index is not very useful in making reliable decisions. An interval estimation approach is in fact more appropriate and widely accepted but the frequency distributions of these estimators are often very complicated which means that the calculation of exact confidence intervals will be difficult.

An alternative approach to the problem of making inference about capability indices is the Bayesian approach. As it is well known in the Bayesian approach the information contained in the prior is combined with the likelihood to obtain the posterior distribution of the parameters. Inferences about the unknown parameters are based on the posterior distributions.

2. Definitions and notation

Four of the commonly used capability indices are

$$C_p = \frac{u-l}{6\sigma}, \qquad C_{pu} = \frac{u-\mu}{3\sigma}, \qquad C_{pl} = \frac{\mu-l}{3\sigma} \qquad \text{and} \qquad C_{pk} = \min\left(C_{pu}, C_{pl}\right).$$

C_{pk} is the normalized distance between the process mean and its closest specification limit. It can easily be verified that $C_{pk} = C_p(1-w)$, where $w = (2|m-\mu|)/(u-l)$ and $m = (u+l)/2$ is the midpoint of the specification limits (u and l). Thus, C_{pk} modifies C_p by a standardized measure w of non-centrality of the process and $C_{pk} = C_p$ if and only if the process is centered at m.

The larger the value of C_{pk}, the more capable is the process. In general, if the value of a process capability index is greater than 1, the process is said to be capable. According to Niverthi and Dey (2000), the thrust these days in the manufacturing industry is to achieve a C_{pk} value of at least 1.33. The definition of C_{pk} includes as special case those processes for which only one limit exists, by letting either $l \to -\infty$ or $u \to \infty$, in which case it reduces to the appropriate standardized measure.

Let $y_1, y_2, \ldots, y_n$ be an independent sample from a manufacturing process. In this paper it will be assumed that the y_i $(i = 1, 2, \ldots, n)$ are independent, identically normally distributed random variables with mean μ and variance σ^2. Since both μ and σ^2 are unknown and no prior information is available, the conventional non-informative, Jeffreys' prior

$$p(\mu, \sigma^2) \propto \sigma^{-2} \tag{1}$$

will be specified for μ and σ^2 in this section. Using (1), it is well known (see for example Zellner, 1971) that the conditional posterior density function of μ is normal:

$$\mu|\sigma^2, \underline{y} \sim N\left(\bar{y}, \frac{\sigma^2}{n}\right), \tag{2}$$

and in the case of the variance component σ^2, the posterior density function is given by

$$p(\sigma^2|\underline{y}) = K(\sigma^2)^{-\frac{1}{2}(n+1)} \exp\left\{-\frac{1}{2}\frac{(n-1)\,s^2}{\sigma^2}\right\}, \quad \sigma^2 > 0, \tag{3}$$

an inverted-gamma density function, where $\underline{y} = [y_1, y_2, \cdots, y_n]'$, $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ is the sample mean, $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$ is the sample variance, and

$$K = \left\{\frac{(n-1)s^2}{2}\right\}^{\frac{1}{2}(n-1)} \frac{1}{\Gamma((n-1)/2)} \tag{4}$$

is the normalizing constant. From (3) it follows that

$$k = \frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1} = \chi^2_v \tag{5}$$

for a given s^2.

As mentioned, these indices are used in process evaluation. From a Bayesian point of view the posterior distributions are of importance. One of the aims of this paper is therefore to derive the exact and in some cases the conditional posterior distributions of the indices. The method proposed by Ganesh (2009) for multiple testing will be applied using a Bayesian procedure for C_{pl}, C_{pu} and C_{pk} to determine whether significant differences between four suppliers exist. This method is a Bayesian version of Tukey's multiple comparisons procedure. A Bayesian method for testing all possible contrasts is also given. In section 9, a Bayesian simulation procedure is illustrated to find the best supplier or group of suppliers. In section 10, a sensitivity analysis regarding the prior choice is considered and in the last section, t-distributed data are analysed.

An estimated index will be denoted by "hat" (^). For example $\hat{C}_p = (u-l)/(6s)$, $\hat{C}_{pl} = (\bar{y}-l)/(3s)$, $\hat{C}_{pu} = (u-\bar{y})/(3s)$ and $\hat{C}_{pk} = \min(\hat{C}_{pu}, \hat{C}_{pl})$.

3. The posterior distribution of the Lower Process Capability Index $C_{pl} = (\mu - l)/(3\sigma)$

Theorem 1. *The posterior distribution of $t = C_{pl}$ is given by*

$$p(t|\tilde{t}) = \frac{3\sqrt{n}\exp\left\{-\frac{9nt^2}{2}\right\}}{\Gamma\left(\frac{v}{2}\right)\sqrt{2\pi}} \sum_{j=0}^{\infty}\left(\frac{9nt\tilde{t}}{\sqrt{v}}\right)^j \frac{1}{j!} \frac{\Gamma\left(\frac{v+j}{2}\right)2^{\frac{1}{2}j}}{\left(1+\frac{9n}{v}\tilde{t}^2\right)^{\frac{v+j}{2}}}, \qquad -\infty < t < \infty, \tag{6}$$

where

$$\tilde{t} = \frac{\bar{y}-l}{3s} = \hat{C}_{pl} \qquad \text{and} \qquad v = n-1.$$

Proof. The proof is given in the Mathematical Appendix to this paper. ■

Note. Chou and Owen (1989) derived the distribution of $\tilde{t}$, which is given by

$$f(\tilde{t}|t) = \frac{3\sqrt{n}\exp\left\{-\frac{9nt^2}{2}\right\}}{\sqrt{v}\sqrt{2\pi}\Gamma\left(\frac{v}{2}\right)} \sum_{j=0}^{\infty}\left(\frac{9n\tilde{t}t}{\sqrt{v}}\right)^j \frac{1}{j!} \frac{\Gamma\left(\frac{v+j+1}{2}\right)2^{\frac{1}{2}j}}{\left(1+\frac{9n}{v}\tilde{t}^2\right)^{\frac{v+j+1}{2}}}. \tag{7}$$

The density in (7) is that of a non-central t distribution with v degrees of freedom and non-centrality parameter δ, where $\delta^2 = 9nt^2$.

Table 1. $\hat{C}_{pl}$, $\hat{C}_{pu}$, and $\hat{C}_{pk}$ values for the four suppliers.

Supplier (i)	**1**	**2**	**3**	**4**
Sample size (n_i)	50	75	70	75
Estimated mean ($\bar{y}_i$)	2.7048	2.7019	2.6979	2.6972
Estimated standard deviation (s_i)	0.0034	0.0055	0.0046	0.0038
$\hat{C}_{pl}^{(i)} = (\bar{y}_i - l)/(3s_i)$	2.4804	1.3576	1.3333	1.5526
$\hat{C}_{pu}^{(i)} = (u - \bar{y}_i)/(3s_i)$	1.5392	1.1273	1.6377	2.0439
$\hat{C}_{pk}^{(i)} = \min(\hat{C}_{pl}^{(i)}, \hat{C}_{pu}^{(i)})$	1.5392	1.1273	1.3333	1.5526

4. The posterior distribution of $C_{pk} = \min(C_{pl}, C_{pu})$

When both specification limits are given, the C_p and C_{pk} indices can be used where

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

Unlike C_p, C_{pk} depends on both μ and σ. The C_{pk} index has been used in Japan and in the U.S. automotive companies (see Kane, 1986; Chou and Owen, 1989).

In Theorem 2 the posterior distribution of $c = C_{pk}$ will be derived.

Theorem 2. *The posterior distribution of $c = C_{pk}$ is given by*

$$p\left(c|\underline{y}\right) = \frac{3\sqrt{n}}{\sqrt{2\pi}} \int_{c^2 v/\tilde{b}^2}^{\infty} \left\{ \exp\left(-\frac{9n}{2}\left[c - t^*\sqrt{\frac{k}{v}}\right]^2\right) + \exp\left(-\frac{9n}{2}\left[c - \tilde{t}\sqrt{\frac{k}{v}}\right]^2\right)\right\} \frac{1}{2^{\frac{v}{2}}\Gamma\left(\frac{v}{2}\right)} k^{\frac{v}{2}-1} \exp\left(-\frac{k}{2}\right) dk, \quad (8)$$

where $v = n - 1$,

$$t^* = \hat{C}_{pu} = \frac{u - \bar{y}}{3s}, \qquad \tilde{t} = \hat{C}_{pl} = \frac{\bar{y} - l}{3s} \qquad \text{and} \qquad \tilde{b} = \hat{C}_p = \frac{u - l}{6s}.$$

Proof. The proof is given in the Mathematical Appendix to this paper. ∎

5. Example: piston rings for automotive engines (Polansky, 2006)

Consider a company with $N = 4$ contracted suppliers representing the four processes that produce piston rings for automobile engines studied by Chou (1994). The edge width of a piston ring after the preliminary disk grind is a very important quality characteristic in automobile engine manufacturing. The lower and upper specification limits of the quality characteristic are $l = 2.6795$mm and $u = 2.7205$mm respectively. Four potential suppliers (Supplier 1 to Supplier 4) for such rings are under consideration by one quality control manager. Samples of size $n_1 = 50$, $n_2 = 75$, $n_3 = 70$ and $n_4 = 75$ are taken from the manufacturing processes of the suppliers. A summary of the results from the samples, $\hat{C}_{pl}$, $\hat{C}_{pu}$, $\hat{C}_{pk}$ values and other statistics are given in Table 1.

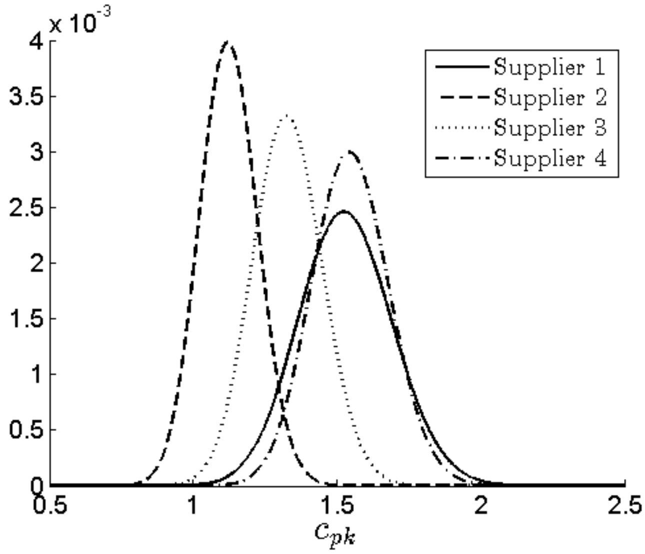

Figure 1. Posterior distributions of C_{pk}.

Table 2. Posterior means and variances.

	Supplier 1	**Supplier 2**	**Supplier 3**	**Supplier 4**
Posterior mean	1.5314	1.1234	1.3285	1.5474
Posterior variance	0.0263	0.0100	0.0144	0.0177

Looking at Table 1, it is clear that Suppliers 4 and 1 give the two largest values for $\hat{C}_{pl}$, $\hat{C}_{pu}$ and $\hat{C}_{pk}$, suggesting that they are the most capable. This may be because they seem to have the smallest variation within the specified range. They therefore represent the best two choices of suppliers. Suppliers 3 and 2 are not as capable as the former because of their greater variability. Because the estimated $\hat{C}_{pk}$ index for Supplier 1 is close to that of Supplier 4 we might feel that that the difference in capability of the processes between these suppliers is not significant. The same statement may hold true of Suppliers 2 and 3. Statistical methods for the comparison of the suppliers' process capability indices are required for the quality control manager to draw intelligent conclusions from this data. A Bayesian simulation procedure will be considered to determine which processes are significantly different from one another. The potential performance of the proposed method will be compared with the permutation approach by Polansky (2006).

The posterior distributions of the capability indices are displayed in Figure 1. From Table 2 it can be seen that the posterior means are for all practical purposes the same as the $\hat{C}_{pk}$ values given in Table 1. Further inspection of Figure 1 and Table 2 shows that Suppliers 1 and 4 have the largest posterior means, suggesting they are the most capable. In the next section a simple Bayesian solution to the problem of constructing simultaneous credibility intervals for the capability indices will be discussed.

6. Simultaneous credibility intervals

The method proposed by Ganesh (2009) can be compared to multiple testing, also referred to as the multiple comparison problem. In multiple testing, the objective is to control the family wise error rate. Similarly in his paper, Ganesh controls the simultaneous coverage rate. If the interest is in constructing simultaneous credibility intervals for all pairwise differences, a Bayesian version of Tukey's simultaneous confidence intervals can be used. Define

$$T^{(2)} = \max_i \left\{ C^{(i)}_{pk(l)} - \mathrm{E}(C^{(i)}_{pk} | \underline{y}_i) \,\middle|\, \underline{y}_i \right\} - \min_j \left\{ C^{(j)}_{pk(l)} - \mathrm{E}(C^{(j)}_{pk(l)} | \underline{y}_j) \,\middle|\, \underline{y}_j \right\},$$

for

$$l = 1, 2, \dots, \tilde{l}; \quad i = 1, 2, \dots, 4; \quad j = 1, 2, \dots, 4; \quad i \neq j,$$

where $T^{(2)}_\alpha$ is the upper α percentage point of the distribution of $T^{(2)}$. Simultaneous $100\,(1-\alpha)\,\%$ credibility intervals for all pairwise differences are given by

$$\mathrm{E}(C^{(i)}_{pk} | \underline{y}) - \mathrm{E}(C^{(j)}_{pk} | \underline{y}) \pm T^{(2)}_\alpha \quad i = 1, 2, \dots, 4;\ j = 1, 2, \dots, 4;\ i \neq j.$$

100,000 Monte Carlo simulations were used to calculate $\mathrm{E}(C^{(i)}_{pk} | \underline{y})$, $\mathrm{E}(C^{(j)}_{pk} | \underline{y})$ and $T^{(2)}_\alpha$.

The simulation procedure is as follows:

1. Simulate k from a χ^2_{n-1} distribution.
2. Calculate $\sigma^{2*}_i = (n-1)\, s^2_i / k$ in (5), where the asterisk (*) indicates a simulated value ($i = 1, 2, \dots, 4$).
3. Compute $\sigma^*_i = \sqrt{\sigma^{2*}_i}$.
4. By using the fact that $\mu_i | \sigma^2_i, \bar{y}_i \sim N(\bar{y}_i, \sigma^2_i / n)$ as in (2), simulate μ^*_i.
5. From the definition of the capability index it follows that $C^{(i)}_{pk}$ can be simulated as $C^{(i)*}_{pk} = \min((u - \mu^*_i)/(3\sigma^*_i), (\mu^*_i - l)/(3\sigma^*_i))$.
6. Repeat steps 1 to 5 $\tilde{l}$ times. As mentioned, for this example $\tilde{l} = 100{,}000$.

In Figure 2 the posterior distribution of $T^{(2)}$ is given and in Table 3 credibility intervals for differences in C_{pk} are illustrated.

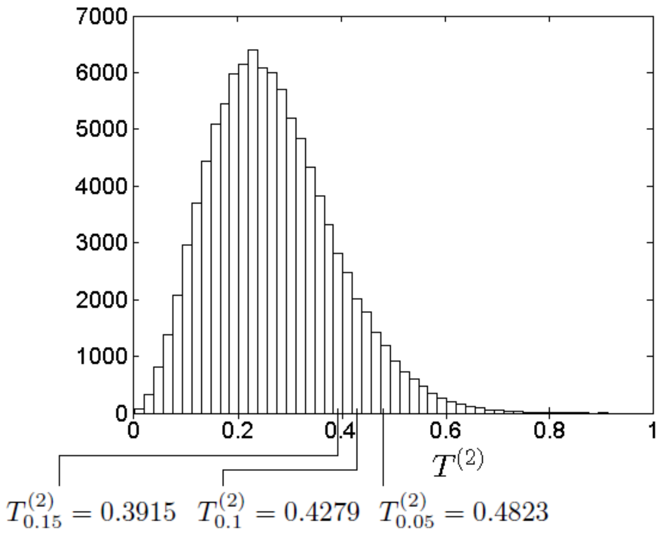

Figure 2. Posterior distribution of $T^{(2)}$.

Table 3. Credibility intervals for differences in C_{pk} – Ganesh Method.

$E(C_{pk}^{(i)} \mid \underline{y}) - E(C_{pk}^{(j)} \mid \underline{y})$	**95% interval**	**90% interval**	**87.47% interval**
Supplier 1 – Supplier 2	(−0.0734; 0.8915)	(−0.0187; 0.8371)	**(0.0000; 0.8155)**
Supplier 1 – Supplier 3	(−0.2779; 0.6867)	(−0.2234; 0.6323)	(−0.2058; 0.6097)
Supplier 1 – Supplier 4	(−0.4971; 0.4675)	(−0.4427; 0.4131)	(−0.4245; 0.3910)
Supplier 2 – Supplier 3	(−0.6871; 0.2775)	(−0.6326; 0.2231)	(−0.6136; 0.2019)
Supplier 2 – Supplier 4	(−0.9063; 0.0583)	(−0.8519; 0.0039)	**(−0.8323; −0.0168)**
Supplier 3 – Supplier 4	(−0.7016; 0.2630)	(−0.6471; 0.2086)	(−0.6265; 0.1890)

For solving the supplier problem Polansky (2006) used multiple comparison techniques in conjunction with permutation tests. The multiple comparisons tests used were:

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

Polansky (2006) came to the conclusion that at the 5% significance level Suppliers 1, 3 and 4 have process capabilities that are not significantly different. Similarly, Suppliers 2 and 3 are not significantly different from one another, but Supplier 2 is significantly different from Suppliers 1 and 4.

According to Table 3 it is only at a significance level of 12.5% that the Bayesian procedure shows a significant difference between Supplier 2 and Suppliers 1 and 4. To see if Ganesh's (2009) version of Tukey's simultaneous confidence intervals is somewhat conservative, the following simulation study has been conducted to evaluate the coverage probability and power of the Bayesian hypothesis testing procedure.

1. (a) Assume that $y \sim N(\mu_1, \sigma_1^2)$ where $\mu_1 = 2.7048$ and $\sigma_1^2 = (0.0034)^2$. The parameters μ_1 and σ_1^2 are obtained from the sample statistics of Supplier 1.
   (b) Simulate the sufficient statistics $\bar{y}_i \sim N(\mu_1, \sigma_1^2/n_1)$ and $(n_1 - 1)\, s_i^2 \sim \sigma_1^2 \chi_{n_1-1}^2$ to represent a data set for the four suppliers where $n_1 = 50$ and $i = 1, 2, 3, 4$.
   (c) By doing $\tilde{l} = 10000$ simulations $T_{0.05}^{(2)}$ can be calculated for our first dataset as well as the credibility intervals as described in Section 6.
   (d) If any one of the six credibility intervals do not contain zero, the null hypothesis

   $$H_0 : C_{pk}^{(1)} = C_{pk}^{(2)} = C_{pk}^{(3)} = C_{pk}^{(4)}$$

   will be rejected. Rejection of H_0 when it is true is called a Type I error.
   (e) Steps (a) - (d) are replicated $l^* = 20{,}000$ times with $\mu_1 = 2.7048$, $\sigma_1^2 = (0.0034)^2$ and $n_1 = 50$ and the estimated Type I error $= \frac{1008}{20000} = 0.0504$ which corresponds well with $\alpha = 0.05$. It means that for 1008 datasets one or more of the six credibility intervals did not contain zero.
2. Assume now that $y \sim N(\mu_2, \sigma_2^2)$ where $\mu_2 = 2.7019$, $\sigma_2^2 = (0.0054)^2$ and $n_2 = 75$. The parameter values are that of the sample statistics of the second supplier. Repeat steps 1(a)–1(e) and also for Suppliers 3 and 4.

In Table 4 the estimated Type I errors for the four cases are given.

The average Type I error = 0.0504 which as mentioned corresponds well with $\alpha = 0.05$. It therefore does not seem that the Ganesh Bayesian version of Tukey's simultaneous confidence interval is too conservative.

Table 4. Estimated type I error for different parameter combinations and sample sizes.

n	μ	σ	**Type I error**
50	2.7048	0.0034	0.0504
75	2.7019	0.0054	0.0483
70	2.6979	0.0046	0.0521
75	2.6972	0.0038	0.0507

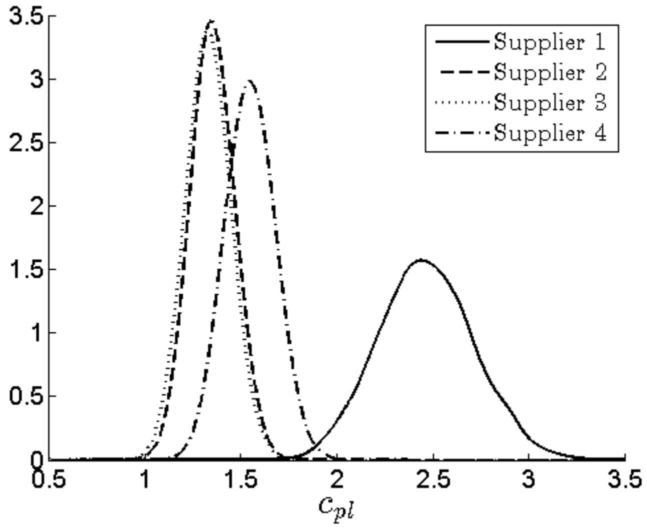

Figure 3. Posterior distributions of C_{pl}.

7. Posterior distributions of C_{pl} and C_{pu}

It might be of interest to also look at the posterior distributions of $C_{pl} = (\mu - l)/(3\sigma)$ and $C_{pu} = (u - \mu)/(3\sigma)$. The posterior distribution of C_{pl} is given in (6) and can be used for illustration purposes. A much easier way to obtain the posterior distribution is to simulate a large number of conditional posterior distributions. The average of these conditional distributions is then the unconditional posterior distribution of C_{pl}. This procedure is called the Rao-Blackwell method.

In Figures 3 and 4 the posterior distributions of C_{pl} and C_{pu} are displayed. In Table 5 the posterior means of C_{pl} and C_{pu} are given for the four suppliers and in Table 6 the 95% credibility intervals for the differences between the suppliers are given using Ganesh method. According to the C_{pl} credibility

Table 5. Posterior means of C_{pl} and C_{pu}.

Supplier	**1**	**2**	**3**	**4**
C_{pl}	2.4920	1.3615	1.3377	1.5585
C_{pu}	1.5460	1.1303	1.6431	2.0521

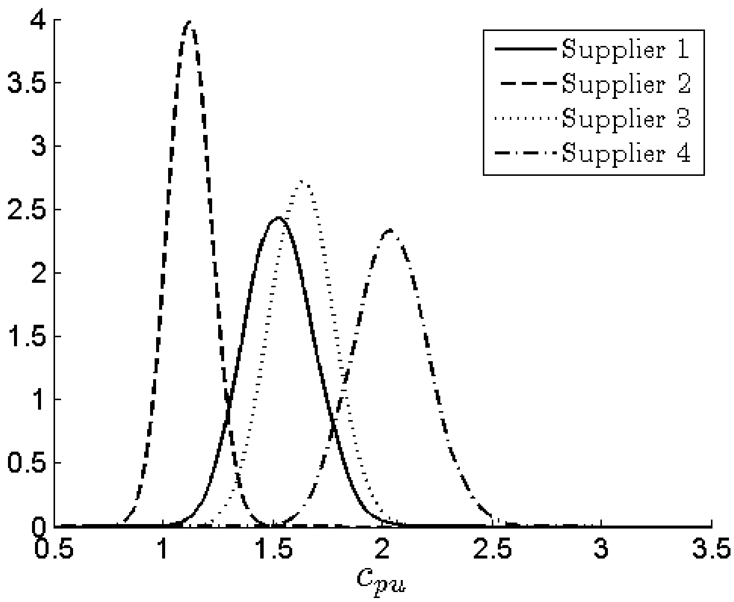

Figure 4. Posterior distributions of C_{pu}.

Table 6. 95% credibility intervals for differences between suppliers.

	C_{pl}	C_{pu}
Supplier 1 – Supplier 2	**(0.4910; 1.7700)**	(-0.1281; 0.9595)
Supplier 1 – Supplier 3	**(0.5148; 1.7938)**	(-0.6408; 0.4467)
Supplier 1 – Supplier 4	**(0.2940; 1.5729)**	(-1.0498; 0.0377)
Supplier 2 – Supplier 3	(-0.6157; 0.6633)	(-1.0565; 0.0310)
Supplier 2 – Supplier 4	(-0.8365; 0.4424)	**(-1.4655; -0.3780)**
Supplier 3 – Supplier 4	(-0.8603; 0.4187)	(-0.9527; 0.1348)

intervals Supplier 1 is significantly different from Suppliers 2, 3 and 4. The other suppliers do not differ significantly from each other. Inspection of the C_{pu} intervals shows that there is a significant difference between Suppliers 2 and 4.

8. Testing all possible contrasts

By extending the method of Ganesh (2009) a Bayesian procedure will be explained for testing all possible contrasts. Let

$$T^{(3)} = \max_{\boldsymbol{l}} \frac{\{\boldsymbol{l}'(\boldsymbol{\theta} - \mathrm{E}(\boldsymbol{\theta}|\boldsymbol{y}))\}^2}{\boldsymbol{l}' \operatorname{Var}(\boldsymbol{\theta}|\boldsymbol{y})\boldsymbol{l}},$$

where $\boldsymbol{l} = [\boldsymbol{l}_1', \boldsymbol{l}_2', \boldsymbol{l}_3']'$. For the supplier example (Section 5), the following are possible contrasts:

$$\boldsymbol{l_1} = \left[\frac{1}{\sqrt{4}}, -\frac{1}{\sqrt{4}}, -\frac{1}{\sqrt{4}}, \frac{1}{\sqrt{4}}\right]', \qquad \boldsymbol{l_2} = \left[\frac{1}{\sqrt{2}}, 0, 0, -\frac{1}{\sqrt{2}}\right]', \qquad \boldsymbol{l_3} = \left[0, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0\right]'.$$

Further, let

$$\boldsymbol{\theta} = \left[C_{pk}^{(1)}, C_{pk}^{(2)}, C_{pk}^{(3)}, C_{pk}^{(4)}\right]',$$

$$\mathrm{E}(\boldsymbol{\theta}|\boldsymbol{y}) = \left[\mathrm{E}\left(C_{pk}^{(1)}\right), \mathrm{E}\left(C_{pk}^{(2)}\right), \mathrm{E}\left(C_{pk}^{(3)}\right), \mathrm{E}\left(C_{pk}^{(4)}\right)\right]'$$

and

$$\operatorname{Var}(\boldsymbol{\theta}|\boldsymbol{y}) = diag\left[\operatorname{Var}\left(C_{pk}^{(1)}\right), \operatorname{Var}\left(C_{pk}^{(2)}\right), \operatorname{Var}\left(C_{pk}^{(3)}\right), \operatorname{Var}\left(C_{pk}^{(4)}\right)\right]'.$$

Therefore,

$$T^{(3)} = \max_{\boldsymbol{l}} \frac{\{\boldsymbol{l}'(\boldsymbol{\theta} - \mathrm{E}(\boldsymbol{\theta}|\boldsymbol{y}))\}^2}{\boldsymbol{l}' \operatorname{Var}(\boldsymbol{\theta}|\boldsymbol{y})\boldsymbol{l}} = \max\left\{\frac{\{\boldsymbol{l}_1'(\boldsymbol{\theta} - \mathrm{E}(\boldsymbol{\theta}|\boldsymbol{y}))\}^2}{\boldsymbol{l}_1' \operatorname{Var}(\boldsymbol{\theta}|\boldsymbol{y})\boldsymbol{l_1}}, \frac{\{\boldsymbol{l}_2'(\boldsymbol{\theta} - \mathrm{E}(\boldsymbol{\theta}|\boldsymbol{y}))\}^2}{\boldsymbol{l}_2' \operatorname{Var}(\boldsymbol{\theta}|\boldsymbol{y})\boldsymbol{l_2}}, \frac{\{\boldsymbol{l}_3'(\boldsymbol{\theta} - \mathrm{E}(\boldsymbol{\theta}|\boldsymbol{y}))\}^2}{\boldsymbol{l}_3' \operatorname{Var}(\boldsymbol{\theta}|\boldsymbol{y})\boldsymbol{l_3}}\right\}.$$

The posterior distribution of $T^{(3)}$ is given in Figure 5. This is obtained by performing the simulation procedure, discussed in Section 6, 500 000 times.

The $100(1-\alpha)\%$ Bayesian confidence (credibility) intervals are given by

$$\boldsymbol{l}_i' \mathrm{E}(\theta|\boldsymbol{y}) \mp \left\{\boldsymbol{l}_i' \operatorname{Var}(\boldsymbol{\theta}|\boldsymbol{y})\boldsymbol{l_i} T_\alpha^{(3)}\right\}^{\frac{1}{2}}, \quad i = 1, 2, 3.$$

For the piston rings example, the 95% credibility for $\boldsymbol{l_1}, \boldsymbol{l_2}$ and $\boldsymbol{l_3}$ are

$$[0.0057, 0.6296], \quad [-0.3613, 0.3465] \quad \text{and} \quad [-0.4098, 0.1172].$$

The hypothesis that the first contrast is zero is therefore rejected at the 5% level. It is therefore clear that on average Suppliers 1 and 4 are definitely better than 2 and 3.

To get an idea of the magnitude of the Type I error of the Bayesian procedure, the sample values of Supplier 1 were taken as parameter values for all four suppliers. By performing the simulation procedure explained in Section 6 it was found that the proportion of rejections for the three contrasts were 0.0161, 0.0154 and 0.0181, which means that the Type I error = 0.0161 + 0.0154 + 0.0181 = 0.0496, which is for all practical purposes 0.05.

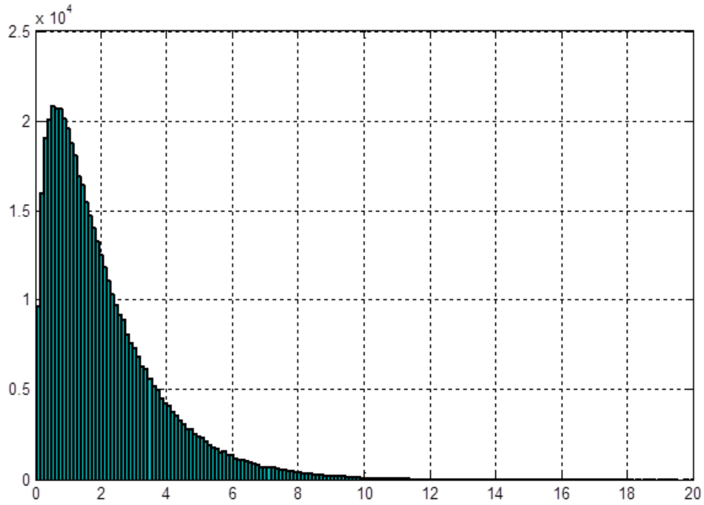

Figure 5. Posterior distributions of $T^{(3)}$.

Table 7. Part of the 1 000 000 simulated C_{pk} values for the four potential suppliers.

1	**Supplier 1**	**Supplier 2**	**Supplier 3**	**Supplier 4**
1	1.59118	1.19072	1.19171	1.37449
2	1.32416	0.90496	1.31725	1.68731
3	1.60787	0.99262	1.23655	1.40261
4	1.31843	1.12253	1.30348	1.75133
5	1.57734	1.12316	1.32383	1.55906
6	1.63339	1.11287	1.3116	1.55693
⋮	⋮	⋮	⋮	⋮
1 000 000	1.38891	1.16121	1.49352	1.65897

9. Selecting the best supplier

The problem of selecting the best supplier has received considerable attention in the literature but mainly from a classical frequentist point of view. In this section a Bayesian simulation procedure is illustrated for finding the best supplier or group of suppliers. This method takes less than three minutes to calculate the probabilities given in Table 9. It therefore seems that our method is much simpler to perform and takes less time than the Monte Carlo integration method given in Wu et al. (2016).

By using the simulation procedure as explained in Section 6 the results in Tables 7–9 were obtained. These probabilities are for all practical purposes the same as those obtained by Wu et al. (2016) using Monte Carlo integration. In (0.451088+0.506793)100% $\approx$ 95% of the cases Supplier 1 or Supplier 4 were selected as the best supplier.

Table 8. Part of the 1 000 000 rankings of the simulated C_{pk} values.

1	**Supplier 1**	**Supplier 2**	**Supplier 3**	**Supplier 4**
1	1	4	3	2
2	2	4	3	1
3	1	4	3	2
4	2	4	3	1
5	1	4	3	2
6	1	4	3	2
⋮	⋮	⋮	⋮	⋮
1 000 000	3	4	2	1

Table 9. Probabilities for the four potential suppliers.

Probability	**Supplier 1**	**Supplier 2**	**Supplier 3**	**Supplier 4**
P($Supplier_i$) is 1^{st}	0.451088	0.000384	0.041735	0.506793
P($Supplier_i$) is 2^{nd}	0.405871	0.006773	0.182797	0.404559
P($Supplier_i$) is 3^{rd}	0.131232	0.100010	0.683410	0.085348
P($Supplier_i$) is 4^{th}	0.011809	0.892833	0.092058	0.003300
Total	1	1	1	1

10. Sensitivity analysis

10.1 Objective priors

In Sections 1–9, the prior $p(\mu, \sigma^2) \propto \sigma^{-2}$ (Jeffreys' independence prior) was specified for μ and σ^2. Using this prior it follows that the conditional posterior density function of μ is normal: $\mu|\sigma^2, \bar{y} \sim N(\bar{y}, \sigma^2/n)$ and in this case of the variance component σ^2, the posterior density function is an inverted-gamma distribution (see (3)), i.e. $(n-1)\, s^2/\sigma^2 \sim \chi^2_{n-1}$.

One form of sensitivity analysis is to vary the power of σ^2. The simulation procedure will be similar to that for the posterior distribution (3) except that (given the choice of distribution) σ^2 is distributed from a central Chi-squared distribution with degrees of freedom as given in Table 10. Prior 1 is the Jeffreys' Rule (Jeffreys' dependence) prior, which is the square root of the determinant of the Fisher information matrix. Prior 3 is Jeffreys' independence prior which is used as prior in this paper and Prior 5 is a uniform prior.

In Table 11 the posterior means and variances of C_{pk} are given for the five priors. Prior 3 is included in Table 11 for completeness sake. It is interesting to note that the posterior means become slightly smaller with decreasing degrees of freedom. The variances remain quite stable. The posterior means in the case of Prior 2 ($df = n - 0.5$) are exactly the same as the sample estimates of C_{pk} given in Table 1.

Table 10. Prior distributions and simulation parameters.

Name	**Specification**	**Simulation**
Prior 1	$p(\mu,\sigma^2) \propto \sigma^{-3}$	$\sigma^2 = (n-1)S^2/\chi^2_n$
Prior 2	$p(\mu,\sigma^2) \propto \sigma^{-2.5}$	$\sigma^2 = (n-1)S^2/\chi^2_{n-0.5}$
Prior 3	$p(\mu,\sigma^2) \propto \sigma^{-2}$	$\sigma^2 = (n-1)S^2/\chi^2_{n-1}$
Prior 4	$p(\mu,\sigma^2) \propto \sigma^{-1}$	$\sigma^2 = (n-1)S^2/\chi^2_{n-2}$
Prior 5	$p(\mu,\sigma^2) \propto$ constant	$\sigma^2 = (n-1)S^2/\chi^2_{n-3}$

Table 11. Posterior means and variances for C_{pk} in the case of objective priors.

Prior	**df.**	**Posterior mean/var.**	**Supplier 1**	**Supplier 2**	**Supplier 3**	**Supplier 4**
1	n	Posterior mean	1.54711	1.13119	1.33821	1.5578
		Posterior variance	0.0262858	0.0100431	0.0144074	0.0176784
2	$n-0.5$	Posterior mean	1.53925	1.12719	1.33327	1.55272
		Posterior variance	0.0262561	0.0100398	0.0144093	0.0177456
3	$n-1$	Posterior mean	1.53142	1.1234	1.32842	1.54744
		Posterior variance	0.026346	0.0100141	0.0144407	0.0177302
4	$n-2$	Posterior mean	1.51558	1.11567	1.3187	1.53696
		Posterior variance	0.026296	0.0100278	0.0144323	0.0177292
5	$n-3$	Posterior mean	1.49936	1.10807	1.30891	1.52618
		Posterior variance	0.0262492	0.0100501	0.0144337	0.0176995

Table 12. Posterior means and variances for C_{pk} in the case of inverse-gamma priors.

Prior	**Posterior mean/var.**	**Supplier 1**	**Supplier 2**	**Supplier 3**	**Supplier 4**
A	Posterior mean	1.30692	1.26928	1.38964	1.45013
	Posterior variance	0.00810945	0.00637278	0.00776587	0.00806702
B	Posterior mean	1.03306	1.01819	1.12544	1.21658
	Posterior variance	0.0109322	0.00746595	0.00944446	0.0101157

10.2 Proper priors

Another form of sensitivity analysis is to look at proper priors for example the normal-inverted gamma prior. Suppose the normal prior $\mu_i \sim N(\mu_0, \sigma^2/k_0)$ is used for μ_i $(i = 1,2,3,4)$. In this case the parameters μ_0 and k_0 can be interpreted as the mean and sample size from a set of prior observations.

It is easy to show that the conditional posterior distribution of μ_i $(i = 1,2,3,4)$ is also normal with mean $\mathrm{E}(\mu_i|\sigma^2, data) = (n_i\bar{y} + \mu_0 k_0)/k_{n_i}$ and variance $V(\mu_i|\sigma_i^2, data) = \sigma_i^2/k_{n_i}$, where $k_{n_i} = n_i + k_0$

For the variance component σ_i^2 we use an inverse-gamma prior: $\sigma_i^2 \sim$ Inverse-Gamma$(\frac{\nu_0}{2}, \frac{\nu_0}{2}\sigma_0^2)$, which means that the posterior distribution of σ_i^2 is

$$\sigma_i^2|data \sim \text{Inverse-Gamma}\left(\frac{\nu_{n_i}}{2}, \frac{\nu_{n_i}\sigma_{n_i}^2}{2}\right), \tag{9}$$

where $\nu_{n_i} = \nu_0 + n_i - 1$ and $\sigma_{n_i}^2 = \nu_{n_i}^{-1}[\nu_0\sigma_0^2 + (n_i - 1)s_i^2 + (k_0 n_i/k_{n_i})(\bar{y} - \mu_0)^2]$

These formulas suggest an interpretation of ν_0 as a prior sample size, from which a prior sample variance of σ_0^2 has been obtained. For further details see Hoff (2009).

From (9) it follows that

$$\sigma_i^2 \sim \frac{\nu_0\sigma_0^2 + (n_i - 1)s_i^2 + \frac{k_0 n_i}{k_{n_i}}(\bar{y} - \mu_0)^2}{\chi_{n_i}^2}, \quad i = 1,2,3,4.$$

In Table 12 the posterior means and variances of C_{pk} for two normal inverse-gamma priors are given. In the case of Prior A the parameter specifications are

$$\nu_0 = 70, \quad \sigma_0^2 = 0.00002, \quad \mu_0 = 2.7 \quad \text{and} \quad k_0 = 10,$$

and for Prior B,

$$\nu_0 = 10, \quad \sigma_0^2 = 0.00009, \quad \mu_0 = 2.7 \quad \text{and} \quad k_0 = 10.$$

From Table 12 it is clear that the subjective priors can have features that have an unexpectedly dramatic effect on the results. In the case of objective priors (see for example Table 9 and Table 11). Suppliers 1 and 4 were selected as the best suppliers in 95% of the cases. On the other hand using subjective priors (Table 12) it is clear that Suppliers 3 and 4 are the two best suppliers. It thus seems to be wrong that Supplier 3 could have been considered as better than Supplier 1. The two proper priors are therefore not recommended for further use.

10.3 Simulation study

In this section, a simulation study is considered to observe if the 95% Bayesian confidence intervals for C_{pk} have the correct frequentist coverage. In doing the simulation study it is assumed that the parameter values μ and σ^2 are unknown. For the simulation study, the samples are drawn from a normal distribution with mean $\mu = 2.7$ and variance $\sigma^2 = (0.004)^2$. Also $l = 2.6795$ and $u = 2.7205$ which means that $C_{pk} = 1.7083$. The parameter values are similar to the sample statistics calculated for the piston rings example. The sample sizes that will be considered are: (i) $n = 10$, (ii) $n = 20$, (iii) $n = 30$, (iv) $n = 40$, (v) $n = 50$ and (vi) $n = 70$. The following priors wil be used: (a) $p(\mu, \sigma^2) \propto \sigma^{-1}$, (b) $p(\mu, \sigma^2) \propto \sigma^{-2}$, (c) $p(\mu, \sigma^2) \propto \sigma^{-2.5}$, (d) $p(\mu, \sigma^2) \propto \sigma^{-3}$ and (e) $p(\mu, \sigma^2) \propto \sigma^{-4}$.

Table 13. Coverage percentages for C_{pk} of ten thousand 95% confidence intervals.

Name	$n = 10$	$n = 20$	$n = 30$	$n = 40$	$n = 50$	$n = 70$
$p(\mu, \sigma^2) \propto \sigma^{-1}$	91.48	92.44	92.49	93.56	93.69	94.11
$p(\mu, \sigma^2) \propto \sigma^{-2}$	93.27	93.61	93.78	94.02	94.24	94.38
$p(\mu, \sigma^2) \propto \sigma^{-2.5}$	93.35	93.82	93.88	94.04	94.31	94.45
$p(\mu, \sigma^2) \propto \sigma^{-3}$	94.09	94.17	94.26	94.31	94.58	94.61
$p(\mu, \sigma^2) \propto \sigma^{-4}$	94.31	94.53	94.49	94.71	94.76	94.83

(I) (i) To conduct the simulation study the sufficient statistics $(n-1)s^2 \sim \sigma^2 \chi^2_{n-1}$ and $\bar{y} \sim N(\mu, \sigma^2/n)$ will be simulated to represent a random sample of n observation from a normal population with mean μ and variance σ^2.

(ii) After the data set has been simulated it is assumed that the values of the parameters μ and σ^2 are unknown. If the prior $p(\mu, \sigma^2) \propto \sigma^{-3}$ (Jeffreys' Rule prior) is for example used, then the posterior distributions are $p(\sigma^2|data) \sim (n-1)s^2/\chi_n$ and $p(\mu|\sigma^2, data) \sim N(\bar{y}, \sigma^2/n)$.

(iii) By simulating σ^2 and μ from the posterior distributions, a C_{pk} value can be calculated. This C_{pk} value is called $C_{pk(1)}$.

(iv) For this data set simulate new parameter values and calculate $C_{pk(l)}$ $(l = 1, \ldots, 10\,000)$.

(v) Calculate the mean and the 95% confidence interval for the 10 000 C_{pk} values and observe if the confidence interval contains the true $C_{pk} = 1.70833$ value.

(II) Repeat (I)(i) by simulating a second random sample (sufficient statistics) and do (I)(ii)–(I)(v) for this sample.

(III) Repeat the procedure 10 000 times. In other words 10 000 samples are drawn and for each sample 10 000 C_{pk} values are simulated.

(IV) Calculate the mean of the 10 000 C_{pk} means. These values as well as the coverage percentages of the 95% confidence intervals are given in the following tables.

From Table 13 it is clear that except for $p(\mu, \sigma^2) \propto \sigma^{-1}$, the coverage percentage is at least 94% if $n \geq 40$. The best coverage percentages are given by priors $p(\mu, \sigma^2) \propto \sigma^{-3}$ and $p(\mu, \sigma^2) \propto \sigma^{-4}$. $p(\mu, \sigma^2) \propto \sigma^{-3}$ is Jeffreys' dependence prior.

In Table 14 the averages of the 10 000 sample means of C_{pk} are illustrated. The true parameter value is $C_{pk} = 1.70833$. From the table it seems that the priors $p(\mu, \sigma^2) \propto \sigma^{-3}$ and $p(\mu, \sigma^2) \propto \sigma^{-4}$ give the best results.

It is well known that Jeffreys' independence prior $p(\mu, \sigma^2) \propto \sigma^{-2}$ gives the correct point estimates and confidence intervals for the parameters μ and σ^2. From Tables 13 and 14 it however seems that priors of the form $p(\mu, \sigma^2) \propto \sigma^{-a}$ $(a > 2)$ will give better results for the parameter C_{pk} than Jeffreys' independence prior.

Other priors for μ and σ^2 should be considered to improve point estimates and coverage percentages for C_{pk}. The prior $p(\mu, \sigma^2) \propto \sigma^{-3}\{1 + \mu^2/(2\sigma^2)\}^{-0.5}$ is a reference as well as a probability-matching prior for μ/σ, the standardised mean and should be considered.

Table 14. Average of the 10 000 sample means of C_{pk}.

Name	$n = 10$	$n = 20$	$n = 30$	$n = 40$	$n = 50$	$n = 70$
$p(\mu, \sigma^2) \propto \sigma^{-1}$	1.58240	1.62296	1.63963	1.6447	1.65268	1.66451
$p(\mu, \sigma^2) \propto \sigma^{-2}$	1.69060	1.66962	1.66824	1.6712	1.67445	1.68032
$p(\mu, \sigma^2) \propto \sigma^{-2.5}$	1.74702	1.68872	1.68026	1.67998	1.68199	1.68164
$p(\mu, \sigma^2) \propto \sigma^{-3}$	1.75201	1.71845	1.70234	1.69062	1.69128	1.68746
$p(\mu, \sigma^2) \propto \sigma^{-4}$	1.79177	1.76331	1.73041	1.71162	1.70888	1.68926

11. A Comparison of Bayesian and frequentist approaches on data with heavy tails

11.1 Bayesian Procedure

To accommodate the possibility of outlying measurements, the assumption of Gaussian observations will be relaxed in the direction of the Student-t distribution family.

Consider a series of n independent observations: $y_i|\mu, \sigma^2, \lambda_i \sim N(\mu, \sigma^2/\lambda_i)$ for $i = 1, 2, 3, \ldots, n$. By placing a prior distribution on λ_i enables a wide variety of distributions $f(y_i|\mu, \sigma^2)$ to emerge as scale mixtures of normal distributions (Andrews and Mallows, 1974; Carlin and Polson, 1991; Wakefield, Smith, Racine-Poon and Gelfand, 1994):

$$f(y_i|\mu, \sigma^2) = \int_0^\infty f(y_i|\mu, \sigma^2, \lambda_i) p(\lambda_i) d\lambda_i.$$

In this section it is assumed that $\nu\lambda_i \sim \chi^2_\nu$ so that $y_i|\mu, \sigma^2 \sim t_\nu(\mu, \sigma^2)$ has a Student t-distribution with mean μ, variance $\sigma^2\nu/(\nu - 2)$ and degrees of freedom ν. Using this method, the likelihood function can be written as

$$\begin{aligned} L(\mu, \sigma^2, \underline{\lambda}, \nu) &= \prod_{i=1}^n f(y_i, \lambda_i|\mu, \sigma^2, \nu) \\ &= \frac{\prod_{i=1}^n (\lambda_i)^{0.5}}{(2\pi)^{n/2}(\sigma^2)^{n/2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n \lambda_i (y_i - \mu)^2} \frac{\nu^{n\nu/2}}{2^{n\nu/2}[\Gamma(\frac{\nu}{2})]^n} \prod_{i=1}^n \lambda_i^{\nu/2-1} e^{-\frac{1}{2}\nu\sum_{i=1}^n \lambda_i}, \end{aligned}$$

where $\underline{\lambda} = [\ \lambda_1 \quad \lambda_2 \quad \cdots \quad \lambda_n]'$.

The prior distribution that will be used is $p(\mu, \sigma^2, \nu) \propto \sigma^{-2} e^{-\xi\nu}$, $\nu > 2$. As the prior for ν (the degrees of freedom) a truncated ($\nu > 2$) exponential distribution with parameter $\xi = 0.1$ is assumed. The truncation assures the finiteness of the mean and variance of the associated t-distribution. For further details see Geweke (1993).

The joint posterior distribution is obtained by multiplying the likelihood function with the prior distribution and is given by

$$\begin{aligned} p(\mu, \sigma^2, \underline{\lambda}, \nu|\underline{y}) \propto & \left(\frac{1}{\sigma^2}\right)^{\frac{1}{2}(n+2)} |H|^{\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma^2}(\underline{y} - \mu\underline{1})' H (\underline{y} - \mu\underline{1})\right\} \\ & \times \frac{\nu^{n\nu/2}}{2^{n\nu/2}[\Gamma(\frac{\nu}{2})]^n} \prod_{i=1}^n \lambda_i^{\nu/2-1} e^{-\frac{1}{2}\nu\sum_{i=1}^n \lambda_i} e^{-\xi\nu}, \quad \nu > 2, \end{aligned} \tag{10}$$

where $H = \text{diag}[\lambda_1, \lambda_2, \ldots, \lambda_n]$, $\underline{y} = [y_1, y_2, \ldots, y_n]'$ and $\mu\underline{1} = \mu[1, 1, \ldots, 1]'$. Therefore, $\prod_{i=1}^{n} \lambda_i^{\frac{1}{2}} = |H|^{\frac{1}{2}}$ and $\sum_{i=1}^{n} \lambda_i (y_i - \mu)^2 = (\underline{y} - \mu\underline{1})' H (\underline{y} - \mu\underline{1})$.

To implement the Gibbs sampler, the full conditional posterior distributions of the unknown parameters are needed.

From (10) it follows that

$$\mu | \sigma^2, H, \underline{y} \sim N\left[(\underline{1}'H\underline{1})^{-1}\underline{1}'H\underline{y};\ \sigma^2\left(\underline{1}'H\underline{1}\right)^{-1}\right], \tag{11}$$

and the posterior distribution of σ^2 is

$$p(\sigma^2|\mu, H, \underline{y}) \propto \left(\frac{1}{\sigma^2}\right)^{\frac{1}{2}(n+2)} \exp\left\{-\frac{1}{2\sigma^2}(\underline{y} - \mu\underline{1})'H(\underline{y} - \mu\underline{1})\right\}, \tag{12}$$

an inverse-gamma distribution. Therefore, $\sigma^2|\mu, H, \underline{y} \sim (\underline{y} - \mu\underline{1})'H(\underline{y} - \mu\underline{1})/\chi_n^2$. Also,

$$p(\lambda_i|\mu, \sigma^2, \nu, y_i) \propto \lambda_i^{\frac{1}{2}(\nu-1)} \exp\left\{-\frac{1}{2}\lambda_i\left[\frac{(y_i - \mu)^2}{\sigma^2} + \nu\right]\right\}. \tag{13}$$

This is a gamma distribution, which means that $\lambda_i|\mu, \sigma^2, \nu, y_i \sim \chi_{\nu+1}^2(\nu + (y_i - \mu)^2/\sigma^2)^{-1}$, $i = 1, 2, \ldots n$. In the case of ν, the degrees of freedom, it follows that

$$p(\nu|\underline{\lambda}) \propto \frac{\nu^{n\nu/2}}{2^{n\nu/2}[\Gamma(\frac{\nu}{2})]^n} \prod_{i=1}^{n} \lambda_i^{\frac{1}{2}(\nu-1)} \exp\left\{-\nu\left[\frac{1}{2}\sum_{i=1}^{n}\lambda_i + \xi\right]\right\}. \tag{14}$$

Example 1

Ten thousand data sets of size $n = 20$ were simulated from a t-distribution with $\nu = 3$ degrees of freedom, $\mu = 30$ and $\sigma^2 = 4$. Also for the capability index $u = 40$ and $l = 20$ which means that $C_{pk} = 1.6667$. By using (11)–(14) and Gibbs sampling, the unconditional posterior distributions $p(\mu|\underline{y})$, $p(\sigma^2|\underline{y})$, $p(\lambda_i|\underline{y})$, $p(\nu|\underline{y})$ and $p(C_{pk}|\underline{y})$ can be obtained.

In Figure 6, the posterior distribution $p(\nu|\underline{y})$ for the first data set is illustrated. The histogram is obtained from 10 000 simulations. The mode of 3.35 is in the vicinity of $\nu = 3$, the true parameter value. It also clear from the figure that the posterior distribution of ν is quite skew. The posterior distributions of ν for most of the 10 000 data sets are of a similar form as $p(\nu|y_1)$ with modes not too far from 3.

As mentioned, 10 000 samples each of size $n = 20$ were drawn from a t-distribution with $\nu = 3$ degrees of freedom, $\mu = 30$ and $\sigma^2 = 4$ and for each sample 10 000 simulations were executed to obtain the posterior distributions of the parameters μ, σ^2, $\underline{\lambda} = [\lambda_1, \lambda_2, \ldots, \lambda_{20}]'$, ν and C_{pk}.

For each sample the mean value of the posterior distributions of the parameters were calculated and it was observed if the 95% Bayes confidence intervals contain the true parameter value. The mean values are illustrated in Figures 7–11.

In Figure 10, for example, the posterior means $\text{E}(\nu|\underline{y}_1)$, $\text{E}(\nu|\underline{y}_2)$,$\ldots$, $\text{E}(\nu|\underline{y}_{10000})$ for the 10000 samples are illustrated. Also $Mean = 11.997 = \frac{1}{10000}\Sigma(\nu|\underline{y}_i)$ with $\text{E}(\nu|\underline{y}_1) = 11.9672$ as shown in Figure 6. From Figure 6 it is also clear that the coverage percentage for ν is very good (94.87%). This means that in the case of 9487 samples, the 95% confidence intervals contain the true parameter value $\nu = 3$.

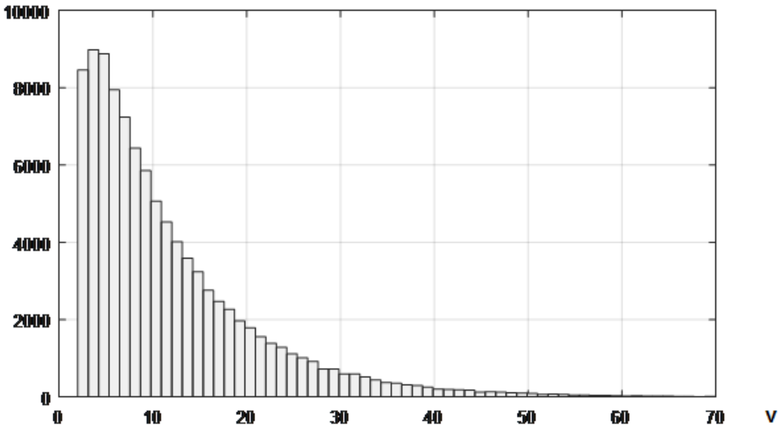

Figure 6. $p(v|y_1)$: Posterior distribution of v for the first dataset. $\mathrm{E}(v|\underline{y}_1) = 11.9672$, mode = 3.35 and 95% interval = $(2.31, 31.29)$.

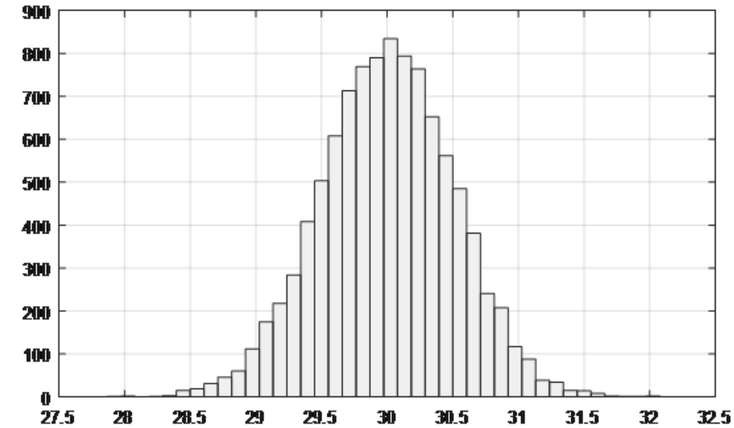

Figure 7. Mean(μ). Mean = 30.006, median = 30.009, mode = 30.002, variance = 0.2544, coverage = 95.13%.

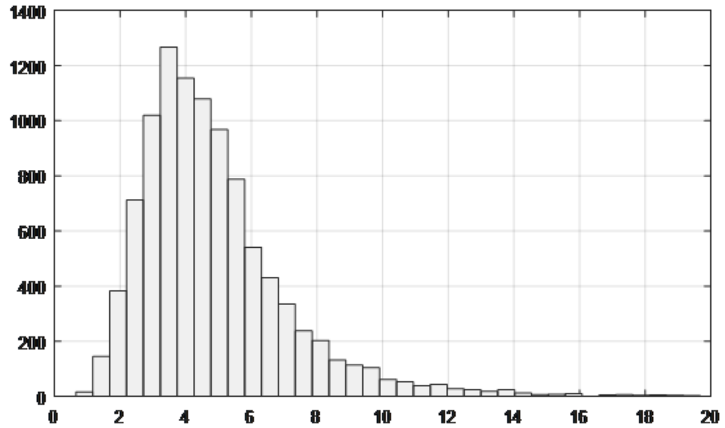

Figure 8. Mean(σ^2). Mean = 4.874, median = 4,397, mode = 3.49, variance = 5.9227, coverage = 94.72%.

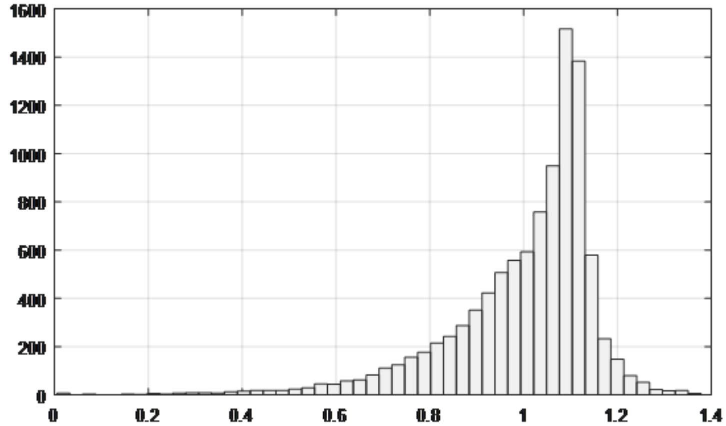

Figure 9. Mean(λ_{20}). Mean = 1.001, median = 1.049, mode = 1.09, variance = 0.0251, coverage = 99.54%.

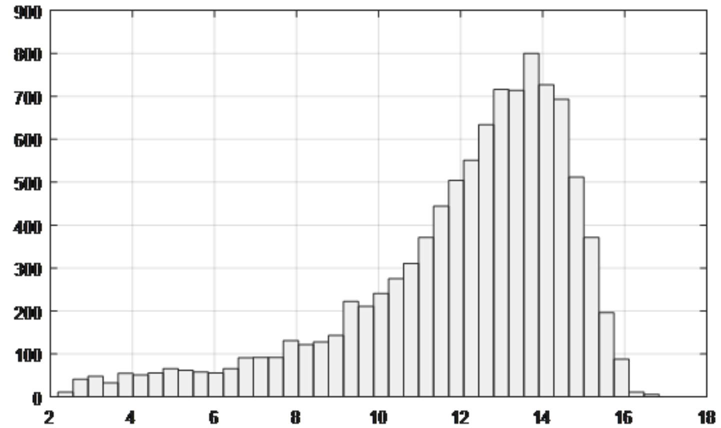

Figure 10. Mean(ν). Mean = 11.997, median = 12.705, mode = 13.74, variance = 7.5527, coverage = 94.87%.

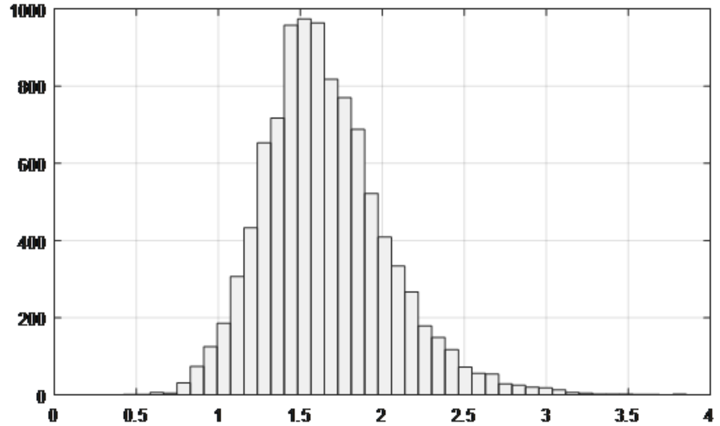

Figure 11. Mean(C_{pk}). Mean = 1.652, median = 1.608, mode = 1.52, variance = 0.1481, 95% interval = (0.942 ;2.439), coverage = 93.79%.

Our main interest is however C_{pk}. According to Figure 11, $Mean = 1.6519$ which is for all practical purposes the same as $C_{pk} = 1.6667$. The coverage percentage of the 95% Bayes confidence intervals is 93.79 which is also quite good.

One way to improve the coverage percentage for C_{pk} is to look at other priors for ν. In Fonseca, Ferreira and Migon (2008) two prior distributions were defined for ν, both based on Jeffreys' prior, the independence Jeffreys' prior and the Jeffreys-rule prior. Villa and Walker (2014) on the other hand constructed an objective prior for ν when the parameter is taken to be discrete. They found an objective criterion, based on loss functions instead of trying to define objective probabilities directly.

11.2 Frequentist Procedure

Consider again a sample of n observations $y_1, y_2, \ldots, y_n$ from a t-distribution with parameters ν, μ and σ^2. The likelihood function is given by

$$L(\mu, \sigma^2, \nu, \underline{y}) = \frac{\left[\Gamma(\frac{\nu+1}{2})\right]^n \nu^{\frac{n\nu}{2}}}{\left[\Gamma(\frac{\nu}{2})\right]^n \left[\Gamma(\frac{1}{2})\right]^n \sigma^n} \prod_{i=1}^{n} \left\{\nu + \left(\frac{y_i - \mu}{\sigma}\right)^2\right\}^{-\frac{1}{2}(\nu+1)}. \tag{15}$$

According to Fonseca et al. (2008), maximum likelihood estimation for the Student t-distribution model is very problematic because the likelihood function is ill-behaved for ν close to zero and may be ill-behaved when $\nu \to \infty$. The conditional maximum likelihood estimators for μ and σ^2 however exists and is given by $\hat{\mu}_\nu = \bar{y} = n^{-1} \sum_{i=1}^{n} y_i$ and $\hat{\sigma}_\nu^2 = ((\nu - 2)/\nu) \operatorname{Var}(Y)$, where $\operatorname{Var}(Y) = (n-1)^{-1} \sum_{i=1}^{n} (y_i - \bar{y})^2$. Substituting $\hat{\mu}_\nu$ and $\hat{\sigma}_\nu^2$ in (15) it follows that the profile likelihood function of ν is defined as

$$L(\nu, \underline{y}) = \frac{\left[\Gamma(\frac{\nu+1}{2})\right]^n \nu^{n\nu/2}}{\left[\Gamma(\frac{\nu}{2})\right]^n \left[\Gamma(\frac{1}{2})\right]^n \left[\left(\frac{\nu-2}{\nu}\right) \operatorname{Var}(y)\right]^{n/2}} \prod_{i=1}^{n} \left\{\nu + \frac{(y_i - \bar{y})^2}{\left(\frac{\nu-2}{\nu}\right) \operatorname{Var}(y)}\right\}^{-\frac{1}{2}(\nu+1)}, \quad \nu > 2. \tag{16}$$

Example 2

As in Example 1, ten thousand datasets are simulated and for each dataset $\hat{\mu}_\nu$ and $\hat{\sigma}_\nu^2$ are calculated. By using (16) the maximum likelihood estimates of ν are obtained. The percentage of maximum likelihood estimates of ν with values less than 30 is 65.98%. In Figure 12 the distribution of these maximum likelihood estimates are illustrated for values of $\hat{\nu}$ between 2 and 100. The percentage of estimates between 2 and 100 is 68.56%.

If for a certain dataset $\hat{\nu} > 100$, this dataset is considered to be drawn from a normal distribution.

The maximum likelihood estimates are then used as the parameter values for the parametric bootstrap procedure. For each dataset with parameter values $\tilde{\mu} = \hat{\mu}_\nu$, $\tilde{\sigma}^2 = \hat{\sigma}_\nu^2$ and $\tilde{\nu} = \hat{\nu}$, 10000 bootstrap samples are generated and are used to estimate the parameter values μ, σ^2, ν and C_{pk}. The mean values are calculated and it is also observed if the 95% confidence intervals contain the true parameter values. The mean values, variance and coverage percentage for C_{pk} are illustrated in Figure 13.

Although the parametric bootstrap procedure works quite well with normal data, it does not seem to do too well in the case of t-distributed data. From Figure 13 it can be seen that the coverage percentage of the 95% confidence intervals for C_{pk} is less than 85%. The mean C_{pk} value is 1.5259 which also does not compare well with the true parameter value of 1.6667. As shown in the Bayesian

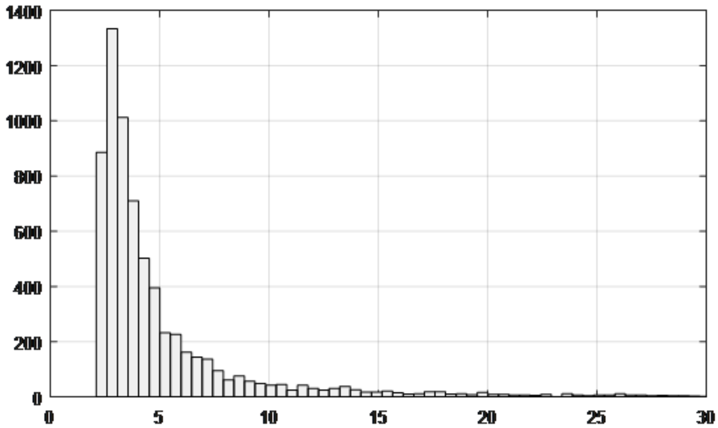

Figure 12. Distribution of the maximum likelihood estimates of ν ($2 < \hat{\nu} < 100$). Mean = 6.3202, median = 3.700, mode = 2.70, variance = 87.0416.

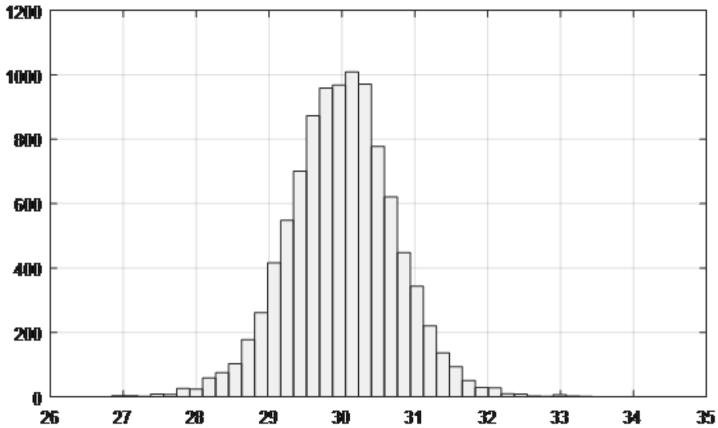

Figure 13. Mean(C_{pk}) – bootstrap method. Mean = 1.5259, median = 1.4747, mode = 1.37, variance = 0.1870, coverage = 84.78%.

case, the coverage percentage of the 95% confidence intervals for C_{pk} is 93.79% and the mean value is 1.6519.

It therefore seems that the Bayesian approach is more robust to non-normality than its frequentist alternative. The reason for this might be the ill-behaviour of the profile likelihood function of ν.

12. Conclusion

This paper developed a Bayesian method to analyse process capability indices C_{pl}, C_{pu} and C_{pk}. Multiple Bayesian testing strategies have been implemented on data representing four suppliers that produce piston rings for automobile engines studied by Chou (1994). A Bayesian version of Tukey's method is used for constructing simultaneous credibility intervals for all pairwise differences. A Bayesian procedure for testing all possible contrasts is also given. It is concluded that the magnitude of the Type I errors of the Bayesian method are correct. A Bayesian simulation procedure is also illustrated to find the best supplier or groups of suppliers. This method gives the same results as the Monte Carlo Integration method in Wu et al. (2016). It however seems that our method is much easier to perform and will probably take less time. In section 10, a sensitivity analysis regarding the prior choice is considered. It is concluded that priors of the form $p(\mu, \sigma^2) \propto \sigma^{-a}$ $(a > 2)$ will give better results for C_{pk} than Jeffreys' independence prior. In the last section, t-distributed data are analysed and it seems that the Bayesian approach is more robust to non-normality than its frequentist alternative.

Mathematical Appendix

Proof of Theorem 1

Since

$$\mu|\sigma^2, \underline{y} \sim N\left(\bar{y}, \frac{\sigma^2}{n}\right)$$

and

$$k = \frac{\nu S^2}{\sigma^2} \sim \chi^2_\nu,$$

it follows that

$$t|\tilde{t}, k \sim N\left(a\sqrt{k}, \frac{1}{9n}\right),$$

where $a = \tilde{t}/\sqrt{\nu}$. Therefore,

$$\begin{aligned}
p\left(t|\tilde{t}\right) &= \int_0^\infty f\left(t|\tilde{t}, k\right) f\left(k\right) dk \\
&= \frac{3\sqrt{n}}{2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)\sqrt{2\pi}} \int_0^\infty \exp\left[-\frac{9n}{2}\left(t - a\sqrt{k}\right)^2\right] k^{\frac{\nu}{2}-1} \exp\left[-\frac{k}{2}\right] dk \\
&= \frac{3\sqrt{n}\exp\left(-9nt^2/2\right)}{2^{\frac{\nu}{2}}\sqrt{2\pi}\Gamma\left(\frac{\nu}{2}\right)} \int_0^\infty k^{\frac{\nu}{2}-1}\left(\sum_{j=0}^\infty \frac{(9nta\sqrt{k})^j}{j!} \exp\left[-\frac{k}{2}\left(1 + 9na^2\right)\right]\right) dk.
\end{aligned}$$

Since

$$\int_0^\infty k^{\frac{1}{2}(v+j)-1}\exp\left[-\frac{k}{2}\left(1+9na^2\right)\right]dk = \frac{2^{\frac{1}{2}(v+j)}\Gamma\left(\frac{v+j}{2}\right)}{\left(1+9na^2\right)^{\frac{1}{2}(v+j)}},$$

and substituting $a = \tilde{t}/\sqrt{v}$, the posterior distribution of $t = (\mu - l)/(3\sigma) = C_{pl}$ follows as

$$p\left(t|\tilde{t}\right) = \frac{3\sqrt{n}\exp\left(-\frac{9nt^2}{2}\right)}{\Gamma\left(\frac{v}{2}\right)\sqrt{2\pi}}\sum_{j=0}^{\infty}\left(\frac{9nt\tilde{t}}{\sqrt{v}}\right)^j\frac{1}{j!}\frac{\Gamma\left(\frac{v+j}{2}\right)2^{\frac{1}{2}j}}{\left(1+\frac{9n}{v}\tilde{t}^2\right)^{\frac{1}{2}(v+j)}} \quad -\infty < t < \infty.$$

Proof of Theorem 2

The C_{pk} index can also be written as

$$C = C_{pk} = \frac{u-l-2\left|\mu-M\right|}{6\sigma},$$

where $M = (u+l)/2$. Since

$$\mu|\sigma^2, \underline{y} \sim N\left(\bar{y}, \frac{\sigma^2}{n}\right),$$

it follows that

$$\mu - M \sim N\left(\zeta, \frac{\sigma^2}{n}\right),$$

where $\zeta = \bar{y} - M$. Let $w = |\mu - M|$, then

$$p\left(w|\sigma^2, \underline{y}\right) = \frac{\sqrt{n}}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{n}{2}\frac{(w-\zeta)^2}{\sigma^2}\right\} + \frac{\sqrt{n}}{\sigma\sqrt{2\pi}}\exp\left(-\frac{n}{2}\frac{(w+\zeta)^2}{\sigma^2}\right);$$

see Kotz and Johnson (1993, p. 26). Now $C = b - \tilde{a}w$, where $\tilde{a} = 1/(3\sigma)$ and $b = C_p = (u-l)/(6\sigma)$. Also $w = -(C-b)\,1/\tilde{a}$ and $|dw/dc| = 1/\tilde{a}$. From this it follows that

$$p\left(C|\sigma^2, \underline{y}\right) = \frac{\sqrt{n}}{\tilde{a}\sigma\sqrt{2\pi}}\left\{\exp\left(-\frac{n}{2\tilde{a}^2\sigma^2}\left[C-b+\tilde{a}\zeta\right]^2\right) + \exp\left(-\frac{n}{2\tilde{a}^2\sigma^2}\left[C-b-\tilde{a}\zeta\right]^2\right)\right\}, \quad C < \tilde{b}\frac{S}{\sigma},$$

where $\tilde{b} = \hat{C}_p = (u-l)/(6s)$. Substituting for $\tilde{a}$, b and ζ and making use of the fact that $k = vS^2/\sigma^2 \sim \chi^2_v$, it follows that

$$p\left(C|k, \underline{y}\right) = \frac{3\sqrt{n}}{\sqrt{2\pi}}\left\{\exp\left(-\frac{9n}{2}\left[C - t^*\sqrt{\frac{k}{v}}\right]^2\right) + \exp\left(-\frac{9n}{2}\left[C - \tilde{t}\sqrt{\frac{k}{v}}\right]^2\right)\right\}, \quad C < \tilde{b}\sqrt{\frac{k}{v}}.$$

Therefore,

$$p\left(C|\underline{y}\right) = \frac{3\sqrt{n}}{\sqrt{2\pi}}\int_{\frac{C^2v}{\tilde{b}^2}}^{\infty}\left\{\exp\left(-\frac{9n}{2}\left[C - t^*\sqrt{\frac{k}{v}}\right]^2\right) + \exp\left(-\frac{9n}{2}\left[C - \tilde{t}\sqrt{\frac{k}{v}}\right]^2\right)\right\}\frac{k^{\frac{v}{2}-1}e^{-\frac{k}{2}}}{2^{\frac{v}{2}}\Gamma\left(\frac{v}{2}\right)}dk.$$

References

ANDREWS, D. F. AND MALLOWS, C. L. (1974). Scale mixtures of normal distributions. *Journal of the Royal Statistical Society: Series B*, **36**, 99–102.

CARLIN, B. P. AND POLSON, N. G. (1991). Inference for nonconjugate Bayesian models using the Gibbs sampler. *Canadian Journal of Statistics*, **19**, 399–405.

CHEN, K. AND PEARN, W. (1997). An application of non-normal process capability indices. *Quality and Reliability Engineering International*, **13**, 355–360.

CHOU, Y. (1994). Selecting a better supplier by testing process capability indices. *Quality Engineering*, **6**, 427–438.

CHOU, Y.-M. AND OWEN, D. B. (1989). On the distributions of the estimated process capability indices. *Communications in Statistics – Theory and Methods*, **18**, 4549–4560.

FONSECA, T. C., FERREIRA, M. A., AND MIGON, H. S. (2008). Objective Bayesian analysis for the Student-*t* regression model. *Biometrika*, **95**, 325–333.

GANESH, N. (2009). Simultaneous credible intervals for small area estimation problems. *Journal of Multivariate Analysis*, **100**, 1610–1621.

GEWEKE, J. (1993). Bayesian treatment of the independent Student-*t* linear model. *Journal of Applied Econometrics*, **8**, S19–S40.

HOFF, P. D. (2009). *A First Course in Bayesian Statistical Methods*, volume 580. Springer, New York.

HOSKINS, J., STUART, B., AND TAYLOR, J. (1988). *A Motorola Commitment: A Six Sigma Mandate. The Motorola Guide to Statistical Process Control for Continuous Improvement Towards Six Sigma Quality.* Motorola, Schaumburg, Illinois.

HUBELE, N. F., SHAHRIARI, H., AND CHENG, C. S. (1991). A bivariate process capability vector. *In* KEATS, J. B. AND MONTGOMERY, D. C. (Editors) *Statistical Process Control in Manufacturing*. Marcel Dekker, New York, 299–310.

KANE, V. E. (1986). Process capability indices. *Journal of Quality Technology*, **18**, 41–52.

KOTZ, S. AND JOHNSON, N. L. (1993). *Process Capability Indices*. Chapman & Hall, London.

LYTH, D. M. AND RABIEJ, R. J. (1995). Critical variables in wood manufacturing's process capability: Species, structure, and moisture content. *Quality Engineering*, **8**, 275–281.

NIVERTHI, M. AND DEY, D. (2000). Multivariate process capability: A Bayesian perspective. *Communication in Statistics – Simulation and Computation*, **29**, 667–687.

POLANSKY, A. M. (2006). Permutation methods for comparing process capabilities. *Journal of Quality Technology*, **38**, 254–266.

VILLA, C. AND WALKER, S. G. (2014). Objective prior for the number of degrees of freedom of a *t* distribution. *Bayesian Analysis*, **9**, 197–220.

WAKEFIELD, J. C., SMITH, A. F. M., RACINE-POON, A., AND GELFAND, A. E. (1994). Bayesian analysis of linear and non-linear population models by using the Gibbs sampler. *Journal of the Royal Statistical Society: Series C*, **43**, 201–221.

WU, C. C., SHIAU, J. J. H., PEARN, W. L., AND HUNG, H. N. (2016). A Bayesian approach for group supplier selections based on the popular process capability index C_{pk}. *Quality Technology & Quantitative Management*, **13**, 109–123.

Wu, C. W. and Pearn, W. L. (2005). Measuring manufacturing capablility for couples and wavelenth division multiplexers. *The International Journal of Advanced Manufacturing Technology*, **25**, 533–541.

Wu, C. W. and Pearn, W. L. (2006). A Bayesian approach for measuring EEPROM process capability based on one-sided indices c_{PU} and c_{PL}. *The International Journal of Advanced Manufacturing Technology*, **31**, 135–144.

Zellner, A. (1971). *An Introduction to Bayesian Inference in Econometrics*. Wiley, New York.

Manuscript received 2018-06-07, revised 2019-05-26, accepted 2019-07-18.