The Korean Communications
in Statistics Vol. 9, No. 1, 2002
pp. 221-227

A Study on the Role of Pivots in Bayesian Statistics[1)]

Hyungtae Hwang[2)]

Abstract

The concept of pivot has been widely used in various classical inferences. In this paper, it is proved by use of pivotal quantities that the Bayesian inferences can be arrived at the same results of classical inferences for the location-scale parameters models under the assumption of non-informative prior distributions. Some theorems are proposed in which the posterior distribution and the sampling distribution of a pivotal quantity coincide. The theorems are applied illustratively to some statistical models.

Keywords : pivotal quantity, sampling distribution, posterior distribution, location-scale parameters, shape parameter, non-informative prior distributions.

1. Introduction

It is not an unusual case that a result of Bayesian inference coincide with that of classical inference for parameters of interest. In such a case, it is commonly observed that the prior distribution is supposed to be a non-informative prior distribution of Jeffrey's type. Such cases can be easily found in Box and Tiao (1973) or Lee (1989) for the estimation problems, and in works of Hwang (2001) for the hypothesis testing problems.

More careful observation reveals that the result of Bayesian inference under the non-informative prior distribution frequently coincides with that of classical inference when the inferences are performed through pivotal quantities. A pivotal quantity is defined as a statistical quantity, a function of sample and parameters, of which sampling distribution does not depend on the parameter values. Here, sampling distribution means the distribution of a statistical quantity, possibly a function of sample and parameters, given the parameter values.

In this paper, we consider location-scale parameters models with suitable pivotal quantities and Jeffrey's non-informative prior distributions. In such cases, it is shown by change of variables methods that the posterior distributions of the pivotal quantities exactly coincide with their sampling distributions regardless of sample sizes. Our results are applied to some typical

1) The present research was conducted by the research fund of Dankook University in 2001.
2) Division of Natural Science, Dankook University, Seoul, 140-714, KOREA,
E-mail : hthwang@dankook.ac.kr

statistical models including the simple regression model and the analysis of variance model, illustratively. An example is also presented to show that our results can be applied to non-location-scale parameters models.

2. Main results

In this section, some theorems are introduced which show the coincidence of sampling distribution and posterior distribution of pivotal quantities from location-scale parameters models. We assume the Jeffrey's non-informative prior distributions in these theorems. A result for location parameter model is shown in the following **Theorem 1**.

Theorem 1. (The case of location parameter model)
Assume the followings:

$$p(x \mid \theta) = f(x-\theta),$$
$$\pi(\theta) \propto 1, \ -\infty < \theta < +\infty.$$

Let $w = w(x, \theta) = x - \theta$. Then, we have $p(w \mid \theta) = p(w \mid x) = f(w)$, that is, the sampling distribution and the posterior distribution of a pivotal quantity, w, are coincident to each other.

Proof. It is clear that $p(w \mid \theta) = f(w)$. Since $p(x, \theta) = p(x \mid \theta)\pi(\theta) \propto f(x-\theta)$, it can be shown that $p(x, w) \propto f(w)$ by substituting $x = x,\ w = x - \theta$. Thus we have $p(w \mid x) \propto p(x, w) \propto f(w)$, and hence the theorem follows.

The following **Theorem 2** shows the result for scale parameter model.

Theorem 2. (The case of scale parameter model)
Assume the followings:

$$p(x \mid \eta) = \frac{1}{\eta} f\left(\frac{x}{\eta}\right),$$
$$\pi(\eta) \propto \frac{1}{\eta}, \quad \eta > 0.$$

Let $w = w(x, \eta) = \frac{x}{\eta}$. Then, we have $p(w \mid \eta) = p(w \mid x) = f(w)$, that is, the sampling distribution and the posterior distribution of a pivotal quantity, w, are coincident to each other.

Proof. It is clear that $p(w \mid \eta) = f(w)$. Since $p(x, \eta) = p(x \mid \eta)\pi(\eta) \propto \frac{1}{\eta^2} f\left(\frac{x}{\eta}\right)$, it can be easily shown that $p(x, w) \propto \frac{1}{x} f(w)$ by substituting $x=x,\ w=\frac{x}{\eta}$. Thus we have $p(w \mid x) \propto p(x, w) \propto \frac{1}{x} f(w) \propto f(w)$, and hence the theorem follows.

The following **Theorem 3** shows the result for location-scale parameter model.

Theorem 3. (The case of location-scale parameter model)
Assume the followings:

$$p(x, y \mid \theta, \eta) = \frac{1}{\eta} f_1\left(\frac{x-\theta}{\eta}\right) \cdot \frac{1}{\eta} f_2\left(\frac{y}{\eta}\right),$$

$$\pi(\theta, \eta) \propto \frac{1}{\eta},\quad -\infty < \theta < +\infty,\ \eta > 0.$$

Let $w = w(x, y, \theta, \eta) = \frac{x-\theta}{y}$. Then, we have $p(w \mid \theta, \eta) = p(w \mid x, y)$, that is, the sampling distribution and the posterior distribution of a pivotal quantity, w, are coincident to each other.

Proof. From the assumptions, we have

$$p(w, y, \theta, \eta) \propto \frac{y}{\eta^3} f_1\left(\frac{wy}{\eta}\right) f_2\left(\frac{y}{\eta}\right), \tag{1}$$

Hence it follows that

by substituting $w=\frac{x-\theta}{y},\ y=y,\ \theta=\theta,\ \eta=\eta$.

Therefore, we have

$$p(w \mid \theta, \eta) \propto p(w, \theta, \eta) \propto \frac{1}{\eta} \int t f_1(wt) f_2(t) dt \propto \int t f_1(wt) f_2(t) dt. \tag{2}$$

On the other hand, it follows from (1) that

$$p(x, y, w, \eta) \propto \frac{y}{\eta^3} f_1\left(\frac{wy}{\eta}\right) f_2\left(\frac{y}{\eta}\right),$$

by substituting $x=x,\ y=y,\ w=\frac{x-\theta}{y},\ \eta=\eta$.

Hence we have by integration,

$$p(x, y, w) \propto \frac{1}{y} \int t f_1(wt) f_2(t) dt.$$

Thus we have

$$p(w \mid x, y) \propto p(x, y, w) \propto \frac{1}{y} \int t f_1(wt) f_2(t) dt \propto \int t f_1(wt) f_2(t) dt. \tag{3}$$

Then the theorem follows from (2) and (3).

In fact, it can be shown by the analogous methods that the result of **Theorem 3** holds when x, θ, w are multivariate quantities. Hence we state the result without proof in the following **Theorem 4**.

Theorem 4. (The multivariate case)
Assume the followings:

$$p(x_1, \cdots, x_k, y \mid \theta_1, \cdots, \theta_k, \eta) = \frac{1}{\eta^k} f_1(\frac{x_1-\theta_1}{\eta}, \cdots, \frac{x_k-\theta_k}{\eta}) \cdot \frac{1}{\eta} f_2(\frac{y}{\eta}),$$

$$\pi(\theta_1, \cdots, \theta_k, \eta) \propto \frac{1}{\eta}, \quad -\infty < \theta_i < +\infty, \quad \eta > 0,$$

$$(w_1, \cdots, w_k) = (\frac{x_1-\theta_1}{y}, \cdots, \frac{x_k-\theta_k}{y})$$

Then we have $p(w_1, \cdots, w_k \mid \theta_1, \cdots, \theta_k, \eta) = p(w_1, \cdots, w_k \mid x_1, \cdots, x_k, y)$.

3. Some applications

In this section, we provide some applications of theorems in section 2. The results of the following examples 1 ~4 can be found in Box and Tiao (1973) or Lee (1989), but the proofs of them generally contain long series of substitutions and integrals. Here, we provide much simpler derivations of them by use of our theorems. In the following examples, there are sufficient statistics for each problem. Hence we may start with these sufficient statistics for the following examples.

Example 1. (A normal population)

Let $x_1, \cdots, x_n$ be a random sample from the normal distribution $N(\mu, \sigma^2)$, and assume the joint prior distribution, $\pi(\mu, \sigma) \propto 1/\sigma$. The sample mean and sample variance are denoted by $\overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$ and $s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x})^2$, respectively.

Let $w_1 = \frac{\overline{x}-\mu}{s/\sqrt{n}}$. Then w_1 is a pivotal quantity whose sampling distribution is $t(n-1)$. Note that the sampling distributions of $\frac{\overline{x}-\mu}{\sigma/\sqrt{n}}$ and $\frac{(n-1)s^2}{\sigma^2}$ are $N(0,1)$ and $\chi^2(n-1)$, respectively, and are independent. Hence it follows by applying **Theorem 3** that the posterior distribution of w_1 is also given by $t(n-1)$.

Now let $w_2 = \frac{s}{\sigma}$. Then the posterior distribution of w_2 is identical to its sampling distribution by **Theorem** 2. Note that $\frac{(n-1)s^2}{\sigma^2}=(n-1)\ w_2^{\ 2}$ and its sampling distribution is given by $\chi^2(n-1)$. Hence the posterior distribution of $\frac{(n-1)s^2}{\sigma^2}$ is also given by $\chi^2(n-1)$.

Example 2. (Two normal populations with equal variances)

Let $x_1,\cdots,x_m$ and $y_1,\cdots,y_n$ be independent random samples from $N(\mu_1,\sigma^2)$ and $N(\mu_2,\sigma^2)$, respectively. Suppose $\pi(\mu_1,\mu_2,\sigma)\propto 1/\sigma$ for joint prior distribution. Let $\bar{x}$, $\bar{y}$, s_1^2, s_2^2 denote the corresponding sample means and sample variances respectively, and $s_p^2=\frac{(m-1)\ s_1^{\ 2}+(n-1)\ s_2^{\ 2}}{(m+n-2)}$ denote the pooled sample variance.

First, letting $w_1=\frac{(\bar{x}-\bar{y})-(\mu_1-\mu_2)}{s_p(1/m+1/n)^{1/2}}$, w_1 is a pivotal quantity whose sampling distribution is $t(m+n-2)$. Hence the posterior distribution of w_1 is also $t(m+n-2)$ by applying **Theorem 3**.

Next, let $w_2=s_p/\sigma$. Then w_2 is a pivotal quantity. Thus the posterior distribution of w_2 is identical to its sampling distribution by **Theorem** 2. Since the sampling distribution of $(m+n-2)\ s_p^{\ 2}/\sigma^2$ is given by $\chi^2(m+n-2)$ and it is a function of w_2, it follows that the posterior distribution of $(m+n-2)\ s_p^{\ 2}/\sigma^2$ is also given by $\chi^2(m+n-2)$.

Example 3. (Simple linear regression model)

Consider the following simple linear regression model:

$$y_i \sim \textit{independently}\ N(\alpha+\beta x_i,\sigma^2),\quad i=1,\cdots,n,$$

where $x_1,\cdots,x_n$ are known constants. Suppose $\pi(\alpha,\beta,\sigma)\propto 1/\sigma$, and let a, b and s^2 denote the ordinary least squares estimators of α and β, and the mean squared errors, respectively.

First, let $w_1=\frac{(b-\beta)}{s/\,S_x}$, where $S_x^{\ 2}=\sum_{i=1}^{n}(x_i-\bar{x})^2$. Then w_1 is a pivotal quantity whose sampling distribution is $t(n-2)$. Hence the posterior distribution of w_1 is shown to be $t(n-2)$ by applying **Theorem 3**.

Next, let $w_2=\frac{y^*-\alpha-\beta x^*}{s\sqrt{1/n+(x^*-\bar{x})^2/\,S_x^{\ 2}}}$ where $y^*=a+bx^*$. Then w_2 is a pivotal quantity whose sampling distribution is $t(n-2)$. Hence the posterior distribution of w_1 is shown to be $t(n-2)$ by applying **Theorem 3**.

Finally, let $w_3 = s/\sigma$. Then, it can be shown as in **Example 2** that the posterior distribution of $(n-2)s^2/\sigma^2 = (n-2)w_3^2$ is given by $\chi^2(n-2)$.

Example 4. (Analysis of variance model)

Let $\overline{x_1}, \cdots, \overline{x_k}$ be independent treatment means with replication numbers $n_1, \cdots, n_k$, respectively. The sampling distribution of $\overline{x_i}$ is assumed to be $N(\theta_i, \sigma^2/n_i)$, for $i = 1, \cdots, k$, respectively. Let s^2 be independent of $\overline{x_1}, \cdots, \overline{x_k}$, of which sampling distribution is given by $\frac{rs^2}{\sigma^2} \mid \sigma^2 \sim \chi^2(r)$, where r represents the error degrees of freedom. For prior distribution, we assume $\pi(\theta_1, \cdots, \theta_k, \sigma) \propto 1/\sigma$.

Let $(w_1, \cdots, w_k) = (\frac{\overline{x_1} - \theta_1}{s/\sqrt{n_1}}, \cdots, \frac{\overline{x_k} - \theta_k}{s/\sqrt{n_k}})$. Then, $(w_1, \cdots, w_k)$ is a pivotal vector whose joint sampling distribution is known to be $t_k(r)$, that is, *k-variate t* distribution with r degrees of freedom. Thus it follows by **Theorem 4** that the joint posterior distribution of $(w_1, \cdots, w_k)$ is also given by $t_k(r)$.

Let $F = \frac{\sum_{i=1}^{k} n_i [(\overline{x_i} - \overline{x_.}) - (\theta_i - \overline{\theta})]^2}{(k-1)s^2}$, where $\overline{x_.} = \frac{1}{n} \sum_{i=1}^{k} n_i \overline{x_i}$ and $\overline{\theta} = \frac{1}{n} \sum_{i=1}^{k} n_i \theta_i$ for $n = n_1 + \cdots + n_k$. From the fact that F is a function of pivotal vector $(w_1, \cdots, w_k)$, we know that the posterior distribution of F is identical to its sampling distribution, and F itself is a pivotal quantity. Hence the sampling distribution of F is shown to be $F(k-1, r)$ regardles of $\theta_1, \cdots, \theta_k$, since the sampling distribution of F is $F(k-1, r)$ when $\theta_1 = \cdots = \theta_k$ Therefore, it follows that the posterior distribution of F is given by $F(k-1, r)$.

As a final example of this section, we consider the Pareto probability model with a shape parameter. Although our theorems are restricted to location-scale parameters models, the following **Example 5** shows that our theorems can be applied to find the posterior distribution of a pivotal quantity from non-location-scale parameters models.

Example 5. (Pareto probability model with a shape parameter)

Let $x_1, \cdots, x_n$ be a random sample from Pareto distribution whose distribution function is given as follows:

$$F(x:\theta)=1-(1+x)^{-\theta}, \quad x>0,$$

where $\theta>0$ is a shape parameter. Let $w=2\theta\sum_{i=1}^{n}\ln(1+x_i)$. Then the sampling distribution of w is shown to be $\chi^2(2n)$ by applying the method of probability integral transformation (Bain and Engelhardt(1992), p366). Hence letting $\eta=1/\theta$, η can be considered as a scale parameter for $2\sum_{i=1}^{n}\ln(1+x_i)$. Now assume the prior distribution of θ to be $\pi(\theta)\propto 1/\theta$, $\theta>0$, so that $\pi(\eta)\propto 1/\eta$, $\eta>0$ by change of variable. Thus it follows from **Theorem 2** that the posterior distribution of w is also given by $\chi^2(2n)$.

4. Concluding remarks

In this paper, we studied the properties of pivotal quantities in Bayesian Statistics, which say that the results of classical Statistics and Bayesian Statistics may be eventually coincide by somewhat systematic methods. The applications of them may be considered for more various problems, and the asymptotic versions of them for more general probability models are left for a future study.

References

[1] Bain, L.J. and Engelhardt, M. (1992). *Introduction to Probability and Mathematical Statistics*, PWS-KENT Publishing Company, Boston.

[2] Box, G.E.P. and Tiao, G.C. (1973). *Bayesian inference in statistical analysis*, Addison -Wesley Publishing Company, Inc.

[3] Hwang, H.T. (2001). A Bayesian hypothesis testing procedure possessing the concept of significance level, submitted to *The Korean Communications in Statistics*.

[4] Lee, P.M. (1989). *Bayesian Statistics : an Introduction*, Edward Arnold, London.

[2001년 10월 접수, 2002년 2월 채택]