

Noninformative Priors for Stress-Strength System in the Burr-Type X Model

Dal Ho Kim, Sang Gil Kang¹ and Jang Sik Cho²

ABSTRACT

In this paper, we develop noninformative priors that are used for estimating the reliability of stress-strength system under the Burr-type X model. A class of priors is found by matching the coverage probabilities of one-sided Bayesian credible interval with the corresponding frequentist coverage probabilities. It turns out that the reference prior as well as the Jeffreys prior are the second order matching prior. The propriety of posterior under the noninformative priors is proved. The frequentist coverage probabilities are investigated for small samples via simulation study.

Key Words : Jeffreys Prior; Reference Prior; Matching Priors; Burr-Type X Stress-Strength Model; Frequentist Coverage Probability.

1. INTRODUCTION

Consider the following stress-strength system, where Y is the strength of a system subject to stress X . The system fails at any moment the applied stress is greater than its strength. Then reliability of the system is given by

$$w_1 = P(X < Y). \quad (1.1)$$

This model was first considered by Birnbaum (1956), and has since then found applications in many areas, especially in structural and aircraft industries. Basu (1985) and Johnson (1988) contain comprehensive reviews of frequentist inference for stress-strength models, although Johnson (1988) contains a small Bayesian component.

¹Department of Statistics, Kyungpook National University, Taegu, 702-701, Korea.

²Department of Statistics, Kyungsoo University, Pusan, 608-736, Korea.

The present paper focuses exclusively on Bayesian inference for w_1 . The emphasis is on developing noninformative priors. Although subjective Bayesian are often critical of such priors, these priors have clear pragmatic appeal especially when prior information is vague in nature.

The most frequently used noninformative prior is Jeffreys' (1961) prior, which is proportional to the positive square root of the determinant of the Fisher information matrix. In the one-parameter case, Welch and Peers (1963) proved that a one-sided credible interval from Jeffreys' prior matches the corresponding frequentist coverage probability up to $o(n^{-\frac{1}{2}})$, where n is the sample size.

In spite of its success in one-parameter problems, Jeffreys' prior frequently runs into serious difficulties in the presence of nuisance parameter(s). To overcome these difficulties, Stein (1985) extended the results in Welch and Peers (1963) and Peers (1965) and introduced a method to find a prior which requires the frequentist coverage probability of the posterior region of a real-valued parametric function to match the nominal level with a remainder of $o(n^{-\frac{1}{2}})$. Tibshirani (1989) reconsidered the case when the real-valued parameter of interest is orthogonal to the nuisance parameter vector in the sense of Cox and Reid (1987). These priors, as usually referred to as 'first order' matching priors, were further studied in Datta and Ghosh (1995a). Recently, Mukerjee and Ghosh (1997) developed a 'second order', that is, $o(n^{-1})$, matching prior. They extended the finding in Mukerjee and Dey (1993) to the case of multiple nuisance parameters based on quantiles, and also developed a second order matching prior based on distribution function.

On the other hand, Berger and Bernardo (1989, 1992) extended Bernardo's (1979) reference prior approach, giving a general algorithm to derive a reference prior by splitting the parameters into several groups according to their order of inferential importance. This approach is very successful in various practical problems. Quite often reference priors satisfy the matching criterion described earlier.

Thompson and Basu (1993) derived reference priors when the stress and strength are both exponentially distributed. It turns out that in such cases, the reference priors agree with Jeffreys' prior. Lee (1998) investigated matching priors in exponential stress-strength model. Lee, Sun and Basu (1997) derived matching priors and Ghosh and Yang (1996) derived matching priors as well as reference priors when the stress and strength are both normally distributed. Sun, Ghosh and Basu (1998) derived Jeffreys' prior, reference priors and matching priors when the stress and strength have both Weibull distributions. It turns

out that none of the Jeffreys' prior and the reference priors is a matching prior. Their study shows that the matching prior performs better than Jeffreys' prior and reference priors in meeting the target coverage probabilities.

Ahmad, Fakhry and Jaheen (1997) studied the empirical Bayes estimation of $w_1 = P(X < Y)$ when X and Y are independent Burr-type X random variables. Our interest is to derive noninformative priors for Burr-type X stress-strength models when w_1 is the parameter of interest. It turns out that reference priors, second order matching prior and Jeffreys' prior are the same.

The outline of the remaining sections is as follows. In Section 2, we derive Fisher information matrices under original parameterization and orthogonal reparametrization. Then we derive Jeffreys prior, the second order matching prior and the reference prior. In Section 3, the propriety of posterior under the derived noninformative prior is proved. Also the marginal density of w_1 under this prior is given. In Section 4, simulated frequentist coverage probabilities under the derived noninformative prior are provided for small samples.

2. NONINFORMATIVE PRIORS

We denote the Burr-type X distribution with probability density function(pdf)

$$f(x; \eta) = 2\eta x e^{-x^2} (1 - e^{-x^2})^{\eta-1}, x > 0, \eta > 0, \tag{2.1}$$

as Burr-type X (η). Sartawi and Abu-Salih (1991) obtained the Bayesian prediction bounds for the order statistics in the one sample and two sample cases under the Burr-type X model. Jaheen (1996) considered the Burr-type X (η) distribution as a lifetime model and obtained Bayes and empirical Bayes estimates of reliability and failure rate functions of the model in one sample case.

Now let X be a random variable with Burr-type X (η_1) and Y is another Burr-type X (η_2) random variable where X and Y are independent. One can see easily that

$$w_1 = P(X < Y) = \frac{\eta_2}{\eta_1 + \eta_2}. \tag{2.2}$$

It is interesting to note that

(i) If W is exponential random variable with failure rate η_1 and Z is exponential random variable with failure rate η_2 where W and Z are independent, then

$$P(Z < W) = \frac{\eta_2}{\eta_1 + \eta_2} = P(X < Y) = w_1.$$

(ii) Let U be Pareto with parameters η_1 and ξ whose pdf is

$$f(u) = \eta_1 \frac{\xi^{\eta_1}}{u^{\eta_1+1}}, \quad u > \xi, \quad \xi > 0,$$

and V be Pareto with parameters η_2 and ξ . When U and V are independent we have

$$P(V < U) = P\left[\ln\left(\frac{V}{\xi}\right) < \ln\left(\frac{U}{\xi}\right)\right] = P(Z < W) = \frac{\eta_2}{\eta_1 + \eta_2} = P(X < Y) = w_1,$$

where $W = \ln\left(\frac{U}{\xi}\right)$ has exponential distribution with failure rate η_1 . Therefore, the results obtained in this paper can be generalized to the exponential and Pareto distributions. However, unlike the exponential distribution, when $\eta = 1$ the failure rate function of Burr-type X (η) model is linear and an increasing function of time. The linear failure rate function has many applications. Another lifetime model that has a linear failure rate is the Rayleigh distribution, which is a special case of the Weibull family. This type of situation would exist if failure occurs randomly and from wear out or deterioration.

Suppose that X_1, \dots, X_m are independent random samples from Burr-type X (η_1), and independently, Y_1, \dots, Y_n are independent random samples from Burr-type X (η_2). The log-likelihood function of (η_1, η_2) is

$$l(\eta_1, \eta_2) \propto m \log(\eta_1) + n \log(\eta_2) - \eta_1 \sum_{i=1}^m \log(1 - e^{-x_i^2}) - \eta_2 \sum_{j=1}^n \log(1 - e^{-y_j^2}).$$

By a simple algebra, the Fisher information matrix of (η_1, η_2) is given by

$$I(\eta_1, \eta_2) = \begin{pmatrix} \frac{m}{\eta_1^2} & 0 \\ 0 & \frac{n}{\eta_2^2} \end{pmatrix}$$

so that $|I(\eta_1, \eta_2)|^{1/2} \propto (\eta_1 \eta_2)^{-1}$. Then Jeffreys' prior is

$$\pi_J(\eta_1, \eta_2) \propto \frac{1}{\eta_1 \eta_2}. \quad (2.3)$$

Now let $\mathbf{X} = (X_1, \dots, X_m)$ and $\mathbf{Y} = (Y_1, \dots, Y_n)$. Under the Burr-type X stress-strength model, the parameter of interest is $w_1 = \eta_2/(\eta_1 + \eta_2)$. Then our interest is to find the probability matching prior for w_1 .

For a prior π , let $w_1^{1-\alpha}(\pi; \mathbf{X}, \mathbf{Y})$ denote the 100(1- α)th percentile of the posterior distribution of w_1 , that is,

$$P^\pi[w_1 \leq w_1^{1-\alpha}(\pi; \mathbf{X}, \mathbf{Y}) | \mathbf{X}, \mathbf{Y}] = 1 - \alpha. \quad (2.4)$$

We want to find priors π for which

$$P[w_1 \leq w_1^{1-\alpha}(\pi; \mathbf{X}, \mathbf{Y})|\eta_1, \eta_2] = 1 - \alpha + o(n^{-u}) \quad (2.5)$$

for some $u > 0$, as n goes to infinity. Priors π satisfying (2.5) are called matching priors. If $u = 1/2$, then π is referred to as a first order matching prior, while if $u = 1$, π is referred to as a second order matching prior.

In order to find such matching priors π , it is convenient to introduce orthogonal parametrization (Cox and Reid, 1987; Tibshirani, 1989). To this end, let

$$w_1 = \frac{\eta_2}{\eta_1 + \eta_2}, \quad w_2 = \eta_1^m \eta_2^n. \quad (2.6)$$

With this parametrization, the likelihood has the alternate representation.

$$\begin{aligned} L(w_1, w_2) \propto & w_2 \prod_{i=1}^m [1 - \exp(-x_i^2)]^{w_2^{1/(m+n)} \left(\frac{w_1}{1-w_1}\right)^{-n/(m+n)}} \\ & \times \prod_{j=1}^n [1 - \exp(-y_j^2)]^{w_2^{1/(m+n)} \left(\frac{w_1}{1-w_1}\right)^{m/(m+n)}} \end{aligned} \quad (2.7)$$

Based on (2.7), the Fisher information matrix is given by

$$I(w_1, w_2) = \begin{pmatrix} \frac{mn}{m+n} \frac{1}{w_1^2(1-w_1)^2} & 0 \\ 0 & \frac{1}{(m+n)w_2^2} \end{pmatrix}.$$

Thus w_1 is orthogonal to w_2 in the sense of Cox and Reid (1987). Following Tibshirani (1989), the class of first order matching prior is characterized by

$$\pi_M^{(1)}(w_1, w_2) \propto \frac{1}{w_1(1-w_1)} d(w_2), \quad (2.8)$$

where $d(\cdot)$ is an arbitrary function differentiable in its arguments.

Clearly the class of prior given in (2.8) is quite large and it is important to narrow down this class of priors. To this end, we consider the class of second order probability matching priors as given in Mukerjee and Ghosh (1997) (see also Mukerjee and Dey, 1993). A second order matching prior is also of the form (2.8), but the function d must satisfy an additional differential equation (cf. (2.10) of Mukerjee and Ghosh (1997)), namely

$$\frac{1}{6} d(w_2) \frac{\partial}{\partial w_1} (I_{11}^{-\frac{3}{2}} L_{1,1,1}) + \frac{\partial}{\partial w_2} \{I_{11}^{-\frac{1}{2}} L_{112} I^{22} d(w_2)\} = 0, \quad (2.9)$$

where

$$L_{1,1,1} = E\left[\left(\frac{\partial \log L}{\partial w_1}\right)^3\right] = \frac{c}{w_1^3(1-w_1)^3}, c = \text{a constant}$$

$$L_{112} = E\left[\frac{\partial^3 \log L}{\partial^2 w_1 \partial w_2}\right] = -\frac{nm}{(m+n)^2} \frac{1}{w_1^2(1-w_1)^2 w_2},$$

and

$$\begin{pmatrix} I^{11} & I^{12} \\ I^{21} & I^{22} \end{pmatrix} = \begin{pmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{pmatrix}^{-1} = \begin{pmatrix} (m+n)w_2^2 & 0 \\ 0 & \frac{m+n}{mn} w_1^2 (1-w_1)^2 \end{pmatrix}.$$

Then (2.9) simplifies to

$$\frac{\partial}{\partial w_2} \left\{ -\sqrt{\frac{mn}{m+n}} \frac{1}{w_1(1-w_1)} w_2 d(w_2) \right\} = 0. \quad (2.10)$$

Hence the set of solution of (2.10) is of the form

$$d(w_2) = \frac{1}{w_2}.$$

Thus the unique second order matching prior is given by

$$\pi_M^{(2)}(w_1, w_2) \propto \frac{1}{w_1(1-w_1)w_2}.$$

Back to (η_1, η_2) formulation the above second order matching prior transforms to

$$\pi_M^{(2)}(\eta_1, \eta_2) \propto \frac{1}{\eta_1 \eta_2} \quad (2.11)$$

which is Jeffreys' prior. The invariance of the first order matching priors is proved in Datta and Ghosh (1996), and that of first and second order matching priors is proved in Mukerjee and Ghosh (1997).

Other possible noninformative prior is the reference prior of Bernardo (1979). Choosing rectangular compacts for each one of w_1 and w_2 , when w_1 is the parameter of interest, due to the orthogonality of w_1 with w_2 , from Datta and Ghosh (1995b), the reference prior as well as the reverse reference prior is given by,

$$\pi_R(w_1, w_2) \propto \frac{1}{w_1(1-w_1)w_2}.$$

This prior is clearly a second order probability matching prior.

Thus it turns out that the Jeffreys' prior, the reference priors and the second order matching prior for (η_1, η_2) are the same in the Burr-type X case. Therefore we denote by

$$\pi(\eta_1, \eta_2) \equiv \pi_J(\eta_1, \eta_2) = \pi_R(\eta_1, \eta_2) = \pi_M^{(2)}(\eta_1, \eta_2). \quad (2.12)$$

Note that the same phenomenon was observed in Ghosh and Sun (1998) for the exponential case.

3. IMPLEMENTATION OF THE BAYESIAN PROCEDURE

We now prove that the posterior is proper under the noninformative prior given in (2.12).

Theorem 1. The posterior distribution of (η_1, η_2) under the prior π , (2.12), is proper.

Proof. Note that

$$\begin{aligned} & \int_0^\infty \int_0^\infty L(\eta_1, \eta_2) \frac{1}{\eta_1 \eta_2} d\eta_1 d\eta_2 \\ &= 2^{m+n} \prod_{i=1}^m \frac{x_i}{(1 - \exp(-x_i^2))} \prod_{j=1}^n \frac{y_j}{(1 - \exp(-y_j^2))} \exp\left(-\sum_{i=1}^m x_i^2 - \sum_{j=1}^n y_j^2\right) \\ &\times \int_0^\infty \int_0^\infty \eta_1^{m-1} \eta_2^{n-1} \left[\prod_{i=1}^m (1 - \exp(-x_i^2))\right]^m \left[\prod_{j=1}^n (1 - \exp(-y_j^2))\right]^n d\eta_1 d\eta_2 \\ &= 2^{m+n} \prod_{i=1}^m \frac{x_i}{(1 - \exp(-x_i^2))} \prod_{j=1}^n \frac{y_j}{(1 - \exp(-y_j^2))} \exp\left(-\sum_{i=1}^m x_i^2 - \sum_{j=1}^n y_j^2\right) \\ &\times \frac{\Gamma(m)}{(\sum_{i=1}^m \log[1 - \exp(-x_i^2)])^m} \frac{\Gamma(n)}{(\sum_{j=1}^n \log[1 - \exp(-y_j^2)])^n} \\ &< \infty. \end{aligned}$$

This completes the proof.

Next, we provide the marginal density of w_1 under the above noninformative prior.

Theorem 2. Under the prior π , (2.12), the marginal posterior density of $w_1 = \eta_2/(\eta_1 + \eta_2)$, is given by

$$\pi(w_1|\mathbf{X}, \mathbf{Y}) \propto w_1^{n-1}(1-w_1)^{m-1}\left(\frac{1}{h(w_1)}\right)^{m+n}, \quad (3.1)$$

where

$$h(w_1) = \log\left[\prod_{i=1}^m(1 - \exp(-x_i^2))\right] + w_1 \log\left[\frac{\prod_{j=1}^n(1 - \exp(-y_j^2))}{\prod_{i=1}^m(1 - \exp(-x_i^2))}\right].$$

The normalizing constant for the marginal density of w_1 requires a one dimensional integration. Therefore we have the marginal posterior density of w_1 , and so it is easy to compute the marginal moment of w_1 .

4. SMALL SAMPLE SIMULATION STUDY

We evaluate the frequentist coverage probability by investigating the credible interval of the marginal posterior density of w_1 under the noninformative prior π given in (2.12) when m and n are small. That is to say, the frequentist coverage of a $(1 - \alpha)$ th posterior quantile should be close to $1 - \alpha$. This is done numerically. Table 1 gives numerical values of the frequentist coverage probabilities of 0 to 0.05 (0 to 0.95) posterior quantiles for the our prior. The computation of these numerical values is based on the following algorithm for any fixed true (η_1, η_2) and any prespecified probability value α . Here α is 0.05 (0.95). Let $w_1^\pi(\alpha|\mathbf{X}, \mathbf{Y})$ be the posterior α -quantile of w_1 given (\mathbf{X}, \mathbf{Y}) . That is to say, $F(w_1^\pi(\alpha|\mathbf{X}, \mathbf{Y})|\mathbf{X}, \mathbf{Y}) = \alpha$, where $F(\cdot|\mathbf{X}, \mathbf{Y})$ is the marginal posterior distribution of w_1 . Then the frequentist coverage probability of this one sided credible interval of w_1 is

$$P_{(\eta_1, \eta_2)}(\alpha; w_1) = P_{(\eta_1, \eta_2)}(0 < w_1 \leq w_1^\pi(\alpha|\mathbf{X}, \mathbf{Y})). \quad (4.1)$$

The estimated $P_{(\eta_1, \eta_2)}(\alpha; w_1)$ when $\alpha = 0.05(0.95)$ is shown in Table 1.

In particular, for fixed (η_1, η_2, m, n) , we take 10,000 independent random samples of m stresses $\mathbf{X} = (X_1, \dots, X_m)$ from Burr-type X (η_1) and n strengths $\mathbf{Y} = (Y_1, \dots, Y_n)$ from Burr-type X (η_2) . Note that under the prior π , for fixed \mathbf{X} and \mathbf{Y} , $w \leq w_1^\pi(\alpha|\mathbf{X}, \mathbf{Y})$ if and only if $F(w|\mathbf{X}, \mathbf{Y}) \leq \alpha$. Under the prior π , $P_{(\eta_1, \eta_2)}(\alpha; w_1)$ can be estimated by the relative frequency of $F(w_1^\pi|\mathbf{X}, \mathbf{Y}) \leq \alpha$. For the cases presented in Table 1, we see that the noninformative prior π meet

Table 1: Frequentist Coverage Probability of 0.05 (0.95) Posterior Quantiles of w_1

η_1	η_2	m	n	π	
				0.05	0.95
1	1	2	2	0.0508	0.9472
1	1	2	3	0.0501	0.9486
1	1	5	5	0.0483	0.9465
1	1	10	10	0.0474	0.9499
3	2	2	2	0.0468	0.9529
3	2	2	3	0.0519	0.9527
3	2	5	5	0.0491	0.9517
3	2	10	10	0.0524	0.9493
5	5	2	2	0.0486	0.9509
5	5	2	3	0.0454	0.9525
5	5	5	5	0.0492	0.9478
5	5	10	10	0.0516	0.9526

very well the target coverage probabilities. Also note that the results in Table 1 are not much sensitive to the change of the values of (η_1, η_2) .

REFERENCES

- Ahmad, K. E., Fakhry, M. E. and Jaheen, Z. F. (1997). "Empirical Bayes estimation of $P(X < Y)$ and characterizations of Burr-type X model," *Journal of Statistical Planning and Inference*, **64**, 297-308.
- Basu, A. P. (1985). "Estimation of the reliability of complex system - a survey," *The Frontiers of Modern Statistical Inference Procedure*, Ed. E. J. Dudewicz, American Science Press, Columbus, 271-287.
- Berger, J. O. and Bernardo, J. M. (1989). "Estimating a product of means : Bayesian analysis with reference priors," *Journal of the American Statistical Association*, **84**, 200-207.
- Berger, J. O. and Bernardo, J. M. (1992). "On the development of reference

- priors (with discussion)," *Bayesian Statistics 4*, Eds. J. M. Bernardo, et al., Oxford University Press, Oxford, 35-60.
- Bernardo, J. M. (1979). "Reference posterior distributions for Bayesian inference (with discussion)," *Journal of Royal Statistical Society*, B, **41**, 113-147.
- Birnbaum, Z. M. (1956). "On a use of the Mann-Whitney statistic," In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Volume I: Contributions to the Theory of Statistics and Probability*, University of California Press, Berkeley, 13-17.
- Cox, D. R. and Reid, N. (1987). "Orthogonal parameters and approximate conditional inference (with discussion)," *Journal of Royal Statistical Society*, B, **49**, 1-39.
- Datta, G. S. and Ghosh, J. K. (1995a). "On priors providing frequentist validity for Bayesian inference," *Biometrika*, **82**, 37-45.
- Datta, G. S. and Ghosh, M. (1995b). "Some remarks on noninformative priors," *Journal of the American Statistical Association*, **90**, 1357-1363.
- Datta, G. S. and Ghosh, M. (1996). "On the invariance of noninformative priors," *The Annals of Statistics*, **24**, 141-159.
- Ghosh, M. and Sun, D. (1998). "Recent development of Bayesian inference for stress-strength models," *Frontiers of Reliability*, Eds. A. P. Basu, S. K. Basu and S. Mukhopadhyay, World Scientific, New Jersey, 143-158.
- Ghosh, M. and Yang, M. C. (1996). "Noninformative priors for the two sample normal problem," *Test*, **5**, 145-157.
- Jaheen, Z.F. (1996). "Empirical Bayes estimation of the reliability and failure rate functions of the Burr type X failure model," *Journal of Applied Statistical Science*, **3**, 281-288.
- Jeffreys, H. (1961). *Theory of Probability*, Oxford University Press, New York.
- Johnson, R. A. (1988). "Stress-strength models for reliability," *Handbook of Statistics, Volume 7: Quality Control and Reliability*, Eds. P. R. Krishnaiah and C. R. Rao, Elsevier Science Pub., New York, 27-54.
- Lee, G. (1998). "Development of matching priors for $P(X < Y)$ in exponential distributions," *Journal of the Korean Statistical Society*, **27**, 421-433.

- Lee, G., Sun, D. and Basu, A. P. (1997). "Noninformative priors for estimating $P(X < Y)$ in normal distributions," preprint.
- Mukerjee, R. and Dey, D. K. (1993). "Frequentist validity of posterior quantiles in the presence of a nuisance parameter : higher order asymptotics," *Biometrika*, **80**, 499-505.
- Mukerjee, R. and Ghosh, M. (1997). "Second order probability matching priors," *Biometrika*, **84**, 970-975.
- Peers, H. W. (1965). "On confidence sets and Bayesian probability points in the case of several parameters," *Journal of Royal Statistical Society*, B, **27**, 9-16.
- Sartawi, H.A. and Abu-Salih, M.S. (1991). "Bayesian prediction bounds for the Burr type X model," *Communications in Statistics-Theory and Methods*, **20**, 2370-2330.
- Stein, C. (1985). "On the coverage probability of confidence sets based on a prior distribution," *Sequential Methods in Statistics*, Banach Center Publications, **16**, 485-514.
- Sun, D., Ghosh, M. and Basu, A. P. (1998). "Bayesian analysis for a stress-strength system under noninformative priors," *Canadian Journal of Statistics*, **26**, 323-332.
- Thompson, R. D. and Basu, A. P. (1993). "Bayesian reliability of stress-strength system," *Advances in Reliability*, Ed. Basu, A., North-Holland, Amsterdam, 411-421.
- Tibshirani, R. (1989). "Noninformative priors for one parameter of many," *Biometrika*, **76**, 604-608.
- Welch, B. N. and Peers, B. (1963). "On formulae for confidence points based on integrals of weighted likelihood," *Journal of Royal Statistical Society*, **35**, 318-329.