BAYESIAN RELIABILITY ESTIMATION OF A TWO PARAMETER CAUCHY DISTRIBUTION

H.A. Howlader & G. Weiss

University of Winnipeg, Winnipeg, Manitoba

Key Words: Bayesian estimation; Cauchy distribution; Method of maximum likelihood; Mean squared-error; Reliability function; Monte Carlo simulation; Posterior mode; Prior/posterior distribution.

ABSTRACT

This paper gives an approximate Bayes procedure for the estimation of the reliability function of a two parameter Cauchy distribution using Jeffreys' non-informative prior under a squared-error loss function. A numerical example is given. Based on a Monte Carlo simulation, two such Bayes estimators of reliability are compared with the maximum likelihood estimator.

1. INTRODUCTION

The Cauchy distribution is a symmetric distribution with a bell-shaped density similar to the normal but with greater probability mass in the tails. The distribution is often used in extreme cases to model heavy-tail distributions, such as those which arise in outlier analyses. The Cauchy distribution arises naturally as the ratio of two independent normal variates. A standard Cauchy random variable has a Student's t-distribution with one degree of freedom, hence the general location-scale Cauchy (μ,σ) density is

$$f(x|\mu,\sigma) = \sigma[\pi\{\sigma^2 + (x-\mu)^2\}]^{-1}, \quad -\infty < x,\ \mu < \infty,\ \sigma > 0 \qquad (1)$$

where μ, the location parameter, is the population median, and σ, the scale parameter, is the inter-quartile range.

Estimation of the location and scale parameters by linear order statistics has been considered by Chan (1970), Cane (1974), Balmer, Boulton and Sack (1974) and Howlader and Weiss (1984). Maximum likelihood estimation in the two-parameter case was considered by Ferguson (1978) and Hinkley (1978).

Franck (1981) considers the problem of testing of normal versus the Cauchy and Spiegelhalter (1983) makes use of some of Franck's results to obtain exact Bayes estimator for μ and σ under a non-informative prior.

Howlader and Weiss (1984) apply Lindley's (1980) procedure to obtain approximate expressions for the Bayes estimators of the parameters of this distribution.

In this paper we consider the problem of estimation of the reliability function,

$$R_t = 1 - F(t) = \int_t^\infty f(x)dx = \frac{1}{2} - \frac{1}{\pi} \tan^{-1}\left(\frac{t-\mu}{\sigma}\right), \tag{2}$$

for various (fixed) values of t.

Although the reliability is usually modeled by positive distributions, since we generally observe only positive lifetimes, this is a minor consideration since for μ sufficiently large, the probability of negative observations is negligible. Indeed, Sinha (1983) has considered reliability estimation under a Normal lifetime model. The use of the Cauchy can be justified as an alternative to the normal, as in Franck (1981), especially when one is concerned with data contamination. Bain (1978) also cited the importance of the Cauchy distribution in life-testing whenever a density with heavier tails than the normal is sought.

2. ESTIMATION OF THE RELIABILITY FUNCTION

A. MLE: Let $x = x_1, x_2, \ldots, x_n$ be a random sample from (1), then the log likelihood function is

$$L = -n \log \pi + n \log \sigma - \sum_{i=1}^{n} \log\{\sigma^2 + (x_i - \mu)^2\}. \tag{3}$$

Setting $\frac{\partial L}{\partial \mu}$ and $\frac{\partial L}{\partial \sigma}$ respectively to 0 yield the likelihood equations as

$$\left.\begin{aligned} \sum_{i=1}^{n} \frac{x_i - \mu}{\sigma^2 + (x_i - \mu)^2} &= 0 \\ \sum_{i=1}^{n} \frac{\sigma^2}{\sigma^2 + (x_i - \mu)^2} &= \frac{n}{2}\,. \end{aligned}\right\} \tag{4}$$

We will use an iterative method to solve (4) for μ and σ. By the invariance property, the MLE of R_t is

$$\hat{R}_t = \frac{1}{2} - \frac{1}{\pi} \tan^{-1}\left(\frac{t - \hat{\mu}}{\hat{\sigma}}\right)\,. \tag{5}$$

B. BAYES ESTIMATOR OF R_t: In a situation where little is known about the parameters, Jeffreys (1983) proposes use of a prior density which is invariant under parametric transformations and which is termed a vague or non-informative prior.

In this study we consider the vague prior

$$p(\mu, \sigma) \propto \frac{1}{\sigma}\,.$$

Combining the likelihood function and the prior density, the joint posterior density of μ and σ is

$$\pi(\mu, \sigma | x) \propto \sigma^{n-1} \prod_{i=1}^{n} [\sigma^2 + (x_i - \mu)^2]^{-1}. \tag{6}$$

Hence, the Bayes estimator of R_t under a squared-error loss function,

$$R_t^* = E(R_t | x) = \frac{\int_0^\infty \int_{-\infty}^\infty [\frac{1}{2} - \frac{1}{\pi} \tan^{-1}(\frac{t-\mu}{\sigma})]\sigma^{n-1} \prod_{i=1}^{n} [\sigma^2 + (x_i - \mu)^2]^{-1} d(\mu, \sigma)}{\int_0^\infty \int_{-\infty}^\infty \sigma^{n-1} \prod_{i=1}^{n} [\sigma^2 + (x_i - \mu)^2]^{-1} d(\mu, \sigma)}\,. \tag{7}$$

The ratio of integrals in (7) does not seem to take a closed form. Lindley (1980) developed an asymptotic expansion for the evaluation of the ratio of integrals of the form

$$\int w(\theta)\exp\{L(\theta)\}d\theta/\int p(\theta)\exp\{L(\theta)\}d\theta \qquad (8)$$

where $\theta = (\theta_1,\theta_2,\ldots,\theta_m)$, $L(\theta)$ is the logarithm of the likelihood function, and $w(\theta)$ and $p(\theta)$ are arbitrary functions of θ. If $w(\theta) = u(\theta)p(\theta)$ and $p(\theta)$ is the prior density of θ, then (8) yields the posterior expectation of $u(\theta)$, i.e.,

$$E[u(\theta)|x] = \frac{\int u(\theta)\exp\{L(\theta) + \rho(\theta)\}d\theta}{\int \exp\{L(\theta) + \rho(\theta)\}d\theta}$$

where $\rho(\theta) = \ell n[p(\theta)]$.

Expanding $w(\theta)$ and $L(\theta)$ in (7) in a Taylor series expansion about MLE of θ, Lindley obtained the required expression for $E[u(\theta)|x]$. For details, see Lindley (1980). Denoting $R_t = u$, and using Lindley's method, Bayes estimator of R_t in (7) takes the form

$$R_t^* = u + \frac{1}{2}\sum(u_{ij}+2u_i\rho_j)\tau_{ij} + \frac{1}{2}\sum L_{ijk}u_\ell\tau_{ij}\tau_{k\ell}. \qquad (9)$$

All functions of the right-hand side of (9) are to be evaluated at MLE of θ. The summations are over all suffixes and from 1 to m, and each suffix denotes differentiation once w.r.t. the variable having that suffix, i.e.,

$$u_{ij} = \partial^2 u/\partial\theta_i\partial\theta_j, \quad L_{ijk} = \partial^3 L/\partial\theta_i\partial\theta_j\partial\theta_k, \quad \text{etc.}$$

The τ's are defined as the (i,j)-th element of the inverse of the matrix formed by the negative of the second derivatives of L, $\{L_{ij}\}$.

In our case $m = 2$, $\theta = (\mu,\sigma)$ and from (2),

$$u_1 = \frac{\partial u}{\partial\mu} = \frac{\sigma}{\pi}\frac{1}{\sigma^2+(t-\mu)^2}\ ; \quad u_2 = \frac{\partial u}{\partial\sigma} = \frac{1}{\pi}\frac{t-\mu}{\sigma^2+(t-\mu)^2}$$

$$u_{11} = \frac{\partial^2 u}{\partial \mu^2} = \frac{2\sigma}{\pi} \frac{t-\mu}{\{\sigma^2+(t-\mu)^2\}^2} \;; \quad u_{22} = \frac{\partial^2 u}{\partial \sigma^2} = -\frac{2\sigma}{\pi} \frac{t-\mu}{\{\sigma^2+(t-\mu)^2\}^2}$$

$$u_{12} = u_{21} = \frac{\partial^2 u}{\partial \mu \partial \sigma} = -\frac{1}{\pi} \frac{\sigma^2-(t-\mu)^2}{\{\sigma^2+(t-\mu)^2\}^2} .$$

From the prior density, we have

$$\rho_1 = \frac{\partial \rho}{\partial \mu} = 0, \quad \rho_2 = \frac{\partial \rho}{\partial \sigma} = -\frac{1}{\sigma} .$$

Hence, (9) takes the form:

$$\begin{aligned} R_t^* = u(\hat\theta) &+ \frac{1}{2} [u_{11}(\hat\theta)\tau_{11} + u_{12}(\hat\theta)\tau_{12} + \{u_{12}(\hat\theta) - 2u_1(\hat\theta)/\hat\sigma\}\tau_{12} \\ &+ \{u_{22}(\hat\theta) - 2u_2(\hat\theta)/\hat\sigma\}\tau_{22}] + \frac{1}{2} L_{111}\{u_1(\hat\theta)\tau_{11}^2 + u_2(\hat\theta)\tau_{22}\tau_{12}\} \\ &+ \frac{1}{2} L_{112}\{3u_1(\hat\theta)\tau_{11}\tau_{12} + u_2(\hat\theta)(\tau_{11}\tau_{22} + 2\tau_{12}^2)\} \\ &+ \frac{1}{2} L_{122}\{u_1(\hat\theta)(\tau_{11}\tau_{22} + 2\tau_{12}^2) + 3u_2(\hat\theta)\tau_{12}\tau_{22}\} \\ &+ \frac{1}{2} L_{222}\{u_1(\hat\theta)\tau_{12}\tau_{22} + u_2(\hat\theta)\tau_{22}^2\}. \end{aligned} \tag{10}$$

Although Lindley (1980) suggests that the L_{ijk} be calculated through finite differencing of the log-likelihood on a 5×5 grid about the joint MLE, this procedure requires great precision. In the case of the Cauchy, the partial derivates are analytic and easy to compute directly. Also, since the expectations of the mixed partial derivatives are zero [$EL_{12\cdots} = 0$], the parameters are locally orthogonal and (10) on further simplification takes the form:

$$\begin{aligned} R_t^* = u(\hat\theta) &+ \frac{1}{2} \{u_{11}(\hat\theta)\tau_{11} + (u_{22}(\hat\theta) - 2u_2(\hat\theta)/\hat\sigma)\tau_{22}\} \\ &+ \frac{1}{2} L_{111}u_1(\hat\theta)\tau_{11}^2 + \frac{1}{2} L_{222}u_2(\hat\theta)\tau_{22}^2 \end{aligned} \tag{11}$$

where τ_{11} and τ_{22} are now simply $-L_{11}^{-1}$ and $-L_{22}^{-1}$ respectively.

An alternative way of evaluating the ratio of integrals of the form (8) is by denoting the logarithm of the posterior density, except for the normalizing constant, $\Lambda(\theta)$, as

$$\Lambda(\theta) = L(\theta) + \rho(\theta)$$

and expanding $\Lambda(\theta)$ in a Taylor series expansion about posterior mode, $\tilde{\theta}$ of θ. It can easily be shown that the Bayes estimator of R_t in this development takes the form

$$\tilde{R}_t^* = u(\tilde{\theta}) + \frac{1}{2}\sum u_{ij}(\tilde{\theta})\psi_{ij} + \frac{1}{2}\sum \Lambda_{ijk}u_\ell(\tilde{\theta})\psi_{ij}\psi_{k\ell} \tag{12}$$

where ψ_{ij} = (i,j)-th element of the inverse of the matrix formed by the negative of the second derivatives of Λ, $\{\Lambda_{ij}\}$.

Under the orthogonality (12) takes the form

$$\tilde{R}_t^* = u(\tilde{\theta}) + \frac{1}{2}\{u_{11}(\tilde{\theta})\psi_{11} + u_{22}(\tilde{\theta})\psi_{22}\}$$
$$+ \frac{1}{2}\Lambda_{111}u_1(\tilde{\theta})\psi_{11}^2 + \frac{1}{2}\Lambda_{222}u_2(\tilde{\theta})\psi_{22}^2. \tag{13}$$

ψ_{11} and ψ_{22} are now simply $-\Lambda_{11}^{-1}$ and $-\Lambda_{22}^{-1}$ respectively.

Letting $s_i = [\sigma^2+(x_i-\mu)^2]^{-1}$, the kernel of the log-likelihood function (3) and the log-posterior density (6) takes the form

$$\ell = v \log \sigma + \sum_{i=1}^{n} \log s_i \tag{14}$$

$$\text{where } v = \begin{cases} n & \text{for the log-likelihood function} \\ n-1 & \text{for the log-posterior density function.} \end{cases}$$

Then (14) gives

$$\ell_{11} = \frac{\partial^2 L}{\partial \mu^2} = 2\sum_{i=1}^{n} s_i - 4\sigma^2 \sum_{i=1}^{n} s_i^2,\quad \ell_{12} = \frac{\partial^2 L}{\partial\mu\partial\sigma} = 4\sigma\sum_{i=1}^{n}(x_i-\mu)s_i^2 = \ell_{21}$$

$$\ell_{22} = \frac{\partial^2 L}{\partial \sigma^2} = -v/\sigma^2 - \ell_{11}$$

$$\ell_{111} = \frac{\partial^3 L}{\partial \mu^3} = 4\sum_{i=1}^{n}(x_i-\mu)s_i^2 - 16\sigma^2\sum_{i=1}^{n}(x_i-\mu)s_i^3$$

$$\ell_{222} = \frac{\partial^3 L}{\partial \sigma^3} = 2v/\sigma^3 + 12\sigma\sum_{i=1}^{n} s_i^2 - 16\sigma^3\sum_{i=1}^{n} s_i^3.$$

3. AN EXAMPLE

In order to illustrate Lindley's method, consider Darwin's data on the difference in heights of self- and cross-fertilized plants given in Box and Tiao (1973, p.153). The data consists of measurements on 15 pairs of plants. Each pair contained a self-fertilized and a cross-fertilized plant grown in the same pot and from same seed. The observations are the 15 differences:

-67 -48 6 8 14 16 23 24 28

29 41 49 56 60 75

The MLE of the parameters in (1) obtained by using a Newton-Raphson iteration are

$\hat{\mu} = 24.9705,\ \hat{\sigma} = 15.7059.$

Hence, the MLE of R_t at an arbitrary t = 8 is

$$\hat{R}_8 = R_8(\hat{\mu},\hat{\sigma}) = \frac{1}{2} - \frac{1}{\pi}\tan^{-1}\left(\frac{8-24.9705}{15.7059}\right) = .7623.$$

We also obtained the posterior mode, as the maximum of the posterior likelihood, by the same optimization which produced the MLE. This gives

$\tilde{\mu} = 24.6429,\ \tilde{\sigma} = 13.7103.$

We then apply the procedure of Lindley (1980) to correct for the bias when we evaluate R_t at the MLE, or at the posterior mode. The values of R_8 and the derivatives of the likelihood, evaluated at the MLE and posterior mode are given below.

Likelihood	Bayesian
$u(8) = R_8(\hat{\mu},\hat{\sigma}) = .7623$	$u(8) = R_8(\tilde{\mu},\tilde{\sigma}) = .7837$
$L_{11} = -.0297$	$\Lambda_{11} = -.1271$
$L_{22} = -.0311$	$\Lambda_{22} = -.7479$
$L_{111} = .0007$	$\Lambda_{111} = .0453$
$L_{222} = .0055$	$\Lambda_{222} = .3849$
$u_1(8) = .0094$	$u_1(8) = .0094$
$u_2(8) = -.0101$	$u_2(8) = -.0114$
$u_{11}(8) = -.0006$	$u_{11}(8) = -.0007$
$u_{22}(8) = .0006$	$u_{22}(8) = .0007$
$\hat{R}^*_8 = .7575$	$\tilde{R}^*_8 = .7877$

A comparison of $\hat{R}_t$, $\hat{R}_t^*$ and $\tilde{R}_t^*$ for t = 5 (5) 45 is given in Table 1.

TABLE 1: MLE and Bayes Estimates of R_t for Darwin's Data

t	$\hat{R}_t$	$\hat{R}_t^*$	$\tilde{R}_t^*$
5	.7879	.7827	.8112
10	.7424	.7380	.7691
15	.6800	.6779	.7097
20	.5976	.5998	.6277
25	.4994	.5077	.5247
30	.4014	.4140	.4156
35	.3191	.3330	.3220
40	.2570	.2700	.2528
45	.2117	.2231	.2041

4. MONTE CARLO STUDY

In order to compare the ML and Bayes estimators of the reliability function, we performed a Monte Carlo simulation study in which we generated 1000 samples of sizes n = 7, 15 and 30. The simulations were performed on a Bytec Hyperion (IBM-compatible) 16-bit micro-computer. Standard Cauchy variates were obtained as the ratio of standard normals which were generated by a modified Box-Mueller algorithm as given in Kennedy and Gentle (1980), which uses the pseudo-uniform random variates obtained from the micro-computer's built-in function.

The MLE of (5) was compared with two forms of approximate Bayes estimators given in (11) and (13);— the MLE corrected for bias, and the posterior mode estimator corrected for bias. The means and mean-squared errors for these estimators of R_t, t = 1 (1) 9 for 1000 samples with $\mu = 5$ and $\sigma = 1$ are given in Table 2. We observe the following:

1. The MLE, in general, does quite well; the Bayes procedures are slightly better at times near μ, where the variances of the sampling distributions are large.

2. The two Bayes procedures are hardly distinguishable.

Table II: Means and MSE's of estimators of the reliability for 1000 simulated Cauchy samples

			Time								
			t = 1	t = 2	t = 3	t = 4	t = 5	t = 6	t = 7	t = 8	t = 9
True Reliability			0.9220	0.8976	0.8524	0.7500	0.5000	0.2500	0.1476	0.1024	0.0780
n = 7	M.L.E.	Mean	0.9198	0.8952	0.8506	0.7497	0.5091	0.2622	0.1551	0.1074	0.0815
		MSE	0.0031	0.0049	0.0087	0.0183	0.0324	0.0209	0.0096	0.0051	0.0030
	Bayes (MLE)	Mean	0.9087	0.8818	0.8341	0.7327	0.5073	0.2785	0.1721	0.1212	0.0928
		MSE	0.0037	0.0056	0.0093	0.0171	0.0274	0.0201	0.0107	0.0060	0.0037
	Bayes (PME)	Mean	0.9130	0.8863	0.8380	0.7320	0.5087	0.2841	0.1719	0.1181	0.0089
		MSE	0.0038	0.0059	0.0101	0.0212	0.0366	0.0251	0.0128	0.0065	0.0039
n = 15	M.L.E.	Mean	0.9223	0.8978	0.8525	0.7498	0.4948	0.2460	0.1456	0.1011	0.0770
		MSE	0.0010	0.0017	0.0034	0.0076	0.0131	0.0074	0.0030	0.0015	0.0009
	Bayes (MLE)	Mean	0.9171	0.8913	0.8525	0.7405	0.4956	0.2558	0.1540	0.1076	0.0082
		MSE	0.0011	0.0019	0.0036	0.0073	0.0116	0.0071	0.0032	0.0017	0.0010
	Bayes (PME)	Mean	0.9186	0.8931	0.8461	0.7415	0.4957	0.2545	0.1520	0.1058	0.0806
		MSE	0.0012	0.0020	0.0039	0.0079	0.0118	0.0076	0.0035	0.0018	0.0011
n = 30	M.L.E.	Mean	0.9214	0.8968	0.8513	0.7484	0.4990	0.2501	0.1479	0.1028	0.0078
		MSE	0.0005	0.0008	0.0016	0.0039	0.0073	0.0036	0.0015	0.0008	0.0005
	Bayes (MLE)	Mean	0.9189	0.8936	0.8472	0.7435	0.4991	0.2552	0.1522	0.1060	0.0081
		MSE	0.0005	0.0009	0.0017	0.0039	0.0068	0.0036	0.0016	0.0009	0.0005
	Bayes (PME)	Mean	0.9192	0.8940	0.8475	0.7435	0.4990	0.2551	0.1517	0.1056	0.0081
		MSE	0.0054	0.0009	0.0017	0.0039	0.0068	0.0037	0.0016	0.0009	0.0005

ACKNOWLEDGEMENT

This research was supported by grants from the Natural Sciences and Engineering Research Council of Canada, Grant numbers A8739 and A8757. We are grateful to Don Grant for computer programming assistance.

REFERENCES

Bain, L.J. (1978), "Statistical Analysis of Reliability and Life-Testing Models - Theory and Methods", Marcel Dekker, Inc., New York.

Balmer, D.W., Boulton, M. and Sack, R.A. (1974), "Optimal solutions in parameter estimation problems for the Cauchy distribution", JASA 69: 238-242.

Box, G.E.D. and Tiao, G.S. (1978), "Bayesian Inference in Statistical Analysis", Addison-Wesley, Reading, Massachusetts.

Cane, G.J. (1974), "Linear estimation of parameters of the Cauchy distribution based on sample quantiles", JASA 69: 243-245.

Chan, L.K. (1970), "Linear estimation of location and scale parameters of the Cauchy distribution based on sample quantiles", JASA 65: 851-859.

Ferguson, T.S. (1978), "Maximum likelihood estimates of the parameters of the Cauchy distribution for samples of size 3 and 4", JASA 73: 211-213.

Franck, W.E. (1981), "The most powerful invariant test of normal vs. Cauchy with applications to stable alternatives", JASA 76: 1002-1005.

Hinkley, D.V. (1978), "Likelihood inference about location and scale parameters", Biometrika 65: 253-261.

Howlader, H.A. and Weiss, G. (1984), "A comparative study of the estimation of location and scale of a Cauchy distribution for small samples", Tech. Report #1, Department of Statistics, University of Winnipeg, Winnipeg, Manitoba.

Jeffreys, H. (1983), "Theory of Probability", 3rd ed., Clarendon Press, Oxford.

Kennedy, W.J., Jr. and Gentle, J.E. (1980), "Statistical Computing", Marcel Dekker, New York.

Lindley, D.V. (1980), "Approximate Bayesian Methods", (with discussants), Trabajos de Estadistica y de Investigacion Operativa, 31: 232-245.

Sinha, S.K. (1983), "Bayes estimation of the reliability function of normal life-time distribution", Tech. Report #122, Department of Statistics, University of Manitoba, Winnipeg, Manitoba.

Spiegelhalter, D.J. (1983), "Exact Bayesian inference on the parameters of a Cauchy distribution with vague prior information", presented at the Second International Bayesian Statistics Conference in Valencia, Spain.