THE EGYPTIAN STATISTICAL JOURNAL
ISSR, Cairo Univ., Vol.(37), No.2, 1993.

COMPARISON OF ESTIMATORS OF LOCATION MEASURES OF AN INVERSE RAYLEIGH DISTRIBUTION

M. K. GHARRAPH
Department of Mathematics
Zagazig University, Zagazig
EGYPT

This paper derives five measures of location of an inverse Rayleigh distribution which depend on an unknown paramter θ. The parameter θ is estimated by the method of maximum likelihood, and the method of moments. Some other estimators are also obtained. Some applications of the inverse Rayleigh distribution in Reliability studies are indicated. The results of a simulation study are presented for the comparison of biases and mean square errors of various estimators of θ.

I. INTRODUCTION

One of the main features of a distribution is its measure of location. For a given distribution there are several measures of location. In this paper the main interest is estimating the location measures of an inverse Rayleigh distribution which is usually used in reliability studies.

A random variable X has a gamma distribution if its probability density function (p.d.f) is of the form

$$(1) \qquad f(x) = \left(\beta^{\alpha}\Gamma(\alpha)\right)^{-1}(x-\gamma)^{\alpha-1} \exp\left(-(x-\gamma)/\beta\right), \qquad (\alpha,\beta > 0,\ x > \gamma)$$

This distribution is of type III in Pearson's system. It depends on three parameters α,β and γ. The standared form of the gamma distribution is obtained by putting $\beta=1$ and $\gamma=0$, giving

$$(2) \qquad f(x) = \left(\Gamma(\alpha)\right)^{-1} x^{\alpha-1} e^{-x}, \qquad (x > 0),$$

For $\alpha = 1$, it is an exponential distribution. If α is a positive integer, it is an Erlangian distribution [1] .

Suppose that X has the standard gamma distribution and we define a random variable Z by the transformation $[(Z-\gamma)/\beta]^c = X$ (with $c > 0$)), then the p.d.f. of Z is

(3) $$f(z) = c(z-\gamma)^{c\alpha-1} (\beta^{c\alpha}\Gamma(\alpha))^{-1} \exp\left\{-[(z-\gamma)/\beta]^c\right\} , \quad (z > \gamma)$$

This p.d.f was defined by Stacy [2] for $\gamma = 0$ as the family of generalized gamma distributions. It includes Weibull distribution (γ=1), half-normal distribution ($\alpha=\frac{1}{2}$, c=2, γ=0), and ordinary gamma distribution (c=1).

Assuming $\gamma = 0$, Stacy and Mihran [3], extended the definition of generalized gamma distribution to include negative (non-zero) value of c, by replacing the multiplier c in (3) by |c| . Then the p.d.f. of Z is

(4) $$f(z) = |c|\, z^{\alpha c-1}[\beta^{\alpha c}\Gamma(\alpha)]^{-1} \exp[-(z/\beta)^c] , \quad (z > 0;\ \alpha, \beta > 0, c \in R).$$

In particular, if $\alpha = 1$, $\beta^c = \theta$, and $c = -2$, then

(5) $$f(z) = \frac{2}{\theta z^3} \exp [-1/\theta z^2], \qquad (z > 0,\ \theta > 0).$$

which is the inverse Rayleigh p.d.f. with parameter θ. We denote it by X:IR(θ).

Generalized gamma distribution were discussed as early as 1925 [1] by Amoros who fitted such a distribution to an observed distribution of income rates.

The p.d.f. (5) can also be obtained in the following manner:

Let Y be a random variable with p.d.f.

(6) $$f(y) = \theta^{-1} \exp [-y/\theta] , \qquad (y > 0, \theta > 0)$$

It is the p.d.f of an exponential random variable with parameter θ and it is denoted by Y: Exp(θ) .

Making the transformation $X = Y^{-\frac{1}{2}}$, the the p.d.f. of X is

(7) $$f(x) = \frac{2}{\theta x^3} \exp(-1/\theta x^2), \quad (x > 0, \theta > 0),$$

which is the p.d.f of the inverse Rayliegh distribution

According to Johnson [1] Voda [5] quoted Treyer [4] as having stated that the distribution of lifetimes of several types of experimental units can be approximated by the inverse Rayleigh distribution

Now, If T is a random variable, then some of its measures of location are

(i) the mean of T, $\theta_1 = E(T)$,

(ii) the harmonic mean of T, $\theta_2 = [E(T^{-1})]^{-1}$,

(iii) the geometric mean of T, $\theta_3 = \exp[E(\ln T)]$,

(iv) the mode of T, the point θ_4 at which the p.d.f. of T has a local maximum , and

(v) the median of T, the point θ_5 at which

$$P(T \leq \theta_5) = P(T \geq \theta_5) = 0.5$$

In the next section, these measures of location are derived for the inverse Rayleigh distribution. In section III the maximum likelihood estimator (MLE) of θ is obtained. An estimator is obtained by the method of moments in section IV. Some adhoc estimators of θ are obtained in section V. Some applications of IR(θ) are presented in section VI. A simulation study is carried out to compare various estimators of θ, their biases and mean square errors for some selected values of θ. The results are presented in tables I and II .

I. MEASURES OF LOCATION OF THE INVERSE RAYLEIGH DISTRIBUTION

If X:IR(θ), then for any real nanber $r < 2$, it is easy to get

$$E(X^r) = \theta^{-\frac{r}{2}} \Gamma(1 - \frac{r}{2}).$$

In particular, if $r = -2$, then

$$E(X^{-2}) = E(Y) = \theta ,$$

where Y:Exp(θ). Hence, we can say that θ is the mean of X^{-2}. Moreover, if $r = -1$, then

$$E(X^{-1}) = \frac{1}{2} \sqrt{\theta \pi} ,$$

Hence, the harmonic mean of X is

$$\theta_2 = [E(X^{-1})]^{-1} = 2/\sqrt{\theta \pi}$$

The mean of X is

$$\theta_1 = E(X) = \sqrt{\pi/\theta}$$

but the variance of X does not exist .

Note that

$$E(\ln X) = \int_0^\infty \frac{2 \ln x}{\theta x^3} e^{-\frac{1}{\theta x^2}} dx$$

let $x = \theta^{-\frac{1}{2}} y^{-\frac{1}{2}}$

Therefore,

$$E(\ln X) = \int_0^\infty e^{-y} \ln(\theta^{-\frac{1}{2}} y^{-\frac{1}{2}}) dy$$

$$= - \frac{1}{2} \ln \theta - \frac{1}{2} \int_0^\infty e^{-y} \ln y \, dy$$

$$= \frac{1}{2} (\gamma - \ln \theta)$$

where $\gamma = - \int_0^\infty e^{-y} \ln y \, dy$ is the Euler constant which is approximately equal to

0.57722. Hence, the geometric mean of X is

$$\theta_3 = e^{\frac{\gamma}{2}} / \sqrt{\theta}$$

This result is needed to evaluate the Shannon entropy of X, H(X), where

$$H(X) = - \int_0^\infty f(x) \ln f(x)\, dx$$

$$= (1 + \ln \theta) + \frac{3}{2} (\gamma - \ln \theta) - \ln 2$$

$$= H(Y) + \frac{3}{2} (\gamma - \ln \theta) - \ln 2,$$

where H(Y) is the Shannon entropy of Y : Exp(θ).

This remark implies that the uncertainty in X is smaller than that of Y if θ is greater than $e^{\gamma} / (4)^{\frac{1}{3}}$ (i.e. if $\theta > 1.122$).

The distribution function of X is

$$P(X \leq x) = F(x;\theta) = \exp(-1/\theta x^2).$$

Put $F(x,\theta) = \frac{1}{2}$ to get the median

$$\theta_5 = 1/ \sqrt{\theta \ln 2}$$

Finally to find the mode, note that

$$\ln f(x) = \ln 2 - \ln \theta - 3 \ln x - 1/\theta x^2$$

$$\frac{\partial f(x)}{\partial x} = -3/x + 2/\theta x^3 ,$$

$$\frac{\partial^2 f(x)}{\partial x^2} = 3/x^2 - 6/\theta x^4$$

put $\frac{\partial f(x)}{\partial x} = 0$ to get $x = \sqrt{2/3\theta}$

and $\frac{\partial^2 f(x)}{\partial x^2} < 0$ for $x = \sqrt{2/3\theta}$

So, the mode $\theta_4 = \sqrt{2/3\theta}$

It is interesting to note that the distribution of inverse Rayleigh is skewed to the right since

$$\theta_4 < \theta_5 < \theta_1$$

Moreover, since Var(X) does not exist, one may use quartile coeficient of skewness

$$\gamma_3 = (Q_3 - 2Q_2 + Q_1)/(Q_3 - Q_1),$$

where $Q_j = 1/\sqrt{\theta \ln\frac{4}{j}}$ is the jth quartile of X (j=1,2,3), or the 10-90 percentile coefficient of skewness

$$\gamma_4 = (P_{90} - 2P_{50} + P_{10})/(P_{90} - P_{10})$$

where P_j is the jth percentile of X (j=1,...,99). However, since γ_3 and γ_4 are free of θ, both γ_3 and γ_4 are not reasonable measures of skewness. Alternatively, it is possible to use the quartile deviation,

$$\frac{Q_3 - Q_1}{2} = \frac{1}{2\sqrt{\theta}} \left(1/\sqrt{\ln \frac{4}{3}} - 1/\sqrt{\ln 4}\right)$$

whioh is approximately equal to $1/2\sqrt{\theta}$

The mean deviation about the mean, as a measure of dispersion, is given by

$$\delta = E|X-\sqrt{\pi/\theta}| = \int_0^\infty |x-\sqrt{\pi/\theta}|\, f(x)\, dx$$

$$= \int_0^{\sqrt{\pi/\theta}} (\sqrt{\pi/\theta} - x)\, f(x)\, dx + \int_{\sqrt{\pi/\theta}}^\infty (x - \sqrt{\pi/\theta})\, f(x)\, dx$$

$$= \sqrt{\pi/\theta}\; F(\sqrt{\pi/\theta}\,;\theta) - \int_0^{\sqrt{\pi/\theta}} x\, f(x)dx + \int_{\sqrt{\pi/\theta}}^\infty x\, f(x)\, dx$$

$$- \sqrt{\pi/\theta}\,[1-F(\sqrt{\pi/\theta};\theta)]$$

$$= 2\sqrt{\pi/\theta}\,(e^{-1/\pi} - 0.42494)$$

$$= 1/\sqrt{0.87\theta}$$

The integrals in the above expression have been computed by using CASIO FX-850p calculator library.

I. THE MAXIMUM LIKELIHOOD ESTIMATOR

Let $X_1,\ldots,X_n$ be a random sample from $IR(\theta)$, then the likelihood function is

$$L(\underset{\sim}{x};\theta) = (2/\theta)^n \exp(-\sum_{j=1}^n x_j^{-2}/\theta) \Big/ \prod_{j=1}^n x_j^3$$

$$= (1/\theta)^n \exp(-\sum_{j=1}^n x_j^{-2}/\theta).\, (2^n/\prod_{j=1}^n x_j^3) = g(t,\theta)\, h(\underset{\sim}{x}),$$

where $t = \sum_{j=1}^n x_j^{-2}$,

which shows that t is a sufficient statistic for θ. Therefore, the search for the minimum variance unbiased estimator of θ can be confined to functions of t, since $E(X^{-2}) = E(Y) = \theta_1$, $\frac{1}{n} \Sigma x_j^{-2}$ is the minimum variance unbiased estimator of θ.

To obtain the maximum likelihood estimator, we find

$$L^*(\theta) = \ln L(\underline{x},\theta) = c - n \ln \theta - \sum_{j=1}^{n} x_j^{-2} / \theta$$

where c is free of θ. Note that

$$\frac{\partial L^*(\theta)}{\partial \theta} = - n/\theta + \sum_{j=1}^{n} x_j^{-2}/\theta^2$$

and

$$\frac{\partial^2 L^*(\theta)}{\partial \theta^2} = n/\theta^2 - 2 \sum_{j=1}^{n} x_j^{-2}/\theta^3$$

So, the maximum likelihood estimator (MLE) of θ, obtained by solving $\frac{\partial L^*(\theta)}{\partial \theta} = 0$, is

$$\hat{\theta} = \frac{1}{n} \sum_{n=1}^{n} x_j^{-2}$$

Note that $\frac{\partial^2 L^*(\theta)}{\partial \theta^2} < 0$ at $\theta = \hat{\theta}$

Since X:IR(θ) and $Y = X^{-2}$: Exp(θ), then $u = 2y/\theta : \chi^2_{(2)}$ and hence

$W = 2 \sum_{j=1}^{n} x_j^{-2}/\theta = 2n\,\hat{\theta}/\theta : \chi^2_{(2n)}$. Thus for any real number r

$$E(\hat{\theta}^r) = \left(\frac{\theta}{2n}\right)^r E(W^r) = \frac{\Gamma(n+r)}{n^r \Gamma(n)} \theta^r$$

So, $E(\hat{\theta}) = \theta$ and $Var(\hat{\theta}) = \frac{\theta^2}{n}$, i.e. the MLE of θ is an unbiased and consistent estimator of θ.

Using the stirling's formula, it can be shown that

$$\frac{\Gamma(n+r)}{n^r \Gamma(n)} \to 1 \qquad \text{as} \qquad n \to \infty .$$

Now, since each θ_j, $j = 1,2,3,4$ and 5 takes the form

$\theta_j = C_j/\sqrt{\theta}$ for some constents C_j, then

$$\hat{\theta}_j = C_j\, \hat{\theta}^{-\frac{1}{2}},$$

$$E(\hat{\theta}_j) = C_j\, E(\hat{\theta}^{-\frac{1}{2}})$$

$$= C_j \frac{\Gamma(n-\frac{1}{2})}{n^{-\frac{1}{2}}\Gamma(n)}\, \theta^{-\frac{1}{2}} \rightarrow \quad C_j\ \theta^{-\frac{1}{2}} = \theta_j \text{ as } n \rightarrow \infty$$

and

$$Var(\hat{\theta}_j) = C_j^2\ Var(\hat{\theta}^{-\frac{1}{2}})$$

$$= C_j^2\ \left\{E(\hat{\theta}^{-1}) - ((\hat{\theta}^{-\frac{1}{2}}))^2\right\}$$

$$= C_j^2\ \ \theta^{-1}\ \left\{\frac{\Gamma(n-1)}{n^{-1}\Gamma(n)} - \left\{\frac{\Gamma(n-\frac{1}{2})}{n^{-\frac{1}{2}}\ \Gamma n}\right\}^2\right\} \rightarrow\ 0 \quad \text{as} \quad n \rightarrow \infty$$

$j = 1,\ldots,5$ $\hat{\theta}_j$ is asymptotically unbiased and consistent for θ_j .

IV ESTIMATION OF θ BY THE METHOD OF MOMENTS

The method of moments consists of equating a suitable number of sample moments to their population counterparts.

In order to estimate θ, replace $E(X)$ by $\bar{x}$ to get

$$\bar{x} = \tilde{\theta}_1 = \sqrt{\pi/\theta_E}$$

Hence

$$\theta_E = \pi/\bar{x}^2$$

V SOME ADHOC ESTIMATORS OF θ

Next, we can obtain some other reasonable estimaotrs of θ in the following manner:-

Equating the sample harmonic mean to its population conterpart θ_2 obtained in section II, we have

$$\tilde{\theta}_2 = n / \sum_{j=1}^{n} x_j^{-1} = 2/\sqrt{\pi\, \theta_H}$$

and hence the estimator of θ through the harmonic mean is

$$\theta_H = 4\,\left(\sum_{j=1}^{n} x_j^{-1} \right)/\pi\, n^2$$

Similarly, equating the sample geometric mean to the population geometric mean, we have

$$\tilde{\theta}_3 = (x_1, x_2, \ldots, x_n)^{\frac{1}{n}} = (\theta_G\, e^{-\gamma})^{-\frac{1}{2}}$$

and hence the estimator of θ through the geometric mean is

$$\theta_G = e^{\gamma} \prod_{j=1}^{n} x_j^{2} / n$$

The sample mode is

$$\tilde{\theta}_4 = \text{mode}\,(x_1, x_2, \ldots, x_n)$$

Equating it to the population mode, we obtain

$$\theta_{Mo} = \frac{2}{3}\, \tilde{\theta}_4^{-2}$$

Finally, let

$$\tilde{\theta}_5 = \text{median } (x_1, x_2, \ldots, x_n)$$

On equating it to the population median, we have

$$\theta_{Me} = (\tilde{\theta}_5^2 \ln 2)^{-1}$$

The properties of the estimators $\theta_H, \theta_G, \theta_{Mo}$ and θ_{Me} are not easy to derive analyticaly, so a simulation study is carried out to compare their biases and mean square errors along with those of $\hat{\theta}$, the maximum likelihood estimator of θ (which is also the minimum variance unbiased estimator of θ).

VI SOME APPLICATIONS

1. Consider a system of n components plased in parallel. Assume the life time of each component X is $IR(\theta)$. Then the life time of the system is $X_{(n)} = \max_{1 \le i \le n} X_i$. It can be shown that

   $X_{(n)}$ is $IR(\frac{\theta}{n})$. Therefore the expected life time of the system is $E(X_{(n)}) = \sqrt{\frac{n\pi}{\theta}}$

2. If we wish to have a similar new system with m components, we can obtain a prediction interval for the future life time of the new system in the following manner:

Note that the maximum life time for the new system $Y_{(m)}$ is also $IR(\frac{\theta}{m})$, and therefore

$$X_{(n)}^{-2} \text{ is } Exp(\tfrac{\theta}{n}) \text{ and } \frac{2n}{\theta} X_{(n)}^{-2} \text{ is } \chi^2_{(n)}$$

Also $\frac{2m}{\theta} Y_{(m)}^{-2}$ is $\chi^2_{(2)}$. Since $X_{(n)}$ and $Y_{(m)}$ are independt, then $\frac{nY_{(m)}}{m X_{(n)}^2}$ has an $F_{(2,2)}$-distribution, and $100(\alpha-1)\%$ prediction interval for $Y_{(m)}$ given $X_{(n)}$ is

$$\left\{ X_{(n)} \sqrt{\frac{m}{n} F_{(\frac{\alpha}{2};2,2)}} \quad , \quad X_{(n)} \sqrt{\frac{m}{n} F_{(\frac{\alpha}{2};2,2)}} \right\}$$

If we have a system of n components in series, then the life time of the system is $X_{(1)} = \min_{1 \leq i \leq n} X_i$. The expected life time of such a system is

$$E\left(X_{(1)}\right) = \sum_{j=0}^{n-1} (-1)^j \binom{n}{j+1} E\left(X \mid X \text{ is } IR\left(\frac{\theta}{j+1}\right)\right)$$

$$= \sum_{j=0}^{n-1} (-1)^j \binom{n}{j+1} \sqrt{\frac{\pi (j+1)}{\theta}}$$

REFERENCES

[1] Johnson, N. L. and Kotz, S (1970). Continuous Univariate distributions. John Wiley and Sous, 166

[2] Stacy, E. W. (1962). A generalization of the gamma distribution, A nnals of Mathematical Statsitics, 33, 1187-1192.

[3] Stacy, E. W. and Mihram, G.A. (1965). Parameter estimation for the generalized gamma distribution, Technometrics, 7, 349-358.

[4] Treyer, V.N. (1964). In Doklady Acad. Nauk. Beloras, SSR (Russian).

[5] Voda, V. Gh. (1972), On the "Inverse Rayliegh variable", Rep. Stat. Appl. Res, Juse, 19,(4).

TABLE (1)

Sample Size (n)	Estimator	θ = 5 Estimate	Biase	MSE	θ = 10 Estimate	Biase	MSE
20	$\hat{\theta}$	4.8173	0.1827	1.1250	10.4228	-0.4228	6.4365
	θ_E	5.8876	-0 .8876	6.9402	13.6716	-3.6716	40.4339
	θ_H	4.8967	0.1033	1.3417	10.6814	-0.6814	6.3677
	θ_G	5.0528	-0.0528	2.0588	11.1914	-1.1914	10.0208
	θ_{Mo}	3.1013	1.8987	6.5678	7.3698	2.6302	19.1507
	θ_{Me}	4.9159	0.0841	2.4131	11.0980	-1.0980	12.5501
30	$\hat{\theta}$	4.9975	0.0025	0.8424	9.9575	0.0425	3.0531
	θ_E	5.8497	-0 .8497	5.0214	11.1791	-1.1791	18.0182
	θ_H	5.0724	-0..0724	0.9403	10.0274	-0.0274	3.2033
	θ_G	5.1796	-0.1796	1.4161	10.1169	-0.1169	4.5219
	θ_{Mo}	3.2289	1.7711	5.9794	6.0438	3.9562	28.2235
	θ_{Me}	5.2241	-0.2241	1.5438	10.2703	-0.2703	6.9263
40	$\hat{\theta}$	4.9483	0.0517	0.6505	10.0119	0.0119	3.0208
	θ_E	5.7240	- 0.7240	4.0933	11.9509	1.9509	21.6833
	θ_H	4.9910	0 ..0090	0.6896	10.1407	-0.1407	3.2045
	θ_G	5.0895	-0.0895	1.0351	10.4029	-0.4029	5.0597
	θ_{Mo}	3.5228	1.4772	6.4249	7.4902	2.5098	21.7951
	θ_{Me}	5.1291	-0.1291	1.0583	10.1768	-0.1768	5.1301

Table (1) sumarizes the results of a simulation study for 100 sample of sizes 20,30 and 40 from IR(θ), θ = 5, and 10. The estimators $\hat{\theta}, \theta_E, \theta_H, \theta_G, \theta_{Mo}, \theta_{Me}$ are computed alongwith the biases in these estimators as well as their mean square errors.

It is clear that the maximum likelihood estimator performs relatively better than some of the other estimators.

TABLE (2)

Sample Size n	Measure	θ = 5 exact	θ = 5 Estimate	θ = 5 Biase	θ = 5 MSE	θ = 10 exact	θ = 10 Estimate	θ = 10 Biase	θ = 10 MSE
20	mean	0.7927	0.7639	0.0287	0.0868	0.5605	0.5803	-0.0198	0.0502
	H.mean	0.5046	0.5103	-0.0056	0.0043	0.3567	0.3681	-0.0113	0.0022
	G.mean	0.5968	0.5956	0.0012	0.0071	0.4220	0.4371	-0.0150	0.0049
	mode	0.3651	0.5172	-0.1521	2.6755	0.2592	0.3961	-0.1379	0.1059
	median	0.5372	0.5560	-0.0188	0.0085	0.3798	0.3916	-0.0118	0.0044
30	mean	0.7927	0.7922	0.0005	0.0336	0.5605	0.5414	0.0191	0.0146
	H.mean	0.5046	0.5080	-0.0034	0.0021	0.3568	0.3596	-0.0029	0.0012
	G.mean	0.5968	0.6030	-0.0061	0.0043	0.4220	0.4224	-0.0003	0.0023
	mode	0.3651	0.5297	-0.1646	0.1195	0.2582	0.3546	-0.0964	0.0427
	median	0.5372	0.5473	-0.0101	0.0048	0.3798	0.3947	-0.0049	0.0024
40	mean	0.7927	0.7821	0.0106	0.0189	0.5605	0.5782	-0.0177	0.0339
	H.mean	0.5046	0.5058	-0.0012	0.0015	0.3568	0.3625	-0.0057	0.0012
	G.mean	0.5968	0.5971	-0.0002	0.0029	0.4220	0.4300	-0.0080	0.0020
	mode	0.3651	0.5347	-0.1696	0.0361	0.2582	0.4101	-0.1519	0.2015
	median	0.5372	0.5380	-0.0009	0.0032	0.3798	0.3859	-0.0061	0.0022

Table (2) extends the results of the simulation study .