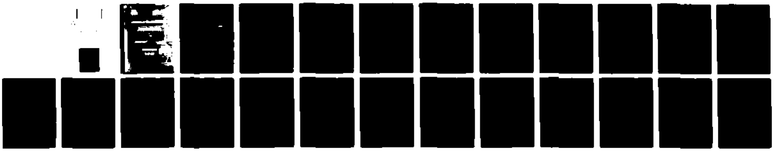
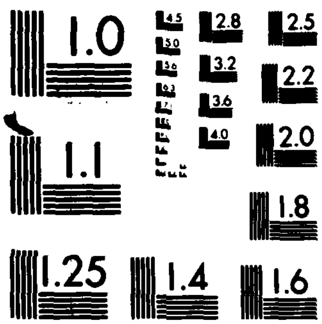
AD-A099 427 MISSOURI UNIV-COLUMBIA DEPT OF STATISTICS F/G 12/1
THE ESTIMATION OF P(X < Y) FOR DISTRIBUTIONS USEFUL IN LIFE TES--ETC(U)
SEP 80 A P BASU N00014-78-C-0655
UNCLASSIFIED TR-98 NL



END
DATE
FILMED
8 81
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER 98	2. GOVT ACCESSION NO. AD-A099 427	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) 6 The Estimation of P(X < Y) for Distributions Useful in Life Testing.		5. TYPE OF REPORT & PERIOD COVERED 9 Technical Rep[redacted]t.
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) 10 Asit [redacted] P./Basu		8. CONTRACT OR GRANT NUMBER(s) 15 N00014-78-C-0655
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics University of Missouri-Columbia Columbia MO 65211	12 25	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 14 TR-98
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Department of the Navy Arlington, VA 22217	11	12. REPORT DATE Sep[redacted]80 13. NUMBER OF PAGES 22
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Reliability, stress-strength model, confidence interval, maximum likelihood estimators, gamma, exponential, bivariate exponential, simulation.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In this paper the reliability function R = P(X < Y) has been estimated when X and Y follow gamma, exponential or bivariate exponential distributions. The paper is partly expository.

DD FORM 1473, 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

402 660

The Estimation of $P(X < Y)$ for Distributions

Useful in Life Testing*

by

A. P. Basu

University of Missouri, Columbia

In this paper the reliability function $R = P(X < Y)$ has been estimated when X and Y follow gamma, exponential or bivariate exponential distributions. The paper is partly expository.

Key words: reliability, stress-strength model, confidence interval, maximum likelihood estimators, gamma, exponential, bivariate exponential, simulation.

*Research supported in part under contract N00014-78-C-0655 for the Office of Naval Research.

1. Introduction.

Let X and Y be two random variables with cumulative distribution functions $F(x)$ and $G(y)$ respectively. Suppose Y is the strength of a component subject to a stress X . Then the component fails if at any moment the applied stress (or load) is greater than its strength. The stress is a function of the environment to which the component is subjected. Strength depends on material properties, manufacturing procedures, and so on. The reliability of a component is the probability that its strength exceeds the stress. From practical considerations it is desirable to draw inference about the reliability function.

The above model was first considered by Birnbaum (1956) and has since found an increasing number of applications in many different areas, especially in the structural and aircraft industries. For a bibliography of available results see Basu (1977).

In many situations, the distribution of X (or of both X and Y) will be completely known except possibly for a few unknown parameters and it is desired to obtain parametric solutions. Thus, in case of missile flights, the stress may be expensive to sample, but the physical characteristics of the missile system, such as the propulsive force, angle of elevation, changes in atmospheric condition, and so on, may all have known distributions; consequently, the distribution of stresses can be calculated. Church and Harris (1970), Owen, Craswell and Hanson (1964), and Govindarajulu (1968) have considered the above problem under the assumption that X and Y have normal distributions. Since in many physical situations, especially in reliability, exponential and other distributions provide more realistic models, it is desirable to obtain estimators of R for distributions useful in life testing. In section 2 we consider gamma and exponential distributions under the assumption that X and Y are independently distributed. The case of a bivariate exponential distribution is studied in section 3. Two distribution free procedures are mentioned in section 4. The effect of misspecifying the model is considered in section 5 and a numerical example is given in section 6.

2. Gamma and Exponential Distribution.

Let X and Y be independently distributed with density functions

$$f(x) = \frac{\alpha^p}{\Gamma(p)} e^{-\alpha x} x^{p-1} , \quad x > 0 , \quad p > 0 \tag{2.1}$$

$$g(y) = \frac{\beta^q}{\Gamma(q)} e^{-\beta y} y^{q-1} , \quad y > 0 , \quad q > 0 \tag{2.2}$$

respectively. Then

$$R_{pq} = P(X < Y) = \int_0^\infty \left[\int_x^\infty \frac{\beta^q}{\Gamma(q)} e^{-\beta y} y^{q-1} dy \right] \frac{\alpha^p}{\Gamma(p)} e^{-\alpha x} x^{p-1} dx$$

$$= \sum_{k=0}^{q-1} \frac{\Gamma(p+k)}{\Gamma(p)\Gamma(k+1)} \frac{\alpha^p \beta^k}{(\alpha+\beta)^{p+k}} . \tag{2.3}$$

Here p and q are assumed to be known integers.

Note $R_{11} = \alpha/(\alpha + \beta)$, and $R_{p1} = R_{11}^p$ for all p . Also,

$$R_{1q} = \sum_{k=0}^{q-1} (\beta/\alpha)^k R_{11}^{k+1} = \frac{\alpha}{\beta} \sum_{k=0}^{q-1} (1 - R_{11})^{k+1} .$$

In particular,

$$R_{21} = R_{11}^2 \ , \quad R_{31} = R_{11}^3 \ ,$$

$$R_{12} = R_{11} + (\beta/\alpha) R_{11}^2$$

$$R_{13} = R_{11} + (\beta/\alpha) R_{11}^2 + (\beta/\alpha) R_{11}^3 \ ,$$

$$R_{22} = R_{11}^2 (3 - 2R_{11}) \ .$$

If R_{11} is close enough to one, as is expected for items with high reliability,

$$R_{1q} \simeq \frac{\alpha}{\beta} (1 - R_{11}) \ .$$

Expressions for R_{p1} and R_{1q} indicate that in this case the expression for the reliability is not strongly dependent on the choice of the parameter p (especially if p is small) and the distribution of X can be approximated by the exponential distribution without much distortion in the value of R_{p1}. However, so far as the parameter q is concerned, the situation is quite reversed. The value of reliability is heavily dependent on the choice of the underlying distribution of Y and one has to choose the value of q more carefully. Later in section 5, we shall further study the effect of misspecifying the parameters p and q and confirm our conclusions by numerical studies.

If two independent random samples $(X_1, X_2, \ldots, X_m)$ and $(Y_1, Y_2, \ldots, Y_n)$ from the two gamma populations are available maximum likelihood estimators (MLE) of α and β are given by $\hat{\alpha} = \frac{p}{\bar{X}}$ and $\hat{\beta} = \frac{q}{\bar{Y}}$. Hence MLE of R_{pq} is

$$\hat{R}_{pq} = \sum_{k=0}^{q-1} \frac{\Gamma(p+k)}{\Gamma(p)\Gamma(k+1)} \frac{\hat{\alpha}^p \hat{\beta}^k}{(\hat{\alpha}+\hat{\beta})^{p+k}} . \tag{2.4}$$

As special cases, if $q = 1$, that is it X follows the gamma distribution and Y follows the exponential distribution

$$\hat{R}_{p1} = \left(\frac{\hat{\alpha}}{\hat{\alpha}+\hat{\beta}}\right)^p . \tag{2.5}$$

Finally, if both p and q are equal to 1, we have the case of two independent exponential distributions and we have

$$\hat{R}_{11} = \frac{\hat{\alpha}}{\hat{\alpha}+\hat{\beta}} = \frac{\bar{Y}}{\bar{X}+\bar{Y}} . \tag{2.6}$$

Since $\hat{R}$ in (2.6) is a function of $\frac{\bar{X}}{\bar{Y}}$, the exact distribution $\hat{R}$ can be obtained in the exponential case. It is well known that $\frac{\bar{X}}{\bar{Y}}\frac{\alpha}{\beta}$ follows the F distribution with $(2m,2n)$ degrees of freedom. Thus the distribution of $\hat{R}_{11}$ follows. The result will be used later in section 5 to compare the performance of independent exponentials with those of dependent exponential models. Using a theorem in Rao (1965, Thm. 6a.2, page 321), the distribution of $\hat{R}_{pq}$ in each of the above cases, for large m and n, can be shown to be normal and hence an estimate of the asymptotic confidence interval for R can be obtained. Thus, in this case with $m = n$ we have for large n, $\sqrt{n}(\hat{R}_{pq} - R_{pq}) \sim N(0,\sigma^2_{pq})$. Expressions for R_{pq}, $\hat{R}_{pq}$ and σ^2_{pq}, for a few selected values of p and q are given below.

$$R_{11} = \frac{\alpha}{\alpha+\beta}, \quad \hat{R}_{11} = \frac{\bar{Y}}{\bar{X}+\bar{Y}}, \quad \sigma^2_{11} = \frac{2\alpha^2\beta^2}{(\alpha+\beta)^4}$$

$$R_{21} = \left(\frac{\alpha}{\alpha+\beta}\right)^2, \quad \hat{R}_{21} = \frac{4\bar{Y}^2}{(\bar{X}+2\bar{Y})^2}, \quad \sigma^2_{21} = \frac{6\alpha^4\beta^2}{(\alpha+\beta)^6}$$

$$R_{12} = \frac{\alpha}{\alpha+\beta} + \frac{\alpha\beta}{(\alpha+\beta)^2}, \quad \hat{R}_{12} = \frac{\bar{Y}}{(\bar{Y}+2\bar{X})} + \frac{2\bar{X}\bar{Y}}{(\bar{Y}+2\bar{X})^2}, \quad \sigma^2_{12} = \frac{6\alpha^2\beta^4}{(\alpha+\beta)^6}.$$

$$R_{22} = \frac{\alpha^2}{(\alpha+\beta)^2} + \frac{2\alpha^2\beta}{(\alpha+\beta)^3}, \quad \hat{R}_{22} = \frac{\bar{Y}^2}{(\bar{X}+\bar{Y})^2} + \frac{2\bar{X}\bar{Y}^2}{(\bar{X}+\bar{Y})^3},$$

$$\sigma^2_{22} = 36(\alpha\beta)^4/(\alpha+\beta)^4.$$

3. Bivariate Exponential Distribution.

Since the exponential distribution is considered a useful model in life testing problems, it is desirable to consider bivariate analogues of univariate exponential distributions which will have properties similar to those of the univariate exponential distribution. Marshall and Olkin (1967) have proposed a very important bivariate exponential distribution (BVE), which is given by

$$(3.1)\qquad \begin{aligned} &\bar{F}(x,y) = P(X>x, Y>y) = e^{-\lambda_1 x - \lambda_2 y - \lambda_{12}\max(x,y)}, \\ &0 \le \lambda_1, \lambda_2, \lambda_{12} < \infty,\ \lambda_1+\lambda_{12} > 0,\ \lambda_2+\lambda_{12} > 0\ (x>0, y>0) . \end{aligned}$$

The BVE arises in several natural ways and is considered a useful model in reliability with appealing properties.

Let (X_i, Y_i), $i = 1,2,\ldots,n$ be a random sample from (3.1). We shall estimate $P(X<Y)$ when (X,Y) follows the BVE. It can be readily obtained from (3.1) that

$$(3.2)\qquad R = P(X<Y) = \frac{\lambda_1}{\lambda_1+\lambda_2+\lambda_{12}} .$$

Hence R is estimated by

$$(3.3)\qquad \hat{R} = \frac{\hat{\lambda}_1}{\hat{\lambda}_1+\hat{\lambda}_2+\hat{\lambda}_{12}}$$

where $\hat{\lambda}_1$, $\hat{\lambda}_2$, and $\hat{\lambda}_{12}$ are the maximum likelihood estimators of λ_1, λ_2, and λ_{12} respectively. Various authors have considered maximum likelihood estimation of the parameters $\lambda_1, \lambda_2, \lambda_{12}$. However, no explicit solutions of these parameters are available. In order to obtain an explicit form for $\hat{R}$, we replace the mle by some special ad hoc estimators called the "INT" estimators of Prochan and Sullo (1976) which have very high asymptotic relative efficiency compared with the mle estimators. Let n_1 = number of pairs such that $X_i < Y_i$, n_2 = number of pairs with $X_i > Y_i$ and n_0 = number of pairs with $X_i = Y_i$.

The "INT" estimators are given by

$$\hat{\hat{\lambda}}_1 = \left[\frac{n_1}{n_1 + n_0}\right] \frac{n}{\sum_{i=1}^{n} X_i}, \quad \hat{\hat{\lambda}}_2 = \left[\frac{n_2}{n_2 + n_0}\right] \frac{n}{\sum_{i=1}^{n} Y_i} \tag{3.4}$$

and

$$\hat{\hat{\lambda}}_{12} = n_0\left[1 + \frac{n_2}{n_1 + n_0} + \frac{n_1}{n_2 + n_0}\right] \Bigg/ \sum_{i=1}^{n} \max(X_i, Y_i) .$$

Proschan and Sullo also prove the following theorem.

<u>Theorem</u> (Proschan and Sullo). $n^{\frac{1}{2}}(\hat{\hat{\underline{\lambda}}}_n - \underline{\lambda})$ is asympototically trivariate normal with mean 0 and dispersion matrix $\sum = (\sigma_{ij})$ where, using the suffix n to denote dependence on the sample size, $\hat{\hat{\underline{\lambda}}}_n = (\hat{\hat{\lambda}}_1, \hat{\hat{\lambda}}_2, \hat{\hat{\lambda}}_{12})'$, $\underline{\lambda} = (\lambda_1, \lambda_2, \lambda_{12})'$,

$$
(3.5)\qquad
\begin{aligned}
\sigma_{11} &= \lambda_1(\lambda-\lambda_1\lambda_2\gamma_1^{-1}) ,\\
\sigma_{12} &= \lambda_1^2\lambda_2^2\lambda_{12}(\lambda\gamma_1\gamma_2)^{-1} ,\\
\sigma_{13} &= -\lambda_1\lambda_2\lambda_{12}(\lambda_1\lambda_2\gamma_2^{-2}+\lambda\gamma_1^{-1})/(\theta_0\gamma_1) ,\\
\sigma_{22} &= \lambda_2(\lambda-\lambda_1\lambda_2\gamma_2^{-1}) ,\\
\sigma_{23} &= -\lambda_1\lambda_2\lambda_{12}(\lambda_1\lambda_2\gamma_1^{-2}+\lambda\gamma_2^{-1})/(\theta_0\gamma_2) ,\\
\sigma_{33} &= \lambda_{12}[\lambda-\lambda_0\theta_0^{-2}(\lambda_1\gamma_2^{-3}+\lambda_2\gamma_1^{-3}+2\lambda_1\lambda_2\gamma_1^{-2}\gamma_2^{-2})] ,
\end{aligned}
$$

and
$$
\begin{aligned}
\lambda &= \lambda_1+\lambda_2+\lambda_{12} , \quad \gamma_1 = \lambda_1+\lambda_{12} ,\\
\gamma_2 &= \lambda_2+\lambda_{12}\\
\theta_0 &= E[\max(X,Y)] = \gamma_1^{-1}+\gamma_2^{-1}-\lambda^{-1} .
\end{aligned}
$$

R is then estimated by $\hat{\hat{R}} = \hat{\hat{\lambda}}_1/(\hat{\hat{\lambda}}_1+\hat{\hat{\lambda}}_2+\hat{\hat{\lambda}}_{12})$ rather than by $\hat{R}$. Since $\hat{\hat{R}}$ is a totally differential function of $\hat{\hat{\lambda}}_1$, $\hat{\hat{\lambda}}_2$ and $\hat{\hat{\lambda}}_{12}$ by Theorem 6a.2(ii) of Rao (page 321), $\sqrt{n}(\hat{\hat{R}}-R)$ is asympototically normally distributed with mean zero and variance σ^2, where

$$
(3.6)\qquad \sigma^2 = \frac{\gamma_2^2}{\lambda^4}\sigma_{11}+\frac{\lambda_1^2}{\lambda^4}(\sigma_{22}+\sigma_{33}) - 2\,\frac{\lambda_1\gamma_2}{\lambda^4}(\sigma_{12}+\sigma_{13}) + 2\,\frac{\lambda_1^2}{\lambda^4}\sigma_{23} .
$$

Thus, from (3.3), (3.4) and (3.6), one sided and two sided confidence limits for R can readily be obtained.

To check the adequacy of the large sample approximation when the underlying distribution is BVE, computer simulations were made. For various values of $(\lambda_1, \lambda_2, \lambda_{12})$ 500 sets of random samples of size n $(n = 10, 15, 20)$ from bivariate exponential distributions were obtained and the empirical distribution of $\hat{\hat{R}}$ was obtained. Even for sample size as low as 10 the exact distribution is found to be well approximated by the normal approximation. The situation improves as the sample size increases.

4. Distribution free procedures.

In all the cases considered in the previous section to check adequacy of normal approximation, the value of R is rather small ($R \le .75$), whereas applications of interest would be for systems with high reliability ($R > .90$). Unfortunately in all the cases with $R > .90$ considered, the sample estimate of the variance of R came out to be negative or very close to zero. (The situation is similar to the problem of having "negative" estimator of variances in analysis of variance problems.)

To study the cases for which $R > .90$, we therefore consider the following estimators of R :

For bivariate data, a rather natural way to estimate $R = P(X < Y)$ would be based on the binomial distribution. Let T be the number of cases for which $X < Y$. Then T is a binomial random variable with mean R and variance $nR(1-R)$. We therefore can obtain an exact binomial confidence interval or an approximate two sided $100(1-\gamma)\%$ confidence interval given by

$$\left(\hat{R} - z_{\gamma}\sqrt{\frac{\hat{R}(1-\hat{R})}{n}}\ ,\ \ \hat{R} + z_{\gamma}\sqrt{\frac{\hat{R}(1-\hat{R})}{n}}\right) , \tag{4.1}$$

where $\hat{R} = T/n$ and z_{γ} is such that $\Phi(z_{\gamma}) = (1-\gamma/2)$, where $\Phi(\cdot)$ is the standard normal distribution.

In case X and Y are assumed independent, a second estimator is the following nonparametric confidence interval proposed by Govindarajulu (1968) and is based on the Wilcoxon-Mann-Whitney statistic.

Let $(X_1, X_2, \ldots, X_m)$ and $(Y_1, Y_2, \ldots, Y_n)$ be two independent samples of measurements from populations with distribution functions $F(x)$ and $G(y)$ respectively. Let

$$\psi(X_i, Y_j) = \begin{cases} 1 & \text{if } X_i < Y_j \\ 0 & \text{otherwise} \end{cases}$$

then $U = \sum_{i=1}^{m} \sum_{j=1}^{n} \psi(X_i, Y_j)$ is the well known two sample Mann-Whitney statistic. That is, U = number of pairs (X_i, Y_j) such that $X_i < Y_j$. Govindarajulu (1968) has explicitly derived one-sided and two-sided distribution free confidence bounds for R (actually Govindarajulu derived confidence bounds for $1-R$) based on the asymptotic normality of $\tilde{R} = U/mn$. In particular, for the two-sided case, Govindarajulu showed that for all F and G and large m or n, the solution ε_γ of the inequality

$$(4.2) \qquad P(|\tilde{R} - R| \le \varepsilon_\gamma) \ge 1-\gamma, \quad 0 < \gamma < 1$$

is given by

$$\varepsilon_\gamma \ge (4\nu)^{-\frac{1}{2}} \Phi^{-1}(1 - \gamma/2) .$$

Here $\nu = \min(m,n)$ and $\Phi(\cdot)$ is the cdf of standard normal distribution. In particular, if $m = n$, a $100(1-\gamma)\%$ confidence interval is given by

$$(4.3) \qquad (\tilde{R} - \varepsilon_\gamma, \tilde{R} + \varepsilon_\gamma) ,$$

where $\tilde{R}$ and Z_γ are as defined before, and

$$\varepsilon_\gamma = \frac{Z_\gamma}{2\sqrt{n}} .$$

5. Effect of misspecifying the model.

In section 2 it was pointed out that for values of $P(X < Y)$ close to unity, R_{pq} is less sensitive to the variation of p and varies considerably for varying values of q. To study the effect of misspecifying the model we carry out the following Monte Carlo experiment.

Let $G(\alpha,p)$ and $G(\beta,q)$ be two given gamma populations with known parameters (α,p) and (β,q) respectively, where $\alpha = 19$ and $\beta = 1$. In this case, if $p = q = 1$, $R_{11} = .95$. By chosing different pairs (p,q) we get different pairs of gamma distributions.

Let $(X_1,X_2,\ldots,X_n)$ and $(Y_1,Y_2,\ldots,Y_n)$ be two random samples from $G(\alpha,p_1)$ and $G(\beta,q_1)$. Since p_1 and q_1, the true values are not known, there is a possibility that we will choose a different pair of distributions as the true model. Let us assume that the above samples have come from populations with distributions $G(\alpha,p_2)$ and $G(\beta,q_2)$. Thus we would estimate $R_{p_2q_2}$ instead of estimating the true value $R_{p_1q_1}$, and compute a confidence interval based on $\hat{R}_{p_2q_2}$. We would not commit much specification error if this confidence interval contains the true unknown value $R_{p_1q_1}$. For a given n and two pairs of values (p_1,q_1) and (p_2,q_2) the above procedure is repeated 1000 times and a count is made of how many times the true value $R_{p_1q_1}$ is contained in the confidence interval based on $\hat{R}_{p_2q_2}$. We repeat the procedure for different values of n and different combinations of (p_1,q_1) and (p_2,q_2). The results are given in

Table 1. Here n is chosen to be 5, 10 and 25. All Combinations of the following pairs of values are chosen for (p_1, q_1) and (p_2, q_2): (1,1), (1,2), (2,1) and (2,2) .

From Table 1 we can make a number of conclusions. First note that no parametric method performs well in all situations. For all the gamma models considered the procedure, as anticipated in section 2, is robust for small variation in p . However, it is sensitive to variation in q . The nonparametric confidence intervals based on the Wilcoxon-Mann-Whitney statistic performs well in all cases, however, in each case the width of the confidence interval for an assumed model is too large compared with the corresponding width based on the parametric models. In each case the parametric method is to be preferred, especially if the parameter q is reasonably well specified.

TABLE 1

Number of counts for N = 5

ASSUMED MODEL	TRUE MODEL			
	G(1,1) ≡ Exponential	G(2,1)	G(1,2)	G(2,2)
$R_{p_1 q_1}$	.95	.9025	.9975	.99275
G(1,1) exact	933	967	46	77
G(1,1) normal approx	855	934	1000	1000
G(2,1)	763	819	66	94
G(1,2)	287	487	828	938
G(2,2)	209	371	725	871
*NP pr	998	991	1000	1000

*Nonparametric Procedure

TABLE 1 (continued)

Number of counts for N = 10

ASSUMED MODEL		TRUE MODEL G(1,1) ≡ Exponential	G(2,1)	G(1,2)	G(2,2)
	$R_{p_1 q_1}$	.95	.9025	.9975	.99275
	G(1,1) exact	947	976	0	2
	G(1,1) normal approx	913	952	188	348
	G(2,1)	775	838	0	1
	G(1,2)	133	293	869	968
	G(2,2)	58	192	750	906
	*NP pr	999	996	1000	1000

*Nonparametric Procedure

TABLE 1 (continued)

Number of counts for N = 25

ASSUMED MODEL		TRUE MODEL G(1,1) ≡ Exponential	G(2,1)	G(1,2)	G(2,2)
	$R_{p_1 q_1}$	.95	.9025	.9975	.99275
	G(1,1) exact	945	964	0	0
	G(1,1) normal approx	934	964	0	0
	G(2,1)	768	812	0	0
	G(1,2)	6	82	925	982
	G(2,2)	1	27	754	928
	*NP pr	1000	998	1000	999

*Nonparametric Procedure

6. An Example.

To illustrate the computation of confidence intervals let us consider the following example. Fifteen items of random strengths $Y_1,\dots,Y_{15}$ are subject to random stresses $X_1,X_2,\dots,X_{15}$. To estimate the reliability function $P(X < Y)$ random samples of 15 pairs of (X,Y) values are drawn and given below.

Pair No.		
1	.0352	1.7700
2	.0397	.9457
3	.0677	1.8985
4	.0233	2.6121
5	.0873	1.0929
6	.1156	.0362
7	.0286	1.0615
8	.0200	2.3895
9	.0793	.0982
10	.0072	.7971
11	.0245	.8316
12	.0251	3.2304
13	.0469	.4373
14	.0838	2.5648
15	.0796	.6377

From past record it is known that $R = .95$. Estimates of $R_{p_1q_1}$ and corresponding confidence interval will depend on the model chosen and the method used to compute the confidence interval. Thus if X and Y are assumed independent, we have two independent samples of size 15 each. Table 2 lists the values of $\hat{R}_{p_2q_2}$ and confidence interval for the cases considered in section 5. We also use the notation of section 5 for convenience.

Let us illustrate the calculation for a couple of cases. From the above data we have $\bar{X} = .0509$ and $\bar{Y} = 1.3602$. If we assume $p_2 = q_2 = 1$ we have, using results in section 2 ,

$$R_{11} = \frac{\overline{Y}}{\overline{X} + \overline{Y}} = .9639$$

and corresponding 95% confidence interval for R_{11} , using normal approximation, is (.9300, .9978) .

On the other hand, if we used the exact distribution of $\hat{R}_{11}$, we can obtain the required confidence interval from an F-table since $U = \frac{\overline{Y}}{\overline{X}} \frac{\beta}{\alpha}$ follows the F-distribution with (2n,2n) degrees of freedom we obtain, after some simplification (.9280, .9823) as the required 95% confidence interval.

From Table 2 it seems any of the first three procedures based on two independent exponential distributions or G(2,1)(X gamma with shape parameter p = 2 and Y independent exponential) will be quite satisfactory.

TABLE 2

Comparison of Various Confidence Intervals
and Estimates of R

Model Used	$\hat{R}$	Confidence Interval
G(1,1) exact	.9639	(.9280,.9823)
G(1,1) normal approximation	.9639	(.9300,.9978)
G(2,1)	.9639	(.9416,.9856)
G(1,2)	.9952	(.9896,1.000*)
G(2,2)	.9962	(.9925,.9999)
Binomial Procedure	.9333	(.8071,1.000*)
NP Procedure	.9600	(.7070,1.000*)

* Computed as being larger than 1.

Acknowledgements

The author would like to thank Dr. Dennis Weier, of the University of South Carolina, Columbia, for his programming help. The author would also like to thank the referee for his constructive criticisms.

References

[1] Basu, A. P. (1977). Estimation of Reliability in the Stress-Strength Model. Proceedings of the 22nd conference on the Design of Experiments.

[2] Birnbaum, Z. W. (1956). On a use of the Mann-Whitney Statistic. Proc. 3rd Berkeley Symposium. Vol. 1, 13-17.

[3] Block, Henry and Basu, A. P. (1974). A Continuous Bivariate Exponential Extension. J. Amer. Stat. Assn. 69, 1031-1037.

[4] Church, J. D. and Harris, Bernard (1970). The Estimation of Reliability from Stress-Strength Relationships. Technometrics 12, 49-54.

[5] Freund, J. E. (1967). A Bivariate Extension of the Exponential Distribution. J. Amer. Stat. Assn. 56, 971-977.

[6] Govindarajulu, Z. (1968). Distribution Free Confidence Bounds for $P(X < Y)$. Ann. Inst. Stat. Math. 20, 299-238.

[7] Marshall, A. W. and Olkin, I. (1967). A Multivariate Exponential Case. J. Amer. Stat. Assn. 62, 30-44.

[8] Owen, D. B., Crasewell, R. J., and Hanson, D. L. (1964). Nonparametric Upper Confidence Bounds for $Pr\{Y<X\}$ and Confidence Limits for $Pr\{Y<X\}$ when X and Y are Normal. J. Amer. Stat. Assn. 59, 906-924.

[9] Proschan, Frank and Sullo, P. (1976). Estimating the Parameters of a Multivariate Exponential Distribution. J. Amer. Stat. Assn. 71, 465-472.

[10] Rao, C. R. (1965). Linear Statistical Inference and its Applications. John Wiley & Sons, Inc.

DATE
FILMED