Journal of Statistial
Theary & Methods
1994, Vol. 5, pp11 ~ 18

Sample Size Choice for Strength-Stress Models in the Exponential Case

Young Joon Cha[1] and Yeon-Woong Hong[2]

ABSTRACT Sample sizes for experiments concerned with inference on $R=\Pr(X<Y)$ are given in an acceptance-sampling theory framework, when X and Y are independent two-parameter exponential random variables with a known(unknown) common location parameter. The choices of sample sizes are based on the Wilson-Hilferty approximation for a known common location parameter, and are based on the asymptotic normality for an unknown common location parameter.

1. INTRODUCTION

The estimation of $R=\Pr(X<Y)$ has been considered by serveral authors. A recent review of the subject can be found in Johnson(1988). Usually R is regarded as a measure of reliability or the performance of an item of strength Y subject to stress X. The literature has focused on point and interval estimation under parameteric and nonparameteric assumptions on the distributions of X and Y (see for example, Reiser and Guttman 1986, 1987; Simonoff, Hochberg, and Reiser 1986). Reiser and Guttman(1989) considered the problem of sample size choice for

1 Dept. of Statistics, Andong National University
2 Dept. of Industrial Engineering, Dongyang University of Technology

reliability verification. They discussed this problem when designing experiments for reliability in a strength-stress situation in order to conservatively meet requirements on the producer's and consumer's risks. These results are derived assuming normality.

In this paper we consider the problem of sample size choice in the independent exponential case with a common location parameter. We use the Wilson-Hilferty approximation to determine sample sizes when a common location parameter is known, and use the asymptotic normality when a common location parameter is unknown.

2. ASSUMPTIONS AND NOTATIONS

Assumptions

(i) X and Y are independent random variables with means $\theta_1+\mu$ and $\theta_2+\mu$, respectively.

(ii) Sample sizes for X and Y are equal, say n .

Notations

i	sample number, $i=1,2$
n	sample size for $X(i=1)$ and $Y(i=2)$
n'	n-1
α, β	producer, consumer risks
μ	common location parameter for X and Y
θ_i	scale parameter of X and Y
θ	θ_1/θ_2
R	$\Pr(X< Y)=1/(1+\theta)$
R_a	acceptable(high) reliability
R_r	rejectable(low) reliability
R_c	critical value
θ_a	$1/R_a-1$

θ_r	$1/R_r-1$
θ_c	$1/R_c-1$
$x_1,\cdots,x_n,\ y_1,\cdots,y_n$	random samples for X and Y
w_1, w_2	$\sum_{i=1}^{n} x_i,\ \sum_{j=1}^{n} y_j$
$x_{(1)},\ y_{(1)}$	$\min(x_1,\cdots,x_n),\ \min(y_1,\cdots,y_n)$
z	$\min(x_{(1)}, y_{(1)})$
t_1, t_2	$w_1-nz,\ w_2-nz$
$t'_1,\ t'_2$	$w_1-nx_{(1)},\ w_2-ny_{(1)}$
$\hat{}$	the maximum likelihood estimate(MLE)
$\sim$	is distributed as
$\cong$	is approximately equal to
df	degrees of freedom
iid	independent and identically distributed
$F_{a,b}$	F distribution with df a and b
rv	random variable

3. DESIGN PROCEDURE

Our procedure for sampling size problem can be stated as follows :
Given R_a, α, R_r, β with $0<R_r<R_a<1$, $0<\alpha<1$, and $0<\beta<1$, we want to find an *acceptance rule* and n such that

(i) the probability of acceptance is $1-\alpha$ if $R=R_a$,

(ii) the probability of acceptance is β if $R=R_r$.

3.1 μ is known

Without loss of generality, we let $\mu=0$. Under the predescribed assumptions, inference on R is equivalent to inference on $\theta=\theta_1/\theta_2$. Inference on θ can be based on the relationship

$$\frac{2w_2/2n\theta_2}{2w_1/2n\theta_1} = \theta/\hat{\theta} \sim F_{2n,2n} , \qquad (1)$$

where $\hat{\theta} = w_1/w_2$.

Instead of working with R, one can equivalently work with θ. Then our problem is now to find θ_c and n such that

$$\Pr(\hat{\theta} < \theta_c \mid \theta = \theta_a) \geq 1 - \alpha \qquad (2)$$

and

$$\Pr(\hat{\theta} < \theta_c \mid \theta = \theta_r) \leq \beta. \qquad (3)$$

Let z_p and $F_{p;a,b}$ denote the p percentile points of the standard normal distribution and the F distribution with df a and b , respectively. From (1) this is equivalent to solving

$$\Pr(F_{2n,2n} > \theta_a/\theta_c) \geq 1 - \alpha \qquad (4)$$

and

$$\Pr(F_{2n,2n} > \theta_r/\theta_c) \leq \beta \qquad (5)$$

for θ_c and n.

Thus it follows from (4) and (5) that

$$\theta_a/\theta_c = F_{\alpha;2n,2n} \qquad (6)$$

and

$$\theta_r/\theta_c = F_{1-\beta;2n,2n} . \qquad (7)$$

In order to solve (6) and (7) we use the following Wilson-Hilferty approximation with $F_p = F_{p;a,b}$:

$$\left[F_p^{1/3}\left(1-\frac{2}{9b}\right)-\left(1-\frac{2}{9a}\right)\right] \cdot \left[\frac{2}{9a}+\frac{2}{9b}F_p^{2/3}\right]^{-1/2} \cong z_p . \qquad (8)$$

In our case $a = b = 2n$, and

$$\frac{9n-1}{\sqrt{9n}} \cdot (F_p^{1/3}-1)\ (1+F_p^{2/3})^{-1/2} \cong z_p . \qquad (9)$$

Appling (9) to (6) and (7), it follows that

$$\frac{9n-1}{\sqrt{9n}} \cdot [\ (\theta_a/\theta_c)^{1/3}-1]\ [\ 1+\ (\theta_a/\theta_c)^{2/3}]^{\ -1/2} \cong z_\alpha \qquad (10)$$

and

$$\frac{9n-1}{\sqrt{9n}} \cdot [\ (\theta_r/\theta_c)^{1/3}-1]\ [\ 1+\ (\theta_r/\theta_c)^{2/3}]^{\ -1/2} \cong -z_\beta\ . \qquad (11)$$

Solving (10) and (11) for θ_c gives

$$\begin{aligned}&\left[\ \left(\frac{z_\alpha}{z_\beta}\right)^2-1\right]\ \theta_c^{4/3}-2\cdot\left[\ \left(\frac{z_\alpha}{z_\beta}\right)^2\ \theta_r^{1/3}-\ \theta_a^{1/3}\right]\ \cdot\ \theta_c^{3/3}\\ &+\left[\ \left(\frac{z_\alpha}{z_\beta}\right)^2-1\right]\ (\ \theta_a^{2/3}+\ \theta_r^{2/3})\cdot\ \theta_c^{2/3}\\ &-2\cdot\left[\ \left(\frac{z_\alpha}{z_\beta}\right)^2\theta_a^{1/3}-\ \theta_r^{1/3}\right]\ (\theta_r\theta_a)^{1/3}\cdot\ \theta_c^{1/3}\\ &+\left[\ \left(\frac{z_\alpha}{z_\beta}\right)^2-1\right]\ (\theta_r\theta_a)^{2/3}\ =\ 0\quad .\end{aligned} \qquad (12)$$

First we consider the case of $\alpha=\beta$. Then (12) reduces to

$$(\ \theta_r^{1/3}-\ \theta_a^{1/3})\cdot\ \theta_c^{2/3}+(\ \theta_a^{1/3}-\ \theta_r^{1/3})(\ \theta_a\theta_r)^{1/3}\ =\ 0$$

which yields the approximate closed form solution

$$\theta_c=\ (\theta_a\theta_r)^{1/2}, \qquad (13)$$

and thus

$$R_c=\frac{\sqrt{R_aR_r}}{\sqrt{R_aR_r}+\sqrt{(1-R_a)(1-R_r)}}\ . \qquad (14)$$

Assuming $9n-1\cong 9n$, solving (10) and (11) for n gives

$$n=\frac{\theta_a^{2/3}+\theta_c^{2/3}}{9\ (\theta_a^{1/3}-\theta_c^{1/3})^2}\ \cdot z_\alpha^2=\frac{\theta_r^{2/3}+\theta_c^{2/3}}{9\ (\theta_r^{1/3}-\theta_c^{1/3})^2}\ \cdot z_\beta^2. \qquad (15)$$

And when $\alpha \neq \beta$, (12) is a quadratic equation of $\theta_c^{1/3}$. We know that (12) has unique real zero between $\theta_a^{1/3}$ and $\theta_r^{1/3}$. Therefore, n is obtained by substituting this solution into (15).

3.2 μ is unknown

When a common location parameter μ is unknown, $\hat{R}=t_2/(t_1+t_2)$ is the MLE and uniformly minimum variance unbiased estimator of R, and $(1+\theta/\hat{\theta})^{-1}$ has a mixed beta distribution. Because of the complexity of the normal approximation formula to a mixed beta distribution, we consider two approximate methods based on

(i) the distribution of t'_1 and t'_2 (rather than the t_1 nad t_2),

(ii) the large sample theory of $\hat{R}$.

Case 1. Method based on distribution of t'_1 and t'_2

Since $2t'_i/\theta_i$, $i=1,2$, have iid chi-square rv with df $2n'$,

$$\frac{2t'_2/2n'\theta_2}{2t'_1/2n'\theta_1} = \theta/\hat{\theta}' \sim F(2n', 2n') \ , \tag{16}$$

where $\hat{\theta}' = t'_1/t'_2$, and from the result of μ unknown case, we know that R_c is the same as formula (13) and sample size is formula (15) plus 1, respectively.

Case 2. The large sample theory of $\hat{R}$.

In the following, large sample theory is used to derive equations for n and R_c for a given two points $(R_a, 1-\alpha)$ and (R_r, β). Then the standardized variate $U=\frac{\sqrt{n}(\hat{R}-R)}{2R(1-R)}$ is parameter-free and asymptotically standard normally distributed. Thus our problem is to find U_c and n such that

$$\Pr(U > U_c \mid R=R_a) = 1-\alpha \tag{17}$$

and

$$\Pr(U > U_c \mid R = R_r) = \beta . \tag{18}$$

It follows from (17) and (18) that

$$U = \frac{\sqrt{n}(R_c - R_a)}{2R_a(1-R_a)} = z_\alpha \tag{19}$$

and

$$U = \frac{\sqrt{n}(R_c - R_r)}{2R_r(1-R_r)} = -z_\beta . \tag{20}$$

Solving (18) and (19) for R_c and n yields the solution

$$R_c = \frac{R_a R_r(1-R_a)z_\alpha + R_r R_a(1-R_r)z_\beta}{R_a(1-R_a)z_\alpha + R_r(1-R_r)z_\beta} , \tag{21}$$

and

$$n = \frac{4[\, z_\alpha R_a(1-R_a) + z_\beta R_r(1-R_r)]^{\,2}}{(R_r - R_a)^2} . \tag{22}$$

REFERENCES

(1) Johnson, R.A. (1988). Stress-strength models for reliability. *Handbook of Statistics*, 7 (P.R. Kirshnaiah and C.R. Rao, ed.), North Holland.

(2) Reiser, B., and Guttman, I. (1986). Statistical inference for $P(Y < X)$: the normal case. *Technometrics*, 28, 253-257.

(3) Reiser, B., and Guttman, I. (1987). A comparison of three-point estimators for $P(Y < X)$ in the normal case. *Comput. Statist. Data Anal.* 5, 59-66.

(4) Reiser, B., and Guttman, I. (1989). Sample size choice for reliability verification in strength-stress models. *The Canadian Journal of Statistics*, No. 3, 253-259.