Defence Science Journal, Vol. 54, No. 1, January 2004, pp. 17-19

SHORT COMMUNICATION

Estimation of Probability Density Function of a Random Variable for Small Samples

V.S. Srinivasan

C-41/3, *Phase-2, DRDO Township, CV Raman Nagar, Bangalore*–560 093

ABSTRACT

A theoretical method of estimating the most general form of probability density functions of random variables has been described. The estimated probability density function depends on the mean value of exponential distribution of the initial random variable. By the most general form of the probability density function, it is meant that the survivor functions are of exp $(-\lambda t^2)$ or exp $(-\gamma t^3)$ or combinations of these functions for different values of λ and γ.

Keywords: Survivor function, hazard function, hotelling-T distribution, moments, parent-variable, probability density function, random variables

NOMENCLATURE

t_1'	A Gaussian random variable with mean 0 and variance ½
t_1''	An exponential random variable with mean 1
t_1	A hotelling-T³ random variable with mean 3
t_2	A Gaussian random variable with mean 0 and variance 3/128
t_3	Same as t_1''
N	Sample size (small); N=2, 3
$R(.)$	Survivor function
$E(.)$	Expected value of (.)
r.v.	Random variable
pdf	Probability density function
t	Parent-variable

1. INTRODUCTION

It is known that the probability density function (*pdf*) $f(t)$ of any random variable is given by

$$f(t) = h(t) \cdot R(t), \tag{1}$$

where $h(t)$ is the hazard function and $R(t)$ is the survivor function given by

$$R(t) = \exp\left[-\int_0^t h(x)\,\mathrm{d}x\right] \tag{2}$$

The variable t is called parent variable. The hazard function can be written in the following two ways:

$$h_1(x) = \mu + \lambda x \tag{3}$$

$$h_2(x) = \mu + \lambda x + \gamma x^2 \tag{4}$$

Revised 22 April 2003

The Eqn (3) is justified if one assumes a linear relationship between x and $h_1(x)$. The equation (4) is justified if one assumes a quadratic relationship between x and $h_2(x)$.

When the hazard function satisfies the Eqn (3), the survivor function is:

$$R_1(t) = \exp\left(-(\lambda t^2)/2 + \mu t\right) \tag{5}$$

When the hazard function satisfies the Eqn(4), the survivor function is:

$$R_2(t) = \exp\left[-\left(\gamma t^3/3 + \lambda t^2/2 + \mu t\right)\right] \tag{6}$$

The problem considered here is that for which values of (λ and μ) or (γ, λ, μ), the most general representation [Eqn (6)] is true? It is true that different forms of probability density function for different equipment are available. Two models are considered here.

2. VARIOUS MODELS

The model I describes the case (i), where λ is so chosen that the variance of the parent variable, t, is kept unaltered at the first and the second stages. Without loss of generality, it is assumed that $\mu = 1$. The model II describes the case (ii), where γ is so chosen that the third moment of the parent variable, t, is kept unaltered at the third stage.

2.1 Model I

2.1.1 *Problem*

The study undertaken here is to determine the parameter values (λ, μ) in Eqn (5). To do this, it was noted that the survivor function in Eqn (5) is a product of two probability density functions which is of the form:

$$f(x) = f(x/t)\ f_\theta(t) \tag{7}$$

where the parent variable, x, is a Gaussian random variable with mean 0 and variance $1/\lambda$. The Eqn(7) indicates that $f(x/t)$ serves as a scaling factor for the exponential probability density function.

2.1.2 *Method*

The method consists of taking n random samples $t_1, t_2, \ldots\ldots\ldots\ldots t_n$ from exponential probability density function with mean 1 and substituting these sample values in the following equations:

$$\hat{f}(x,t) = 1/\bar{t} \tag{8}$$

$$\hat{\lambda} = \left(2/x^2\right) \ln\left(\bar{t}\right) \tag{9}$$

where $\bar{t}$ is the sample mean and $\hat{\gamma}$ is the estimate of λ as a function of x.

2.2 Model II

2.2.1 *Problem*

The problem in this model is to determine the parameter values (λ, μ, γ) in Eqn (6). To do this, it was noted that the survivor function is a product of two probability density functions which is of the form:

$$f(x) = g(T,\ t/x)\ f_\lambda(x) \tag{10}$$

where the parent variable is hotelling-T^3 random variable. The above equation indicates that $g(T, t/x)$ serves as a scaling factor for the Gaussian probability density function. The symbols T, t and x represent variables t_1, t_3 and t_2.

2.2.2 *Method*

To illustrate the method, one should consider only the case of three samples. The method consists of taking three samples x_1, x_2, x_3 from Gaussian probability density function with mean 0 and variance (3/128) and substituting the sample values in the following equations:

$$\sqrt{2\pi}\ \hat{g}(T,\ t) = 1/s_1 \tag{11}$$

$$\sqrt{2\pi}\ \hat{g}(T, 0) = 1/s_0 \tag{12}$$

$$\hat{\gamma} = (3/T^3)\ \left\{\ln(s_1) - \ln(s_0)\right\} \tag{13}$$

$$\hat{\mu} = (1/t)\ \ln\left\{\sqrt{2\pi}\ s_0\right\} \tag{14}$$

$$s_0^2 = \sum_{i=1}^{2}(x_i - \bar{x})^2 \quad (15)$$

$$s_1^2 = (1/2) \cdot \sum_{i=1}^{3}(x_i - \bar{x})^2 \quad (16)$$

where $\bar{x}$ is the sample mean, $\hat{\gamma}$ is the estimate of γ as a function of T, and $\hat{\mu}$ is the estimate of μ as a function of t. When $\lambda = 2$ and $\mu = 1$, the effective value of γ is found to be 3.5 approx.

3. CONCLUSION

This paper describes a general method of splitting the parent random variable into two linear combinations of sub-variables (t_1', t_1'') where t_1' is a Gaussian random variable with mean 0 and variance 1/2, and t_1'' is an exponential random variable with mean 1 and (t_1, t_2, t_3), where t_1 is a hotelling-T^3 random variable with mean 3, t_2 is a Gaussian random variable with mean 0 and variance 3/128, and t_3 is same as t_1''. For determination of values (λ, μ) or (γ, λ, μ), it is considered that all moments of the random variables are to be unaltered. This study can be applied in reliability theory because the proper choice of (λ, μ) or (γ, λ, μ) will reduce the cost of maintenance in plant reliability problems.

ACKNOWLEDGEMENTS

The author thanks Shri S. Kannan, Scientist, Centre for Airborne Systems, Bangalore and Shri S. Sundararajan of Aeronautical Development Agency, Bangalore, for their help in the preparation of this paper.

REFERENCES

1. Modarres, Mohammad; Kaminskiy, Mark & Krivtsov, Vasiliy. Reliability engineering and risk analysis–A practical guide. Marcel Dekker Inc, New York, Basel, 1999.
2. Kalman, R.E. A new approaches to linear filtering and prediction problems. *J.Basic Engg.*, 1960, **82**, 35-45.
3. Mudholkar, G..S.; Srivastava, D.K. &Freimer, M. The exponentiated Weibull family; A re-analysis of bus-motor failure data. *Technometrics*, 1995, **37**(4), 436-45.

Contributor

Dr VS Srinivasan received his PhD (Operations Research) from the University of Delhi, Delhi, in 1974. He joined DRDO in 1963 and worked at its various laboratories as Scientist. He worked at the Electronics & Radar Development Establishment, (LRDE), Bangalore, from 1969 to 1978. Later, he joined Naval Physical & Oceanographic Laboratory (NPOL), Kochi, from 1978 to 1986, till his retirement. His areas of research include: Reliability, statistics, and signal detection. He has published 37 papers in various national/international journals. He served as a referee for the *IEEE Transactions on Reliability*. He has also presented many papers in seminars. At LRDE, he acted as leader/member for a number of projects.