KYBERNETIKA — VOLUME 25 (1989), NUMBER 3

NEW MODELS IN DURABILITY TOOL-TESTING: PSEUDO-WEIBULL DISTRIBUTION

VIOREL GH. VODĂ

In this paper a new class of distribution functions is introduced as a model for time to failure in tool durability testing. The distribution has been called "Pseudo-Weibull". Some inferences on its properties are presented and a peculiar case is extensively investigated.

1. INTRODUCTION

The problem of tool durability is considered to be a most important factor in mechanical engineering. In the last three decades various models for tool life have been proposed. Among these models the statistical ones play a major role, since almost all tool characteristics are regarded as random variables.

Classical distribution functions – as for instance log-normal, exponential or Weibull ones – have proved their utility in this matter, but the large variety of tools themselves have imposed the need to find other distributions to describe the behaviour of tool life.

Two Soviet engineers (Družhinin [1] and Katzev [2]) have introduced the so-called "Alpha" distribution, initially to describe merely the time for performing a given operation. A detailed statistical analysis of the Alpha model has been performed by the present author in [3].

In the present paper we shall derive a new distribution function which has proved its utility in durability tool-testing.

2. MATHEMATICAL DERIVATION

It is known that in the reliability theory there are constructed various classes of distribution functions, using a given distribution function as generator. The important thing is the following: does the new distribution function preserve some properties of the former one? For instance, it is known that if $F(x)$ is an IFR (increasing failure rate) distribution, then the distribution of

(2.1) $F_1(x) = \mu \int_0^x \bar{F}(u)\,du$ where $\bar{F}(x) = 1 - F(x)$ and $\mu = \int_0^\infty x\,dF(x)$

is also an IFR distribution (cf. [5]).

In the following, we shall find an analogy for the densities and we shall study the conservability of the property: if $f(x)$ belongs to a certain class of densities, then $f_1(x)$ – constructed below – belongs to the same class of densities.

By analogy to the above written relationship, we define

$$(2.2) \qquad f_1(x) = \frac{x}{\mu_1} f(x) \quad \text{where} \quad \mu_1 = \int_0^\infty x f(x)\,dx\,.$$

Obviously, f_1 is a density function. Let us consider now the following cases:

1) *Gamma distribution*

$$X: f(x; \theta, k) = [\theta^k\,\Gamma(k)]^{-1} \exp(-x/\theta)\,, \quad x, \theta, k > 0\,.$$

Since $\mathsf{E}(X) = \mu_1 = k\theta$, we have immediately

$$X_1: f_1(x) = [k\theta^{k+1}\,\Gamma(k)]^{-1}\,x^k \exp(-x/\theta) = [\theta^{k+1}\,\Gamma(k+1)]^{-1}\,x^k \exp(-x/\theta)$$

Hence, if $X \in GA(x; \theta, k)$, then $X_1 \in GA(x; \theta, k + 1)$. Conservative.

2) *Exponential distribution*

$$X: f(x; \theta) = \theta^{-1} \exp(-x/\theta)\,, \quad x, \theta > 0\,.$$

We have $\mathsf{E}(X) = \theta$ and therefore

$$X_1: f_1(x) - \theta^{-2} x \exp(-x/\theta)\,, \quad x, \theta > 0\,.$$

It follows that if $X \in EXP(x; \theta)$, then $X_1 \in GA(x; \theta, 2)$. Non-conservative.

3) *Log-normal distribution*

$$X: f(x; \mu, \sigma^2) = (\sigma x \sqrt{(2\pi)}^{-1} \exp\{-(\ln x - \mu)^2/2\sigma^2\}\,.$$

We have $\mathsf{E}(X) = \exp(\mu + \sigma^2/2)$ and hence

$$X_1: f_1(x) = (\sigma x \sqrt{(2\pi)})^{-1} \exp\{-(\ln x - \mu - \sigma^2)^2/2\sigma^2\}\,, \quad x, \sigma > 0\,,$$
$$\mu \in \mathbb{R}\,.$$

Therefore if $X \in LN(x; \mu, \sigma^2)$ then $X_1 \in LN(x; \mu + \sigma^2, \sigma^2)$. The distribution is conservative.

4) *Generalized Rayleigh distribution* (see [6])

$$X: f(x; \theta, k) = 2\theta^{k+1}[\Gamma(k+1)]^{-1}\,x^{2k+1} \exp(-\theta x^2)\,, \quad x, \theta > 0\,, \quad k > -1\,.$$

Since $\mathsf{E}(X) = \theta^{-1/2}\,\Gamma(k + \frac{3}{2})/\Gamma(k + 1)$ we deduce

$$X_1: f_1(x) = 2\theta^{k+3/2}[\Gamma(k + \tfrac{3}{2})]^{-1}\,x^{2k+2} \exp(-\theta x^2)\,, \quad x > 0\,, \quad \theta > 0\,, \quad k > -\tfrac{3}{2}.$$

Hence, if $X \in GRV(x; \theta, k)$, then $X_1 \in GRV(x; \theta, k + \frac{1}{2})$. Conservability is preserved.

5) The *Weibull* case seems to be the most interesting.

$$X: f(x; \theta, k) = k\theta^{-1}x^{k-1} \exp(-x^k/\theta), \quad x > 0, \quad \theta, k > 0.$$

We have $\mathsf{E}(X) = \theta^{1/k}\,\Gamma(1 + 1/k)$ and it follows immediately that

$$(2.3) \qquad X_1 : f_1(x) = kx^k[\theta^{1+1/k}\,\Gamma(1 + 1/k)]^{-1} \exp(-x^k/\theta), \quad x, \theta, k > 0.$$

The distribution is not conservative. Due to this fact, we shall call this distribution a "Pseudo-Weibull" one. Let us notice also that if $k = 1$, we obtain

$$f_1(x) = \theta^{-2}x \exp(-x/\theta), \quad \theta > 0, \quad x > 0,$$

which is just $GA(x; \theta, 2)$. Observe also that (2.3) is not a generalization of the Gamma distribution since it does not contain Weibull or Gamma distribution as particular cases.

If we make now a little bit of statistical history, it is interesting to notice that the generalized Gamma distribution studied by Stacy [3] was introduced in fact by L. Amoroso in 1925 in the Italian journal "Annali di Matematica Pura ed Applicata" in order to model the evalution of income in Italy for a period of 20 years (cf. [3]).

3. STATISTICAL INFERENCE ON PSEUDO-WEIBULL VARIATE

We give now some straightforward results:

$$\mathsf{E}(X_1^j) = \int_0^\infty x^j f_1(x)\,dx = \theta^{j/k}\,\Gamma\left(1 + \frac{1+j}{k}\right)\Big/\Gamma(1 + 1/k)$$

which leads to

$$\mathsf{E}(X_1) = \theta^{1/k}\,\Gamma(1 + 2/k)/\Gamma(1 + 1/k), \quad \mathsf{E}(X_1^2) = \theta^{2/k}\,\Gamma(1 + 3/k)/\Gamma(1 + 1/k)$$

providing

$$\mathsf{Var}(X_1) = \theta^{2/k}[\Gamma(1 + 3/k)\,\Gamma(1 + 1/k) - \Gamma^2(1 + 2/k)]/\Gamma^2(1 + 1/k)$$

$$\mathsf{STD}(X_1) = \sqrt{\mathsf{Var}(X_1)} = \theta^{1/k}[\Gamma(1 + 3/k)\,\Gamma(1 + 1/k) - \Gamma^2(1 + 2/k)]^{1/2}/\Gamma(1 + 1/k)$$

$$\mathsf{CV}(X_1) = \{\Gamma(1 + 3/k)\,\Gamma(1 + 1/k) - \Gamma^2(1 + 2/k)\}^{1/2}/\Gamma(1 + 2/k)$$

The coefficient of variation can be used for estimation purposes. The distribution function has the form

$$F(x; \theta, k) = \Gamma_{(\theta x)^{1/2}}(1 + 1/k)/\Gamma(1 + 1/k)$$

where by $\Gamma_*(\cdot)$ we understand the incomplete Gamma function (see for details [3]).

We shall prove now the following result.

Lemma. If X_1 is a Pseudo-Weibull variable with the density

$$f_1(x; \theta, k) = k[\theta^{1+1/k}\,\Gamma(1 + 1/k)]^{-1}\,x^k \exp(-x^k/\theta)$$

then the variable X_1^k belongs to the class $GA(x; \theta, 1 + 1/k)$.

Proof. Let us write the distribution of the variable X_1^k. We have

$$F(y) = \text{Prob}\{X_1^k < y\} = \text{Prob}(X_1 < y^{1/k}) = (\text{const.}). \int_0^{y^{1/k}} x^k \, e^{-x^k/\theta} \, dx .$$

Put now $u = x^k/\theta$ and hence $x = (\theta u)^{1/k}$, $dx = k^{-1}\theta^{1/k}u^{1/k-1}$. We have hence

$$F(y) = \frac{1}{\Gamma(1 + 1/k)} \int_0^{y/\theta} u^{(1+1/k)-1} \, e^{-u} \, du = \frac{\Gamma_{y/\theta}(1 + 1/k)}{\Gamma(1 + 1/k)}$$

from which it follows immediately

$$f(y; \theta, k) = F'(y; \theta, k) = \frac{y^{1/k}}{\theta^{1+1/k} \, \Gamma(1 + 1/k)} \exp(-y/\theta) . \qquad \square$$

4. PARAMETER ESTIMATION

Consider that the parameter k is known. Write now the likelihood function

$$L(x_1, x_2, \ldots, x_n; \theta) = \frac{k^n \prod_{i=1}^{n} x_i^k}{\theta^{n+n/k} \, \Gamma^n(1 + 1/k)} \exp\left[-(1/\theta) \sum_{i=1}^{n} x_i^k\right]$$

and taking logarithms, we obtain

$$\ln L = n \ln k + k \sum_{i=1}^{n} \ln x_i - (n + nk) \ln \theta - n \ln \Gamma(1 + 1/k) - (1/\theta) \sum_{i=1}^{n} x_i^k ;$$

$$\frac{\partial \ln L}{\partial \theta} = -(n + n/k)\, \theta^{-1} + \theta^{-2} \sum_{i=1}^{n} x_i^k = 0$$

with the solution

$$\hat{\theta} = [n(1 + 1/k)]^{-1} \sum_{i=1}^{n} x_i^k .$$

Since the variables x_i^k are Gamma distributed, it follows that it is unbiased (the mean value of x_i^k is $\theta(1 + 1/k)$).

5. PROPERTIES OF THE VARIABLE $PW(x; \theta, 1)$

In this section we shall study some properties of the variable $PW(x; \theta, 1)$ with the density

$$X: f(x; \theta) = \theta^2 x \exp(-\theta x) , \quad x \geqq 0 , \quad \theta > 0 .$$

The reliability function is

$$R(x; \theta) = 1 - F(x; \theta) = (1 + \theta x) \, e^{-\theta x} , \quad x \geqq 0 , \quad \theta > 0 .$$

The failure rate has the expression

$$h(x; \theta) = \frac{f(x; \theta)}{R(x; \theta)} = \theta\left(1 - \frac{1}{1 + \theta x}\right), \quad x \geqq 0 , \quad \theta > 0$$

that is the variable X is of IFR type. Our variable is unimodal one, the mode being

$X_{m0} = 1/\theta$. Now, it is interesting to observe that

$$\mathsf{E}(X^{j-1}) = \Gamma(j + 1)/\theta^{j-1}$$

and hence

$$\mathsf{E}(X) = 2/\theta\,, \quad \mathsf{E}(X^2) = 6/\theta^2\,, \quad \mathsf{CV}(X) = 1/\sqrt{2}\,.$$

This is an important property to have the coefficient of variation constant. We see also that $X_{m0} = \mathsf{E}(X)/2$, therefore an estimation of the mode is $X_{m0} = \sum x_i/2n$, where x_i, $i = 1, 2, \ldots, n$ is a random sample drawn from X.

The moment generating function and the characteristic function are respectively

$$G(t) = \mathsf{E}(e^{tx}) = \theta^2 \int_0^\infty x\, e^{-(\theta - t)x}\, dx = \frac{\theta^2}{(\theta^2 - t)^2}\,(1 - \theta^{-1}t)^{-2}$$

$$\varphi(t) = \mathsf{E}(e^{itx}) = (1 - \theta^{-1}\, it)^{-2}$$

The median of the distribution is proved by the equation $F(x) = \frac{1}{2}$ and leads to the transcendental equation

$$\theta X_{mc} = \ln 2 + \ln(1 + \theta X_{mc})\,.$$

An approximation of the median is therefore

$$X_{mc} \approx \theta^{-1}\sqrt{\ln 2}\,.$$

The estimation of the parameter is interesting since MLE estimator is easily given as

$$\widehat{\left(\frac{1}{\theta}\right)} = \frac{1}{2n}\sum_{i=1}^{n} x_i$$

which is in fact an estimator of the mode. This estimate is unbiased and with minimum variance, since we have the relationship

$$\mathsf{E}\widehat{\left(\frac{1}{\theta}\right)} = \frac{1}{2n}\sum_{i=1}^{n}\mathsf{E}(X_i) = \frac{1}{\theta}$$

$$\mathsf{Var}\widehat{\left(\frac{1}{\theta}\right)} = \frac{1}{4n^2}\sum_{i=1}^{n}\mathsf{Var}(X_i) = \frac{1}{2n\theta^2}$$

(this last expression is the Rao-Cramer bound).

We shall now proceed to test a simple statistical hypothesis H_0: $\theta = \theta_0$ versus H_1: $\theta = \theta_1$ where $\theta_0 < \theta_1$. The test used will be the maximum likelihood ratio one, with only one observation. We have immediately

$$r(x) = \frac{f(x;\theta_1)}{f(x;\theta_0)} = \left(\frac{\theta_1}{\theta_0}\right)^2 \exp\{-(\theta_1 - \theta_0)\,x\}\,, \quad x \geqq 0\,.$$

Accordingly with the Neyman-Pearson lemma, the critical region is determined for such values of x for which $r(x) > K$. Therefore, the critical region W is the set of points $W = (0, x_c)$. Setting the significance level of the test – let it be α – the decision constant x_c is derived from the equation

$$\int_0^{x_c} f(x;\theta_0)\, dx = \alpha\,, \quad \text{or} \quad 1 - (1 + \theta_0 x_c)\, e^{-\theta_0 x_c} = \alpha$$

Table 1.

No.	Name	$f(x)$	μ_1	$f_1(x)$	Conservability	
					yes	no
1	$GA(x;\theta,k)$	$\dfrac{x^{k-1}}{\theta^k\,\Gamma(k)}\exp\{-x/\theta\}$	$k\theta$	$\dfrac{x^k}{\theta^{k+1}\,\Gamma(k+1)}\exp\{-x/\theta\}$	$GA(x;\theta,k+1)$	–
2	$EXP(x;\theta)$	$\theta^{-1}\exp\{-x/\theta\}$	θ	$\theta^{-2}x\exp\{-x/\theta\}$	–	$GA(x;\theta,2)$
3	$LN(x;\mu,\sigma^2)$	$\dfrac{1}{\sigma x\sqrt{(2\pi)}}\exp\left\{-\dfrac{(\ln x-\mu)^2}{2\sigma^2}\right\}$	$e^{\mu+\sigma^2/2}$	$\dfrac{1}{\sigma x\sqrt{(2\pi)}}\exp\left\{-\dfrac{(\ln x-\mu-\sigma^2)^2}{2\sigma^2}\right\}$	$LN(x;\mu+\sigma^2,\sigma^2)$	–
4	$GRV(x;\theta,k)$	$\dfrac{2\theta^{k+1}}{\Gamma(k+1)}x^{2k+1}e^{-\theta x^2}$	$\dfrac{\theta^{-\frac{1}{2}}\Gamma(k+\frac{3}{2})}{\Gamma(k+1)}$	$\dfrac{2\theta^{k+\frac{3}{2}}}{\Gamma(k+\frac{3}{2})}x^{2k+2}e^{-\theta x^2}$	$GRV(x;\theta,k+\frac{1}{2})$	–
5	$P(x;\delta,b)$	$\delta b^{-\delta}x^{\delta-1}$	$\dfrac{\delta b}{\delta+1}$	$(\delta+1)\,b^{-(\delta+1)}\,x^{\delta}$	$P(x;\delta+1,b)$	–
6	$W(x;\theta,k)$	$k\theta^{-1}x^{k-1}\exp\{-x^{k/\theta}\}$	$\theta^{\frac{1}{k}}\Gamma\left(1+\dfrac{1}{k}\right)$	$\dfrac{kx^k}{\theta^{1+1/k}\,\Gamma(1+1/k)}\exp\{-x^{k/\theta}\}$	–	$PW(x;\theta,k)$

$f_1 = \dfrac{x}{\mu_1}f(x)$, $\mu_1 = \int_0^\infty x\,f(x)\,dx < +\infty$, $GA(x;\theta,2) \equiv PW(x;\theta,1)$ $P(x;\delta,b) \sim$ power distribution which leads to the solution (by certain approximations)

$$x_c \approx \sqrt{(2)}\,(\theta_0^{-1})\,[-\ln(1-\alpha)]^{1/2}\,.$$

The behaviour of $r(x)$ function is given in Figure 1. The decision is taken usually,

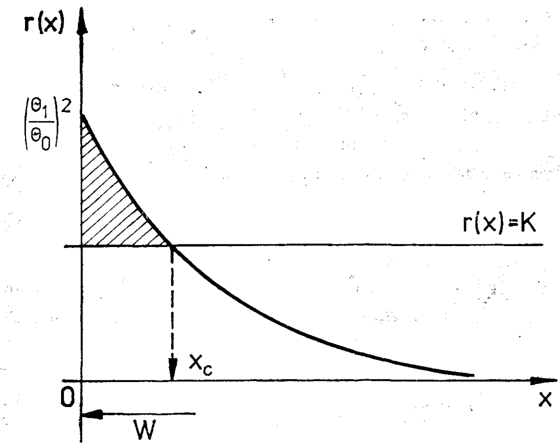

Fig. 1. The behaviour of LR function.

namely if $x \leqq x_c$ then accept H_0 and if $x > x_c$ then accept H_1. The power of the test is deduced in the following way:

$$\pi(\theta_1) = \int_0^{x_c} f(x; \theta_1)\,dx \approx$$

$$\approx 1 - \left\{1 + \sqrt{2}\frac{\theta_1}{\theta_0}[-\ln(1-\alpha)]^{1/2}\right\}\exp\left\{-\sqrt{(2)}\frac{\theta_1}{\theta_0}[-\ln(1-\alpha)]^{1/2}\right\}$$

If we denote θ_1/θ_0 by δ we have

$$\pi(\delta) \approx 1 - [1 + \delta\sqrt{(2)}\,(-\ln(1-\alpha))^{1/2}]\exp[-\delta\sqrt{(2)}\,(-\ln(1-\alpha))^{1/2}].$$

(Received November 21, 1986.)

REFERENCES

[1] G. V. Družhinin: Reliability of Automatic Systems (in Russian). Energhya, Moscow 1967.

[2] P. G. Katzev: Statistical Methods for Studying Cutting Tools (in Russian). Mashinostroyenie, Moscow 1974.

[3] V. Gh. Vodă: New Statistical Models in Durability Analysis (in Romanian). Academy Publ. House, Bucharest 1980.

[4] Al. C. Dorin and V. Gh. Vodă: Reliability of cutting tools (in Romanian). Quality and Metrology *III* (1973), 11, 651—655.

[5] R. E. Barlow and F. Proschan: Statistical Analysis of Reliability and Life Test Data. Holt, Rinehart and Winston, New York 1975.

[6] V. Gh. Vodă: "Inverse Rayleigh" distributed random variable. Rep. Stat. Appl. Res. (JUSE), Tokyo, vol. *19* (1972), no. 4, 13—21 and vol. *21* (1974), no. 1, 1—8.

Dr. Viorel Gh. Vodă, Center of Mathematical Statistics, Magheru Blvd. No 22, Bucharest. Romania.