Zeitschrift: IABSE reports = Rapports AIPC = IVBH Berichte
Band: 66 (1992)

Artikel: Classical and Bayesian analysis of fatigue strength data
Autor: Arnold, Barry C. / Castillo, Enrique / Sarabia, José María
DOI: https://doi.org/10.5169/seals-50688

Classical and Bayesian Analyses of Fatigue Strength Data

Analyses classique et bayesienne appliquées aux données de résistance à la fatigue

Klassische und Bayes'sche Analysen von Ermüdungsfestigkeitsdaten

Barry C. ARNOLD
Professor
University of California
Riverside, CA, USA

Enrique CASTILLO
Professor
University of Cantabria
Santander, Spain

José María SARABIA
Assistant Professor
University of Cantabria
Santander, Spain

Barry C. Arnold graduated from Stanford University in 1965 with a Ph. D. in Statistics.

Enrique Castillo graduated from the Polytechnical University of Madrid in 1969. Awarded a Ph. D. in Civil Engineering from Northwestern University in 1972 and in Mathematics from Complutensis University Madrid in 1974.

José María Sarabia graduated from Valladolid University in 1986 with a Ph. D. in Mathematics from the University of Cantabria.

SUMMARY
Functional equations and basic engineering principles suggest that fatigue length of longitudinal elements may be modelled using a Weibull distribution with a virtual length function entering as a scale parameter. When applied to yarn data, a quadratic virtual length function is supported by both classical and Bayesian analyses. Related simplified models and models exhibiting asympotic independence are also investigated.

RÉSUMÉ
Les équations fonctionelles et les principes technologiques fondamentaux permettent d'envisager une modélisation de l'effet de la longueur sur la fatigue des élements longitudinaux à l'aide d'une distribution de Weibull, en introduisant une fonction de longueur virtuelle en tant que paramètre d'échelle. Dans le cas des données de fibres textiles une fonction virtuelle quadratique peut être admise pour la longueur tant par l'analyse classique que par l'analyse bayesienne. Des modèles simplifiés et des modèles qui mettent en évidence une indépendance asymptotique sont également étudiés.

ZUSAMMENFASSUNG
Funktionalgleichungen und grundlegende Ingenieurprinzipien deuten darauf hin, daß der Einfluß der Länge auf die Ermüdung von Längselementen unter Anwendung einer Weibull-Verteilung mit einer virtuellen Längenfunktion als Maßstabparameter modelliert werden kann. Auf Ergebnisse an Garnen angewandt, wird eine quadratische virtuelle Längenfunktion sowohl durch traditionelle als auch durch Bayes'sche Analyse unterstützt. Ähnliche vereinfachte Modelle und solche, die asymptotische Unabhängigkeit aufweisen, werden ebenfalls untersucht.

1. INTRODUCTION

Longer wires are weaker than short ones of the same diameter. That much has been "obvious" for centuries. How does this relationship between length and strength manifest itself? A standard experimental paradigm involves the use of samples of the material in question of equal diameters but varying lengths. The samples are subjected to repeated vibrational stress and failure times are observed. There will be considerable variability but, in general, the short samples will have longer survival times. Development of an appropriate stochastic model will be especially desirable. A important use of the model will be to predict the strength (as manifested by failure times) of elements of lengths different from those actually studied. Extrapolation will inevitably be involved. Typically only short samples can be studied but, in fact, much longer elements are of interest. We study one meter long samples of cable and try to predict how a bridge cable will behave. Such enthusiastic extrapolation is risky. It is however inevitable. Small errors in the fitted parameters have little effect on predictions for short samples but may result in enormous ranges of uncertainty when we extrapolate to longer samples. This cannot be avoided. All we can do is provide the engineer with the available information, sparse and uncertain though it may be. Basically our goal is to provide predictions and estimated reliabilities of predictions for long elements based on experiments using short elements. If the predictions are too crude, then it will be quite appropriate to call for new experiments, undoubtedly involving longer samples. To illustrate these ideas, we will reanalyze the Picciotto yarn data. Short lengths of yarn (less than or equal to a meter in length) were studied. Extrapolation to longer lengths is desirable. In particular, it is of interest to know whether an assumption of asymptotic independence is tenable; i.e. for long elements, if one is twice as long as another, is it twice as weak?

2. DEVELOPMENT OF THE MODEL

The survival function for an element of length x will be denoted by $\bar{F}(t,x)$. It represents the probability that an element of length x will survive at least t units of time (often measured in cycles). An accelerated failure model would be one in which the distribution of failure times depended on the length x only through a scale factor. That is

$$\bar{F}(t,x) = \bar{F}_0(h(x)t) \tag{2.1}$$

where F_0 is the distribution of failure times for an item of unit length and so, by convention, $h(1) = 1$. Castillo and Ruiz-Cobo [5] use an argument involving functional equations to arrive at (2.1) but many experimenters will be happy to accept such a model in which the scale of but not the shape of the survival distribution depends on element length.

Bogdanoff and Kozin [2] suggest a model of the form

$$\bar{F}(t,x) = [\bar{F}(t,y)]^{N(y,x)} \tag{2.2}$$

for some function $N(y,x)$. Castillo et al [4] show that this necessarily implies that the model is of the form

$$\bar{F}(t,x) = [\bar{F}_0(t)]^{q(x)} \tag{2.3}$$

for some base survival function $F_0(t)$ and arbitrary non-negative function $q(x)$. This is recognizable as a proportional hazards model in the sense of Cox [6].

Assuming that we find the accelerated risk paradigm, (2.1), and the proportional hazards paradigm, (2.3), to both be compelling we are forced to conclude that a Weibull model is appropriate with some function of length entering as a scale parameter. Specifically our model is of the form

$$\bar{F}(t,x) = [e^{-t^c}]^{q(x)} \tag{2.4}$$

where $q(x)$ is a non-negative function. Finally a parametrically parsimonious model might assume that $q(x)$ is a low degree polynomial.

3. INDEPENDENCE, ASYMPTOTIC INDEPENDENCE AND VIRTUAL LENGTH

The simplest model for fatigue strength corresponds to the choice $q(x) = mx$ in (2.3). In this model an element of length x behaves as if, when it divided into k subelements of length x/k, the subelements act independently and the full element survives if and only if all k subelements survive. This is the independence model. It undoubtedly only applies to long elements and then only approximately. Following Arnold, Castillo and Sarabia [1] we will say that the model exhibits asymptotic independence if

$$\lim_{x\to\infty}[q(\lambda x) - \lambda q(x)] = 0 \quad \forall\, \lambda > 1 \tag{3.1}$$

and exhibits strong asymptotic independence if

$$\lim_{x\to\infty} q(x)/mx = 1 \text{ for some } m > 0. \tag{3.2}$$

Both cases correspond to situations in which, for large x, $q(x)$ behaves like mx.

The function $q(x)/q(1)$ will, under independence, correspond to the length of the element (namely x). We will call this function the virtual length function. An element of the material of length x acts as if (under an independence assumption) it were of length $q(x)/q(1)$. A key restriction on $q(x)$ is that it be non-negative over the range of observed values of x. More importantly it should remain non-negative over the range of x values to which we wish to extrapolate.

4. THE PICCIOTTO DATA; LIKELIHOOD ANALYSIS

Yarn samples of length's 0.3(0.1)1.0 meters were studied by Picciotto [7]. A total of 797 observations were made; 99 or 100 for each value of x. The full data set is reported in [1] and [4] as well as in Picciotto [7]. The previous discussion suggests a model of the form

$$P(T > t|X = x) = \exp[-(\alpha + \beta x + \gamma x^2)t^\delta] \tag{4.1}$$

where a convenient normalization has been invoked so that T = (# of cycles to failure)/1000 and X corresponds to length measured in meters. The virtual length function is

$$q(x) = \alpha + \beta x + \gamma x^2. \tag{4.2}$$

Our observed levels of x cover the range 0.3 - 1.0. Our interest is in extrapolation to large values of x. Consequently it is reasonable to restrict α, β, γ in (4.2) so that $q(x)$ remains positive over the range $(0.3, \infty)$. A more stringent requirement that $q(x)$ remain positive over the half line $(0, \infty)$ would require $\alpha = 0$ and $\beta = 0$ so that $q(x)$ would assume the particularly simple form γx^2. This quadratic virtual life model is appealing in its simplicity but we must be wary, since this simplicity may be bought at a high price of reduced explanatory power in the region of interest for extrapolation. If, for some reason we were very interested in small, rather than large, values of x then of course a restriction that $q(x)$ remain positive for small values of x would be essential. For our present purposes, it is judged to not be essential. The model (4.1) may be fitted to the Picciotto data by standard maximum likelihood routines. Using a BMDP package the following estimated values of the four parameters were obtained.

parameter	estimate	asymptotic standard error	estimate/s.e.
α	-6.21	18.98	-0.33
β	-6.84	86.03	-0.08
γ	320.23	101.29	3.16
δ	1.97	0.05	37.99

The corresponding value of the log-likelihood is 1413.27. Note the negative values of α and β. If these are kept in the virtual length function then, for very small values of x (less than $0.3 - 1$, the observed range), the virtual length will be negative. Note also that the estimated value of δ namely 1.97 is remarkably close to 2, the value corresponding to a Rayleigh distribution.

Physical considerations do suggest that we set $\alpha = 0$ (so that $g(0) = 0$). The maximum likelihood estimates subject to this restriction are found to be

parameter	estimate	asymptotic standard error	estimate/s.e.
α	0		
β	-34.56	17.22	-2.01
γ	347.84	59.32	5.86
δ	1.97	0.05	38.26

The corresponding log-likelihood is 1413.21, only a tiny reduction from the value obtained for the four parameter model. Setting $\beta = 0$ (to guarantee $g(x) > 0$ on $(0, \infty)$) is a bit more costly. We find

parameter	estimate	asymptotic standard error	estimate/s.e.
α	0		
β	0		
γ	247.52	23.66	10.47
δ	1.91	0.04	45.40

with a corresponding log-likelihood of 1410.70. This is a barely significant reduction and the gain in parsimony and in the ability to extrapolate to small values of x may be judged to be worth the price. Over the observed range (0.3–1.0), the two fitted virtual length functions $-34.56x + 347.84x^2$ and $247.82x^2$ are very similar.

If we set $\alpha = 0$ and $\delta = 2$ and fit the resulting Rayleigh model, our maximum likelihood estimates of β and γ are given by

parameter	estimate	asymptotic standard error	estimate/s.e.
α	0		
β	-41.98	13.03	-3.22
γ	380.60	27.70	13.74
δ	2		

with a log-likelihood of 1413.04. This is remarkably close to the value 1413.27 obtained for the full four parameter model. Based on the available data, such a parsimonious Rayleigh model would be recommended. Further simplification by setting $\beta = 0$, though desirable in that it forces $g(x)$ to be positive over $(0, \infty)$, will be bought at a significant reduction in the log-likelihood (down to 1408.23). This extremely parsimonious model, namely

$$P(T > t|X = x) = \exp[-302.13(xt)^2],$$

merits reporting and would gain in stature if some physical interpretation could be found for the belief that $(xT)^2$ should be exponentially distributed. Further comment on this model will be found in Section 7.

5. BAYESIAN ANALYSIS

In practice it is to be expected that, not only would the engineer be able to argue in favor of a particular parametric model such as (4.1), but also he would have some insights into plausible values for the parameters in the model. If we denote the elicited prior for $(\alpha, \beta, \gamma, \delta)$ by $\varphi(\alpha, \beta, \gamma, \delta)$ (most likely of the form $\varphi_1(\alpha)\varphi_2(\beta)\varphi_3(\gamma)\varphi_4(\delta)$), then our posterior will be proportional to the following expression

$$h(\alpha, \beta, \gamma, \delta) = \varphi(\alpha, \beta, \gamma, \delta) \prod_{i=1}^{797} (\alpha + \beta x_i + \gamma x_i^2)\delta t_i^{\delta-1} \exp[-(\alpha + \beta x_i + \gamma x_i^2)t_i^\delta]. \tag{5.1}$$

Posterior means will serve as reasonable point estimates of the four parameters. Thus we would need to evaluate, for example

$$\tilde{\alpha} = \int_\Theta \alpha h(\alpha, \beta, \gamma, \delta) d\alpha d\beta d\gamma d\delta / \int_\Theta h(\alpha, \beta, \gamma, \delta) d\alpha d\beta d\gamma d\delta \tag{5.2}$$

where $\Theta = \{(\alpha, \beta, \gamma, \delta) : \delta > 0, \alpha + \beta x + \gamma x^2 > 0 \ \forall\ x \geq 0.3\}$. Such four dimensional numerical integrations may use up considerable amounts of computer time, which may be expensive. An alternative approximate approach involves iterative one dimensional averaging. Provided that the posterior is unimodal this will provide a reasonable approximation to the posterior means. In this approach admissible initial values $(\alpha_0, \beta_0, \gamma_0, \delta_0)\epsilon\Theta$ are chosen and are updated as follows

$$\alpha_1 = \int \alpha h(\alpha, \beta_0, \gamma_0, \delta_0) d\alpha / \int h(\alpha, \beta_0, \gamma_0, \delta_0) d\alpha$$

$$\beta_1 = \int \beta h(\alpha_1, \beta, \gamma_0, \delta_0) d\beta / \int h(\alpha_1, \beta, \gamma_0, \delta_0) d\beta$$

$$\gamma_1 = \int \gamma h(\alpha_1, \beta_1, \gamma, \delta_0) d\gamma / \int h(\alpha_1, \beta_1, \gamma, \delta_0) d\gamma$$

$$\delta_1 = \int \delta h(\alpha_1, \beta_1, \gamma_1, \delta) d\delta / \int h(\alpha_1, \beta_1, \gamma_1, \delta) d\delta \tag{5.3}$$

Then the process is repeated, now using $(\alpha_1, \beta_1, \gamma_1, \delta_1)$ as initial values. The integration in (5.3) is over the ranges of values for which the integrands are positive. The process is continued until stable values are obtained.

Implementation of a simplified version of the scheme (5.3) (in which we set $\alpha = 0$) with the Picciotto data, assuming improper uniform priors, yielded estimates of the form

$$\tilde{\beta} = -34.49$$

$$\tilde{\gamma} = 349.94$$

$$\tilde{\delta} = 1.973$$

As expected these are good approximations to the maximum likelihood estimates. More disparity would be encountered if a more precise prior were to be used.

6. WHAT SHOULD THE VIRTUAL LENGTH FUNCTION BE?

Data driven analysis suggested that a quadratic virtual length function is appropriate. For predictive purposes that may be adequate but, since many non-quadratic functions on the real line are well approximated by quadratic functions on the interval $(0.3, 1)$, extrapolation will be dangerous. Existence of a plausible physical explanation for quadratic virtual length would lessen the dangers of such projections. One possible explanation argues that failures occur at faults and that faults are distributed along the element according to a possibly non-homogenous Poisson process. An element with k faults will have a survival time function given by $[\bar{G}_0(t)]^k$. If the non-homogenous fault generating Poisson process has rate function $\lambda(x)$ then the expected number of faults in an element of length x will be $\int_0^x \lambda(y)dy$. This will be quadratic if λ is a linear function. Under such a scenario a Weibull model with a quadratic virtual length function might provide a good approximation to reality. The problem with this model is that our elements of length say $\frac{1}{4}$ were not made that length. They were cut from a longer piece. The fault generating process model is difficult to justify under such circumstances.

Quadratic virtual length has another negative feature. It manifestly fails to exhibit asymptotic independence. Yet asymptotic independence is surely a reasonable requirement for a model. A long element will fail if it fails in the first or second half of its length. These events should have equal probabilities and should be (roughly) independent. It doesn't take much arguing to convince yourself that at least asymptotic independence is appriopriate. Consequently the correct virtual length function will behave for small lengths in an approximately quadratic fashion and will be approximately linear for large x. In fact many researchers (see e.g. Castillo and Fernandez-Canteli [3]) argue that our main task is to determine a threshhold beyond which the linear virtual length model can be assumed to hold.

A three parameter virtual length function of the following form

$$g(x) = ax + b(e^{cx} - 1) \tag{6.1}$$

was considered in Arnold, Castillo and Sarabia [1]. They assumed a Rayleigh distribution so the survival model was of the form

$$P(T > t|X = x) = \exp\left[-[ax + b(e^{cx} - 1)]t^2\right]. \tag{6.2}$$

This virtual length function behaves roughly as a quadratic for $x < 1$ but, for large x, behaves like ax. For extrapolation purposes, particular interest will be focussed on the parameter a. If we apply this model to the Picciotto data, the likelihood surface exhibits a ridge. The likelihood can be made large by picking very small values of c together with appropriate large values of a and b. Three sets of parameter values which yield approximately equal values of the likelihood (and which are consequently essentialy equally plausible) are

a	b	c	log-likelihood
1011.5	1089.90	-1.00	1412.63
7859	79048	-0.10	1413.05
76290	7633183	-0.01	1413.07

The data based on lengths in the interval $(0.3, 1)$ support the plausibility of arbitrarily large values of a. Perhaps predictably, they are unable to assist us in determining the appropriate asymptotic slope.

7. THE SIMPLE MODEL

The simplest model with explanatory power corresponds to the choice $\alpha = 0, \beta = 0, \delta = 2$ in (4.1). For this model, as remarked in Section 4, $Y = (xT)^2$ will have an exponential distribution. The sample c.d.f. of the 797 Y_i's from the Picciotto data should look like $1 - e^{-\lambda y}$. Consequently a plot of $\log \bar{F}_n(y)$ (the log of the empirical survival function) vs y should be a line through the origin with negative slope $(= -\lambda)$. The actual situation for the Picciotto data is as shown in Figure 1.

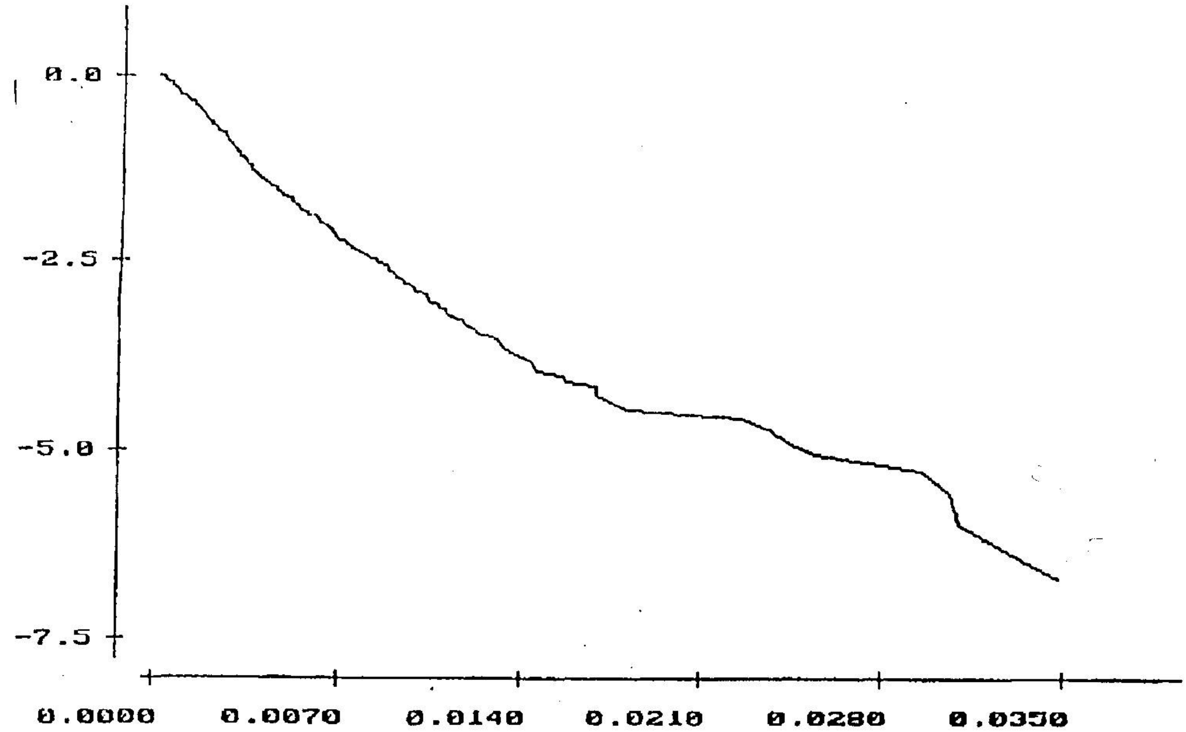

Figure 1: Plot of the logarithm of the empirical survival function $(\log \bar{F}_n(y))$ versus $y(= (xT)^2)$ for Picciotto data.

Elimination of the 8 largest Y_i's, regarded as potential outliers, improves the picture considerably. The empirical plot of the remaining 789 points, shown in Figure 2, might be judged to be acceptably linear. This may be advocated as a simplistic straw-man model for small values of the length x. We are still lacking a theoretical argument in favor of the quadratic virtual length which is implicit in such a model. We recognize that extrapolation to large lengths using such a model will be questionable since it will fly in the face of our belief of the plausibility of asymptotic independence.

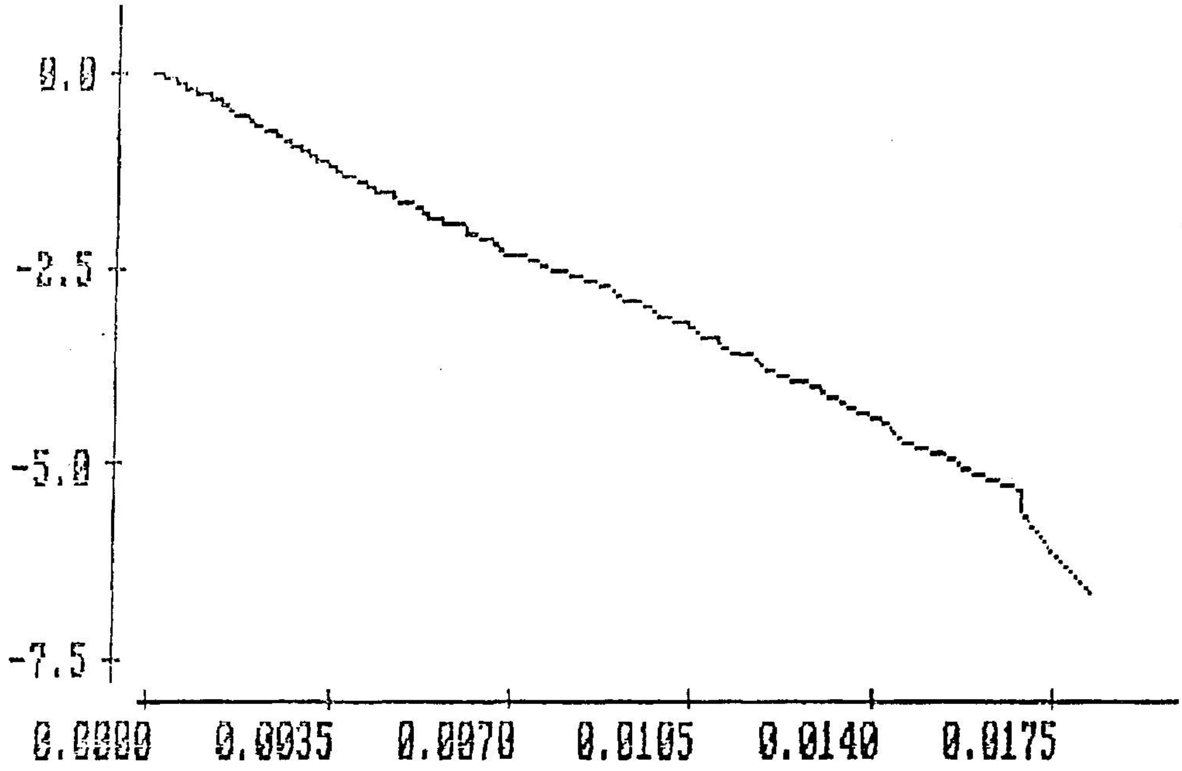

Figure 2: Plot of log $\bar{F}_n(y)$ versus y for the 789 Picciotto points remaining after deleting eight outliers.

8. REFERENCES

[1] Arnold, B.C., Castillo, E. and Sarabia, J.M. "Modelling the fatigue life of longitudinal elements", Technical Report, Department of Statistics, University of California, Riverside, 1992.

[2] Bogdanoff, J.L. and Kozin, F. "Effect of length on fatigue life of cables." *Journal of Engineering Mechanics*, ASCE, 1987, Vol. 113(6), pages 925-940.

[3] Castillo, E. and Fernandez-Canteli, A. "Statistical models for fatigue analysis of long elements." This proceedings, 1992.

[4] Castillo, E. and Fernandez-Canteli, A., Ruiz-Tolosa, J.R. andSarabia, J.M. "Statistical models for analysis of fatigue life of long elements." *Journal of Engineering Mechanics*, ASCE, 1990, Vol. 116(5), pages 1036-1049.

[5] Castillo, E. and Ruiz-Cobo, R. *Functional Equations and Modelling in Science and Engineering.* Marcel Dekker, New York, 1992.

[6] Cox, D.R. "Regression models and life tables." *Journal of Royal Statistical Society*, 1972, Vol 34, pp. 187-220.

[7] Piccioto, R. "Tensile fatigue characteristics of sized Polyester/viscose yarn and their effect on weaving performance." Thesis presented to North Carolina State University, at Raleigh, N.C. in partial fulfillment of the requirements for the degree of Master of Science, 1970.

Leere Seite
Blank page
Page vide