9th National Convention on Statistics
Edsa Shangri-La Hotel
October 4-5, 2004

Some Aspects of the Design and Analysis of Simple Accelerated Life Tests

by
Eduardo Cruz and Adolfo de Guzman

For additional information, please contact:

Author's Name:	Eduardo Cruz and Adolfo de Guzman
Designation:	Assistant Professor/Professor
Agency:	School of Statistics
	University of the Philippines Diliman
Address:	Quezon City
Telefax:	(02) 928-0881
Email:	edcruz2@up.edu.ph

Some Aspects of the Design and Analysis of Simple Accelerated Life Tests

by

Eduardo L. Cruz and Adolfo M. de Guzman[1]

Abstract

The main objective of accelerated life tests in this setting is the recovery of the distribution of a random variable which is difficult to observe at a certain level of a given factor. To solve this problem the inverse problem approach is utilized, that is, the variable is observe at a different level and the transfer function is used to recover the elusive random variable (life). The problem then is reduced to finding the transfer function. The Arrhenius model for accelerated life tests is studied in particular.

Keywords and phrases: Accelerated life test, U-Statistics, design of experiments, inverse problem.

I. Introduction

From the point of view of production and reliability engineers, accelerated life testing is an important aspect of product development, quality control and improvement. Accelerated life testing (ALT) is accomplished by applying increased stress on the product or product component. It is intended to produce data on strength and on lifetime for material components and systems to validate models. For an excellent but elementary exposition on this, the reader is referred to Nelson (1971).

Accelerated life testing is also very interesting from the point of view of theory: most lifetime data suffer from the censoring problem of statistics. Lifetime observations usually exceed the interest time of the observer or even his own lifetime. ALTs, however, are performed to destruction thereby eliminating censored observations. The analysis of censored or incomplete data requires a modification of the usual statistical methodology of "IID observations and their uncensored generalizations". Recent developments point to the counting process approach with the use of martingales as introduced by Aalen (1978) for analyzing survival or lifetime data.

The audience is referred to Anderson, Borgen, Gill and Keiding (1993) and Fleming and Harrington (1992). The first volume covers extensively

[1]E. L. Cruz is Assistant Professor and A. M. de Guzman is Professor of Statistics at the University of the Philippines at Diliman, Quezon City Philippines 1101.

both theory and and applications of the counting process approch to survival analysis while the second volume treats the analysis of clinical data via the martingale approach. Both books require some amount of mathematical as well as statistical sophistication. It will, therefore, take some more time for this approach to be a practical and widely accepted statistical technology.

A way out of this problem is provided by the ALT is some interesting cases. This is so because ALT, if it does result in complete data, usually reduces the number of incomplete observations. However, there is a price to pay. This paper is an installment. We hope others will make a contribution.

Section 2 of this paper describes a Generalized Arrhenius Lifetime Model for accelerated testing. Its properties are discussed in section 3 and ends by showing two theorems that indeed generalize both the Arrhenius model and the Cox proportional hazards model in the Generalized Arrhenius accelerated lifetime model. Section 4 then reviews U-Statistics and discussed its application the generalized accerated model. The final section acknowledges people who contributed to the ideas presented here and some references.

II. The Arrhenius-Based Model for Accelerated Tests

We first formulate the problem. Let X be a nonnegative random variable that represents lifetime in the normal condition, that is, without any stress applied to it, with unknown distribution F. Suppose stress is applied on X and the observations $Y_1, Y_2, \ldots, Y_n$ are made from F_s. It may be noted that this is an example of an "inverse problem" where stress S is a known operator. That is,

$$S : X \to Y$$

or

$$S : F \to F_s$$

in some space of random variables.

One approach to accelerated life test model is given by Cox and Oakes, 1984 and Barlow and Sheuer, 1971. Denote a failure time distribution function under a normal stress condition by $F_0(\cdot)$. The accelerated life time transformation is given in terms of $F(t; z)$ and $F_0(\cdot)$ by the relationship

$$F(t; z) = F_0[(t\psi(z, A)],$$

where $y(z, A)$ is a positive function connecting Time to Failure with a stress factor z; and A is a vector of unknown parameters; for $z = 0$, $y(z, A)$ is assumed equal to 1. This relationship is a scale transformation. It means that a change in stress does not result in a change in the shalpe of the distribution function but changes its scale only. This relationship can be written in terms of the acceleration function as

$$g(t) = \psi(z, A)t.$$

In other words, the relationship above is equivalent to the linear one with a time acceleration function.

We now propose a model that would enable us solve the above problem. Let ϑ be an nonempty set whose elements we will refer to as stress space. For example, in the Arrhenius model ϑ is the set of nonnegative reals representing temperature. The stochastic process $\{X_s : s \in \vartheta\}$ will be called a general lifetime model with stress space ϑ. This model says the random lifetime X_s depends on the stress S, where stress is a state or configuration of stress factors belonging to some known stress space ϑ. Information on the stress space should be available in the underlying field such as medicine for clinical trials, psychology for behavioral studies, chemistry for phenomena depending on clinical reactions. Two situations arise from this model: first, if S is fixed or controlled, this model is interpreted as the general accelerated life test model. If S is random, the Generalized Life Test (GLT) model describes the survival model with covariates. In this paper we consider only the first case when S is fixed. Our model is given by

$$X_s = A(s, \theta)X$$

where θ is a vector of parameters, $s \in \vartheta$ and $A(s, \theta)$ is monotone and continuous from the right. We will refer to this model as the Generalized Arrhenius Accelerated Life Testing (GAALT) model. The function $A(s, \theta)$ is called the Arrhenius function. It is easy to show that when the stress space ϑ is the set of nonnegative reals and $A(s, \theta) = exp(\beta/s)$, $s \in \vartheta$, (1) is equivalent to the Arrhenius model described in Nelson, 1971.

We now show the equivalence of our approach to the one given by Cox and Oakes. Let $X \sim F$, a distribution function and $Y = A(s, \theta)X$. We have,

$$
\begin{aligned}
P(Y \leq y) &= P(A(s,\theta)X \leq y) \\
&= P(X \leq \tfrac{y}{A(s,\theta)}) \\
&= F(\tfrac{y}{A(s,\theta)})
\end{aligned}
$$

Taking $A(s,\theta) = \frac{1}{\psi(s,\theta}$ gives $F(y,s) = F(\psi(s,\theta)y)$ which is the model shown in Cox and Oakes (1984).

III. Analytic Properties of the GAALT

We now derive some analytic properties of the generalized Arrhenius life testing model and state them as propositions.

Proposition 1. If X_s is a GAALT with Arrhenius function $A(s,\theta)$ and stress space S, then the mean function $M_{(s)}$ and the covariance function $\sigma_{(s,t)}$ are given by,

$$
\begin{aligned}
M_{(s)} &= A(s,\theta)E(X) \\
&= A(s,\theta)\mu, \quad s \in \vartheta \quad \text{and} \\
\sigma_{(s,t)} &= E\{[A(s,\theta)X - A(s,\theta)\mu][A(t,\theta)X - A(t,\theta)\mu]\} \\
&= A(s,\theta)A(t,\theta)\sigma^2.
\end{aligned}
$$

for some μ and $\sigma^2 > 0$.
In particular, the variance at any stress $s \in \vartheta$ is $\sigma_{(s)} = A(s,\theta)^2\sigma^2$. The proof of this is straightforward.

Proposition 2. If $X_s = A(s)X$, then $\log(X_s)$ has constant variance $\sigma^2_{\log}$ and mean $\mu_{\log} = \log[A(s) + \mu]$, for some $\mu > 0$.

The log transformation results in a constant variance but a changing mean. If $A(s) \in (0,1]$, the variance decreases with a decreasing A since $\log A(s) \leq 0$.

The next proposition will show the effect of log transformation in a GAALT model.

Proposition 3. If Y_s is lognormal with parameters μ_{log} and σ^2_{log}, then X is lognormal with mean $\mu_{\log} - \log A(s)$ and variance σ^2_{log}.

This simple theorem will enable us to make inferences about X which is not observed in the GAALT model via the Arrhenius function $A(s)$, when it is known.

The folowing lemma is useful in interpreting the log mean of Y and X.

Lemma 1. If Z is log symmetric random variable with log mean $\mu_{\log}$, then $\exp[E(\log(Z)] = \exp[\mu_{\log}] = \text{median}[Z]$.
Proof: We have that

$$\begin{aligned} P[\log Z \leq \mu_{\log}] &= 1/2 \quad \text{from symmetry. Also,} \\ P[\log Z \leq \mu_{\log}] &= P[Z \leq \exp[\mu_{\log}]. \text{ Hence,} \\ F(\exp[\mu_{\log})]) &= F[\text{median} Z]. \end{aligned}$$

where F is the disribution of Z. The result follows for absolutely continuous distribution functions F.

We next relate the mean $\mu_{\log}$ and the variance $\sigma^2_{\log}$ to the mean μ and variance σ^2 of the untransformed X. We need the next lemma.

Lemma 2. Let X be a nonnegative random variable with d.f F and a finite second moment $E(X^2)$. Then there exists positive numbers δ_F^2 and γ_F such that,

$$\begin{aligned} &\text{(i.)} \quad E[\log X] \leq \gamma_F \quad \text{and} \\ &\text{(ii.)} \quad Var[\log X] \geq \delta_F^2 + [\gamma_F - \mu^2{}_F]^2 \end{aligned} .$$

In fact,

$$\begin{aligned} \gamma_F &= \int_1^\infty x dF(x) \quad \text{and} \\ \delta_F^2 &= \int_0^{x=-\log(x)} x^2 dF(x) \end{aligned}$$

Here $\mu_F = E[X]$.

Proposition 4. In the GAALT model with Arrhenius function $A(s) \in (0, \infty)$ and stress space ϑ, the expectations and variance have the properties,

$$E[\log Y_s] \leq \log A(s) + \gamma_F \leq \gamma_F$$
$$Var[\log Y_s] \geq \delta_F^2 - \gamma_F,$$

where the random variable X has distribution F, satisties the usual regularity conditions and γ_F and δ_F^2 are defined in the previous lemma.
The next theorem states the equivalence of our approach (scaling) to the classical Arrhenius model where the Arrhenius function is given by,

$$A(s) = \exp[\frac{\beta}{s}], \qquad s \in (0, \infty).$$

Theorem 1. If the Arrhenius equation function A(s) is given by,

$$A(s) = \exp\left(\frac{\beta}{s}\right)$$

where $s \in (0, \infty)$, then X has a lognormal distribution with parameters $\mu_{\log}$ and $\sigma_{\log}^2$, then the GAALT model is equivalent to the Arrhenius model with Arrhenius equation,

$$\mu(s) = \alpha + \frac{\beta}{s}$$

where $\alpha = E[\log X]$ and $\beta > 0$ is constant.

Proof: We need to show that,

$$E[\log Y_s] = \alpha + \frac{\beta}{s}.$$

To show this we have,

$$\begin{aligned} E[\log Y_s] &= E[\log A(s) + \log X] \\ &= \tfrac{\beta}{s} + E[\log X] \\ &= \tfrac{\beta}{s} + \mu_{\log} \\ &= \alpha + \tfrac{\beta}{s}. \end{aligned}$$

In the next theorem we show the equivalence of the GAALT model and the Cox regression model or the proportional hazards model. The reader is refered to Miller (1981) for an exposition. This result shows that the dual of the GAALT model is the Cox proportional hazards in the counting process approach to survival analysis, the GAALT being the lifetime approach to survival analysis.

Theorem 2. In the GAALT model with Arrhenius function $A(s)$, and stress space ϑ which is a space of Euclidean r-space, the hazard function at any stress $s \in \vartheta$ is given by,

$$\begin{aligned} \lambda(x,s)A(s) &= \lambda_0(x) \quad \text{for all} x \in (0,\infty) \\ \text{if} \quad A(s) &= \exp[-\beta' s], \quad \text{we have,} \\ \lambda(x,s) &= \lambda_0(x)\exp[-\beta' s], \end{aligned}$$

which is the Cox proportional hazards model.

Proof:
Let $P(Y_s \leq y) = F_s(y)$ Then,

$$\begin{aligned} F_s(y) &= P(A(s)X \leq y) \\ &= P(X \leq y/A(s)) \\ &= F(y/A(s)) \end{aligned}$$

where F is the disribution function of X.
Also,

$$dF(y/A(s)) = f(y/A(s)/A(s),$$

where f is the density of X.

The hazard function at stress level s is given by,

$$\begin{aligned} \lambda(x,s) &= f_s(y)/[1 - F_s(y)] \\ &= \{[f(y/A(s))][1/A(s)]\}/[1 - F_s(y/A(s))] \\ &= \{[1/A(s)][f(y/A(s))]\}/[1 - F_s(y/A(s))] \\ &= a/A(s)\lambda_0(x), \qquad \text{putting} \quad x = yA(s). \end{aligned}$$

IV. U-Statistics for Accelerated Tests

Let **P** be a family of probability measures on an arbitrary measurable space. In the nonparametric setting, **P** is defined to be a large family of distributions subject only to mild restrictions like continuity or existence of moments. The basic theory of U-Statistics is due to Hoeffding (1948). Detailed expositions can be found in Randles and Wolfe (1979).

Definition 1. Let γ denote a real-valued function defined for $P \in \mathbf{P}$. We say that γ is an estimable parameter within **P**, if for some integer r there exists an unbiased estimator of γ based on i.i.d. random variables distributed according to P; that is, if there exists a real-valued measureable function $h(x_1, x_2, \ldots, x_r)$ such that

$$E_P(h(x_1, x_2, \ldots, x_r) = \gamma$$

for all $P \in \mathbf{P}$, where $X_1, X_2, \ldots, X_r$ are i.i.d with distribution P.

The function h is assumed to be a symmetric function of its arguments. The reason for this is if f is an unbiased estimator of γ, then the average of f applied to all permutations of the variables is still unbiased and symmetric. That is,

$$h(x_1, x_2, \ldots, x_r) = \frac{1}{n!} \sum_{\pi \in \Pi_r} f(x_1, x_2, \ldots, x_r),$$

Definition 2. For a real-valued function $h(x_1, x_2, \ldots, x_n)$ and for a sample, $X_1, \ldots, X_n$, of size $n \geq r$ from a distribution P, a U-Statistic with kernel h is defined as

$$U_n = U_n(h) = \frac{1}{\binom{n}{r}} \sum_{\binom{n}{r}} h(X_{i_1}, \ldots, X_{i_r})$$

where the summation is over the set of all $\binom{n}{r}$ combinations of r integers, $i_1 < i_2, \ldots, < i_r$ chosen from $(1, 2, \ldots, n)$. We have the following lemma which gives the expected value and variance of the U-Statistic in definition 2.

Lemma 3. The expected value and variance of U_n is given by,

$$E(U_n) = \gamma,$$

$$Var(U_n) = \frac{1}{\binom{n}{r}} \sum_{c=1}^{r} \binom{r}{c} \binom{n-r}{r-c} {\sigma_c}^2$$

where $\sigma_c = \mathrm{Cov}[h(X_1, \ldots, X_c], h(X_1, \ldots, X_r]$.
Here c is the number of integers common to $(i_1, \ldots, i_r) \in \binom{n}{r}$ and $(j_1, \ldots, i_r) \in \binom{n}{r}$.
If $\sigma_m^2 < \infty$, then $Var(U_n) \cong r^2 \sigma_1^2 / n$ for large n. This will be useful in deriving the asymptotic distribution of U_n. The proof is omitted.

Now, if **P** is large enough, then the order statistics form a complete sufficent statistics from $P \in \mathbf{P}$. And U_n, being symmetric in $(X_1, \ldots, X_n)$, is a function of the order statistics, is UMVUE. We do not deal further with this because our interest is in the asymptotic distribution of U_n which is presented below. This result is due to Hoeffding who first stated it in his 1948 paper.

Theorem 2. Let γ be an estimable parameter of degree r with symmetric kernel h and let U_n be the corresponding U-Statistic with Var(U_n)$< \infty$. Then

$$\sqrt{n}[U_n - \gamma] \xrightarrow{d} N(0, r^2\sigma_1^2)$$

provided $\sigma_1 > 0$.

IV.1 U-Statistics in the GAALT

Our GAALT model is given by

$$X(s) = A(s)X,$$

where $X \sim F$ (unknown).

In warranty analysis, the concern is in estimating the probability $P(X > x) = \gamma$ which represents the reliability of a product with lifetime X. The number x provides the warranty for the product, is given, s is fixed, A(s) is known, and Y is available random sample at stress level s. We construct the U-Statistic estimate of γ and compute for the mean and variance. Finally we verify the Central Limit Theorem for this statistic .

IV.2 U-Statistics for a Reliability Estimate

To estimate $P(X > x) = \gamma$ we conduct a life test at fixed stress level s where n products are subjected to stress s. Let $Y_1, Y_2, \ldots, Y_n$ be the lifetime observations of the products, then

$$\gamma = P[Y > y] = P[A(s)X > y]$$

Let Ψ =1 if $t > x$ and 0 otherwise. Then, $E[\Psi(Y)] = \gamma$ so that h(x) is a symmetric kernel of degree $r = 1$ for estimating γ.

We then have,

$$\begin{aligned} E(\Psi(Y) &= \gamma \\ h(y) &= \Psi(y) \end{aligned}$$

with degree $r = 1$. The U-Statistic is given by

$$\begin{aligned} U_n &= U_n(Y_1, \ldots, Y_n) \\ &= \tfrac{1}{n} \sum \Psi(Y_i) \\ &= \tfrac{1}{n}[\text{number of observations} > y.] \end{aligned}$$

The mean $E(U_n) = \gamma$ is the reliability of the device at time x. The variance is given by

$$Var(U_n) = \frac{1}{\binom{n}{1}} \left[\binom{1}{1} \binom{n-1}{1-1} \right] \sigma_1 = \frac{1}{n}(\gamma - \gamma^2)$$

since $\sigma_1 = E[h(X_1)h(X_1) - \gamma^2 = \gamma - \gamma^2 > 0$. We now have,

$$\sqrt{n}[U_n - \gamma] \sim AN[0, (\gamma - \gamma^2)]$$

since r=1. (Here, AN is short for Asymptotically Normal.) It now follows that,

$$\frac{(U_n - \gamma)}{\sqrt{\frac{U_n(1-U_n)}{n}}} \sim AN(0, 1) \text{ since}$$

$$Var[U_n) = \frac{(\gamma - \gamma^2)}{n} \to 0 \quad \text{as} \quad n \to \infty.$$

These results provide what are needed in computing confidence intervals and tests of hypothesis concerning γ.

V. Conclusion

In this paper the use of U-Statistics in life tests where the time transformation is known is proposed. The estimates of reliability and variance turn out be the usual estimators even under the nonparametric setting. For the reliability estimate, we have avoided the problem of censoring because the estimators require only that at a given time we count only the number of devise still functioning.

An important consideration is the choice of stress level. This is where the test of hypothesis $H_0 : A(s) = 1$ applies. Rejection of this hypothesis means we can proceed to the test or increase stress at a certain level. For the null hypothesis $H_0 : A(s) = 1/2$, rejection in favor of the alternative $H_0 : A(s) > 1/2$ indicates that we can start our accelerated test model at level s where $A(s) > 1/2$. For products where $A(s)$ in not known, the search for $A(s)$ is the subject of many papers.

VI. Acknowledgements

We wish to acknowledge the contribution of our colleages and students when the second author taught nonparametric statistics in the School of Statistics to this piece of research. Our thanks goes to Dr. Daniel Bonzo for suggesting the problem, and to Dr. Welfredo Patungan for going over the manuscript.

VII. Bibliography

Andersen, P.K., Borgan, O., Gill, R.D., Keiding, N. (1993). Statistical Model Based on Counting Processes. Springer-Verlag, New York.

Barlow, Richard and Scheuer, Ernest (1971). Estimation from Accelerated Life Tests. Technometrics, Vol.13, Number 1, 145-160.

Cox, D.R. and Oakes (1984). The Analysis of Survival Data. Chapman and Hall, London.

Gnedenko, B., Pavlov, I. and Ushakov, I. (1999). Statistical Reliability Engineering. Wiley, New York.

Hoeffding, W.,(1948). A class of distributions with asymptotic normal distributions. Ann. Math. Statist., Vol. 19, 293-325.

Miller, R.G. Jr. (1981). Survival Analysis. John Wiley and Sons, New York.

Fleming, T. and Harrington, D. (1992). Counting Processes and Survival Analysis. Wiley, New York.

Nelson, Wayne (1971). Analysis of Accelerated Life Tests Data - Part I: The Arrhenius model and Graphical Tools. IEEE Transactions of Electrical Insulator, Vol. E1 -6, No. 4. Dec 1971.