MODELS FOR DEPENDENT LIFELENGTHS INDUCED BY COMMON ENVIRONMENTS

BY NOZER D. SINGPURWALLA[1] AND MARK A. YOUNGREN

George Washington University and US Army Concepts Analysis Agency

Multivariate distributions for the lifelengths of the components of a system operating under a common environment, when the environment has a different effect on each component, and when the environment is dynamic, are derived. Modelling of the dynamic environment is by a gamma process.

1. Introduction. Multivariate distributions for the lifelengths of biological and engineering systems have been proposed by Freund (1961), Downton (1970), Marshall and Olkin (1967), Lindley and Singpurwalla (1986), and Lee and Gross (1990). In this paper we build upon the theme proposed by Lindley and Singpurwalla, and generate classes of multivariate distributions which may lead to improved assessments of system reliability.

As a motivating scenario, suppose that we have an m-component, parallel redundant system, and suppose that the lifelengths of these components are judged exponential with known scale parameters $\lambda_{10}, \lambda_{20}, \ldots, \lambda_{m0}$ when they are tested in a laboratory individually. The λ_{io}'s [or more generally, the $\lambda_{io}(t)$'s, if the lifelengths are judged to be other than exponential] will be referred to as the *baseline failure rates* of the m components. Suppose that the effect of the common operating environment—when assumed to be static over time—is to modulate each λ_{io} by a common factor η, where η is unknown and has distribution G, so that the reliabilities become $\exp\{-\int_0^t \eta\lambda_{io}(u)du\}$. Uncertainty about η induces dependence among the component lifelengths $T_1, \ldots, T_m$. The T_i's, $i = 1, \ldots, m$, have a multivariate distribution whose nature is prescribed by the form of G. When the operating environment is dynamic, η becomes a function of time t, say $\eta(t)$; we will refer to η, or more generally $\eta(t)$, $t \geq 0$, as the *environmental factor function*—henceforth EFF. It is important to bear in mind that the EFF is merely a parameter that has little, if any, physical meaning. It is introduced for convenience with the aim of capturing our opinion about the effects of the environment on the failure rate of each component.

[1]Research supported by Grant DAAL03-87-K-0056, the U.S. Army Research Office and Contract N00014-85-K-0202, Project NR 042-372, Office of Naval Research.

AMS 1980 subject classifications. Primary 60K10; secondary 62N05.

Key words and phrases. Multivariate life distributions, gamma process, failure rate, environmental effects.

When $m = 2$, $\eta(t) = \eta$ and G is a gamma distribution with scale β and shape α, T_1 and T_2 will have a bivariate Pareto distribution (Johnson and Kotz [1972] p. 285) with a joint survival function

$$(1) \quad \bar{F}(t_1,t_2) \stackrel{\text{def}}{=} P(T_1 > t_1, T_2 > t_2) = \left(\frac{\beta}{\lambda_{10}t_1 + \lambda_{20}t_2 + \beta}\right)^{\alpha}, \quad t_1, t_2 > 0.$$

The above distribution, which can be transformed to a bivariate logistic distribution, was motivated by Lindley and Singpurwalla; it can be shown to be a special case of the Dirichlet distribution. Currit and Singpurwalla (1988) compared the behavior of $\bar{F}(t_1,t_2)$ with $\exp(-(\lambda_{10}t_1+\lambda_{20}t_2))$, the survival function obtained under the assumption that T_1 and T_2 are independent and exponentially distributed with parameters λ_{10} and λ_{20}, respectively, and showed that the two could lead to drastically different results. The aim of this paper is to consider extensions of (1) along the several lines described below.

2. Multiple Environmental Factor Functions with Dependence. A natural way to expand upon the previous theme is to assume that each λ_{io} is modulated by η_i, $i = 1, \ldots, m$, and that the uncertainty about the η_i's is described by a meaningful multivariate distribution. Dependencies between the η_i's can be motivated when some factors which constitute the environment—such as temperature—may have an identical effect on all the components, whereas the other factors—such as humidity—may have different effects on the different components. A plausible model for describing dependencies among the η_i's is due to Cherian (1941) and David and Fix (1961)—henceforth C–D–F.

Let $m = 2$ and assume that $\eta_i = X_0 + X_i$, $i = 1, 2$, where the random quantity X_0 captures the contribution of the common factors on both the components, and X_i captures the contribution of the other factors on component i. In the C–D–F model, X_0, X_1 and X_2 are assumed to be independent, each having a gamma distribution with scale (shape) parameter $\beta_i(\alpha_i)$, $i = 0, 1, 2$, respectively. Clearly, η_1 and η_2 are dependent and have a joint density which may be easily derived (see Johnson and Kotz (1972), pp. 216–220).

It is easy to verify that under the above scenario,

$$(2) \quad \bar{F}_{\text{CDF}}(t_1,t_2) = \left(\frac{\beta_0}{\lambda_{10}t_1 + \lambda_{20}t_2 + \beta_0}\right)^{\alpha_0} \Pi_{i=1}^{2} \left(\frac{\beta_i}{\lambda_{10}t_i + \beta_i}\right)^{\alpha_i}, \quad t_1, t_2 > 0.$$

2.1. Inequalities for Survival Functions with Increasing Degrees of Dependence. The nature of the dependence between T_1 and T_2 depends on the dependence between η_1 and η_2. In the case of (1), $\eta_1 = \eta_2 = \eta$ and so the dependence between η_1 and η_2 is the strongest possible. The C–D–F case and the independent case are increasingly less dependent. To facilitate the construction of three pairs of random quantities (η_1, η_2), (η_1, η_2') and (η_1, η_2'') with decreasing degrees of dependence, four mutually independent random quantities X_0, X_0', X_1 and X_1' are introduced with

$X_0\ (X_1) \stackrel{d}{=} X_0'(X_1')$, where the notation "$X \stackrel{d}{=} Y$" indicates that X has the same distribution as Y. Let $\eta_1 = \eta_2 = X_0 + X_1$, $\eta_2' = X_0 + X_1'$ and $\eta_2'' = X_0' + X_1'$ and suppose that X_i and X_i' have a gamma distribution with shape α_i and scale β_i, $i = 0,1$. Clearly, $\eta_2 \stackrel{d}{=} \eta_2' \stackrel{d}{=} \eta_2''$ but the pairs (η_1, η_2), (η_1, η_2'), (η_1, η_2'') are increasingly less dependent.

It is now easy to verify that the pair (η_1, η_2) [where η_1 and η_2 are identical] will result in the bivariate survival function of the form given by (1); specifically

$$\bar{F}_{\mathrm{LS}}(t_1, t_2) = \Pi_{i=0}^{1} \left(\frac{\beta_i}{\beta_i + \lambda_{10} t_1 + \lambda_{20} t_2} \right)^{\alpha_i}, \quad t_1, t_2 \geq 0.$$

When $\alpha_1 = \alpha_2$, the pair (η_1, η_2') will lead to the bivariate survival function (2), which because of its derivation via the C–D–F distribution will be denoted $\bar{F}_{\mathrm{CDF}}(t_1, t_2)$. Finally, since η_1 and η_2'' are independent, the resulting survival function is

$$\bar{F}_I(t_1, t_2) = \Pi_{\ell=1}^{2} \left(\frac{\beta_0}{\beta_0 + \lambda_{\ell 0} t_\ell} \right)^{\alpha_0} \Pi_{m=1}^{2} \left(\frac{\beta_1}{\beta_1 + \lambda_{m0} t_m} \right)^{\alpha_1}, \quad t_1, t_2 \geq 0.$$

Let "$(X_1, Y_1) \stackrel{D}{>} (X_2, Y_2)$" denote the fact that the pair (X_1, Y_1) is more dependent than the pair (X_2, Y_2). Then, by construction, $(\eta_1, \eta_2) \stackrel{D}{>} (\eta_1, \eta_2') \stackrel{D}{>} (\eta_1, \eta_2'')$, and now it is easy to verify

THEOREM 2.1.

$$\bar{F}_I(t_1, t_2) < \bar{F}_{\mathrm{CDF}}(t_1, t_2) < \bar{F}_{\mathrm{LS}}(t_1, t_2), \text{ for } t_1, t_2 > 0.$$

Thus, for any fixed t_1, $t_2 > 0$, the bivariate survival function of 2 component parallel redundant systems increases as the degree of dependence between their EFF's increases. The inequality generalizes for the case of m components. When we set $t_1 = t_2 = \ldots = t_m$, we obtain inequalities for the system reliability function of series systems.

3. Dependencies Induced By Dynamic Environments. The material in the previous two sections assumed that the EFF is constant over time, so that $\eta_i(t) = \eta_i$, $i = 1, \ldots, m$. This assumption is not meaningful when the environment is dynamic as is often the case. As a starting scenario, suppose that $\eta_i(t) = \eta(t)$, $t \geq 0$ and $i = 1, \ldots, m$, and suppose that our uncertainty about $\eta(t)$ is described by a continuous time stochastic process, called the *gamma process*. The gamma process for the EFF produces useful results, and can be motivated as the limit of a piecewise constant EFF with independent gamma distributed innovations.

3.1. Motivating the Gamma Process. Suppose that $\eta(t)$ is a piecewise constant right continuous function over specified time intervals $[t_j, t_{j+1})$, $j = 0, 1, \ldots$, where $t_0 \equiv 0$. Specifically, let $\eta(t) = \eta_j$, $t \in [t_j, t_{j+1})$, with the η_j's unknown. Suppose that the environment is composed of a known number of at most $s + 1$ distinct stresses, each having the same effect on all the m components, with the k–th stress contributing an innovation C_k to η_j. The parameter η_j changes to η_{j+1} when one or more of the innovations C_k appears or disappears. Suppose further that the effects of the innovations C_k are additive, so that $\eta_j = \sum_{k=0}^{s} I_k(t_j)C_k$, $j = 0, 1, \ldots$, where $I_k(t_j) = 1$, if the k–th stress is present in $[t_j, t_{j+1})$, and is 0 otherwise. The innovations C_k and the variables $I_k(t_j)$ are assumed to be mutually independent for all $j = 0, 1, \ldots$ and all $k = 0, 1, \ldots, s$. If N_j, the number of stresses during $[t_j, t_{j+1})$ is known, but their identities are unknown, and if each C_k is assumed to have a gamma distribution with parameters α and β, then the η_j's are independent gamma distributed variables with parameter $N_j\alpha$ and β. It can now be shown [cf. Youngren (1988), p. 54], that in the limit, as $\Delta t_j \stackrel{\text{def}}{=} (t_{j+1} - t_j) \to 0$, the cumulative failure rate of the i-th component at time t, $0 \le t_n < t \le t_{n+1}$, is a gamma process. We denote the cumulative failure rate of the i-th component as

$$\Lambda_i(t) = \lambda_{io}\left[\sum_{j=0}^{n-1} \eta_j(t_{j+1} - t_j) + \eta_n(t - t_n)\right];$$

recall that λ_{io} is the baseline failure rate of the i-th component.

Instead of assuming that the η_j's are independent as is done above, suppose that the η_j's have a time dependent structure as follows. Let

$$\begin{aligned} I_k(t_j) &= 0,\ 0 \le j < k, \quad j = 0, 1, \ldots, j; \\ &= 1,\ j \ge k, \qquad\ k = 0, 1, \ldots, s; \text{ then} \end{aligned}$$

$\eta_j = \sum_{k=0}^{j} C_k$, and if one's uncertainty about the C_k's is described via a gamma distribution with parameters α_j and β, then here again it can be shown [cf. Youngren (1988), p. 60], that when $\Delta t_j \to 0$, $\eta(t)$ is a gamma process for any $t \ge 0$.

3.2. Preliminaries on Gamma Processes. The gamma process is nonnegative, nondecreasing in time and possesses independent increments. It has been studied by Ferguson and Klass (1972), Çinlar (1980), and Dykstra and Laud (1981); the use of gamma processes in survival analysis is primarily due to Ferguson (1973), Ferguson and Phadia (1979), and Kalbfleisch (1978).

DEFINITION 3.1. Let $\alpha(t)$ be a nondecreasing left-continuous real valued function on $[0, \infty)$ with $\alpha(0) = 0$, and let $\beta \in (0, \infty)$. A stochastic process $(Y(t),\ t \ge 0)$ is said to be a *gamma process* with parameters $\alpha(t)$ and β, denoted "$Y(t) \in G_{pr}(\alpha(t), \beta)$", if:

1. $Y(0) = 0$

2. $Y(t)$ has independent increments, and

3. $Y(t) - Y(s) \sim \gamma(\alpha(t) - \alpha(s),\ \frac{1}{\beta})$ for any $0 \le s \le t$.

Dykstra and Laud (1981) extend the gamma process to include a time-varying scale parameter $\beta(t)$.

DEFINITION 3.2. Let $\beta(t)$, $t \geq 0$ be a positive right-continuous real valued function, and let $Y(t) \in G_{pr}(\alpha(t), 1)$. The process $Z(t) \stackrel{\text{def}}{=} \int_0^t \beta(s)\, dY(s)$ is an *extended gamma process* denoted "$Z(t) \in G_{pr}(\alpha(t), \beta(t))$."

Note that the gamma process is a special case of the extended gamma process, where $\beta(t) = \beta\ \forall t$.

Dykstra and Laud (1981) give the following properties of the extended gamma process. Let $Z(t) \in G_{pr}(\alpha(t), \beta(t))$. Then

$$\begin{aligned} E[Z(t)] &= \int_0^t \beta(u)\, d\alpha(u), \quad \operatorname{Var}[Z(t)] = \int_0^t \beta^2(u)\, d\alpha(u), \text{ and} \\ G^*_{Z(t)}(s) &= \exp\left[-\int_0^t \log(1 + s\beta(u))\, d\alpha(u)\right], \ s > 0, \end{aligned}$$

where $G^*_{Z(t)}$ is the Laplace Stieltjes transform (LST) of the distribution of $Z(t)$.

3.3. Modelling the EFF as a Gamma Process. Suppose that $\eta(t)$ is described by a gamma process with parameters $(\alpha(t), \frac{1}{\beta})$. If the baseline failure rate is a continuous, positive, real valued function of time, then the following theorem is used to derive the bivariate and marginal survival functions.

THEOREM. *Let $\eta(t) \in G_{pr}(\alpha(t), \frac{1}{\beta})$, let $\lambda_0(t)$ be a known continuous positive real valued function and let $\Lambda(t) = \int_0^t \lambda_0(u)\, \eta(u)\, du$. Then the univariate survival function is*

$$\bar{F}(t) = \exp\left\{-\int_0^t \log[1 + \frac{1}{\beta}\int_u^t \lambda_0(s)\, ds] d\alpha(u)\right\}.$$

The proof of this theorem is based on Dykstra and Laud (1981), and is given by Youngren (1988).

The bivariate survival function for $0 \leq t_1 \leq t_2$ follows directly from the above theorem. Specifically, $\bar{F}(t_1, t_2) =$

$$\begin{aligned} & \exp\left\{-\int_0^{t_1} \log[1 + \frac{1}{\beta}\int_u^{t_1} (\lambda_{10}(s) + \lambda_{20}(s))\, ds] d\alpha(u)\right\} \\ \times\ & \exp\left\{-\int_{t_1}^{t_2} \log[1 + \frac{1}{\beta}\int_u^{t_2} \lambda_{20}(s)\, ds] d\alpha(u)\right\}. \end{aligned}$$

We can choose plausible functional forms for $\alpha(t)$ and $\lambda_{io}(t)$, suggested by the physical model of the environment, that enable us to obtain closed form solutions for the survival functions. If we use our time-dependent model of Section 3.1, wherein $\eta(t) = \sum_{k=0}^{j} C_k$, for $t \in [t_j, t_{j+1})$, then for $s \in [t_\ell, t_{\ell+1})$, $\eta(t) - \eta(s) = \sum_{k=\ell+1}^{j} C_k$ is distributed as gamma with a shape parameter that depends on the length of the interval $(t - s)$. This leads, in the limit as $\Delta t \to 0$, to a gamma process with a linear shape function $\alpha(t)$, say $\alpha(t) = \alpha_1 t$, for some $\alpha_1 > 0$, $t \geq 0$.

Assume that $\eta(t) \in G_{pr}(\alpha_1 t, \frac{1}{\beta})$, which implies that the component failure rate is $\lambda_i(t) = \lambda_{io}\eta(t) \in G_{pr}(\alpha_1 t, \frac{\lambda_{io}}{\beta})$. For convenience let $\beta_{11} \overset{\text{def}}{=} \frac{\lambda_{1o}}{\beta}$ and $\beta_{21} \overset{\text{def}}{=} \frac{\lambda_{2o}}{\beta}$; then the bivariate survival function for $0 \leq t_1 \leq t_2$ is

$$\begin{aligned} \bar{F}(t_1, t_2) &= (1 + (\beta_{11} + \beta_{21})t_1)^{\frac{\alpha_1}{(\beta_{11}+\beta_{21})}} (1 + \beta_{21}(t_2 - t_1))^{\frac{\alpha_1}{\beta_{21}}} \\ &\times \left[\frac{1 + \beta_{21}(t_2 - t_1)}{1 + (\beta_{11} + \beta_{21})t_1}\right]^{\alpha_1 t_1} \cdot \left[\frac{e}{1 + \beta_{21}(t_2 - t_1)}\right]^{\alpha_1 t_2}, \text{ with} \end{aligned}$$

marginal failure rate functions $r_i(t_i) = \alpha_1 \log[1 + \beta_{i1} t_i]$, $i = 1, 2$.

It is interesting to note that the marginal distributions can also be obtained using an extended gamma process for $\lambda_i(t)$ with shape parameter $\alpha(t) = \alpha_1 t$, and scale parameter $\frac{\lambda_{io}(t)}{\beta} = \beta_{i1} t$.

REFERENCES

CHERIAN, K.C. (1941). A bivariate correlated gamma type distribution function. *J. Ind. Math. Soc.* **5** 133–144.

ÇINLAR, E. (1980). On a generalization of gamma processes. *J. Appl. Prob.* **17** 467–480.

CURRIT, A. and SINGPURWALLA, N.D. (1988). On the reliability function of a system of components sharing a common environment. *J. Appl. Prob.* **25**, No. 4, 763–771.

DAVID, F.N. and FIX, E. (1961). Rank correlation and regressions in a non-normal surface. In *Fourth Berkeley Symposium on Mathematical Statistics and Probability* (J. Neyman, ed.), University of California Press, **1** 177–197.

DOWNTON, F. (1970). Bivariate exponential distributions in reliability theory. *J. Roy. Statist. Soc.* B **32** 408–417.

DYKSTRA, R.L. and LAUD, P. (1981). A Bayesian nonparametric approach to reliability. *Ann. Statist.* **9**, No. 2, 356–367.

FERGUSON, T.S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1**, No. 2, 209–230.

FERGUSON, T.S. and KLASS, M.J. (1972). A representation of independent increment processes without Gaussian components. *Ann. Math. Statist.* **43**, No. 5, 1634–1643.

FERGUSON, T.S. and PHADIA, E.G. (1979). Bayesian nonparametric estimation based on censored data. *Ann. Statist.* **7**, No. 1, 163–186.

FREUND, J.E. (1961). A bivariate extension of the exponential distribution. *J. Amer. Statist. Assoc.* **56** 971–977.

JOHNSON, N.L. and KOTZ, S. (1972). *Distributions in Statistics: Continuous Multivariate Distributions.* John Wiley and Sons, Inc., New York.

KALBFLEISCH, J.D. (1978). Nonparametric Bayesian analysis of survival time data. *J. Roy. Statist. Soc.* B **40**, No. 2, 214–221.

LEE, M.-L.T. and GROSS, A.J. (1990). Lifetime distributions under unknown environment. *J. Statist. Plan. Infer.*, to appear.

LINDLEY, D.V. and SINGPURWALLA, N.D. (1986). Multivariate distributions for the lifelengths of components of a system sharing a common environment. *J. Appl. Prob.* **23**, No. 2, 418–431.

MARSHALL, A.W. and OLKIN, I. (1967). A multivariate exponential distribution. *J. Amer. Statist. Assoc.* **62** 30–44.

YOUNGREN, M.A. (1988). Dependent lifelengths induced by dynamic environments. D.Sc. dissertation, George Washington University.

DEPARTMENT OF OPERATIONS RESEARCH
GEORGE WASHINGTON UNIVERSITY
WASHINGTON, DC 20052

U.S. ARMY CONCEPTS ANALYSIS AGENCY
8120 WOODMONT AVE.
BETHESDA, MD 20814–2797