Goodness-of-Fit of the Distribution of Time-to-First-Occurrence in Recurrent Event Models

MA. ZENIA N. AGUSTIN (zagusti@siue.edu)
Department of Mathematics and Statistics, Southern Illinois University at Edwardsville, Edwardsville, IL 62026

EDSEL A. PEÑA (pena@stat.sc.edu) *
Department of Statistics, University of South Carolina, Columbia, SC 29208

Abstract. Imperfect repair models are a class of stochastic models that deal with recurrent phenomena. This article focuses on the Block, Borges, and Savits (1985) age-dependent minimal repair model (the BBS model) in which a system that fails at time t undergoes one of two types of repair: with probability $p(t)$, a perfect repair is performed, or with probability $1-p(t)$, a minimal repair is performed. The goodness-of-fit problem of interest concerns the initial distribution of the failure ages. In particular, interest is on testing the null hypothesis that the hazard rate function of the time-to-first event occurrence, $\lambda(\cdot)$, is equal to a prespecified hazard rate function $\lambda_0(\cdot)$. This paper extends the class of hazard-based smooth goodness-of-fit tests introduced in Peña (1998a) to the case where data accrual is from a BBS model. The goodness-of-fit tests are score tests derived by reformulating Neyman's idea of smooth tests in terms of hazard functions. Omnibus as well as directional tests are developed and simulation results are presented to illustrate the sensitivities of the proposed tests for certain types of alternatives.

Keywords: Counting process; directional test; imperfect and perfect repairs; minimal repair; nonhomogeneous Poisson process; omnibus test; repairable system; score test; smooth goodness-of-fit.

1. Introduction

Stochastic models for recurrent phenomena are of interest in a variety of fields of study such as in biomedicine, public health, engineering, reliability, economics, actuarial science, and sociology. For concrete examples, see for instance Keiding, Andersen and Fledelius (1998), Lawless (1998), Oakes (1998), and Prentice, Williams and Peterson (1981). The development of statistical inference procedures appropriate for analyzing such data is therefore of prime importance. In the reliability and engineering settings, recurrent events translate to failures or repairs in repairable systems. Various models have been proposed for repairable systems, cf., Brown and Proschan (1983), Ascher and Feingold (1984), Block, Borges and Savits (1985), Crowder, Kimber,

* E. Peña's research is partially supported by NIH/NIGMS Grant 1 R01 GM56182.

Smith and Sweeting (1991), and Dorado, Hollander and Sethuraman (1997). This paper focuses primarily on the age-dependent minimal repair model introduced by Block, Borges and Savits (1985), hereon referred to as the BBS model. In the BBS model, a component or system whose distribution function of the time-to-first-failure is F is put on test at time 0. If the component fails at time t, then with probability $p(t)$, a perfect repair is performed; otherwise, with probability $q(t) = 1 - p(t)$, a minimal or imperfect repair is performed. A perfect repair restores the component to the good-as-new state, while a minimal repair returns the component to a working state in such a way that the time until its next failure is stochastically equivalent to that of a working component of the same age. To formalize the BBS model, consider a sequence of failure ages $\{W_0 \equiv 0, W_1, W_2, \ldots\}$ obtained under a model of minimal repair with W_1 having absolutely continuous distribution F with associated density function f and hazard rate function $\lambda = f/\bar{F}$, where $\bar{F}$ is the survival function. The sequence $(W_k)_{k=1}^{\infty}$ is a Markov process with the conditional survival function of W_k, given $W_0, W_1, \ldots, W_{k-1}$, being $\bar{F}(t|W_{k-1}) = \bar{F}(t)/\bar{F}(W_{k-1}),\ t \geq W_{k-1},\ k \geq 1$. Following Hollander, Presnell and Sethuraman (1992), to determine the mode of repair at each failure time, independent Uniform$[0,1]$ random variables $\{U_1, U_2, \ldots\}$, which are independent of $(W_k)_{k=1}^{\infty}$, are generated. If $U_k < p(W_k)$, then a perfect repair is performed at W_k, otherwise a minimal repair is instituted. Letting $\nu = \inf\{k : U_k < p(W_k)\}$, where $\inf \emptyset = \infty$, the sequence $(W_1, \ldots, W_\nu)$ is referred to as an epoch of a BBS model. From Block, et al. (1985), W_ν is almost surely finite provided that $\int_0^\infty p(t)\lambda(t)\mathrm{d}t = \infty$. Since a perfect repair restores a component to the good-as-new state, it suffices to observe a component only up to the time of its first perfect repair. Hence, the statistical inference procedures presented in this paper are based on epochs of independent components.

Aside from its utility in reliability modeling, the BBS model also subsumes other models used in biomedical, engineering, actuarial, social science, and economic settings. For instance, by setting $p(t) = 0$, the nonhomogeneous Poisson process (NHPP) model is obtained, which is a common model for non-life insurance claims, periods of recession or stock market crashes in economics, recurrence of tumors in the medical arena, and incidence of domestic abuse in a sociological setting. The much-studied independent and identically distributed (i.i.d.) model on the other hand is obtained from the BBS model by letting $p(t) = 1$.

The literature dealing with minimal repair models have been mostly confined to examining their probabilistic properties, cf., Brown and Proschan (1983), Block, et al. (1985), Kijima (1989), and Arjas and Norros (1989). Only a few papers, such as Whitaker and Samaniego

(1989), Hollander, et al. (1992), and Presnell, Hollander and Sethuraman (1994), have explored inference procedures for the BBS model. Whitaker and Samaniego (1989) dealt with the estimation of the reliability of systems subject to imperfect repair, while Hollander, et al. (1992) and Presnell, et al. (1994) proposed nonparametric tests for the minimal repair assumption. Our main interest in the present paper is not on validating the minimal repair assumption but on goodness-of-fit tests concerning the initial distribution of the failure ages assuming that data accrual follows the BBS model. Determination of the correct initial distribution of the failure ages is important if one is interested in prediction or development of optimal maintenance policies. Our primary interest is testing the null hypothesis $H_0 : \lambda(\cdot) = \lambda_0(\cdot)$, where $\lambda_0(\cdot)$ is a completely specified hazard function. A separate paper will deal with the composite hypothesis problem of testing the null hypothesis $H_0 : \lambda(\cdot) \in \mathcal{C} \equiv \{\lambda_0(\cdot;\xi) : \xi \in \Xi\}$.

The goodness-of-fit procedures proposed in this paper extend the class of hazard-based smooth goodness-of-fit tests in Peña (1998a) to recurrent events under the no-covariate setting. The proposed smooth goodness-of-fit tests are score tests, which were shown to be powerful against a wide range of alternatives, cf., see Rayner and Best (1989) for the density-based or Neyman-based formulation, and Peña (1998ab) for the hazard-based formulation. These papers, however, dealt with the case where the data involves only the time to first failure, that is, for $p(t) = 1$. To our knowledge, the present work is the first one dealing with smooth goodness-of-fit tests for recurrent data.

2. The Goodness-of-Fit Test

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be the basic measurable space on which the random entities are defined. Let $\mathcal{T}$ be an index set of times which may be $[0,\tau)$ or $[0,\tau]$, where $\tau \le \infty$ is known. Consider observing n independent BBS processes so the observables are $\{W_{jk} : 1 \le j \le n; 1 \le k \le \nu_j\}$. To facilitate the development of the theory, we adopt the stochastic process formulation of the BBS model as set forth in Hollander, et al. (1992). Define the multivariate counting process $\mathbf{N}^* = \{(N_1^*(t), \ldots, N_n^*(t)) : t \in \mathcal{T}\}$ by $N_j^*(t) = \sum_{k=1}^{\infty} I\{W_{jk} \le t\}$, $j = 1, \ldots, n$, and the filtration $\mathbf{F}^* = \{\mathcal{F}_t^* : t \in \mathcal{T}\}$ by $\mathcal{F}_t^* = \mathcal{F}_0 \vee \bigvee_{j=1}^{n} \mathcal{F}_{jt}^*$, where $\mathcal{F}_{jt}^* = \sigma\left\{\{N_j^*(s) : s \le t\} \cup \{U_{jk} : k \ge 1\}\right\}$, and with $\mathcal{F}_0$ containing all null sets of $\mathcal{F}$. The multivariate counting process of interest is $\mathbf{N} = \{(N_1(t), \ldots, N_n(t)) : t \in \mathcal{T}\}$ with

$$N_j(t) = N_j^*(t \wedge W_{j\nu_j}), \quad j = 1, \ldots, n,$$

and the corresponding observable filtration $\mathbf{F} = \{\mathcal{F}_t : t \in \mathcal{T}\}$ is given by $\mathcal{F}_t = \bigvee_{j=1}^{n} \mathcal{F}^*_{j(t \wedge W_{j\nu_j})}$. The $\mathbf{F}$ compensator of $\mathbf{N}$ is $\mathbf{A} = \{(A_1(t), \ldots, A_n(t)) : t \in \mathcal{T}\}$ with

$$A_j(t) = \int_0^t Y_j(s)\lambda(s)\,\mathrm{d}s, \quad j = 1, \ldots, n,$$

where $Y_j(s) = I\{W_{j\nu_j} \geq s\}$ and $\lambda(\cdot)$ is some unknown baseline hazard function.

Recall that the main problem of interest involves testing the null hypothesis $H_0 : \lambda(\cdot) = \lambda_0(\cdot)$, where $\lambda_0(\cdot)$ is a completely specified hazard function. The main idea behind smooth goodness-of-fit tests is the embedding of the hypothesized hazard rate $\lambda_0(\cdot)$ into a larger parametric family of hazard rate functions. This larger family is obtained by smoothly transforming $\lambda_0(\cdot)$. To this end, let us define the family of order k smooth alternatives via

$$\mathcal{A}_k = \{\lambda_k(\cdot;\boldsymbol{\theta}) = \lambda_0(\cdot)\exp[\boldsymbol{\theta}'\boldsymbol{\Psi}(\cdot)] : \boldsymbol{\theta} \in \mathbb{R}^k\}, \tag{2.1}$$

where k is some fixed positive integer, and $\boldsymbol{\Psi}(\cdot)$ is a $k \times 1$ vector of locally bounded predictable processes. Note that by setting $\boldsymbol{\theta} = \mathbf{0}$ in (2.1), we recover the hypothesized hazard rate function. Hence, the null hypothesis $H_0 : \lambda(\cdot) = \lambda_0(\cdot)$ can be restated as $H_0^* : \boldsymbol{\theta} = \mathbf{0}$. In order to derive the score test for this hypothesis, we first obtain the score process associated with $\boldsymbol{\theta}$.

Under the model in (2.1), the compensator of $\mathbf{N}(\cdot)$ is $\mathbf{A}(\cdot;\boldsymbol{\theta}) = (A_1(\cdot;\boldsymbol{\theta}), \ldots, A_n(\cdot;\boldsymbol{\theta}))$, where $A_j(\cdot;\boldsymbol{\theta}) = \int_0^{\cdot} Y_j(s)\lambda_0(s)\exp[\boldsymbol{\theta}'\boldsymbol{\Psi}(s)]\,\mathrm{d}s$. It is then straightforward to see that the score process associated with $\boldsymbol{\theta}$ is

$$\boldsymbol{U}_{\boldsymbol{\theta}}(t;\boldsymbol{\theta}) = \sum_{j=1}^{n} \int_0^t \boldsymbol{\Psi}(s)\;\mathrm{d}M_j(s;\boldsymbol{\theta}),$$

where $M_j(s;\boldsymbol{\theta}) = N_j(s) - A_j(s;\boldsymbol{\theta}),\ j = 1, 2, \ldots, n$. In order to come up with the appropriate test procedure, one needs to obtain the distribution of $\boldsymbol{U}_{\boldsymbol{\theta}}(t;\boldsymbol{\theta})$ under the null hypothesis. The following regularity conditions are needed.

(I) $\int_0^{\tau} \lambda_0(s)\;\mathrm{d}s < \infty$.

(II) There exists a $k \times k$ matrix function $\boldsymbol{D}$ such that as $n \to \infty$,

$$\sup_{t \in \mathcal{T}} \left\| \frac{1}{n} \sum_{j=1}^{n} \boldsymbol{\Psi}(t)\boldsymbol{\Psi}(t)' Y_j(t) - \boldsymbol{D}(t) \right\| \xrightarrow{\text{pr}} 0.$$

(III) The matrix $\boldsymbol{\Sigma}(\tau) = \int_0^{\tau} \boldsymbol{D}(t)\lambda_0(t)\;\mathrm{d}t$ is positive definite.

(IV) For each $\epsilon > 0$, $\ell = 1, \ldots, k$, and for every $t \in \mathcal{T}$,

$$\frac{1}{n}\sum_{j=1}^{n}\int_0^t \psi_\ell(s)^2 I\{|\psi_\ell(s)| \geq \sqrt{n}\epsilon\} Y_j(s)\lambda_0(s)\ \mathrm{d}s \ \stackrel{\text{pr}}{\longrightarrow}\ 0.$$

We now present the asymptotic result which anchors the test procedure.

Theorem 2.1: *Under the BBS model, if conditions (I) – (IV) hold and H_0^* is true, then as $n \to \infty$, $\frac{1}{\sqrt{n}}\boldsymbol{U}_{\boldsymbol{\theta}}(\tau;\mathbf{0})$ converges in distribution to a zero-mean normal variable with covariance matrix $\boldsymbol{\Sigma}(\tau)$.*

Proof : Under H_0^*, $\mathbf{M}(\cdot) = (M_1(\cdot;\mathbf{0}), \ldots, M_n(\cdot;\mathbf{0}))$ is a vector of local square integrable martingales with quadratic variation process

$$\frac{1}{n}\sum_{j=1}^{n}\int_0^{\cdot} \boldsymbol{\Psi}(s)\boldsymbol{\Psi}(s)' Y_j(s)\lambda_0(s)\ \mathrm{d}s. \tag{2.2}$$

By conditions (I) and (II), (2.2) converges in probability to $\boldsymbol{\Sigma}(\cdot)$ as $n \to \infty$. On the other hand, the Lindeberg-type condition for $\frac{1}{\sqrt{n}}\boldsymbol{U}_{\boldsymbol{\theta}}(\cdot;\mathbf{0})$ is guaranteed by condition (IV). Hence, by invoking Rebolledo's Martingale Central Limit theorem (cf., Andersen, Borgan, Gill and Keiding (1993)), the weak convergence of the score process to a zero-mean Gaussian process with covariance matrix $\boldsymbol{\Sigma}(\cdot)$ follows. The desired result finally obtains by taking $t = \tau$. ||

The remainder of this paper concentrates on the case where $\boldsymbol{\Psi}(\cdot)$ is deterministic. In such a situation, by using the Glivenko-Cantelli Theorem, we easily see that the limiting covariance matrix is given by

$$\boldsymbol{\Sigma}(\tau) = \int_0^{\tau} \boldsymbol{\Psi}(s)\boldsymbol{\Psi}(s)' \exp\left[-\int_0^s p(u)\lambda_0(u)\ \mathrm{d}u\right]\lambda_0(s)\ \mathrm{d}s.$$

We also point out that the results in Theorem 2.1 can also be obtained from Proposition 1 in Peña (1998a) by taking in that paper the covariate vector to be $\boldsymbol{X} \equiv \mathbf{0}$ and the at-risk process to be $Y_j(s) = I\{W_{j\nu_j} \geq s\}$.

The asymptotic α-level smooth goodness-of-fit test of $H_0^* : \boldsymbol{\theta} = \mathbf{0}$, or equivalently $H_0 : \lambda(\cdot) = \lambda_0(\cdot)$, therefore

"Rejects H_0 whenever

$$S(\tau) \equiv \frac{1}{n}\boldsymbol{U}_\theta(\tau;\mathbf{0})'\boldsymbol{\Sigma}^-(\tau)\boldsymbol{U}_\theta(\tau;\mathbf{0}) \ \geq\ \chi^2_{k^*;\alpha}\text{"} \tag{2.3}$$

where $\boldsymbol{\Sigma}^-(\cdot)$ is a generalized inverse of $\boldsymbol{\Sigma}(\cdot)$ and $\chi^2_{k^*;\alpha}$ is the $(1-\alpha)100^{\text{th}}$ percentile of the chi-square distribution with degrees of freedom $k^* = \text{rank}[\boldsymbol{\Sigma}(\tau)]$. Note that since we are not assuming that the probability of perfect repair $p(\cdot)$ is known, $\mathbf{\Sigma}(\tau)$ needs to be estimated. A possible consistent estimator is

$$\hat{\mathbf{\Sigma}}(\tau) = \frac{1}{n}\sum_{j=1}^{n}\int_0^{\tau} \mathbf{\Psi}(s)\mathbf{\Psi}(s)' Y_j(s)\lambda_0(s)\ \mathrm{d}s.$$

It is apparent in the form of the test statistic that the choice of the process $\mathbf{\Psi}(\cdot)$ is crucial. Indeed, the $\mathbf{\Psi}(\cdot)$ process determines the family of alternatives for which the test will have good power. For the case where units are observed only up to the time of the first failure, Peña (1998ab) examined several choices for $\mathbf{\Psi}(\cdot)$. In this paper, we focus on a polynomial specification. A polynomial specification is appealing since they form a basis for most functions, aside from being of simple form. The use of polynomials is prevalent in the literature originating with Neyman (1937) and subsequently considered by Thomas and Pierce (1979) and Gray and Pierce (1985). Since these papers dealt with the density-based formulation, powers of the distribution function were used. In our hazard-based formulation, we consider $\mathbf{\Psi}(\cdot)$ of form

$$\mathbf{\Psi}(t:\mathrm{PW}_k) = [1,\quad \Lambda_0(t),\quad \ldots,\quad \Lambda_0(t)^{k-1}]', \tag{2.4}$$

where $k \in \{1, 2, \ldots\}$ is a specified order and $\Lambda_0(t) = \int_0^t \lambda_0(s)\mathrm{d}s$. The label PW_k is adopted to distinguish the polynomial specification from other forms of $\mathbf{\Psi}$ explored in the literature (cf., Peña, 1998ab; Agustin and Peña, 1999). The choice of the smoothing parameter k will be discussed in the next section when we present the results of a simulation study. In the traditional density-based formulation of the smooth alternatives, a data-dependent choice of k has been explored by Ledwina (1994) and Kallenberg and Ledwina (1995).

The smoothing process $\mathbf{\Psi}(\cdot)$ as specified in (2.4) yields the score statistic vector

$$\frac{1}{\sqrt{n}}\boldsymbol{U}_{\boldsymbol{\theta}}(\tau;\mathbf{0}) \equiv \boldsymbol{Q}(\tau:\mathrm{PW}_k) = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\left[\sum_{i=1}^{N_j(\tau)}(R_{ji})^{\ell-1} - \frac{(R^{\tau}_{j\nu_j})^{\ell}}{\ell}\right]_{\ell=1,\ldots,k},$$

where

$$R_{ji} = \Lambda_0(W_{ji}), (i = 1, 2, \ldots, \nu_j), \quad \text{and} \quad R^{\tau}_{j\nu_j} = \Lambda_0(\tau \wedge W_{j\nu_j}),$$

which can be viewed as the generalized residuals in this BBS model. The limiting covariance matrix is obtained to be

$$\mathbf{\Sigma}(\tau:\mathrm{PW}_k) = \left[\left(\int_0^{\tau} \Lambda_0(t)^{\ell+\ell'-2}\exp\{-\Lambda_0^*(t)\}\ \mathrm{d}\Lambda_0(t)\right)_{\ell,\ell'=1,\ldots,k}\right],$$

where $\Lambda_0^*(t) = \int_0^t p(u)\lambda_0(u)\mathrm{d}u$. If we take the simpler case where the probability of perfect repair is constant, the limiting covariance matrix simplifies to

$$\mathbf{\Sigma}(\tau : \mathrm{PW}_k) = \left[\left(\frac{(\ell+\ell'-2)!}{p^{\ell+\ell'-1}}\mathrm{IG}[p\Lambda_0(\tau);\ell+\ell'-1]\right)_{\ell,\ell'=1,\ldots,k}\right],$$

where $\mathrm{IG}(t;m) = \int_0^t \frac{1}{\Gamma(m)}u^{m-1}\exp(-u)\mathrm{d}u$. If we let $\tau \to \infty$, and using the fact that $\int_0^\infty \lambda_0(t)\mathrm{d}t = \infty$, then the limiting covariance matrix reduces to

$$\mathbf{\Sigma}(\mathrm{PW}_k) = \left[\left(\frac{(\ell+\ell'-2)!}{p^{\ell+\ell'-1}}\right)_{\ell,\ell'=1,\ldots,k}\right].$$

For notation, from hereon, whenever we evaluate limiting quantities as $\tau \to \infty$, we shall suppress writing the argument $\tau = \infty$. Of course, in such situations, the limiting values should be viewed as approximations to the case when τ is large. A consistent estimator of the limiting covariance matrix is given by

$$\hat{\mathbf{\Sigma}}(\tau : \mathrm{PW}_k) = \frac{1}{n}\sum_{j=1}^{n}\left[\left(\frac{(R_{j\nu_j}^{\tau})^{\ell+\ell'-1}}{\ell+\ell'-1}\right)_{\ell,\ell'=1,\ldots,k}\right].$$

The asymptotic α-level "polynomial" test of H_0 is

$$\text{"Reject } H_0 \text{ if } S(\tau : \mathrm{PW}_k) \equiv \boldsymbol{Q}(\tau : \mathrm{PW}_k)'\hat{\mathbf{\Sigma}}(\tau : \mathrm{PW}_k)^{-}\boldsymbol{Q}(\tau : \mathrm{PW}_k) \geq \chi^2_{k;\alpha}.\text{"} \tag{2.5}$$

To demonstrate some special cases of this test, if the smoothing parameter is $k = 1$, then we obtain the score statistic $Q(\tau : \mathrm{PW}_1) = \frac{1}{\sqrt{n}}\sum_{j=1}^n[N_j(\tau) - R_{j\nu_j}^{\tau}]$ and the estimated variance $\hat{\Sigma}(\tau : \mathrm{PW}_1) = \frac{1}{n}\sum_{j=1}^{n} R_{j\nu_j}^{\tau}$. Thus the resulting test statistic is

$$S(\tau : \mathrm{PW}_1) = \frac{\left[\sum_{j=1}^n[N_j(\tau) - R_{j\nu_j}^{\tau}]\right]^2}{\sum_{j=1}^n R_{j\nu_j}^{\tau}} \tag{2.6}$$

which can be viewed as a generalization of the Pearson-type test statistic studied by Akritas (1988). Furthermore, suppose we allow for right-censoring and set $p(t) = 1$. This results in the randomly right-censored model without covariates. Denote the minimum of the failure time and the censoring variable for the j^{th} unit by Z_j, and let δ_j be the correponding censoring indicator. If there are no ties among the Z_j's, then (2.6) simplifies to

$$S(\tau : \mathrm{PW}_1) = \frac{\left[\sum_{j=1}^{n}(\delta_j - R_j^\tau)\right]^2}{\sum_{j=1}^{n} R_j^\tau}, \tag{2.7}$$

where $R_j^\tau = \Lambda_0(Z_j \wedge \tau)$. The expression given in (2.7) is the test statistic for right-censored data proposed by Hyde (1977).

The individual components of $S(\tau : \mathrm{PW}_k)$ are asymptotically χ_1^2-distributed and can be used as directional tests. For $i = 1, \ldots, k$, the i^{th} directional test statistic is

$$S_i(\tau : \mathrm{PW}_k) = \frac{Q_i^2(\tau : \mathrm{PW}_k)}{\hat{\sigma}_i^2(\tau : \mathrm{PW}_k)},$$

where $\hat{\sigma}_i^2(\tau : \mathrm{PW}_k)$ is the $(i,i)^{\mathrm{th}}$ element of $\hat{\boldsymbol{\Sigma}}(\tau : \mathrm{PW}_k)$. Note that these directional test statistics need not be independent of each other. If one desires asymptotically independent directional tests, an alternative choice for the $\boldsymbol{\Psi}(\cdot)$ process is obtained by replacing the polynomial-type specification by orthogonal polynomials. In the classical density-based formulation, Neyman (1937) obtained orthogonal polynomials by choosing the components of $\boldsymbol{\Psi}$ to be orthonormal with respect to the density specified under the null hypothesis. In the hazard-based formulation, this corresponds to choosing the vector $\boldsymbol{\Psi}$ such that

$$\int_0^\tau \boldsymbol{\Psi}(w)\boldsymbol{\Psi}(w)' \exp\left(-\int_0^w p(u)\lambda_0(u)\,\mathrm{d}u\right)\lambda_0(w)\,\mathrm{d}w = \boldsymbol{I}_k,$$

where $\boldsymbol{I}_k$ is the identity matrix of order k. In the case of a constant probability of perfect repair, i.e., $p(t) \equiv p$, then the vector of interest is $\boldsymbol{\Psi}^*$ which satisfies the condition

$$\int_0^{\Lambda_0(\tau)} \boldsymbol{\Psi}^*(w)\boldsymbol{\Psi}^*(w)' \exp(-pw)\,\mathrm{d}w = \boldsymbol{I}_k.$$

The Gram-Schmidt orthogonalization procedure can be applied to obtain the elements of $\boldsymbol{\Psi}^*$. In the limiting case $\tau \to \infty$, the Gram-Schmidt procedure produces

$$\psi_i^*(w) = (-1)^{i-1}\sqrt{p}\sum_{\ell=0}^{i-1}\binom{i-1}{\ell}\frac{(-wp)^\ell}{\ell!}, \quad i = 1, \ldots, k, \tag{2.8}$$

where $\binom{i}{\ell}$ is the combination of i objects taken ℓ at a time. The functions in (2.8) are the scaled Laguerre polynomials. Note that $\psi_i^*(\cdot)$

depends on the unknown parameter p, which needs to be estimated. A consistent estimator of p is $\hat{p} = n/N_\bullet$, where $N_j = N_j(\infty)$ and $N_\bullet = \sum_{j=1}^{n} N_j$. Consequently, the score statistic is given by $\boldsymbol{Q}(\mathrm{OR}_k) = (Q_1(\mathrm{OR}_k), Q_2(\mathrm{OR}_k), \ldots, Q_k(\mathrm{OR}_k))'$ with

$$Q_h(\mathrm{OR}_k) = \frac{1}{\sqrt{n}}(-1)^{h-1}\sqrt{\hat{p}} \times$$
$$\sum_{\ell=0}^{h-1}\left\{\binom{h-1}{\ell}\frac{(-\hat{p})^\ell}{\ell!}\sum_{j=1}^{n}\left[\sum_{i=1}^{N_j}(R_{ji})^\ell - \frac{(R_{j\nu_j})^{\ell+1}}{\ell+1}\right]\right\}.$$

A consistent estimator of the limiting covariance matrix is $\hat{\boldsymbol{\Sigma}}(\mathrm{OR}_k) = \left[(\hat{\sigma}_{h_1,h_2}(\mathrm{OR}_k))_{h_1,h_2=1,\ldots,k}\right]$, where

$$\hat{\sigma}_{h_1,h_2}(\mathrm{OR}_k) = \frac{(-1)^{h_1+h_2-2}\hat{p}}{n}\times$$
$$\sum_{m_1=0}^{h_1-1}\sum_{m_2=0}^{h_2-1}\left\{\binom{h_1-1}{m_1}\binom{h_2-1}{m_2}\frac{(-\hat{p})^{m_1+m_2}}{m_1!m_2!}\left[\sum_{j=1}^{n}\frac{(R_{j\nu_j})^{m_1+m_2+1}}{m_1+m_2+1}\right]\right\}.$$

The asymptotic α-level "orthogonal" test of H_0 when $\tau \to \infty$ becomes

$$\text{“ Reject } H_0 \text{ if } S(\mathrm{OR}_k) \equiv \boldsymbol{Q}(\mathrm{OR}_k)'\hat{\boldsymbol{\Sigma}}(\mathrm{OR}_k)^{-}\boldsymbol{Q}(\mathrm{OR}_k) \geq \chi^2_{k;\alpha}.\text{''} \quad (2.9)$$

One might wonder why the limiting covariance matrix is replaced by an estimator when the limiting covariance matrix $\boldsymbol{I}_k$ is completely known. Our simulation studies revealed that for finite sample sizes, the tests which use the estimator performed better than the ones based on $\boldsymbol{I}_k$. This is similar to other testing situations where the use of the observed information matrix has advantages over those using the expected information matrix. In the simulation results presented in Section 3, one would also notice that for the same value of the smoothing parameter k, the achieved powers of $S(\mathrm{PW}_k)$ and $S(\mathrm{OR}_k)$ are identical. This phenomenon, also observed in Thomas and Pierce (1979) for the composite hypothesis case when units are observed only up to the time of the first failure, is due to the fact that the orthogonal polynomials are simply linear transformations of the polynomial-type specification, rendering the $S(\mathrm{PW}_k)$ and $S(\mathrm{OR}_k)$ to be equivalent test statistics.

A question is whether one gains by using orthogonal polynomials. Analogous to the case of the "polynomial" test, the components of $S(\mathrm{OR}_k)$ are χ^2_1-distributed and can be used as directional tests. For $i = 1, \ldots, k$, the i^{th} directional test statistic is

$$S_i(\mathrm{OR}_k) = \frac{Q_i^2(\mathrm{OR}_k)}{\hat{\sigma}_i^2(\mathrm{OR}_k)}, \quad (2.10)$$

where $\hat{\sigma}_i^2(\mathrm{OR}_k)$ is the $(i,i)^{\mathrm{th}}$ element of $\hat{\mathbf{\Sigma}}(\mathrm{OR}_k)$. Whereas the directional tests based on the polynomial specification are asymptotically dependent, the test statistics in (2.10) are asymptotically independent under this no-covariate setting. Furthermore, the simulation results show that the directional components of the test based on the polynomial specification tend to be anticonservative when the sample size is small, a problem that is not encountered when we use the directional components of the test based on the orthogonal specification. Finally, each component of this test based on the orthogonal specification is capable of detecting specific departures from the hypothesized hazard. For example, the simulation results revealed that $S_1(\mathrm{OR}_4)$ is quite sensitive against scale changes, while the other components are not sensitive to these alternatives. It is not yet definitive to us, however, which class of alternatives each component of the orthogonal test will be able to detect powerfully. This issue calls for further study.

3. Monte Carlo Studies

The results in the preceding section were developed under the asymptotic set-up, so in this section we present results of Monte Carlo simulations which provide information regarding the finite-sample properties of the polynomial and orthogonal tests. We consider the case where the probability of perfect repair $p(\cdot)$ is constant, thereby reducing the BBS (1985) model to the Brown and Proschan (1983) model. For the BBS model, in general, $p(\cdot)$ may depend on unknown parameters which need to be estimated. Hence, this scenario falls in the realm of composite hypothesis testing and will be dealt with in a separate paper.

Of particular interest in these simulations are the achieved levels and powers of the tests. The issue concerning the appropriate value of the smoothing parameter k will also be addressed in this section. For the simulations, we considered $k = 1, 2, 3, 4$. Moreover, all tests were evaluated as $\tau \to \infty$, so we omit the value of τ in the notation. We reiterate that for each value of k, $S(\mathrm{PW}_k)$ and $S(\mathrm{OR}_k)$ are equivalent. It should also be noted that $S_1(\mathrm{PW}_4)$ and $S_1(\mathrm{OR}_4)$ are identical. The simulation programs were coded in Fortran, and subroutines from the IMSL (1987) Math/Stat Library were used to generate random numbers and invert matrices. The simulations were run using Microsoft Fortran PowerStation 4.0.

3.1. Achieved Significance Levels

For determining the achieved significance levels of the tests, we tested the null hypothesis $H_0 : \lambda(t) = \lambda_0(t) = 1$, i.e., the unit exponential hazard rate function. The following experiment was replicated 2000 times. For each combination of sample size $n \in \{20, 30, 50, 100, 200\}$ and probability of perfect repair $p \in \{0.2, 0.5\}$, record values $\{W_{ji} : j = 1, \ldots, n; i \geq 1\}$ for a Brown and Proschan (1983) imperfect repair model were generated. The initial record values $\{w_{j1} : j = 1, \ldots, n\}$ were generated from a unit exponential distribution. Because of the memoryless property of the exponential distribution, succeeding record values were generated according to $w_{ji} = \sum_{\ell=1}^{i} e_{j\ell},\ i \geq 2$, where $e_{j\ell}$ is a unit exponential variate. To determine the type of repair to be performed at each failure time, Uniform$[0, 1]$ variates u_{ji} were generated corresponding to each W_{ji}. If $u_{ji} < p$, a perfect repair is performed at W_{ji}; otherwise, an imperfect repair is done. Since we assumed $\tau = \infty$, each unit was observed up to the time of its first perfect repair $W_{j\nu_j}$. The sample realization $\{w_{ji} : j = 1, \ldots, n; i = 1, \ldots, \nu_j\}$ was used to test H_0 using the various tests. The simulations were done at the 5% and 10% asymptotic level tests. Since the results lead to similar conclusions for both levels, we present only the results for the 5% asymptotic level tests.

The results are summarized in Table I. An examination of this table reveals that for a sample size of 30 and a probability of perfect repair of 0.50, the achieved levels of the directional tests $S_i(\mathrm{PW}_4)$ and $S_i(\mathrm{OR}_4)$, with the exception of $S_3(\mathrm{PW}_4)$ and $S_4(\mathrm{PW}_4)$, are consistent with the 5% asymptotic level. It should be noted that while the directional components of $S(\mathrm{PW}_4)$ tend to be anticonservative when n is small, the corresponding components of $S(\mathrm{OR}_4)$ are consistent with the predetermined significance level. For the omnibus tests, on the other hand, $S(\mathrm{PW}_3)$, $S(\mathrm{PW}_4)$, and consequently $S(\mathrm{OR}_3)$ and $S(\mathrm{OR}_4)$, are anticonservative when n is small. This anticonservatism, however, decreases as n increases and more so as p decreases. In fact, for $p = 0.2$ and $n = 100$, the achieved levels of all the tests, except for $S(\mathrm{PW}_4)$ and hence, $S(\mathrm{OR}_4)$, are consistent with the 5% asymptotic level. The decrease in anticonservatism as p decreases is due to the fact that a decrease in the probability of perfect repair translates to a larger number of observations. The simulation results therefore suggest that in this no covariate recurrent event setting, $k = 2$ or 3 will be appropriate.

Let us now turn to the achieved powers of the tests. The power simulations in subsection 3.2 used $n = 100$ and $p = 0.2$ since for this combination the achieved levels of the tests, except for $S(\mathrm{PW}_4)$, equivalently $S(\mathrm{OR}_4)$, are consistent with the 5% asymptotic level.

3.2. Achieved Powers

To examine the powers of the tests we again used the null hypothesis $H_0 : \lambda(t) = \lambda_0(t) = 1$, and considered Weibull-type and gamma-type alternatives. Record values from a distribution with hazard function $\Lambda(\cdot)$ can be generated according to $W_{ji} = \Lambda^{-1}(X_{j1} + X_{j2} + \ldots + X_{ji})$, where $X_{j\ell}$'s are unit exponential random variables. Hence, for the Weibull-type alternatives, record values were generated according to $W_{ji} = \beta(X_{j1} + X_{j2} + \ldots + X_{ji})^{1/\gamma}$, where β and γ are scale and shape parameters, respectively. To determine the type of repair to be performed at each failure time, the same scheme described in the previous subsection was used. For the gamma-type alternatives, on the other hand, we utilized the relationship between the hazard function Λ and the distribution function F to obtain the record values $W_{ji} = F^{-1}(1 - \exp[-(X_{j1} + X_{j2} + \ldots + X_{ji})]; \beta, \gamma)$. An IMSL (1987) subroutine was used to compute the inverse of the gamma distribution function. The values of (β, γ) were chosen so that the resulting alternatives represent scale and shape changes.

For the Weibull-type alternatives, the achieved powers of the tests against scale changes are sumarized in Table II. All the omnibus tests considered had very good power against these alternatives. $S(\mathrm{PW}_1)$, and hence $S(\mathrm{OR}_1)$, came out to have the highest power, followed closely by $S(\mathrm{PW}_2)$, equivalently $S(\mathrm{OR}_2)$. For the directional tests, $S_1(\mathrm{PW}_4)$ and its equivalent $S_1(\mathrm{OR}_4)$ stood out from the rest in terms of power. The other directional components of the orthogonal test had virtually no power to detect scale changes in Weibull-type alternatives. Table III summarizes the powers of the tests against shape changes in the Weibull alternatives. Taking $\gamma < 1$ results in decreasing failure rate alternatives, while $\gamma > 1$ leads to increasing failure rate alternatives. All the omnibus tests had good powers against this type of alternative. Based on the simulation results, if one suspects this type of alternative, then a value of k equal to 1 or 2 would suffice. Though the powers of the tests corresponding to $k = 3, 4$ are comparable, these tests tend to be anticonservative when the sample size is small. For the directional tests, the components of the polynomial test all had fairly good powers with $S_2(\mathrm{PW}_4)$ coming out to be the best, followed closely by $S_1(\mathrm{PW}_4)$. The directional components of the orthogonal test behaved quite differently. $S_1(\mathrm{OR}_4)$ emerged to be the best overall. The test based on $S_2(\mathrm{OR}_4)$ seems unable to detect increasing failure rate alternatives but had good power against decreasing failure rate alternatives. The opposite, however, is true for $S_3(\mathrm{OR}_4)$ and $S_4(\mathrm{OR}_4)$. Note that the powers of these two tests for decreasing failure rate alternatives are significantly smaller than those for increasing failure rate alternatives.

For the shape changes in gamma-type alternatives, the results are summarized in Table IV. $S(\mathrm{PW}_3)$ [equivalently, $S(\mathrm{OR}_3)$] had the highest power among the omnibus tests considered, while $S_1(\mathrm{PW}_4)$ [equivalently $S_1(\mathrm{OR}_4)$] turned out to be the best directional test. The other components of the polynomial test did not have any power at all to detect these alternatives. Notice that the powers for the shape changes in the gamma alternatives are considerably lower than the corresponding powers for the shape changes in the Weibull alternatives for the same value of the shape parameter γ. This is due to the fact that equal values of the shape parameters for the Weibull and gamma distributions do not mean comparable deviations from the exponential distribution. To illustrate this point, Figures 1 and 2 present the overlaid graphs of the density, distribution, hazard rate, and cumulative hazard functions of the exponential with mean unity, Weibull with $\gamma = 0.95$, and gamma with $\gamma = 0.95$. The scale parameters for the Weibull and gamma are both unity. Notice that the discrepancy between the Weibull and gamma, relative to the exponential, is very noticeable when viewed in terms of the hazards; while they seem insignificant in the context of the densities or distributions. The observed high simulated powers of the tests when the alternative is the Weibull can now be explained by the larger difference between the exponential hazard and the Weibull hazard relative to the difference between the exponential hazard and the gamma hazard. These figures seem to indicate also the advantage of using the hazard-based formulation over a density-based formulation for developing testing procedures because, when viewed through the hazards, differences of the distributions seem to be magnified.

Based on the different alternatives considered in these simulations, the polynomial, and hence the orthogonal, tests with $k = 1, 2$, or 3 show great promise as omnibus tests. Their individual components, on the other hand, are capable of detecting specific departures from the null distribution and thus will make good directional tests.

4. An Application

In this section, we illustrate the applicability of the proposed family of tests by analyzing the air conditioner data originally presented as Table 1 in Proschan (1963). The data consist of the times between failures of the air conditioning system of 13 Boeing 720 jet airplanes, with major overhauls indicated by **. Following Presnell, et al. (1994), we considered the intervals between failures as failure times between minimal repairs and the age at which a major overhaul is undertaken as the time of the first perfect repair. For those planes that were not overhauled, we treated the last observed failure age as the time of the first perfect repair. We assumed that τ is large enough so that we observe all failures until the time of the first perfect repair. Under this set-up, the total number of failures observed is $\sum_{j=1}^{13} N_j = 192$. In Presnell, et al. (1994), the minimal repair assumption was tested and no evidence against it was found. Hence, assuming that the Brown and Proschan (1983) imperfect repair model holds, we tested $H_0 : \lambda(t) = 1/94.34$, that is, the exponential distribution with mean 94.34 hours. The value 94.34 was computed solely from the interfailure times, assuming that they are independent and identically distributed exponential random variables. We computed the values of the test statistics based on both the polynomial and orthogonal specifications for $k = 1, 2, 3, 4$. The resulting values and their corresponding p-values are presented in Table V. Of the global tests, the ones based on $S(\mathrm{PW}_2/\mathrm{OR}_2)$ and $S(\mathrm{PW}_4/\mathrm{OR}_4)$ rejected the exponential assumption at the 10% level of significance. Moreover, the directional tests based on $S_3(\mathrm{PW}_4)$, $S_4(\mathrm{PW}_4)$, $S_3(\mathrm{OR}_4)$, and $S_4(\mathrm{OR}_4)$ also rejected the hypothesized exponential distribution. The fact that the value of $S(\mathrm{PW}_1/\mathrm{OR}_1)$ (equivalently, $S_1(\mathrm{PW}_4)$, $S_1(\mathrm{OR}_4)$) is 0 should not come as a surprise since the value 94.34 was computed from the data. A more appropriate test would have been to consider a composite hypothesis setting which would warrant further adjustments to the variance function. This will be done in another paper which focuses on the composite hypothesis case.

5. Concluding Remarks

In this paper, we extended the family of hazard-based smooth goodness-of-fit tests to models pertaining to recurrent events, in particular, those arising from the BBS model. Furthermore, by virtue of the fact that the BBS model is a general model which subsumes other models used in biomedical, engineering, economics, and sociological settings, the proposed goodness-of-fit tests will also have applicability in these special cases. For example, since the NHPP is a special case of the BBS model, then a goodness-of-fit test for the intensity function of a NHPP is therefore immediately obtained by simply specializing the proposed tests to the case where $p(t) = 0$. In addition, goodness-of-fit tests for the failure-time distribution in right-censored models can be generated by taking $p(t) = 1$ and introducing a censoring variable. In particular, we demonstrated that generalizations of tests considered by Hyde (1977) and Akritas (1988) can be obtained from the proposed class of tests.

The appeal of the proposed class of tests lies in the fact that a rich family of tests can be generated by varying the smoothing process $\mathbf{\Psi}(\cdot)$. This paper focused on the polynomial and orthogonal choices. The simulation studies showed that these choices lead to powerful omnibus tests and their individual components show potential for being good directional tests. The appropriate value of the smoothing parameter k was partially addressed. An examination of the simulation results revealed that $k = 2$ or 3 would be appropriate for the setting of interest. The simulation results also highlighted an advantage of the hazard-based formulation of test procedures over the traditional density-based formulation. Differences between distributions tend to be magnified when viewed via hazard functions. Other choices for the smoothing process have been explored and some of these choices lead to generalizations of existing tests for independent and identically distributed observations to recurrent events. For instance, Agustin and Peña (1999) dealt with the case where $p(\cdot)$ is completely known or constant and the smoothing process $\mathbf{\Psi}(\cdot)$ given by

$$\psi_j(s) = \frac{B_j(s)}{B_j(\tau)} - \frac{1}{2}, \quad j = 1, \ldots, n$$

where $B_j(t) = \int_0^t Y_j(u)[1-p(u)]\lambda_0(u)\, du$. This choice of smoothing process, which is a stochastic process, resulted in a test that can be viewed as a generalization of the test proposed by Barlow, Bartholomew, Bremmer and Brunk (1972) applied to generalized residuals. Furthermore, a data-driven version of the hazard-based smooth goodness-of-fit tests will be the focus of future research. Choosing k based on the observed data will make the implementation of the proposed procedures more appealing to practitioners.

Acknowledgements

The authors wish to thank the Editor, Associate Editor, and the three reviewers for their comments and suggestions. E. Peña also wishes to thank the faculties of the Department of Biostatistics and the Department of Statistics at the University of Michigan for their hospitality during his visit, and Bowling Green State University where part of this research was performed while he was professor there.

References

M. Akritas, "Pearson-type goodness-of-fit tests: the univariate case," *Journal of the American Statistical Association* vol. 83 pp. 222-230, 1988.

P. Andersen, O. Borgan, R. Gill and N. Keiding, *Statistical Models Based on Counting Processes,* Springer-Verlag: New York, 1993.

Z. Agustin and E. Peña, "Order statistic properties, random generation, and goodness-of-fit testing for a minimal repair model," *Journal of the American Statistical Association* vol. 94 pp. 266–272, 1999.

E. Arjas and I. Norros, "Change of life distribution via hazard transformation: an inequality with application to minimal repair," *Mathematics of Operations Research* vol. 14 pp. 355–361, 1989.

H. Ascher and H. Feingold, *Repairable Systems Reliability: Modeling, Inference, Misconceptions and Their Causes,* Lecture Notes in Statistics, Marcel-Dekker: New York, 1984.

R. Barlow, D. Bartholomew, J. Bremmer and H. Brunk, *Statistical Inference under Order Restrictions,* John Wiley: New York, 1972.

H. Block, W. Borges and T. Savits, "Age-dependent minimal repair," *Journal of Applied Probability* vol. 22 pp. 370–385, 1985.

M. Brown and F. Proschan, "Imperfect repair," *Journal of Applied Probability* vol. 20 pp. 851–859, 1983.

M. Crowder, A. Kimber, R. Smith and T. Sweeting, *Statistical Analysis of Reliability Data,* Chapman and Hall: London, 1991.

C. Dorado, M. Hollander and J. Sethuraman, "Nonparametric estimation for a general repair model," *The Annals of Statistics* vol. 25 pp. 1140–1160, 1997.

R. Gray and D. Pierce, "Goodness-of-fit for censored survival data," *The Annals of Statistics* vol. 13 pp. 552–563, 1985.

M. Hollander, B. Presnell and J. Sethuraman, "Nonparametric methods for imperfect repair models," *The Annals of Statistics* vol. 20 pp. 879–896, 1992.

J. Hyde, "Testing survival under right censoring and left truncation," *Biometrika* vol. 64 pp. 225-230, 1977.

IMSL, Inc., *Math/Stat Library,* Authors: Houston, 1987.

W. Kallenberg and T. Ledwina, "Consistency and Monte Carlo simulation of a data driven version of smooth goodness-of-fit tests," *The Annals of Statistics* vol. 23 pp. 1594–1608, 1995.

N. Keiding, C. Andersen and P. Fledelius, "The Cox regression model for claims data in non-life insurance," *Astin Bulletin* vol. 28 pp. 95–118, 1998.

M. Kijima, "Some results for repairable systems with general repair," *Journal of Applied Probability* vol. 26 pp. 89–102, 1989.

J. Lawless, "Repeated Events," *Encyclopedia of Biostatistics* vol. 5 pp. 3783–3787, 1998.

T. Ledwina, "Data driven version of Neyman's smooth test of fit," *Journal of the American Statistical Association* vol. 89 pp. 1000–1005, 1994.

J. Neyman, "“Smooth test” for goodness of fit," *Skand. Aktuarietidskrift* vol. 20 pp. 149–199, 1937.

D. Oakes, "Duration dependence," *Encyclopedia of Biostatistics* vol. 2 pp. 1248–1252, 1998.

E. Peña, "Smooth goodness-of-fit tests for the baseline hazard in Cox's proportional hazards model," *Journal of the American Statistical Association* vol. 93 pp. 673–692, 1998a.

E. Peña, "Smooth goodness-of-fit tests for composite hypothesis in hazard based models," *The Annals of Statistics* vol. 26 pp. 1935–1971, 1998b.

R. Prentice, B. Williams and A. Peterson, "On the regression analysis of multivariate failure time data," *Biometrika* vol. 68 pp. 373–379, 1981.

B. Presnell, M. Hollander and J. Sethuraman, "Testing the minimal repair assumption in an imperfect repair model," *Journal of the American Statistical Association* vol. 89 pp. 289–297, 1994.

F. Proschan, "Theoretical explanation of observed decreasing failure rate," *Technometrics* vol. 5 pp. 375–383, 1963.

J. Rayner and D. Best, *Smooth Tests of Goodness of Fit,* Oxford University Press, 1989.

D. Thomas and D. Pierce, "Neyman's smooth goodness-of-fit test when the hypothesis is composite," *Journal of the American Statistical Association* vol. 74 pp. 441–445, 1979.

L. Whitaker and F. Samaniego, "Estimating the reliability of systems subject to imperfect repair," *Journal of the American Statistical Association* vol. 84 pp. 301–309, 1989.

Table I. Simulated levels of 5%-asymptotic level tests for different sample sizes (n) and probability of perfect repair (p). The failure times under the null hypothesis were generated according to the Brown and Proschan imperfect repair model with initial distribution EXP(1).

p	0.50					0.20				
n	20	30	50	100	200	20	30	50	100	200
Test										
$S(\mathrm{PW/OR}_1)$	4.60	4.25	5.45	4.75	4.20	5.60	5.35	4.30	4.30	5.00
$S(\mathrm{PW/OR}_2)$	6.80	4.60	5.15	5.10	4.30	6.00	5.90	4.75	4.45	4.40
$S(\mathrm{PW/OR}_3)$	9.35	7.15	7.05	6.70	4.65	6.60	7.05	5.40	5.20	5.30
$S(\mathrm{PW/OR}_4)$	12.10	10.15	10.55	8.80	7.10	9.00	7.60	6.85	6.05	6.35
$S_1(\mathrm{PW}_4)$	4.60	4.25	5.45	4.75	4.20	5.60	5.35	4.30	4.30	5.00
$S_2(\mathrm{PW}_4)$	6.55	4.55	5.50	4.85	4.45	5.70	5.65	4.60	4.75	4.80
$S_3(\mathrm{PW}_4)$	7.95	6.85	5.80	5.55	4.55	6.40	6.40	4.85	5.40	6.30
$S_4(\mathrm{PW}_4)$	10.75	8.45	7.75	6.95	5.95	8.35	7.30	5.80	5.50	6.10
$S_1(\mathrm{OR}_4)$	4.60	4.25	5.45	4.75	4.20	5.60	5.35	4.30	4.30	5.00
$S_2(\mathrm{OR}_4)$	4.35	3.70	4.25	4.65	4.15	4.80	5.65	4.70	4.50	4.50
$S_3(\mathrm{OR}_4)$	6.30	5.10	5.05	4.60	4.45	5.75	5.85	4.50	5.15	4.80
$S_4(\mathrm{OR}_4)$	5.20	4.60	4.85	4.80	4.95	5.10	5.80	4.75	5.20	4.75

Table II. Simulated powers of the 5%-asymptotic level tests when the failure times were generated according to the Brown and Proschan imperfect repair model with initial distribution Weibull(β, γ), probability of perfect repair $p = 0.20$, and sample size $n = 100$.

Test	β	0.85	0.90	0.95	1.00	1.05	1.10	1.15
	γ	1	1	1	1	1	1	1
$S(\mathrm{PW/OR}_1)$		94.65	64.90	22.05	5.15	18.75	55.30	89.15
$S(\mathrm{PW/OR}_2)$		91.55	56.70	18.95	4.85	13.90	43.75	81.10
$S(\mathrm{PW/OR}_3)$		88.60	51.85	17.15	5.55	12.10	38.75	74.30
$S(\mathrm{PW/OR}_4)$		85.60	48.55	16.85	5.95	10.75	34.50	69.80
$S_1(\mathrm{PW}_4)$		94.65	64.90	22.05	5.15	18.75	55.30	89.15
$S_2(\mathrm{PW}_4)$		75.45	43.05	16.85	5.25	10.90	29.40	59.30
$S_3(\mathrm{PW}_4)$		48.30	27.30	13.60	5.60	4.85	10.55	23.30
$S_4(\mathrm{PW}_4)$		34.80	21.45	12.45	5.50	3.65	4.35	9.15
$S_1(\mathrm{OR}_4)$		94.65	64.90	22.05	5.15	18.75	55.30	89.15
$S_2(\mathrm{OR}_4)$		11.90	5.70	4.70	5.30	4.70	5.30	7.45
$S_3(\mathrm{OR}_4)$		7.45	7.00	5.35	4.80	4.60	5.35	6.15
$S_4(\mathrm{OR}_4)$		7.65	6.25	5.75	4.40	4.85	4.85	5.30

Table III. Simulated powers of the 5%-asymptotic level tests when the failure times were generated according to the Brown and Proschan imperfect repair model with initial distribution Weibull(β, γ), probability of perfect repair $p = 0.20$, and sample size $n = 100$.

Test	β	1	1	1	1	1	1	1
	γ	0.85	0.90	0.95	1.00	1.05	1.10	1.15
$S(\mathrm{PW}/\mathrm{OR}_1)$		100.00	99.20	67.30	5.00	56.95	98.70	99.95
$S(\mathrm{PW}/\mathrm{OR}_2)$		100.00	99.40	63.45	4.65	60.10	99.35	99.95
$S(\mathrm{PW}/\mathrm{OR}_3)$		100.00	99.50	61.85	5.55	54.25	99.00	99.95
$S(\mathrm{PW}/\mathrm{OR}_4)$		100.00	99.15	55.70	5.85	55.10	99.00	99.95
$S_1(\mathrm{PW}_4)$		100.00	99.20	67.30	5.00	56.95	98.70	99.95
$S_2(\mathrm{PW}_4)$		100.00	99.65	68.85	5.25	67.65	99.45	99.95
$S_3(\mathrm{PW}_4)$		100.00	95.35	39.05	5.15	50.35	93.60	99.65
$S_4(\mathrm{PW}_4)$		99.10	75.20	17.45	5.50	39.85	81.65	96.15
$S_1(\mathrm{OR}_4)$		100.00	99.20	67.30	5.00	56.95	98.70	99.95
$S_2(\mathrm{OR}_4)$		100.00	88.35	27.55	4.85	12.40	18.95	14.20
$S_3(\mathrm{OR}_4)$		49.95	27.75	15.35	4.65	21.80	63.60	86.70
$S_4(\mathrm{OR}_4)$		29.25	10.85	7.00	6.05	14.50	54.70	88.00

Table IV. Simulated powers of the 5%-asymptotic level tests when the failure times were generated according to the Brown and Proschan (1983) imperfect repair model with initial distribution Gamma(γ, β), probability of perfect repair $p = 0.20$, and sample size $n = 100$.

Test	β / γ	1 / 0.85	1 / 0.90	1 / 0.95	1 / 1.00	1 / 1.05	1 / 1.10	1 / 1.15
$S(\mathrm{PW/OR}_1)$		30.50	15.65	8.50	5.55	6.20	13.85	22.25
$S(\mathrm{PW/OR}_2)$		29.80	13.90	8.30	5.85	7.05	14.45	25.70
$S(\mathrm{PW/OR}_3)$		35.65	18.10	9.85	5.40	6.10	11.95	21.80
$S(\mathrm{PW/OR}_4)$		32.70	16.65	9.30	6.65	8.75	15.25	25.50
$S_1(\mathrm{PW}_4)$		30.50	15.65	8.50	5.55	6.20	13.85	22.25
$S_2(\mathrm{PW}_4)$		8.25	6.40	5.55	5.10	4.55	4.65	4.40
$S_3(\mathrm{PW}_4)$		6.35	5.80	5.65	5.35	4.85	4.75	4.70
$S_4(\mathrm{PW}_4)$		6.60	6.70	5.75	5.75	5.90	5.40	4.95
$S_1(\mathrm{OR}_4)$		30.50	15.65	8.50	5.55	6.20	13.85	22.25
$S_2(\mathrm{OR}_4)$		15.05	7.10	5.05	4.85	6.30	10.40	15.45
$S_3(\mathrm{OR}_4)$		18.20	11.55	6.55	4.75	5.70	7.80	10.20
$S_4(\mathrm{OR}_4)$		16.85	10.65	6.50	4.65	5.75	6.95	9.25

Table V. Computed test statistics and their corresponding p-values for testing that the initial failure distribution of the Proschan (1963) air conditioner data set is exponential with mean 94.34 hours.

Test	Test Statistic	p-value
$S(\mathrm{PW}_1/\mathrm{OR}_1)$	0.00	1.0000
$S(\mathrm{PW}_2/\mathrm{OR}_2)$	5.84	0.0539[a]
$S(\mathrm{PW}_3/\mathrm{OR}_3)$	5.89	0.1171
$S(\mathrm{PW}_4/\mathrm{OR}_4)$	8.30	0.0812[a]
$S_1(\mathrm{PW}_4)$	0.00	1.0000
$S_2(\mathrm{PW}_4)$	1.72	0.1897
$S_3(\mathrm{PW}_4)$	2.98	0.0843[a]
$S_4(\mathrm{PW}_4)$	3.77	0.0522[a]
$S_1(\mathrm{OR}_4)$	0.00	1.0000
$S_2(\mathrm{OR}_4)$	2.40	0.1213
$S_3(\mathrm{OR}_4)$	5.40	0.0201[b]
$S_4(\mathrm{OR}_4)$	5.02	0.0251[b]

[a] Significant at the 10% level.
[b] Significant at the 5% level.

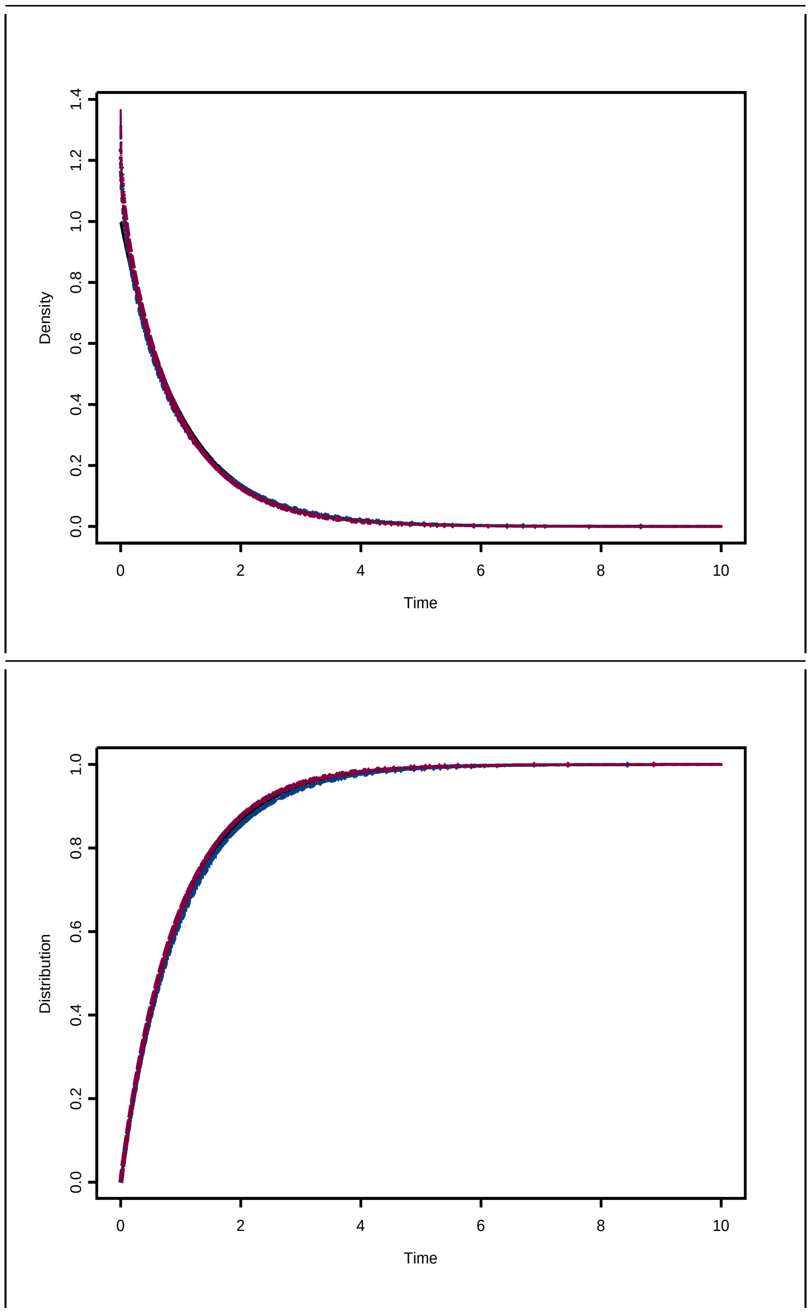

Figure 1. Overlaid plots of the density and distributionfunctions associated with the unit exponential, Weibull with $\gamma = 0.95, \beta = 1$, and gamma with $\gamma = 0.95, \beta = 1$. Legend: Solid = Exponential; Dash-Dot = Weibull; and Dashes = Gamma.

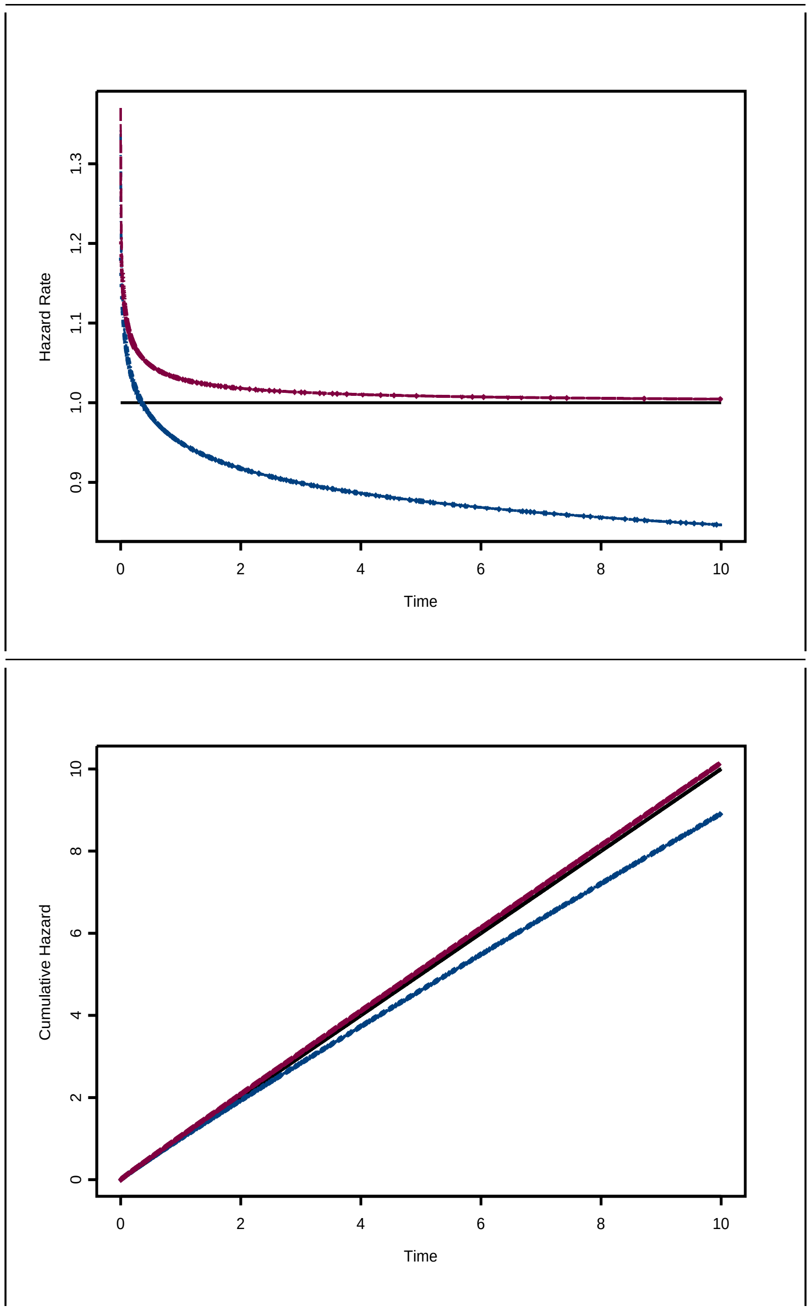

Figure 2. Overlaid plots of the hazard rate and cumulative hazard functions associated with the unit exponential, Weibull with $\gamma = 0.95, \beta = 1$, and gamma with $\gamma = 0.95, \beta = 1$. Legend: Solid = Exponential; Dash-Dot = Weibull; and Dashes = Gamma.