# Intensity-Based Approach to Goodness-of-Fit Testing with Discrete Right-Censored Data

Edsel A. Peña*

January 31, 2002

**Abstract**

A general class of intensity-based goodness-of-fit tests for right-censored discrete data is presented. The procedures are applicable to data which are presented in a life-table, or to interval-censored data. Finite and asymptotic properties of the test statistics, which were derived as partial likelihood score statistics, are obtained, with the main asymptotic result facilitated through the discrete martingale central limit theorem. Specific members of this class of tests are illustrated. In particular, tests for the null hypothesis that the discrete failure times are geometrically distributed are obtained. Under this geometric setting, simulation results pertaining to the achieved levels and powers of the tests are presented. Through these simulation studies, tests based on a polynomial-type specification perform well against a wide variety of alternatives, so such tests could serve as omnibus goodness-of-fit tests.

KEYWORDS AND PHRASES: Hazard odds; martingale difference array; Neyman's smooth test; partial likelihood; score test; test for geometric distribution.

AMS SUBJECT CLASSIFICATION: Primary: 62N03 Secondary: 62G10

## 1 Introduction

Statistical goodness-of-fit (gof) testing has always been an active research area as evidenced by entering the phrase "goodness of fit" in the `MathSciNet` search engine. In its simplest form a random sample $T_1, T_2, \ldots, T_n$ from an unknown distribution function $F$ is observed, and it is desired to determine if $F = F_0$, where $F_0$ is a specified distribution. The most well-known gof procedure is Pearson's (1900) chi-square test which utilizes the statistic

$$\chi^2 = \sum_{j=1}^{K} \frac{(O_j - E_j)^2}{E_j}, \tag{1}$$

*E. Peña is Professor in the Department of Statistics, University of South Carolina, Columbia, SC 29208. He acknowledges NIH/NIGMS Grant 1 R01 GM56182 which partially supported this research and NSF Grant DMS 0102870 which partially supports his research. He also wishes to thank Prof. V. Nair for discussions that benefited this work.

where $K$ is the size of the partition of the support of $F_0$, $O_j$ is the number of $T_i$'s in the $i$th member of the partition, and $E_j$ is the number of $T_i$'s expected to be in the $j$th member of the partition when $F_0$ holds. The popularity of this test is partly due to its simplicity and the fact that it requires only critical values from the family of chi-square distributions. There are other tests for the simple gof problem, such as Kolmogorov-Smirnov (KS) type tests, Neyman's (1937) smooth gof tests, Cramer-von Mises (CVM) type tests, and those by Khamaladze (1981, 1993). A review of some of these procedures could be found in Stephens (1992). Many of these tests have extensions to the composite null hypothesis setting, where the problem is to test whether $F \in \mathcal{C}$, with $\mathcal{C}$ a specified (parametric) family of distributions, cf., Chernoff and Lehmann (1954), Rao and Robson (1974), D'Agostino and Stephens (1986), and Greenwood and Nikulin (1996). Except for Pearson's test, most of the above-mentioned procedures imposes the restriction that $F$ is continuous, with this assumption typically made in order to facilitate the derivations of distributional results.

Though not as prevalent as the case with continuous distributions, gof tests for discrete distributions, or when data arose from grouping of continuous data, have also been considered. Kulperger and Singh (1982) examined $\chi^2$ gof tests for discrete distributions and considered the issue of random grouping. Cressie and Read (1984) introduced the family of power divergence statistics for performing gof with multinomial data. Best and Rayner (1989, 1999) proposed Neyman smooth gof tests for the null hypothesis that $F$ is geometric and Poisson, respectively; while Eubank (1997) proposed Neyman smooth-type tests for dealing with multinomial data. In Choulakian, Lockhart and Stephens (1994) a test for the discrete uniform was presented; while in Spinelli and Stephens (1997) CVM-type procedures were developed for testing a Poisson distribution. Kocherlakota and Kocherlakota (1986), Rueda, Perez-Abreau and O'Reilly (1991), Baringhaus and Henze (1992), and Nakamura and Perez-Abreau (1993) examined gof procedures for discrete data using the empirical probability generating function; in particular, tests for the Poisson distribution were developed. Empirical distribution-based methods were also considered for discrete models. Among papers adopting this approach were Henze (1996) and Klar (1999). However, all of these papers dealing with goodness-of-fit for discrete models assume that $T_1, T_2, \ldots, T_n$ are completely observed.

In biomedical, engineering, reliability, and in other areas where the primary variable of interest is the time-to-occurrence of an event, hereon referred to as a failure time, it is typical that some of the failure times will be right-censored due to time contraints, limited resources, withdrawal from the study, loss to follow-up, etc. Numerous papers have appeared dealing with the modeling and analysis of failure times in the presence of incomplete observations. For continuous failure times, the problem of gof testing has been addressed in several papers with the aim of extending to censored data those procedures that were developed for complete data. Among these papers are those of Koziol and Green (1976), Hyde (1977), Hollander and Proschan (1979), Nair (1981, 1982, 1984), Gatsonis, Hsieh and Korwar (1985), Habib and Thomas (1986), Akritas (1988), Hjort (1990), Hollander and Peña (1992), Li and Doss (1993), and Kim (1993). An interesting goal in gof testing with censored data is to extend Pearson's test. The difficulty underlying such an extension is that the exact number of failures in a member of the partition is not observable . An attempt to extend Neyman's smooth gof procedure in the presence of right-censored data has also been made by Gray and Pierce (1985). Their approach parallels that of Neyman (1937) where the density function is embedded in a wider class. A different extension of the smooth gof tests with continuous failure times, which adapts naturally to censored data and enables point process theory, was that in Peña (1998ab) and Agustin and Peña (2001), the latter dealing with reliability models for recurrent events.

Except for the test proposed in Hyde (1977) which is a special case of the class of tests proposed in the present paper, the gof problem with right-censored discrete failure times does not seem to have been investigated extensively in the literature. The existing gof procedures for discrete and complete data mentioned earlier have not yet been extended for discrete and censored data, which is a rather surprising situation since discrete failure times are ubiquitous in many studies. For instance, discrete failure times occur because of the intrinsic nature of the failure time process such as when failure is measured in terms of counts, or due to an inherent limitation in the measurement process forcing subjects to be observed only at the end of specified intervals, such as when the failure times are interval-censored or when data is presented in a life-table format as is done in actuarial settings.

The present paper aims to partially remedy this situation by providing a general class of gof tests for discrete failure times and in the presence of right-censoring. The procedures presented in this paper are discrete analogs of the intensity-based smooth goodness-of-fit tests developed in Peña (1998ab) for continuous failure times. In this formulation, the sequence of odds associated with the hazard rates are embedded in a wider class, in contrast to the usual Neyman formulation where the sequence of probabilities are the ones which are embedded, cf., Rayner and Best (1989). This intensity-based embedding facilitates the derivation of the smooth goodness-of-fit tests as score tests, thereby endowing the tests with certain local optimality properties. Furthermore, the framework allows for more general test statistics owing to the use of the discrete martingale central limit theorems in Helland (1982) to obtain distributional results. In addition, in contrast to the development of Pearson's test in which the vantage point is the time origin and the underlying question is: 'How many observations are expected to have values in a member of the partition of the support of $F_0$?' the current approach's vantage point is dynamic in that the relevant question is: 'Given that just before a certain time point there are a certain number of units at risk, how many are expected to fail at this time point?' Consequently, instead of dealing with global probabilities, the main focus are conditional probabilities, hazards, or intensities, which are the natural quantities when dealing with dynamic or time-evolving systems.

This paper will focus mainly on the goodness-of-fit problem with a simple null hypothesis. The case with a composite null hypothesis will be dealt with in another paper. We now outline the contents of this paper. In Section 2 we introduce notation and state formally the goodness-of-fit problem. Section 3 presents the development of the class of smooth goodness-of-fit tests, and in Section 4, special cases of the test statistic are illustrated by varying the smoothing matrix. In particular, we demonstrate that an analog of the Pearson statistic can be obtained from the class of test statistics by choosing an appropriate smoothing matrix. Finite and asymptotic properties of the test statistics under the null hypothesis are derived in Section 5. The advantage of using the intensity-based formulation for deriving the smooth goodness-of-fit tests becomes evident here as the problem of showing convergence in distribution of the statistics is facilitated through the use of central limit theorems for martingale difference arrays. In Section 6 we apply the asymptotic results to the special cases of the test statistics discussed in Section 4. As a concrete application of the proposed class of tests, in Section 7 special cases of the proposed class of tests are obtained for testing that the failure times are geometrically-distributed. Section 8 presents the results of simulation studies to ascertain the achieved levels and powers of the tests presented in Section 7. These simulation studies help in ascertaining whether the asymptotic approximations are acceptable, and pinpoint, at least for the geometric setting, which tests tend to have good power in detecting a wide range of alternatives, as well as which tests could be considered as directional tests in the sense of having good power for specific departures from the null hypothesis, but poor power against other alternatives. So as to facilitate a smooth exposition, most formal proofs of lemmas and theorems are relegated in technical Appendix A.

## 2 Basic Entities and the GOF Problem

Let $T$ be a discrete failure-time variable taking values in the ordered countable set $\mathcal{A} = \{a_1, a_2, a_3, \ldots\}$ with $a_j < a_{j+1}$, and denote by $F(t) = \mathbf{P}\{T \leq t\}$ its distribution function. Let

$$\Lambda(t) = \int_{-\infty}^{t} \frac{\mathrm{d}F(w)}{1 - F(w-)}, \quad t \in \Re,$$

be the associated hazard function. Denote by $\lambda_j = \Delta\Lambda(a_j) = \Lambda(a_j) - \Lambda(a_j-), j = 1, 2, 3, \ldots$, the hazard rate at $a_j$, so that $\lambda_j = \mathbf{P}\{T = a_j | T \geq a_j\}$. Then the following identities are well-known (cf., Kalbfleisch and Prentice (1980)):

$$\bar{F}(t) = 1 - F(t) = \mathbf{P}\{T > t\} = \prod_{\{j:\ a_j \le t\}} (1 - \lambda_j); \tag{2}$$

$$p_j = \mathbf{P}\{T = a_j\} = \left\{ \prod_{\{i:\ a_i < a_j\}} (1 - \lambda_i) \right\} \lambda_j = \left\{ \prod_{i=1}^{j} (1 - \lambda_i) \right\} \left( \frac{\lambda_j}{1 - \lambda_j} \right). \tag{3}$$

Let $\{\lambda_1^0, \lambda_2^0, \ldots, \lambda_j^0, \ldots\}$ be a specified sequence of hazard probabilities.

$$H_0 : \lambda_j = \lambda_j^0, \quad j = 1, 2, \ldots, \quad \text{versus} \quad H_1 : \lambda_j \neq \lambda_j^0 \ \text{ for some } j \in \{1, 2, \ldots, \}. \tag{4}$$

Let $T_1, T_2, \ldots, T_n$ be independent and identically distributed (iid) random variables from $F$. Furthermore, let $C_1, C_2, \ldots, C_n$ be another set of random variables such that $C_i \in \mathcal{A}$ and for each $i = 1, 2, \ldots, n$,

$$\mathbf{P}\{T_i = a_j | T_i \ge a_j, C_i \ge a_j\} = \mathbf{P}\{T_i = a_j | T_i \ge a_j\} = \lambda_j \text{ for all } j = 1, 2, \ldots. \tag{5}$$

Condition (5) is the so-called independent censoring condition, which postulates that the hazard rate of the unit at time $a_j$, given the information that it is at risk at $a_j$, equals its unconditional hazard rate $\lambda_j$. This condition holds for instance in the special case where $T_i$'s and $C_i$'s are mutually independent, but (5) is weaker than this outright independence between failure and censoring times. For more discussion of this notion, see Kalbfleisch and Prentice (1980) and Fleming and Harrington (1991) for the case where failure times are continuous. The observable variables are the right-censored vectors

$$(Z_1, \delta_1), (Z_2, \delta_2), \ldots, (Z_n, \delta_n) \tag{6}$$

with $Z_i = \min(T_i, C_i)$, $\delta_i = I\{T_i \le C_i\}$, and $I\{A\}$ is the indicator function of event $A$. The test of $H_0$ versus $H_1$ in (4) is to be based on the data in (6).

The framework described above could be applied to grouped continuous failure times. The discrete values $a_j$'s may represent an interval of values of the continuous variable. To amplify, if $T^*$ is the continuous failure time variable, then there may be prespecified time points $0 \equiv t_0^* <$

$t_1^* < t_2^* < \ldots < t_J^*$, and $T^* \in (t_{j-1}^*, t_j^*]$ if and only if the observable variable $T$ takes the value $a_j$. If we let

$$R_j = \sum_{i=1}^n I\{Z_i \geq a_j\}, \quad O_j = \sum_{i=1}^n I\{Z_i = a_j, \delta_i = 1\}, \quad \text{and} \quad W_j = \sum_{i=1}^n I\{Z_i = a_j, \delta_i = 0\},$$

the resulting data could be summarized in the form of a life-table as depicted below.

| $a_j$ | $R_j$ | $O_j$ | $W_j$ |
|---|---|---|---|
| $a_1$ | $R_1$ | $O_1$ | $W_1$ |
| $a_2$ | $R_2$ | $O_2$ | $W_2$ |
| $a_3$ | $R_3$ | $O_3$ | $W_3$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $a_J$ | $R_J$ | $O_J$ | $W_J$ |

In this representation, $R_j$ is the number of units at risk, $O_j$ is the number of units that failed, and $W_j$ is the number of units censored or withdrawn in the $j$th interval, respectively. The discrete hazard rate $\lambda_j$ could be viewed as the total hazard of the interval $(t_{j-1}^*, t_j^*]$, that is, if $\Lambda^*(\cdot)$ is the cumulative hazard function of $T^*$, then

$$\lambda_j = \Lambda^*((t_{j-1}^*, t_j^*]) = \Lambda^*(t_j^*) - \Lambda^*(t_{j-1}^*), \quad j = 1, 2, \ldots.$$

# 3 The General Class of GOF Tests

## 3.1 Hazard Odds Embedding and Likelihoods

For $j = 1, 2, \ldots, J$, define the true and null hazard odds, respectively, via

$$\rho_j = \frac{\lambda_j}{(1 - \lambda_j)} \quad \text{and} \quad \rho_j^0 = \frac{\lambda_j^0}{(1 - \lambda_j^0)}.$$

In terms of these hazard odds the gof problem amounts to testing $H_0 : \rho_j = \rho_j^0, j = 1, 2, \ldots, J$, versus $H_1 : \rho_j \neq \rho_j^0$ for some $j \in \{1, 2, \ldots, J\}$. Let $p$ be a pre-specified positive integer, and let

$$\mathbf{\Psi} = (\mathbf{\Psi}_1, \mathbf{\Psi}_2, \ldots, \mathbf{\Psi}_J) = \begin{bmatrix} \psi_1 \\ \psi_{\mathbf{2}} \\ \vdots \\ \psi_{\mathbf{p}} \end{bmatrix} \tag{7}$$

be a $p \times J$ matrix, possibly random. The $\boldsymbol{\Psi}_j, j = 1, 2, \ldots, J$, are $p \times 1$ vectors; while $\psi_k, k = 1, 2, \ldots, p$, are $1 \times J$ vectors. Assume that $\psi_1, \psi_2, \ldots, \psi_p$ are linearly independent. Specific choices of these vectors will be discussed later. The idea in the development of the class of tests is to embed the hypothesized value $(\rho_1^0, \rho_2^0, \ldots, \rho_J^0)$ in the parameterized class

$$\mathcal{C}_p = \left\{(\rho_1(\theta), \rho_2(\theta), \ldots, \rho_J(\theta)) : \theta = (\theta_1, \theta_2, \ldots, \theta_p)^{\mathrm{t}} \in \Re^p\right\} \tag{8}$$

where

$$\rho_j(\theta) = \rho_j^0 \exp\{\theta^{\mathrm{t}} \boldsymbol{\Psi}_j\}, \quad j = 1, 2, \ldots, J. \tag{9}$$

This is equivalent to assuming that the logarithms of the hazard odds ratios satisfy

$$\log \left\{ \frac{\rho_j(\theta)}{\rho_j^0} \right\} = \theta^{\mathrm{t}} \boldsymbol{\Psi}_j = \sum_{k=1}^{p} \theta_k \Psi_{kj}, \quad j = 1, 2, \ldots, J. \tag{10}$$

One may justify this embedding as follows. The mapping $j \in \{1, 2, \ldots, J\} \mapsto h_j \equiv \log\{\rho_j/\rho_j^0\}$ may be viewed as having range $\mathcal{H}$, a subspace of $\Re^J$. Let $\{\psi_1^{\mathrm{t}}, \psi_2^{\mathrm{t}}, \ldots, \psi_J^{\mathrm{t}}\}$ be a collection of $J \times 1$ vectors which forms a basis of $\mathcal{H}$, so that any member of $\mathcal{H}$ is a linear combination of these $\psi_j$'s. In particular,

$$\left( \log \left\{ \frac{\rho_j}{\rho_j^0} \right\}_{j=1,2,\ldots,J} \right)^{\mathrm{t}} = \sum_{j=1}^{J} \theta_j \psi_j^{\mathrm{t}}. \tag{11}$$

If the $\psi_j$'s are properly ordered according to some criterion, such as their frequency, then one may truncate the summation in the right-hand side of (11) up to a certain prespecified order $p$ to obtain the approximation

$$\left( \log \left\{ \frac{\rho_j}{\rho_j^0} \right\}_{j=1,2,\ldots,J} \right)^{\mathrm{t}} \approx \sum_{j=1}^{p} \theta_j \psi_j^{\mathrm{t}}. \tag{12}$$

Solving for $\rho_j$ in (12) yields precisely the embedding specified in (8) and (9). With this embedding at hand, the gof problem now reduces to testing $H_0^* : \theta = \mathbf{0}$ versus $H_1^* : \theta \neq \mathbf{0}$. The members of the proposed class of gof tests are the score tests of $H_0^* : \theta = \mathbf{0}$ generated by varying the smoothing matrix $\boldsymbol{\Psi}$ in (7) and the smoothing order $p$.

The resulting class of tests are the discrete failure time analog of the intensity-based smooth goodness-of-fit tests introduced in Peña (1998ab) which were developed for continuous failure times. Those procedures were in turn generalizations of the smooth tests by Neyman (1937) (see also Rayner and Best (1989)), but whereas Neyman dealt with the embedding of the density functions, these intensity-based formulations instead embed the hazard rates. This approach offers an advantage in cases where the failure times are censored. See for instance Gray and Pierce's (1985) attempt to mimic Neyman's procedure in the situation with censored continuous failure times. A further difference between the continuous and discrete failure time settings is that whereas in the continuous setting the hazard rate function was the main focus of the embedding, in the discrete case we deal instead with the embedding of the hazard odds. This is somewhat analogous to Cox's (1972) discrete generalization of his proportional hazards model where an odds model was utilized.

We obtain the likelihood function given the data $(Z_1, \delta_1), (Z_2, \delta_2), \ldots, (Z_n, \delta_n)$ for the period $[0, a_J]$. We note that for the given $J$, the effective censoring variables are $\min(C_i, a_J), i = 1, 2, \ldots, n$. We first state an intermediate result whose proof is straightforward hence omitted.

**Lemma 1** *The independent censoring condition in (5) implies that, for every* $j = 1, 2, \ldots$,

$$\mathbf{P}(T_i > a_j | T_i \geq a_j, C_i = a_j) = \mathbf{P}(T_i > a_j | T_i \geq a_j) \left\{ \frac{\mathbf{P}(C_i = a_j | T_i > a_j, C_i \geq a_j)}{P(C_i = a_j | T_i \geq a_j, C_i \geq a_j)} \right\}.$$

We now obtain the relevant (partial) likelihood of $\lambda = (\lambda_1, \ldots, \lambda_J)$. Given $(Z_i, \delta_i) = (z_i, d_i), i = 1, 2, \ldots, n$, we have

$$\begin{aligned} L \equiv L(\lambda) &= \prod_{i=1}^{n} \mathbf{P}(T_i = z_i, C_i \geq z_i)^{d_i} \mathbf{P}(T_i > z_i, C_i = z_i)^{1-d_i} \\ &= \prod_{i=1}^{n} \Big[ \{\mathbf{P}(T_i \geq z_i, C_i \geq z_i) \mathbf{P}(T_i = z_i, C_i \geq z_i | T_i \geq z_i, C_i \geq z_i)\}^{d_i} \times \\ &\quad \{\mathbf{P}(T_i \geq z_i, C_i \geq z_i) \mathbf{P}(T_i > z_i, C_i = z_i | T_i \geq z_i, C_i \geq z_i)\}^{1-d_i} \Big] \\ &= \prod_{i=1}^{n} \Big[ \mathbf{P}(T_i \geq z_i, C_i \geq z_i) \mathbf{P}(T_i = z_i | T_i \geq z_i, C_i \geq z_i)^{d_i} \times \\ &\quad \{\mathbf{P}(C_i = z_i | T_i \geq z_i, C_i \geq z_i) \mathbf{P}(T_i > z_i | T_i \geq z_i, C_i = z_i)\}^{1-d_i} \Big]. \end{aligned}$$

Applying Lemma 1 and condition (5), we obtain

$$L(\lambda) = \prod_{i=1}^{n} \Big[ \mathbf{P}(T_i \geq z_i, C_i \geq z_i) \mathbf{P}(T_i = z_i | T_i \geq z_i)^{d_i} \times \\ \{ \mathbf{P}(T_i > z_i | T_i \geq z_i) \mathbf{P}(C_i = z_i | T_i > z_i, C_i \geq z_i) \}^{1-d_i} \Big] = L_1(\lambda) L_2(\lambda)$$

with

$$L_1(\lambda) = \prod_{i=1}^{n} \mathbf{P}(T_i \geq z_i) \mathbf{P}(T_i = z_i | T_i \geq z_i)^{d_i} \mathbf{P}(T_i > z_i | T_i \geq z_i)^{1-d_i} \tag{13}$$

$$L_2(\lambda) = \prod_{i=1}^{n} \mathbf{P}(C_i \geq z_i | T_i \geq z_i) \mathbf{P}(C_i = z_i | T_i > z_i, C_i \geq z_i)^{1-d_i}. \tag{14}$$

For purposes of developing our procedures, we focus on the partial likelihood $L_1(\lambda)$. Notice that if $T_i$'s and $C_i$'s are independent, and the distribution of the $C_i$'s does not involve the $\lambda_j$'s, then the factor $L_2(\lambda)$ will be uninformative about $\lambda$. With $\lambda(z) = \sum_{j=1}^{\infty} \lambda_j I\{z = a_j\}$, by using (2), we obtain from (13)

$$\begin{aligned} L_1(\lambda) &= \prod_{i=1}^{n} \left\{ \prod_{\{j:\ a_j < z_i\}} (1 - \lambda_j) \right\} \lambda(z_i)^{d_i} [1 - \lambda(z_i)]^{1-d_i} \\ &= \prod_{i=1}^{n} \left\{ \prod_{\{j:\ a_j \leq z_i\}} (1 - \lambda_j) \right\} \left( \frac{\lambda(z_i)}{1 - \lambda(z_i)} \right)^{d_i}. \end{aligned}$$

For $j = 1, 2, \ldots$, let

$$O_j = \sum_{i=1}^{n} I\{Z_i = a_j, \delta_i = 1\}; \tag{15}$$

$$R_j = \sum_{i=1}^{n} I\{Z_i \geq a_j\}. \tag{16}$$

Thus, $O_j$ is the number of units that failed at $a_j$, while $R_j$ is the number of units that are still in the study just before $a_j$, that is, the number at risk at $a_j$. It follows that

$$L_1(\lambda) = \prod_{j=1}^{J} \left( \frac{\lambda_j}{1 - \lambda_j} \right)^{O_j} (1 - \lambda_j)^{R_j} = \prod_{j=1}^{J} \lambda_j^{O_j} (1 - \lambda_j)^{R_j - O_j}.$$

We have therefore established the following theorem.

**Theorem 1** *Under condition (5), the partial likelihood of* $\lambda$*, given* $(Z_i, \delta_i), i = 1, 2, \ldots, n$*, is*

$$L_1(\lambda) = \prod_{j=1}^{J} \lambda_j^{O_j}(1 - \lambda_j)^{R_j - O_j},$$

*where* $O_j$*'s and* $R_j$*'s are defined in (15) and (16), respectively.*

Note that since $\lambda_j = \rho_j/(1 + \rho_j)$ and $1 - \lambda_j = 1/(1 + \rho_j)$, then $L_1(\lambda)$ can be expressed in terms of $\rho = (\rho_1, \rho_2, \ldots, \rho_J)$ via

$$L_1(\rho) = \prod_{j=1}^{J} \frac{\rho_j^{O_j}}{(1 + \rho_j)^{R_j}}. \tag{17}$$

We remark that the partial likelihood $L_1(\lambda)$ is the full likelihood for $\lambda$ if the $T_i$'s and $C_i$'s are independent and the $C_i$'s are uninformative about $\lambda$ in the sense of having the distribution of $C_i$ not depending on $\lambda$. Without these conditions, $L_1(\lambda)$ is simply a partial likelihood, and there could be a consequent loss in efficiency by utilizing $L_1(\lambda)$. This situation is analogous to the use of the partial likelihood when dealing with Cox's proportional hazards model.

### 3.2 Score Function

By using the embedding of $\rho^0$ into $\mathcal{C}_p$ defined in (8) and (9), the partial likelihood of $\theta = (\theta_1, \theta_2, \ldots, \theta_p)^{\mathrm{t}}$ becomes $L_1^*(\theta) = \prod_{j=1}^{J} \left\{[\rho_j(\theta)]^{O_j}/[1 + \rho_j(\theta)]^{R_j}\right\}$, where $\rho_j(\theta) = \rho_j^0 \exp(\theta^{\mathrm{t}} \mathbf{\Psi}_j), j = 1, 2, \ldots, J$. Therefore, $l_1^*(\theta) = \log L_1^*(\theta) = \sum_{j=1}^{J} \{O_j \log \rho_j(\theta) - R_j \log[1 + \rho_j(\theta)]\}$. Since, for $j = 1, 2, \ldots, J$, we have $\partial/\partial\theta \rho_j(\theta) = \mathbf{\Psi}_j \rho_j(\theta)$ and $\partial^2/\partial\theta\partial\theta^{\mathrm{t}} \rho_j(\theta) = \mathbf{\Psi}_j^{\otimes 2} \rho_j(\theta)$, where, for a vector $\mathbf{a}$, $\mathbf{a}^{\otimes 2} = \mathbf{a}\mathbf{a}^{\mathrm{t}}$, then

$$\mathbf{U}^*(\theta) \equiv \frac{\partial}{\partial\theta} l_1^*(\theta) = \sum_{j=1}^{J} \mathbf{\Psi}_j \left\{O_j - R_j \left(\frac{\rho_j(\theta)}{1 + \rho_j(\theta)}\right)\right\} = \sum_{j=1}^{J} \mathbf{\Psi}_j \{O_j - R_j \lambda_j(\theta)\}, \tag{18}$$

where we used the identity $\lambda_j(\theta) = \rho_j(\theta)/[1+\rho_j(\theta)]$. Also, since $\partial/\partial\theta^{\mathrm{t}} \lambda_j(\theta) = \mathbf{\Psi}_j^{\mathrm{t}} \rho_j(\theta)/[1 + \rho_j(\theta)]^2$, then

$$-\frac{\partial^2}{\partial\theta\partial\theta^{\mathrm{t}}} l_1^*(\theta) = \sum_{j=1}^{J} \mathbf{\Psi}_j^{\otimes 2} R_j \left[\frac{\rho_j(\theta)}{1 + \rho_j(\theta)}\right] \left[\frac{1}{1 + \rho_j(\theta)}\right] = \sum_{j=1}^{J} \mathbf{\Psi}_j^{\otimes 2} R_j \lambda_j(\theta)[1 - \lambda_j(\theta)]. \tag{19}$$

Therefore, we have established the following theorem:

**Theorem 2** *Under condition (5), the score vector and observed information matrix associated with the partial likelihood* $\theta \mapsto L_1^*(\theta)$ *are given, respectively, by (18) and (19). In particular, under* $H_0^* : \theta = \mathbf{0}$*, the score vector and information matrix are given, respectively, by*

$$\mathbf{U}_0^*(\mathbf{\Psi}) = \sum_{j=1}^{J} \mathbf{\Psi}_j [O_j - R_j \lambda_j^0] \quad \textit{and} \quad \mathbf{I}_0^*(\mathbf{\Psi}) = \sum_{j=1}^{J} \mathbf{\Psi}_j^{\otimes 2} R_j \lambda_j^0 (1 - \lambda_j^0).$$

The score test statistic for $H_0^* : \theta = \mathbf{0}$ versus $H_1^* : \theta \neq \mathbf{0}$ is

$$S^2(\mathbf{\Psi}) = \left\{ \sum_{j=1}^{J} \mathbf{\Psi}_j [O_j - R_j \lambda_j^0] \right\}^{\mathrm{t}} \left\{ \sum_{j=1}^{J} \mathbf{\Psi}_j^{\otimes 2} R_j \lambda_j^0 (1 - \lambda_j^0) \right\}^{-} \left\{ \sum_{j=1}^{J} \mathbf{\Psi}_j [O_j - R_j \lambda_j^0] \right\} \tag{20}$$

where $\mathbf{A}^-$ is a generalized inverse of $\mathbf{A}$. We remark that $O_j - R_j \lambda_j^0$ has the interpretation of being the observed frequency minus the *conditional* expected frequency at the value $a_j$. Thus, the statistic $\sum_{j=1}^{J} \mathbf{\Psi}_j [O_j - R_j \lambda_j^0]$ is a weighted statistic associated with the observed minus conditionally expected frequencies. Let $\mathbf{O} = (O_1, O_2, \ldots, O_J)^{\mathrm{t}}$ and $\mathbf{E}_0 = (E_1^0, E_2^0, \ldots, E_J^0)^{\mathrm{t}}$, where $E_j^0 = R_j \lambda_j^0, j = 1, 2, \ldots, J$, and let $\mathbf{V}_0$ be the $J \times J$ diagonal matrix whose diagonal elements are $V_{jj}^0 = R_j \lambda_j^0 (1 - \lambda_j^0), j = 1, 2, \ldots, J$. Then, in matrix notation, $\mathbf{U}_0^*(\mathbf{\Psi}) = \mathbf{\Psi}(\mathbf{O} - \mathbf{E}_0)$ and $\mathbf{I}_0^*(\mathbf{\Psi}) = \mathbf{\Psi} \mathbf{V}_0 \mathbf{\Psi}^{\mathrm{t}}$. Therefore, the score statistic in (20) can be re-expressed in matrix form via

$$S^2(\mathbf{\Psi}) = (\mathbf{O} - \mathbf{E}_0)^{\mathrm{t}} \mathbf{\Psi}^{\mathrm{t}} \left( \mathbf{\Psi} \mathbf{V}_0 \mathbf{\Psi}^{\mathrm{t}} \right)^{-} \mathbf{\Psi} (\mathbf{O} - \mathbf{E}_0). \tag{21}$$

In Section 5, we will establish that under $H_0^*$ and certain regularity conditions, as $n \to \infty$, $S^2(\mathbf{\Psi})$ converges in distribution to a central chi-squared distribution with degrees-of-freedom $k_0^* = \mathrm{rank}(\mathcal{I}_0)$, where $\mathcal{I}_0$ is the in-probability limit of $\frac{1}{n}\mathbf{I}_0^*$. Consequently, an asymptotic $\alpha$-level test of $H_0$ is of the form:

$$\text{“Reject } H_0 \text{ whenever } S^2(\mathbf{\Psi}) \geq \chi^2_{k^*;\alpha}\text{”} \tag{22}$$

where $k^*$ is the rank of $\mathbf{I}_0^*$ and $\chi^2_{k^*;\alpha}$ is the $100(1-\alpha)$ percentile of a central chi-squared distribution with degrees-of-freedom $k^*$.

## 4 Specific Choices of $\boldsymbol{\Psi}$

Before proceeding with the asymptotic theory of the test statistic, we first discuss in this section specific choices of the smoothing matrix $\boldsymbol{\Psi}$ and their induced test statistic. We categorize these choices according to whether $p = 1$ or $p > 1$. Those specifications with $p = 1$ could be viewed as directional ones in the sense that the resulting tests will have good power against specific types of alternatives. For our notation, we denote by $\mathcal{J} = \{1, 2, \ldots, J\}$, and for a subset $A \subseteq \mathcal{J}$, we denote by $\mathbf{1}_A$ the $J \times 1$ vector whose $j$th element is $I\{j \in A\}$. We also adopt the convention that $0/0 = 0$.

### 4.1 Specifications with $p = 1$

For $\boldsymbol{\Psi}$-specifications with $p = 1$, the simplest nonrandom choice is $\psi_1 = \mathbf{1}^{\mathrm{t}}_{\mathcal{J}} = (1, 1, \ldots, 1)$. The resulting statistic is

$$S^2(\psi_1) = \frac{[\sum_{j=1}^J (O_j - E_j^0)]^2}{\sum_{j=1}^J R_j \lambda_j^0 (1 - \lambda_j^0)} = \left[\frac{O_\bullet - E_\bullet^0}{\sqrt{V_\bullet^0}}\right]^2, \tag{23}$$

where $O_\bullet = \mathbf{1}^{\mathrm{t}}_{\mathcal{J}} \mathbf{O} = \sum_{j=1}^J O_j$, $E_\bullet^0 = \mathbf{1}^{\mathrm{t}}_{\mathcal{J}} \mathbf{E}_0 = \sum_{j=1}^J E_j^0$, and $V_\bullet^0 = \mathbf{1}^{\mathrm{t}}_{\mathcal{J}} \mathbf{V}_0 \mathbf{1}_{\mathcal{J}} = \sum_{j=1}^J V_{jj}^0 = \sum_{j=1}^J R_j \lambda_j^0 (1 - \lambda_j^0)$. This statistic compares the total number of observed failures ($O_\bullet$) and the total conditionally expected number of failures ($E_\bullet$) over the whole study duration. Note that

$$O_\bullet = \sum_{j=1}^J \sum_{i=1}^n I\{Z_i = a_j, \delta_i = 1\} = \sum_{i=1}^n \left[\sum_{j=1}^J I\{Z_i = a_j, \delta_i = 1\}\right] = \sum_{i=1}^n \delta_i;$$

$$E_\bullet^0 = \sum_{j=1}^J R_j \lambda_j^0 = \sum_{j=1}^J \left[\sum_{i=1}^n I\{Z_i \geq a_j\}\right] \lambda_j^0 = \sum_{i=1}^n \left[\sum_{j=1}^J I\{Z_i \geq a_j\}\right] \lambda_j^0 = \sum_{i=1}^n \sum_{j=1}^{K_i^*} \lambda_j^0$$

where $K_i^* = \max\{j \in \{1, 2, \ldots, J\} : a_j \leq Z_i\}$. Consequently, $O_\bullet - E_\bullet^0 = \sum_{i=1}^n (\delta_i - \sum_{j=1}^{K_i^*} \lambda_j^0)$, which is the numerator of Hyde's (1977, eq. (1)) test statistic. As such, Hyde's gof test for discrete censored data is a special case of the proposed class of tests. It is also interesting to note that Hyde used martingale arguments to obtain the asymptotic distribution of his statistic. In Section 5, we will develop asymptotic properties of the class of statistics via the martingale central limit theorem of Helland (1982).

Another choice, anchored on the intuitive idea of putting more weight on time points with more information, is to let

$$\psi_2 = \sqrt{n}\left(\frac{I\{V_{11}^0 > 0\}}{\sqrt{V_{11}^0}}, \frac{I\{V_{22}^0 > 0\}}{\sqrt{V_{22}^0}}, \ldots, \frac{I\{V_{JJ}^0 > 0\}}{\sqrt{V_{JJ}^0}}\right)^{\mathrm{t}}$$

resulting in the test statistic

$$S^2(\psi_2) = \left\{\frac{1}{\sqrt{J^*}}\sum_{j=1}^{J}\sqrt{R_j}\left[\frac{\hat{\lambda}_j - \lambda_j^0}{\sqrt{\lambda_j^0(1-\lambda_j^0)}}\right]\right\}^2, \tag{24}$$

where $\hat{\lambda}_j = O_j/R_j, j = 1, 2, \ldots, J$, and $J^* = \sum_{j=1}^{J} I\{V_{jj}^0 > 0\}$. Observe that $\psi_2$ is a non-deterministic specification, and the form of the statistic is an average of binomial-type statistics.

A more general choice is to let, for a prespecified nonrandom $\gamma \in \Re$,

$$\psi_3^{\gamma} = \left[\left(\frac{R_1}{n}\right)^{\gamma} I\{R_1 > 0\}, \left(\frac{R_2}{n}\right)^{\gamma} I\{R_2 > 0\}, \ldots, \left(\frac{R_J}{n}\right)^{\gamma} I\{R_J > 0\}\right]^{\mathrm{t}},$$

which leads to test statistics of form

$$S^2(\psi_3^{\gamma}) = \left\{\frac{\sum_{j=1}^{J} I\{R_j > 0\} R_j^{1+\gamma}(\hat{\lambda}_j - \lambda_j^0)}{\sqrt{\sum_{j=1}^{J} I\{R_j > 0\} R_j^{1+2\gamma}\lambda_j^0(1-\lambda_j^0)}}\right\}^2. \tag{25}$$

Note that $S^2(\psi_1) = S^2(\psi_3^0)$, while $S^2(\psi_2)$ is related to $S^2(\psi_3^{-1/2})$. Other values of $\gamma$ of potential interest are $\gamma \in \{-1, 1/2, 1\}$. For instance, $\gamma = -1$ produces

$$S^2(\psi_3^{-1}) = \left\{\frac{\sum_{j=1}^{J} I\{R_j > 0\}(\hat{\lambda}_j - \lambda_j^0)}{\sqrt{\sum_{j=1}^{J} I\{R_j > 0\}\lambda_j^0(1-\lambda_j^0)/R_j}}\right\}^2,$$

a statistic which computes differences between hazard estimates and null hazard values at the $a_j$'s, and then forms an unweighted linear combination of these differences.

## 4.2 Specifications with $p > 1$

The first multidimensional choice provides a generalization of the classical Pearson test statistic. Let $p$ be a positive integer with $p \leq J$, and let $A_1, A_2, \ldots, A_p$ with $A_i \neq \emptyset, i = 1, 2, \ldots, p$, be a (disjoint) partition of $\mathcal{J}$. Define $\mathbf{\Psi}_5$ to be the (nonrandom) $p \times J$ matrix

$$\mathbf{\Psi}_5 = \left[\mathbf{1}_{A_1}, \mathbf{1}_{A_2}, \ldots, \mathbf{1}_{A_p}\right]^{\mathrm{t}}.$$

For $A \subseteq \mathcal{J}$, let $O_\bullet(A) = \mathbf{1}_A^{\rm t}\mathbf{O} = \sum_{j \in A} O_j$, $E_\bullet^0(A) = \mathbf{1}_A^{\rm t}\mathbf{E}_0 = \sum_{j\in A} E_j^0$, and $V_\bullet^0(A) = \mathbf{1}_A^{\rm t}\mathbf{V}_0\mathbf{1}_A = \sum_{j\in A} V_{jj}^0 = \sum_{j \in A} E_j^0(1-\lambda_j^0)$. Then $\mathbf{\Psi}_5$ induces

$$S^2(\mathbf{\Psi}_5) = \sum_{i=1}^{p} \frac{[O_\bullet(A_i) - E_\bullet^0(A_i)]^2}{V_\bullet^0(A_i)}, \tag{26}$$

which is analogous to the Pearson statistic except for the fact that the divisors are $V_\bullet^0(A_i)$'s instead of $E_\bullet^0(A_i)$'s. It should be noted that the expected frequencies $E_\bullet^0(A_i)$'s are *dynamic* expected frequencies in the sense that they are not the expected frequencies for the sets $A_i$'s when viewed at the time origin, but rather are the sums of the *dynamic* expected frequencies at each $a_j, j \in A_i$, with the dynamic expected frequency at $a_j$ being a *conditional* expected frequency given information just prior to $a_j$. A potential obstacle to the use of this specification is the problem of choosing the partition $A_1, A_2, \ldots, A_p$. An interesting special case of the specification $\mathbf{\Psi}_5$, which circumvents the above mentioned problem, is to set $p = J$ and to take $A_i = \{i\}, i = 1, 2, \ldots, J$, so $\mathbf{\Psi}_5$ is the $J \times J$ identity matrix $\mathbf{I}_J$. In such a situation, the test statistic simplifies to

$$S^2(\mathbf{I}_J) = \sum_{j=1}^{J} \frac{(O_j - E_j^0)^2}{E_j^0(1-\lambda_j^0)}, \tag{27}$$

a test statistic which has intuitive appeal because of its simplicity and similarity with the Pearson statistic. However, we will see from the simulation studies that it does not perform well in practical settings.

To describe a more elaborate multidimensional specification, let $\mathbf{R} = (R_1, R_2, \ldots, R_J)^{\rm t}$ and, for $k \in \mathcal{Z}_+$, denote by $\mathbf{R}^k = (R_1^k, R_2^k, \ldots, R_J^k)^{\rm t}$. Given a fixed $p \in \mathcal{Z}_+$ with $p \leq J$, define the $p \times J$ random matrix $\mathbf{\Psi}_6^p$ via

$$\mathbf{\Psi}_6^p = \left[ \left(\frac{\mathbf{R}}{n}\right)^0, \left(\frac{\mathbf{R}}{n}\right)^1, \ldots, \left(\frac{\mathbf{R}}{n}\right)^{p-1} \right]^{\rm t}.$$

The $j$th column of $\mathbf{\Psi}_6^p$ is therefore $\mathbf{\Psi}_j^p = \left(1, R_j/n, (R_j/n)^2, \ldots, (R_j/n)^{p-1}\right)^{\rm t}$. The appeal of this specification is that its rows could serve as (random) polynomial basis for the logarithm of the odds ratios. For $i, i_1, i_2 = 1, 2, \ldots, p$, define

$$U_i^*(\mathbf{\Psi}_6^p) = \left[\left(\frac{\mathbf{R}}{n}\right)^{i-1}\right]^{\mathrm{t}} (\mathbf{O} - \mathbf{E}_0) = \sum_{j=1}^{J} \left(\frac{R_j}{n}\right)^{i-1} (O_j - E_j^0);$$

$$I_{i_1 i_2}^*(\mathbf{\Psi}_6^p) = \left[\left(\frac{\mathbf{R}}{n}\right)^{i_1-1}\right]^{\mathrm{t}} \mathbf{V}_0 \left[\left(\frac{\mathbf{R}}{n}\right)^{i_2-1}\right] = \sum_{j=1}^{J} \left(\frac{R_j}{n}\right)^{i_1+i_2-2} V_{jj}^0.$$

The test statistic generated by $\mathbf{\Psi}_6^p$ becomes

$$S^2(\mathbf{\Psi}_6^p) = \left[(U_i^*(\mathbf{\Psi}_6^p))_{i=1,\ldots,p}\right]^{\mathrm{t}} \left[(I_{i_1 i_2}^*(\mathbf{\Psi}_6^p))_{i_1,i_2=1,\ldots,p}\right]^{-} \left[(U_i^*(\mathbf{\Psi}_6^p))_{i=1,\ldots,p}\right]. \qquad (28)$$

## 5 Asymptotics

In this section we examine the asymptotic properties of the general class of test statistics. For this analysis, we consider a sequence of models, indexed by $n$, so we will have $J = J^{(n)}$, $p = p^{(n)}$, and the $p \times J$ matrices $\mathbf{\Psi} = \mathbf{\Psi}^{(n)}$ with $i$th row vector denoted by $\psi_i = \psi_i^{(n)}, i = 1, 2, \ldots, p$. The $(i, j)$th component of the $\mathbf{\Psi}$ matrix will be denoted by $\Psi_{ij}$. We also have $O_j = O_j^{(n)}$ and $R_j = R_j^{(n)}$. We first establish asymptotic results as $n \to \infty$ but with $J$ fixed at $J_0$. For this purpose, define for $i = 1, 2, \ldots$ and $j = 1, 2, \ldots, J_0$,

$$V_{ij} = I\{Z_i = a_j, \delta_i = 1\} = I\{T_i = a_j, C_i \geq a_j\}; \quad W_{ij} = I\{Z_i \geq a_j\} = I\{T_i \geq a_j, C_i \geq a_j\},$$

and

$$\mathcal{F}_j = \bigvee_{i=1}^{n} \sigma\left\{W_{i1}, V_{i1}, W_{i2}, \ldots, V_{ij}, W_{ij+1}\right\}.$$

To shorten notation, let $\mathbf{E}_{j-1}[\cdot] = \mathbf{E}[\cdot|\mathcal{F}_{j-1}]$, $\mathbf{Var}_{j-1}[\cdot] = \mathbf{Var}[\cdot|\mathcal{F}_{j-1}]$ and $\mathbf{Cov}_{j-1}[\cdot,\cdot] = \mathbf{Cov}[\cdot,\cdot|\mathcal{F}_{j-1}]$. Furthermore, for $i = 1, 2, \ldots, n$, let $\mathbf{U}_i = (U_{i1}, U_{i2}, \ldots, U_{iJ_0})^{\mathrm{t}}$ where $U_{ij} = V_{ij} - \lambda_j^0 W_{ij}$. Then we have the representation $\mathbf{O} - \mathbf{E}_0 = \sum_{i=1}^n \mathbf{U}_i$. Consequently, if $\mathbf{\Psi} = [\mathbf{\Psi}_1, \mathbf{\Psi}_2, \ldots, \mathbf{\Psi}_{J_0}]$ is the $p \times J_0$ (possibly random) smoothing matrix, then

$$\mathbf{\Psi}(\mathbf{O} - \mathbf{E}_0) = \sum_{i=1}^{n} [\mathbf{\Psi}\mathbf{U}_i] = \sum_{i=1}^{n} \left[\sum_{j=1}^{J_0} \mathbf{\Psi}_j U_{ij}\right].$$

Under $H_0$, the true sequence of hazard probabilities is denoted by $\{\lambda_j^0 : j = 1, 2, 3, \ldots\}$. Also, define the $J_0 \times J_0$ random matrix $\mathbf{V}_i^0$ via

$$\mathbf{V}_i^0 = \text{Diag}\left(W_{i1}\lambda_1^0(1-\lambda_1^0), \ldots, W_{iJ_0}\lambda_{J_0}^0(1-\lambda_{J_0}^0)\right).$$

Since $R_j = \sum_{i=1}^n W_{ij}$, then $\mathbf{V}_0 = \sum_{i=1}^n \mathbf{V}_i^0$.

**Lemma 2** *If $\mathbf{\Psi}$ is such that for each $i = 1, \ldots, n$ and $j = 1, \ldots, J_0$, $\mathbf{\Psi}_j$ is measurable with respect to $\mathcal{F}_{j-1}$ and $\mathbf{E}[(\Psi_{ij}U_{ij})^2] < \infty$, then under $H_0$, $\mathbf{E}[\mathbf{\Psi}\mathbf{U}_i] = \mathbf{0}$ and $\mathbf{Cov}\{\mathbf{\Psi}\mathbf{U}_i, \mathbf{\Psi}\mathbf{U}_i\} = \mathbf{E}\left[\mathbf{\Psi}\mathbf{V}_i^0\mathbf{\Psi}^{\mathrm{t}}\right]$.*

**Proof:** See Appendix A. $\|$

Because of the representation $\mathbf{\Psi}(\mathbf{O}-\mathbf{E}_0) = \sum_{i=1}^n(\mathbf{\Psi}\mathbf{U}_i)$, if $\mathbf{\Psi}$ is non-random, it follows by the Central Limit Theorem for independent and identically distributed random vectors that, under $H_0$, $\frac{1}{\sqrt{n}}\mathbf{\Psi}(\mathbf{O}-\mathbf{E}_0) \xrightarrow{\mathrm{d}} N_p(\mathbf{0}, \mathbf{\Xi})$, where

$$\mathbf{\Xi} = \mathbf{\Psi}\mathbf{\Sigma}_0\mathbf{\Psi}^{\mathrm{t}} \quad \text{and} \quad \mathbf{\Sigma}_0 = \mathbf{E}\{\mathbf{V}_0\} = \text{Diag}\{\mathbf{P}(Z_1 \geq a_j)\lambda_j^0(1-\lambda_j^0), j = 1, 2, \ldots, J_0\}.$$

We formalize this result as a theorem.

**Theorem 3** *If $\mathbf{\Psi}$ is such that for each $i = 1, \ldots, n$ and $j = 1, \ldots, J_0$, $\mathbf{\Psi}_j$ is measurable with respect to $\mathcal{F}_{j-1}$ and $\mathbf{E}[(\Psi_{ij}U_{ij})^2] < \infty$, then under $H_0$, as $n \to \infty$, $n^{-1/2}\mathbf{\Psi}(\mathbf{O}-\mathbf{E}_0)$ converges in distribution to a $p$-dimensional normal with mean vector $\mathbf{0}$ and covariance matrix $\mathbf{\Xi} = \mathbf{\Psi}\mathbf{\Sigma}_0\mathbf{\Psi}$.*

To be able to use the above asymptotic result for testing purposes, we need to be able to estimate the matrix $\mathbf{\Sigma}_0$. An obvious consistent estimator of $\mathbf{\Sigma}_0$ is $\frac{1}{n}\mathbf{V_0}$, so consequently, under $H_0$ and the additional assumptions imposed above to obtain Theorem 3, we have:

$$S^2(\mathbf{\Psi}) = (\mathbf{O}-\mathbf{E}_0)^{\mathrm{t}}\mathbf{\Psi}^{\mathrm{t}}\left(\mathbf{\Psi}\mathbf{V}_0\mathbf{\Psi}^{\mathrm{t}}\right)^{-}\mathbf{\Psi}(\mathbf{O}-\mathbf{E}_0) \xrightarrow{\mathrm{d}} \chi^2_{k^*},$$

with $k^*$ being the rank of $\mathbf{\Xi}$.

Note however that the asymptotic result in Theorem 3 is restricted on two grounds: first, we had to assume that the smoothing matrix $\mathbf{\Psi}$ is nonrandom, and second, we had to fix the value of $J$ to be $J_0$ so in particular the result does *not* cover the case where $J$ increases as $n$ increases. The following theorem is more general in that it covers the case where the smoothing matrix may be random or when $J$ changes with $n$, but with $p$ remaining fixed. For notation, given a square matrix $\mathbf{A}$ of order $p$, we let $\text{trace}(\mathbf{A}) = \sum_{i=1}^{p} a_{ii}$.

**Theorem 4** *If the conditions of Lemma 2 hold, and if in addition, $p$ does not change with $n$, and there exists a $p \times p$ positive definite matrix $\mathbf{\Sigma}_0$ such that, as $n \to \infty$,*

*(i)* $\frac{1}{n}\mathbf{I}_0^* = \frac{1}{n}\mathbf{\Psi}\mathbf{V}_0\mathbf{\Psi}^{\mathrm{t}} \xrightarrow{\mathrm{pr}} \mathbf{\Xi}_0$;

*(ii)* $\max_{1\leq j\leq J} \text{trace}\left\{\left(\mathbf{\Psi}\mathbf{V}_0\mathbf{\Psi}^{\mathrm{t}}\right)^{-1}\left(\mathbf{\Psi}_j V_{jj}^0 \mathbf{\Psi}_j^{\mathrm{t}}\right)\right\} \xrightarrow{\mathrm{pr}} 0$*; and*

*(iii)* $\max_{1\leq j\leq J} ||\mathbf{\Psi}_j||^2 = O_p(1)$,

*then* $\frac{1}{\sqrt{n}}\mathbf{U}_0^* = \frac{1}{\sqrt{n}}\mathbf{\Psi}(\mathbf{O}-\mathbf{E}) \xrightarrow{\mathrm{d}} N_p(\mathbf{0}, \mathbf{\Xi}_0)$, *so* $S^2(\mathbf{\Psi}) = (\mathbf{O}-\mathbf{E}_0)^{\mathrm{t}}\mathbf{\Psi}^{\mathrm{t}}\left(\mathbf{\Psi}\mathbf{V}_0\mathbf{\Psi}^{\mathrm{t}}\right)^{-}\mathbf{\Psi}(\mathbf{O}-\mathbf{E}_0)$ *converges in distribution to a (central) chi-square distribution with $k^*$ degrees-of-freedom.*

**Proof:** See Appendix A. $\|$

We remark that condition (ii) of Theorem 4 is *not* a necessary condition for the convergence in distribution to occur. Indeed, this condition seems strong as it usually requires $J \to \infty$ whenever $n \to \infty$. So far, we have not been able to weaken this condition.

# 6 Asymptotics for the Specific $\mathbf{\Psi}$'s

In this section we apply the asymptotic results in the preceding section to the specific choices of $\mathbf{\Psi}$ mentioned in Section 4. We note that each of these choices of $\mathbf{\Psi}$ satisfy the measurability condition, that is, that $\mathbf{\Psi}_j$ is measurable with respect to $\mathcal{F}_{j-1}$.

For the choice $\psi_1 = \mathbf{1}_{\mathcal{J}}^{\mathrm{t}}$ leading to the test statistic (23), if $J$ is fixed, then Theorem 3 immediately applies. On the other hand, if we allow $J$ to vary with $n$, then condition (i) of

Theorem 4 becomes $\frac{1}{n}\sum_{j=1}^{J} R_j\lambda_j^0(1-\lambda_j^0) \xrightarrow{\text{pr}} \sigma_0^2$ for some positive and finite $\sigma_0^2$; while condition (ii) of Theorem 4 becomes

$$\frac{\max_{1\le j\le J} R_j\lambda_j^0(1-\lambda_j^0)}{\sum_{l=1}^{J} R_l\lambda_l^0(1-\lambda_l^0)} \xrightarrow{\text{pr}} 0. \tag{29}$$

This last condition forces $J$ to increase to infinity as $n$ increases to infinity since a lower bound for the quantity in the left-hand side of (29) is $1/J$. For specific situations regarding the $\lambda_j^0$'s and the censoring mechanism, these conditions need to be verified.

For the choice $\psi_2$ leading to the statistic (24), Theorem 3 does not cover this situation because $\mathbf{\Psi}$ is random. Resorting to Theorem 4, with $J^* = \sum_{j=1}^{J} I\{V_{jj}^0 > 0\}$, conditions (i) and (ii) of this theorem become $J^*/n \xrightarrow{\text{pr}} \sigma_0^2 \in (0,\infty)$ and $J^* \xrightarrow{\text{pr}} \infty$. In particular, this implies that $J \to \infty$ as $n \to \infty$.

The specification $\psi_3^\gamma$ leading to the test statistic in (25) is also not covered by Theorem 3. Applying Theorem 4, the sufficient conditions for asymptotic normality that need to be checked are

$$\frac{1}{n}\sum_{j=1}^{J} I\{R_j > 0\}\left(\frac{R_j}{n}\right)^{2\gamma} R_j\lambda_j^0(1-\lambda_j^0) \xrightarrow{\text{pr}} \sigma_0^2 \in (0,\infty);$$

$$\frac{\max_{1\le j\le J} I\{R_j > 0\}R_j^{1+2\gamma}\lambda_j^0(1-\lambda_j^0)}{\sum_{l=1}^{J} I\{R_l > 0\}R_l^{1+2\gamma}\lambda_l^0(1-\lambda_l^0)} \xrightarrow{\text{pr}} 0.$$

For the specification $\mathbf{\Psi}_5$, which induces the Pearson-type statistic in (26), when $J$ is fixed and the partition is also fixed, then Theorem 3 applies to yield the asymptotic normality. This asymptotic normality holds in particular for the Pearson-type test statistic in (27). When $J$ changes with $n$, then by Theorem 4, the resulting sufficient conditions for asymptotic normality are, for each $i = 1, 2, \ldots, p$,

$$\frac{1}{n}V_\bullet^0(A_i) \xrightarrow{\text{pr}} \sigma_i^2 \in (0,\infty) \quad \text{and} \quad \frac{\max_{j\in A_i} V_{jj}^0}{V_\bullet^0(A_i)} \xrightarrow{\text{pr}} 0,$$

where we recall that for $A \subseteq \mathcal{J}$, $V_\bullet^0(A) = \sum_{j\in A} V_{jj}^0$.

For the polynomial-type specification $\mathbf{\Psi}_6$ leading to (28), not much simplification in notation could be achieved relative to the statements of conditions (i) and (ii) of Theorem 4, and the actual verification in specific situations may also be hard.

## 7 Tests for the Geometric Distribution

The geometric distribution is the discrete analog of the exponential distribution and is a common model for discrete failure-time data. For a failure-time variable $T$ which has a geometric distribution with success probability $\eta$, its probability and distribution functions are given, respectively, by

$$p(j|\eta) = (1-\eta)^{j-1}\eta, j = 1, 2, \ldots, \quad \text{and} \quad F(j|\eta) = 1 - (1-\eta)^j, j = 1, 2, \ldots.$$

The hazard rates are therefore $\lambda(j|\eta) = \eta, j = 1, 2, \ldots$. Suppose it is desired to test the hypothesis that the failure-time distribution is geometric with success probability $\eta = \eta_0$ based on a right-censored data $(Z_1, \delta_1), (Z_2, \delta_2), \ldots, (Z_n, \delta_n)$ with $Z_i = \min(T_i, C_i \wedge J_0), i = 1, 2, \ldots, n$, where $C_1, C_2, \ldots, C_n$ are censoring variables, and $J_0$ is the upper limit of the observation period. For this testing problem, we could apply the tests presented in the preceding sections. Under this geometric setting, we write down the forms of these statistics below, without going into the algebraic simplifications. These are the test statistics considered in the simulation study in the next section to ascertain the levels and powers of these gof tests.

The test statistic induced by $\psi_1$ reduces, under this geometric setting, to the global 'binomial-type' statistic given by

$$S^2(\psi_1) = \left[\frac{(O_\bullet/R_\bullet - \eta_0)}{\sqrt{\eta_0(1-\eta_0)/R_\bullet}}\right]^2,$$

where $O_\bullet$ is the total number of observed failures, and $R_\bullet$ may be viewed as the total exposure time. Thus, this statistic amounts to comparing the global estimate of the rate of failure and $\eta_0$, and also recall that this coincides with Hyde's (1977) statistic. In contrast, the statistic induced by $\psi_2$ becomes an average of statistics based on the hazard rates at each time point and is given by

$$S^2(\psi_2) = \left\{\frac{1}{\sqrt{J^*}}\sum_{j=1}^{J_0} I\{R_j > 0\}\left[\frac{\hat{\lambda}_j - \eta_0}{\sqrt{\eta_0(1-\eta_0)/R_j}}\right]\right\}^2, \tag{30}$$

where $\hat{\lambda}_j = O_j/R_j, j = 1, 2, \ldots, J_0$, and $J^* = \sum_{j=1}^{J_0} I\{R_j > 0\}$. Two special cases, corresponding to $\gamma = 1$ and $\gamma = -1$, of the class of statistics induced by $\psi_3^\gamma$ are

$$S^2(\psi_3^1) = \left\{ \frac{\sum_{j=1}^{J_0} R_j^2(\hat{\lambda}_j - \eta_0)}{\sqrt{\sum_{j=1}^{J_0} R_j^3 \eta_0(1-\eta_0)}} \right\}^2 \quad \text{and} \quad S^2(\psi_3^{-1}) = \left\{ \frac{\sum_{j=1}^{J_0} I\{R_j > 0\}(\hat{\lambda}_j - \eta_0)}{\sqrt{\sum_{j=1}^{J_0} I\{R_j > 0\}\eta_0(1-\eta_0)/R_j}} \right\}^2.$$

For the specification $\mathbf{\Psi}_5 = [\mathbf{1}_{A_1}, \mathbf{1}_{A_2}, \ldots, \mathbf{1}_{A_p}]^{\mathrm{t}}$ where $A_1, A_2, \ldots, A_p$ is a partition of $\mathcal{J}_0$, the resulting test statistic, under this geometric setting, becomes

$$S^2[\mathbf{\Psi}_5] = \left[\frac{1}{1-\eta_0}\right] \left[\sum_{i=1}^{p} \frac{[O_\bullet(A_i) - R_\bullet(A_i)\eta_0]^2}{R_\bullet(A_i)\eta_0}\right], \tag{31}$$

which is almost the Pearson statistic. In particular, the statistic in (27) is given by

$$S^2(I_{\mathcal{J}_0}) = \sum_{j=1}^{J_0} I\{R_j > 0\} \left[\frac{\hat{\lambda}_j - \eta_0}{\sqrt{\eta_0(1-\eta_0)/R_j}}\right]^2. \tag{32}$$

The statistic induced by the specification $\mathbf{\Psi}_6^p$ simplifies, under this geometric setting, to

$$S^2(\mathbf{\Psi}_6^p) = \frac{n}{\eta_0(1-\eta_0)} \left[ \left( \sum_{j=1}^{J_0} \left(\frac{R_j}{n}\right)^i (\hat{\lambda}_j - \eta_0) \right)_{i=1,2,\ldots,p} \right]^{\mathrm{t}} \times$$
$$\left[ \left( \sum_{j=1}^{J_0} \left(\frac{R_j}{n}\right)^{i_1+i_2-1} \right)_{i_1,i_2=1,2,\ldots,p} \right]^{-} \left[ \left( \sum_{j=1}^{J_0} \left(\frac{R_j}{n}\right)^i (\hat{\lambda}_j - \eta_0) \right)_{i=1,2,\ldots,p} \right].$$

Note that $S^2(\mathbf{\Psi}_6^1)$ coincides with $S^2(\psi_1)$. In the simulation studies, we set $p = 1, 2, 3, 4$.

An existing problem that needs to be addressed in future research is a dynamic, i.e., data-dependent, method for determining the smoothing order $p$. For the Neyman formulation of smooth gof test with complete data, inroads into this problem have been obtained by Ledwina (1994) and Kallenberg and Ledwina (1995). See also subsection 7.6.1 of Hart (1997) for discussion on the Mallows and Schwarz criteria for selecting $p$ in this Neyman formulation. However, for the intensity-based formulation with continuous data, and for the method for discrete data proposed in the present paper, this order selection problem calls for a resolution.

## 8 Simulation Studies

To examine the performance and acceptability of the asymptotic approximations for some of the members of the proposed class of smooth goodness-of-fit tests, we performed simulation studies under the geometric null hypothesis. The tests are those associated with the statistics $S^2(\psi_1)$, $S^2(\psi_2)$, $S^2(\psi_3^1)$, $S^2(\psi_3^{-1})$, $S^2(\mathbf{\Psi}_5)$, $S^2(\mathcal{I}_{J_0})$, and $S^2(\mathbf{\Psi}_6^p)$ for $p = 1, 2, 3, 4$. The expressions of these statistics were explicitly given or described in the preceding section. Also, recall that $S^2(\psi_1)$ and $S^2(\mathbf{\Psi}_6^1)$ coincide. For the statistic $S^2(\mathbf{\Psi}_5)$ with a general partition $A_1, A_2, \ldots, A_p$, we choose the following four partitions of $\mathcal{J}_0$: for $S^2(\mathbf{\Psi}_5^1)$, $A_1 = \{2j-1 : j = 1, 2, \ldots, [J_0/2]\}, A_2 = \{2j : j = 1, 2, \ldots, [J_0/2]\}$; for $S^2(\mathbf{\Psi}_5^2)$, $A_1 = \{j : j = 1, 2, \ldots, [J_0/2]\}, A_2 = \{j : j = [J_0/2] + 1, \ldots, J_0\}$; for $S^2(\mathbf{\Psi}_5^3)$, $A_1 = \{j : j = 1, 2, \ldots, [J_0/3]\}, A_2 = \{j : j = [J_0/3] + 1, \ldots, 2[J_0/3]\}, A_3 = \{j : j = 2[J_0/3] + 1, \ldots, J_0\}$; and for $S^2(\mathbf{\Psi}_5^4)$, $A_1 = \{j : j = 1, 2, \ldots, [J_0/4]\}, A_2 = \{j : j = [J_0/4]+1, \ldots, 2[J_0/4]\}, A_3 = \{j : j = 2[J_0/4]+1, \ldots, 3[J_0/4]\}, A_4 = \{j : j = 3[J_0/4]+1, \ldots, J_0\}$.

### 8.1 Achieved Levels of Tests

For the simulation study to assess the achieved levels of the tests, the failure times $T_1, T_2, \ldots, T_n$ were generated according to a geometric distribution with mean $1/\eta_0$. The censoring variables $C_1, C_2, \ldots, C_n$ were also generated according to a geometric distribution with parameter chosen so the probability of an uncensored observation (UCP), $\mathbf{P}\{T_1 \leq C_1\}$, was equal to a specified value of UCP. We also specified $J_0$, the upper limit of the observation period, so the effective probability of an uncensored observation is $\mathbf{P}\{T_1 \leq C_1 \wedge J_0\}$. The parameters of the simulation runs were therefore $n$, $\eta_0$, UCP, $J_0$, and M, the number of replications, which was set to 1000 for all the runs. The simulation program was coded in `S-Plus`. As outputs of each run, we obtained the simulated mean and variance of the test statistic, the proportion of uncensored failure times, the proportion of rejections of $H_0$ by the (asymptotic) 5% level tests, which is an estimate of the achieved level of the 5% level test, as well as that for the (asymptotic) 10% level tests. We present in Tables 1, 2, 3, and 4 the achieved levels of the 5% asymptotic test for different combinations of the simulation parameters with $n \in \{30, 50, 100, 200\}$, $J_0 \in \{30, 50\}$, $\eta_0 \in \{.03, .10\}$, and UCP $\in \{.50, .75\}$. The presentation of the results was designed to reveal effects of the sample size on the level of the tests.

Examining Tables 1, 2, 3, and 4, we immediately notice that the intuitive-looking Pearson-type test statistic $S^2(\mathcal{I}_{J_0})$ is extremely anticonservative even when $n = 200$ which makes it practically useless. For the other test statistics, when $\eta_0 = .03$, $n = 100$, $J_0 = 50$, and UCP $=$ .75, the achieved levels of the tests are consistent (i.e., to within two standard errors) with the nominal asymptotic level of .05, except for $S^2(\mathbf{\Psi}_6^3)$ whose achieved level is 6.9%. Looking into the other information from the simulation outputs for this particular test, the achieved level of the (asymptotic) 10% level test was 10.9%, which is within the nominal value, while the simulated mean and variance were 3.12 and 7.31, respectively, which are a bit larger than the mean and variance of a chi-squared distribution with three degrees-of-freedom. Other runs for the $S^2(\psi_6^3)$-based test at $n = 100$, $J_0 = 50$, $\eta_0 = .03$, and UCP $= .75$ yielded achieved levels [e.g., 4.9% for the 5%-level test and 10.5% for the 10%-level test] consistent with the nominal values, so we are inclined to conclude that at $n = 100$, $J_0 = 50$ and $\eta_0 = .03$, the test associated with $S^2(\mathbf{\Psi}_6^3)$ has an acceptable Type I error rate. Because of the close agreement between the achieved and nominal levels for most of the tests when the simulation parameter values are $n = 100$, $J_0 = 50$, $\eta_0 = .03$, and UCP $= .75$, the power simulations were performed at these values.

When the sample size is either $n = 30$ or $n = 50$, and with $J_0 = 50$, we notice that the tests based on $S^2(\mathbf{\Psi}_6^3)$ and $S^2(\mathbf{\Psi}_6^4)$ are a bit anticonservative. The anticonservatism of most of the tests becomes very apparent when $J_0 = 30$ and when the observed proportions of censored failure times are high. When $\eta_0$ is changed from .03 to .10, which translates into a smaller mean for the failure times, the tests tend to be a bit more anticonservative as evidenced by looking at Table 3 with $J_0 = 50$ and $n = 100$. This anticonservatism could be due to the fact that under this situation, in contrast to the case where $\eta_0 = .03$, the numbers at risks in the larger times tend to be small or become zero, hence there is very little or no information at these times. In particular, the Pearson-type statistics $S^2(\mathbf{\Psi}_5^p), p = 2, 3, 4$, which are based on equal partitioning of $\mathcal{J}_0$ into 2, 3, and 4 groups, respectively, are drastically affected by at-risk sets at these larger times becoming small or empty.

## 8.2 Achieved Powers of the Tests

To examine the powers of these tests, we considered several alternatives to see which type of alternative a particular test is sensitive, and to see which tests could be considered as omnibus in the sense of having good power against a wide variety of alternatives. The null hypothesis tested in these power simulations was the geometric distribution with mean of $1/\eta_0$ where $\eta_0 = .03$, except for the case when the alternative is Poisson in which case we took $\eta_0 = .10$. The censoring variables were taken to be geometrically distributed with parameter chosen to achieve a UCP $= .75$ under the null hypothesis. As in the level simulation, the effective proportion of uncensored observations will be lower than the UCP because of the effect of $J_0$, the upper limit of the observation period. In these power simulations, $n = 100$ and $J_0 = 50$, since for these parameters, the tests, except $S^2(\mathcal{I}_{J_0})$, achieve the nominal asymptotic level. Though the test based on $S^2(\mathcal{I}_{J_0})$ was also included in the power simulation, we did not include their results in the summary tables since it did not have the appropriate level.

The first set of alternatives was that the failure times were generated according to a geometric distribution with parameter

$$\eta = 2\eta_0 \frac{\exp(c)}{1 + \exp(c)}$$

where $c \in \{-1, -.5, -.25, -.1, 0, .1, .25, .5, 1\}$ with $c = 0$ coinciding with the null hypothesis. The simulated powers of the different tests are summarized in Table 5. Examining this table, we note that the test based on $S^2(\psi_1) = S^2(\mathbf{\Psi}_6^1)$ is the most sensitive for detecting these geometric alternatives, followed closely by the tests based on $S^2(\psi_2)$ and $S^2(\psi_3^1)$. The achieved powers of the tests based on $S^2(\mathbf{\Psi}_5)$, $S^2(\mathbf{\Psi}_6^2)$, and $S^2(\mathbf{\Psi}_6^3)$ are also good, but clearly dominated by the tests based on $S^2(\mathbf{\Psi}_6^1)$ and $S^2(\mathbf{\Psi}_6^2)$. The test based on $S^2(\psi_3^{-1})$ is good when $c < 0$ but deteriorates slightly when $c > 0$.

The second set of alternatives was the negative binomial distribution. More precisely, the failure times were generated according to $T_i = Y_i + 1, i = 1, 2, \ldots, n$, where the $Y_i$'s are distributed according to

$$\mathbf{P}\{Y = y|r, p\} = \binom{y + r - 1}{r - 1} (1 - p)^y p^r, \quad y = 0, 1, 2, \ldots.$$

The parameters $(r, p)$ were set to $(2, .03)$, $(2, .04)$, $(2, .05)$, $(2, .0583)$, $(2, .06)$, $(2, .07)$, and $(2, .08)$. The parameter value $(2, .0583)$ leads to the same mean failure time as the null specification. We also ran the simulation for $r = 3$, but the results were uninteresting as almost all the tests yielded powers of 100%. Table 6 summarizes the observed powers of the tests for this negative binomial alternative. Evidently, the tests based on $S^2(\mathbf{\Psi}_6^p)$ for $p = 2, 3, 4$, are best for detecting this alternative. The tests based on $S^2(\mathbf{\Psi}_5^p), p = 2, 3, 4$, performed well, though not as well as those based on $S^2(\mathbf{\Psi}_6^p), p = 2, 3, 4$. The other tests all have low powers for at least one of the choices of $(r, p)$. In particular, notice that the tests $S^2(\psi_1) = S^2(\mathbf{\Psi}_6^1)$ and $S^2(\psi_2)$ which were excellent for the geometric alternatives have very low powers when the mean failure times is close to the mean under the null specification. This makes us surmise that these particular tests are good for detecting a location or mean change but not other types of alternatives. This suspicion was further strengthened when we changed the alternatives to Poisson-type probabilities. For this set of alternatives, the failure times were generated according to $T_i = V_i + 1, i = 1, 2, \ldots, n$, where the $V_i$'s have probabilities given by

$$\mathbf{P}\{V = k|\lambda\} = \frac{\exp(-\lambda)\lambda^k}{k!}, \quad k = 0, 1, 2, \ldots.$$

We decided to set $\eta_0 = .10$ since when we used $\eta_0 = .03$, most of the tests achieved 100% power, which does not allow comparison of their relative strengths. The values of $\lambda$ considered in the simulation were chosen to be $\lambda \in \{7, 8, 9, 10, 11\}$, with $\lambda = 9$ leading to the same mean as the null specification. The observed powers of the tests are summarized in Table 7. From this table we note that the tests associated with $S^2(\mathbf{\Psi}_6^p), p = 2, 3, 4$, are excellent, as well as those based on $S^2(\psi_3^1)$ and $S^2(\psi_3^{-1})$. The other tests, including the Pearson-type statistics, have very low powers for some of the chosen values of $\lambda$.

The next set of alternatives differs from the previous ones in that instead of using a well-known discrete distribution, we specified the hazard sequence according to the following function: $\lambda_j = 2\eta_0 G_j(\mathbf{a})/[1 + G_j(\mathbf{a})], j = 1, 2, \ldots, J_0$, where with $q$ being the dimension of the vector $\mathbf{a}$,

$$G_j(\mathbf{a}) = \exp\left\{\sum_{k=1}^{q} a_k \left(\frac{j}{J_0}\right)^{k-1}\right\}. \tag{33}$$

We refer to this class of alternatives as *polynomially*-generated. We chose four values of $\mathbf{a}$ to achieve different shapes of the hazard sequence. These hazard sequences are plotted (plots were made continuous for aesthetic purposes) in Table 8, with this table also containing a summary of the observed powers of the tests. We note in passing that in generating the failure times, we cannot simply call a pre-existing object in the `S-Plus` libraries. Rather, each failure time was generated according to the following `S-Plus` pseudocode, with Ber($p$) representing the Bernoulli random generator with success probability of $p$ [we used the `S-Plus` object `rbinom`$(1, 1, p)$ for this purpose]: M $\leftarrow$ 0; j $\leftarrow$ 1; `while`(M = 0) {R $\leftarrow$ Ber($\lambda_j$); `if` (R == 1) {T $\leftarrow$ j; M $\leftarrow$ 1}; `else` {j $\leftarrow$ j+1} }; `return`(T). Obviously, for the four chosen values of the vector $\mathbf{a}$, the tests based on $S^2(\mathbf{\Psi}_6^p), p = 2, 3, 4$, dominated the other tests. The tests based on $S^2(\mathbf{\Psi}_5^p), p = 3, 4$, also performed acceptably but not as well as the those based on $S^2(\mathbf{\Psi}_6^p)$.

The last set of alternatives is similar to the preceding specification except that we used trigonometric functions to define the function $G_j$ above. For this set of alternatives, we set

$$G_j(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \exp\left\{\sum_{k=1}^{q} [a_k \sin(2\pi c_k(j/J_0)) + b_k \cos(2\pi c_k(j/J_0))]\right\}, \tag{34}$$

and refer to the resulting alternatives as *trigonometrically*-generated. We chose four sets of values of $\mathbf{a}, \mathbf{b}, \mathbf{c}$, with the values chosen so as to obtain varied sequences of the $\lambda_j$'s. The choices are summarized in Table 9, with the corresponding hazard sequences also plotted. We observe from this table that the test based on $S^2(\mathbf{\Psi}_6^4)$, followed by that based on $S^2(\mathbf{\Psi}_6^3)$, performed very well. In particular, notice that the domination of $S^2(\mathbf{\Psi}_6^4)$ becomes very pronounced when the hazard sequence plot has high frequency. It is surprising to observe, however, that the test based on $S^2(\mathbf{\Psi}_5^4)$ also performed very well and was not dominated by $S^2(\mathbf{\Psi}_6^4)$. In fact, for two sets of choices of the alternative parameters $(\mathbf{a}, \mathbf{b}, \mathbf{c})$, the test based on $S^2(\mathbf{\Psi}_5^4)$ had better power performance than the test based on $S^2(\mathbf{\Psi}_6^4)$.

As a way of concisely summarizing the power properties of these tests under the five classes of alternatives considered in the simulations, we present in Table 10 'letter grades' for these tests, with a grade of A indicating very good power, while a grade of E indicating very poor power. In assigning the ratings, if a test did poorly in at least one of the choices within an alternative class, then it is assigned a poor rating even though it has good power against the other choices for that particular class of alternatives. On the basis of this summary, unless one has specific knowledge of the type of alternative that could have generated the data if the null hypothesis is false, it appears prudent to utilize either of the tests based on $S^2(\mathbf{\Psi}_6^p)$ for $p = 2, 3, 4$, with higher values of $p$ preferred when the hazard sequence is expected to have high frequency. The tests based on unidimensional smoothing functions may have good powers against specific types of departures from the null hypothesis, but they have poor performances for other types of alternatives; while those based on $S^2(\mathbf{\Psi}_6^p)$ with $p = 2, 3, 4$, though not always achieving the highest powers, always have competitive powers for all the alternative classes considered. It also appears that even though the Pearson-type tests have decent powers, they are still bettered by those based on the $\mathbf{\Psi}_6$-specification.

# A Appendix: Proofs

**Proof of Lemma 2:** First, note that $\mathbf{E}_{j-1}\{U_{ij}\} = \mathbf{E}_{j-1}\{V_{ij}\} - \lambda_j^0 W_{ij} = 0$ since $\mathbf{E}_{j-1}\{V_{ij}\} = W_{ij}\mathbf{E}\{V_{ij}|T_i \geq a_j, C_i \geq a_j\} = W_{ij}\lambda_j^0$ by virtue of the independent censoring condition (5). Using the same reasoning, note also that $\mathbf{Var}_{j-1}\{U_{ij}\} = W_{ij}\lambda_j^0(1-\lambda_j^0)$, while for $j_1 < j_2$,

$\mathbf{Cov}_{j_1-1}\{U_{ij_1}, U_{ij_2}\} = 0$. To see this result, we have $\mathbf{Cov}_{j_1-1}\{U_{ij_1}, U_{ij_2}\} = \mathbf{E}_{j_1-1}\{U_{ij_1}U_{ij_2}\} = \mathbf{E}_{j_1-1}\{U_{ij_1}\mathbf{E}_{j_2-1}\{U_{ij_2}\}\} = 0$.

To establish the first assertion of the lemma, we have: $\mathbf{E}\left\{\sum_{j=1}^{J_0} \mathbf{\Psi}_j U_{ij}\right\} = \sum_{j=1}^{J_0} \mathbf{E}\{\mathbf{\Psi}_j U_{ij}\} = \sum_{j=1}^{J_0} \mathbf{E}\{\mathbf{E}_{j-1}\{\mathbf{\Psi}_j U_{ij}\}\} = \sum_{j=1}^{J_0} \mathbf{E}\{\mathbf{\Psi}_j \mathbf{E}_{j-1}\{U_{ij}\}\} = 0$, where we used the fact that $\mathbf{\Psi}_j \in \mathcal{F}_{j-1}$ and the results established in the preceding paragraph. On the other hand, note that

$$\mathbf{Cov}\left\{\sum_{j=1}^{J_0} \mathbf{\Psi}_j U_{ij}\right\} = \sum_{j=1}^{J_0} \mathbf{Var}\{\mathbf{\Psi}_j U_{ij}\} + 2\sum_{j_1<j_2} \mathbf{Cov}\{\mathbf{\Psi}_{j_1} U_{ij_1}, \mathbf{\Psi}_{j_2} U_{ij_2}\}.$$

But

$$\begin{aligned}
\mathbf{Var}\{\mathbf{\Psi}_j U_{ij}\} &= \mathbf{E}\{\mathbf{Var}_{j-1}\{\mathbf{\Psi}_j U_{ij}\}\} + \mathbf{Var}\{\mathbf{E}_{j-1}\{\mathbf{\Psi}_j U_{ij}\}\} \\
&= \mathbf{E}\{\mathbf{\Psi}_j W_{ij}\lambda_j^0(1-\lambda_j^0)\mathbf{\Psi}^{\mathrm{t}}\} + \mathbf{Var}\{\mathbf{\Psi}_j \mathbf{E}_{j-1}\{U_{ij}\}\} = \mathbf{E}\{\mathbf{\Psi}_j W_{ij}\lambda_j^0(1-\lambda_j^0)\mathbf{\Psi}^{\mathrm{t}}\}.
\end{aligned}$$

For $j_1 < j_2$,

$$\begin{aligned}
\mathbf{Cov}\{\mathbf{\Psi}_{j_1} U_{ij_1}, \mathbf{\Psi}_{j_2} U_{ij_2}\} &= \mathbf{E}\{\mathbf{Cov}_{j_1-1}\{\mathbf{\Psi}_{j_1} U_{ij_1}, \mathbf{\Psi}_{j_2} U_{ij_2}\}\} + \mathbf{Cov}\{\mathbf{E}_{j_1-1}\{\mathbf{\Psi}_{j_1} U_{ij_1}\}, \mathbf{E}_{j_1-1}\{\mathbf{\Psi}_{j_2} U_{ij_1}\}\} \\
&= \mathbf{E}\{\mathbf{\Psi}_{j_1}\mathbf{E}_{j_1-1}\{U_{ij_1}, U_{ij_2}\mathbf{\Psi}_{j_2}^{\mathrm{t}}\}\} = \mathbf{E}\{\mathbf{\Psi}_{j_1}\mathbf{E}_{j_1-1}\{\mathbf{E}_{j_2-1}\{U_{ij_1}U_{ij_2}\mathbf{\Psi}_{j_2}^{\mathrm{t}}\}\}\} \\
&= \mathbf{E}\{\mathbf{\Psi}_{j_1}\mathbf{E}_{j_1-1}\{U_{ij_1}\mathbf{E}_{j_2-1}\{U_{ij_2}\mathbf{\Psi}_{j_2}^{\mathrm{t}}\}\}\} = \mathbf{0}
\end{aligned}$$

since $\mathbf{E}_{j_2-1}\{U_{ij_2}\mathbf{\Psi}_{j_2}^{\mathrm{t}}\} = \mathbf{E}_{j_2-1}\{U_{ij_2}\}\mathbf{\Psi}_{j_2}^{\mathrm{t}} = \mathbf{0}$. Therefore,

$$\mathbf{Cov}\left\{\sum_{j=1}^{J_0} \mathbf{\Psi}_j U_{ij}\right\} = \sum_{j=1}^{J_0} \mathbf{E}\left\{\mathbf{\Psi}_j W_{ij}\lambda_j^0(1-\lambda_j^0)\mathbf{\Psi}^{\mathrm{t}}\right\} = \mathbf{E}\left\{\sum_{j=1}^{J_0} \mathbf{\Psi}_j W_{ij}\lambda_j^0(1-\lambda_j^0)\mathbf{\Psi}^{\mathrm{t}}\right\} = \mathbf{E}\left\{\mathbf{\Psi}\mathbf{V}_0\mathbf{\Psi}^{\mathrm{t}}\right\}.$$

||

**Proof of Theorem 4:** Let $\mathbf{c} \in \Re_p$ be fixed but arbitrary and with $\mathbf{c} \neq \mathbf{0}$. We consider the sequence $\{T_n(\mathbf{c}) : n = 1, 2, 3, \ldots\}$ with $T_n(\mathbf{c}) = \frac{1}{\sqrt{n}}(\mathbf{c}^{\mathrm{t}}\mathbf{\Psi})(\mathbf{O} - \mathbf{E}_0)$. If we could show that $T_n(\mathbf{c})$ converges in distribution to a $N(0, \mathbf{c}^{\mathrm{t}}\mathbf{\Xi}\mathbf{c})$ as $n \to \infty$, then by the Cramer-Wold device it will follow that $\frac{1}{\sqrt{n}}\mathbf{U}_0^* = \frac{1}{\sqrt{n}}\mathbf{\Psi}(\mathbf{O} - \mathbf{E}_0) \xrightarrow{\mathrm{d}} N_p(\mathbf{0}, \mathbf{\Xi})$. Let $\sigma_n^2(\mathbf{c}) = \frac{1}{n}\mathbf{c}^{\mathrm{t}}\mathbf{\Psi}\mathbf{V}_0\mathbf{\Psi}^{\mathrm{t}}\mathbf{c}$. By condition (i) it follows that $\sigma_n^2(\mathbf{c}) \xrightarrow{\mathrm{pr}} \sigma_0^2(\mathbf{c}) = \mathbf{c}^{\mathrm{t}}\mathbf{\Xi}\mathbf{c}$, which is positive since $\mathbf{\Xi}$ is positive definite. We may rewrite $T_n(\mathbf{c})$ via $T_n(\mathbf{c}) = \frac{1}{\sqrt{n}}\sum_{j=1}^{J}(\mathbf{c}^{\mathrm{t}}\mathbf{\Psi}_j)(O_j - E_j^0)$. Let $\{\xi_{nj}(\mathbf{c}) : j = 1, 2, \ldots, J; n = 1, 2, \ldots\}$

be defined via

$$\xi_{nj}(\mathbf{c}) = \frac{n^{-1/2}(\mathbf{c}^t\mathbf{\Psi}_j)(O_j - E_j^0)}{\sigma_0(\mathbf{c})}.$$

Then, since $\mathbf{\Psi}_j \in \mathcal{F}_{j-1}$ and $\mathbf{E}\{O_j|\mathcal{F}_{j-1}\} = \mathbf{E}_{j-1}(O_j) = R_j\lambda_j^0$, we obtain $\mathbf{E}\{\xi_{nj}(\mathbf{c})\} = 0$. Therefore, $\{\xi_{nj}(\mathbf{c}) : j = 1, 2, \ldots, J\}$ is a martingale difference array. Furthermore, we have $\mathbf{Var}\{\xi_{nj}(\mathbf{c})|\mathcal{F}_{j-1}\} = \mathbf{Var}_{j-1}\{\xi_{nj}(\mathbf{c})\} = n^{-1}(\mathbf{c}^t\mathbf{\Psi}_j)^2 R_j\lambda_j^0(1-\lambda_j^0)/\mathbf{c}^t\mathbf{\Xi}\mathbf{c}$. Therefore,

$$\begin{aligned}\sum_{j=1}^J \mathbf{E}_{j-1}\{\xi_{nj}^2(\mathbf{c})\} &= \sum_{j=1}^J \mathbf{Var}_{j-1}\{\xi_{nj}(\mathbf{c})\} = \frac{n^{-1}\sum_{j=1}^J(\mathbf{c}^t\mathbf{\Psi}_j)^2 R_j\lambda_j^0(1-\lambda_j^0)}{\mathbf{c}^t\mathbf{\Xi}\mathbf{c}} \\ &= \frac{n^{-1}\mathbf{c}^t\mathbf{\Psi}\mathbf{V}_0\mathbf{\Psi}^t\mathbf{c}}{\mathbf{c}^t\mathbf{\Xi}\mathbf{c}} \xrightarrow{\text{pr}} \frac{\mathbf{c}^t\mathbf{\Xi}\mathbf{c}}{\mathbf{c}^t\mathbf{\Xi}\mathbf{c}} = 1.\end{aligned}$$

By Helland's Theorem (Helland, 1982), if we could show that for every $\epsilon > 0$,

$$\sum_{j=1}^J \mathbf{E}_{j-1}\left\{\xi_{nj}^2(\mathbf{c})I\{|\xi_{nj}(\mathbf{c})| > \epsilon\}\right\} \xrightarrow{\text{pr}} 0, \tag{35}$$

then we could conclude that $T_n(\mathbf{c})/\sigma_0(\mathbf{c}) = \sum_{j=1}^J \xi_{nj}(\mathbf{c}) \xrightarrow{\text{d}} N(0,1)$ for every $\mathbf{c} \in \Re_p$ with $\mathbf{c} \neq \mathbf{0}$. From this it will follow that $n^{-1/2}\mathbf{\Psi}(\mathbf{O} - \mathbf{E}_0) \xrightarrow{\text{d}} N_p(\mathbf{0}, \mathbf{\Xi})$, which will complete the proof of the theorem. We establish condition (35) by establishing Lemma 3 and Lemma 4. ||

**Lemma 3** *If condition (iii) of Theorem 4 holds, and for every* $\mathbf{c} \in \Re_p$ *with* $\mathbf{c} \neq 0$,

$$\frac{\max_{1\le j\le J}\{(\mathbf{c}^t\mathbf{\Psi}_j)^2 R_j\lambda_j^0(1-\lambda_j^0)\}}{\sum_{l=1}^J(\mathbf{c}^t\mathbf{\Psi}_l)^2 R_l\lambda_l^0(1-\lambda_l^0)} \xrightarrow{\text{pr}} 0, \tag{36}$$

*then condition (35) holds.*

**Proof:** For $j = 1, 2, \ldots, J$, let

$$Z_j = \frac{O_j - E_j^0}{\sqrt{R_j\lambda_j^0(1-\lambda_j^0)}} \quad \text{and} \quad \Gamma_j(\mathbf{c}) = \frac{n^{-1/2}(\mathbf{c}^t\mathbf{\Psi}_j)\sqrt{R_j\lambda_j^0(1-\lambda_j^0)}}{\sigma_0(\mathbf{c})}.$$

Then,

$$\sum_{j=1}^J \mathbf{E}_{j-1}\left\{\xi_{nj}^2(\mathbf{c})I\{|\xi_{nj}(\mathbf{c})| > \epsilon\}\right\} = \sum_{j=1}^J \Gamma_j^2(\mathbf{c})\mathbf{E}_{j-1}\left\{Z_j^2 I\{|Z_j| > \epsilon/|\Gamma_j(\mathbf{c})|\}\right\} \le \frac{1}{\epsilon^2}\sum_{j=1}^J \Gamma_j^4(\mathbf{c})\mathbf{E}_{j-1}\{Z_j^4\}.$$

Note that $Z_j^4[R_j\lambda_j^0(1-\lambda_j^0)]^2 = (\sum_{i=1}^n U_{ij})^4$. Furthermore, we have

$$\begin{aligned}
\mathbf{E}_{j-1}\left[\left(\sum_{i=1}^n U_{ij}\right)^4\right] &= \sum_{i=1}^n \mathbf{E}_{j-1}(U_{ij}^4) + 2\sum_{i_1\neq i_2}\mathbf{E}_{j-1}(U_{i_1j}^2U_{i_2j}^2) \\
&= \left(\sum_{i=1}^n W_{ij}\right)\lambda_j^0(1-\lambda_j^0)[(1-\lambda_j^0)^3 + \\
&\quad (\lambda_j^0)^3] + 2\left(\sum_{i_1\neq i_2} W_{i_1j}W_{i_2j}\right)[\lambda_j^0(1-\lambda_j^0)]^2 \\
&\leq R_j\lambda_j^0(1-\lambda_j^0) + (R_j^2 - R_j)[\lambda_j^0(1-\lambda_j^0)]^2 \\
&\leq 3R_j\lambda_j^0(1-\lambda_j^0) + 2[R_j\lambda_j^0(1-\lambda_j^0)]^2.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\sum_{j=1}^J \Gamma_j^4(\mathbf{c})\mathbf{E}_{j-1}(Z_j^4) &= \sum_{j=1}^J \frac{\Gamma_j^4(\mathbf{c})}{[R_j\lambda_j^0(1-\lambda_j^0)]^2}\mathbf{E}_{j-1}\left[\left(\sum_{i=1}^n U_{ij}\right)^4\right] \\
&\leq \frac{1}{n^2\sigma_0^4(\mathbf{c})}\sum_{j=1}^J(\mathbf{c}^{\mathrm{t}}\mathbf{\Psi}_j)^4\left\{3R_j\lambda_j^0(1-\lambda_j^0) + 2[R_j\lambda_j^0(1-\lambda_j^0)]^2\right\} \\
&\leq 3||\mathbf{c}||^2\left\{\max_{1\leq j\leq J}\frac{||\mathbf{\Psi}_j||^2}{n}\right\}\frac{1}{\sigma_0^4(\mathbf{c})n}\sum_{j=1}^J(\mathbf{c}^{\mathrm{t}}\mathbf{\Psi}_j)^2R_j\lambda_j^0(1-\lambda_j^0) + \\
&\quad 2\frac{1}{\sigma_0^4(\mathbf{c})n^2}\sum_{j=1}^J\left[(\mathbf{c}^{\mathrm{t}}\mathbf{\Psi}_j)^2R_j\lambda_j^0(1-\lambda_j^0)\right]^2.
\end{aligned}$$

The first term of the upper bound on the right-hand side is $o_p(1)$ because $\max_{1\leq j\leq J}||\mathbf{\Psi}_j||^2 = o_p(n)$, while the second term of this bound is $o_p(1)$ whenever condition (36) holds. Thus, the lemma is established. ||

**Lemma 4** *Condition (ii) of Theorem 4 is sufficient for (36).*

**Proof:** First, note that

$$\begin{aligned}
(\mathbf{c}^{\mathrm{t}}\mathbf{\Psi}_j)^2R_j\lambda_j^0(1-\lambda_j^0) &= \left(\sum_{i=1}^p c_i\Psi_{ij}\right)^2 R_j\lambda_j^0(1-\lambda_j^0) \\
&= \left(\sum_{i_1=1}^p\sum_{i_2=1}^p c_{i_1}c_{i_2}\Psi_{i_1j}\Psi_{i_2j}\right)R_j\lambda_j^0(1-\lambda_j^0) \\
&= \sum_{i_1=1}^p\sum_{i_2=1}^p\left[\Psi_{i_1j}\Psi_{i_2j}R_j\lambda_j^0(1-\lambda_j^0)\right]c_{i_1}c_{i_2} \\
&= \mathbf{c}^{\mathrm{t}}\mathbf{\Psi}\mathbf{J}_j\mathbf{V}_0\mathbf{J}_j\mathbf{\Psi}^{\mathrm{t}}\mathbf{c},
\end{aligned}$$

where $\mathbf{J}_j$ is the $J \times J$ matrix whose elements are all zeros except for the $(j,j)$th element which equals 1. It then follows that

$$\sum_{l=1}^{J} (\mathbf{c}^{\mathrm{t}}\mathbf{\Psi}_l)^2 R_l \lambda_l^0 (1-\lambda_l^0) = \mathbf{c}^{\mathrm{t}}\mathbf{\Psi}\left(\sum_{l=1}^{J} \mathbf{J}_l \mathbf{V}_0 \mathbf{J}_l\right)\mathbf{\Psi}^{\mathrm{t}}\mathbf{c} = \mathbf{c}^{\mathrm{t}}\mathbf{\Psi}\mathbf{V}_0\mathbf{\Psi}^{\mathrm{t}}\mathbf{c}.$$

Therefore,

$$\frac{(\mathbf{c}^{\mathrm{t}}\mathbf{\Psi}_j)^2 R_j \lambda_j^0 (1-\lambda_j^0)}{\sum_{l=1}^{J} (\mathbf{c}^{\mathrm{t}}\mathbf{\Psi}_l)^2 R_l \lambda_l^0 (1-\lambda_l^0)} = \frac{\mathbf{c}^{\mathrm{t}}\mathbf{\Psi}\mathbf{J}_j\mathbf{V}_0\mathbf{J}_j\mathbf{\Psi}^{\mathrm{t}}\mathbf{c}}{\mathbf{c}^{\mathrm{t}}\mathbf{\Psi}\mathbf{V}_0\mathbf{\Psi}^{\mathrm{t}}\mathbf{c}},$$

which is bounded above by the *largest* solution, which we shall denote by $\alpha_j$, of the determinantal equation (in $\alpha$) given by

$$\det\left[\mathbf{\Psi}\mathbf{J}_j\mathbf{V}_0\mathbf{J}_j\mathbf{\Psi}^{\mathrm{t}} - \alpha\mathbf{\Psi}\mathbf{V}_0\mathbf{\Psi}^{\mathrm{t}}\right] = \det\left[\mathbf{\Psi}\left\{\mathbf{J}_j\mathbf{V}_0\mathbf{J}_j - \alpha\mathbf{V}_0\right\}\mathbf{\Psi}^{\mathrm{t}}\right] = 0. \tag{37}$$

[See, for instance, Anderson (1984, Th. A.2.4 on p. 590).] Consequently,

$$\frac{\max_{1\le j\le J}(\mathbf{c}^{\mathrm{t}}\mathbf{\Psi}_j)^2 R_j \lambda_j^0 (1-\lambda_j^0)}{\sum_{l=1}^{J} (\mathbf{c}^{\mathrm{t}}\mathbf{\Psi}_l)^2 R_l \lambda_l^0 (1-\lambda_l^0)} \le \max_{1\le j\le J} \alpha_j,$$

so if $\max_{1\le j\le J}\alpha_j \xrightarrow{\mathrm{pr}} 0$, then (36) holds. By straightforward manipulations, we have

$$\mathbf{\Psi}\mathbf{J}_j\mathbf{V}_0\mathbf{J}_j\mathbf{\Psi}^{\mathrm{t}} - \alpha\mathbf{\Psi}\mathbf{V}_0\mathbf{\Psi}^{\mathrm{t}} = -\alpha\left(\mathbf{\Psi}\mathbf{V}_0\mathbf{\Psi}^{\mathrm{t}}\right) + \left(\sqrt{V_{jj}^0}\mathbf{\Psi}_j\right)\left(\sqrt{V_{jj}^0}\mathbf{\Psi}_j\right)^{\mathrm{t}}.$$

Using the well-known identity $\det[\mathbf{C} + \mathbf{y}\mathbf{y}^{\mathrm{t}}] = \det(\mathbf{C})[1 + \mathbf{y}^{\mathrm{t}}\mathbf{C}^{-1}\mathbf{y}]$ for a nonsingular matrix $\mathbf{C}$, it follows that

$$\det\left[\mathbf{\Psi}\left\{\mathbf{J}_j\mathbf{V}_0\mathbf{J}_j - \alpha\mathbf{V}_0\right\}\mathbf{\Psi}^{\mathrm{t}}\right] = (-\alpha)^p \det\left(\mathbf{\Psi}\mathbf{V}_0\mathbf{\Psi}^{\mathrm{t}}\right)\left\{1 + \left(-\frac{1}{\alpha}\right)V_{jj}^0\mathbf{\Psi}_j^{\mathrm{t}}\left(\mathbf{\Psi}\mathbf{V}_0\mathbf{\Psi}^{\mathrm{t}}\right)^{-1}\mathbf{\Psi}_j\right\}.$$

Therefore, the solutions of (37) are $\alpha = 0$ with multiplicity $p-1$, and the largest solution is

$$\alpha_j = V_{jj}^0\mathbf{\Psi}_j^{\mathrm{t}}\left(\mathbf{\Psi}\mathbf{V}_0\mathbf{\Psi}^{\mathrm{t}}\right)^{-1}\mathbf{\Psi}_j = \mathrm{trace}\left\{\left(\mathbf{\Psi}\mathbf{V}_0\mathbf{\Psi}^{\mathrm{t}}\right)^{-1}\left(\mathbf{\Psi}_j V_{jj}^0\mathbf{\Psi}_j^{\mathrm{t}}\right)\right\}.$$

Thus, the condition given in (ii) of the statement of Theorem 4 is sufficient for (36) to hold. This completes the proof of Lemma 4, and hence Theorem 4. ||

## References


1. Anderson, T. (1984). *An Introduction to Multivariate Statistical Analysis.* 2nd edition. John Wiley & Sons: New York.

2. Agustin, Z. and Peña, E. (2001). Goodness-of-Fit of the Distribution of Time-to-First-Occurrence in Recurrent Event Models. To appear in *Lifetime Data Analysis.*

3. Akritas, M. (1988). Pearson-type goodness-of-fit tests: the univariate case. *J. Amer. Statist. Assoc.*, **83**, 222-230.

4. Baringhaus, L. and Henze, N. (1992). A goodness of fit test for the Poisson distribution based on the empirical generating function. *Statist. Probab. Lett.*, **13**, 269–274.

5. Best, D. and Rayner, J. (1989). Goodness of fit for the geometric distribution. *Biom. J.*, **31**, 307–311.

6. Best, D. and Rayner, J. (1999). Goodness of fit for the Poisson distribution. *Statist. Probab. Lett.*, **44**, 259–265.

7. Chernoff, H. and Lehmann, E. (1954). The use of maximum likelihood estimates in $\chi^2$ tests for goodness of fit. *Ann. Math. Statist.*, **25**, 579–586.

8. Choulakian, V., Lockhart, R. and Stephens, M. (1994). Cramer-von Mises statistics for discrete distributions. *Canad. J. Statist.*, **22**, 125–137.

9. Cox, D. (1972). Regression models and life tables (with discussion). *J. Royal Statist. Soc. Ser. B*, **34**, 187–220.

10. Cressie, N. and Read, T. (1984). Multinomial goodness-of-fit tests. *J. Royal Statist. Soc. Ser. B*, **46**, 440–464.

11. D'Agostino, R. and Stephens, M. (Editors). *Goodness-of-Fit Techniques.* Marcel Dekker, Inc.: New York.

12. Eubank, R. (1997). Testing goodness of fit with multinomial data. *J. Amer. Statist. Assoc.*, **92**, 1084–1093.

13. Fleming, T. and Harrington, D. (1991). *Counting Processes & Survival Analysis.* Wiley: New York.

14. Gatsonis, C., Hsieh, H. and Korwar, R. (1985). Simple nonparametric tests for a known standard survival based on censored data. *Comm. Statist. Theory Methods*, **14**, 2137–2162.

15. Greenwood, P. and Nikulin, M. (1996). *A Guide to Chi-Squared Testing.* John Wiley & Sons, Inc.: New York.

16. Habib, M. and Thomas, D. (1986). Chi-squared goodness-of-fit tests for randomly censored data. *Ann. Statist.*, **14**, 759–765.

17. Hart, J. (1997). *Nonparametric Smoothing and Lack-of-Fit Tests.* Springer: New York.

18. Helland, I. (1982). Central Limit Theorems for Martingales with Discrete or Continuous Time. *Scand. J. Statist.*, **9**, 79–94.

19. Henze, N. (1996). Empirical-distribution-function goodness-of-fit tests for discrete models. *Canad. J. Statist.*, **24**, 81–93.

20. Hjort, N. (1990). Goodness-of-fit tests in models for life history data based on cumulative hazard rates. *Ann. Statist.*, **18**, 1221–1258.

21. Hollander, M. and Proschan, F. (1979). Testing to determine the underlying distribution using randomly censored data. *Biom. J.*, **35**, 193–401.

22. Hollander, M. and Peña, E. (1992). A chi-squared goodness-of-fit test for randomly censored data. *J. Amer. Statist. Assoc.*, **87**, 458–463.

23. Hyde, J. (1977). Testing survival under right censoring and left truncation. *Biometrika*, **64**, 225–230.

24. Kalbfleisch, J. and Prentice, R. (1980). *The Statistical Analysis of Failure Time Data.* Wiley: New York.

25. Kallenberg, W. and Ledwina, T. (1995). Consistency and Monte Carlo simulation of a data driven version of smooth goodness-of-fit tests. *Ann. Statist.*, **23**, 1594-1608.

26. Khamaladze, E. (1981). A martingale approach in the theory of goodness-of-fit tests. *Teor. Veroyatnost. i Primenen*, **26**, 246–265 (English translation in *Theory Probab. Appl.*, **26**, 240–257).

27. Khamaladze, E. (1993). Goodness of fit problem and scanning innovation martingales. *Ann. Statist.*, **21**, 798–829.

28. Kim, J. (1993). Chi-square goodness-of-fit tests for randomly censored data. *Ann. Statist.*, **21**, 1621–1639.

29. Klar, B. (1999). Goodness-of-fit tests for discrete models based on the integrated distribution function. *Metrika*, **49**, 53–69.

30. Kocherlakota, S. and Kocherlakota, K. (1986). Goodness of fit tests for discrete distributions. *Comm. Statist. A — Theory Methods*, **15**, 815–829.

31. Koziol, J. and Green, S. (1976). A Cramer-von Mises statistic for randomly censored data. *Biometrika*, **63**, 139-156.

32. Kulperger, R. and Singh, A. (1982). On random grouping in goodness of fit tests of discrete distributions. *J. Statist. Plann. Inference*, **7**, 109–115.

33. Ledwina, T. (1994). Data driven version of Neyman's smooth test of fit. *J. Amer. Statist. Assoc.*, **89**, 1000–1005.

34. Li, G. and Doss, H. (1993). Generalized Pearson-Fisher chi-square goodness-of-fit tests, with applications to models with life history data. *Ann. Statist.*, **21**, 772–797.

35. Nair, V. (1981). Plots and tests for goodness of fit with randomly censored data. *Biometrika*, **68**, 99–103.

36. Nair, V. (1982). Goodness of fit tests for multiply right censored data. Nonparametric statistical inference, Vol. I, II (Budapest, 1980), 653–666, *Colloq. Math. Soc. Jnos Bolyai*, **32**, North-Holland, Amsterdam-New York.

37. Nair, V. (1984). Confidence Bands for Survival Functions with Censored Data: A Comparative Study. *Technometrics*, **26**, 265–275.

38. Nakamura, M. and Perez-Abreau, V. (1993). Use of an empirical probability generating function for testing a Poisson model. *Canad. J. Statist.*, **21**, 149–156.

39. Neyman, J. (1937). "Smooth" test for goodness of fit. *Skand. Aktuarietidskr.*, **20**, 150–199.

40. Pearson, K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can reasonably be supposed to have arisen from random sampling. *Philos. Mag.*, 5ht ser., **50**, 157–175.

41. Peña, E. (1998a). Smooth Goodness-of-Fit Tests for the Baseline Hazard in Cox's Proportional Hazards Model. *J. Amer. Statist. Assn.*, **93**, 673–692.

42. Peña, E. (1998b). Smooth Goodness-of-Fit Tests for Composite Hypothesis in Hazard-Based Models. *Ann. Statist.*, **26**, 1935–1971.

43. Rao, K. and Robson, D. (1974). A chi-square statistic for goodness-of-fit tests within the exponential family. *Comm. Statist.*, **3**, 1139–1153.

44. Rueda, R., Perez-Abreu, V. and O'Reilly, F. (1991). Goodness of fit for the Poisson distribution based on the probability generating function. *Comm. Statist. Theory Methods A*,**20**, 3093–3110.

45. Rayner, J. and Best, D. (1989). *Smooth Tests of Goodness of Fit.* Oxford University Press: New York.

46. Spinelli, J. and Stephens, M. (1997). Cramer-von Mises tests of fit for the Poisson distribution. *Canad. J. Statist.*, **25**, 257–268.

47. Stephens, M. (1992). Introduction to Kolmogorov (1933). On the empirical determination of a distribution. In *Breakthroughs in Statistics* **2:** *Methodology and Distribution* (S. Kotz and N. Johnson, eds.) Springer: New York.

Table 1: Simulated levels of the asymptotic 5%-level tests for the geometric distribution when $\eta_0 = 0.03$ and a 75% probability of an uncensored failure time.

| $J_0$ | 30 | | | | 50 | | | |
|---|---|---|---|---|---|---|---|---|
| $n$ | 30 | 50 | 100 | 200 | 30 | 50 | 100 | 200 |
| % Uncensored | 52.8 | 53.1 | 52.9 | 52.9 | 65.1 | 65.3 | 65.5 | 65.3 |
| Statistic | | | | | | | | |
| $S^2(\psi_1)$ | 5.8 | 4.8 | 5.5 | 5.5 | 5.4 | 5.3 | 4.9 | 5.8 |
| $S^2(\psi_2)$ | 5.2 | 4.7 | 6.1 | 5.6 | 4.9 | 5.1 | 4.8 | 5.8 |
| $S^2(\psi_3^1)$ | 5.7 | 5.3 | 4.5 | 4.5 | 6.0 | 5.4 | 5.9 | 5.4 |
| $S^2(\psi_3^{-1})$ | 4.7 | 5.0 | 5.9 | 6.1 | 3.3 | 4.0 | 4.8 | 5.3 |
| $S^2(\mathbf{\Psi}_5^1)$ | 5.4 | 5.0 | 4.9 | 6.1 | 5.9 | 6.1 | 5.5 | 5.9 |
| $S^2(\mathbf{\Psi}_5^2)$ | 5.5 | 4.9 | 5.1 | 5.1 | 5.3 | 5.9 | 4.6 | 5.5 |
| $S^2(\mathbf{\Psi}_5^3)$ | 7.7 | 5.2 | 5.0 | 5.0 | 7.2 | 6.9 | 5.5 | 6.1 |
| $S^2(\mathbf{\Psi}_5^4)$ | 6.8 | 5.3 | 6.0 | 5.2 | 7.3 | 6.2 | 5.8 | 4.6 |
| $S^2(\mathcal{I}_{J_0})$ | 11.0 | 8.9 | 7.5 | 5.6 | 13.3 | 10.2 | 8.5 | 6.3 |
| $S^2(\mathbf{\Psi}_6^1)$ | 5.8 | 4.8 | 5.5 | 5.5 | 5.4 | 5.3 | 4.9 | 5.8 |
| $S^2(\mathbf{\Psi}_6^2)$ | 6.3 | 6.2 | 5.3 | 5.1 | 5.8 | 5.9 | 5.6 | 6.0 |
| $S^2(\mathbf{\Psi}_6^3)$ | 7.0 | 5.4 | 5.7 | 5.1 | 8.0 | 6.3 | 6.9 | 5.8 |
| $S^2(\mathbf{\Psi}_6^4)$ | 8.2 | 6.4 | 5.9 | 4.6 | 8.2 | 6.7 | 6.2 | 5.2 |

Table 2: Simulated levels of the asymptotic 5%-level tests for the geometric distribution when $\eta_0 = 0.03$ and a 50% probability of an uncensored failure time.

| $J_0$ | 30 | | | | 50 | | | |
|---|---|---|---|---|---|---|---|---|
| $n$ | 30 | 50 | 100 | 200 | 30 | 50 | 100 | 200 |
| % Uncensored | 42.5 | 42.3 | 42.1 | 42.0 | 47.8 | 47.2 | 47.8 | 47.9 |
| Statistic | | | | | | | | |
| $S^2(\psi_1)$ | 6.4 | 4.1 | 6.1 | 6.6 | 5.9 | 4.5 | 6.0 | 4.4 |
| $S^2(\psi_2)$ | 5.6 | 3.9 | 5.5 | 6.0 | 7.0 | 5.4 | 5.2 | 4.9 |
| $S^2(\psi_3^1)$ | 6.3 | 5.3 | 6.2 | 6.0 | 5.5 | 4.0 | 5.7 | 4.8 |
| $S^2(\psi_3^{-1})$ | 4.8 | 4.0 | 5.5 | 5.2 | 8.7 | 5.2 | 4.0 | 4.9 |
| $S^2(\mathbf{\Psi}_5^1)$ | 5.0 | 5.2 | 5.4 | 6.6 | 5.4 | 5.9 | 5.3 | 5.7 |
| $S^2(\mathbf{\Psi}_5^2)$ | 5.3 | 5.5 | 5.1 | 5.0 | 6.5 | 6.6 | 6.2 | 5.3 |
| $S^2(\mathbf{\Psi}_5^3)$ | 6.4 | 4.7 | 3.9 | 5.9 | 8.1 | 6.5 | 6.8 | 6.1 |
| $S^2(\mathbf{\Psi}_5^4)$ | 6.5 | 5.1 | 4.8 | 5.7 | 7.9 | 7.2 | 5.9 | 5.5 |
| $S^2(\mathcal{I}_{J_0})$ | 13.5 | 9.8 | 9.0 | 6.6 | 19.4 | 13.0 | 10.0 | 10.0 |
| $S^2(\mathbf{\Psi}_6^1)$ | 6.4 | 4.1 | 6.1 | 6.6 | 5.9 | 4.5 | 6.0 | 4.4 |
| $S^2(\mathbf{\Psi}_6^2)$ | 6.4 | 5.5 | 5.9 | 5.9 | 6.0 | 5.1 | 6.0 | 4.4 |
| $S^2(\mathbf{\Psi}_6^3)$ | 6.9 | 4.8 | 5.8 | 5.1 | 6.8 | 6.3 | 5.8 | 4.4 |
| $S^2(\mathbf{\Psi}_6^4)$ | 8.5 | 5.0 | 5.9 | 4.8 | 8.5 | 7.1 | 6.1 | 5.6 |

Table 3: Simulated levels of the asymptotic 5%-level tests for the geometric distribution when $\eta_0 = 0.10$ and a 75% probability of an uncensored failure time.

| $J_0$ | 30 | | | | 50 | | | |
|---|---|---|---|---|---|---|---|---|
| $n$ | 30 | 50 | 100 | 200 | 30 | 50 | 100 | 200 |
| % Uncensored | 74.1 | 74.0 | 73.9 | 73.8 | 74.9 | 74.9 | 74.6 | 74.9 |
| Statistic | | | | | | | | |
| $S^2(\psi_1)$ | 6.6 | 5.4 | 4.2 | 4.8 | 4.9 | 5.7 | 5.9 | 4.9 |
| $S^2(\psi_2)$ | 7.1 | 5.1 | 5.3 | 4.4 | 6.0 | 6.0 | 6.5 | 6.0 |
| $S^2(\psi_3^1)$ | 5.5 | 6.1 | 3.3 | 4.1 | 6.0 | 5.7 | 5.5 | 4.7 |
| $S^2(\psi_3^{-1})$ | 8.4 | 6.9 | 8.2 | 5.0 | 8.5 | 9.7 | 7.5 | 6.9 |
| $S^2(\mathbf{\Psi}_5^1)$ | 6.0 | 6.2 | 4.5 | 6.1 | 5.6 | 5.6 | 3.9 | 4.5 |
| $S^2(\mathbf{\Psi}_5^2)$ | 8.1 | 5.6 | 6.3 | 5.4 | 9.6 | 9.2 | 8.0 | 5.8 |
| $S^2(\mathbf{\Psi}_5^3)$ | 9.5 | 7.2 | 6.0 | 5.8 | 11.6 | 10.1 | 10.0 | 10.7 |
| $S^2(\mathbf{\Psi}_5^4)$ | 10.6 | 7.8 | 6.6 | 5.8 | 12.3 | 12.9 | 11.4 | 9.6 |
| $S^2(\mathcal{I}_{J_0})$ | 11.8 | 10.6 | 8.0 | 5.4 | 13.3 | 14.2 | 10.9 | 9.5 |
| $S^2(\mathbf{\Psi}_6^1)$ | 6.6 | 5.4 | 4.2 | 4.8 | 4.9 | 5.7 | 5.9 | 4.9 |
| $S^2(\mathbf{\Psi}_6^2)$ | 6.8 | 6.2 | 4.0 | 5.8 | 6.3 | 5.5 | 5.1 | 5.5 |
| $S^2(\mathbf{\Psi}_6^3)$ | 6.2 | 5.4 | 4.1 | 5.4 | 7.5 | 6.0 | 5.0 | 5.6 |
| $S^2(\mathbf{\Psi}_6^4)$ | 6.8 | 6.2 | 4.8 | 4.5 | 8.2 | 6.9 | 5.9 | 5.9 |

Table 4: Simulated levels of the asymptotic 5%-level tests for the geometric distribution when $\eta_0 = 0.10$ and a 50% probability of an uncensored failure time.

| $J_0$ | 30 | | | | 50 | | | |
|---|---|---|---|---|---|---|---|---|
| $n$ | 30 | 50 | 100 | 200 | 30 | 50 | 100 | 200 |
| % Uncensored | 50.0 | 49.8 | 49.9 | 49.9 | 49.8 | 50.0 | 50.2 | 49.8 |
| Statistic | | | | | | | | |
| $S^2(\psi_1)$ | 4.9 | 5.3 | 4.2 | 6.0 | 4.0 | 4.9 | 6.1 | 5.4 |
| $S^2(\psi_2)$ | 6.0 | 6.1 | 4.2 | 4.7 | 4.3 | 5.6 | 5.0 | 5.4 |
| $S^2(\psi_3^1)$ | 4.7 | 5.0 | 5.2 | 5.6 | 5.6 | 5.6 | 5.7 | 4.7 |
| $S^2(\psi_3^{-1})$ | 8.0 | 7.1 | 6.7 | 6.6 | 6.5 | 7.1 | 7.4 | 7.8 |
| $S^2(\mathbf{\Psi}_5^1)$ | 4.2 | 5.9 | 3.7 | 3.8 | 4.4 | 4.1 | 6.2 | 3.9 |
| $S^2(\mathbf{\Psi}_5^2)$ | 8.1 | 9.0 | 6.5 | 4.5 | 5.7 | 4.8 | 7.7 | 7.7 |
| $S^2(\mathbf{\Psi}_5^3)$ | 9.2 | 10.2 | 10.3 | 7.1 | 8.3 | 8.4 | 10.5 | 7.9 |
| $S^2(\mathbf{\Psi}_5^4)$ | 9.5 | 11.7 | 9.8 | 8.7 | 8.4 | 9.7 | 11.0 | 11.3 |
| $S^2(\mathcal{I}_{J_0})$ | 13.6 | 9.8 | 9.5 | 10.9 | 11.6 | 12.5 | 11.3 | 11.3 |
| $S^2(\mathbf{\Psi}_6^1)$ | 4.9 | 5.3 | 4.2 | 6.0 | 4.0 | 4.9 | 6.1 | 5.4 |
| $S^2(\mathbf{\Psi}_6^2)$ | 5.6 | 6.0 | 4.7 | 5.8 | 5.4 | 5.4 | 5.6 | 5.4 |
| $S^2(\mathbf{\Psi}_6^3)$ | 6.9 | 6.6 | 6.3 | 5.6 | 6.0 | 6.0 | 6.3 | 4.9 |
| $S^2(\mathbf{\Psi}_6^4)$ | 7.5 | 6.2 | 5.6 | 5.9 | 6.9 | 7.2 | 5.4 | 5.6 |

Table 5: Simulated powers of the asymptotic 5%-level tests for a geometric null with $\eta_0 = 0.03$ and a 75% probability (under the null) of an uncensored failure time under geometric alternatives with 'success' probability of $\eta$.

| $\eta$ | .0161 | .0227 | .0263 | .0285 | **.03 ($H_0$)** | .0315 | .0337 | .0373 | .0438 |
|---|---|---|---|---|---|---|---|---|---|
| % Uncensored | 45.1 | 56.0 | 60.9 | 63.6 | **65.3** | 66.7 | 69.1 | 72.1 | 76.6 |
| Statistic | | | | | | | | | |
| $S^2(\psi_1)$ | 99.6 | 56.9 | 15.9 | 7.4 | **6.0** | 7.2 | 17.3 | 46.5 | 91.9 |
| $S^2(\psi_2)$ | 99.4 | 56.6 | 14.9 | 7.5 | **5.3** | 6.4 | 17.0 | 40.7 | 87.3 |
| $S^2(\psi_3^1)$ | 99.1 | 50.4 | 14.0 | 6.4 | **6.0** | 7.9 | 17.8 | 42.4 | 86.3 |
| $S^2(\psi_3^{-1})$ | 98.6 | 50.4 | 14.3 | 7.6 | **4.6** | 5.4 | 11.6 | 25.9 | 65.3 |
| $S^2(\mathbf{\Psi}_5^1)$ | 97.9 | 44.8 | 11.9 | 5.7 | **5.8** | 7.7 | 16.3 | 43.1 | 85.0 |
| $S^2(\mathbf{\Psi}_5^2)$ | 98.6 | 45.3 | 11.9 | 6.2 | **4.9** | 8.1 | 17.6 | 42.5 | 85.0 |
| $S^2(\mathbf{\Psi}_5^3)$ | 97.6 | 35.9 | 8.6 | 5.0 | **4.8** | 9.0 | 16.2 | 38.4 | 81.6 |
| $S^2(\mathbf{\Psi}_5^4)$ | 95.1 | 29.7 | 8.2 | 5.3 | **5.1** | 8.4 | 16.1 | 38.5 | 79.4 |
| $S^2(\mathbf{\Psi}_6^1)$ | 99.6 | 56.9 | 15.9 | 7.4 | **6.0** | 7.2 | 17.3 | 46.5 | 91.9 |
| $S^2(\mathbf{\Psi}_6^2)$ | 98.7 | 45.8 | 10.8 | 5.8 | **5.7** | 6.7 | 16.0 | 40.3 | 87.1 |
| $S^2(\mathbf{\Psi}_6^3)$ | 97.9 | 36.3 | 8.3 | 5.9 | **5.4** | 7.3 | 16.5 | 40.3 | 84.6 |
| $S^2(\mathbf{\Psi}_6^4)$ | 95.7 | 29.7 | 6.3 | 5.9 | **5.4** | 8.1 | 16.5 | 39.0 | 83.3 |

Table 6: Simulated powers of the asymptotic 5%-level tests for a geometric null with $\eta_0 = 0.03$ and a 75% probability (under the null) of an uncensored failure time under negative binomial alternatives with parameters $(r, p)$.

| $(r, p)$ | (2,.03) | (2,.04) | (2,.05) | (2,.0583) | (2,.06) | (2,.07) | (2,.08) |
|---|---|---|---|---|---|---|---|
| % Uncensored | 34.1 | 46.9 | 57.2 | 63.6 | 65.0 | 71.5 | 75.7 |
| Statistic | | | | | | | |
| $S^2(\psi_1)$ | 100 | 100 | 75.5 | 17.8 | 7.4 | 10.1 | 63.9 |
| $S^2(\psi_2)$ | 100 | 98.9 | 48.7 | 5.8 | 3.1 | 44.1 | 89.9 |
| $S^2(\psi_3^1)$ | 100 | 100 | 97.8 | 66.7 | 50.5 | 4.5 | 2.0 |
| $S^2(\psi_3^{-1})$ | 100 | 91.6 | 21.6 | 7.8 | 11.9 | 58.9 | 89.6 |
| $S^2(\mathbf{\Psi}_5^1)$ | 100 | 99.5 | 62.5 | 9.0 | 6.3 | 8.7 | 55.6 |
| $S^2(\mathbf{\Psi}_5^2)$ | 100 | 99.9 | 90.9 | 61.8 | 58.2 | 57.9 | 84.9 |
| $S^2(\mathbf{\Psi}_5^3)$ | 100 | 100 | 93.6 | 76.6 | 73.8 | 78.5 | 92.9 |
| $S^2(\mathbf{\Psi}_5^4)$ | 100 | 99.8 | 92.6 | 78.9 | 73.4 | 82.8 | 94.9 |
| $S^2(\mathbf{\Psi}_6^1)$ | 100 | 100 | 75.5 | 17.8 | 7.4 | 10.1 | 63.9 |
| $S^2(\mathbf{\Psi}_6^2)$ | 100 | 100 | 97.7 | 89.6 | 89.2 | 91.7 | 97.6 |
| $S^2(\mathbf{\Psi}_6^3)$ | 100 | 99.9 | 96.9 | 87.9 | 88.1 | 93.1 | 98.1 |
| $S^2(\mathbf{\Psi}_6^4)$ | 100 | 99.9 | 94.1 | 82.7 | 83.3 | 91.0 | 97.1 |

Table 7: Simulated powers of the asymptotic 5%-level tests for a geometric null with $\eta_0 = 0.10$ and a 75% probability (under the null) of an uncensored failure time when the failure times are generated according to Poisson probabilities for different means ($\lambda$).

| $\lambda$ | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|
| % Uncensored | 77.1 | 74.4 | 71.6 | 68.9 | 75.0 |
| Statistic | | | | | |
| $S^2(\psi_1)$ | 2.9 | 0 | 10.4 | 85.5 | 99.9 |
| $S^2(\psi_2)$ | 100 | 99.4 | 83.9 | 17.8 | 0.1 |
| $S^2(\psi_3^1)$ | 88.4 | 100 | 100 | 100 | 100 |
| $S^2(\psi_3^{-1})$ | 99.8 | 99.9 | 98.6 | 99.0 | 95.3 |
| $S^2(\mathbf{\Psi}_5^1)$ | 10.6 | 2.7 | 9.8 | 62.8 | 98.7 |
| $S^2(\mathbf{\Psi}_5^2)$ | 1.8 | 0 | 10.8 | 85.4 | 100 |
| $S^2(\mathbf{\Psi}_5^3)$ | 9.3 | 22.4 | 55.5 | 95.2 | 100 |
| $S^2(\mathbf{\Psi}_5^4)$ | 83.6 | 95.7 | 99.7 | 100 | 100 |
| $S^2(\mathbf{\Psi}_6^1)$ | 2.9 | 0 | 10.4 | 85.5 | 99.9 |
| $S^2(\mathbf{\Psi}_6^2)$ | 100 | 100 | 100 | 100 | 100 |
| $S^2(\mathbf{\Psi}_6^3)$ | 100 | 100 | 100 | 100 | 100 |
| $S^2(\mathbf{\Psi}_6^4)$ | 100 | 100 | 100 | 100 | 100 |

Table 8: Simulated powers of the asymptotic 5%-level tests for a geometric null with $\eta_0 = 0.03$ and a 75% probability (under the null) of an uncensored failure time when the failure times are generated according to the specification $\lambda_j = 2\eta_0 G_j(\mathbf{a})/[1 + G_j(\mathbf{a})]$ where $G_j(\mathbf{a}) = \exp\{\sum_{k=1}^{q} a_k (j/J_0)^{k-1}\}$.

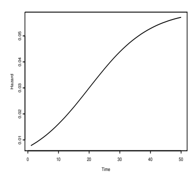
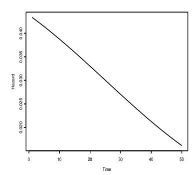
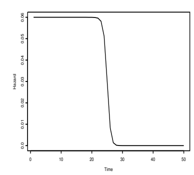
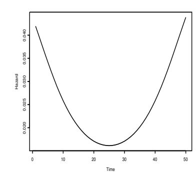
| Parameter: **a** | (-2,5) | (1,-2) | (20,10,-100) | (1,-8,8) |
|---|---|---|---|---|
| Plot of True Hazard Sequence |  |  |  |  |
| % Uncensored | 65.2 | 67.3 | 71.5 | 60.4 |
| Statistic | | | | |
| $S^2(\psi_1)$ | 14.0 | 24.8 | 88.6 | 19.1 |
| $S^2(\psi_2)$ | 3.5 | 8.7 | 38.5 | 22.9 |
| $S^2(\psi_3^1)$ | 83.1 | 48.2 | 99.2 | 11.1 |
| $S^2(\psi_3^{-1})$ | 21.4 | 3.0 | 1.1 | 22.0 |
| $S^2(\mathbf{\Psi}_5^1)$ | 7.2 | 21.5 | 85.1 | 16.2 |
| $S^2(\mathbf{\Psi}_5^2)$ | 94.2 | 43.3 | 100 | 16.5 |
| $S^2(\mathbf{\Psi}_5^3)$ | 96.8 | 40.9 | 100 | 34.7 |
| $S^2(\mathbf{\Psi}_5^4)$ | 97.0 | 38.9 | 100 | 42.4 |
| $S^2(\mathbf{\Psi}_6^1)$ | 14.0 | 24.8 | 88.6 | 19.1 |
| $S^2(\mathbf{\Psi}_6^2)$ | 99.0 | 51.3 | 100 | 34.0 |
| $S^2(\mathbf{\Psi}_6^3)$ | 98.2 | 49.6 | 100 | 53.8 |
| $S^2(\mathbf{\Psi}_6^4)$ | 97.0 | 48.0 | 100 | 54.6 |

Table 9: Simulated powers of the asymptotic 5%-level tests for a geometric null with $\eta_0 = 0.03$ and a 75% probability (under the null) of an uncensored failure time when the failure times are generated according to the specification $\lambda_j = 2\eta_0 G_j(\mathbf{a},\mathbf{b},\mathbf{c})/[1+G_j(\mathbf{a},\mathbf{b},\mathbf{c})]$ where $G_j(\mathbf{a},\mathbf{b},\mathbf{c}) = \exp\{\sum_{k=1}^{q}[a_k\sin(2\pi c_k(j/J_0)) + b_k\cos(2\pi c_k(j/J_0))]\}$.

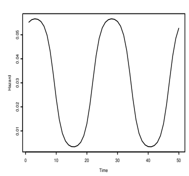
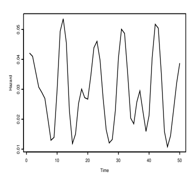
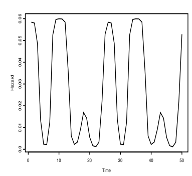
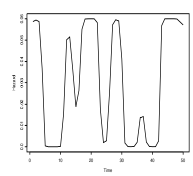
| Parameters $(\mathbf{a},\mathbf{b},\mathbf{c})$ | $\mathbf{a}=2$ <br> $\mathbf{b}=2$ <br> $\mathbf{c}=2$ | $\mathbf{a}=(.2,1,-.5)$ <br> $\mathbf{b}=(-.3,.4,.5)$ <br> $\mathbf{c}=(10,5,8)$ | $\mathbf{a}=(2,1,-3)$ <br> $\mathbf{b}=(-1,2,1)$ <br> $\mathbf{c}=(2,-4,-6)$ | $\mathbf{a}=(-5,3,2)$ <br> $\mathbf{b}=(4,-4,3)$ <br> $\mathbf{c}=(2,6,-3)$ |
|---|---|---|---|---|
| Plot of True Hazard Sequence |  |  |  |  |
| % Uncensored | 67.3 | 65.3 | 63.3 | 66.3 |
| Statistic | | | | |
| $S^2(\psi_1)$ | 19.1 | 5.8 | 6.3 | 5.0 |
| $S^2(\psi_2)$ | 7.4 | 5.3 | 7.7 | 5.0 |
| $S^2(\psi_3^1)$ | 50.1 | 8.9 | 14.6 | 9.7 |
| $S^2(\psi_3^{-1})$ | 3.2 | 5.6 | 10.2 | 6.3 |
| $S^2(\mathbf{\Psi}_5^1)$ | 18.7 | 4.5 | 6.3 | 5.4 |
| $S^2(\mathbf{\Psi}_5^2)$ | 21.1 | 5.9 | 7.2 | 6.9 |
| $S^2(\mathbf{\Psi}_5^3)$ | 32.7 | 5.3 | 28.9 | 32.5 |
| $S^2(\mathbf{\Psi}_5^4)$ | 94.0 | 5.8 | 89.6 | 36.9 |
| $S^2(\mathbf{\Psi}_6^1)$ | 19.1 | 5.8 | 6.3 | 5.0 |
| $S^2(\mathbf{\Psi}_6^2)$ | 65.5 | 10.9 | 39.9 | 11.9 |
| $S^2(\mathbf{\Psi}_6^3)$ | 74.3 | 13.7 | 43.4 | 31.1 |
| $S^2(\mathbf{\Psi}_6^4)$ | 72.6 | 16.8 | 44.9 | 64.8 |

Table 10: Assigned letter grades of the different tests under the five classes of alternatives based on their simulated powers with A = Best and E = Worst.

| Type of Alternative | Geometric | Negative Binomial | Poisson | Polynomially Generated | Trigonometrically Generated |
|---|---|---|---|---|---|
| Statistic | | | | | |
| $S^2(\psi_1)$ | A | E | E | E | E |
| $S^2(\psi_2)$ | A- | E | E | E | E |
| $S^2(\psi_3^1)$ | A- | E | B | D | D |
| $S^2(\psi_3^{-1})$ | B | E | A- | E | E |
| $S^2(\mathbf{\Psi}_5^1)$ | B | E | E | E | E |
| $S^2(\mathbf{\Psi}_5^2)$ | B | B | E | C | E |
| $S^2(\mathbf{\Psi}_5^3)$ | B | B | D | B | D |
| $S^2(\mathbf{\Psi}_5^4)$ | B- | B | B | B | B |
| $S^2(\mathbf{\Psi}_6^1)$ | A | E | E | E | E |
| $S^2(\mathbf{\Psi}_6^2)$ | B | A | A | B | C |
| $S^2(\mathbf{\Psi}_6^3)$ | B | A | A | A | B- |
| $S^2(\mathbf{\Psi}_6^4)$ | B- | A- | A | A | B |