Chi-Squared Goodness of Fit Test Based on Random Cells With Recurrent Event Data

Withanage A. De Mel * Akim Adekpedjou † Gideon K.D. Zamba ‡

Abstract

We consider a recurrent event wherein the inter-event time distribution F is assumed to belong to some parametric family of the distributions $\mathcal{F}$, where the unknown parameter θ is q-dimensional. This work deals with the problem of goodness-of-fit test for F. We develop a chi-square type test where the k nonoverlapping cell boundaries are randomly chosen. Our test used a Kaplan Meier type nonparametric maximum likelihood estimator (NPMLE) of F to obtain the observed frequencies. The minimum distance estimator of θ is obtained by minimizing the quadratic form that resulted from the properly scaled vector of differences between the observed and expected cell frequencies. The proposed chi-square test statistic is constructed by using the NPMLE of F and the minimum distance estimator. We show that the proposed test statistic is asymptotically chi-square with $k-q-1$ degrees of freedom. Results for specific families of distributions such as Weibull is presented. We also discuss results of a simulation study as well as application to a biomedical data set.

Key Words: Recurrent events; Random cells boundaries; Chi-square test; Minimum distance estimator; Goodness-of-fit

1. Introduction

Consider a recurrent event process for $i = 1, ..., n$ units where the jth event occur at calendar time $S_{i,j}$. Suppose that for unit i, the recurrent event is observed over a random interval $[0, \tau_i]$ where the τ_is are independent and identically distributed (i.i.d.) with an absolutely continuous distribution function $G(t) = P(\tau \leq t)$. Let $T_{i,j} = S_{i,j} - S_{i,j-1}$ be the time between two occurrences of the event, the so called gap time or inter-event time-and these are assumed to be i.i.d. with absolutely continuous distribution function $F(t) = P(T_{i,j} \leq t)$. For the ith unit, the $T_{i,j}$s could be viewed as the time elapsed between the $(j-1)th$ and the jth occurrences in an experimental unit in a reliability or engineering study or that of a subject in a biomedical study. If K_i is the total number of occurrences per unit, then the observable for n units is n i.i.d. copies $\mathcal{O}_1$, $\mathcal{O}_2$,...,$\mathcal{O}_n$ with

$$\mathcal{O}_i = (K_i, \tau_i, T_{i,1}, \ldots, T_{i,K_i}, \tau_i - S_{i,K_i}, [\mathbf{X}_i(s) : s \leq \tau_i]), \tag{1}$$

where $\tau_i - S_{i,K_i}$ is the right censoring random variable for the inter-event time T_{i,K_i+1}, and $\mathbf{X}_i(s)$ is a possibly m-dimensional time dependent covariates vector.

Data of the type in (1) are prevalent in a variety of disciplines and settings including the biomedical, engineering, and social sciences-to only name those. In reliability, the data often represent successive failures of a repairable system. In biomedical studies, the data could be successive occurrence of a chronic disease, the recurrence of tumors in cancer patients etc. In sociology, absenteeism rate of employees and the recurrence of war and conflict in geographical regions are potential example of these type of data. Due to its high

*Department of Mathematics and Statistics, Missouri University of Science and Techonology, Rolla, MO 65409

†Department of Mathematics and Statistics, Missouri University of Science and Techonology, Rolla, MO 65409

‡Department of Biostatistics, University of Iowa, Iowa

prevalence and importance in many diverse areas, it is essential there exists appropriate statistical methodology to analyze them. These analysis include, but not limited to: estimation of model parameters such as the survivor function $\bar{F}(t) = 1 - F(t)$, the cumulative hazard rate function $\Lambda(t) = \int_0^t \lambda(w)dw$, where $\lambda(t)$ is the hazard rate function of F, or quantiles of $T_{i,j}$. Other major inferential problems include goodness of fit tests pertaining to the distribution function F, such as Kolmogorov-Smirnov, Cramér-von Mises, or the Pearson's type of tests. There has been sustain interest in testing goodness of fit for a parametric family of distribution, especially the development of chi-square types of test since the pioneering work of Fisher in 1922 and Fisher in 1924. Those interests come from the fact that in survival analysis, for instance, there may be physical reasons that indicate a parametric family for the underlying failure time distribution. In reliability studies, extreme value distributions such as Gumbel, Fréchet, or Weibull come as limit of distributions of parallel or series systems. If the parametric distributions are good models, they could be used for modeling large claims in actuarial studies. In the area of failure time data analysis-under valid assumptions-parametrically driven estimates of relative hazard, survival time or their functionals such as mean or median, tend to have smaller standard errors than they would in non-parametric settings. Another benefit derived from the method developed herein is its ability to be used for study design purposes. In the single events, the seminal papers dealing with the problem of chi-square goodness of fit with fixed or data-dependent cells include those of Hjort(1990), Kim(1993) , Li and Doss(1993), Habib and Thomas(1986), Akritas(1988), Hollander and Peña (1992), Moore and Spruill (1975), Pollard(1979), Mihalko and Moore(1980) to name a few. The situation where the event is recurrent has also been dealt with, albeit not as thoroughly yet as in the single event. In the recurrent event settings, the goodness of fit problem has been considered by Presnell, Hollander, and Sethuraman(1994), Agustin and Peña(2001) and Agustin and Peña(2005), Stocker and Adekpedjou(2011), Adekpedjou and Zamba(2012). Presnell et al.(1994) proposed tests for the minimal repair assumption in the imperfect repair model. Agustin and Peña(2001) proposed goodness of fit test for the Block, Borges, and Savits(1985) model whereas Agustin and Peña(2005) developed goodness of fit test for an extended Block et al.(1985) model that include covariates. Stocker and Adekpedjou(2011) developed a class of tests for the hazard rate function that include chi-squared, Kolmogorov-Smirnov, Cramẽr-von Mises and obtained asymptotic properties of their tests using empirical process techniques and Khmaladze transformation. Adekpedjou and Zamba(2012) developed a chi-squared goodness of fit for testing the hypothesis of completely known distribution with fixed cells based on a NPMLE of F. All the tests developed in the above manuscripts in the recurrent event settings are based on fixed cell boundaries and are different in many ways from the techniques developed herein. Standard statistics of the chi-square types are defined in terms of cells which are fixed prior to taking the observations. Furthermore, their distributions are based on multinomial distributions. To test a composite hypothesis, the unknown parameter must be estimated. If the estimator used is the maximum likelihood based on the cell frequencies, the resulting test is Pearson-Fisher χ^2. If instead the maximum likelihood used in the construction of the χ^2 statistic is based on the original data, then the resulting statistic does not have a limiting χ^2 distribution and the limiting distribution in general depends on the unknown true value θ (cf. Chernoff and Lehman(1954)). To overcome this, we allow the cell frequencies to be data dependent, that is depend on the estimated value of the unknown parameter and require that the cells settle down as the sample size increases. Cells obtained following this technique are called *random cells*. By doing so, the limiting distribution of the χ^2 statistic will not depend on the unknown parameter. This approach of constructing χ^2 statistic is more flexible, guarantees that the cells probabilities will not be small and is increasingly practicable. The cell frequencies are no longer multinomial and

the limiting distribution of the vector of standardized frequencies is now obtained using sophisticated techniques such empirical process techniques and the Skorohod construction. Li and Doss developed a χ^2 test based on random cells for right-censored and left truncated data in the single event settings. In fact, their test is applicable whenever there exists an estimator $\hat{F}(t)$ of $F(t)$ satisfying the asymptotic property $\sqrt{n}(\hat{F} - F) \Rightarrow W$, where $\Rightarrow$ denote weak convergence and W is a zero-mean continuous Gaussian process whose variance-covariance matrix is non-singular. Other chi-square tests based on random cells have also been developed by Moore(1971), Pollard(1979) among others-and these are not developed for recurrent events. The major goals of this article is to develop chi-square goodness of fit for testing the null hypothesis that F belongs to some parametric family of distributions. In the test we propose, the cells are random and data-driven. The proposed test generalized the work of Li and Doss(1993) to the situation where the event is recurrent. Furthermore, it encompasses a wide range of tests including-fixed null with random cells, fixed cells with composite hypothesis- and is different from those proposed in the literature of recurrent events. We use the NPMLE of the distribution function of the inter-event time developed in Peña et al.(2001) to obtain the observed frequencies. The NPMLE estimator is a generalized Kaplan-Meier type (cf. Kaplan and Meier(1958)). The expected frequencies are obtained using the estimator of θ that minimizes a quadratic form obtained from the suitably standardized vector of " observed - expected " frequencies.

1.1 Background on recurrent event modeling

In this subsection, we briefly review relevant stochastic processes that are used in the estimation of the NPMLE of F. Following Peña et al.(2001), we begin by defining some relevant *calendar* time stochastic processes given by $N_i^\dagger = \{N_i^\dagger(s) : s \leq \tau_i\}$, $Y_i^\dagger = \{Y_i^\dagger(s) : s \geq 0\}$, where $N_i^\dagger(s) = \sum_{j=1}^{\infty} I\{S_{i,j} \leq s \wedge \tau_i\}$, and $Y_i^\dagger(s) = I\{\tau_i \geq s\}$. For the ith subject, the $N_i^\dagger$ process determines the event occurrences up to time τ_i whereas the $Y_i^\dagger$ process determines if the unit is at-risk for a recurrent event. Let the backward recurrence time process-that is the time elapsed since the last event- for unit i be defined by $R_i = \{R_i(s) : s \geq 0\}$ with $R_i(s) = s - S_{i,N_i^\dagger(s-)}$, where $s-$ is the time just before s. From stochastic integration theory, the compensator process of $N_i^\dagger$ is $A_i^\dagger = \{A_i^\dagger(s) : s \geq 0\}$ where $A_i^\dagger(s) = \int_0^s Y_i^\dagger(v)\lambda[R_i(v), \theta]dv$ where $\lambda(\cdot, \theta)$ is the hazard rate function of $F(\cdot, \theta)$. The martingale process with respect to the natural filtration $\mathcal{F} = \{\mathcal{F}_s : s \geq 0\}$ generated by $\{[(N_i^\dagger(s), Y_i^\dagger(s+)) : s \geq 0], i = 1, 2, \ldots, n\}$ is $M_i^\dagger = \{M_i^\dagger(s) : s \geq 0\}$ with $M_i^\dagger(s) = N_i^\dagger(s) - A_i^\dagger(s)$ being a square integrable martingale with respect to the filtration $\mathcal{F}_s$. Using the aggregate of processes that keep track of both calendar time and gap-time- $N(s,t)$, $A(s,t)$, and $M(s,t)$, Peña et al.(2001) developed, based on the data in (1), a NPMLE of the survivor function $\bar{F}(t)$ denoted by $\hat{\bar{F}}(s,t)$- which is a Kaplan-Meier type (cf. Kaplan and Meier(1958)) and given by

$$\hat{\bar{F}}(s^*, t) = \prod_{w \leq t} \left[1 - \frac{N(s^*, dw)}{Y(s^*, w)}\right], \tag{2}$$

where $s^* = \max_i \tau_i$, the maximum observation window. Furthermore, they showed that, over an appropriate Skorohod space

$$\sqrt{n}[\hat{\bar{F}}(s^*, t) - \bar{F}(t)] \rightarrow_d W, \tag{3}$$

where W is a zero-mean Gaussian process with some variance-covariance matrix $\Sigma_1(s^*, t)$.

1.2 Notation and assumptions

We consider the data in (1). We assume that the inter-event times are i.i.d. with a common absolutely continuous distribution function F. The problem of goodness of fit with recurrent event data we consider here is to test the null hypothesis H_0 that F is a member of the family $\mathcal{F}_\theta = \{F(\cdot,\theta) : \theta \in \Theta \subseteq \Re^q\}$. The NPMLE of F is as given in (2). In what follows, we let the estimate of the survivor function be $\hat{\bar{F}}(s^*,\cdot)$ and will assume the asymptotic property (3) is in force. Let θ_0 be the true parameter value of θ. We introduce the following notations.

Let $T = [0, t^*]$, where $t^* = \max_{i,j} T_{i,j}$ is the largest gap-time. The set T could also be taken to be $[0, s^*]$. We consider a subdivision of $[0, t^*]$ given by $0 = t_0^n < t_1^n < \cdots < t_k^n = t^\star$ where the end-points are functional of the data, namely $t_j^n = t_j^n(\mathcal{O}_i; i = 1, ..., n)$. The random cells are given by $I_i^n = [t_{l-1}^n, t_l^n)$ for $l = 1, \ldots, k$, and we require them to settle down as sample size increases, that is $t_l^n \to_p t_l$ and $I_l^n \to_p I_l = [t_{l-1}, t_l)$ as $n \to \infty$, under $F(\cdot, \theta_0)$, where $t_l \in [0, t^*]$. Here the notation $\to_p$ means convergence in probability.

Set $\mathbf{t}^n = (t_1^n, \ldots, t_{k-1}^n)$ and $\mathbf{t} = (t_1, ..., t_{k-1})$. The number of $T_{i,j}$ falling in the lth random cell I_l^n $(l = 1, ..., k)$ by calendar time s, that is the observed cells frequencies using the NPMLE is defined by

$$\hat{p}_l^n(s) = \int_{I_l^n} \hat{F}(s, dw) = \hat{F}(s, t_l^n) - \hat{F}(s, t_{l-1}^n). \tag{4}$$

Under the null hypothesis, the expected random cell frequencies, that is the expected number of $T_{i,j}$ falling in I_l^n are given by

$$p_l^n(\theta) = \int_{I_l^n} F(dw, \theta) = F(t_l^n, \theta) - F(t_{l-1}^n, \theta), \tag{5}$$

and these are expected, as $n \to \infty$, to stabilize to

$$p_l(\theta) = \int_{I_l} F(dw, \theta) = F(t_l, \theta) - F(t_{l-1}, \theta). \tag{6}$$

In the sequel, we introduce the corresponding vector of observed cells frequencies, expected random cells frequencies, and limiting values of expected random cells frequencies by

$$\hat{\mathbf{p}}^n(s) = [\hat{p}_l^n(s)]_{k\times 1}, \quad \mathbf{p}^n(\theta) = [p_l^n(\theta)]_{k\times 1}, \quad \mathbf{p}(\theta) = [p_l(\theta)]_{k\times 1}, \tag{7}$$

respectively. Let the lth element of a $k \times 1$-vector $\mathbf{U}_n(s, t; \theta)$ of "observed-expected" frequencies over the random cells I_l^n be defined by

$$U_n^l(s, t; \theta) = \sqrt{n}[\hat{p}_l^n(s) - p_l^n(\theta)], \quad l = 1, ..., k. \tag{8}$$

In general, a chi-square statistic has the form $\mathbf{U}_n'(s, t; \theta_n)'\hat{\mathbf{\Sigma}}\mathbf{U}_n(s, t; \theta_n)$, where $\mathbf{a}'$ denote the transpose of a vector $\mathbf{a}$, θ_n is an estimator of θ having some nice asymptotic properties, and $\hat{\mathbf{\Sigma}}$ is a $k \times k$ matrix that could possibly depends on θ_n. The matrix $\hat{\mathbf{\Sigma}}$ is-most of the time- an estimate of the Moore-Penrose generalized inverse of a consistent estimator of the in-probability limit of the variance-covariance matrix of the limiting distribution of $\mathbf{U}_n(s, t; \theta)$. At any rate, the true limiting matrix $\mathbf{\Sigma}$ is in general assumed to satisfy some regularity conditions such as positive definite and non-singularity. We now impose some assumptions that are crucial for the proof of our asymptotic results. These are the classical conditions imposed on the hypothesized distribution and the expected frequencies under H_0 for chi square type tests -however slightly changed to accommodate recurrent events.
Assumption I. There exists a neighborhood $\mathcal{N}(\theta_0)$ of θ_0 on which $F(t, \theta)$ is continuous and differentiable on $[0, t^*] \times \mathcal{N}(\theta_0)$ and the derivative of all orders are continuous in t and θ.
Assumption II. The $k \times q$ matrix

$$\nabla_{\theta'}\mathbf{p}(\theta) = \begin{bmatrix} \nabla_{\theta_1}p_1(\theta) & \cdots & \nabla_{\theta_q}p_1(\theta) \\ \vdots & \ddots & \vdots \\ \nabla_{\theta_1}p_k(\theta) & \cdots & \nabla_{\theta_q}p_k(\theta) \end{bmatrix}_{k\times q}$$

where $\nabla_\theta = \frac{\partial}{\partial\theta} \equiv (\partial/\partial\theta_j, j = 1, 2, \ldots, q)^t$ and $\nabla_{\theta_j}p_i(\theta) = \frac{\partial}{\partial\theta_j}p_i(\theta)$ is of rank q for all $\theta \in \Theta$. Other assumptions will be added as needed as we progress in the manuscript.

2. Preliminary results

Our first result in this subsection pertains to the asymptotic distribution of $\mathbf{U}_n(s, t; \theta_0)$.

Theorem 1 *Under H_0, $\boldsymbol{U}_n(s, t; \theta_0)$ converges in distribution to $N_k(\mathbf{0}, \Sigma)$ where $\Sigma = J\Sigma_1(s, t; \theta_0)J'$ and the matrix J is given by*

$$J = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & -1 \end{bmatrix}_{k\times(k-1)}$$

Furthermore, rank$(J) = k - 1$.

Proof: Define the product-limit type process by $W_n(s, t; \theta_0) = \sqrt{n}[\hat{\bar{F}}(s, t) - \bar{F}(t, \theta_0)]$, where as before θ_0 is the true value of θ. With $\xi_n = [W_n(s, t_j^n; \theta_0)]_{(k-1)\times 1}$, it is easily shown that $\mathbf{U}_n(s, t; \theta_0) = -J\xi_n$. Define $\xi_n^{(1)} = [W_n(s, t_j; \theta_0)]_{(k-1)\times 1}$ and $\xi_n^{(2)} = [W_n(s, t_j^n; \theta_0) - W_n(s, t_j; \theta_0)]_{(k-1)\times 1}$. Then $\xi_n = \xi_n^{(1)} + \xi_n^{(2)}$. Observe that $W_n(s, t; \theta_0)$ is a type of process given in (3) and its weak convergence to say $W(s, t; \theta_0)$ is given in (3). Using the Cramér-Wold device, Peña, Strawderman, and Hollander(2000) proved convergence of finite dimensional distributions of $W(s, t; \theta_0)$ to Gaussian distributions for any $t_1 < t_2 < \cdots < t_k$. The proof is complete if we can show that $\xi_n^{(2)}$ converges in probability to a $(k - 1)$- dimensional vector $\mathbf{0}_{(k-1)\times 1}$. By the representation theorem of Pollard (cf. Pollard(1879)), there exists a new probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$, new processes $\tilde{W}_n(s, t; \theta_0)$ and $\tilde{W}(s, t; \theta_0)$ such that $W_n(s, t; \theta_0) =_{ed} \tilde{W}_n(s, t; \theta_0)$ [where $=_{ed}$ means "equal in distribution"] and $W(s, t; \theta_0) =_{ed} \tilde{W}(s, t; \theta_0)$. Moreover, $\tilde{W}_n(s, t; \theta_0)$ converges weakly to $\tilde{W}(s, t; \theta_0)$ and the new processes have the same finite distributions as the old ones on their respective probability spaces. Since $\tilde{W}(s, t; \theta_0)$ has continuous sample paths, the distribution equalities imply

$$\sup_{0\le t\le t^\star} |\tilde{W}_n(s, t; \theta_0) - \tilde{W}(s, t; \theta_0)| \to_p 0$$

as $n \to \infty$. Next, for $l = 1, ..., k$, introduce $\tilde{t}_l^n = t_l^n(\tilde{\mathcal{O}}_i : i = 1, ..., n)$, the counterparts of t_l^n on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ such that $\tilde{t}_l^n =_{ed} t_l^n$. In addition, $\tilde{t}_l^n \to_p t_l$ since $t_l^n \to_p t_l$ and $\tilde{t}_l^n =_{ed} t_l^n$. It then follows, for large n that (we drop the argument θ_0)

$$\begin{aligned} |\tilde{W}_n(s, \tilde{t}_l^n) - \tilde{W}_n(s, t_l)| &\le |\tilde{W}_n(s, \tilde{t}_l^n) - \tilde{W}(s, \tilde{t}_l^n)| + |\tilde{W}(s, \tilde{t}_l^n) - \tilde{W}(s, t_l)| \\ &\quad + |\tilde{W}(s, t_l) - \tilde{W}_n(s, t_l)| \\ &\le 2 \sup_{0\le t\le t^\star} |\tilde{W}_n(s, t) - \tilde{W}(s, t)| + |\tilde{W}(s, \tilde{t}_l^n) - \tilde{W}(s, t_l)| \quad (9) \\ &\to_p 0 \end{aligned}$$

as $n \to \infty$, where the last inequality is obtained by using the continuous sample path property of $\tilde{W}(s,t;\theta_0)$. Because $\tilde{W}(s,t;\theta_0)$ has continuous sample paths, an application of the continuous mapping theorem implies that the right hand side of (9) is negligible. Therefore, $\xi_n^{(2)} \to_p 0$ and the result in the statement of the theorem follows by applying Slustky theorem. ‖

Let $\Xi(\theta;s,t)$ be the square root of the Moore-Penrose generalized inverse of $\Sigma(s,t;\theta)$. Then $\Xi(\theta;s,t)$ is a $(k \times k$ symmetric matrix whose elements are function of $(\theta, \mathbf{t})$ for fixed s. The variance-covariance matrix Σ_1 of W in (3) is a non-singular matrix. We now impose some conditions on $\Xi(\theta;s,t)$.
Condition 1: $\Xi(\theta;s,t)$ is continuous at (θ,t) for a fixed s
Condition 2: $\Xi^{-1}(\theta;s,t)$ exists and bounded on $(\Theta \times [0,s^*] \times [0,s^*])$
For brevity, let $\Xi(\theta;s,\mathbf{t}^n) \equiv \Xi_n(s,\theta)$ be the matrix obtained with $\mathbf{t}$ replaced by the random boundaries vector $\mathbf{t}^n$. Define

$$\mathbf{V}_n(\theta;s,t) = \Xi_n(s,\theta)\mathbf{U}_n(s,t;\theta). \tag{10}$$

The limiting distribution of $\mathbf{V}_n(\theta;s,t)$ under H_0 is straightforward from Theorem 1. We now imposed our third assumption about the minimum chi-square estimator of θ_0.
Assumption III Let $\{\bar{\theta}_n(s^*,t^*) : n = 1,2,...\}$ be the sequence of θ-values minimizing the sequence of quadratic forms $\{\mathbf{V}_n'(\theta;s^*,t^*)\mathbf{V}_n(\theta;s^*,t^*) : n = 1,2,...\}$. Then under the assumptions and conditions given above, $\bar{\theta}_n(s^*,t^*) \to_p \theta_0$ as $n \to \infty$.

In the above theorem, $\bar{\theta}_n(s^*,t^*)$ is the modified minimum chi square estimator of θ. When $t^* \to \infty$, the modified minimum chi square reduces to the minimum chi square estimator $\bar{\theta}_n(s^*)$. In cases were closed form solution for zeros of $\bar{\theta}_n(s^*,t^*)\ \mathbf{V}_n'(\theta;s^*,t^*)\mathbf{V}_n(\theta;s^*,t^*)$ does not exist, numerical methods such as the Newton-Raphson algorithm will be needed to minimize the quadratic form.

From now on, we abbreviate $\bar{\theta}_n(s^*,t^*)$ by $\bar{\theta}_n$. The next two lemmas pertain to a Taylor-type expansion around $\bar{\theta}_n$ for $\mathbf{p}^n(\bar{\theta}_n)$ and $\nabla_{\theta'}\mathbf{p}^n(\bar{\theta}_n)$. These two lemmas prove to be crucial in most of our asymptotic proofs.

Lemma 1 *Under the assumptions and conditions enumerated above, we have:* $[i]\ \boldsymbol{p}^n(\bar{\theta}_n) = \boldsymbol{p}(\theta_0) + o_p(1)$ *and* $[ii]\ \nabla_{\theta'}\boldsymbol{p}^n(\bar{\theta}_n) = \nabla_{\theta'}\boldsymbol{p}(\theta_0) + o_p(1)$.

Introduce the $k \times q$ matrix $\mathbf{B}(s,t;\theta_0)$ with (i,j)th entry equals to

$$\Xi(s,t;\theta_0)_{i,j}\frac{\partial p_i(t,\theta_0)}{\partial \theta_j} \tag{11}$$

for $i = 1,...,k$ and $j = 1,...,q$, so that $\mathbf{B}(s,t;\theta_0) = \Xi(s,t;\theta_0)\nabla_{\theta_0'}\mathbf{p}(t,\theta_0)$.

Lemma 2 *Under the regularity condition and assumptions, we have:*

$$\boldsymbol{B}'(s,t;\theta_0)V_n(s,t;\bar{\theta}_n) = o_p(1)$$

Proof: Since $\bar{\theta}_n$ is the value of θ minimizing the quadratic form $\mathbf{V}_n'(\theta;s,t)\mathbf{V}_n(\theta;s,t)$, it follows that for $j = 1,...,q$

$$\nabla_{\theta_j}[\mathbf{V}_n'(s,t;\bar{\theta}_n)\cdot\mathbf{V}_n(s,t;\bar{\theta}_n)] = 0, \tag{12}$$

where the symbol $\cdot$ represents the dot operator. Using the definition of $\mathbf{V}_n(s,t;\theta)$, (12) is equivalent to

$$\nabla_{\theta_j}[\mathbf{U}_n'(s,\bar{\theta}_n)\Xi^2(s,\bar{\theta}_n)\mathbf{U}_n(s,\bar{\theta}_n)] = 0. \tag{13}$$

Expending (13) and differentiating $\mathbf{U}_n'(s,\bar{\theta}_n)$ with respect to θ_j, we obtain

$$-2\sqrt{n}\nabla_{\theta_j}[\mathbf{p}'_n(\bar{\theta}_n)]\Xi^2(s,\bar{\theta}_n)\mathbf{V}_n(s,\bar{\theta}_n) + \mathbf{U}'_n(s,\bar{\theta}_n)\nabla_{\theta_j}(\Xi^2(s,\bar{\theta}_n))\mathbf{U}_n(s,\bar{\theta}_n) = 0.$$

By Condition II, $\Xi^{-1}(s,\bar{\theta}_n)$ exists and is bounded, therefore, $\mathbf{U}_n(s,\bar{\theta}_n) = \Xi^{-1}(s,\bar{\theta}_n)\mathbf{V}_n(s,\bar{\theta}_n)$. Furthermore, the asymptotic distribution of $\mathbf{U}_n(s,\theta_0)$ yields

$$\|\mathbf{U}_n(s,\theta_0)\| = O_p(1). \tag{14}$$

That the statement of the proposition holds follows using (14) and the rules of multiplication of little o and big O by bounded elements. ‖

Theorem 2 *Under the null hypothesis H_0 and Assumptions I and II above, we have*

$$\sqrt{n}(\bar{\theta}_n - \theta_0) = \boldsymbol{B}(s^*,t^*;\theta_0)[\boldsymbol{B}'(s^*,t^*;\theta_0)\boldsymbol{B}(s^*,t^*;\theta_0)]^{-1}\boldsymbol{B}'(s^*,t^*;\theta_0)\boldsymbol{V}_n(s^*,t^*;\theta_0) + o_p(1) \tag{15}$$

Proof: Start out by adding and subtracting $\mathbf{p}(\theta_0)$ in (10) to obtain (we drop the gap time argument in $\mathbf{V}$ and $\mathbf{U}$ for simplicity)

$$\begin{aligned}\mathbf{V}(s,\bar{\theta}_n) &= \Xi(s,\bar{\theta}_n)\sqrt{n}[\hat{\mathbf{p}}^n(s) - \mathbf{p}^n(\theta_0)] - \Xi(s,\bar{\theta}_n)\sqrt{n}[\hat{\mathbf{p}}^n(\bar{\theta}_n) - \mathbf{p}^n(\theta_0)] \\ &= \Xi(s,\bar{\theta}_n)\mathbf{U}(s,\theta_0) - \Xi(s,\bar{\theta}_n)\sqrt{n}[\mathbf{p}^n(\bar{\theta}_n) - \mathbf{p}^n(\theta_0)].\end{aligned}$$

Using (4), Assumption I, and Lemma 1, we obtain

$$\mathbf{V}_n(s,\bar{\theta}_n) = \Xi(s,\bar{\theta}_n)\mathbf{U}_n(s,\theta_0) - \Xi(s,\bar{\theta}_n)[\nabla_{\theta'}\mathbf{p}(\theta_0) + o_p(1)]\sqrt{n}(\bar{\theta}_n - \theta_0). \tag{16}$$

An application of Condition I, Lemma 1 to $\Xi(s;\bar{\theta}_n)$, and the rules of multiplication of little o by bounded elements yields

$$\begin{aligned}\mathbf{V}(s,\bar{\theta}_n) &= [\Xi(s,\theta_0) + o_p(1)]\mathbf{U}_n(s,\theta_0) - [\Xi(s,\theta_0) + o_p(1)][\nabla_{\theta'}\mathbf{p}(\theta_0) + o_p(1)]\sqrt{n}(\bar{\theta}_n - \theta_0) \\ &= \mathbf{V}_n(s,\theta_0) + o_p(1)\mathbf{U}_n(s,\theta_0) - [\Xi(s,\theta_0)\nabla_{\theta'}\mathbf{p}(\theta_0) + o_p(1)]\sqrt{n}(\bar{\theta}_n - \theta_0) \\ &= \mathbf{V}_n(s,\theta_0) - [\mathbf{B}(s,\theta_0) + o_p(1)]\sqrt{n}(\bar{\theta}_n - \theta_0) + o_p(1)\end{aligned} \tag{17}$$

Multiplying (17) by $\mathbf{B}'(s,t;\theta_0)$ and using the fact that $\mathbf{B}'(s,t;\theta_0)\mathbf{V}_n(s,\bar{\theta}_n)$ is negligible by Lemma 2 gives the desired result. ‖

3. Construction of the test statistic and applications

3.1 Construction of the test and large sample properties

With $\bar{\theta}_n$ being the minimum chi-square estimator, let $\mathbf{V}_n(s,\bar{\theta}_n)$ be the value of $\mathbf{V}_n(s,\theta)$ at $\bar{\theta}_n$. Furthermore, let $\mathbf{A}(s,t;\theta)$ be the $k \times k$ matrix defined by

$$\mathbf{A}(s,t;\theta) = I_k - [\mathbf{B}'(s,t;\theta)\mathbf{B}(s,t;\theta)]^{-1}\mathbf{B}'(s,t;\theta).$$

Theorem 3 *Under the regularity condition and assumptions stated above, and under H_0, we have*

$$\boldsymbol{V}_n(s^*,t^*;\bar{\theta}_n) \to_d N_k(\mathbf{0},\Gamma(s^*,t^*;\theta_0))$$

where $\Gamma(s,t;\theta_0) = \boldsymbol{A}'(s,t;\theta_0)\Xi(s,t;\theta_0)\Sigma\Xi(s,t;\theta_0)\boldsymbol{A}(s,t;\theta_0)$.

Proof: From Lemma 1 and the definition of $\mathbf{V}_n(s,t;\theta)$, the asymptotic distribution of $\mathbf{V}_n(s,t;\theta_0)$ under H_0 is given by

$$\mathbf{V}(s^*,t;\theta_0) \rightarrow_d N(\mathbf{0},\Omega(s^*,t;\theta_0)) \tag{18}$$

where $\Omega(s,t;\theta_0) = \Xi(s,t;\theta_0)\Sigma\Xi'(s,t;\theta_0)$. Taylor expending of $\mathbf{V}_n(s,t;\bar{\theta}_n)$ around θ_0 and an application of Theorem 2 to $\sqrt{n}(\bar{\theta}_n - \theta_0)$ successively yields

$$\begin{aligned}
\mathbf{V}_n(s,\bar{\theta}_n) &= \mathbf{V}_n(s,\theta_0) - [\mathbf{B} + o_p(1)]\sqrt{n}(\bar{\theta}_n - \theta_0) + o_p(1) \\
&= \mathbf{V}_n(s,\theta_0) - [\mathbf{B} + o_p(1)][(\mathbf{B}\mathbf{B}')^{-1}\mathbf{B}'\mathbf{V}_n(s,\theta_0)] + o_p(1) \\
&= [I - \mathbf{B}(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}']\mathbf{V}_n(s,\theta_0) + o_p(1).
\end{aligned}$$

Thus the limiting distribution of $\mathbf{V}_n(s^*,t;\bar{\theta}_n)$ follows upon applying standard results of multivariate normal distributions. ‖

To obtain a statistic with limiting chi-squared distribution, we first provide a consistent estimator of $\Sigma_1(s,t;\theta_0)$. The limiting variance-covariance matrix $\Sigma_1(t_1,t_2,...,t_{k-1})$ of the finite dimensional distributions of W for any $0 < t_1 < t_2 < \cdots < t_{k-1}$ under H_0 is given by

$$\Sigma_1(s,t;\theta_0) = \begin{bmatrix}
\Sigma_1(s,t_1;\theta_0) & \Sigma_1(s,t_1;\theta_0) & \cdots & \Sigma_1(s,t_1;\theta_0) \\
\Sigma_1(s,t_1;\theta_0) & \Sigma_1(s,t_2;\theta_0) & \cdots & \Sigma_1(s,t_2;\theta_0) \\
\vdots & \vdots & \ddots & \vdots \\
\Sigma_1(s,t_1;\theta_0) & \Sigma_1(s,t_2;\theta_0) & \cdots & \Sigma_1(s,t_{k-1};\theta_0)
\end{bmatrix};$$

where each $\Sigma_1(s,t_j;\theta_0)$, $j = 1,...,k-1$ is given by (cf. Peña et al.[**?**])

$$\Sigma_1(s,t_j;\theta_0) = \bar{F}^2(t_j,\theta_0)\int_0^{t_j} \frac{\log -\bar{F}(dw,\theta_0)}{y(s,w;\theta_0)}. \tag{19}$$

With $Y(s,t) = \sum_{i=1}^n Y_i(s,t)$ being the generalized at-risk process, it has been shown in the aforementioned paper that a uniformly consistent estimator of $y(s,t)$ is $\bar{Y}(s,t)$. Therefore, a natural estimator of $\Sigma_1(s,t_j;\theta_0)$ under H_0 for each j is

$$\hat{\Sigma}_1(s,t_j;\bar{\theta}_n) = \bar{F}^2(t,\bar{\theta}_n)\int_0^{t_j} \frac{\log -\bar{F}(dw,\bar{\theta}_n)}{\bar{Y}(s,t)}. \tag{20}$$

The next result pertains to the rank of the matrices $\Gamma(s,t;\theta_0)$ and $\hat{\Gamma}(s,t;\bar{\theta}_n)$. This will be used later to obtain the number of degrees of freedom of our chi-square statistic.

Theorem 4 *Under the regularity condition and assumptions, we have*
$[i]$ $rank(\Gamma(s,t;\theta_0)) = k - q - 1$
$[ii]$ $P(rank(\hat{\Gamma}(s,t;\bar{\theta}_n)) = k - q - 1) \rightarrow 1$ *as* $n \rightarrow \infty$.

Proof: $[i]$ Abbreviate $\Sigma_1(s,t;\theta_0)$, $\Xi(s,t;\theta_0)$, $\mathbf{A}(s,t;\theta_0)$ and $\Gamma(s,t;\theta_0)$ by Σ_1, Ξ, $\mathbf{A}$ and Γ respectively. We begin by noting that because Σ_1 is a positive definite matrix, there exists a matrix T via Cholesky decomposition such that $\Sigma_1 = TT'$. Therefore,

$$rank(\Gamma) = rank[\mathbf{A}\Xi JTT'J'\Xi\mathbf{A}] = rank[(\mathbf{A}\Xi JT)(\mathbf{A}\Xi JT)'],$$

and $rank(\Gamma)$ reduces to $rank[\mathbf{A}\Xi JT]$. The latter can be further reduced to $rank[\mathbf{A}\Xi J]$. The proof of $[i]$ will be completed if we can show that $rank[\mathbf{A}\Xi J] = k - q - 1$. Using results from linear models theory, that amounts to showing that

$$rank(\mathbf{A}\Xi J) = dim(\mathcal{C}(\mathbf{A}\Xi J)) = k - dim(\mathcal{C}^{\perp}(\mathbf{A}\Xi J)),$$

where $\mathcal{C}(\mathbf{A}\Xi J)$ is the column space spanned by the columns of the matrix $\mathbf{A}\Xi J$ and $\mathcal{C}^{\perp}(\mathbf{A}\Xi J)$ is the space of all vectors orthogonal to $\mathcal{C}(\mathbf{A}\Xi J)$. Let $\mathbf{B}(s,\theta)$ be the matrix with (i,j)th defined in (11). It is clear that $rank(\mathbf{B}) = q$ and $dim\,\mathcal{C}(\mathbf{B}) = q$. Let $\mathbf{e} = \Xi^{-1}\mathbf{1}$, where $\mathbf{1}$ is a $k \times 1$ vector with $(\mathbf{1})_i = 1$, for $i = 1, ..., k$. Then it is straightforward to see that using matrix multiplication

$$\mathbf{B}'\mathbf{e} = \nabla_{\theta'}[\mathbf{1}' \cdot \mathbf{p}(\theta)]' = 0. \tag{21}$$

From (21), it follows that $\mathbf{e}$ is orthogonal to $\mathcal{C}(\mathbf{B})$ and consequently $dim\,\mathcal{C}[\mathbf{B}, \mathbf{e}] = q + 1$. To complete the proof, we need to show that the space of all vectors orthogonal to $\mathcal{C}[\mathbf{B}, \mathbf{e}]$ is $\mathcal{C}[\mathbf{A}\Xi J]$. Observe that

$$\begin{aligned} (\mathbf{A}\Xi J)'\mathbf{B} &= J'\Xi\mathbf{A}\mathbf{B} = J'\Xi\mathbf{0} = \mathbf{0}, \\ (A\Xi J)'\mathbf{e} &= J'\mathbf{1} = \mathbf{0}. \end{aligned}$$

Therefore $\mathcal{C}(\mathbf{A}\Xi J)$ is orthogonal to $\mathcal{C}[\mathbf{B}, \mathbf{e}]$, hence $\mathcal{C}^{\perp}(\mathbf{A}\Xi)J) \supseteq \mathcal{C}[\mathbf{B}, \mathbf{e}]$. In a similar way, we can prove the inverse inclusion, and the result follows.
$[ii]$ This part follows from part $[i]$ and standard results on rank of uniformly consistent estimator of matrices. $\|$ We are now ready to construct our test statistic. Let $\Gamma^{-}(s^*, t; \theta_0)$ and $\hat{\Gamma}^{-}(s^*, t; \bar{\theta}_n)$ denote the Moore-Penrose generalized inverse of $\Gamma(s^*, t; \theta_0)$ and $\hat{\Gamma}^{-}(s^*, t; \bar{\theta}_n)$ respectively. From $[i]$ and $[ii]$ of Theorem 4, it follows that

$$\hat{\Gamma}^{-}(s^*, t; \bar{\theta}_n) \rightarrow_p \Gamma^{-}(s^*, t; \theta_0). \tag{22}$$

The result in (22) is key to obtaining the asymptotic distribution of our chi-square statistic, given in the next theorem.

Theorem 5 *Define* $\bar{Q}(s^*, t) = \boldsymbol{V}'(s^*, t; \bar{\theta}_n)\hat{\Gamma}^{-}(s^*, t; \bar{\theta}_n)\boldsymbol{V}(s^*, t; \bar{\theta}_n)$. *Then, under* H_0

$$\bar{Q}(s^*, t) \rightarrow_p \chi^2(k - q - 1),$$

and the test reject the hypothesized family of distributions at level α *if* $\bar{Q}(s^*, t) \geq \chi^2(k - q - 1, \alpha)$, *where* $\chi^2(k - q - 1, \alpha)$ *is the upper* α *point of* $\chi^2(k - q - 1)$.

Proof: Equation (22) and Theorem 3 together with results on multivariate normal distribution provide the result. $\|$

4. Simulation Study

4.1 Simulation Design

We perform a Monte Carlo simulation study using the **R** software on Linux platform. The LINUX server, `ph - nmx`, at the College of Public Health, The University of Iowa is used for the simulations. The goal of the simulation study is to assess estimation using minimum χ^2 methods, and to gauge the performance of our proposed random cell test with respect to nominal and achieved significance levels. For the sake of brevity, we design our simulations using 3 random partitions (I_l) of the monitoring period; although more than three partitions can be considered. In order to carry the simulation, recurrent event data must be generated during a study monitoring period [0, τ]. To generate the recurrent event data, we use a structure that reconciles the interocurrence time survivor function $\bar{F}$ and the length of the monitoring period τ–by implication, the censoring distribution $\bar{G}$. A well-known structure that reconciles $\bar{F}$ and $\bar{G}$ in the presence of recurrent event data is the generalized Koziol-Green (KG) model (Koziol & Green(1976)). The generalized KG model for a recurrent event settings postulates the existence of a monitoring parameter

$\beta > 0$ such that $\bar{G}(t) = \bar{F}(t)^{\beta}$. The parameter β controls the events intensity over the monitoring period and is reasonably constrained to (0, 1] for practical relevance– constraint that leads to more observed recurrences. We set the value of β to be 0.3 and consider estimating and testing within parametric model: the Weibull parametric lifetime models. To find the minimum chi square estimator, the quadratic form of theorem 2 was used. It is to be noted that any other estimator that is asymptotically equivalent to the MCSE provides the same asymptotic result as the MCSE (Li and Doss, 1993). In the case of recurrent events, among the class of MCSE, the minimum Hellinger distance estimator (Beran 1977) has proved to have consistently provided unbiased estimators of the parametric family.
Model: True inter event time survivor function follows a Weibull distribution; $\bar{F}(t, \theta) = \exp(-t^{\theta})$; with null survivor function $\bar{F}(t, 1)$.

Estimation
For model, we vary θ in $\{2.00, 1.50, 1.25, 1.00, .50, .25\}$. We view these values as 'true' parameters for the sake of simulation and to gauge how well they are recuperated by our estimation methods. We consider small sample as well as large sample estimation (n in $\{$ 30, 50, 100, 200$\}$). For each combination of (θ, n), we run 100 replications to estimate the parameter and carry out the test. Table 1 displays the value of a 'true' parameter (i.e. θ as set by the simulated data), the average value of the parameter across 100 replications ($\bar{\theta}$), and the standard deviation around the estimation. The true parameter value is what has been used to generate the recurrent event data. $\bar{\theta}$ and $\dot{\theta}$ are the estimated values from the simulation after we apply the minimum χ^2 framework. We found out that for the Weibull model, the more concave shape densities (i.e $\theta \gg 1$) are prone to bigger estimation errors than the convex counterparts.

Table 1: Weibull Parametric Family: $\bar{F}(t,\theta) = \exp(-t^{\theta})$

n	θ	$\bar{\theta}$	SD	$\dot{\theta}$	n	θ	$\bar{\theta}$	SD	$\dot{\theta}$
30	0.25	0.25	0.01	0.25	100	0.25	0.25	0.01	0.25
	0.50	0.50	0.03	0.50		0.50	0.50	0.01	0.50
	0.75	0.76	0.07	0.76		0.75	0.75	0.03	0.74
	1.00	1.02	0.09	1.02		1.00	1.01	0.05	1.01
	1.25	1.28	0.15	1.26		1.25	1.26	0.07	1.27
	1.50	1.53	0.20	1.52		1.50	1.53	0.11	1.54
	1.75	1.87	0.26	1.81		1.75	1.79	0.14	1.79
	2.00	2.10	0.33	2.08		2.00	2.07	0.16	2.04
	3.00	3.15	0.58	3.11		3.00	3.10	0.26	3.09
	5.00	5.42	1.13	5.40		5.00	5.21	0.50	5.18
50	0.25	0.25	0.01	0.25	200	0.25	0.25	0.01	0.25
	0.50	0.50	0.02	0.50		0.50	0.50	0.01	0.50
	0.75	0.76	0.04	0.76		0.75	0.75	0.02	0.75
	1.00	1.02	0.08	1.02		1.00	1.00	0.04	1.00
	1.25	1.28	0.10	1.28		1.25	1.25	0.05	1.25
	1.50	1.55	0.14	1.54		1.50	1.51	0.08	1.51
	1.75	1.81	0.18	1.79		1.75	1.76	0.09	1.76
	2.00	2.08	0.22	2.08		2.00	2.02	0.13	2.02
	3.00	3.18	0.43	3.15		3.00	3.04	0.18	3.05
	5.00	5.29	0.90	5.14		5.00	5.07	0.34	5.07

Table 2: Observed Significance

θ	n	$Test_1$	$Test_2$
0.5	20	0.030	0.001
	30	0.010	0.010
	50	0.001	0.001
	75	0.001	0.001
	100	0.001	0.001
	200	0.001	0.001
1.0	20	0.110	0.100
	30	0.120	0.100
	50	0.070	0.070
	75	0.050	0.050
	100	0.060	0.050
	200	0.030	0.040
3.0	20	0.190	0.100
	30	0.140	0.110
	50	0.070	0.060
	75	0.040	0.020
	100	0.020	0.010
	200	0.020	0.001

Testing:
For testing the parametric family against the NPMLE, the estimated values $\bar{\theta}$ obtained from theorem 2, $\bar{\theta}_n$ say, were plugged into the quadratic form of theorem 6 to obtain test statistics that are compared to the χ^2 distribution with one degree of freedom. The test has been set to reach a nominal significance level of 0.05. The achieved significance is represented by the proportion of these quadratic forms that cross the upper .95 quantile of the limiting χ^2 distribution (i.e. 3.8415). Table 2 displays such results for selected values of θ. We use two different tests based on two different estimations of $\Gamma(s, t, \bar{\theta}_n)$. The first estimator is based on the parametric cumulative hazard functions as outlined in equation 20. The resulting test is labeled $Test_1$. The second estimator substitutes the parametric estimation of equation 20 by its non-parametric equivalent (see for example estimator I of Adekpedjou & Zamba , 2010). The second test is labeled $Test_2$ on table 2. For Weibull parametric model, the tests are anti-conservative in small samples and tend to be conservative as sample size increases. But the test built around a non-parametric estimation of the integrated hazard is very conservative when the parameter is small. The Weibull family with decreasing hazard reacts extremely conservatively to the test in large sample than the Weibull family with increasing hazard. Anti-conservativeness remains an issue in small sample; though more pronounced when the hazard increases in time. By and large, reliability growth suffers from small and large sample conservativeness while reliability deterioration suffers from small sample anti-conservativeness and a large sample conservativeness.

5. Concluding remarks

In this manuscript, a goodness of fit test for testing whether or nor the distribution of the time between failure belongs to some parametric family of distribution is developed. The chi-square test developed is adaptive in the sense that the boundaries are data dependent and those are expected to stabilize are sample size increases. The test developed is more flexible and guaranties cells probabilities will not be small as would be the case if fixed cells probabilities were chosen. Moreover, the chi-square test fails when many cells have small expected number of observations. The test statistics is shown to be asymptotically chi-square. There are many avenues for estimating the unknown parameter of the underlying distribution. However, we found the minimum chi-square estimator more appealing because of its nice asymptotic prosperities and its equivalence to other estimator of θ for large samples.

REFERENCES

Adekpedjou, A. and Zamba, K. D. (2012), "A chi-squared goodness of fit test for recurrent event data," *Nonparametric estimation with recurrent event data*, 11, 97–119.

Li, Gang and Doss, Hani (1993), "Generalized Pearson-Fisher chi-square goodness-of-fit tests, with applications to models with life history data," *The Annals of Statistics*, 21, 772–797.

Block, H.W., Borges, W.S., and Savits (1993), "Age-sependent minimal repair," *J. Appl, Probab.* , 22, 370–385.

Koziol, J.A., and Green, S.B. (1976), "A Cramér-von Mises statistic for randomly censored data," *Biometrika* , 63, 465–474.

Kaplan, E.L. and Meier, P. (1958), "Nonparametric estimation from incomplete observations," *Journal of the American Statistical Association*, 53, 457–481.

Akritas, Michael G. (1988), "Pearson-type goodness-of-fit tests: the univariate case," *Journal of the American Statistical Association*, 83, 222–230.

Hollander, Myles and Peña, Edsel A. (1992), "A chi-squared goodness-of-fit test for randomly censored data," *Journal of the American Statistical Association* , 87, 458–463.

Kim, Joo Han (1993), "Chi-square goodness-of-fit tests for randomly censored data," *The Annals of Statistics*, 21, 1621–1639.

Habib, M. G. and Thomas, D. R. (1986), "Chi-square goodness-of-fit tests for randomly censored data," *The Annals of Statistics*, 14, 759–765.

Stocker, IV, Russell S. and Adekpedjou, A. (2011), "Optimal goodness-of-fit tests for recurrent event data," *Lifetime Data Analysis*, 17, 409–432.

Peña, E. A. and Strawderman, R. L. and Hollander, M. (2001), "Nonparametric estimation with recurrent event data," *Journal of the American Statistical Association*, 96, 1299–1315.

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

Agustin, Ma. Z. N. and Peña, E. A. (2001), "Goodness-of-fit of the distribution of time-to-first-occurrence in recurrent event models," *Lifetime Data Analysis*, 7, 289–306.

Agustin, Ma. Z. N. and Peña, E. A. (2001), "A basic approach to goodness-of-fit testing in recurrent event models," *Journal of Statistical Planning and Inference*, 133, 285–303.

Spiekerman, C. F. and Lin, D. Y. (1998), "Marginal regression models for multivariate failure time data," *Journal of the American Statistical Association*, 93, 1164–1175.

Moore, D. S. and Spruill, M. C. (1975), "Unified large-sample theory of general chi-squared statistic for tests of fit," *The Annals of Statistics*, 3, 599–616.

Mihalko, D. P. and Moore, D. S. (1980), "Chi-square tests of fit for type II censored data," *The Annals of Statistics*, 8, 625–644.

Chernoff, H. and Lehmann, E.L. (1954), "The use of maximum likelihood estimates in Chi-square tests for goodness of fit," *The Annals of Statistics*, 25, 578–586.

Moore, D. S. (1972), "A Chi-Square Statistic with Random Cell Boundaries," *The Annals of Statistics*, 42, 147–156.

Harris, R. R., and Kanji, G. K (1983), "On the Use of Minimum Chi-Square Estimation," *Journal of the Royal Statistical Society. Series D (The Statistician)*, 32, 379–394.

Pollard, D. (1979), "General Chi-square Goodness-of-fit Tests with Data-dependent Cells," *Z. Wahrscheinlichkeitheorie veiw. Gebiete*, 50, 317–331.

Dahiya, R. C. and Gurland, J. (1972), "Pearson chi-squared test of fit with random intervals," *Biometrika*, 59,

147–153.

Fisher, R.A. (1922), "On the interpretation of chi-square from contingency tables, and the calculation of P," *Journal of the Royal Statistical Society*, 85, 87–94.

Fisher, R.A. (1924), "The condition under which chi-square measures the discrepancey between observation and hypothesis," *Journal of the Royal Statistical Society*, 87, 442–450.