Estimating Load-Sharing Properties in a Dynamic Reliability System

Paul H. Kvam* **Edsel A. Peña†**

February 21, 2003

Abstract

An estimator for the load share parameters in an equal load-share model is derived based on observing k-component parallel systems of identical components that have distribution function $F(\cdot)$ and failure rate $r(\cdot)$. In an equal load share model, after the first of k components fails, failure rates for the remaining components change from $r(t)$ to $\gamma_1 r(t)$, then to $\gamma_2 r(t)$ after the next failure, and so on. On the basis of observations on n independent and identical systems, a nonparametric estimator of the component baseline cumulative hazard function $R = -\log(1-F)$ is presented, and its asymptotic limit process is established to be a Gaussian process. The effect of estimation of the load-share parameters is considered in the derivation of the limiting process. Potential applications can be found in diverse areas, including materials testing, software reliability and power plant safety assessment.

Keywords and Phrases: Dependent systems, Nelson-Aalen estimator, proportional hazards.

*P. Kvam is Associate Professor in the School of Industrial Engineering, Georgia Institute of Technology, Atlanta, GA, USA. He acknowledges NSF Grant DMI-9908035. His e-mail address is *pkvam@isye.gatech.edu*.

†E. Peña is Professor in the Department of Statistics, University of South Carolina, Columbia, SC 29208, USA. He acknowledges NSF Grant DMS 0102870. His e-mail address is *pena@stat.sc.edu*.

1 Introduction

Most reliability methods are intended for components that operate independently within a system. It is more realistic, however, to develop models that incorporate stochastic dependencies among the system's components. In many systems, the performance of a functioning component will be affected by how the other components within the system are operating or not operating (cf., Hollander and Peña, 1995). Statistical methods for analyzing systems with dependent components are not yet well developed. Real examples of dependent systems include fiber composites, software and hardware systems, power plants, automobiles, and materials subject to failure due to crack growth, to name just a few.

In the nuclear power industry, for example, components are redundantly added to systems to safeguard against core meltdown. If the failure of one back-up system adversely affects the operation of another, the probability of core meltdown can increase significantly. If four to eight motor operated valves can be employed to ensure the circulation of cooling water around the reactor, the failure of one or two valves can induce a higher rate of failure of the remaining valves due to increased water pressure, thus diminishing the effects of the component redundancy.

Unfortunately, analysts have few options for modeling dependent systems. Existing methods for such systems studied in engineering and the physical sciences are typically based on two classes of models: shock models and load-share models. Shock models, such as Marshall and Olkin's (1967) bivariate exponential model, enable the user to model component dependencies by incorporating latent variables to allow simultaneous component failures.

Load share models dictate that component failure rates depend on the operating status of the other system components and the effective system structure function. Daniels (1945) originally adopted this model to describe how the strain on yarn fibers increases as individual fibers within a bundle break. Freund (1961) formalized the probability theory for a bivariate exponential load share model. In most applications, the shock model provides an easier avenue for multivariate modeling of system component lifetimes. However, dynamic models such as the load-share model are deemed more realistic in environments where a component's performance can change once another component in the system fails or degrades.

Perhaps the most important element of the load-share model is the rule that governs how failure rates change after some components in the system fail. This rule depends on the reliability application and how the components within the system interact, i.e., through the structure function. For researchers in the textile industry who deal with the reliability of composite materials, a bundle of fibers can be considered as a parallel system subject to a steady tensile load. The rate of failure for individual fibers depends on how the unbroken fibers within the bundle share the load of this overall stress. The load share rule of such a system depends on the physical properties of the fiber composite. Yarn bundles or untwisted cables tend to spread the stress load uniformly after individual failures. This leads to an *equal load-share rule*, which implies the existence of a constant system load that is distributed equally among the working components.

In more complex settings, a bonding matrix joins the individual fibers as a composite material, and an individual fiber failure affects the load of certain surviving fibers (e.g., neighbors) more than others. This characterizes a *local load sharing rule*, where a failed component's load is transferred to adjacent components; the proportion of the load the surviving components inherit depends on their 'distance' to the failed component. A more general *monotone load sharing rule* assumes only that the load on any individual component is nondecreasing as other items fail. Lynch (1999) characterized some relationships between the failure rate and the load-share rule based on a monotone load share rule. Relationships for some specific load share rules are studied in Durham and Lynch (2000).

Past research has stressed reliability estimation based on known load share rules. To our knowledge, statistical methods have not been developed for characterizing systems with dependent components by estimating *unknown* parameters of the load-share rule. In this paper, we consider estimating the component baseline lifetime distribution based on observing dynamic systems of identical components. Dependence between system components is modeled through a load-share framework, with the load-sharing rule containing unknown parameters. Of primary interest in the model is the baseline distribution, but the parameters of the load-sharing rule may also be of importance such as when an estimate of the system reliability is desired, or they could just be viewed as nuisance parameters. We focus on the *equal load-share rule*, where the failure rate of the remaining functioning components within the system change uniformly after each component failure, but with the magnitudes of change being unknown.

2 Examples of Load-Share Systems

The load share rule has obvious potential for application in modeling systems with interdependent components, as described in the preceding section. The load-sharing framework also applies to problems of detecting members of a finite population. Suppose the resources allocated toward finding a finite set of items are defined globally, rather than assigned individually. Once items are detected, resources can be redistributed for the problem of detecting the remaining items, and this action gives rise to a load sharing model. In most cases, the items are identical to the observer, and an equal load-share rule is appropriate for characterizing the system dependence.

Unlike load-share models for fiber strength, these more general models give no indication of how load share parameters might change as other components fail. In this case, inference based on known load share parameters seems unrealistic and the problem of estimating those parameters becomes crucial.

We have already discussed two examples for which the load-share rule might apply: risk assessment in power plants, and the study of fiber strength in relation to fiber composites in textile engineering. Other important examples include the following:

Software Reliability: The load-sharing model generalizes the dynamic model suggested by Jelinski and Moranda (1972), among others, for software reliability. The most basic problem is to assume that an unknown number of faults exist in the system (i.e., software). After a fixed time, some number of faults are found, and the number of remaining faults is to be estimated. The load-share model represents a more flexible and realistic method of predicting the detection of faults by acknowledging the dynamic nature of fault detection when some faults have already been found. For instance, in problems where the number of software bugs is relatively small, the discovery of a major defect can help conceal or reveal other existing bugs in the software.

Civil Engineering: With a large structure supported by welded joints, the structure fails only after a series of supporting joints fail. The failure of one or two welded joints in a bridge support, for instance, might cause the stress on remaining joints to increase, thus causing earlier subsequent failures. Static reliability models fail to consider the changing stress in this setting, which constitutes a load-sharing model.

Materials Testing: Fatigue and material degradation is often characterized by crack growth, especially in large structures such as an airplane engine turbine or a commercial airplane fuselage. At the microscopic level, these materials have an intractable number of cracks, with only a few becoming large enough to be measurable, usually at stress centers such as edges, rivets, etc. It is known that the largest crack in a (predefined) local area will inherit much of the test stress, and thus will grow at a faster rate than the other measurable cracks; see Carlson and Kardomateas (1996) for instance. This provides a platform for extending the load-share model to degradation data. Certainly, the interdependence between crack growths cannot be modeled using simple physical principles, thus a nonparametric load-share model has potential application.

A similar approach, used in modeling the incubation period for the Human Immunodeficiency Virus (HIV) in Jewell and Kalbfleisch (1996), is based on marker processes. Rate changes can be incorporated into the model via time-dependent stochastic markers that carry covariate information. Marker processes are based on the shock model approach to describing component dependence, but are closely related to load-sharing models. As an illustrative reliability example, a car's odometer serves as an obvious marker for the car's chronological lifetime. This approach serves as a natural one for modeling crack growth in materials using observed degradation (e.g., crack size) as a stochastic marker.

Population Sampling: In wildlife studies, population sizes are estimated from relatively small samples. Capture/recapture methods can be used for these estimation methods, and involve finding previously tagged animals in order to deduce the sample's size relative to the larger population. In some cases, the detection of a tagged animal may affect the detection rate of the remaining sample. When recapture probabilities are significantly nonzero, the load-share framework allows the experimenter to modify the detection model after a recapture occurs.

Combat Modeling: The attrition of military hardware and personnel in combat situations is highly dynamic, and the loss of one component in combat can easily change the success rate (or death rate) of the remaining components in the field; see Kvam and Day (2001). Specific load share models could be used to model the natural dependence between components within the system as well as their relative status within the group (e.g., even with combat machines, the components are not generally identical in effectiveness or constitution).

3 Estimation of Load-Share Model Parameters

Consider a system with k identical components for which stochastic component dependencies are induced via a load sharing model. Suppose we observe n independent and identical systems over an observation period $[0, \tau]$, where τ is possibly random and could be the time of the last component failure among all nk components. We monitor the times of component failures of these systems.

For $i = 1, 2, 3, \ldots$, let $S_{i,1} < S_{i,2} < \ldots$ be the successive component failure times for the ith system whose values are less than or equal to τ, so that $S_{i,j}$ is the jth smallest component failure time for the ith system. Denote by F the baseline component failure time distribution. The hazard function (or cumulative hazard rate) corresponding to F is $R(x) = -\log(1 - F(x))$, and the hazard rate is $r(x) = f(x)/[1 - F(x)]$, where $f(x)$ is the density of F. Thus, the hazard function can be expressed as $R(x) = \int_0^x r(u)du$.

Inter-component dependencies are due to the fact that the system's environment can possibly become more or less harsh on the remaining functioning components upon failure of other components. This framework is based on applications for which failure rates or detection rates of all items within the system are equal, but the change in rate after a component failure depends on the set of functioning components in the system. Note that upon a component failure, the effective system structure function also changes (cf., Hollander and Peña, 1995). For the specific model considered in the present paper, until the first component failure, the failure rate of each of k components in the system equals the baseline rate $r(x)$. Upon the first failure within a system, the failure rates of the $k - 1$ remaining components jump to $\gamma_1 r(x)$, and remain at that rate until the next component failure. After this failure, the failure rates of the $k - 2$ surviving components jump to $\gamma_2 r(x)$, and so on. The failure rate of the last remaining component is $\gamma_{k-1} r(x)$. The (equal) load share rule can be characterized by the $k - 1$ unknown parameters $\gamma_1, \gamma_2, \ldots, \gamma_{k-1}$ and the unknown baseline distribution or hazard function. For example, a system with a constant load would assign $\gamma_j = k/(k - j), \quad j = 1, ..., k - 1$. In the sequel, we let

$$\boldsymbol{\gamma} = (\gamma_0 \equiv 1, \gamma_1, \ldots, \gamma_{k-1})'.$$

Estimating the underlying baseline functions F or R may be of primary interest. For this nonparametric estimation problem, we can adopt the generalized maximum likelihood approach introduced by Kiefer and Wolfowitz (1956). In some situations such as when estimation of the *system* reliability is desired, estimation of the load share parameters $\boldsymbol{\gamma}$ will also be of interest; otherwise it may be viewed as a vector of nuisance parameters. In the ith system, the conditional hazard function of the $(j+1)$ smallest component lifetime $S_{i,j+1}$, given the first i component failure times $S_{i,1},\ldots,S_{i,j}$, is

$$R^*(s|S_{i,1},\ldots,S_{i,j}) = \gamma_j R(s) + (\gamma_{j-1}-\gamma_j)R(S_{i,j}) + \ldots + (1-\gamma_1)R(S_{i,1})$$

for $s > S_{i,j}$. This is immediately evident if we express R^* in terms of r and note that r changes at times of component failures: $r(x)$ to $\gamma_1 r(x)$ to $\gamma_2 r(x)$, and so on. We write $dR^*(s) = R^*(s) - R^*(s-)$, so $dR^*(S_{i,j}) = \gamma_{j-1}[R(S_{i,j}) - R(S_{i,j}-)]$, where for a function h, $h(s-) = \lim_{a\downarrow 0} h(s-a)$. In terms of R^*, the observed likelihood corresponding to the n systems and assuming that all nk components fail can be expressed via

$$\prod_{i=1}^{n}\prod_{j=1}^{k} dR^*(S_{ij})\exp\{-R^*(S_{ij})\}.$$

In terms of the baseline hazard function R and the unknown load share parameters,

$$\begin{aligned} & L(R,\boldsymbol{\gamma}|\{S_{i,j}, i=1,\ldots,n, j=1,\ldots,k\}) \\ &= \left(\prod_{i=1}^{n}\prod_{j=1}^{k} dR(S_{i,j})\right)\left(\prod_{j=1}^{k-1}\gamma_j^n\right)\times \\ &\quad \exp\left(-\sum_{i=1}^{n}\sum_{j=1}^{k}\left[\gamma_{j-1}R(S_{i,j}) + \sum_{r=1}^{j-1}(\gamma_{r-1}-\gamma_r)R(S_{i,r})\right]\right). \end{aligned} \quad (1)$$

A standard approach to obtaining the estimator of $R(\cdot)$ from (1) is to first fix $\boldsymbol{\gamma}$, and then maximize the likelihood with respect to R to obtain $\hat{R}(\cdot;\boldsymbol{\gamma})$. This $\hat{R}(\cdot;\boldsymbol{\gamma})$ is then plugged into (1) to obtain the profile likelihood for $\boldsymbol{\gamma}$, which is maximized in $\boldsymbol{\gamma}$ to obtain the estimator $\hat{\boldsymbol{\gamma}}$. The final estimator of $R(\cdot)$ is $\hat{R}(\cdot;\hat{\boldsymbol{\gamma}})$. Instead of directly proceeding as described, we approach the problem using point process theory which allows us later to derive asymptotic properties of the estimator of $R(\cdot)$ in a broader framework. The resulting estimators of $\boldsymbol{\gamma}$ and $R(\cdot)$ in this stochastic process approach are, however, the same estimators that arise by proceeding in the approach outlined above.

To shift into this stochastic process framework, in the sequel we sometimes write $\gamma[j]$ for γ_j, and define the counting processes

$$N_i(t) = \sum_{j=1}^{k} I(S_{i,j} \leq t), \quad i = 1, 2, \ldots, n,$$

where $I(W)$ equals one if event W occurs, and is zero otherwise. $N_i(t)$ represents the number of component failures for the ith system that occurred on or before time t. Then, we may write

$$\gamma[N_i(w)] = \sum_{j=0}^{k-1} \gamma_j I(N_i(w) = j).$$

To express the likelihood in terms of stochastic processes, we let

$$Y_i(w) = (k - N_i(w-))\, I(\tau \geq w). \tag{2}$$

Define $\mathcal{F}_{it} = \sigma\{(N_i(w), Y_i(w+)); w \leq t\}$ to be the filtration generated by the ith system up to time t, and let $\mathcal{F}_t = \bigvee_{i=1}^{n} \mathcal{F}_{it}$. The load-share model can now be equivalently described by specifying the intensities of the $N_i(\cdot)$'s to be

$$\Pr\{dN_i(t) = 1 | \mathcal{F}_{it-}\} = r(t) Y_i(t) \gamma[N_i(t-)] dt, \qquad i = 1, ..., n. \tag{3}$$

If we denote by

$$A_i(t) = \int_0^t \gamma[N_i(u-)] r(u) Y_i(u) du, \tag{4}$$

then $\boldsymbol{M} = \{(M_i(t) = N_i(t) - A_i(t), 0 \leq t \leq \tau), i = 1, .., n\}$ is a vector of orthogonal square-integrable zero-mean martingales (cf., Andersen, et al., 1993). We will also make use of the process $J(w) = I(\sum_{i=1}^{k} Y_i(w) > 0)$. In particular, $J(w) = 0$ indicates all nk components have already failed at time $w-$. Note also that $J(\cdot)$ is a predictable and bounded process.

If $\boldsymbol{\gamma}$ is known, by using the zero-mean property of the martingale $\sum_{i=1}^{n} M_i(\cdot)$ analogously to the derivation of the Nelson-Aalen estimator (Aalen (1978)), we immediately obtain the estimator

$$\hat{R}(s; \boldsymbol{\gamma}) = \int_0^s \frac{J(w) dN(w)}{\sum_{i=1}^{n} Y_i(w) \gamma[N_i(w-)]}. \tag{5}$$

The estimator in (5) is similar in structure to the hazard function estimator for tensile strengths derived by Ryden (1999). To obtain the estimator of $R(\cdot)$ for the more general case where $\boldsymbol{\gamma}$ is unknown, we first obtain the profile likelihood for $\boldsymbol{\gamma}$ by plugging in $\hat{R}(\cdot;\boldsymbol{\gamma})$ from (5) into the likelihood function in (1). From (1) and (3) we obtain this profile likelihood for $\boldsymbol{\gamma}$ to be

$$L_p(s;\boldsymbol{\gamma}) = \prod_{i=1}^{n} \prod_{0\leq w\leq s} \left[\frac{Y_i(w)\gamma[N_i(w-)]}{\sum_{l=1}^{n} Y_l(w)\gamma[N_l(w-)]}\right]^{dN_i(w)}, \tag{6}$$

where the second product in (6) denotes the product-integral. This profile likelihood may also be viewed as a partial likelihood process. Once $\hat{\boldsymbol{\gamma}}$ is obtained from (6), the estimator of R becomes

$$\hat{R}(s) = \hat{R}(s;\hat{\gamma}).$$

By virtue of the product representation of $\bar{F} = 1 - F$ in terms of R given by $\bar{F}(s) = \prod_{0\leq w\leq s}[1 - R(dw)]$, we then obtain an estimator of $\bar{F}$ via

$$\hat{\bar{F}}(s) = \prod_{0\leq w\leq s}\left[1 - \hat{R}(dw)\right].$$

To find $\hat{R}(s)$ and to facilitate the presentation of asymptotic properties of the estimators $\hat{R}$ and $\hat{\boldsymbol{\gamma}}$, we introduce the following notation:

- $Q_{i,j}(t) = Y_i(t)I(N_i(t-) = j)$, $1 \leq i \leq n, 0 \leq j \leq k-1$;
- $\boldsymbol{Q}_i(t) = (Q_{i,0}(t), ..., Q_{i,k-1}(t))'$, $1 \leq i \leq n$;
- $\boldsymbol{Q}(t) = (\sum_{i=1}^{n} Q_{i,0}(t), ..., \sum_{i=1}^{n} Q_{i,k-1}(t))'$;
- $\boldsymbol{\delta}_i(t) = (\delta_{i,0}(t), ..., \delta_{i,k-1}(t))'$, with $\delta_{i,j}(t) = I(Q_{i,j}(t) > 0), 1 \leq i \leq n$;
- $\boldsymbol{\gamma}^{-1} \equiv (1/\gamma_0, ..., 1/\gamma_{k-1})$;
- $\boldsymbol{q}(s) = (q_0(s), \ldots, q_{k-1}(s))$, with

$$q_j(w) = E(Q_{i,j}(w)) = (k-j)P(\tau \geq w, N_1(w-) = j);$$

- $\hat{\boldsymbol{\rho}}(t;\boldsymbol{\gamma}) = \sum_{i=1}^{n} \boldsymbol{\gamma} * \boldsymbol{Q}_i(t)/\boldsymbol{\gamma}'\boldsymbol{Q}(t)$;

- $\boldsymbol{\rho}(t;\boldsymbol{\gamma}) = E[\sum_{i=1}^{n} \boldsymbol{\gamma} * \boldsymbol{Q}_i(t)]/E[\boldsymbol{\gamma}'\boldsymbol{Q}(t)] = \boldsymbol{\gamma} * \boldsymbol{q}(t)(\boldsymbol{\gamma}'\boldsymbol{q}(t))^{-1}$.

Here, $*$ represents component-by-component multiplication. With this notation, $\hat{R}(\cdot,\boldsymbol{\gamma})$ becomes

$$\hat{R}(s;\boldsymbol{\gamma}) = \int_0^s J(w)(\boldsymbol{\gamma}'\boldsymbol{Q}(w))^{-1} dN(w),$$

while the profile log-likelihood becomes

$$\begin{aligned}\ell_p(s;\boldsymbol{\gamma}) &= \log L_p(s;\boldsymbol{\gamma})\\ &= \sum_{i=1}^{n}\int_0^s \log[\boldsymbol{\gamma}'\boldsymbol{Q}_i(w)]dN_i(w) - \int_0^s \log[\boldsymbol{\gamma}'\boldsymbol{Q}(w))]dN(w).\end{aligned}$$

The corresponding profile (partial) score function for $\boldsymbol{\gamma}$ is

$$\boldsymbol{U}(s;\boldsymbol{\gamma}) \equiv \nabla_{\boldsymbol{\gamma}}\ell_p(s;\boldsymbol{\gamma}) = \sum_{i=1}^{n}\int_0^s \left[\frac{\boldsymbol{Q}_i(w)}{\boldsymbol{\gamma}'\boldsymbol{Q}_i(w)} - \frac{\boldsymbol{Q}(w)}{\boldsymbol{\gamma}'\boldsymbol{Q}(w)}\right] dN_i(w) \tag{7}$$

and the profile information matrix is

$$\boldsymbol{I}(s;\boldsymbol{\gamma}) \equiv -\nabla_{\boldsymbol{\gamma}'}\nabla_{\boldsymbol{\gamma}}\ell_p(s;\boldsymbol{\gamma}) = \sum_{i=1}^{n}\int_0^s \left[\frac{\boldsymbol{Q}_i(w)\boldsymbol{Q}_i(w)'}{(\boldsymbol{\gamma}'\boldsymbol{Q}_i(w))^2} - \frac{\boldsymbol{Q}(w)\boldsymbol{Q}(w)'}{(\boldsymbol{\gamma}'\boldsymbol{Q}(w))^2}\right] dN_i(w). \tag{8}$$

If we ignore differentiation by the known constant $\gamma_0 = 1$, $\boldsymbol{U}$ is a vector of length $(k-1)$, and $\boldsymbol{I}$ is a $(k-1)\times(k-1)$ matrix.

Using the notation defined earlier, (7) and (8) can be further simplified by noting that $\boldsymbol{Q}_i(w)(\boldsymbol{\gamma}'\boldsymbol{Q}_i(w))^{-1} = \boldsymbol{\gamma}^{-1} * \boldsymbol{\delta}_i(w)$ and $\boldsymbol{Q}(w)(\boldsymbol{\gamma}'\boldsymbol{Q}(w))^{-1} = \boldsymbol{\gamma}^{-1} * \hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})$. In terms of $\hat{\boldsymbol{\rho}}$,

$$\boldsymbol{U}(s;\boldsymbol{\gamma}) = \boldsymbol{\gamma}^{-1} * \sum_{i=1}^{n}\int_0^s [\boldsymbol{\delta}_i(w) - \hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})]\, dN_i(w).$$

We let the symbol $\boldsymbol{D}(\boldsymbol{\eta})$ represent a diagonal matrix with diagonal elements $\boldsymbol{\eta}$. Because $\boldsymbol{Q}_i(w)\boldsymbol{Q}_i(w)' = \boldsymbol{D}(Q_{i,j}(w))$, $j = 0,...,k-1$, the first term in the integrand in (8) can be written as $\boldsymbol{D}(\boldsymbol{\gamma}^{-1})\boldsymbol{D}(\boldsymbol{\delta}(w))\boldsymbol{D}(\boldsymbol{\gamma}^{-1})$, and the second term as $\boldsymbol{D}(\boldsymbol{\gamma}^{-1})\hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})\hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})'\boldsymbol{D}(\boldsymbol{\gamma}^{-1})$, so equations (7) and (8) become

$$\boldsymbol{U}(s;\boldsymbol{\gamma}) = \boldsymbol{D}(\boldsymbol{\gamma}^{-1}) \textstyle\sum_{i=1}^{n}\int_0^s [\boldsymbol{\delta}_i(w) - \hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})]\, dN_i(w),$$
$$\boldsymbol{I}(s;\boldsymbol{\gamma}) = \boldsymbol{D}(\boldsymbol{\gamma}^{-1}) \left(\textstyle\sum_{i=1}^{n}\int_0^s [\boldsymbol{D}(\boldsymbol{\delta}_i(w)) - \hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})\hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})']\, dN_i(w)\right) \boldsymbol{D}(\boldsymbol{\gamma}^{-1}).$$

Solving the set of $k-1$ nonlinear equations

$$\boldsymbol{U}(\tau;\boldsymbol{\gamma}) = \sum_{i=1}^{n} \int_0^{\tau} \left[\frac{\boldsymbol{Q}_i(w)}{\boldsymbol{\gamma}'\boldsymbol{Q}_i(w)} - \frac{\boldsymbol{Q}(w)}{\boldsymbol{\gamma}'\boldsymbol{Q}(w)} \right] dN_i(w) = \mathbf{0} \tag{9}$$

does not lead to a closed form solution for the MLE of $\boldsymbol{\gamma}$. However, solving the set of equations is not a difficult numerical problem. For instance, a Newton-Raphson method could be implemented, which has the iterations

$$\boldsymbol{\gamma}_{new} \leftarrow \boldsymbol{\gamma}_{old} + \boldsymbol{I}(\tau, \boldsymbol{\gamma}_{old})^{-1} \boldsymbol{U}(\tau; \boldsymbol{\gamma}_{old}).$$

Other approaches could also be used to solve (9); a similar set of equations is in Kvam and Samaniego (1993) for solving the likelihood equations in an exponential factorial model, and as in that paper, applying Theorem 2.1 of Mäkeläinen, Schmidt and Styan (1981) establishes that there exists a unique solution $\hat{\boldsymbol{\gamma}} \geq \mathbf{0}$ satisfying $\boldsymbol{U}(\tau;\hat{\boldsymbol{\gamma}}) = 0$. In Kvam and Samaniego (1993), a nonlinear Gauss-Seidel iterative method (see Ortega and Rheinboldt (1970), for example) was applied to solve the set of equations.

4 Asymptotic Properties

To obtain the asymptotic properties of the estimator $\hat{\boldsymbol{\gamma}}$ which solves $\boldsymbol{U}(\tau, \hat{\boldsymbol{\gamma}}) = \mathbf{0}$ in (9), we re-express the martingale $\boldsymbol{M}$ in terms of $\{\boldsymbol{Q}_i, \boldsymbol{Q}\}$. First, note that from (4), the compensator of M_i is

$$A_i(s) = \sum_{j=0}^{k-1} \gamma_j \int_0^s Q_{i,j}(w) dR(w) = \int_0^s \boldsymbol{\gamma}'\boldsymbol{Q}_i(w) dR(w), \tag{10}$$

so the quadratic variation process of M_i is

$$\langle M_i(\cdot;\boldsymbol{\gamma}) \rangle (s) = \int_0^s \boldsymbol{\gamma}'\boldsymbol{Q}_i(w) dR(w).$$

Result 1: In terms of $\boldsymbol{M}$, the score function in (9) can be simplified to

$$\boldsymbol{U}(s;\boldsymbol{\gamma}) = \sum_{i=1}^{n} \int_0^s \left[\boldsymbol{\delta}_i(w) - \hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma}) \right] dM_i(w).$$

The proof of this result is presented in the Appendix. This simplification leads us to the following asymptotic properties for the score process. Their proofs are also relegated to the Appendix.

Lemma 1 *If $\{N_i(\cdot), i = 1, ..., n\}$ are independent and identically distributed, and*

$$\inf_{0 \leq w \leq \tau} \sum_{j=0}^{k-1} (k-j)\gamma_j P(N_1(w-) = j) > 0,$$

then the score function in (7)*, which equals*

$$\boldsymbol{U}(s;\boldsymbol{\gamma}) = \sum_{i=1}^{n} \int_0^s [\boldsymbol{\delta}_i(w) - \hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})]\, dM_i(w), \tag{11}$$

is a square-integrable martingale with quadratic variation process $\langle \boldsymbol{U}(\cdot;\boldsymbol{\gamma})\rangle(s)$ whose in-probability limit is

$$\boldsymbol{\Upsilon}(s;\boldsymbol{\gamma}) \equiv \int_0^s [\boldsymbol{D}(\boldsymbol{\rho}(w;\boldsymbol{\gamma})) - \boldsymbol{\rho}(w;\boldsymbol{\gamma})\boldsymbol{\rho}(w;\boldsymbol{\gamma})']\, \boldsymbol{\gamma}'\boldsymbol{q}(w) dR(w).$$

Furthermore, $n^{-1/2}\boldsymbol{U}(\cdot;\boldsymbol{\gamma})$ converges weakly to a zero-mean Gaussian process with covariance matrix function $\boldsymbol{\Upsilon}(\cdot;\boldsymbol{\gamma})$.

Based on the asymptotic properties of $\boldsymbol{U}(s;\boldsymbol{\gamma})$, the asymptotic properties of $\hat{\boldsymbol{\gamma}}$ and $\hat{R}(s)$ are presented in the following theorems.

Theorem 1 *Under the conditions of Lemma 1,*

(i) $\hat{\boldsymbol{\gamma}}$ converges in probability to $\boldsymbol{\gamma}$; and

(ii) $\sqrt{n}(\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}) \xrightarrow{d} N(\boldsymbol{0}, \boldsymbol{\Sigma}(\tau,\boldsymbol{\gamma}))$ where $\boldsymbol{\Sigma}(\tau,\boldsymbol{\gamma}) = \boldsymbol{D}(\boldsymbol{\gamma})\boldsymbol{\Upsilon}(\tau,\boldsymbol{\gamma})^{-1}\boldsymbol{D}(\boldsymbol{\gamma})$, and with

$$\boldsymbol{\Upsilon}(\tau,\boldsymbol{\gamma}) \equiv \int_0^\tau [\boldsymbol{D}(\boldsymbol{\rho}(w;\boldsymbol{\gamma})) - \boldsymbol{\rho}(w;\boldsymbol{\gamma})\boldsymbol{\rho}(w;\boldsymbol{\gamma})']\, \boldsymbol{\gamma}'\boldsymbol{q}(w) dR(w).$$

Theorem 2 *Under the conditions of Lemma 1, if τ is such that $\boldsymbol{\gamma}'\boldsymbol{q}(\tau) > 0$, then*

$$\left\{\sqrt{n}(\hat{R}(s) - R(s)) : 0 \leq s \leq \tau\right\}$$

converges weakly to a zero-mean Gaussian process with variance function

$$\Xi(s;\boldsymbol{\gamma}) \equiv \int_0^s \{\boldsymbol{\gamma}'\boldsymbol{q}(w)\}^{-1} dR(w) + \boldsymbol{\varrho}(s;\boldsymbol{\gamma})'[\boldsymbol{\Upsilon}(\tau;\boldsymbol{\gamma})]^{-1}\boldsymbol{\varrho}(s;\boldsymbol{\gamma}),$$

where $\boldsymbol{\varrho}(s;\boldsymbol{\gamma}) = \int_0^s \boldsymbol{\rho}(w;\boldsymbol{\gamma}) dR(w)$.

Corollary 1 *Under the conditions of Theorem 2,*

$$\left\{\sqrt{n}(\hat{\hat{F}}(s) - \bar{F}(s)) : 0 \leq s \leq \tau\right\}$$

converges weakly to a zero-mean Gaussian process $\{Z(s) : 0 \leq s \leq \tau\}$ *whose variance function is* $\mathbf{Var}\{Z(s)\} = \bar{F}(s)^2\,\Xi(s; \boldsymbol{\gamma})$.

We attempt to provide an explicit expression for the limiting variance functions. To try to do so, an expression for $\Pr\{N_1(w-) = j\}$ is needed in order to get an expression for

$$\begin{aligned} q_j(w) = E\{Q_j(w)\} &= (k-j)\Pr\{\tau \geq w, N_1(w-) = j\} \\ &= I\{\tau \geq w\}(k-j)\Pr\{N_1(w-) = j\} \end{aligned}$$

when τ is fixed. We now note that

$$\begin{aligned} \Pr\{N_1(w-) = j\} &= \Pr\{N_1(w-) \geq j\} - \Pr\{N_1(w-) \geq j+1\} \\ &= \Pr\{S_j < w\} - \Pr\{S_{j+1} < w\} \\ &= \Pr\{S_{j+1} \geq w\} - \Pr\{S_j \geq w\}. \end{aligned}$$

Theorem 5.1 of Hollander and Peña (1995) then provides an expression for $\Pr\{S_j \geq w\}$. Introducing the notation

$$\zeta_{i,j}(\boldsymbol{\gamma}, k) = \prod_{l=0;l\neq i}^{j} \left[\frac{\gamma_l(k-l)}{\gamma_l(k-l) - \gamma_i(k-i)}\right]$$

for $i \leq j$ and $i, j \in \{0, 1, 2, \ldots, k-1\}$, with the convention $\prod_{\emptyset} = 1$, from Hollander and Peña (1995) we have

$$\Pr\{S_j \geq w\} = I\{j \geq 1\} \sum_{i=0}^{j-1} \zeta_{i,j-1}(\boldsymbol{\gamma}, k) \exp\{-\gamma_i(k-i)R(w)\}. \qquad (12)$$

Consequently, for $j = 0, 1, \ldots, k-1$,

$$q_j(w) = I\{\tau \geq w\}(k-j) \left\{\sum_{i=0}^{j} \zeta_{i,j}(\boldsymbol{\gamma}, k) \exp\{-\gamma_i(k-i)R(w)\} - I\{j \geq 1\} \sum_{i=0}^{j-1} \zeta_{i,j-1}(\boldsymbol{\gamma}, k) \exp\{-\gamma_i(k-i)R(w)\}\right\}.$$

Unfortunately, this does not yield a simple expression for $\boldsymbol{\Xi}(s;\boldsymbol{\gamma})$. For instance, the first term of this limiting variance function is given by

$$\int_0^s \frac{dR(w)}{\sum_{j=0}^{k-1}\gamma_j q_j(w)}$$

$$= \int_0^s \left[I\{\tau \geq w\} \sum_{j=0}^{k-1} \gamma_j (k-j) \left\{ \sum_{i=0}^{j} \zeta_{i,j}(\boldsymbol{\gamma},k) \exp\{-\gamma_i (k-i) R(w)\} - \right.\right.$$

$$\left.\left. I\{j \geq 1\} \sum_{i=0}^{j-1} \zeta_{i,j-1}(\boldsymbol{\gamma},k) \exp\{-\gamma_i(k-i)R(w)\} \right\} \right]^{-1} dR(w)$$

$$= \int_0^{R(s\wedge\tau)} \left[k\exp(-kv) + \sum_{j=1}^{k-1} \gamma_j (k-j) \times \right.$$

$$\left. \left\{ \sum_{i=0}^{j} \zeta_{i,j}(\boldsymbol{\gamma},k) \exp\{-\gamma_i(k-i)v\} - \sum_{i=0}^{j-1} \zeta_{i,j-1} \exp\{-\gamma_i(k-i)v\} \right\} \right]^{-1} dv.$$

Note that this expression is at most equal to

$$\boldsymbol{\Xi}^*(s) = \int_0^{R(s\wedge\tau)} \frac{dv}{k\exp(-kv)} = \frac{1}{k^2}\left(\frac{1-\exp\{-kR(s\wedge\tau)\}}{\exp\{-kR(s\wedge\tau)\}} \right),$$

which is the asymptotic variance function of the Nelson-Aalen estimator of $R(s)$ which utilizes *only* the first component failure for each system and given by

$$\tilde{R}(s) = \frac{1}{k} \sum_{\{i:\ S_{i1} \leq s\}} \left[\frac{1}{\sum_{j=1}^{n} I\{S_{j1} \geq S_{i1}\}} \right]. \tag{13}$$

This particular result demonstrates that if the $\boldsymbol{\gamma}$ is known, then the estimator $\hat{R}(s)$ is more efficient than the estimator $\tilde{R}(s)$, which of course is not a surprising result. However, since $\boldsymbol{\gamma}$ is not known and is estimated to form the estimator $\hat{R}(s)$, the second term in $\boldsymbol{\Xi}(s;\boldsymbol{\gamma})$, which is given by

$$\left(\int_0^s \rho(w;\boldsymbol{\gamma})dR(w)\right)' \{\boldsymbol{\Upsilon}(\tau;\boldsymbol{\gamma})\}^{-1} \left(\int_0^s \rho(w;\boldsymbol{\gamma})dR(w)\right),$$

need to be taken into account in comparing the asymptotic variance of $\hat{R}(s)$ with that of $\tilde{R}(s)$. Recall that this term is the effect of the estimation of $\boldsymbol{\gamma}$ by $\hat{\boldsymbol{\gamma}}$.

We now prove at this point that indeed, for two-component parallel systems, i.e., $k=2$, the estimator $\hat{R}(\cdot)$ improves on the estimator $\tilde{R}(\cdot)$ by showing that the asymptotic variance of the former is at most that of the latter. We have been unsuccessful thus far, but certainly have not given up, in our attempt to formally establish this domination result for the case where $k > 2$.

For notation, let us define

$$\Delta(s;\tau) = \boldsymbol{\Xi}^*(s) - \boldsymbol{\Xi}(s).$$

Since $\boldsymbol{\Xi}^*(s) = \int_0^s \{q_0(w)\}^{-1} dR(w)$, then it follows that

$$\Delta(s;\tau) = \int_0^s \left(\frac{\boldsymbol{\gamma}'\boldsymbol{q}}{q_0(q_0 + \boldsymbol{\gamma}'\boldsymbol{q})} \right) dR - \left(\int_0^s \boldsymbol{\rho} dR \right)' \{\boldsymbol{\Upsilon}(\tau)\}^{-1} \left(\int_0^s \boldsymbol{\rho} dR \right). \quad (14)$$

Theorem 3 *For $k = 2$, $\Delta(s;\tau) \geq 0$ for $s \leq \tau$, implying that the estimator $\hat{R}(\cdot)$ is asymptotically never less efficient than the estimator $\tilde{R}(\cdot)$.*

Proof: First we note that when $k = 2$, and since $\rho_1 = \gamma_1 q_1/(q_0 + \gamma_1 q_1)$, then

$$\begin{aligned}
\boldsymbol{\Upsilon}(\tau) &= \int_0^\tau \rho_1(1-\rho_1)(q_0 + \gamma_1 q_1) dR \\
&= \int_0^\tau \left(\frac{\gamma_1 q_1}{q_0 + \gamma_1 q_1} \right) \left(\frac{q_0}{q_0 + \gamma_1 q_1} \right) (q_0 + \gamma_1 q_1) dR \\
&= \int_0^\tau \left(\frac{\gamma_1 q_1}{q_0 + \gamma_1 q_1} \right) q_0 dR \\
&\geq \int_0^s \left(\frac{\gamma_1 q_1}{q_0 + \gamma_1 q_1} \right) q_0 dR.
\end{aligned}$$

Therefore, when $k = 2$,

$$\begin{aligned}
\Delta(s;\tau) &= \int_0^s \left(\frac{\gamma_1 q_1}{q_0 + \gamma_1 q_1} \right) \frac{dR}{q_0} - \frac{\left(\int_0^s \left(\frac{\gamma_1 q_1}{q_0 + \gamma_1 q_1} \right) dR \right)^2}{\boldsymbol{\Upsilon}(\tau)} \\
&\geq \int_0^s \left(\frac{\gamma_1 q_1}{q_0 + \gamma_1 q_1} \right) \frac{dR}{q_0} - \frac{\left(\int_0^s \left(\frac{\gamma_1 q_1}{q_0 + \gamma_1 q_1} \right) dR \right)^2}{\int_0^s \left(\frac{\gamma_1 q_1}{q_0 + \gamma_1 q_1} \right) q_0 dR}.
\end{aligned}$$

But by Cauchy-Schwartz Inequality, we have the inequality for positive functions f, g and measure ν,

$$\left(\int f d\nu \right)^2 = \left(\int \sqrt{fg}\sqrt{f/g} d\nu \right)^2 \leq \left(\int fg d\nu \right) \left(\int \frac{f}{g} d\nu \right).$$

Applying this result, we have

$$\left(\int_0^s \left(\frac{\gamma_1 q_1}{q_0+\gamma_1 q_1}\right) dR\right)^2 \leq \left(\int_0^s \left(\frac{\gamma_1 q_1}{q_0+\gamma_1 q_1}\right) q_0 dR\right)\left(\int_0^s \left(\frac{\gamma_1 q_1}{q_0+\gamma_1 q_1}\right) \frac{dR}{q_0}\right)$$

so that

$$\frac{\left(\int_0^s \left(\frac{\gamma_1 q_1}{q_0+\gamma_1 q_1}\right) dR\right)^2}{\int_0^s \left(\frac{\gamma_1 q_1}{q_0+\gamma_1 q_1}\right) q_0 dR} \leq \int_0^s \left(\frac{\gamma_1 q_1}{q_0+\gamma_1 q_1}\right) \frac{dR}{q_0},$$

from which it follows that $\Delta(s;\tau) \geq 0$, thereby completing the proof of the theorem. $\|$

For practical purposes, we need consistent estimators of the variance functions of these limiting processes. An obvious estimator of $\boldsymbol{\Upsilon}(\tau, \boldsymbol{\gamma})$ is provided by

$$\hat{\boldsymbol{\Upsilon}}(\tau;\hat{\boldsymbol{\gamma}}) = \frac{1}{n}\sum_{i=1}^{n}\int_0^\tau \left[\boldsymbol{D}(\hat{\boldsymbol{\rho}}(w;\hat{\boldsymbol{\gamma}})) - \hat{\boldsymbol{\rho}}(w;\hat{\boldsymbol{\gamma}})\hat{\boldsymbol{\rho}}(w;\hat{\boldsymbol{\gamma}})'\right]\hat{\boldsymbol{\gamma}}'\boldsymbol{Q}_i(w)d\hat{R}(w).$$

To estimate the limiting covariance matrix for $\hat{\boldsymbol{\gamma}}$, we can use

$$\hat{\boldsymbol{\Sigma}}(\tau,\hat{\boldsymbol{\gamma}}) = \boldsymbol{D}(\hat{\boldsymbol{\gamma}})\hat{\boldsymbol{\Upsilon}}(\tau,\hat{\boldsymbol{\gamma}})^{-1}\boldsymbol{D}(\hat{\boldsymbol{\gamma}}), \tag{15}$$

and an estimator of the limiting variance function of $\sqrt{n}(\hat{R}-R)$ is provided by

$$\hat{\Xi}(s;\hat{\boldsymbol{\gamma}}) = \frac{1}{n}\sum_{i=1}^{n}\int_0^s \left\{\hat{\boldsymbol{\gamma}}'\boldsymbol{Q}_i(w)\right\}^{-1} d\hat{R}(w) + \hat{\boldsymbol{\varrho}}(s;\hat{\boldsymbol{\gamma}})'[\hat{\boldsymbol{\Upsilon}}(\tau;\hat{\boldsymbol{\gamma}})]^{-1}\hat{\boldsymbol{\varrho}}(s;\hat{\boldsymbol{\gamma}}) \tag{16}$$

where $\hat{\boldsymbol{\varrho}}(s;\hat{\boldsymbol{\gamma}}) = \int_0^s \hat{\boldsymbol{\rho}}(w;\hat{\boldsymbol{\gamma}})d\hat{R}(w)$. Finally, an estimator of the limiting variance function of $\sqrt{n}(\hat{\bar{F}}-\bar{F})$ is given by

$$\widehat{\mathbf{Var}}\{Z(s)\} = \hat{\bar{F}}(s)^2\,\hat{\Xi}(s;\hat{\boldsymbol{\gamma}}). \tag{17}$$

The results in Theorem 2 and Corollary 1 are analogous to the asymptotic results in Andersen and Gill (1982) which consider the estimation of the baseline hazard and distribution functions in the multiplicative intensity model. The Andersen and Gill model subsumes the Cox (1972) proportional hazards model. The difference between the load share problem and the regular set up is that the data structure and the stochastic model under load sharing is more complicated; they arise from observing several components in a system combined with the evolution of the failure rates of the components being governed by the component histories.

5 Discussion and Example

To show how prevalent such dependent systems are in many facets of life, consider the following sports example. Table 1 contains data from the National Basketball Association franchise Boston Celtics obtained during the second half of the 2001-2002 season. During this season, the Celtics nucleus of three star players (Antoine Walker, Paul Pierce, Kenny Anderson) compose a kind of system, and a component failure is defined as the event when a player fouled out or was removed from the game due to accumulating a high number of fouls. A player is ejected from the game after committing six fouls, but they are usually taken out (if only for a few minutes) once they reach two fouls, especially if this occurs in the first half of the basketball game. Perhaps the player with two fouls is likely to play conservatively after returning to the game; that is, he makes a concerted effort to not foul while on the court. These three players were chosen because they comprised the chief ”system” of players for the Boston Celtics in 2001-2002, and their foul rates per game were similar over the course of the season. If we measure the amount of game time until each player receives two personal fouls, we found the distributions are not significantly different.

Listed in Table 1 are the game times for each player's 2nd personal foul for the games in the second half of the NBA season in which all three players started the game and committed at least two fouls by the end. Once a player commits two fouls (and is likely benched for a period of time), it might be conjectured that the foul rate of the other star players will change. The foul rate might go down if the team chooses to play more conservatively, trying to decrease the chance another star player gets into foul trouble. On the other hand, if the player with two fouls plays conservatively, the other two star players may shoulder more responsibility on defense, and hence be prone to foul more frequently. In either case, the time-until-second-foul for the three players constitutes a load-share system, with the load share parameters characterizing this change in foul rate due to one or more players being in “foul trouble”.

The estimate of the cumulative hazard function for the time-to-second-foul is illustrated in Figure 1, while Figure 2 presents the corresponding estimate of the survivor function of the time-to-second-foul. In this example, the load-share parameters are of primary interest, and the underlying hazard function represents an infinite set of nuisance parameters. Estimates for the load share parameters are ($\hat{\gamma}_1 = 1.091, \hat{\gamma}_2 = 0.998$). Confidence regions for the estimates are shown in Figure 3 using confidence levels of 0.50, 0.90 and 0.95, which all include the point $(1, 1)$. Thus, for this example, the data set does not provide sufficient evidence that a player's foul rate does not change after one or more of the key Celtic's players get into foul trouble.

Examples in other fields of application can be analyzed and illustrated in the same manner. One fundamental conjecture would be that the system is under constant load, or $H_0 : \gamma_i = k/(k - i), i = 1, ..., k - 1$. In other applications, $1 \leq \gamma_1 \leq \gamma_2 \leq \cdot \leq \gamma_{k-1}$ might be a reasonable assumption. This is the monotone load share rule mentioned in Section 1.

6 Appendix

Proof of Result 1: The score function in (9) can be decomposed into two parts:

$$\begin{aligned} \boldsymbol{U}(s;\gamma) &= \boldsymbol{U}_1(s;\gamma) + \boldsymbol{U}_2(s;\gamma) \\ &= \sum_{i=1}^{n} \int_0^s [\boldsymbol{\delta}_i(w) - \hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})]\, dM_i(w) + \\ &\quad \sum_{i=1}^{n} \int_0^s [\boldsymbol{\delta}_i(w) - \hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})]\, (\boldsymbol{\gamma}'\boldsymbol{Q}_i(w)) dR(w). \end{aligned} \tag{18}$$

It turns out that the second term in (18) is

$$\boldsymbol{U}_2(s;\boldsymbol{\gamma}) = \int_0^s \left(\sum_{i=1}^{n} \boldsymbol{\delta}_i(w)(\boldsymbol{\gamma}'\boldsymbol{Q}_i(w)) - \hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma}) \sum_{i=1}^{n} (\boldsymbol{\gamma}'\boldsymbol{Q}_i(w)) \right) dR(w) = 0.$$

This is true because $Q_{i,j}(w) > 0$ implies $Q_{i,j'}(w) = 0, j \neq j'$, so that

$$\sum_{i=1}^{n} \boldsymbol{\delta}_i(\boldsymbol{\gamma}'\boldsymbol{Q}_i(w)) = (\gamma_0 Q_{i,0}(w), ..., \gamma_{k-1} Q_{i,k-1}(w))' = \boldsymbol{D}(\boldsymbol{\gamma})\boldsymbol{Q}(w).$$

But $\hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})(\boldsymbol{\gamma}'\boldsymbol{Q}_i(w)) = \boldsymbol{D}(\boldsymbol{\gamma})\boldsymbol{Q}(w)$, so that $\boldsymbol{U}_2(s;\boldsymbol{\gamma}) = 0$.

Proof of Lemma 1: By stochastic integration theory, the score process $\{\boldsymbol{U}(s;\boldsymbol{\gamma}) : 0 \le s \le \tau\}$ is clearly a square-integrable martingale with quadratic variation process

$$
\begin{aligned}
&\langle \boldsymbol{U}(\cdot;\boldsymbol{\gamma}), \boldsymbol{U}(\cdot;\boldsymbol{\gamma})\rangle(s) \\
&= \sum_{i=1}^{n} \int_0^s (\boldsymbol{\delta}_i(w) - \hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma}))(\boldsymbol{\delta}_i(w) - \hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma}))'(\boldsymbol{\gamma}'\boldsymbol{Q}_i(w))dR(w) \\
&= \sum_{i=1}^{n} \int_0^s (\boldsymbol{D}(\boldsymbol{\delta}_i(w)) - 2\boldsymbol{\delta}_i\hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})' + \hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})\hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})')(\boldsymbol{\gamma}'\boldsymbol{Q}_i(w))dR(w) \\
&= \sum_{i=1}^{n} \int_0^s [\boldsymbol{D}(\boldsymbol{\gamma} * \boldsymbol{Q}_i) - 2\boldsymbol{D}(\boldsymbol{\gamma})\boldsymbol{Q}_i(w)\hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})' + \\
&\quad \hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})\hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})'(\boldsymbol{\gamma}'Q_i(w))]\, dR(w) \\
&= \int_0^s (\boldsymbol{D}(\boldsymbol{\gamma})\boldsymbol{D}(\boldsymbol{Q}(w)) - 2\boldsymbol{D}(\boldsymbol{\gamma})\boldsymbol{Q}(w)\hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})' + \\
&\quad \hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})\hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})'(\boldsymbol{\gamma}'\boldsymbol{Q}(w))\, dR(w) \\
&= \int_0^s (\boldsymbol{D}(\hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})) - \hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})\hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})')\,(\boldsymbol{\gamma}'\boldsymbol{Q}(w))dR(w).
\end{aligned}
$$

It follows from the Glivenko-Cantelli type strong law of large numbers that if $\{(N_i(w), 0 \le w \le \tau), i = 1, ..., n\}$ are independent and identically distributed, then for $j = 0, ..., k-1$,

$$\sup_{0\le w\le\tau} \left| \frac{1}{n}\sum_{i=1}^{n} Q_{i,j}(w) - q_j(w) \right| \xrightarrow{\text{Pr}} 0,$$

Therefore, provided that $\inf_{0\le w\le\tau} \sum_{j=0}^{k-1}(k-j)\gamma_j P(N_1(w-) = j) > 0$, then

$$\sup_{0\le w\le\tau} |\hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma}) - \boldsymbol{\rho}(w;\boldsymbol{\gamma})| \xrightarrow{\text{Pr}} \mathbf{0},$$

where the jth element of $\boldsymbol{\rho}(w;\boldsymbol{\gamma})$ is $\gamma_j q_j(w)/\sum_{j'=0}^{k-1}\gamma_{j'}q_{j'}(w)$ for $j = 0, ..., k-1$. By Robolledo's martingale central limit theorem (see Andersen, et al. (1993), Theorem II.5.1), it follows that $n^{-1/2}\boldsymbol{U}(\cdot;\boldsymbol{\gamma})$ converges weakly to a zero-mean Gaussian process with covariance matrix function $\boldsymbol{\Upsilon}(s;\boldsymbol{\gamma})$. ‖

Proof of Theorem 1: The establishment of the consistency of $\hat{\boldsymbol{\gamma}}$ follows the usual route of consistency proofs for partial likelihood MLEs. We therefore refer the reader for such standard proofs to Andersen, et. al. (1993). With consistency of $\hat{\boldsymbol{\gamma}}$ established, observe first that for all $\boldsymbol{\eta} > \mathbf{0}$, $\boldsymbol{D}(\boldsymbol{\eta})$ has full rank, and from (9), we have

$$\boldsymbol{U}(s;\boldsymbol{\gamma}) - \boldsymbol{U}(s;\hat{\boldsymbol{\gamma}}) = \sum_{i=1}^{n} \int_0^s (\hat{\boldsymbol{\rho}}(w;\hat{\boldsymbol{\gamma}}) - \hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma}))dN_i(w).$$

Since $\boldsymbol{U}(\tau;\hat{\boldsymbol{\gamma}}) = \mathbf{0}$, we can therefore write

$$\boldsymbol{U}(\tau;\boldsymbol{\gamma}) = \sum_{i=1}^{n} \int_0^\tau (\hat{\boldsymbol{\rho}}(w;\hat{\boldsymbol{\gamma}}) - \hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma}))dN_i(w).$$

A first-order Taylor series expansion of $\hat{\boldsymbol{\rho}}(\cdot;\hat{\boldsymbol{\gamma}})$ about $\boldsymbol{\gamma}$ yields

$$\hat{\boldsymbol{\rho}}(w;\hat{\boldsymbol{\gamma}}) = \hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma}) + \left[\nabla_{\boldsymbol{\gamma}}\hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})|_{\gamma=\xi}\right](\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma})$$

where $\boldsymbol{\xi}$ lies in the line segment connecting $\hat{\boldsymbol{\gamma}}$ and $\boldsymbol{\gamma}$, and the (j,j')th element of $\nabla_{\boldsymbol{\gamma}}\hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})$ is

$$\nabla_{\boldsymbol{\gamma}}\hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})_{(j,j')} = \begin{cases} \hat{\boldsymbol{\rho}}_j(w;\boldsymbol{\gamma})(1-\hat{\boldsymbol{\rho}}_j(w;\boldsymbol{\gamma}))/\gamma_j & \text{for} \quad j = j' \\ -\hat{\boldsymbol{\rho}}_j(w;\boldsymbol{\gamma})\hat{\boldsymbol{\rho}}_{j'}(w;\boldsymbol{\gamma})/\gamma_{j'} & \text{for} \quad j \neq j' \end{cases}.$$

Here $\hat{\boldsymbol{\rho}}_j$ is the j^{th} element of $\hat{\boldsymbol{\rho}}$. This matrix simplifies to

$$\nabla_{\boldsymbol{\gamma}}\hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma}) = \boldsymbol{D}(\hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma}))\boldsymbol{D}(\boldsymbol{\gamma}^{-1}) - \hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})\hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})'\boldsymbol{D}(\boldsymbol{\gamma}^{-1}).$$

Since $\boldsymbol{U}(\tau;\hat{\gamma}) = \mathbf{0}$, then

$$(\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}) = \left(\int_0^\tau \nabla_{\boldsymbol{\gamma}}\hat{\boldsymbol{\rho}}(w;\boldsymbol{\xi})dN(w)\right)^{-1}\boldsymbol{U}(\tau;\boldsymbol{\gamma}).$$

Because $\boldsymbol{\xi} = \boldsymbol{\gamma} + o_p(1)$, and by continuity considerations,

$$\begin{aligned}
&\sqrt{n}(\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}) \\
&= n^{-1/2}\boldsymbol{D}(\boldsymbol{\xi})\left(\frac{1}{n}\int_0^\tau \left(\boldsymbol{D}(\hat{\boldsymbol{\rho}}(w;\boldsymbol{\xi})) - \hat{\boldsymbol{\rho}}(w;\boldsymbol{\xi})\hat{\boldsymbol{\rho}}(w;\boldsymbol{\xi})'\right)dN(w)\right)^{-1}\boldsymbol{U}(\tau;\boldsymbol{\gamma}) \\
&= n^{-1/2}\boldsymbol{D}(\boldsymbol{\gamma})\left(\frac{1}{n}\int_0^\tau \left(\boldsymbol{D}(\hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})) - \hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})\hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})'\right)dN(w)\right)^{-1}\boldsymbol{U}(\tau;\boldsymbol{\gamma}) \\
&\quad + o_p(1).
\end{aligned}$$

The matrix whose inverse is taken converges to $\boldsymbol{\Upsilon}(\tau, \boldsymbol{\gamma})$ because

$$\begin{aligned}
&\frac{1}{n}\int_0^\tau \left(\boldsymbol{D}(\hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})) - \hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})\hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})'\right) dN(w) \\
&= \frac{1}{n}\int_0^\tau \left(\boldsymbol{D}(\hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})) - \hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})\hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})'\right) [dM(w;\boldsymbol{\gamma}) + \boldsymbol{\gamma}'\boldsymbol{Q}(w)dR(w;\boldsymbol{\gamma})] \\
&= O_p(\sqrt{n}) + \frac{1}{n}\int_0^\tau \left(\boldsymbol{D}(\hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})) - \hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})\hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma})'\right) \boldsymbol{\gamma}'\boldsymbol{Q}(w)dR(w;\boldsymbol{\gamma}) \\
&\xrightarrow{\Pr} \int_0^\tau \left(\boldsymbol{D}(\boldsymbol{\rho}(w;\boldsymbol{\gamma})) - \boldsymbol{\rho}(w;\boldsymbol{\gamma})\boldsymbol{\rho}(w;\boldsymbol{\gamma})'\right) \boldsymbol{\gamma}'\boldsymbol{q}(w)dR(w;\boldsymbol{\gamma}) \\
&= \boldsymbol{\Upsilon}(\tau, \boldsymbol{\gamma}).
\end{aligned}$$

This result, along with (1) and (19) establish Theorem 1. ‖

Proof of Theorem 2: Recall that

$$\hat{R}(s) = \int_0^s J(w)(\hat{\boldsymbol{\gamma}}'(\boldsymbol{Q}(w))^{-1} dN(w). \tag{19}$$

We seek a representation of $\hat{R}$ by expanding $(\hat{\boldsymbol{\gamma}}'\boldsymbol{Q}(w))^{-1}$ around $\boldsymbol{\gamma}$ using a first-order Taylor series. First note that

$$\nabla_\gamma(\boldsymbol{\gamma}'\boldsymbol{Q}(w))^{-1} = -\boldsymbol{Q}(w)(\boldsymbol{\gamma}'\boldsymbol{Q}(w))^{-2} = -\boldsymbol{D}(\boldsymbol{\gamma}^{-1})[\boldsymbol{\gamma}'\boldsymbol{Q}(w)]^{-1}\hat{\boldsymbol{\rho}}(w;\boldsymbol{\gamma}).$$

It therefore follows that, for some $\boldsymbol{\xi}$ between $\boldsymbol{\gamma}$ and $\hat{\boldsymbol{\gamma}}$,

$$(\hat{\boldsymbol{\gamma}}'\boldsymbol{Q}(w))^{-1} = (\boldsymbol{\gamma}'\boldsymbol{Q}(w))^{-1} - \left(\boldsymbol{D}(\boldsymbol{\xi}^{-1})\hat{\boldsymbol{\rho}}(w;\boldsymbol{\xi})(\boldsymbol{\xi}'\boldsymbol{Q}(w))^{-1}\right)' (\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}). \tag{20}$$

¿From the proof of Theorem 1, we have the decomposition

$$\begin{aligned}
\sqrt{n}(\hat{R}(s) - R(s)) &= \sqrt{n}\int_0^s (J(w) - 1)dR(w) + \\
&\quad \sqrt{n}\left(\hat{R}(s) - \int_0^s J(w)dR(w)\right),
\end{aligned} \tag{21}$$

where, from (19) and because of (20),

$$\begin{aligned}
\hat{R}(s) &= \int_0^s (\boldsymbol{\gamma}'Q(w))^{-1}J(w)dN(w) - \\
&\quad \left(\int_0^s (\boldsymbol{\xi}'Q(w))^{-1}\hat{\boldsymbol{\rho}}(w;\boldsymbol{\xi})'\boldsymbol{D}(\boldsymbol{\xi}^{-1})J(w)dN(w)\right)(\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}).
\end{aligned}$$

The first term of $\sqrt{n}(\hat{R}(s) - R(s))$ in (21) goes to zero in probability. Using the above representation for $\hat{R}$, the second term in (21) becomes

$$\begin{aligned}
&\sqrt{n}\left(\int_0^s \frac{J(w)dN(w)}{\boldsymbol{\gamma}'Q(w)} - \int_0^s J(w)dR(w)\right) \\
&= \quad -\sqrt{n}\left(\int_0^s \hat{\boldsymbol{\rho}}(w;\boldsymbol{\xi})'\boldsymbol{D}(\boldsymbol{\xi}^{-1})\frac{J(w)dN(w)}{\boldsymbol{\xi}'Q(w)}\right)(\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}).
\end{aligned}$$

To further simplify our notation, let us define

- $\Psi_1(s;\boldsymbol{\eta}) = \int_0^s \hat{\boldsymbol{\rho}}(w;\boldsymbol{\eta})'J(w)(\boldsymbol{\eta}'Q(w))^{-1}dN(w)$,
- $\Psi_2(s;\boldsymbol{\eta}) = \frac{1}{n}\int_0^s(\boldsymbol{D}(\rho(w;\boldsymbol{\eta})) - \rho(w;\boldsymbol{\eta})\rho(w;\boldsymbol{\eta})')dN(w)$,
- $\Psi_3(s,\boldsymbol{\gamma}) = \sqrt{n}\int_0^s J(w)(\boldsymbol{\gamma}'\boldsymbol{Q}(w))^{-1}dM(w)$.

Then, in terms of these processes,

$$\begin{aligned}
&\sqrt{n}(\hat{R}(s) - R(s)) = \\
&\qquad \Psi_3(s;\boldsymbol{\gamma}) - \sqrt{n}\left(\Psi_1(s;\boldsymbol{\xi})\boldsymbol{D}(\boldsymbol{\xi}^{-1})\boldsymbol{D}(\boldsymbol{\gamma})\Psi_2(\tau;\boldsymbol{\gamma})^{-1}\boldsymbol{U}(\tau;\boldsymbol{\gamma})\right) + o_p(1).
\end{aligned}$$

We now describe the limit of $\sqrt{n}(\hat{R}(s) - R(s))$ in terms of the limits of $\Psi_i, i = 1, 2, 3$. We have

$$\frac{1}{n}\Psi_1(s;\boldsymbol{\gamma}) \xrightarrow{\text{Pr}} \int_0^s \boldsymbol{\rho}(w;\boldsymbol{\gamma})'\frac{\boldsymbol{\gamma}'\boldsymbol{q}(w)dR(w)}{\boldsymbol{\gamma}'\boldsymbol{q}(w)} = \int_0^s \boldsymbol{\rho}(w;\boldsymbol{\gamma})'dR(w) = \boldsymbol{\varrho}(s;\boldsymbol{\gamma})'.$$

¿From the proof of Theorem 1, we also have that $\Psi_2(\tau;\boldsymbol{\gamma}) \xrightarrow{\text{Pr}} \boldsymbol{\Upsilon}(\tau,\boldsymbol{\gamma})$, and by Rebolledo's martingale central limit theorem, if $\boldsymbol{\gamma}'\boldsymbol{q}(w) > 0$, then $\Psi_3(s;\boldsymbol{\gamma})$ converges to a zero-mean Gaussian process on $[0,\tau]$ with variance function $s \mapsto \int_0^s\{\boldsymbol{\gamma}'\boldsymbol{q}(w)\}^{-1}dR(w)$. From (18), we have that $\boldsymbol{U}(\tau;\boldsymbol{\gamma}) \xrightarrow{d} N(\boldsymbol{0}, \boldsymbol{\Upsilon}(\tau,\boldsymbol{\gamma}))$.

Because $\boldsymbol{D}(\boldsymbol{\xi}^{-1})\boldsymbol{D}(\boldsymbol{\gamma}) \xrightarrow{\text{Pr}} \boldsymbol{D}(\boldsymbol{1})$ and $\sqrt{n}\int_0^s(J(w) - 1)dR(w)$ is asymptotically negligible, it follows that

$$\sqrt{n}(\hat{R}(s) - R(s)) = \Psi_3(s,\boldsymbol{\gamma}) - \boldsymbol{\varrho}(s;\boldsymbol{\gamma})'\boldsymbol{\Upsilon}(\tau,\boldsymbol{\gamma})^{-1}\left[\frac{1}{\sqrt{n}}\boldsymbol{U}(\tau;\boldsymbol{\gamma})\right] + o_p(1),$$

which converges to a Gaussian process by virtue of the Gaussian process limits of $\Psi_3(\cdot,\boldsymbol{\gamma})$ and $n^{-1/2}\boldsymbol{U}_1(\cdot;\boldsymbol{\gamma})$. The limiting variance function of $\sqrt{n}(\hat{R}(s) - R(s))$

now immediately follows from the above representation by observing that the covariance process between $\Psi_3(\cdot,\boldsymbol{\gamma})$ and $\boldsymbol{U}_1(\cdot,\boldsymbol{\gamma})$ is

$$\begin{aligned}
&\langle \Psi_3(\cdot;\boldsymbol{\gamma}), \boldsymbol{U}_1(\cdot;\boldsymbol{\gamma})\rangle(s) \\
&= \frac{1}{n}\sum_{i=1}^{n}\int_0^s \frac{J(w)\boldsymbol{\gamma}' Q_i(w)}{\frac{1}{n}\boldsymbol{\gamma}' Q(w)}[\boldsymbol{\delta}_i(w) - \boldsymbol{\rho}(w;\boldsymbol{\gamma})]dR(w) \\
&= \int_0^s \frac{J(w)\left(\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{\delta}_i(w)\boldsymbol{\gamma}' Q_i(w)\right)}{\frac{1}{n}\boldsymbol{\gamma}'\boldsymbol{Q}(w)}dR(w) - \int_0^s J(w)\boldsymbol{\rho}(w;\boldsymbol{\gamma})' dR(w).
\end{aligned}$$

Since

$$n^{-1}\sum_{i=1}^{n}\boldsymbol{\delta}_i(w)\boldsymbol{\gamma}'\boldsymbol{Q}_i(w) = n^{-1}\sum_{i=1}^{n}\sum_{j=0}^{k-1}\gamma_j \boldsymbol{Q}_{i,j}(w) = n^{-1}\boldsymbol{D}(\boldsymbol{\gamma})\boldsymbol{Q}(w),$$

then

$$\begin{aligned}
&\langle \Psi_3(\cdot;\boldsymbol{\gamma}), \boldsymbol{U}_1(\cdot;\boldsymbol{\gamma})\rangle(s) \\
&= \int_0^s \frac{J(w)\left(\frac{1}{n}\boldsymbol{D}(\boldsymbol{\gamma})\boldsymbol{Q}(w)\right)}{\frac{1}{n}\boldsymbol{\gamma}'\boldsymbol{Q}(w)}dR(w) - \int_0^s J(w)\boldsymbol{\rho}(w;\boldsymbol{\gamma})' dR(w) \\
&= \int_0^s J(w)\rho(w;\boldsymbol{\gamma})' dR(w) - \int_0^s J(w)\boldsymbol{\rho}(w;\boldsymbol{\gamma})' dR(w) = \boldsymbol{0}.
\end{aligned}$$

This fact completes the proof of Theorem 2. ‖

Proof of Corollary 1: This follows by applying the functional delta-method and invoking the asymptotic result in Theorem 2, since

$$\hat{\bar{F}}(\cdot) = \phi(\hat{R})(\cdot) \equiv \prod_{0\le w\le \cdot}\left[1 - d\hat{R}(w)\right].$$

Thus,

$$\sqrt{n}\left[\hat{\bar{F}}(\cdot) - \bar{F}(\cdot)\right] = \sqrt{n}\left[\phi(\hat{R})(\cdot) - \phi(R)(\cdot)\right].$$

By the functional delta-method (cf., Andersen, et. al. (1993)), it follows that the limiting process is $d\phi(R)\cdot W$ where, with W being the Gaussian limiting process in Theorem 2,

$$d\phi(R)\cdot W(s) = \int_{w\in[0,s]}\left\{\prod_{[0,w)}(1-dR)\right\}W(dw)\left\{\prod_{(w,s]}(1-dR)\right\} = \bar{F}(s)W(s).$$

‖

References

[1] Aalen, O. (1978), "Nonparametric inference for a family of counting processes," *Annals of Statistics*, 6, 701–726.

[2] Andersen, P. and Gill, R. (1982), "Cox's regression model for counting processes: A large sample study," *Annals of Statistics*, 10, 1100–1120.

[3] Andersen, P. K., Borgan, Ø., Gill, R. D. and Keiding, N. (1993), *Statistical Models Based on Counting Processes*, Springer-Verlag, New York.

[4] Carlson, R. L. and Kardomateas, G. A. (1996), *An Introduction to Fatigue in Metals and Composites*, Chapman & Hall, London.

[5] Cox, D. R. (1972), "Regression models and life tables (with discussion)," *Journal of the Royal Statistical Society, B*, 34, 187–220.

[6] Daniels, H. E. (1945), "The statistical theory of the strength bundles of threads," *Proceedings of the Royal Society*, London, A, 83, 405–435.

[7] Durham, S. D. and Lynch, J. D. (2000), "A threshold representation for the strength distribution of a complex load sharing system," *Journal of Statistical Planning and Inference*, Vol. 83, 25–46.

[8] Freund, J. E. (1961), "A bivariate extension of the exponential distribution," *Journal of the American Statistical Association*, Vol. 56, 971–977.

[9] Hollander, M. and Peña, E. (1995), "Dynamic Reliability Models With Conditional Proportional Hazards," *Lifetime Data Analysis*, 1, 377–401.

[10] Jelinski, Z. and Moranda, P. (1972), "Software reliability research," *Statist. Comp. Perform. Evaluation*, 45–484.

[11] Jewell, N. P. and Kalbfleisch, J. D. (1996), "Marker processes in survival analysis," *Lifetime Data Analysis*, Vol. 2, 15–19.

[12] Kiefer, J. and Wolfowitz, J. (1956), "Consistency of the maximum likelihood estimator in the presence of infinitely many incidental parameters," *Annals of Mathematical Statistics*, 27, 887–906.

[13] Kvam, P. H. and Samaniego, F. J. (1993) "Life testing in variably scaled environments," *Technometrics*, Vol. 35, No. 3, 306–314.

[14] Kvam, P. H. and Day, D. (2001) "The multivariate Polya distribution in combat modeling," *Naval Research Logistics*, Vol. 48, No. 1, 1–17.

[15] Lynch, J. D. (1999) "On the joint distribution of component failures for monotone load sharing systems," *Journal of Statistical Planning and Inference*, Vol. 78, 13–21.

[16] Mäkeläinen, T., Schmidt, K. and Styan, G. P. H. (1981) "On the existence and uniqueness of the maximum likelihood estimate of a vector-valued parameter in fixed-size samples," *Annals of Statistics*, Vol. 9, No. 4, 758–767.

[17] Marshall, A. W. and Olkin, I. (1967), "A multivariate exponential distribution," *Journal of the American Statistical Association*, 62, 30–44.

[18] Ortega, J. M. and Rheinboldt, C. R. (1970) *Iterative solution of Nonlinear Equations in Several Variables*, New York: Academic Press.

[19] Rydeń, P. (1999) *Estimation of the Reliability of Systems Described by the Daniels Load-Sharing Model*, Licentiate Thesis, Department of Mathematical Statistics, Umeå University, Sweden.

Table 1: Time until second personal foul in 28 games of Boston Celtics 2001-2002 season.

Date	Pierce	Walker	Anderson	Date	Pierce	Walker	Anderson
1/5	21.02	30.22	43.43	2/27	42.06	23.21	45.36
1/7	24.25	45.54	17.19	3/4	28.51	33.59	16.20
1/9	6.55	19.47	23.28	3/8	34.56	32.53	40.44
1/11	15.35	16.37	25.40	3/10	40.33	15.35	28.33
1/19	39.08	30.32	43.53	3/11	27.56	46.21	28.05
1/23	16.2	4.16	39.52	3/13	9.54	36.21	28.12
1/25	34.59	46.44	16.33	3/16	27.09	11.11	23.33
1/26	19.10	38.40	20.17	3/18	40.36	33.21	17.04
1/29	28.22	37.43	25.41	3/24	41.44	36.28	19.13
1/31	32.00	45.52	39.11	4/5	32.23	8.17	41.27
2/3	11.25	19.09	11.59	4/7	7.53	37.31	13.43
2/12	17.39	25.43	22.51	4/8	28.34	35.58	41.48
2/13	28.47	31.15	2.41	4/10	26.32	28.02	29.33
2/19	23.42	31.28	40.03	4/15	30.47	40.40	42.13

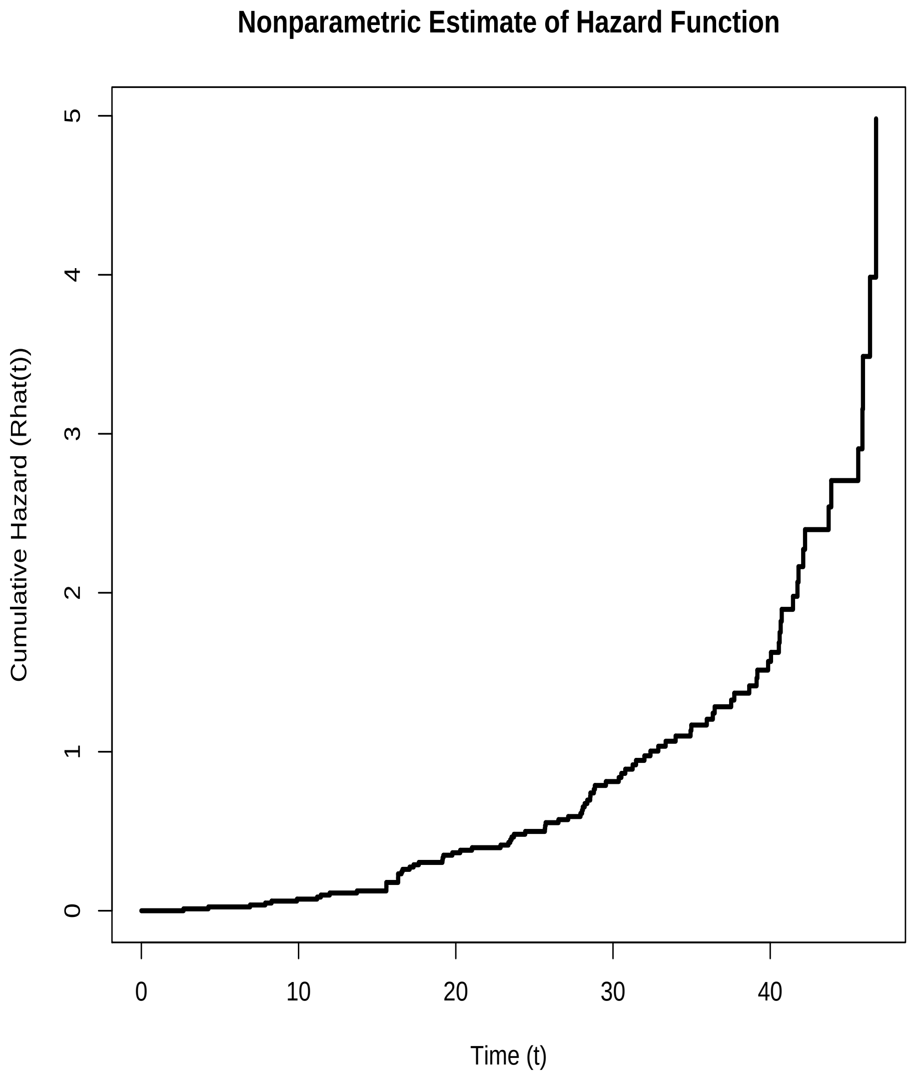

Figure 1: Estimated Cumulative Hazard Function of Minutes Played Until Committing the Second Foul.

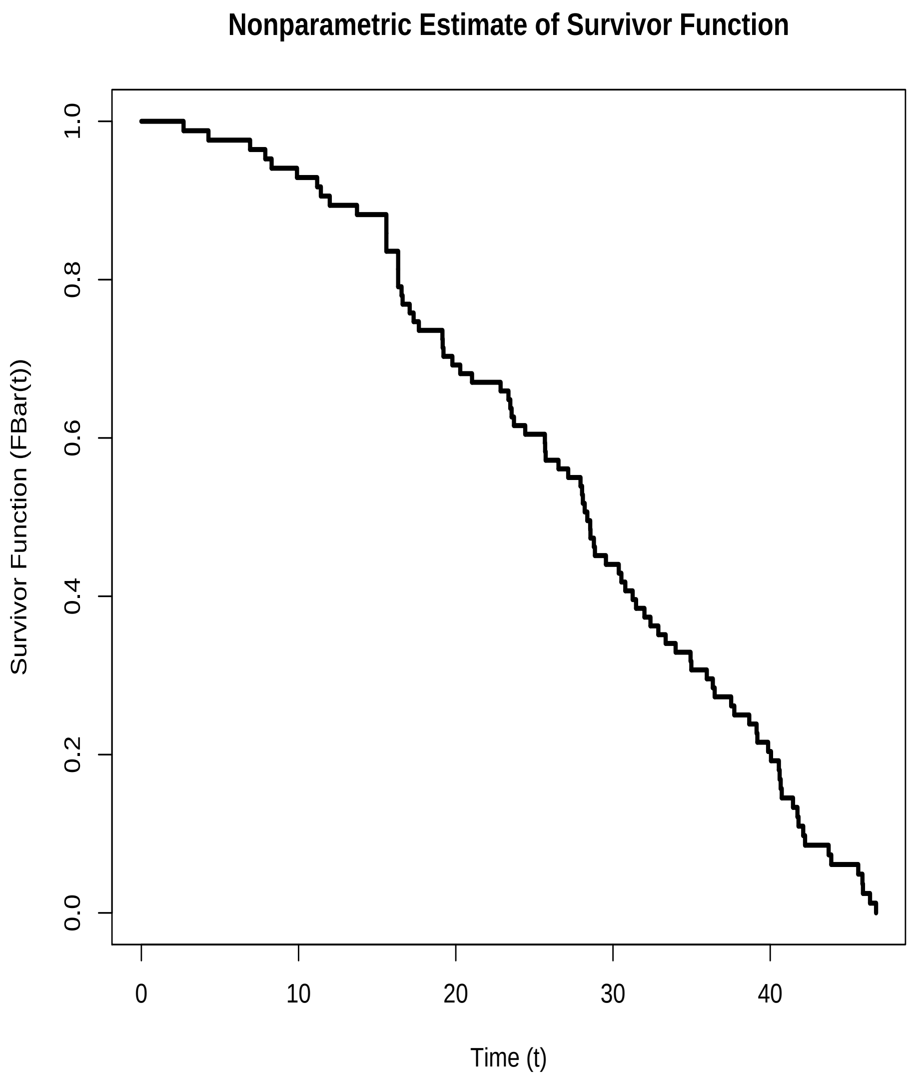

Figure 2: Estimated Survivor Function of Minutes Played Until Committing the Second Foul.

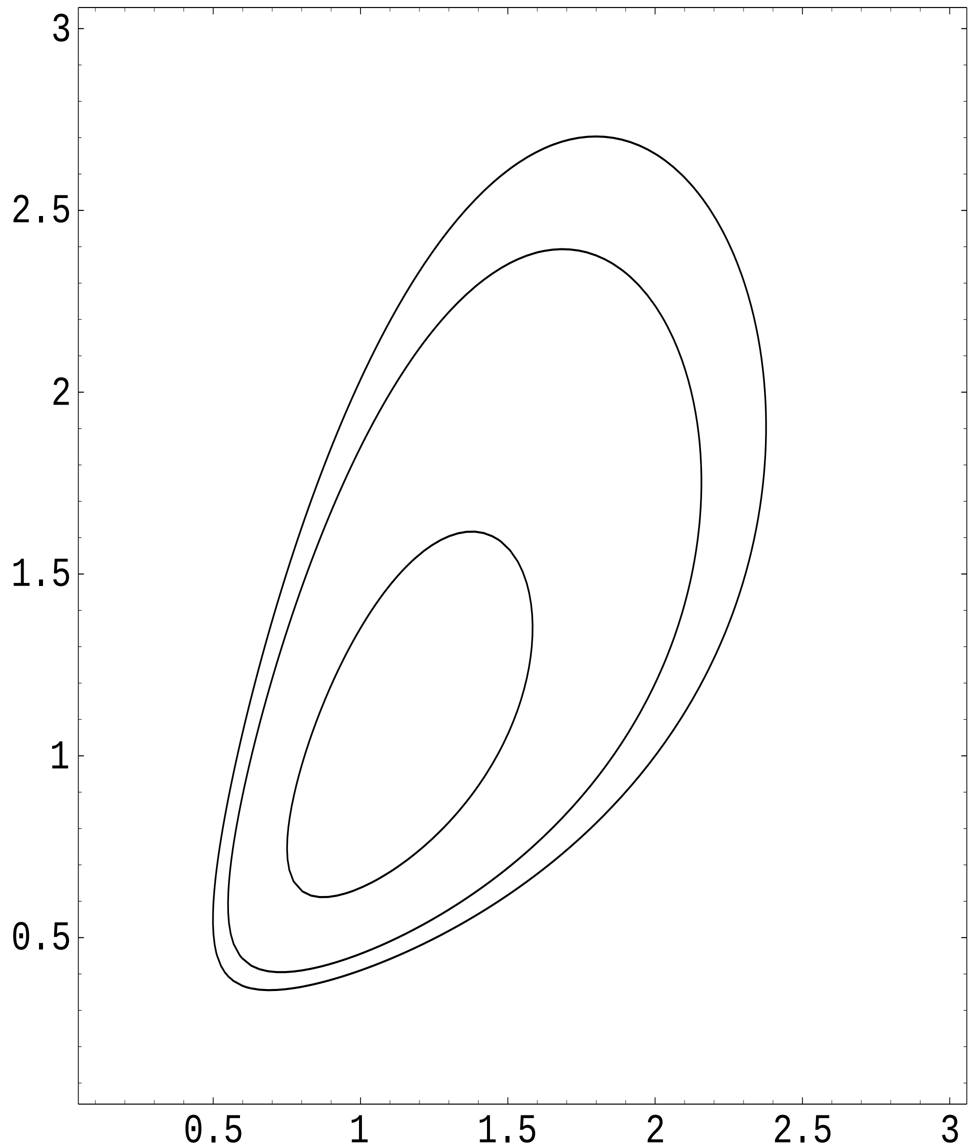

Figure 3: Confidence Regions (50%, 90%, 95%) for (γ_1, γ_2)