Rostock. Math. Kolloq. **47**, 49–62 (1994)

Subject Classification (AMS)
62F12, 62N05

Waltraut Kahle; Friedrich Liese

Consistency and Asymptotic Normality of Minimum Contrast Estimation in Renewal Processes

ABSTRACT. The consistency and asymptotic normality of minimum contrast estimation (which includes the maximum likelihood estimation as a special case) is established if the sample is from a renewal process and the observation time tends to infinity. It is shown, that the conditions for consistency and asymptotic normality for maximum likelihood estimation are fulfilled if the distribution of the time between two renewals is Weibull or Maxwell. This fact is used to construct simultaneous confidence regions for the parameters of the Weibull distribution and a confidence intervall for the parameter of the Maxwell distribution.

1 Introduction

In reliability theory and statistical quality control very often renewal processes are observed. If the distribution of the lifetime is exponential it is possible to calculate the distribution of the point estimation of the unknown parameter and to derive confidence estimations from this distribution (Mann, Schafer, Singpurwalla [4]). In other cases asymptotic methods have to be used to calculate confidence estimations of the unknown parameters. There are two kinds of asymptotics:

1. The number m of realizations of renewal processes tends to infinity and each realization is observed a fixed time $T_i,\ i = 1, \ldots, m$.
2. The observation time of at least one realization of the renewal process tends to infinity and the number of realizations is fixed.

In the first case the consistency and asymptotic normality of maximum likelihood estimations can be proved, using theorems for independent, not necessary identically distributed observations (see a.e. Ibragimov, Hasminski [2]). In the second case we have neither independent nor identically distributed observations and it is impossible to prove the consistency and asymptotic normality of maximum likelihood estimations by classical methods. In this paper conditions which guarantee consistency and asymptotic normality of minimum contrast

estimations are established if the observation time of renewal processes tends to infinity.
It turns out that the asymptotic behaviour of minimum contrast estimators for renewal processes is closely related to the behaviour of the same estimators in the i.i.d. case. The difference is that in our case the number of observations is random. We do not follow PFANZAGL [5] in section 2 where the consistency is established. We require continuity instead of semicontinuity of the contrast function and prove weak consistency. This leads to an essential simplification of the proof. Section 3 contains the proofs of the theorems of section 2. In section 4 the conditions of consistency and asymptotic normality are established for the special case of maximum likelihood estimation if the underlying distribution is Weibull. It is demonstrated, how the asymptotic normality can be used to construct simultaneous confidence regions for the parameters of the Weibull distribution. For the parameter of the Maxwell distribution maximum likelihood estimation and another minimum contrast estimation are compared.

2 Models and results

Suppose $X_1, X_2, \ldots$ are i.i.d. nonnegative random variables with distribution function F_θ, where the parameter θ belongs to the parameter set Θ being a subset of R^p. Introduce the renewal process N_t by

$$N_t = \max\left\{k : \sum_{i=1}^{k} X_i \leq t\right\}.$$

If $F_\theta(0) < 1$ then $N_t < \infty$. The rest time R_t is defined by

$$R_t = t - \sum_{i=1}^{N_t} X_i.$$

The problem now is to estimate the parameter θ using the sample $X_1, \ldots, X_{N_t}$, R_t.
Often, not only one renewal process but m independent processes are observed with different observation times. More precisely let $X_{1,j}, X_{2,j}, \ldots,\ j = 1, \ldots, m$ i.i.d. random variables with distribution function F_θ. Let $T_1, \ldots, T_m$ be the observation times and denote by

$$R_{T_j} = T_j - \sum_{i=1}^{N_{T_j}} X_{i,j}$$

the corresponding rest times. Our available information consists in the sample

$$X_{1,j}, \ldots, X_{N_{T_j},j},\ R_{T_j}, \quad j = 1, \ldots, m.$$

To estimate the parameter θ we use the concept of minimum contrast estimations intensively studied in a series of papers by PFANZAGL, a.e. [5] not for renewal processes but for i.i.d.

samples.

Let $(P_\theta)_{\theta\in\Theta}$ be a family of distributions on $(R, \mathcal{B})$ where R is the real line and $\mathcal{B}$ the σ-algebra of Borel sets. Let $\varrho : R\times\Theta \longrightarrow [-\infty, \infty]$ be a function so that $\varrho(\cdot, \theta)$ is measurable for every $\theta \in \Theta$. ϱ is called a contrast function if for every $\theta_1, \theta_2 \in \Theta$ the integral $\int \varrho(x, \theta_1)P_{\theta_2}(dx)$ exists and

$$\int \varrho(x, \theta_1)P_{\theta_2}(dx) > \int \varrho(x, \theta_2)P_{\theta_2}(dx) \quad \text{if } \theta_1 \neq \theta_2. \tag{1}$$

Example 1: Suppose P_θ is dominated by a σ-finite measure λ. Denote by $f_\theta = \frac{dP_\theta}{d\lambda}$ the Radon-Nikodym derivative and put

$$\varrho(x, \theta) = -\log f_\theta(x).$$

Then ϱ is a contrast function provided that $\int(\log f_{\theta_2})dP_{\theta_2} > -\infty$ and $\int(\log f_{\theta_1})dP_{\theta_2} < \infty$ for every $\theta_1, \theta_2 \in \Theta$ (see PFANZAGL [5]).

Example 2: Suppose $\Theta = (0, \infty)$, μ is the Lebesgue measure and $f(x)$ is a density with respect to μ. Assume $f(x) = 0$ for $x < 0$ and

$$\int_0^\infty x f(x)dx = 1.$$

Introduce a family of densities by $f_\theta(x) = \theta f(\theta x)$. Then obviously $\frac{1}{\theta}$ is the expectation of f_θ. Set

$$\varrho(x, \theta) = -\log\theta + x\theta.$$

Clearly $-\varrho$ is the loglikelihood of the exponential distribution. But we will show that even ϱ is a contrast function for the distribution P_θ with density f_θ. This contrast function will be applied to construct consistent and asymptotically normaly distributed estimators.

$$\begin{aligned}
&\int (\varrho(x, \theta_1) - \varrho(x, \theta_2))\, f_{\theta_2}(x)dx \\
&= \int \left(-\log\frac{\theta_1}{\theta_2} + (\theta_1 - \theta_2)x\right) \theta_2 f(\theta_2 x)dx \\
&= -\log\frac{\theta_1}{\theta_2} + \frac{\theta_1}{\theta_2} - 1 \geq 0.
\end{aligned}$$

As $-\log x - 1 + x \geq 0$ with equality iff $x = 1$ we get that ϱ is a contrast function.

To apply the concept of contrast function to renewal processes we introduce

$$Q(\tau) = \sum_{j=1}^{m}\sum_{i=1}^{N_{T_j}} \varrho(X_{ij}, \tau) + \sum_{j=1}^{m} \tilde{\varrho}(R_{T_j}, \tau) \tag{2}$$

where ϱ is a contrast function and $\tilde{\varrho}(\cdot,\tau)$ is any measurable function which includes the information contained in the rest time.
$X_{ij},\ j = 1,\ldots,m,\ i = 1,2,\ldots$ are i.i.d. with common distribution P_θ. Now, roughly speaking, the minimum contrast estimation is defined by minimization of $Q(\tau)$ over $\tau \in \Theta$. To be more precise we denote by $\Omega = R \times R \times \ldots$ the set of all sequences of real numbers and by $\mathcal{F}$ Kolmogorov's σ-algebra. Write $(\Omega,\mathcal{F})^m$ for the m-times product space. Suppose $\Theta \subseteq R^p$ is a Borel set and denote by $\mathcal{B}_\Theta$ the σ-algebra of Borel subsets of Θ. By $\mathcal{E}$ we shall mean the class of all measurable mappings

$$S : (\Omega,\mathcal{F})^m \longrightarrow (\Theta,\mathcal{B}_\Theta)$$

which depends only on the observations up to the times $T_1,\ldots,T_j$, i.e. which may be represented as

$$S = \sum_{k_1,\ldots,k_m=0}^{\infty} \varphi_{k_1,\ldots,k_m}(X_{1,1},\ldots,X_{k_m,m},R_{T_1},\ldots,R_{T_m}) I_{A_{k_1,\ldots,k_m}}$$

where

$$\varphi_{k_1,\ldots,k_m} :\ R^{N+m} \longrightarrow \Theta\ ,\ N = \sum_{i=1}^{m} k_i$$

are measurable functions and $I_{A_{k_1,\ldots,k_m}}$ is the indicator function of the event

$$A_{k_1,\ldots,k_m} = \{N_{T_1} = k_1,\ldots,N_{T_m} = k_m\}.$$

Now a minimum contrast estimator $\hat{\theta}$ is an estimator $\hat{\theta}$ which belongs to $\mathcal{E}$ and minimizes $Q(\theta)$ over Θ.
Suppose $P(X_{i,j} > 0) > 0$. Then $N_{T_i} < \infty$ a.s. Assume further that Θ is compact and both $\tau \to \varrho(x,\tau)$ and $\tau \to \tilde{\varrho}(x,\tau)$ are lower semicontinuous for every $x \in R$. In this situation one can easily shown by similar methods as in PFANZAGL [5] that a measurable minimum contrast estimation exists.
Assume the distributions P_θ of $X_{i,j}$, $\theta \in \Theta$, are measures theoretical equivalent and put for any fixed $\theta_0 \in \Theta$ $f_\theta = \frac{dP_\theta}{dP_{\theta_0}}$. If F_θ (distribution function of $X_{i,j}$) is continuous, then the loglikelihood of the j-th renewal process in $[0,T_j]$ is given by [3]:

$$\log \frac{dP_\theta^{T_j}}{dP_{\theta_0}^{T_j}} = \sum_{i=1}^{N_{T_j}} \log f_\theta(x_{i,j}) + \log \frac{1 - F_\theta(R_{T_j})}{1 - F_{\theta_0}(R_{T_j})} \tag{3}$$

where $P_\theta^{T_j}$ and $P_{\theta_0}^{T_j}$ are the distributions of the renewal processes on $[0,T_j]$ with parameter θ and θ_0, respectively. Comparing (2) and (3) it is easy to see that the maximum likelihood estimation is a special minimum contrast estimation.
Assume $T_{j,1} \leq T_{j,2} \leq \ldots$ for $j = 1,\ldots,m$ and put $T_n = \sum_{j=1}^m T_{j,n}$.

Theorem 1 *Suppose Θ is compact and both $\varrho(x,\theta)$ and $\tilde{\varrho}(x,\theta)$ are continuous functions on $[0,\infty)\times\Theta$ with values in $(-\infty,+\infty)$. Assume $(P_\theta)_{\theta\in\Theta}$ is the distribution of the nonnegative i.i.d. $X_{i,j}$ and ϱ is a contrast function with respect to $(P_\theta)_{\theta\in\Theta}$. Suppose*

$$\int \sup_{\theta\in\Theta} |\varrho(x,\theta)| P_\theta(dx) < \infty.$$

If $0<\mu=EX_{i,j}<\infty$ and $T_n=\sum_{j=1}^m T_{j,n}\to\infty$ for $n\to\infty$ then every minimum contrast estimation $\hat{\theta}_n$ which minimizes (2) is weakly consistent, i.e. for every $\varepsilon>0$

$$\lim_{n\to\infty} P(|\hat{\theta}_n-\theta|>\varepsilon)=0.$$

We now establish conditions for asymptotic normality of minimum contrast estimates. Assume that $\lim_{n\to\infty}\frac{T_{j,n}}{T_n}=\alpha_j$ exists.

Theorem 2 *Assume the conditions of Theorem 1 are fulfilled. Suppose in addition that the true parameter θ belongs to the interior Θ^0 of Θ and $\tilde{\varrho}(x,\theta)$, $\varrho(x,\theta)$ are twice continuously differentiable with respect to each θ_j, $j=1,\ldots,p$, $\theta=(\theta_1,\ldots,\theta_p)$ in Θ^0 and there exists a neighborhood $U_\theta\subseteq\Theta^0$ of θ so that*

$$\int \sup_{\tau\in U_\theta} M(\tau) P_\theta(dx) < \infty \tag{4}$$

$$\textit{with} \qquad M(\tau)=\sum_{j,k=1}^{p}\left|\frac{\partial^2}{\partial\theta_j\partial\theta_k}\varrho(x,\tau)\right|$$

and

$$\int \left(\frac{\partial}{\partial\theta_j}\varrho(x,\tau)\right)^2 P_\theta(dx) < \infty. \tag{5}$$

If $\det\Sigma(\theta)\neq 0$ with

$$\Sigma(\theta)=((\Sigma_{jk})),\ j,k=1,\ldots,p, \qquad \Sigma_{jk}=\int \frac{\partial^2}{\partial\theta_j\partial\theta_k}\varrho(x,\theta)P_\theta(dx)$$

and $E\frac{\partial}{\partial\theta_k}\varrho(X_{11},\theta)=0$, $k=1,\ldots,p$, then the distribution of

$$\sqrt{T_n}\left(\hat{\theta}_n-\theta\right)$$

converges weakly to a p-dimensional normal distribution with expectation zero and covariance matrix $C(\theta)=\frac{1}{\mu}\Sigma^{-1}(\theta)\tilde{\Sigma}(\theta)\Sigma^{-1}(\theta)$ where

$$\tilde{\Sigma}(\theta)=\left(\int\frac{\partial}{\partial\theta_i}\varrho(x,\theta)\frac{\partial}{\partial\theta_j}\varrho(x,\theta)P_\theta(dx)\right)_{1\le i,j\le p}.$$

We remark that in the case of $\varrho=-\log f_\theta$ it holds $\Sigma=\tilde{\Sigma}$ and $C=\frac{1}{\mu}\Sigma^{-1}=\frac{1}{\mu}I^{-1}$ with the Fisher information matrix I.

3 Proofs

Lemma 3 *Suppose $\Theta \subseteq R^p$ is compact and $\tau \to h(X_1, \tau)$ is a.s. continuous. Assume $X_1, X_2, \ldots$ are i.i.d. and*

$$E \sup_{\tau \in \Theta} |h(X_1, \tau)| < \infty.$$

Then

$$\lim_{n \to \infty} \sup_{\tau \in \Theta} \left| \frac{1}{n} \sum_{i=1}^{n} h(X_i, \tau) - Eh(X_1, \tau) \right| = 0 \quad a.s.$$

Lemma 3 is a consequence of the strong law of large numbers for random variables taking values in a separable Banach space (see P. Revesz [6]).
Set $N_T = \max\{k : \sum_{i=1}^{k} X_i < T\}$.

Lemma 4 *Assume the assumptions of Lemma 3 are fulfilled. If $0 < \mu = EX_1 < \infty$ then for every sequence $T_n \to \infty$*

$$\lim_{T_n \to \infty} \sup_{\tau \in \Theta} \left| \frac{1}{T_n} \sum_{i=1}^{N_{T_n}} h(X_i, \tau) - \frac{1}{\mu} Eh(X_1, \tau) \right| = 0 \quad a.s.$$

Proof: Lemma 3 implies

$$\lim_{n \to \infty} \sup_{\tau \in \Theta} \left| \frac{1}{N_{T_n}} \sum_{i=1}^{N_{T_n}} h(X_i, \tau) - Eh(X_1, \tau) \right| = 0.$$

The rest follows from

$$\frac{1}{T_n} N_{T_n} \xrightarrow[n \to \infty]{} \frac{1}{\mu} \qquad \text{a.s.}$$

Lemma 5 *Let $\{X_{ik},\ k = 1, \ldots, p,\ i \geq 1\}$ be a sequence of i.i.d. random vectors with mean 0 and regular covariance matrix Σ. Further, assume that*

$$\frac{N_{T_n}}{T_n} \xrightarrow{P} \frac{1}{\mu} \quad (0 < \mu < \infty) \quad \text{as} \quad T_n \to \infty.$$

Then the distribution of the vectors

$$\left(\frac{1}{\sqrt{N_{T_n}}} \sum_{i=1}^{N_{T_n}} X_{i,k} \right)_{k=1,\ldots,p} \qquad \text{and} \qquad \left(\frac{1}{\sqrt{\mu T_n}} \sum_{i=1}^{N_{T_n}} X_{i,k} \right)_{k=1,\ldots,p}$$

converges weakly to a normal distribution with mean zero and covariance matrix Σ as $T_n \to \infty$.

This lemma may be obtained from ANSCOMBE's theorem using the CRAMER – WALD device (see a.e. GUT [1] and SERVLING [7]).

Proof of Theorem 1: Put

$$Q_{j,n}(\tau) = \frac{1}{T_{j,n}} \left(\sum_{i=1}^{N_{T_{j,n}}} \varrho(X_{i,j}, \tau) + \tilde{\varrho}(R_{T_{j,n}}, \tau) \right)$$

and

$$M_{T_{j,n}} = \inf \left\{ k : \sum_{j=1}^{k} X_{i,j} > T_{j,n} \right\}.$$

$M_{T_{j,n}}$ is a stopping time with respect to the filtration $\mathcal{F}_{j,n} = \sigma(X_{j,1}, X_{j,2}, \ldots)$. Hence by WALD's identity

$$\begin{aligned} E\left[\sum_{i=1}^{N_{T_{j,n}}} \sup_{\tau} |\varrho(X_{i,j}, \tau)|\right] &\leq E\left[\sum_{i=1}^{M_{T_{j,n}}} \sup_{\tau} |\varrho(X_{i,j}, \tau)|\right] \\ &\leq \left(EM_{T_{j,n}}\right) E \sup_{\tau} |\varrho(X_{1,1}, \tau)| \\ &\leq \left(EN_{T_{j,n}} + 1\right) E \sup_{\tau} |\varrho(X_{1,1}, \tau)|. \end{aligned} \tag{6}$$

Put

$$\frac{1}{T_n} Q_n(\tau) = \sum_{j=1}^{m} \frac{T_{j,n}}{T_n} Q_{j,n}(\tau).$$

and

$$\varphi(x) = \sup_{\tau} |\tilde{\varrho}(x, \tau)|.$$

By the continuity of $\tilde{\varrho}$ we get $\varphi(x) < \infty$ for every x. As

$$P\left(R_{T_{j,n}} > a\right) \leq P\left(X_{1,1} > a\right)$$

we see that the sequence $\varphi(R_{T_{j,n}})$ is stochastically bounded. Hence in probability

$$\frac{1}{T_n} \sum_{j=1}^{m} \varphi(R_{T_{j,n}}) \xrightarrow[n\to\infty]{} 0. \tag{7}$$

If $\lim_{n\to\infty} \frac{T_{j,n}}{T_n} = 0$ then by (6)

$$\lim_{n\to\infty} \sup_{\tau} |Q_{j,n}(\tau)| = 0$$

in probability. If $\lim_{n\to\infty} \frac{T_{j,n}}{T_n} > 0$ then by Lemma 4 and (7)

$$\lim_{n\to\infty} P\left(\sup_{\tau} \left| Q_{j,n}(\tau) - \frac{1}{\mu} E\varrho(X_{1,1}, \tau) \right| > \varepsilon\right) = 0.$$

This means

$$\lim_{n\to\infty} P\left(\sup_{\tau}\left|\frac{1}{T_n}Q_n(\tau) - \frac{1}{\mu}E\varrho(X_{1,1},\tau)\right| > \varepsilon\right) = 0. \tag{8}$$

As $\varrho(X_{1,1},\tau)$ depends continuously on τ and $E\sup_\tau|\varrho(X_{1,1},\tau)| < \infty$ the dominated convergence theorem implies that $E\varrho(X_{1,1},\tau)$ is a continuous function of τ. As ϱ is a contrast function and Θ is compact we get for every $\gamma > 0$

$$\alpha = \inf_{\tau:\,\|\tau-\theta\|>\gamma} \frac{1}{\mu}E\varrho(X_{1,1},\tau) - \frac{1}{\mu}E\varrho(X_{1,1},\theta) > 0.$$

Put for $\beta > 0$

$$A_\beta = \left\{\sup_{\tau}\left|\frac{1}{T_n}Q_n(\tau) - \frac{1}{\mu}E\varrho(X_{1,1},\tau)\right| \le \beta\right\}.$$

For τ with $\|\tau-\theta\| > \gamma$ it holds on the event A_β

$$\begin{aligned}\frac{1}{T_n}Q_n(\tau) &\ge \frac{1}{\mu}E\varrho(X_{1,1},\tau) - \beta\\ &\ge \frac{1}{\mu}E\varrho(X_{1,1},\theta) + \alpha - \beta\\ &\ge \frac{1}{T_n}Q_n(\theta) + \alpha - 2\beta.\end{aligned}$$

Hence

$$\inf_{\|\tau-\theta\|>\gamma}\frac{1}{T_n}Q_n(\tau) \ge \frac{1}{T_n}Q_n(\theta) + \alpha - 2\beta > \frac{1}{T_n}Q_n(\theta) + \frac{2\alpha}{3} - 2\beta$$

and

$$\{\|\hat\theta_n - \theta\| > \gamma\} \subseteq \left\{\inf_{\tau:\,\|\tau-\theta\|>\gamma}\frac{1}{T_n}Q_n(\tau) > \frac{1}{T_n}Q_n(\theta)\right\} \subseteq A_{\alpha/3}$$

the assertion now follows from (8).

Proof of Theorem 2: Introduce the event A_n by

$$A_n = \left\{\hat\theta_n \in \tilde U_\theta\right\}$$

where $\tilde U_\theta \subseteq U_\theta$ is a closed sphere with centre θ and positive radius. Then by the consistency of $\hat\theta_n$

$$P(A_n) \xrightarrow[n\to\infty]{} 1.$$

On the event A_n the function $Q_n(\theta)$ attains the minimum at $\hat\theta_n$ which is a local minimum. Hence on A_n

$$0 = \sum_{j=1}^{m}\left(\sum_{i=1}^{N_{T_{j,n}}}\varrho_k(X_{i,j},\hat\theta_n) + \tilde\varrho_k(R_{T_{j,n}},\hat\theta_n)\right), \quad k = 1,\ldots,p \tag{9}$$

where $\varrho_k(x,\theta) = \frac{\partial}{\partial\theta_k}\varrho(x,\theta)$. Similarly, $\varrho_{i,j}$ denotes the derivative with respect to θ_i and θ_j. Taylor expansion and (9) yield

$$\begin{aligned} o_p(1) &= \sum_{j=1}^{m}\left(\sum_{i=1}^{N_{T_{j,n}}} \varrho_k(X_{i,j},\theta) + \tilde{\varrho}_k(R_{T_{j,n}},\theta)\right) \\ &+ \sum_{j=1}^{m}\sum_{l=1}^{p}(\hat{\theta}_{n,l}-\theta_l)\left[\sum_{i=1}^{N_{T_{j,n}}} \varrho_{k,l}(X_{i,j},\bar{\theta}_n) + \tilde{\varrho}_{k,l}(R_{T_{j,n}},\bar{\theta}_n)\right] \end{aligned}$$

where $\|\bar{\theta}_n - \theta\| < \|\hat{\theta} - \theta\|$. Using the same arguments as in the proof of Theorem 1 we get

$$\frac{1}{\sqrt{T_n}}\sum_{j=1}^{m}\sup_{\tau\in U_\theta}\left|\tilde{\varrho}_k(R_{T_{j,n}},\tau)\right|$$

converges stochastically to zero as $n \to \infty$. Hence

$$\begin{aligned} o_p(1) &= (\sum_{j=1}^{m}\left(\frac{T_{j,n}}{T_n}\right)^{1/2}\cdot\frac{1}{\sqrt{T_{j,n}}}\sum_{i=1}^{N_{T_{j,n}}} \varrho_k(X_{i,j},\theta) \\ &+ \sum_{l=1}^{p}\sqrt{T_n}(\hat{\theta}_{n,l}-\theta_l)\left[\frac{1}{T_n}\sum_{j=1}^{m}\sum_{i=1}^{N_{T_{j,n}}} \varrho_{k,l}(X_{i,j},\bar{\theta}_n)\right]. \end{aligned} \tag{10}$$

In view of Lemma 5 and the assumption $E\varrho_k(X_{1,1},\theta) = 0$ the distribution of the vectors

$$\left(\frac{1}{\sqrt{\mu T_{j,n}}}\sum_{i=1}^{N_{T_{j,n}}} \varrho_k(X_{i,j},\theta)\right)_{k=1,\dots,p}$$

converges weakly to a normal distribution with mean zero and covariance matrix

$$(E\varrho_k(X_{1,1},\theta)\varrho_l(X_{1,1},\theta))_{1\le k,l\le p} = \tilde{\Sigma}(\theta).$$

By the independence of the m renewal processes we see that the distribution of

$$\sum_{j=1}^{m}\left(\frac{T_{j,n}}{T_n}\right)^{1/2}\cdot\frac{1}{\sqrt{T_{j,n}}}\sum_{i=1}^{N_{T_{j,n}}} \varrho_k(X_{i,j},\theta)$$

weakly converges to a normal distribution with mean zero and covariance matrix $\mu\tilde{\Sigma}(\theta)$. The consistency of $\hat{\theta}_n$ implies the stochastic convergence of $\bar{\theta}_n$ to θ. If we applay the law of large numbers in Lemma 4

$$\frac{1}{T_{j,n}}\sum_{i=1}^{N_{T_{j,n}}} \varrho_{k,l}(X_{i,j},\tau), \qquad \tau\in\tilde{U}_\theta,$$

we see that

$$\frac{1}{T_n}\sum_{j=1}^{m}\sum_{i=1}^{N_{T_{j,n}}} \varrho_{k,l}(X_{i,j},\bar{\theta}_n)$$

tends stochastically to $\frac{1}{\mu}E\varrho_{k,l}(X_{1,1},\theta)$ and $\left(\frac{1}{\mu}E\varrho_{k,l}(X_{11},\theta)\right)_{1\le k,l\le p}$ is the matrix $\frac{1}{\mu}\sum(\theta)$. The rest follows from the representation (10).

4 Consistency and Asymptotic Normality of Minimum Contrast Estimations for the Weibull and the Maxwell Distribution

In many practical problems the Weibull distribution is used to describe the random failure time of an item. In this section the model of section 2 is used to show, that the maximum likelihood estimation of parameters of the Weibull distribution is asymptotically normal distributed, if the sample is the result of observation a renewal process.
In the case of the Weibull distribution the density of renewal time has the form

$$f_\theta(x) = k \cdot \vartheta^{-k} \cdot x^{k-1} \cdot \exp\left(-\left(\frac{x}{\vartheta}\right)^k\right), \qquad t > 0. \tag{11}$$

Let be $\theta = (\theta_1, \theta_2) = (k, \vartheta)$, where $k > 0$ is the form parameter and $\vartheta > 0$ is the scale parameter. Θ is assumed to be compact.
From (3) the loglikelihood for observations of m renewal processes in $[0, T_j]$, $j = 1, \ldots, m$ is given by

$$\begin{aligned}\log L &= \sum_{j=1}^{m}\sum_{i=1}^{N_{T_j}} \left(\log k - \log\vartheta + (k-1)(\log x_{i,j} - \log\vartheta) - \left(\frac{x_{i,j}}{\vartheta}\right)^k\right) \\ &\quad - \sum_{j=1}^{m}\left(\frac{R_{T_j}}{\vartheta}\right)^k. \end{aligned} \tag{12}$$

The maximum likelihood estimations for the parameters are defined by

$$\frac{1}{\hat{k}} + \frac{\sum_{j=1}^m \sum_{i=1}^{N_{T_j}} \log x_{i,j}}{\sum_{i=1}^m N_{T_j}} - \frac{\sum_{j=1}^m \sum_{i=1}^{N_{T_j}} x_{i,j}^{\hat{k}} \log x_{i,j} + \sum_{j=1}^m R_{T_j}^{\hat{k}} \log R_{T_j}}{\sum_{j=1}^m \sum_{i=1}^{N_{T_j}} x_{i,j}^{\hat{k}} + \sum_{j=1}^m R_{T_j}^{\hat{k}}} = 0 \tag{13}$$

and

$$\hat{\vartheta} = \left[\frac{\sum_{j=1}^m R_{T_j}^{\hat{k}} + \sum_{j=1}^m \sum_{i=1}^{N_{T_j}} x_{i,j}^{\hat{k}}}{\sum_{j=1}^m N_{T_j}}\right]^{1/\hat{k}}. \tag{14}$$

To apply the concept of minimum contrast estimation to this case we remark, that

$$\varrho(X_{i,j}, \theta) = -\left(\log k - \log\vartheta + (k-1)(\log X_{i,j} - \log\vartheta) - \left(\frac{X_{i,j}}{\vartheta}\right)^k\right)$$

is a contrast function (see example 1) and

$$\tilde{\varrho}(R_{T_j}) = -\left(\frac{R_{T_j}}{\vartheta}\right)^k$$

is a measurable function. Both $\varrho(x, \theta)$ and $\tilde{\varrho}(x, \theta)$ are continuous functions on $[0, \infty) \times \Theta$ with values in $(-\infty, +\infty)$ and twice continuously differentiable with respect to each θ_j, $j = 1, 2$,

in Θ^0.
To prove the other conditions of Theorems 1 and 2 the matrixes Σ and $\tilde{\Sigma}$ must be calculated. In the case of the Weibull distribution the Rao-Cramer conditions of regularity are fulfilled and it is easy to prove, that $\tilde{\Sigma} = \Sigma$. Σ is the Fisher information matrix of one observation and can be calculated in the following way:

$$\begin{aligned}
\varrho_{1,1}(X_{1,1},\theta) &= \frac{1}{k^2}\left(1+\left(\frac{X_{i,j}}{\vartheta}\right)^k \log^2\left(\frac{X_{i,j}}{\vartheta}\right)^k\right) \\
\varrho_{1,2}(X_{1,1},\theta) &= \frac{1}{\vartheta}\left(1-\left(\frac{X_{i,j}}{\vartheta}\right)^k \log\left(\frac{X_{i,j}}{\vartheta}\right)^k - \left(\frac{X_{i,j}}{\vartheta}\right)^k\right) \\
\varrho_{2,2}(X_{1,1},\theta) &= -\frac{k}{\vartheta^2}+\frac{k(k+1)}{\vartheta^2}\left(\frac{X_{i,j}}{\vartheta}\right)^k .
\end{aligned}$$

Using the equations $E(\frac{\partial\varrho(X_{1,1},\theta)}{\partial\theta_k}) = 0,\ k = 1,2$, after a substitution under the integral the information of one realization takes the form

$$\begin{aligned}
\Sigma_{11} &= \frac{1}{k^2}\left(1+\int_0^\infty u\ (\log u)^2 e^{-u}du\right) \\
\Sigma_{12} &= \frac{1}{\vartheta}\left(1-\int_0^\infty ue^{-u}\log u\, du - \int_0^\infty ue^{-u}du\right) \\
\Sigma_{22} &= -\frac{k}{\vartheta^2}+\frac{k(k+1)}{\vartheta^2}\int_0^\infty ue^{-u}du.
\end{aligned}$$

The integrals can be calculated using the gamma function and its derivations. Finally, we get

$$\Sigma = \begin{bmatrix} \dfrac{\pi^2/6+(1-c)^2}{k^2} & -\dfrac{1-c}{\theta} \\ -\dfrac{1-c}{\theta} & \dfrac{k^2}{\theta^2} \end{bmatrix}$$

where c is Euler's constant $c \approx 0,5772$.
Now it is easy to see that the conditions of the Theorems 1 and 2 are fulfilled, i.e. $\hat{\theta}_n$ from (13) and (14) is weakly consistent and the distribution of

$$\sqrt{T_n}\left(\hat{\theta}_n - \theta\right)$$

converges weakly to a 2-dimensional normal distribution with expectation zero and covariance matrix $C(\theta) = \frac{1}{\mu}\Sigma^{-1}(\theta)$ where μ is the expectation of the Weibull distribution

$$\mu = \vartheta\Gamma(1+\frac{1}{k}).$$

The fact that the maximum likelihood estimation is asymptotically normal distributed can be used to construct simultaneous confidence regions for the parameters of the Weibull

distribution.
A level line of the two-dimensional Gaussian distribution can be determined such that inside of this line there is $(1-\alpha)$ of the probability mass of the distribution. The equation of this level line is given by the quadratic form

$$(\hat{\theta}-\theta)^T \Sigma^{-1} (\hat{\theta}-\theta) < a^2 \tag{15}$$

with $\hat{\theta} = (\hat{k}, \hat{\vartheta})$, $\theta = (k, \vartheta)$, where $a^2 = -2\log\alpha$ is the quantil of the χ^2-distribution with 2 degrees of freedom and Σ is the Fisher information matrix. To calculate such a level line numerical methods must be used.

To compare the maximum likelihood estimator with the minimum contrast estimator considered in example 2 we investigate the scale model $f_\theta(x) = \theta f(\theta x)$, $\theta \in [a,b] \subseteq (0,\infty)$. Put $\varrho(x,\theta) = -\log\theta + x\theta$ and $\tilde{\varrho}(x,\theta) = \theta x$. Then

$$Q(\theta) = \sum_{j=1}^{m} \sum_{i=1}^{N_{T_j}} (-\log\theta + x_{i,j}\theta) + \sum_{j=1}^{m} \theta R_{T_j}.$$

The minimum contrast estimation is then given by

$$\tilde{\theta} = \frac{\sum_{j=1}^{m} N_{T_j}}{\sum_{j=1}^{m} T_j}.$$

Suppose f is twice continuously differentiable and $\int_0^\infty t^2 f(t)dt < \infty$. We obtain from Theorem 2 that for $T_n \to \infty$ the distribution of the corresponding $\tilde{\theta}_n$ is asymptotically normal with expectation θ and variance $T_n\sigma^2\theta^5$ where $\sigma^2 = \int_0^\infty (t-\frac{1}{\theta})^2 f(t)dt$. Now we specialize f and set

$$f(x) = \frac{2}{\sqrt{2\pi}} x^2 \exp\left(-\frac{x^2}{2}\right) \quad \text{for} \quad x \geq 0$$

and $f(x) = 0$ for $x < 0$. Calculating σ^2 one easily derives that the distribution of $\sqrt{T_n}(\hat{\theta}_n - \theta)$ converges weakly to a normal distribution with mean zero and variance

$$C(\theta) = \theta^3 \cdot 2\left(\sqrt{\pi/2} - 1\right).$$

Otherwise the loglikelihood function is given by

$$\begin{aligned} \log L \quad &= \quad \sum_{j=1}^{m} \sum_{i=1}^{N_{T_j}} \left(\log\frac{2}{2\pi} + 3\log\theta - 2\log x_{i,j} - \frac{\theta^2 x_{ij}^2}{2}\right) \\ &\quad + \sum_{j=1}^{m} \log[1 - F(R_{T_j})], \end{aligned}$$

where $F(\cdot)$ is the distribution function of the Maxwell distribution. This distribution function can be calculated using normal distribution. But to found the maximum likelihood

estimation observing renewal processes numerical methods have to be used.
To find asymptotic confidence intervals for the parameter θ the asymptotic properties of minimum contrast estimations can be used also. It is easy to see that the conditions of the Theorems 1 and 2 are fulfilled and the asymptotic variance of the maximum likelihood estimation is given by

$$C(\theta) = \theta^3 \frac{\sqrt{\pi/2}}{12}.$$

The maximum likelihood estimation gives a shorter confidence interval but is more complicated to calculate as the minimum contrast estimation from example 2.

References

[1] **Gut, A. :** *Stopped Random Walks. Limit Theorems and Applications.* New York 1988

[2] **Ibragimov, I.A.** and **Has'minskii, R.Z. :** *Statistical Estimation: Asymptotic Theory.* New York 1987

[3] **Liptser, R.S.** and **Shiryayev, A.N. :** *Statistics of random processes II. Applications.* New York 1978

[4] **Mann, N.R., Schafer, R.E.** and **Singpurwalla, N.P. :** *Methods for Statistical Analysis of Reliability and Life Data.* New York 1974

[5] **Pfanzagl, J. :** *On the measurability and consistency of minimum contrast estimates.* Metrika **14**, 249-276 (1969)

[6] **Revesz, P. :** *Die Gesetze der großen Zahlen.* Budapest 1968

[7] **Serfling, R. :** *Approximation Theorems of Mathematical Statistics.* New York 1980

received: September 9, 1991 **revised:** January 9, 1992

Author:

PD Dr. W. Kahle
Technische Universitt „Otto von Guericke“
Fakultt fr Mathematik
Universittsplatz 2
39106 Magdeburg
Germany

Prof. Dr. F. Liese
Universitt Rostock
Fachbereich Mathematik
Universittsplatz 1
18051 Rostock
Germany