Statistical Analysis of Redundant Systems

Vilijandas Bagdonavičius[1], Inga Masiulaitytė[1] and Mikhail Nikulin[2]

[1] Vilnius University, Vilnius, Lithuania;
[2] IMB, Victor Segalen University, Bordeaux, France `nikou@sm.u-bordeaux2.fr`

Summary. Nonparametric and parametric methods of estimation of redundant systems with "warm" stand-by units reliability are given. Asymptotic properties of the estimators and asymptotic confidence intervals are obtained. Power of goodness-of-fit tests from finite samples is investigated by simulation.

Keywords and phrases: AFT model, Confident Interval, Failure time, Maximum Likelihood, Nonparametric estimation, Parameter estimation, Redundant system, Reliability, Sedyakin's model, Warm stand-by unit, Weibull distribution

1 Introduction

Let us consider redundant systems with one main unit and $m-1$ stand-by units operating in "warm" conditions, i.e. under lower stress than the main one. We shall use notation $S(1, m-1)$ for such systems.

The problem is to obtain confidence intervals for the cumulative distribution functions of redundant systems using failure data of two groups of units, the first group functioning in "hot" and second – in "warm" conditions.

We suppose that switching from "warm" to "hot" conditions does not do any damage to units. Bagdonavičius, Masiulaityte and Nikulin [BMN08] give mathematical formulation of "fluent switch on" and propose tests for verification of this hypothesis. The formulation is based on the "principle of Sedyakin", see [Sed66].

Denote by T_1, F_1 and f_1 the failure time, the c.d.f. and the probability density function of the main unit. The failure times of the stand-by units denote by $T_2, \ldots, T_m$. In "hot" conditions their distribution functions are also F_1. In "warm" conditions the c.d.f. of T_i is F_2 and the p.d.f. is f_2, $i = 2, \ldots, m$. If a stand-by unit is switched to "hot" conditions, its c.d.f. is different from F_1 and F_2. For $i = 1, 2$ denote by $S_i = 1 - F_i$, $\lambda_i = f_i / S_i$ and $\Lambda_i = -\ln S_i$ the survival function, hazard rate and cumulative hazard, respectively.

The failure time of the system $S(1, m-1)$ is $T^{(m)} = T_1 \vee T_2 \vee \ldots \vee T_m$. Denote by K_j and k_j the c.d.f. and the p.d.f. of $T^{(j)}$, respectively, $(j = 2, \ldots, m)$, $K_1 = F_1$, $k_1 = f_1$. The c.d.f K_j can be written in terms of the c.d.f. K_{j-1} and F_1:

$$K_j(t) = \mathbf{P}(T^{(j)} \le t) = \int_0^t \mathbf{P}(T_j \le t | T^{(j-1)} = y) dK_{j-1}(y). \qquad (1)$$

The "fluent switch on" hypothesis H_0 formulated in [BMN08] states that

$$f_{T_j|T^{(j-1)}=y}(t) = \begin{cases} f_2(t) & \text{if} \quad t \le y, \\ f_1(t + g(y) - y) & \text{if} \quad t > y; \end{cases}, \quad g(y) = F_1^{-1}(F_2(y)). \qquad (2)$$

This model implies that

$$K_j(t) = \int_0^t F_1(t + g(y) - y) dK_{j-1}(y). \qquad (3)$$

So the distribution function K_m of the system with $m-1$ stand-by units is defined recurrently using the formula (3).

In particular, if we suppose that the distribution of units functioning in "warm" and "hot" conditions differ only in scale, i.e. $F_2(t) = F_1(rt)$ for all $t \ge 0$ and some $r > 0$, then $g(y) = ry$. Combining this assumption and the model (2) we have more strict hypothesis H_0^*. This hypothesis can also be considered as generalization of the accelerated failure time (AFT) model (Bagdonavičius [Bag78]) to the case of stress with random switch on.

2 Point estimators of the c.d.f. of redundant systems

Suppose that the following data are available :

a) complete ordered sample $T_{11}, \dots, T_{1n_1}$ of the failure times of units tested in "hot" conditions;

b) the time to obtain complete data in "warm" conditions may be long, so we suppose that n_2 units are tested time t_1 in "warm" conditions and the ordered first failure times $T_{21}, \dots, T_{2m_2}$ are obtained.

2.1 Nonparametric estimation

Denote by

$$\hat{F}_j(t) = \frac{1}{n_i} \sum_{i=1}^{n_j} \mathbf{1}_{\{T_{ji} \le t\}}, \quad \hat{F}_j^{-1}(y) = \inf\{s : \hat{F}_j(s) \ge y\}$$

the empirical distribution function and its inverse, respectively, for the jth sample.

The estimator of the function $g(t)$ is

$$\hat{g}(t) = \hat{F}_1^{-1}(\hat{F}_2(t)), \quad t \le t_1.$$

Under H_0 for any $t \le t_1$ the value $K_j(t)$ of the c.d.f is estimated recurrently:

$$\hat{K}_j(t) = \int_0^t \hat{F}_1(t + \hat{g}(y) - y) d\hat{K}_{j-1}(y), \quad \hat{K}_1(t) = \hat{F}_1(t). \qquad (4)$$

If we suppose that the distribution of units functioning in "warm" and "hot" conditions differ only in scale, i.e. $g(y) = ry$ then the c.d.f. $K_j(t)$ can be estimated at any point $t \ge 0$ replacing $\hat{g}(y)$ by $\hat{r}y$ in (4), where $\hat{r}$ is a convenient estimator of r. Set

$$N_1(t) = \sum_{i=1}^{n_1} \mathbf{1}_{\{T_{1i} \le t\}}, \quad N_2(t) = \sum_{i=1}^{n_2} \mathbf{1}_{\{T_{2i} \le t, t \le t_1\}},$$

$$Y_1(t) = \sum_{i=1}^{n_1} \mathbf{1}_{\{T_{1i} \ge t\}}, \quad Y_2(t) = \sum_{i=1}^{n_2} \mathbf{1}_{\{T_{2i} \ge t, t \le t_1\}}.$$

Bagdonavičius *et al* (2008) give the following estimator of the parameter r:

$$\hat{r} = \tilde{U}^{-1}(0) = sup\{r : \tilde{U}(r) > 0\};$$

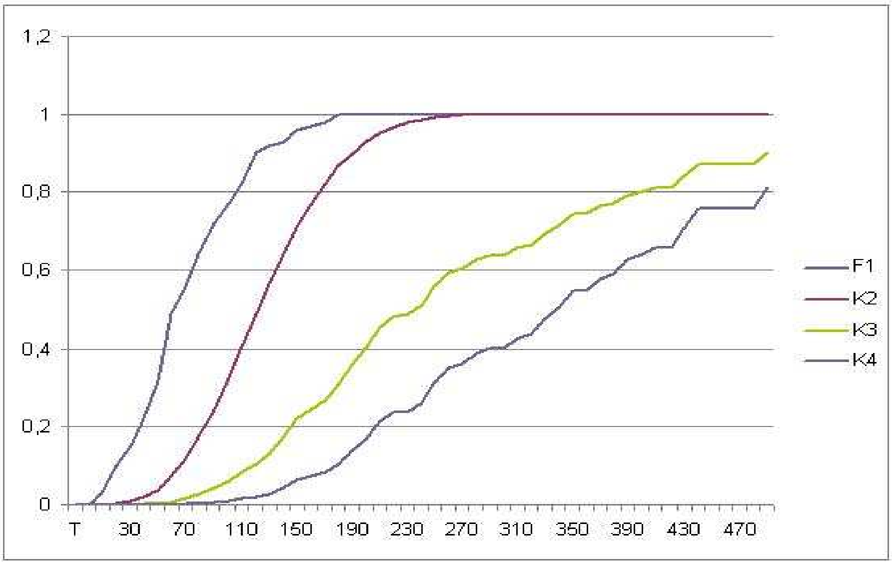
here

$$\tilde{U}(r) = -\int_0^{rt_1} \frac{Y_2(v/r)dN_1(v)}{Y_1(v) + Y_2(v/r)} + \int_0^{t_1} \frac{Y_1(ru)dN_2(u)}{Y_1(ru) + Y_2(u)},$$

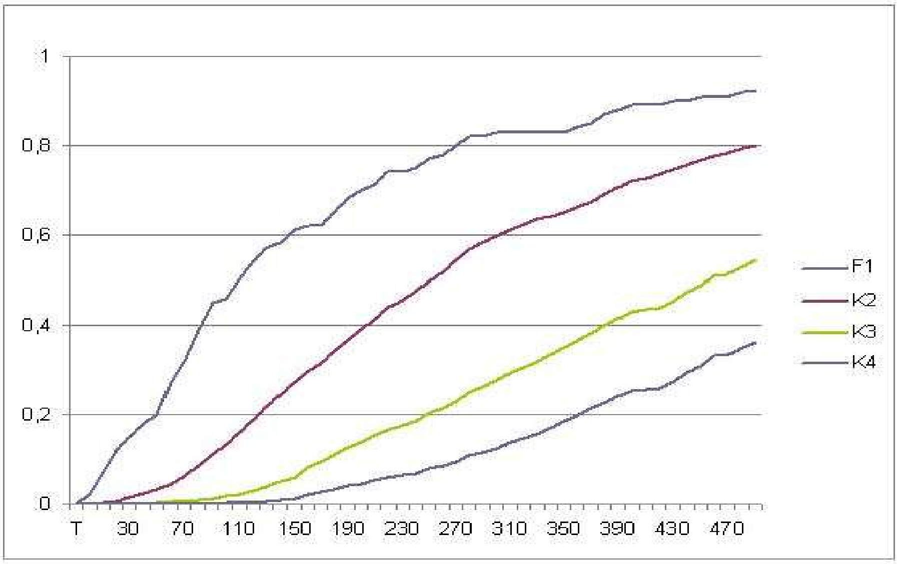
is $\hat{c}$adlag stochastic process with trajectories which are non-increasing step functions satisfying the inequalities $\tilde{U}(0+) > 0$, $\tilde{U}(+\infty) < 0$.

In the pictures 1 and 2 the graphs of the trajectories of the nonparametric estimators $\hat{F}_1, \hat{K}_i$ are given using data simulation when the main unit has Weibull or loglogistic distribution, respectively, and the hypothesis H_0^* is verified. It can be seen from these pictures how the reliability of redundant systems increases including supplementary stand-by units.

Picture 1. Graphs of the trajectories of the nonparametric estimators $\hat{F}_1, \hat{K}_i$ (Weibull distribution)



Picture 2. Graphs of the trajectories of the nonparametric estimators $\hat{F}_1, \hat{K}_i$ (loglogistic distribution)



2.2 Parametric estimation

Suppose that in hot conditions the c.d.f $F_1(t;\theta)$ is absolutely continuous and depends on finite dimensional parameter $\theta \in \Theta \subset \mathbf{R}^k$. Set $\gamma = (r, \theta^T)^T$.

The maximum likelihood estimator $\gamma^* = (r^*, (\theta^*)^T)^T$ of the parameter γ maximizes the loglikelihood function

$$\ell(\gamma) = \sum_{i=1}^{n_1} \ln f_1(T_{1i};\theta) + m_2 \ln r + \sum_{i=1}^{m_2} \ln f_1(rT_{2i};\theta) + (n_2 - m_2) \ln S_1(rt_1;\theta).$$

Under H_0^* for any $t \geq 0$ and $j \geq 2$ the c.d.f. $K_j(t)$ is estimated recurrently:

$$\hat{K}_j(t) = \int_0^t F_1(t + r^* y - y; \theta^*) d\hat{K}_{j-1}(y), \quad \hat{K}_1(t) = F_1(t;\theta^*). \tag{5}$$

3 Asymptotic distribution of $\hat{K}_j$ and confidence intervals for $K_j(t)$

Suppose that

$$\frac{n_i}{n} = l_i + O(\frac{1}{n}), \quad l_i \in (0,1), \quad \text{as} \quad n = n_1 + n_2 \to \infty.$$

3.1 Nonparametric case

The limit distribution of the empirical distribution functions is well known:

$$\sqrt{n}(\hat{F}_i - F_i) \xrightarrow{\mathcal{D}} U_i \tag{6}$$

on $D(A_i)$, where $\xrightarrow{\mathcal{D}}$ means weak convergence, $A_1 = [0,\infty)$, $A_2 = [0,t_1]$, U_1, U_2 are independent Gaussian martingales with $U_i(0)=0$ and the covariances

$$\mathbf{cov}(U_i(u),U_i(v)) = \frac{1}{l_i}\,F_i(u\wedge v)S_i(u\vee v). \tag{7}$$

Let us find the asymptotic distribution of the estimator $\hat r$. Denote by $r_0 \in (0,1)$ the true value of r. Under the model H_0^* it is the ratio of the mean failure times μ_1 and μ_2 of units functioning in "hot" and "warm" conditions, respectively.

Lemma. *Suppose that the c.d.f. F_1 is absolutely continuous with positive p.d.f. f_1 on $(0,\infty)$ and the hypothesis $F_2(t)=F_1(r_0t)$ is true. If*

$$A = -\frac{1}{r_0}\int_0^{r_0t_1} uf_1(u)d\Lambda_1(u) - t_1f_1(r_0t_1) \neq 0, \tag{8}$$

then

$$\sqrt{n}(\hat r - r_0) \xrightarrow{d} Y = -\frac{W}{A}, \tag{9}$$

where

$$W = -\int_0^{t_1}[U_1(r_0u) - U_2(u)d\Lambda_2(u) - U_1(r_0t_1) + U_2(t_1), \tag{10}$$

Remark. *If samples are complete then*

$$W = -\int_0^{\infty}[U_1(r_0u)-U_2(u)]d\Lambda_2(u), \quad A = -\frac{1}{r_0}\int_0^{\infty} uf_1(u)d\Lambda_1(u), \tag{11}$$

and

$$\sqrt{n}(\hat r - r_0) \xrightarrow{\mathcal{D}} Y = -\frac{W}{A} \sim N(0,\frac{1}{l_1l_2A^2}). \tag{12}$$

Theorem. *If F_1 is continuously differentiable on $[0,\infty)$ then under H_0^* for any $t>0$ and any natural $j\geq 2$*

$$\sqrt{n}(\hat K_j(s) - K_j(s)) \xrightarrow{\mathcal{D}}$$

$$W_j(s) = \int_0^s U_1(s+r_0y-y)dK_{j-1}(y) + \mu^{(j-1)}(s)Y + \int_0^s F_1(s+r_0y-y)dW_{j-1}(y) \tag{13}$$

on $D[0,t]$, where $W_1(s)=U_1(s)$, $\mu^{(j-1)}(s) = \int_0^s yf_1(s+r_0y-y)dK_{j-1}(y)$.

The asymptotic variance of $\sqrt{n}(\hat K_j(t) - K_j(t))$, $j\geq 2$ might be estimated recurrently, using the equation (13): the covariances $\mathbf{Cov}(W_j(s),W_j(t)) = \mathbf{E}(W_j(s)W_j(t))$ can be written in terms of the covariances

$$\mathbf{E}(W_{j-1}(u)W_{j-1}(v)), \quad \mathbf{E}(W_{j-1}(u)U_1(v)), \quad \mathbf{E}(W_{j-1}(u)U_2(v)),$$

$$\mathbf{E}(U_1(u)U_1(v)), \quad \mathbf{E}(U_2(u)U_2(v)).$$

Note that for $j=2$ these covariances are

$$\mathbf{E}(W_1(u)W_1(v)) = \mathbf{E}(W_1(u)U_1(v)) = \mathbf{E}(U_1(u)U_1(v)) = \frac{1}{l_1}\,F_1(u\wedge v)S_1(u\vee v),$$

$$\mathbf{E}(W_1(u)V_2(v)) = \mathbf{E}(U_1(u)U_2(v)) = 0, \quad \mathbf{E}(U_2(u)U_2(v)) = \frac{1}{l_2}\,F_2(u\wedge v)S_2(u\vee v).$$

Let us find the asymptotic variance of $\sqrt{n}(\hat{K}_2(t) - K_2(t))$. Suppose first that samples are complete. In the following we skip the index in r_0 .

The formulas (10) and (11) imply that $\sqrt{n}(\hat{K}_2(s) - K_2(s)) \xrightarrow{\mathcal{D}} W_2(s)$, where

$$W_2(t) = F_2(t)U_1(t) + \int_0^t U_1(t+ry-y)dF_1(y) - \int_0^t U_1(y)dF_1(t+ry-y)+$$

$$\frac{\mu(t)}{A}\left(\int_0^\infty U_1(ry)d\Lambda_2(y) - \int_0^\infty U_2(y)d\Lambda_2(y)\right) = (V_1+V_2+V_3+V_4)(t). \qquad (14)$$

$$\mu(t) = \mu^{(1)}(t) = \int_0^t yf_1(t+ry-y)dF_1(y), \quad A = -\frac{1}{r}\int_0^\infty uf_1(u)d\Lambda_1(u). \qquad (15)$$

The random variable $W_2(t)$ has zero mean. Set $\nu(t) = \int_0^t F_1(t+ry-y)\,dF_1(y)$. The variance $\mathbf{Var}(W_2(t))$ is defined by the following formula:

$$l_1\mathbf{Var}(W_2(t)) = -4\nu^2(t) + \int_0^t F_1(t+ry-y)[F_1(t+ry-y) + 2F_1(y)]\,dF_1(y)+$$

$$2F_1(t)\nu(rt) + 2\int_{rt}^t F_1(t+ry-y)F_1((t-y)/(1-r))\,dF_1(y) + \frac{\mu^2(t)}{l_2A^2}+$$

$$\frac{2\mu(t)}{A}\left[\nu(t) + \int_0^t [F_1(t+ry-y)\ln S_1(y) - S_1(t+ry-y)\ln S_1(t+ry-y)]dF_1(y))\right].$$

Set

$$Z_{1i} = \hat{F}_1(t+(\hat{r}-1)T_{1i}-), \quad \hat{F}_1(t-) = \frac{1}{n_1}\sum_{i=1}^{n_1}\mathbf{1}_{\{T_{1i}<t\}}, \quad Z_{2i} = \hat{F}_1(\frac{t-T_{1i}}{1-\hat{r}}-),$$

$$Z_{3i} = \hat{F}_1(T_{1i}-), \quad Z_{4i} = \hat{f}_1(t+(\hat{r}-1)T_{1i}-), \quad \hat{\mu}(t) = \frac{1}{n_1}\sum_{T_{1i}\le t} T_{1i}Z_{4i}, \quad Z_{5i} = \hat{f}_1(T_{1i}-).$$

The variance $\mathbf{Var}(W_2(t))$ is estimated using the statistic

$$\frac{n_1}{n}\hat{\mathbf{Var}}(W_2(t)) = -4\hat{\phi}_1^2(t) + \hat{\phi}_2(t) + \frac{n\hat{\mu}^2(t)}{n_2\hat{A}^2} + \frac{2\hat{\mu}(t)}{\hat{A}}\hat{\phi}_3(t);$$

here

$$\hat{\phi}_1(t) = \frac{1}{n_1}\sum_{T_{1i}\le t} Z_{1i}, \quad \hat{A} = -\frac{1}{\hat{r}n_1}\sum_{i=1}^{n_1}\frac{T_{1i}Z_{5i}}{1-Z_{3i}},$$

$$\hat{\phi}_2(t) = \frac{1}{n_1}\sum_{T_{1i}\le t} Z_{1i}[Z_{1i} + 2Z_{3i} + 2\hat{F}_1(t)\mathbf{1}_{\{T_{1i}\le\hat{r}t\}} + 2Z_{2i}\mathbf{1}_{\{T_{1i}>\hat{r}t\}}],$$

$$\hat{\phi}_3(t) = \frac{1}{n_1} \sum_{T_{1i} \le t} [Z_{1i}(1 + \ln(1 - Z_{3i}) - (1 - Z_{1i}) \ln(1 - Z_{1i})].$$

So the variance $\sigma^2_{\hat{K}_2}$ of the estimator $\hat{K}_2(t)$ is estimated by

$$\hat{\sigma}^2_{\hat{K}_2} = \frac{1}{n} \hat{\mathbf{Var}}(W_2(t)).$$

The asymptotic $1 - \alpha$ confidence interval for $K_2(t)$ is $(\underline{K}_2(t), \overline{K}_2(t)$, where

$$\underline{K}_2(t) = \left(1 + \frac{1 - \hat{K}_2(t)}{\hat{K}_2(t)} \exp\left\{ \frac{\hat{\sigma}_{\hat{K}_2} z_{1-\alpha/2}}{\sqrt{\hat{K}_2(t)(1 - \hat{K}_2(t))}} \right\} \right)^{-1},$$

$$\overline{K}_2(t) = \left(1 + \frac{1 - \hat{K}_2(t)}{\hat{K}_2(t)} \exp\left\{ - \frac{\hat{\sigma}_{\hat{K}_2} z_{1-\alpha/2}}{\sqrt{\hat{K}_2(t)(1 - \hat{K}_2(t))}} \right\} \right)^{-1}. \tag{16}$$

Remark. *In the case of censoring the expression in parenthesis of the term V_4 in (17) is replaced by $\int_0^{t_1} [U_1(ru) - U_2(u) d\Lambda_2(u) + U_1(rt_1) + U_2(t_1)$, so only minor modifications are needed.*

We investigated finite sample confidence level of the proposed asymptotic confidence intervals. The failure times T_{1j} and T_{2j} were simulated from exponential distribution:

$$T_{1j} \sim \mathcal{E}(\lambda_1), \quad T_{2j} \sim \mathcal{E}(\lambda_2), \quad \lambda_1 = \frac{1}{100}; \quad \lambda_2 = \frac{1}{300}.$$

The number of replications was 2000. For various values of t the proportions of confidence interval realizations covering the true value of the distributional function $K_2(t)$ are given in the

Table 1. Confidence level for finite samples ($n_1 = n_2 = 100$)

Time t	50	100	150	200	300	400	500
$K_2(t)$	0.114	0.319	0.513	0.667	0.856	0.941	0.977
Confidence level (%)	91.1	90.8	90.5	90.3	89.8	89.4	89.1

3.2. Parametric case

Denote by $I_n(\gamma) = -\mathbf{E}\ddot{\ell}(\gamma)$ the Fisher information matrix and suppose that $\frac{1}{n} I_n(\gamma) \to i(\gamma)$. Under classical assumptions on the family of distributions $f_1(t, \theta)$ the maximum likelihood estimator γ^* is asymptotically normal:

$$\sqrt{n}(\gamma^* - \gamma) \xrightarrow{d} Y = (Y_1, Y_2^T)^T \sim N_{k+1}(0, i^{-1}(\gamma)).$$

Y_1 is one-dimensional, Y_2 – k-dimensional.

Using delta method we obtain:

$$\sqrt{n}(\hat{K}_2(t) - K_2(t)) \xrightarrow{\mathcal{D}} W_2(t) = C_2^T(t; \gamma) Y,$$

where

$$C_2(t;\gamma) = (C_{21}(t;\gamma), C_{22}^T(t;\gamma))^T, \quad C_{21}(t;\gamma) = \int_0^t \frac{\partial}{\partial r} F_1(t+ry-y;\theta)dF_1(y;\theta),$$

$$C_{22}(t;\gamma) = \int_0^t \frac{\partial}{\partial \theta} F_1(t+ry-y;\theta)dF_1(y;\theta) + F_1(t+ry-y;\theta)d(\frac{\partial}{\partial \theta}F_1(y;\theta)).$$

The random variable $W_2(t)$ is linear function of Y.

If $j \geq 2$ then

$$\sqrt{n}(\hat{K}_j(t) - K_j(t)) \xrightarrow{\mathcal{D}} W_j(t).$$

The random variable $W_j(t)$, $j \geq 2$, is also linear function of Y:

$$W_j(t) = Y^T C_j(t;\gamma), \quad C_j(t;\gamma) \in (C[0,t])^{k+1}.$$

It follows by induction:

$$W_j(t) = Y^T(\int_0^t \frac{\partial}{\partial \gamma} F_1(t+ry-y;\theta)dK_{j-1}(y;\gamma) + F_1(t+ry-y;\theta)dC_{j-1}(t;\gamma)).$$

So the variance

$$\mathbf{Var}(W_j(t)) = \mathbf{Var}(C_j(t;\gamma)^T Y) = C_j^T(t;\gamma)i^{-1}(\gamma)C_j(t;\gamma)$$

is estimated by $nC_2^T(t;\hat{\gamma})I^{-1}(\hat{\gamma})C_j(t;\hat{\gamma})$, and the variance $\sigma^2_{\hat{K}_j(t)}$ of the estimator $\hat{K}_2(t)$ is estimated by

$$\hat{\sigma}^2_{\hat{K}_j(t)} = C_j^T(t;\hat{\gamma})I^{-1}(\hat{\gamma})C_j(t;\hat{\gamma}).$$

The matrix $I(\hat{\gamma})$ may be replaced by $-\ddot{\ell}(\hat{\gamma})$.

The asymptotic $1-\alpha$ confidence interval for $K_j(t)$ is $(\underline{K}_j(t), \overline{K}_j(t)$, where

$$\underline{K}_j(t) = \left(1 + \frac{1-\hat{K}_j(t)}{\hat{K}_j(t)} \exp\left\{\frac{\hat{\sigma}_{\hat{K}_j} z_{1-\alpha/2}}{\sqrt{\hat{K}_j(t)(1-\hat{K}_j(t))}}\right\}\right)^{-1},$$

$$\overline{K}_j(t) = \left(1 + \frac{1-\hat{K}_j(t)}{\hat{K}_j(t)} \exp\left\{-\frac{\hat{\sigma}_{\hat{K}_j} z_{1-\alpha/2}}{\sqrt{\hat{K}_j(t)(1-\hat{K}_j(t))}}\right\}\right)^{-1}. \tag{17}$$

Example 1. *Exponential distribution:* $S_1(t) = e^{-\lambda t}$.

Suppose that samples are complete. The logarithm of the parametric likelihood function is

$$\ell(r,\mu,\nu) = n\ln\lambda - \lambda(\sum_{i=1}^{n_1} T_{1i} + r\sum_{i=1}^{n_2} T_{2i}) + n_2 \ln r,$$

the MLE estimators are $r^* = \hat{\mu}_1/\hat{\mu}_2$, $\lambda^* = 1/\hat{\mu}_1$, where $\hat{\mu}_j = \frac{1}{n_j}\sum_{i=1}^{n_j} T_{ji}$, $j = 1,2$. The Fisher information matrix is

$$I(r,\lambda) = \begin{pmatrix} \frac{n_2}{r^2} & \frac{n_2}{r\lambda} \\ \frac{n_2}{r\lambda} & \frac{n}{\lambda^2} \end{pmatrix}, \quad \frac{1}{n}I(r,\lambda) \to i(r,\lambda) = \begin{pmatrix} \frac{l_2}{r^2} & \frac{l_2}{r\lambda} \\ \frac{l_2}{r\lambda} & \frac{1}{\lambda^2} \end{pmatrix}, \quad i^{-1}(r,\lambda) = \begin{pmatrix} \frac{r^2}{l_1 l_2} & -\frac{r\lambda}{l_1} \\ -\frac{r\lambda}{l_1} & \frac{\lambda^2}{l_1} \end{pmatrix}.$$

The partial derivatives are:

$$\frac{\partial}{\partial r}F_1(t+ry-y;\lambda) = \lambda y e^{-\lambda(t+ry-y)}, \quad \frac{\partial}{\partial \lambda}F_1(t+ry-y;\lambda) = (t+ry-y)e^{-\lambda(t+ry-y)}.$$

Note that $K_1(t) = F_1(t)$.

So $C_2 = (C_{21}, C_{22})^T$, where

$$C_{21}(t;r,\lambda) = \frac{1}{r^2}S_1(t)[F_2(t) - \lambda rt S_2(t)], \quad C_{22}(t;r,\lambda) = S_1(t)F_2(t)\frac{1+r}{r}t.$$

The equality

$$\mathbf{Var}(W_2(t)) = \mathbf{Var}(C_2(t;r,\lambda)^T Y) = \frac{S_1^2(t)}{l_1 l_2 r^2}\{l_1[F_2(t)-\lambda rt S_2(t)]^2 + l_2[(1-\lambda t)F_2(t)-\lambda rt]^2\}$$

implies that

$$\hat\sigma^2_{\hat K_2(t)} = \frac{\hat S_1^2(t)}{n_1 n_2 \hat r^2}\{n_1[\hat F_2(t) - \lambda^* r^* t \hat S_2(t)]^2 + n_2[(1-\lambda^* t)\hat F_2(t) - \lambda^* r^* t]^2\},$$

where

$$\hat S_1(t) = e^{-\lambda^* t}, \quad \hat S_2(t) = e^{-\lambda^* r^* t}, \quad \hat F_i(t) = 1 - \hat S_i(t), \quad i = 1, 2.$$

The asymptotic $(1-\alpha)$ confidence interval for K_2 has the form (19) with

$$\hat K_2(t) = \hat F_1(t) - \frac{\hat S_1(t)\hat F_2(t)}{r^*},$$

Taking into consideration the equality

$$K_2(t) = F_1(t) - \frac{S_1(t)F_2(t)}{r},$$

the weights $C_3 = (C_{31}, C_{32})^T$ and the estimator

$$\hat\sigma^2_{\hat K_3(t)} = C_3^T(t;\hat\gamma)I^{-1}(\hat\gamma)C_3(t;\hat\gamma).$$

can be computed and the asymptotic confidence interval for $K_3(t)$ can be found, etc.

We investigated finite sample confidence level of the proposed asymptotic confidence intervals. The same simulated data was used as in Section 3.1. For various values of t the proportions of confidence interval realizations covering the true value of the distributional function $K_2(t)$ are given in Table 2.

Table 2. Confidence level (parametric estimation) for finite samples ($n_1 = n_2 = 100$)

Time t	50	100	150	200	300	400	500
$K_2(t)$	0.114	0.319	0.513	0.667	0.856	0.941	0.977
Confidence level (%)	91.75	91.0	91.7	91.6	91.1	89.5	88.9

Example 2. *Weibull distribution:* $S_1(t) = e^{-(t/\mu)^\nu}$

The logarithm of the parametric likelihood function is

$$\ell(r,\mu,\nu) = n(\ln\nu - \nu\ln\mu) + n_2\nu\ln r + (\nu-1)(\sum_{i=1}^{n_1}\ln T_{1i} + \sum_{i=1}^{n_2}\ln T_{2i}) - \mu^{-\nu}(\sum_{i=1}^{n_1} T_{1i}^{\nu} + r^{\nu}\sum_{i=1}^{n_2} T_{2i}^{\nu}),$$

and the Fisher information matrix

$$I(r,\mu,\nu) = \begin{pmatrix} \frac{n_2\nu^2}{r^2} & -\frac{n_2\nu^2}{r\nu} & \frac{n_2\Gamma'(2)}{r} \\ -\frac{n_2\nu^2}{r\nu} & \frac{n\nu^2}{\mu^2} & -\frac{n\Gamma'(2)}{\mu} \\ \frac{n_2\Gamma'(2)}{r} & -\frac{n\Gamma'(2)}{\mu} & \frac{n(1+\Gamma''(2))}{\mu^2} \end{pmatrix}.$$

The partial derivatives are:

$$\frac{\partial}{\partial r}F_1(t+ry-y;\mu,\nu) = \frac{\nu y}{\mu}(\frac{t+ry-y}{\mu})^{\nu-1}S_1(t+ry-y),$$

$$\frac{\partial}{\partial \mu}F_1(t+ry-y;\mu,\nu) = -\frac{\nu}{\mu}(\frac{t+ry-y}{\mu})^{\nu}S_1(t+ry-y),$$

$$\frac{\partial}{\partial \nu}F_1(t+ry-y;\mu,\nu) = (\frac{t+ry-y}{\mu})^{\nu}\ln(\frac{t+ry-y}{\mu})S_1(t+ry-y).$$

So $C_2 = (C_{21}, C_{22}, C_{23})^T$, where

$$C_{21}(t;r,\mu,\nu) = \int_0^t \frac{\partial}{\partial r}F_1(t+ry-y;\mu,\nu)dF_1(y),$$

$$C_{22}(t;r,\mu,\nu) = \int_0^t \frac{\partial}{\partial \mu}F_1(t+ry-y;\mu,\nu)dF_1(y) + F_1(t+ry-y;\mu,\nu)d(\frac{\partial}{\partial \mu}F_1(y;\mu,\nu)),$$

$$C_{23}(t;r,\mu,\nu) = \int_0^t \frac{\partial}{\partial \nu}F_1(t+ry-y;\mu,\nu)dF_1(y) + F_1(t+ry-y;\mu,\nu)d(\frac{\partial}{\partial \nu}F_1(y;\mu,\nu)).$$

The last three integrals can be computed numerically.

Let us investigate the efficiency $AE(\hat{r}, r^*)$ of the non-parametric estimator $\hat{r}$ with respect to the efficient parametric maximum likelihood estimator r^* for some parametric models used in reliability (complete samples).

Example 1 (continuation): *exponential distribution.*

In this case $A = -1/r_0$, so by (12)

$$\sqrt{n}(\hat{r} - r_0) \xrightarrow{\mathcal{D}} \hat{Y} \sim N(0, \frac{r_0^2}{l_1 l_2}).$$

The parametric MLE

$$r^* = \frac{\hat{\mu}_1}{\hat{\mu}_2}, \quad \hat{\mu}_j = \frac{1}{n_j}\sum_{i=1}^{n_j} T_{ji}.$$

Here $\hat{\mu}_j$ is the estimator of the mean $\mu_j = \mathbf{E}(T_{ji})$, $i = 1, \ldots, n_i$, $j = 1, 2$. Set $\sigma_j^2 = \mathbf{Var}(T_{ji})$, $j = 1, 2$. The convergence

$$\sqrt{n}(\hat{\mu}_j - \mu_j) \xrightarrow{\mathcal{D}} Y_j^* = -\int_0^\infty U_j(y)dy \sim N(0, \sigma_j^2/l_j)$$

implies

$$\sqrt{n}(r^* - r_0) \xrightarrow{\mathcal{D}} \tilde{Y} = \frac{1}{\mu_2}(rY_2^* - Y_1^*) \sim N(0, \frac{\sigma_1^2}{l_1 l_2 \mu_2^2}) = N(0, \frac{r_0^2}{l_1 l_2}).$$

So $AE(\hat{r}, r^*) = 1$.

Example 2 (continuation): *Weibull distribution.*

The first diagonal element of the inverse matrix I^{-1} is

$$I^{11} = \frac{nr^2}{n_1 n_2 \nu^2},$$

So

$$\sqrt{n}(r^* - r_0) \xrightarrow{\mathcal{D}} Y \sim N(0, \frac{r_0^2}{l_1 l_2 \nu^2}),$$

which means that in the Weibull case the nonparametric estimator $\hat{r}$ has the same asymptotic efficiency as the parametric MLE r^*.

Example 3. *Loglogistic distribution:* $S_1(t) = \frac{1}{1+(t/\mu)^\nu}$.

The logarithm of the parametric likelihood function is

$$\ell(r, \mu, \nu) = n(\ln \nu - \nu \ln \mu) + n_2 \nu \ln r + (\nu - 1)(\sum_{i=1}^{n_1} \ln T_{1i} + \sum_{i=1}^{n_2} \ln T_{2i}) -$$

$$2(\sum_{i=1}^{n_1} \ln(1 + (\frac{T_{1i}}{\mu})^\nu) + \sum_{i=1}^{n_2} \ln(1 + (\frac{rT_{2i}}{\mu})^\nu)),$$

and the Fisher information matrix

$$I(r, \mu, \nu) = \begin{pmatrix} \frac{n_2\nu^2}{3r^2} & -\frac{n_2\nu^2}{3r\nu} & 0 \\ -\frac{n_2\nu^2}{3r\nu} & \frac{n\nu^2}{3\mu^2} & 0 \\ 0 & 0 & \frac{n(3+2\Gamma''(2)-2(\Gamma'(2))^2)}{3\mu^2} \end{pmatrix}.$$

The first diagonal element of the inverse matrix I^{-1} is

$$I^{11} = \frac{3nr^2}{n_1 n_2 \nu^2},$$

So

$$\sqrt{n}(r^* - r_0) \xrightarrow{\mathcal{D}} Y \sim N(0, \frac{3r_0^2}{l_1 l_2 \nu^2})$$

which means that differently from the Weibull case relative asymptotic efficiency of the nonparametric estimator $\hat{r}$ with respect to the parametric MLE r^* is small: $ASE(\hat{r}, r^*) = 1/\sqrt{3} = 0,58$.

Acknowledgements: The authors are grateful to the University Victor Segalen and to "StatXpert" for the financial support.

References

[ABGK93] Andersen,P.K., Borgan, O.,Gill, R.D.and Keiding, N.: *Statistical Models Based on Counting Processes.* Springer: New York, (1993).

[Bag78] Bagdonavičius, V.: Testing the hyphothesis of the additive accumulation of damages. *Probab. Theory and its Appl.*, **23**, No. 2, 403–408 (1978).

[BMN07] Bagdonavicius, V., Masiulaityte,I., Nikulin, M. (2008). Statistical analysis of redundant systems with one stand-by unit. In: Huber, C., Limnios,N., Mesbah, M., Nikulin, M (eds) *Mathematical Methods in Survival Analysis, Reliability and Quality of Life*, ISTE/WILEY, London, 183–192 (2007).

[BMN08] Bagdonavicius, V., Masiulaityte,I., Nikulin, M. (2008). Statistical analysis of redundant systems with "warm" stand-by units. *Stochastics: An International Journal of Probability and Stochastical Processes*, **80**, #2-3, 115–128(2008).

[Sed66] Sedyakin,N.M. (1966). On one physical principle in reliability theory.(in russian). *Techn. Cybernetics*, **3**, 80-87.