Technometrics
2016, Vol. 58, No. 1, pp. 95-103.
Supplementary Materials

Tail estimation for window-censored processes

Holger Rootzén* and Dmitrii Zholud†

Residual life for gamma and Weibull distributions

Gamma distribution

Let $\Gamma(k,x) = \int_x^\infty y^{k-1}e^{-y}dy$ be the upper incomplete gamma function, and let $\Gamma(k) = \Gamma(k,0)$ be the gamma function. The gamma distribution with shape parameter $k > 0$ and scale parameter $\sigma > 0$ has the probability density and tail functions

$$f_0(x) = \frac{1}{\sigma^k \Gamma(k)} x^{k-1} e^{-x/\sigma} \quad \text{and} \quad \bar{F}_0(x) = \frac{1}{\Gamma(k)}\Gamma(k, x/\sigma),$$

respectively, and has mean $\mu_0 = k\sigma$. Thus, using the equality $\Gamma(k+1) = k\Gamma(k)$, the residual life probability density and cumulative distribution functions are

$$f_0^r(x) = \frac{1}{\sigma\Gamma(k+1)}\Gamma(k, x/\sigma) \text{ and } \bar{F}_0^r(x) = \frac{1}{\Gamma(k+1)}\left(\Gamma(k+1, x/\sigma) - \frac{x}{\sigma}\Gamma(k, x/\sigma)\right),$$

where $\bar{F}_0^r$ was obtained by changing the integration order. All the above holds for $x > 0$.

Weibull distribution

The probability density and the cumulative distribution functions of the Weibull distribution with shape parameter $k > 0$ and scale parameter $\sigma > 0$ are given by

$$f_0(x) = \frac{k}{\sigma}\left(\frac{x}{\sigma}\right)^{k-1} e^{-(x/\sigma)^k} \quad \text{and} \quad F_0(x) = 1 - e^{-(x/\sigma)^k},$$

respectively, and the mean is $\mu_0 = \sigma\Gamma(1 + 1/k)$. Thus, using a change of variables to obtain the second expression, and with $\gamma(x,k) = \int_0^x y^{k-1}e^y dy$, the lower incomplete gamma function, the density and distribution functions of the residual life time are

$$f_0^r(x) = \frac{1}{\sigma\Gamma(1+1/k)} e^{-(x/\sigma)^k} \quad \text{and} \quad F_0^r(x) = \frac{1}{\Gamma(1/k)}\gamma\left(\frac{1}{k}, \left(\frac{x}{\sigma}\right)^k\right).$$

Efficient algorithms for numerical computation of the incomplete gamma function are available in e.g. MatLab or Wolfram Mathematica software.

---

*Department of Mathematical Statistics
Chalmers University of Technology and University of Göteborg, Sweden.
E-mail: hrootzen@chalmers.se
†Department of Mathematical Statistics
Chalmers University of Technology and University of Göteborg, Sweden.
E-mail: dmitrii@chalmers.se

Asymptotic normality

Here we establish that the tail estimators $\hat{\theta} = (\hat{\sigma}, \hat{\gamma})$ from Subsection 2.2 satisfy the conditions of Theorem 5.23 of van der Vaart (1998), and hence are asymptotically normal. For this we assume that $\theta = (\sigma, \gamma)$ belongs to a compact set $K = [\sigma_\ell, \sigma_u] \times [\gamma_\ell, \gamma_u]$, with $0 < \sigma_\ell$ and $0 < \gamma_\ell < \gamma_u < 1$. Here the last restriction, as before, is to ensure that means exist. We use notation and equation numbers from the paper and assume that $0 < u < w$ so that the observation window is longer than the threshold.

Since σ is a scale parameter, it is only the ratio between σ and $w - u$ which determines asymptotic behavior, and not their individual values. Hence we fixed $w - u = 1$ in the calculations. Additionally, we did them for $\sigma_\ell = 0.025$ and $\sigma_u = 2.5$, since censoring is too mild, or too severe, for the estimation problem to be practically interesting for σ outside of this range. We further used $\gamma_\ell = 0.01$ and $\gamma_u = 0.99$ However, it is straightforward to consider also wider parameter ranges (and we in fact have done this too). Recall the notation $X = L - u$ for the excesses of the threshold u by the observed 0-intervals, and S for the length of the first 1-interval, and that X and S are assumed to be independent. Let

$$\begin{aligned}
m_{nc,\theta}(x) &= -\log\sigma - \left(\frac{1}{\gamma} + 1\right)\log\left(1 + \frac{\gamma}{\sigma}x\right), \\
m_{rc,\theta}(x) &= -\frac{1}{\gamma}\log\left(1 + \frac{\gamma}{\sigma}x\right), \\
m_{lc,\theta}(x) &= -\log\sigma + \log(1-\gamma) - \frac{1}{\gamma}\log\left(1 + \frac{\gamma}{\sigma}x\right), \\
m_{dc,\theta} &= -\left(\frac{1}{\gamma} - 1\right)\log\left(1 + \frac{\gamma}{\sigma}(w-u)\right),
\end{aligned}$$

and let $1_{nc}(X) = 1$ if X is non-censored and $1_{nc}(X) = 0$ otherwise, and define $1_{rc}, 1_{lc}$ and 1_{dc} similarly. Next, set

$$\begin{aligned}
m_{0,\theta}(X) &= 1_{lc}(X)m_{lc,\theta}(X) + 1_{dc}(X)m_{dc,\theta} \\
m_{1,\theta}(X) &= 1_{nc}(X)m_{nc,\theta}(X) + 1_{rc}(X)m_{rc,\theta}(X) \\
m_\theta(X) &= m_{0,\theta}(X) + m_{1,\theta}(X)
\end{aligned}$$

so that, with $\ell_u(\sigma, \gamma)$ defined in (7) , and with $X_1, \ldots X_n$ the entire set of $n = \bar{n}_{nc} + \bar{n}_{rc} + \bar{n}_{lc} + \bar{n}_{dc}$ observed 0-intervals, we have that

$$\ell_u(\theta) = \sum_{i=1}^{n} m_\theta(X_i).$$

Now, inspection of the partial derivatives of $m_{nc,\theta}(X)$ with respect to σ and γ (computed with the help of Mathematica), shows that they are continuous on the compact set $(\theta, x) \in K \times [0, 1]$, and hence are uniformly bounded. Thus, estimating 1_{nc} by 1, we obtain that

$$|1_{nc}(x)m_{nc,\theta_1}(x) - 1_{nc}(x)m_{nc,\theta_2}(x))| \leq C(|\sigma_1 - \sigma_2| + |\gamma_1 - \gamma_2|),$$

where C denotes a generic constant which does not depend on θ or x. The corresponding inequalities for the rc, lc, and nc cases are obtained in precisely the same way, and, using Cauchy's inequality, it follows that $|m_{\theta_1}(x) - m_{\theta_2}(x))| \leq C\|\theta_1 - \theta_2\|$, with $\|\cdot\|$ denoting Euclidean norm, so that the first condition in Theorem 5.23 of van der Vaart (1998) is satisfied.

Further, write E for expectation taken with respect to the true parameter value $\theta_0 = (\sigma_0, \gamma_0)$. We next use numerical computation to show that the symmetric 2×2 matrix $\frac{\partial^2}{\partial\theta^2} E\left(m_{0,\theta}(X)\right)$, evaluated at θ_0 is strictly negative definite, by showing that the determinant is strictly positive and that the diagonal elements are negative. Inspection of all the second order derivatives show they also are continuous on the compact set $(\theta, x) \in K \times [0,1]$, and hence are uniformly bounded. It follows that the operations of differentiation and taking expectations may be freely interchanged. Thus we start by showing that the determinant

$$D_0(\sigma_0, \sigma, \gamma_0, \gamma) := E\left(\frac{\partial^2}{\partial\sigma^2} m_{0,\theta}(X)\right) E\left(\frac{\partial^2}{\partial\gamma^2} m_{0,\theta}(X)\right) - \left(E\left(\frac{\partial}{\partial\sigma}\frac{\partial}{\partial\gamma} m_{0,\theta}(X)\right)\right)^2 \quad (1)$$

evaluated at $\theta = \theta_0$ is strictly positive. Recall the notation $p_0 = E\left(1_{lc}(X) + 1_{dc}(X)\right)$ and $p_1 = 1 - p_0$ for the probabilities that $X(0) = 0$ and $X(0) = 1$, respectively. Considering the first factor in the righthand side of (1), we have that

$$\begin{aligned} E\left(\frac{\partial^2}{\partial\sigma^2} m_{0,\theta}(X)\right) &= E\left(1_{lc}(X)\frac{\partial^2}{\partial\sigma^2} m_{lc,\theta}(X) + 1_{dc}(X)\frac{\partial^2}{\partial\sigma^2} m_{dc,\theta}\right) \\ &= p_0\left(\int_0^1 \frac{\partial^2}{\partial\sigma^2} m_{lc,\theta}(x) g_0^r(x;\sigma_0,\gamma_0)dx \right. \qquad (2) \\ &\qquad \left. + \bar{G}_0^r(1;\sigma_0,\gamma_0)\frac{\partial^2}{\partial\sigma^2} m_{dc,\theta}\right). \end{aligned}$$

Here $g_0^r(x)$ is given by (5) , and $\frac{\partial^2}{\partial\sigma^2} m_{lc,\theta}(x)$ and $\frac{\partial^2}{\partial\sigma^2} m_{dc,\theta}$ are straightforwardly obtained from symbolic Mathematica differentiation. Setting $\sigma = \sigma_0, \gamma = \gamma_0$, and inserting numerical values for σ_0, γ_0, the integral in the second line of (2) may be obtained by numerical computation, and hence the entire expression in the righthand side may be computed numerically. Similarly the other three expressions in (1) may be computed numerically, and thus also $D_0(\sigma_0, \sigma_0, \gamma_0, \gamma_0)$. Using Mathematica to plot $D_0(\sigma_0, \sigma_0, \gamma_0, \gamma_0)$ as a function of σ_0 and γ_0 for $(\sigma_0, \gamma_0) \in [\sigma_\ell, \sigma_u] \times [\gamma_\ell, \gamma_u]$, shows that it is strictly positive, see online supplement described in Section **??**. These computations also show that the diagonal elements of $\frac{\partial^2}{\partial\theta^2} E\left(m_{0,\theta}(X)\right)$ are negative. Hence $\frac{\partial^2}{\partial\theta^2} E\left(m_{0,\theta}(X)\right)$ is strictly negative definite for $\theta = \theta_0 \in K$.

Next, writing E_S for conditional expectation given S, we as above have that

$$\frac{\partial^2}{\partial\theta^2} E\left(m_{1,\theta}(X)\right) = E\left(\frac{\partial^2}{\partial\theta^2} m_{1,\theta}(X)\right) = E\left(E_S\left(\frac{\partial^2}{\partial\theta^2} m_{1,\theta}(X)\right)\right).$$

Numerical computation of the determinant and diagonal elements of $E_S\left(\frac{\partial^2}{\partial\theta^2} m_{1,\theta}(X)\right)$ similarly showed that $E_S\left(\frac{\partial^2}{\partial\theta^2} m_{1,\theta}(X)\right)$ is non-negative definite for all $(\sigma_0, \gamma_0, S) \in [\sigma_\ell, \sigma_u] \times [\gamma_\ell, \gamma_u] \times [0,1]$. For these computations we additionally used the following observations.

Let $D_1(\sigma_0, \gamma_0|S)$ denote the determinant of $E_S\left(\frac{\partial^2}{\partial\theta^2} m_{1,\theta}(X)\right)$ for $\sigma = \sigma_0, \gamma = \gamma_0$. Then, $D_1(\sigma_0, \gamma_0|S) = D_1\left(\sigma_0/(1-S), \gamma_0|1\right)/(1-S)^2$ and hence computation over (σ_0, γ_0, S) can be transferred to a computation with $S = 0$ and $\sigma_u = 2.5/(1-S)$. We did these computations from $\sigma_\ell = 0.025$ up to $\sigma_u = 25$, corresponding to $S = 0.9$. Since $D_1(\sigma_0, \gamma_0|S) \to 0$ as $\sigma_0 \to \infty$ the values were negligible for $S \in (0.9, 0.1)$. The computations of the diagonal elements used the corresponding scaling properties of the diagonal elements.

Next, $\frac{\partial^2}{\partial\theta^2}E(m_{1,\theta}(X))$ is non-negative definite since integrals of non-negative definite matrices are non-negative definite. Finally, a sum of a strictly negative definite matrix and a non-negative definite matrix is strictly negative definite, and thus $\frac{\partial^2}{\partial\theta^2}E(m_\theta(X)) = \frac{\partial^2}{\partial\theta^2}E\left(m_{0,\theta}(X) + m_{1,\theta}(X)\right)$ evaluated at $\theta = \theta_0$ is strictly negative definite, and the second condition of Theorem 5.23 of van der Vaart (1998) holds.

To obtain last condition of Theorem 5.23 of van der Vaart (1998), consistency of $\hat{\theta}$, we use Theorem 5.9 of this reference. To check the first condition of that theorem, that $\frac{1}{n}\sum_{i=1}^{n} m_\theta(X_i)$ converges to its mean uniformly for $\theta \in K$ is standard, and is left to the reader. The second condition follows by similar, but simpler, arguments as above (and in fact, the local part, that the condition is satisfied in some surrounding of θ_0 is already established above).

As a final comment, since integration and differentiation may be freely exchanged, it is clear that one may use the observed information matrix, instead of the expected one, to measure estimation uncertainty.

Estimation algorithms

We provide MATLAB scripts that estimate the parameters of the distribution for exponential, gamma, and Weibull distributions, under full-parametric model, and for exponential and Generalized Pareto distributions, under semi-parametric model. One may also compute the corresponding 95% confidence intervals: except for exponential distribution in full-parametric case, where the explicit formula is available, confidence intervals are obtained using standard delta method where the inverse of the hessian matrix of the corresponding log-likelihood function is computed numerically. The scripts are located under *MATLAB→Estimators* folder.

Simulation study

MATLAB→Simulation Study folder contains MATLAB scripts that were used to produce Table 1, Table 2, and Table 3 of the Simulation Study. In each of the cases considered in the study, *SimulateGlances*.m* is used to simulate censored glances, and *RunSimulation*.m* produces 10,000 estimates of the parameters of the distribution from simulated data (for various choices of the original distribution parameters and sample sizes, i.e. number of simulated censored glances).

100-Car data analysis

MATLAB→Analysis of 100CarData folder contains MATLAB scripts that were used to produce Figure 3 and Figure 4. Similar figures can be produced for any other choice of the parametric or semi-parametric models considered in the paper.

Numerical verification of asymptotic normality

Wolfram Mathematica Scripts→Asymptotic Normalit.nb is a numerical study where we verify the non-negative definiteness of the expected information matrix.

References

van der Vaart, A. (1998). *Asymptotic Statistics.* Cambridge University Press, New York. 2, 4