Statistics 581, Problem Set 6 Solutions

Wellner; 11/15/2001

1. Suppose that $X_1, \ldots, X_n$ are i.i.d. with the Weibull distribution F_θ given by

$$1 - F_\theta(x) = \exp(-(x/\alpha)^\beta), \qquad x \ge 0$$

where $\theta = (\alpha, \beta) \in (0, \infty) \times (0, \infty)$.
(a) Find the inverse (or quantile function) $F_\theta^{-1}(u)$ corresponding to F_θ in terms of α, β, and $u \in (0,1)$, and show that

$$\log F_\theta^{-1}(u) = \log \alpha + \frac{1}{\beta} \log \log \left(\frac{1}{1-u} \right) .$$

(b) Fix $r \in (0, 1/2)$ and $s \in (1/2, 1)$ Use the $r-$th and $s-$th quantiles of the X_i's, namely $\mathbb{F}_n^{-1}(r)$ and $\mathbb{F}_n^{-1}(s)$, to obtain simple consistent estimators $\hat{\alpha}_n$ and $\hat{\beta}_n$ of α and β. Prove that your estimators are consistent.
(c) Prove that your estimators $\hat{\alpha}_n$ and $\hat{\beta}_n$ satisfy

$$\sqrt{n} \begin{pmatrix} \hat{\alpha}_n - \alpha \\ \hat{\beta}_n - \beta \end{pmatrix} \to_d N_2(0, \Sigma)$$

and identify Σ as a function of α, β, and t.
(d) How would you choose r and s to minimize the asymptotic variance of $\hat{\beta}_n$?

Solution: (a) Since $1 - F_\theta(x) = \exp(-(x/\alpha)^\beta)$, it follows we can solve $F_\theta(x) = u$ for $x = F_\theta^{-1}(u)$. This yields

$$F_\theta^{-1}(u) = \alpha(-\log(1-u))^{1/\beta} ,$$

or

$$\log F_\theta^{-1}(u) = \log \alpha + \frac{1}{\beta} \log \log \left(\frac{1}{1-u} \right) . \tag{0.1}$$

(b) Since we can estimate $F_\theta^{-1}(r)$ and $F_\theta^{-1}(s)$ respectively by $\mathbb{F}_n^{-1}(r)$ and $\mathbb{F}_n^{-1}(s)$ respectively, the relationship in (0.1) suggests that we estimate α and β as the solutions $\hat{\alpha}$ and $\hat{\beta}$ of the pair of equations

$$\log \mathbb{F}_n^{-1}(r) = \log \hat{\alpha} + \frac{1}{\hat{\beta}} \log \log 1/(1-r) , \tag{0.2}$$

$$\log \mathbb{F}_n^{-1}(s) = \log \hat{\alpha} + \frac{1}{\hat{\beta}} \log \log 1/(1-s) . \tag{0.3}$$

Letting $A_r \equiv \log\log 1/(1-r)$, and $B_s \equiv \log\log 1/(1-s)$, we find that

$$\begin{aligned} 1/\hat{\beta} &= \frac{1}{B_s - A_r} (\log \mathbb{F}_n^{-1}(s) - \log \mathbb{F}_n^{-1}(r)) \\ &\equiv a_{r,s} \log \mathbb{F}_n^{-1}(s) - a_{r,s} \log \mathbb{F}_n^{-1}(r) \end{aligned}$$

and

$$\log \hat{\alpha} = \frac{-A_r}{B_s - A_r} \log \mathbb{F}_n^{-1}(s) + \frac{B_s}{B_s - A_r} \log \mathbb{F}_n^{-1}(r))$$
$$\equiv c_{r,s} \log \mathbb{F}_n^{-1}(s) + d_{r,s} \log \mathbb{F}_n^{-1}(r)$$

where

$$a_{r,s} \equiv \frac{1}{B_s - A_r}, \qquad c_{r,s} \equiv -A_r a_{r,s} \qquad d_{r,s} \equiv B_s a_{r,s}\,.$$

Since $(\mathbb{F}_n^{-1}(r), \mathbb{F}_n^{-1}(s)) \rightarrow_{a.s.} (F_\theta^{-1}(r), F_\theta^{-1}(s))$, It follows easily by the continuous mapping theorem that

$$\frac{1}{\hat{\beta}} \rightarrow_{a.s.} a_{r,s} \log F_\theta^{-1}(s) - a_{r,s} \log F_\theta^{-1}(r) = \frac{1}{\beta}\,,$$

and

$$\log \hat{\alpha} \rightarrow_{a.s.} c_{r,s} \log F_\theta^{-1}(s) + d_{r,s} \log F_\theta^{-1}(r) = \log \alpha\,,$$

and hence by the continuous mapping theorem, $(\hat{\alpha}, \hat{\beta}) \rightarrow_{a.s.} (\alpha, \beta)$.
(c) First, we know that

$$\sqrt{n} \begin{pmatrix} \mathbb{F}_n^{-1}(r) - F^{-1}(r) \\ \mathbb{F}_n^{-1}(s) - F^{-1}(s) \end{pmatrix} \rightarrow_d \underline{Z} \sim N_2(0, \Sigma)$$

where

$$\Sigma = \begin{pmatrix} \frac{r(1-r)}{f^2(F^{-1}(r))} & \frac{r(1-s)}{f(F^{-1}(r))f(F^{-1}(s))} \\ \frac{r(1-s)}{f(F^{-1}(r))f(F^{-1}(s))} & \frac{s(1-s)}{f^2(F^{-1}(s))} \end{pmatrix}.$$

This implies that

$$\sqrt{n} \begin{pmatrix} \log \mathbb{F}_n^{-1}(r) - \log F^{-1}(r) \\ \log \mathbb{F}_n^{-1}(s) - \log F^{-1}(s) \end{pmatrix} \rightarrow_d D\underline{Z} \sim N_2(0, D\Sigma D^T)$$

where

$$D = \begin{pmatrix} 1/F^{-1}(r) & 0 \\ 0 & 1/F^{-1}(s) \end{pmatrix}.$$

Hence it follows that

$$\sqrt{n} \begin{pmatrix} 1/\hat{\beta} - 1/\beta \\ \log \hat{\alpha} - \log \alpha \end{pmatrix}$$
$$= M\sqrt{n} \begin{pmatrix} \log \mathbb{F}_n^{-1}(r) - \log F^{-1}(r) \\ \log \mathbb{F}_n^{-1}(s) - \log F^{-1}(s) \end{pmatrix}$$
$$\rightarrow_d M D \underline{Z} \sim N_2(0, M D \Sigma D^T M^T)\,.$$

where

$$M = \begin{pmatrix} -a_{r,s} & a_{r,s} \\ d_{r,s} & c_{r,s} \end{pmatrix} = a_{r,s} \begin{pmatrix} -1 & 1 \\ B_s & -A_r \end{pmatrix}.$$

Finally, with $g(x,y) = (g_1(x), g_2(y))$, $g_1(x) = 1/x$, $g_2(y) = \exp y$, we find, by the delta-method, that

$$\sqrt{n}\begin{pmatrix} \hat{\beta}-\beta \\ \hat{\alpha}-\alpha \end{pmatrix}$$
$$\to_d \quad \nabla g M D \underline{Z} \sim N_2(0, \nabla g M D \Sigma D^T M^T \nabla g^T)$$

where

$$\nabla g = \begin{pmatrix} -\beta^2 & 0 \\ 0 & \alpha \end{pmatrix}.$$

We begin combining all this by noting that $D\Sigma D^T$ involves the function

$$\begin{aligned} F^{-1}(u) f(F^{-1}(u)) &= \alpha\left(\log\left(\frac{1}{1-u}\right)\right)^{1/\beta} \frac{\beta}{\alpha}\left(\log\left(\frac{1}{1-u}\right)\right)^{(\beta-1)/\beta}(1-u) \\ &= \beta(1-u)\log\left(\frac{1}{1-u}\right) \equiv \beta g(u) \end{aligned}$$

at the points $u = r$ and $u = s$. Computing $D\Sigma D^T$ yields

$$D\Sigma D^T = \beta^{-2}\begin{pmatrix} \frac{r(1-r)}{g(r)^2} & \frac{r(1-s)}{g(r)g(s)} \\ \frac{r(1-s)}{g(r)g(s)} & \frac{s(1-s)}{g(s)^2} \end{pmatrix} \equiv \beta^{-2}\begin{pmatrix} c_{11}(r,s) & c_{12}(r,s) \\ c_{12}(r,s) & c_{22}(r,s) \end{pmatrix}.$$

Since the matrix M just depends on r, s, we find that the matrix

$$M D\Sigma D^T M^T = \beta^{-2} a_{r,s}^2 \begin{pmatrix} d_{11}(r,s) & d_{12}(r,s) \\ d_{12}(r,s) & d_{22}(r,s) \end{pmatrix},$$

where

$$\begin{aligned} d_{11}(r,s) &= c_{11}(r,s) - 2c_{12}(r,s) + c_{22}(r,s) \\ d_{12}(r,s) &= B_s(c_{12}(r,s) - c_{11}(r,s)) - A_r(c_{22}(r,s) - c_{12}(r,s)) \\ d_{22}(r,s) &= A_r^2 c_{22}(r,s) - 2A_r B_s c_{12}(r,s) + B_s^2 c_{11}(r,s). \end{aligned}$$

Thus we conclude that the asymptotic covariance matrix of $(\hat{\beta}, \hat{\alpha})$ is given by

$$\nabla g M D\Sigma D^T M^T \nabla g^T = a_{r,s}^2 \begin{pmatrix} \beta^2 d_{11}(r,s) & -\alpha d_{12}(r,s) \\ -\alpha d_{12}(r,s) & (\alpha/\beta)^2 d_{22}(r,s) \end{pmatrix}.$$

(d) The asymptotic variance of $\hat{\beta}$ is

$$\beta^2 a_{r,s}^2 d_{11}(r,s) = \beta^2\left(c_{11}(r,s) - 2c_{12}(r,s) + c_{22}(r,s)\right) a_{r,s}^2.$$

This is minimized by $r = r_0 \approx .1704$, $s = s_0 \approx .97$, and the minimum value is $\beta^2(.917) > \beta^2(6/\pi^2)$. This ad-hoc estimator $\hat{\beta}$ based on quantiles is *inefficient*; its asymptotic variance (for any value of r, s, including the minimizing r_0, s_0) is larger than the best possible asymptotic variance, which is $\beta^2(6/\pi^2)$ as we will see in Chapter 3. In fact the ARE when $r = r_0$ and $s = s_0$ is $(6/\pi^2)/.917 = .663$)

The asymptotic variance of $\hat{\alpha}$ is

$$(\alpha/\beta)^2 a_{r,s}^2 d_{22}(r,s) = (\alpha/\beta)^2\left(B_s^2 c_{11}(r,s) - 2A_r B_s c_{12}(r,s) + B_s^2 c_{22}(r,s)\right).$$

This is minimized by $r = r_0 \approx .398$, $s = s_0 \approx .82$, and the minimum value is $(\alpha/\beta)^2(1.359) > (\alpha/\beta)^2(1.11)$. This ad-hoc estimator $\hat{\alpha}$ based on quantiles is also *inefficient*; its asymptotic variance (for any value of r, s, including the minimizing r_0, s_0) is larger than the best possible asymptotic variance, which is about $(\alpha/\beta)^2(1.11)$ as we will see in Chapter 3. For the estimator based on the optimal r_0, s_0 (for α!), the ARE is $\approx 1.109/1.359 = .816$

2. Suppose that $X, X_1, X_2, \ldots, X_n$ are independent Poisson(λ) random variables:

$$P(X = k) \equiv p_k = e^{-\lambda}\frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \ldots.$$

Note that

$$\frac{p_k}{p_{k-1}} = \frac{\lambda}{k},$$

and hence whole family of alternative estimators $\{\tilde{\lambda}_n^{(k)}\}_{k\ge 1}$ is given by

$$\tilde{\lambda}_n^{(k)} = k\frac{\hat{p}_n(k)}{\hat{p}_n(k-1)}$$

where $\hat{p}_n(k) \equiv n^{-1}\sum_{i=1}^n 1_{[X_i = k]}$.
(a) Show that $\tilde{\lambda}_n \to_p \lambda$ for each $k = 1, 2, \ldots$.
(b) Show that

$$\sqrt{n}(\tilde{\lambda}_n^{(k)} - \lambda) \to_d N(0, \sigma_k^2(\lambda)) \quad \text{as} \quad n \to \infty$$

and compute $\sigma_k^2(\lambda)$ explicitly as a function of k and λ.
(c) What is the asymptotic relative efficiency of $\tilde{\lambda}_n^{(k)}$ to $\widehat{\lambda}_n = \overline{X}_n$ for $k > 1$? (The ARE of $\tilde{\lambda}_n^{(1)}$ with respect to $\widehat{\lambda}_n$ was computed in the Midterm Exam Solutions.)

Solution: (a) First, $(\hat{p}_n(k-1), \hat{p}_n(k)) \to_p (p_k, p_{k-1})$ by the WLLN. The function $g: R^2 \to R$ defined by $g(u, v) = v/u$ is continuous at (u, v) with $u \neq 0$. Hence it follows from the Mann-Wald (or continuous mapping) theorem that

$$g(\hat{p}_n(k-1), \hat{p}_n(k)) \to_p g(p_{k-1}, p_k) = \frac{\lambda^k e^{-\lambda}/k!}{\lambda^{k-1}e^{-\lambda}/(k-1)!} = \frac{\lambda}{k},$$

and hence

$$\tilde{\lambda}_n = kg(\hat{p}_n(k-1), \hat{p}_n(k)) \to_p kg(p_{k-1}, p_k) = \lambda\,.$$

(b) Note that $g(u, v)$ is differentiable at (u, v) with $u \neq 0$ and $\nabla g(u, v) = (-v/u, 1)/u$. Hence $\nabla g(p_{k-1}, p_k) = (-\lambda/k, 1)/p_{k-1}$. From the Multivariate CLT we have

$$\sqrt{n}\begin{pmatrix} \hat{p}_n(k-1) - p_{k-1} \\ \hat{p}_n(k) - p_k \end{pmatrix} \to_d N_2(0, \Sigma)$$

where

$$\Sigma = \begin{pmatrix} p_{k-1}(1 - p_{k-1} & -p_{k-1}p_k \\ -p_{k-1}p_k & p_k(1 - p_k) \end{pmatrix}.$$

Hence it follows from the delta method (or g'-theorem) that

$$\sqrt{n}(g(\hat{p}_n(k-1), \hat{p}_n(k)) - g(p_{k-1}, p_k)) \to_d N(0, \nabla g \Sigma (\nabla g)^T) = N(0, A^2)$$

and we can easily calculate

$$A^2 \equiv A^2(k,\lambda) = \nabla g \Sigma (\nabla g)^T = \frac{\lambda}{k}\left(1+\frac{\lambda}{k}\right)\frac{1}{p_{k-1}}\,.$$

Thus it follows that

$$\begin{aligned}\sqrt{n}(\tilde{\lambda}_n - \lambda) &= k\sqrt{n}(g(\hat{p}_n(k-1),\hat{p}_n(k)) - g(p_{k-1},p_k)) \\ &\to kN(0,A^2) = N(0,k^2A^2)\,.\end{aligned}$$

(c) It follows from (b) that the ARE of $\tilde{\lambda}_n^{(k)}$ with respect to $\widehat{\lambda}_n = \overline{X}_n$ is

$$ARE(\tilde{\lambda}_n^{(k)},\widehat{\lambda}_n) = \frac{\lambda}{k\lambda(1+\lambda/k)/p_{k-1}} = \frac{p_{k-1}}{\lambda+k}\,.$$

See Figure 1 below for a set of plots showing that these estimators get worse as k grows; $k=1$ is the only one which has an ARE anywhere close to 1.

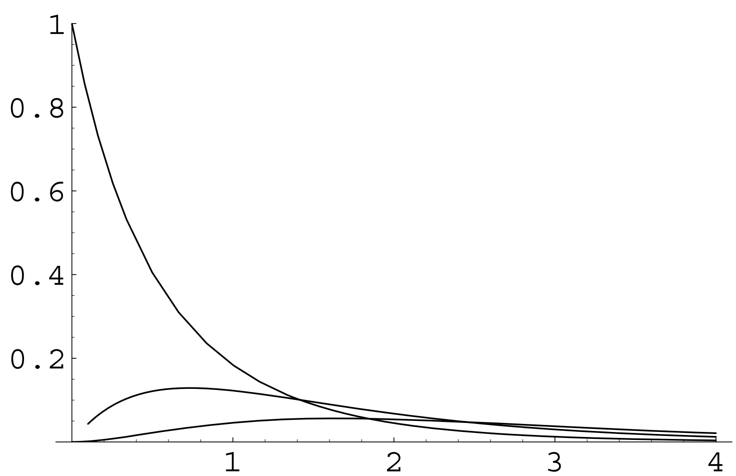

Figure 1: ARE for $k = 1, 2, 3$.

3. Suppose that Suppose that $X_1,\ldots,X_n,\ldots$ are i.i.d. random vectors in R^k with common distribution function F and corresponding probability measure P on $(R^k,\mathcal{B}_k)$. Let $\mathbb{P}_n$ be the empirical measure defined by

$$\mathbb{P}_n = n^{-1}\sum_{i=1}^{n}\delta_{X_i}\,,$$

and consider $\mathbb{P}_n$ and the empirical process $\mathbb{G}_n$ as indexed by a class of sets $\mathcal{C} \subset \mathcal{B}_k$:

$$\{\mathbb{P}_n(C):\ C\in\mathcal{C}\}, \qquad \{\mathbb{G}_n(C):\ C\in\mathcal{C}\}\,,$$

where

$$\mathbb{G}_n \equiv \sqrt{n}(\mathbb{P}_n - P)\,.$$

(a) Show that $\mathbb{G}_n \to_{f.d.} \mathbb{G}_P$ where $\mathbb{G}_P$ is a P-Brownian bridge process indexed by $\mathcal{C}$: i.e. show that for any integer m and sets $C_1, \ldots, C_m \in \mathcal{C}$,

$$(\mathbb{G}_n(C_1), \ldots, \mathbb{G}_n(C_m)) \to_d (\mathbb{G}_P(C_1), \ldots, \mathbb{G}_P(C_m)) \sim N_m(0, \Sigma)$$

where $\Sigma = (\sigma_{jj'})$ is given by

$$\sigma_{jj'} = P(C_j \cap C_{j'}) - P(C_j)P(C_{j'})\,.$$

(b) When $\mathcal{C} = \mathcal{O} \equiv \{(-\infty, x] :\ x \in R^k\}$ specialize the result in (a) and show that it gives the finite-dimensional convergence of the empirical distribution function $\mathbb{F}_n$: i.e.
(i) show that $\mathbb{P}_n((-\infty, x]) = \mathbb{F}_n(x)$;
(ii) show that $P((-\infty, x]) = F(x)$;
(iii) show that $\mathbb{Y}(x) \equiv \mathbb{G}_P((-\infty, x])$ has mean zero and covariance

$$E\{\mathbb{Y}(x)\mathbb{Y}(y)\} = F(x \wedge y) - F(x)F(y), \qquad x, y \in R^k\,.$$

Solution: (a) This follows directly from the multivariate central limit theorem: for any integer m and sets $C_1, \ldots, C_m \in \mathcal{C}$,

$$\begin{pmatrix} \mathbb{G}_n(C_1) \\ \cdot \\ \cdot \\ \cdot \\ \mathbb{G}_n(C_m) \end{pmatrix} = \frac{1}{\sqrt{n}} \sum_{i=1}^n \begin{pmatrix} 1_{C_1}(X_i) - P(C_1) \\ \cdot \\ \cdot \\ \cdot \\ 1_{C_m}(X_i) - P(C_m) \end{pmatrix} \equiv \sqrt{n}\,\underline{\overline{Y}}_n$$

where $\underline{Y}_1, \ldots, \underline{Y}_n$ are i.i.d. with $E(\underline{Y}_i) = 0$,

$$E(\underline{Y}_i \underline{Y}_i') = (P(C_j \cap C_{j'}) - P(C_j)P(C_{j'})) \equiv \Sigma\,.$$

Hence by the multivariate central limit theorem

$$\begin{pmatrix} \mathbb{G}_n(C_1) \\ \cdot \\ \cdot \\ \cdot \\ \mathbb{G}_n(C_m) \end{pmatrix} \to_d \begin{pmatrix} \mathbb{G}_P(C_1) \\ \cdot \\ \cdot \\ \cdot \\ \mathbb{G}_P(C_m) \end{pmatrix} \sim N_m(0, \Sigma)\,. \tag{0.4}$$

In other words, $\mathbb{G}_n \to_{f.d.} \mathbb{G}_P$ as indexed by the class $\mathcal{C}$.
(b) When $\mathcal{C} = \mathcal{O} \equiv \{(-\infty, x] :\ x \in R^k\}$, we compute
(i) $\mathbb{P}_n((-\infty, x]) = n^{-1} \sum_{i=1}^n 1_{(-\infty, x]}(X_i) = \mathbb{F}_n(x)$;
(ii) $P((-\infty, x]) = P(X \le x) = F(x)$;
(iii) and we have, with $\mathbb{Y}(x) \equiv \mathbb{G}_P((-\infty, x])$,

$$\begin{aligned} E\{\mathbb{Y}(x)\mathbb{Y}(y)\} &= E\{\mathbb{G}_P((-\infty, x])\mathbb{G}_P((-\infty, y]) \\ &= P((-\infty, x] \cap (-\infty, y]) - P((-\infty, x])P((-\infty, y]) \\ &= P((-\infty, x \wedge y]) - P((-\infty, x])P((-\infty, y]) \\ &= F(x \wedge y) - F(x)F(y) \quad \text{by three applications of } (ii). \end{aligned}$$