Statistics 581
Problem Set 5 Solutions
Wellner; 10/31/2018

1. van der Vaart, problem 3.8, page 34, modified. Let $X_1, \ldots, X_n$ be i.i.d. Bernoulli(p) with $0 < p < 1$.
(a) Find the limit distribution of $\sqrt{n}(\overline{X}_n^{-1} - p^{-1})$.
(b) Show that $E|\overline{X}_n^{-1}| = \infty$ for every n.
(c) Connect the example in (a) to a result in the 581 Course Notes, Section 2.4.

Solution: (a) By the Lindeberg CLT it follows easily that $\sqrt{n}(\overline{X}_n - p) \to_d Z \sim N(0, p(1-p))$. Furthermore, $g(y) = y^{-1}$ is differentiable at $p > 0$ with derivative $g'(p) = -p^{-2}$. It then follows from the delta-method that

$$\sqrt{n}(g(\overline{X}_n) - g(p)) \to_d g'(p)Z \sim N(0, g'(p)^2 p(1-p)) = N(0, (1-p)/p^3).$$

(b) On the other hand, since $P_p(n\overline{X}_n = 0) = P_p(\text{Bin}(n,p) = 0) = (1-p)^n > 0$, and hence $E_p\{\overline{X}_n^{-1}\} \ge (n/0) \cdot (1-p)^n = \infty$.
(c) Letting $Y_n \equiv \sqrt{n}(\overline{X}_n^{-1} - p^{-1})$ we have $Y_n \to_d Y_0 \sim N(0, (1-p)/p^3)$ while from (b)

$$E|Y_n| \ge E|Y_n^+| \ge \sqrt{n}E(\overline{X}_n^{-1} - p) = \infty,$$

so we have $0 < E|Y_0| = \sqrt{(1-p)/p^3}E|N(0,1)| < \liminf E|Y_n| = \infty$. Thus strict inequality can occur in Proposition 2.4.6 of the Chapter 2 notes, page 25.

2. van der Vaart, problem 3.6, page 34: Let $X_1, \ldots, X_n$ be i.i.d. with expectation μ and variance 1. Find constants a_n and $b_n$such that $a_n(\overline{X}_n^2 - b_n)$ converges in distribution when $\mu = 0$ or $\mu \ne 0$.

Solution: When $\mu = 0$, we can take $b_n = 0$ for all n and $a_n = n$. Then with $Z \sim N(0,1)$,

$$n(\overline{X}_n^2 - 0) = \{\sqrt{n}\overline{X}_n\}^2 = \{\sqrt{n}(\overline{X}_n - 0)\}^2 \to_d Z^2 = \chi_1^2$$

by the Lindeberg (ordinary) CLT and the continuous mapping theorem. When $\mu \ne 0$, then we can take $a_n = \sqrt{n}$ and $b_n = \mu^2$: then we have

$$\sqrt{n}(\overline{X}_n^2 - \mu^2) \to_d 2\mu Z \sim N(0, 4\mu^2)$$

by the Lindeberg CLT (again) followed by the delta-method.

3. van der Vaart, problem 19.4, page 290: Suppose that $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$ are independent samples from distribution functions F and G respectively. The Kolmogorov-Smirnov statistic for testing the null hypothesis $H: \ F = G$ versus $K: \ F \ne G$ is the supremum distance $K_{m,n} \equiv \|\mathbb{F}_m - \mathbb{G}_n\|_\infty$ between the empirical distributions of the two samples.
(a) Find the limiting distribution of $\sqrt{mn/N}K_{m,n}$ under the null hypothesis. Do this first assuming that $\lambda_N \equiv m/N \equiv m/(m+n) \to \lambda \in [0,1]$ as $m \wedge n \to \infty$. What

can you say if the latter hypothesis is dropped?
(b) Show that the Kolmogorov - Smirnov test is asymptotically consistent against every alternative $F \neq G$.
(c) Find the asymptotic power function as a function of (Δ_F, Δ_G) for alternatives (F_m, G_n) where $\{F_m\}$ and $\{G_n\}$ satisfy, much as in our discussion in class on 26 October, $\|F_m - F_0\}_\infty \to 0$, $\|G_n - F_0\|_\infty \to 0$ and, for functions $\Delta_F, \Delta_G : [0,1] \to \mathbb{R}$, $\|\sqrt{m}(F_m - F_0) - \Delta_F(F_0)\|_\infty \to 0$ and $\|\sqrt{n}(G_n - F_0) - \Delta_G(F_0)\|_\infty \to 0$.

Solution: (a) If we assume that $\lambda_N \equiv m/N \to \lambda \in [0,1]$, then we have

$$\begin{aligned}
\sqrt{\frac{mn}{N}} K_{m,n} &= \sqrt{\frac{mn}{N}} \|\mathbb{F}_m - F - (\mathbb{G}_n - F)\|_\infty \\
&= \|\sqrt{\frac{n}{N}} \sqrt{m}(\mathbb{F}_m - F) - \sqrt{\frac{m}{N}} \sqrt{n}(\mathbb{G}_n - F)\|_\infty \\
&\stackrel{d}{=} \|\sqrt{\frac{n}{N}} \mathbb{U}_m^X(F) - \sqrt{\frac{n}{N}} \mathbb{U}_n^Y(F)\|_\infty \\
&\to_d \|\sqrt{\overline{\lambda}} \mathbb{U}^X(F) - \sqrt{\lambda} \mathbb{U}^Y(F)\|_\infty \\
&\stackrel{d}{=} \|\mathbb{U}\|_\infty
\end{aligned}$$

where $\mathbb{U}^X$ and $\mathbb{U}^Y$ are independent Brownian bridge processes on $[0,1]$ and since the process $\mathbb{U} \equiv \sqrt{\overline{\lambda}}\mathbb{U}^X - \sqrt{\lambda}\mathbb{U}^Y$ is again a Brownian bridge process: note that it is clearly Gaussian and it has

$$\begin{aligned}
E\mathbb{U}(t) &= \sqrt{\overline{\lambda}}\mathbb{U}^X(t) - \sqrt{\lambda}\mathbb{U}^Y(t) = 0 - 0 = 0, \quad \text{and} \\
E\mathbb{U}(s)\mathbb{U}(t) &= \overline{\lambda} E\mathbb{U}^X(s)\mathbb{U}^X(t) + \lambda E\mathbb{U}^Y(s)\mathbb{U}^Y(t) \\
&= \overline{\lambda}(s \wedge t - st) + \lambda(s \wedge t - st) = s \wedge t - st.
\end{aligned}$$

Thus under the null hypothesis the limiting distribution of the two sample statistic is just $\|\mathbb{U}\|_\infty$, the same limiting distribution as for the one-sample K-S statistic as in Example 2.5.1 in Chapter 2 of the course notes. If $\lambda_N = m/N$ does not converge to a fixed $\lambda \in [0,1]$ as $m \wedge n \to \infty$, then since λ_N takes values in the compact set $[0,1]$, starting with an arbitrary subsequence $\{\lambda_{N'}\}$ we can always find a further subsequence $\{\lambda_{N''}\}$ such that $\lambda_{N''}$ converges to some $\lambda \in [0,1]$. Then the preceding argument shows that along this subsequence we have $\sqrt{mn/N} K_{m,n} \to_d \|\mathbb{U}\|_\infty$. Since this limit is the same for any initial subsequence $\{N'\}$, we conclude that the convergence holds for the full sequence $\sqrt{mn/N} K_{m,n}$ and that the limit distribution is just $\|\mathbb{U}\|$ for all $m \wedge n \to \infty$.
(b) When the alternative hypothesis holds, i.e. $F \neq G$, then the Glivenko-Cantelli theorem implies that

$$\|\mathbb{F}_m - F\|_\infty \to_{a.s.} 0 \quad \text{and} \quad \|\mathbb{G}_n - G\|_\infty \to_{a.s.} 0.$$

Then we have

$$\|\mathbb{F}_m - \mathbb{G}_n\|_\infty \to_{a.s.} \|F - G\|_\infty > 0.$$

Thus we can write

$$\begin{aligned}
\sqrt{\frac{mn}{N}} K_{m,n} &= \sqrt{\frac{mn}{N}} \|\mathbb{F}_m - \mathbb{G}_n\|_\infty \\
&= \|\sqrt{\overline{\lambda}_N}\sqrt{m}(\mathbb{F}_m - F) - \sqrt{\lambda_N}\sqrt{n}(\mathbb{G}_n - G) + \sqrt{\frac{mn}{N}}(F-G)\|_\infty \\
&\geq \sqrt{\frac{mn}{N}} \Big\|(F-G)\Big\|_\infty - \sqrt{\overline{\lambda}_N}\Big\|\sqrt{m}(\mathbb{F}_m - F)\Big\|_\infty - \sqrt{\lambda_N}\|\sqrt{n}(\mathbb{G}_n - G)\|_\infty \\
&\qquad \text{by the triangle inequality} \\
&= \sqrt{N\lambda_N \cdot \overline{\lambda}_N}(F-G)\|_\infty - O_p(1) \\
&\to_p \infty \quad \text{if} \quad m \wedge n \to \infty \\
&\qquad \text{and either} \quad \limsup_N \lambda_N < 1 \quad \text{or} \quad \liminf_N \ \lambda_N > 0.
\end{aligned}$$

Thus when $F \neq G$ and either $\limsup_N \lambda_N < 1$ or $\liminf_N \ \lambda_N > 0$ we have

$$P_{F,G}\left(\sqrt{\frac{mn}{N}} K_{m,n} > \lambda_\alpha\right) \to 1.$$

(c) Under local alternatives $\{F_m\}$ and $\{G_n\}$ satisfying the hypotheses of the problem statement and assuming that $\lambda_N \to \lambda$, we have, by an argument similar to that of (a),

$$\begin{aligned}
\sqrt{\frac{mn}{N}}(\mathbb{F}_m - \mathbb{G}_n) &= \sqrt{\overline{\lambda}_N}\left(\sqrt{m}(\mathbb{F}_m - F_m) + \sqrt{m}(F_m - F_0)\right) \\
&\quad - \sqrt{\lambda_N}\left(\sqrt{n}(\mathbb{G}_n - G_n) + \sqrt{n}(G_n - F_0)\right) \\
&\Rightarrow \sqrt{\overline{\lambda}}\{\mathbb{U}^X(F_0) + \Delta_X(F_0)\} - \sqrt{\lambda}\{\mathbb{U}^Y(F_0) + \Delta_Y(F_0)\} \\
&\stackrel{d}{=} \mathbb{U}(F_0) + \sqrt{\overline{\lambda}}\Delta_X(F_0) - \sqrt{\lambda}\Delta_Y(F_0),
\end{aligned}$$

Thus with $\Delta \equiv \sqrt{\overline{\lambda}}\Delta_X - \sqrt{\lambda}\Delta_Y$, the power of the two - sample K-S test under these local alternative satisfies

$$P_{F_n,G_n}\left(\sqrt{mn/N}\|\mathbb{F}_m - \mathbb{G}_n\|_\infty > \lambda_\alpha\right) \to P\left(\|\mathbb{U} + \Delta\|_\infty > \lambda_\alpha\right)$$

4. Suppose that $X_1, \ldots, X_n$ are i.i.d. Cauchy$(0,1)$; so the density of each X_i with respect to Lebesgue measure on R is $f(x) = \pi^{-1}(1+x^2)^{-1}$, $x \in R$.
(a) Compute the distribution function F of the X_i's.
(b) Compute and plot the inverse distribution function F^{-1} corresponding to F.
(c) For what values of $r > 0$ is $E|X_1|^r < \infty$?
(d) Find the distribution function of $M_n \equiv \max_{1 \leq i \leq n} X_i$.
(e) For what values of r is $E|M_n|^r < \infty$?
(f) Find a sequence of constants b_n so that $M_n/b_n \ \to_d$ and find the limiting distribution. [Hint: see Ferguson, ACLST, Theorem 14, page 95.]
(g) Find the densities of M_n/b_n with b_n as in (f). Do these densities converge pointwise to a limit density? If so, what can you conclude from Scheffé's theorem?

Solution: (a) $F(x) = (1/\pi)\int_{-\infty}^{x}(1+t^2)^{-1}dt = (1/\pi)\{\arctan(x) + \pi/2\}$.
(b) Setting $F(x) = u$ and solving for $x = F^{-1}(u)$ yields $F^{-1}(u) = \tan(\pi(u - 1/2))$. Note that $F^{-1}(1/2) = \tan(0) = 0$; $F^{-1}(1) = \tan(\pi/2) = \infty$, and $F^{-1}(0) = \tan(-\pi/2) = -\infty$.
(c) We compute

$$\begin{aligned} E|X_1|^r &= \frac{1}{\pi}\int_{-\infty}^{\infty}|x|^r\frac{1}{1+x^2}dx \\ &= \frac{2}{\pi}\left\{\int_0^1 \frac{x^r}{1+x^2}dx + \int_1^{\infty}\frac{x^r}{1+x^2}dx\right\} \\ &\le \frac{2}{\pi}\left\{\int_0^1 \frac{x^r}{1+x^2}dx + \int_1^{\infty}\frac{x^r}{x^2}dx\right\} \\ &= \frac{2}{\pi}\left\{\int_0^1 \frac{x^r}{1+x^2}dx + \frac{1}{1-r}\right\} < \infty \end{aligned}$$

if $r < 1$. Since

$$E|X_1| = \frac{2}{\pi}\int_0^{\infty}\frac{x}{1+x^2}dx = \infty\,,$$

$E|X_1|^r < \infty$ if and only if $r < 1$.
(d) Since the X_i's are i.i.d. with distribution function F,

$$F_{M_n}(x) = P(M_n \le x) = P(X_1 \le x, \ldots, X_n \le x) = F(x)^n.$$

(e) First, note that

$$1 - F_{|M_n|}(x) = P(|M_n| > x) = P(\cup_{i=1}^n[|X_i| > x]) \le \sum_{i=1}^n P(|X_i| > x) = n(1 - F_{|X_1|}(x))$$

where $F_{|X_1|}(x) = P(|X_1| \le x) = F(x) - F(-x)$. Hence

$$\begin{aligned} E|M_n|^r &= \int_0^{\infty} rt^{r-1}(1 - F_{|M_n|}(t))dt \\ &\le \int_0^{\infty} rt^{r-1}n(1 - F_{|X|}(t))dt \\ &= nE|X_1|^r < \infty \end{aligned}$$

if $r < 1$ by part (d). But since $E|M_n|^r \ge E|X_1|^r = \infty$ if $r \ge 1$, we conclude that $E|M_n|^r < \infty$ if and only if $r < 1$.
(f) Note that $1 - F(x) = \pi^{-1}\int_x^{\infty}(1+t^2)^{-1}dt \sim 1/(\pi x)$ in the sense that $x(1 - F(x)) \to 1/\pi$ as $x \to \infty$. [This follows easily by writing the left side as $(1 - F(x))/(x^{-1})$ and using L'Hopital's rule.] Hence for $b_n \to \infty$ and $x > 0$

$$F_{M_n/b_n}(x) = P(M_n \le xb_n) = F(xb_n)^n \qquad \text{by part d}$$

and, with $c_n \equiv xb_n(1 - F(xb_n)) \to 1/\pi$,

$$\begin{aligned} F_{M_n/b_n}(x) &= F(xb_n)^n = (1 - (1 - F(xb_n)))^n \\ &= (1 - [xb_n(1 - F(xb_n))]/(xb_n))^n \\ &= (1 - c_n/xb_n)^n. \end{aligned}$$

From this last expression it becomes clear that the choice $b_n = n$ yields,

$$F_{M_n/b_n}(x) \to \exp(-1/\pi x) \equiv G(x), \qquad \text{for } x > 0,$$

while for $x \le 0$

$$F_{M_n/b_n}(x) \to 0$$

since $F(xb_n) \le 1/2$ for $x \le 0$. Note that $G(0) = \exp(-\infty) = 0$, G is monotone increasing, and $G(\infty) = \exp(0) = 1$. In fact, G is a member of the Weibull family with shape parameter -1, and is one of the three different families that can arise as limit distributions of maxima of independent rv's; see e.g. Ferguson (1996), *A Course in Large Sample Theory*, page 95.
(g) The density of $F_{M_n/b_n} = F(xb_n)^n$ is given by

$$\begin{aligned} f_{M_n/b_n}(x) &= nF(xb_n)^{n-1} f(xb_n) b_n \\ &= \left(1 - \frac{c_n}{xb_n}\right)^{n-1} \frac{1}{\pi(1 + (xb_n)^2)} \cdot nb_n \\ &\to \exp(-1/(\pi x)) \frac{1}{\pi x^2} \equiv g(x) \quad \text{when} \quad b_n = n. \end{aligned}$$

Thus Scheffé's theorem yields

$$d_{TV}(P_{M_n/n}, P_G) = \frac{1}{2} \int_{-\infty}^{\infty} |f_{M_n/n}(x) - g(x)| dx \to 0$$

as $n \to \infty$. It would be interesting to know more about the rates of convergence in theorems of this type.

5. 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 $t \in (0, 1/2)$. Use the $t-$th and $(1-t)-$th quantiles of the X_i's, namely $\mathbb{F}_n^{-1}(t)$ and $\mathbb{F}_n^{-1}(1-t)$, 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 t 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

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

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

$$\log \mathbb{F}_n^{-1}(t) = \log\hat\alpha + \frac{1}{\hat\beta}\log\log 1/(1-t)\,, \tag{2}$$

$$\log \mathbb{F}_n^{-1}(1-t) = \log\hat\alpha + \frac{1}{\hat\beta}\log\log 1/t\,. \tag{3}$$

Letting $A_t \equiv \log\log 1/(1-t)$, and $B_t \equiv \log\log 1/t$, we find that

$$\begin{aligned} 1/\hat\beta &= \frac{1}{B_t - A_t}(\log\mathbb{F}_n^{-1}(1-t) - \log\mathbb{F}_n^{-1}(t)) \\ &\equiv a_t \log\mathbb{F}_n^{-1}(1-t) + b_t\log\mathbb{F}_n^{-1}(t) \end{aligned}$$

and

$$\begin{aligned} \log\hat\alpha &= \frac{-A_t}{B_t - A_t}\log\mathbb{F}_n^{-1}(1-t) + \frac{B_t}{B_t - A_t}\log\mathbb{F}_n^{-1}(t)) \\ &\equiv c_t\log\mathbb{F}_n^{-1}(t) + d_t\log\mathbb{F}_n^{-1}(1-t) \end{aligned}$$

where

$$a_t \equiv \frac{1}{B_t - A_t}, \qquad b_t = -a_t, \qquad c_t \equiv -A_t a_t \qquad d_t \equiv B_t a_t\,.$$

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

$$\frac{1}{\hat\beta} \to_{a.s.} a_t\log F_\theta^{-1}(1-t) + b_t\log F_\theta^{-1}(t) = \frac{1}{\beta}\,,$$

and

$$\log\hat\alpha \to_{a.s.} c_t\log F_\theta^{-1}(1-t) + d_t\log F_\theta^{-1}(t) = \log\alpha\,,$$

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

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

where

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

This implies that

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

where

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

Hence it follows that

$$\begin{aligned} & \sqrt{n}\begin{pmatrix} 1/\hat{\beta} - 1/\beta \\ \log\hat{\alpha} - \log\alpha \end{pmatrix} \\ = \;& M\sqrt{n}\begin{pmatrix} \log \mathbb{F}_n^{-1}(1-t) - \log F^{-1}(1-t) \\ \log \mathbb{F}_n^{-1}(t) - \log F^{-1}(t) \end{pmatrix} \\ \to_d \;& MD\underline{Z} \sim N_2(0, MD\Sigma D^T M^T). \end{aligned}$$

where

$$M = \begin{pmatrix} a_t & b_t \\ c_t & d_t \end{pmatrix} = a_t \begin{pmatrix} 1 & -1 \\ -A_t & B_t \end{pmatrix}.$$

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

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

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

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

at the points $u = t$ and $u = 1-t$. Computing $D\Sigma D^T$ yields

$$\begin{aligned} D\Sigma D^T &= \beta^{-2}\begin{pmatrix} \frac{1-t}{t(\log(1/t))^2} & \frac{t}{(1-t)\log(1/t)\log(1/(1-t))} \\ \frac{t}{(1-t)\log(1/t)\log(1/(1-t))} & \frac{t}{(1-t)(\log(1/(1-t)))^2} \end{pmatrix} \\ &\equiv \beta^{-2}\begin{pmatrix} s_{11}(t) & s_{12}(t) \\ s_{12}(t) & s_{22}(t) \end{pmatrix}. \end{aligned}$$

Since the matrix M just depends on t, we find that the matrix

$$MD\Sigma D^T M^T = \beta^{-2} a_t^2 \begin{pmatrix} r_{11}(t) & r_{12}(t) \\ r_{12}(t) & r_{22}(t) \end{pmatrix},$$

where

$$\begin{aligned} r_{11}(t) &= s_{11}(t) - 2s_{12}(t) + s_{22}(t) \\ r_{12}(t) &= B_t(s_{12}(t) - s_{22}(t)) - A_t(s_{11}(t) - s_{12}(t)) \\ r_{22}(t) &= A_t^2 s_{11}(t) - 2A_t B_t s_{12}(t) + B_t^2 s_{22}(t)\,. \end{aligned}$$

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

$$\nabla g M D \Sigma D^T M^T \nabla g^T = a_t^2 \begin{pmatrix} \beta^2 r_{11}(t) & \alpha r_{12}(t) \\ \alpha r_{12}(t) & (\alpha/\beta)^2 r_{22}(t) \end{pmatrix}.$$

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

$$\beta^2 a_t^2 r_{11}(t) = \beta^2 \left(s_{11}(t) - 2s_{12}(t) + s_{22}(t)\right) a_t^2\,.$$

This is minimized by $t = t_0 \approx .10725$, and the minimum value is $\beta^2(1.13264) > \beta^2(6/\pi^2)$ see Figures 1 and 2 below. This ad-hoc estimator $\hat{\beta}$ based on quantiles is *inefficient*; its asymptotic variance (for any value of t, including the minimizing t_0) is larger than the best possible asymptotic variance, which is $\beta^2(6/\pi^2)$ as we will see in Chapter 3.)

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

$$(\alpha/\beta)^2 a_t^2 r_{22}(t) = (\alpha/\beta)^2 \left(A_t^2 s_{11}(t) - 2A_t B_t s_{12}(t) + B_t^2 s_{22}(t)\right)\,.$$

This is minimized by $t = t_0 \approx .2295$, and the minimum value is $(\alpha/\beta)^2(1.423) > (\alpha/\beta)^2(1.11)$ see Figures 3 and 4 below. This ad-hoc estimator $\hat{\beta}$ based on quantiles is also *inefficient*; its asymptotic variance (for any value of t, including the minimizing t_0) is larger than the best possible asymptotic variance, which is about $(\alpha/\beta)^2(1.11)$ as we will see in Chapter 3.