Statistics 581, Problem Set 8 Solutions

Wellner; 11/17/2010

1. (a) Show that if $\theta_n = cn^{-1/2}$ and T_n is the Hodges super-efficient estimator discussed in class, then the sequence $\{\sqrt{n}(T_n - \theta_n)\}$ is uniformly square-integrable.
(b) Lehmann and Casella, Problem 2.13, page 501.
(c) Let $R_n(\theta) \equiv nE_\theta(T_n - \theta)^2$ where T_n is the Hodges superefficient estimator as in Example 3.3.1 (so $T_n = \delta_n$ of Example 2.5, Lehmann and Casella pages 440 - 443). Show that $R_n(n^{-1/4}) \to \infty$ as $n \to \infty$.

Solution: (a) Recall that when $\theta = \theta_n = cn^{-1/2}$ is true, we showed in class that

$$\sqrt{n}(T_n - \theta_n) \stackrel{d}{=} Z1_{[|Z+c|>n^{1/4}]} + \{aZ + c(a-1)\}1_{[|Z+c|\le n^{1/4}]} \equiv S_n$$

where $Z \sim N(0,1)$. Now by the C_r-inequality (or simply $(a+b)^2 \le 2(a^2+b^2)$),

$$|S_n|^2 \le 2\{Z^2 + (aZ + c(a-1))^2\} \equiv Y$$

where the right side is a fixed integrable random variable. Thus the desired interchange of limit and integration is valid by the dominated convergence theorem or by Vitali's theorem since

$$\limsup_n E\{|S_n|^2 1_{[|S_n|^2\ge\lambda]}\} \le E\{Y1_{[Y\ge\lambda]}\} \to 0$$

as $\lambda \to \infty$. Thus $\{S_n\}$ is uniformly square integrable.

(b) (a') First recall that (with $\delta_n = T_n$) since $\sqrt{n}(\overline{X} - \theta) \stackrel{d}{=} Z \sim N(0,1)$ we can write

$$\begin{aligned}
\sqrt{n}(T_n - \theta) &= \sqrt{n}(\overline{X}_n 1_{[|\overline{X}_n|>n^{-1/4}]} + a\overline{X}_n 1_{[|\overline{X}_n|\le n^{-1/4}]} - \theta) \\
&\stackrel{d}{=} Z1_{[|Z+\theta\sqrt{n}|>n^{1/4}]} + [aZ + \sqrt{n}\theta(a-1)]1_{[|Z+\theta\sqrt{n}|\le n^{1/4}]} \\
&= Z + [(a-1)Z + (a-1)\sqrt{n}\theta]1_{[|Z+\theta\sqrt{n}|\le n^{1/4}]} \\
&= Z - (1-a)[Z + \sqrt{n}\theta]1_{[|Z+\theta\sqrt{n}|\le n^{1/4}]}.
\end{aligned}$$

Thus

$$\begin{aligned}
b_n(\theta) &= E_\theta(T_n) - \theta \\
&= n^{-1/2}\left\{EZ - (1-a)E[Z+\sqrt{n}\theta]1_{[|Z+\theta\sqrt{n}|\le n^{1/4}]}\right\} \\
&= -\frac{1-a}{\sqrt{n}}E[Z+\sqrt{n}\theta]1_{[|Z+\theta\sqrt{n}|\le n^{1/4}]} \\
&= -\frac{1-a}{\sqrt{n}}\int_{-n^{1/4}}^{n^{1/4}} x\phi(x-\sqrt{n}\theta)dx
\end{aligned}$$

since $Z + \theta\sqrt{n} \sim N(\theta\sqrt{n}, 1)$.

(b') Differentiating the result in (a') gives

$$\begin{aligned}
b_n'(\theta) &= -\frac{1-a}{\sqrt{n}}\int_{-n^{1/4}}^{n^{1/4}} x\phi'(x-\sqrt{n}\theta)(-\sqrt{n})\,dx \\
&= -(1-a)\int_{-n^{1/4}}^{n^{1/4}} x(x-\sqrt{n}\theta)\phi(x-\sqrt{n}\theta)\,dx \qquad \text{since } \phi'(x) = -x\phi(x) \\
&\to 0 \qquad \text{if } \theta \ne 0
\end{aligned}$$

by the dominated convergence theorem since $x(x-\sqrt{n}\theta)\phi(x-\sqrt{n}\theta)1_{[-n^{1/4},n^{1/4}]}(x) \to 0$ for each fixed x and is dominated by the integrable function $4e^{-1}\phi(x)/(|\theta| \wedge 1)$ (for $n \geq (3/|\theta|)^4$).
Details of this domination: For $|x| \leq n^{1/4}$ it follows that

$$|x||x-\sqrt{n}\theta| \leq n^{1/4}|-n^{1/4}-\sqrt{n}\theta| \leq n^{1/2}+n^{3/4}|\theta| \leq 2n^{3/4}(|\theta| \vee 1)$$

while

$$\begin{aligned}
\phi(x-\sqrt{n}\theta) &= \phi(x)\exp(\sqrt{n}\theta x - n\theta^2/2) \\
&\leq \phi(x)\exp(|\theta|n^{3/4} - n\theta^2/2) \\
&= \phi(x)\exp(|\theta|n^{3/4}(1-n^{1/4}|\theta|/2)) \\
&\leq \phi(x)\exp(-\frac{1}{2}|\theta|n^{3/4}) \qquad \text{if } \ 1-n^{1/4}|\theta|/2 < -1/2
\end{aligned}$$

or, equivalently, when $n > (3/|\theta|)^4$. Combining these two bounds yields

$$\begin{aligned}
|x||x-\sqrt{n}\theta|\phi(x-\sqrt{n}\theta) &\leq \phi(x)n^{3/4}2(|\theta| \vee 1)\exp(-|\theta|n^{3/4}/2) \\
&= \phi(x)\begin{cases} 2n^{3/4}\exp(-|\theta|n^{3/4}/2) & \text{if } \ |\theta|<1 \\ 2n^{3/4}|\theta|\exp(-|\theta|n^{3/4}/2) & \text{if } \ |\theta| \geq 1 \end{cases} \\
&= \phi(x)\begin{cases} (4/|\theta|)(n^{3/4}|\theta|/2)\exp(-|\theta|n^{3/4}/2) & \text{if } \ |\theta|<1 \\ 4(n^{3/4}|\theta|/2)\exp(-|\theta|n^{3/4}/2) & \text{if } \ |\theta| \geq 1 \end{cases} \\
&\leq \frac{4e^{-1}}{|\theta| \wedge 1}\phi(x).
\end{aligned}$$

When $\theta = 0$

$$b_n'(0) = -(1-a)\int_{-n^{1/4}}^{n^{1/4}} x^2\phi(x)\ dx \to -(1-a)\int_{-\infty}^{\infty} x^2\phi(x)dx = -(1-a).$$

(c') The information inequality implies that

$$Var_\theta(\sqrt{n}(T_n-\theta)) \geq \frac{(b_n'(\theta)+1)^2}{I(\theta)} = (b_n'(\theta)+1)^2$$

since $I(\theta) = 1$. At the point $\theta = 0$ the right side converges to a^2, while the limit inferior of the left side is the variance of the limiting distribution at $\theta = 0$, namely a^2. Thus there is no contradiction with the information inequality.

(c) Using the distributional identity in (b) yields

$$\begin{aligned}
R_n(\theta) &= 1+(1-a)^2E(Z+\sqrt{n}\theta)^2 1_{[|Z+\theta\sqrt{n}| \leq n^{1/4}]} \\
&\qquad - 2(1-a)E\{Z(Z+\sqrt{n}\theta)1_{[|Z+\theta\sqrt{n}| \leq n^{1/4}]}\} \\
&= 1+\{(1-a)^2-2(1-a)\}E(Z+\sqrt{n}\theta)^2 1_{[|Z+\theta\sqrt{n}| \leq n^{1/4}]} \\
&\qquad + 2(1-a)\sqrt{n}\theta E\{(Z+\sqrt{n}\theta)1_{[|Z+\theta\sqrt{n}| \leq n^{1/4}]}\} \\
&= 1-(1-a^2)E(Z+\sqrt{n}\theta)^2 1_{[|Z+\theta\sqrt{n}| \leq n^{1/4}]} \\
&\qquad + 2(1-a)\sqrt{n}\theta E\{(Z+\sqrt{n}\theta)1_{[|Z+\theta\sqrt{n}| \leq n^{1/4}]}\}
\end{aligned}$$

(This confirms the first identity in Lehmann's example 4.7, page 442.) Squaring out the expectation in the second term and writing the third term as the sum of two terms yields, with $\alpha_n \equiv n^{1/4} - \sqrt{n}\theta$, $\beta_n \equiv -n^{1/4} - \sqrt{n}\theta$,

$$\begin{aligned} R_n(\theta) &= 1 - (1-a^2)EZ^2 1_{[|Z+\theta\sqrt{n}| \le n^{1/4}]} \\ &\quad - 2(1-a^2)\sqrt{n}\theta E Z 1_{[|Z+\theta\sqrt{n}| \le n^{1/4}]} \\ &\quad - (1-a^2)n\theta^2(\Phi(\beta_n) - \Phi(\alpha_n)) \\ &\quad + 2(1-a)n\theta^2(\Phi(\beta_n) - \Phi(\alpha_n)) \\ &\quad + 2(1-a)\sqrt{n}\theta E\{Z 1_{[|Z+\theta\sqrt{n}| \le n^{1/4}]}\} \\ &= 1 - (1-a^2)EZ^2 1_{[|Z+\theta\sqrt{n}| \le n^{1/4}]} \\ &\quad + (1-a)^2 n\theta^2(\Phi(\beta_n) - \Phi(\alpha_n)) \\ &\quad - 2a(1-a)\sqrt{n}\theta E(Z 1_{[|Z+\theta\sqrt{n}| \le n^{1/4}]}) \end{aligned}$$

where

$$\begin{aligned} E(Z 1_{[|Z+\theta\sqrt{n}| \le n^{1/4}]}) &= \int_{\alpha_n}^{\beta_n} z\phi(z)dz \\ &= -\int_{\alpha_n}^{\beta_n} \phi'(z)dz \qquad \text{since } \phi'(z) = -z\phi(z) \\ &= -(\phi(\beta_n) - \phi(\alpha_n)). \end{aligned}$$

Thus it follows that

$$\begin{aligned} R_n(\theta) &= 1 - (1-a^2)EZ^2 1_{[|Z+\theta\sqrt{n}| \le n^{1/4}]} \\ &\quad + (1-a)^2 n\theta^2(\Phi(\beta_n) - \Phi(\alpha_n)) \\ &\quad + 2a(1-a)\sqrt{n}\theta(\phi(\beta_n) - \phi(\alpha_n)). \end{aligned}$$

(This confirms the second identity in Lehmann's problem 4.7, page 442.) Now we take $\theta = \theta_n = n^{-1/4}$, and note that $\alpha_n = -2n^{1/4}$, $\beta_n = 0$. Since the expectation of in the second term in the last display is bounded below by zero and above by 1 we find that

$$\begin{aligned} R_n(n^{-1/4}) &\ge a^2 + (1-a)^2 n^{1/2}(1/2 - \Phi(-2n^{1/4}) \\ &\qquad + 2a(1-a)n^{1/4}(\phi(0) - \phi(-2n^{1/4})) \\ &\to a^2 + \infty + \infty = \infty \end{aligned}$$

since $n^{1/2}\Phi(-2n^{1/4}) \to 0$ and $n^{1/4}\phi(-2n^{1/4}) \to 0$.

(c), Second (more elegant) solution: Julia Palacios (and others): from the lecture notes, 3.3 (3), it follows that

$$R_n(\theta) = E[n(T_n - \theta)^2] = nVar[T_n] + nb_n(\theta)^2 \ge a^2 + nb_n(\theta)^2.$$

Using the formula for $b_n(\theta)$ from part (a) above, it follows that it is enough to show that

$$\left| \int_{-n^{1/4}}^{n^{1/4}} x\phi(x - n^{1/4})dx \right| \to \infty.$$

But we have, with $Z \sim N(0,1)$ (and hence $E|Z| < \infty$),

$$\begin{aligned}
\left|\int_{-n^{1/4}}^{n^{1/4}} x\phi(x - n^{1/4})dx\right| &= \left|\int_{-2n^{1/4}}^{0} (y + n^{1/4})\phi(y)dy\right| \\
&\geq \left|n^{1/4}\int_{-2n^{1/4}}^{0} \phi(y)dy\right| - \left|\int_{-2n^{1/4}}^{0} y\phi(y)dy\right| \\
&\geq n^{1/4}(\Phi(0) - \Phi(-2n^{1/4})) - E|Z| \\
&\to \infty .
\end{aligned}$$

2. Suppose that $Z \sim N(0,1)$ and, for $\mu \in R$ and $\sigma > 0$, that $X = \mu + \sigma Z \sim P_{\mu,\sigma} = N(\mu, \sigma^2)$.
(a) Compute the likelihood ratio

$$\frac{dP_{\mu,\sigma}}{dP_{0,\sigma}}(x) = \frac{\sigma^{-1}\phi((x-\mu)/\sigma)}{\sigma^{-1}\phi(x/\sigma)} \qquad \text{and} \qquad Y \equiv \log \frac{dP_{\mu,\sigma}}{dP_{0,\sigma}}(X) .$$

What is the distribution of Y under $P_{0,\sigma}$ and under $P_{\mu,\sigma}$?
(b) Plot the function $l(\mu; X) \equiv \log(dP_{\mu,\sigma}/dP_{0,\sigma})(X)$ as a function of μ.
(c) Find the maximum value of the function $l(\mu; X)$ in B (as a function of μ) and the value of $\mu \equiv \hat{\mu}$ which achieves the maximum.
(d) What is the distribution of $\hat{\mu}$ under $P_{0,\sigma}$ and under $P_{\mu,\sigma}$? What is the distribution of $l(\hat{\mu}; X)$ under $P_{0,\sigma}$ and under $P_{\mu,\sigma}$?

Solution: A. The likelihood ratio

$$\begin{aligned}
\frac{dP_{\mu,\sigma}}{dP_{0,\sigma}}(x) &= \frac{\sigma^{-1}\phi((x-\mu)/\sigma)}{\sigma^{-1}\phi(x/\sigma)} = \frac{\exp(-(x-\mu)^2/(2\sigma^2))}{\exp(-x^2/(2\sigma^2))} \\
&= \exp\left(\frac{\mu}{\sigma^2}x - \frac{1}{2}\frac{\mu^2}{\sigma^2}\right) .
\end{aligned}$$

Hence

$$Y \equiv \log \frac{dP_{\mu,\sigma}}{dP_{0,\sigma}}(X) = \frac{\mu}{\sigma}\frac{X}{\sigma} - \frac{1}{2}\frac{\mu^2}{\sigma^2} .$$

Under $P_{0,\sigma}$ we find that $E(Y) = 0 - \frac{\mu^2}{2\sigma^2}$ and $Var(Y) = \mu^2/\sigma^2 \equiv V^2$ so that

$$Y \sim N(-\frac{1}{2}V^2, V^2) \qquad \text{under} \;\; P_{0,\sigma} .$$

Under $P_{\mu,\sigma}$ a similar computation gives $E(Y) = \mu^2/\sigma^2 - \mu^2/(2\sigma^2) = V^2/2$ and $Var(Y) = V^2$, so

$$Y \sim N(\frac{1}{2}V^2, V^2) \qquad \text{under} \;\; P_{\mu,\sigma} .$$

B and C. The function

$$l(\mu, \sigma; X) \equiv \log \frac{dP_{\mu,\sigma}}{dP_{0,\sigma}}(X) = \frac{\mu}{\sigma}\frac{X}{\sigma} - \frac{\mu^2}{2\sigma^2} = \frac{X^2}{2\sigma^2} - \frac{1}{2}\frac{(X-\mu)^2}{\sigma^2}$$

is quadratic in μ with maximum value $X^2/(2\sigma^2)$ which is achieved at $\mu = \hat{\mu} \equiv X$.
D. Under $P_{0,\sigma}$, $\hat{\mu} = X \sim N(0, \sigma^2)$ and $l(\hat{\mu}, \sigma; X) = X^2/(2\sigma^2) \sim \chi_1^2/2$. Under $P_{\mu,\sigma}$, $\hat{\mu} = X \sim N(\mu, \sigma^2)$ and $l(\hat{\mu}, \sigma; X) = X^2/(2\sigma^2) \sim \chi_1^2(\delta)/2$ with $\delta = \mu^2/\sigma^2$.

3. Suppose that $\theta = (\theta_1, \theta_2) \in \Theta \subset R^k$ where $\theta_1 \in R$ and $\theta_2 \in R^{k-1}$. Show that:
A. $\mathbf{l}_1^* = \dot{\mathbf{l}}_1 - I_{12} I_{22}^{-1} \dot{\mathbf{l}}_2$ is orthogonal to $[\dot{\mathbf{l}}_2] \equiv \{a' \dot{\mathbf{l}}_2 : a \in R^{k-1}\}$ in $L_2(P_\theta)$.
B. $I_{11 \cdot 2} = \inf_{c \in R^{k-1}} E_\theta (\dot{\mathbf{l}}_1 - c' \dot{\mathbf{l}}_2)^2$ and that the minimum is achieved when $c' = I_{12} I_{22}^{-1}$. Thus

$$I_{11 \cdot 2} = E_\theta (\dot{\mathbf{l}}_1 - I_{12} I_{22}^{-1} \dot{\mathbf{l}}_2)^2 = E_\theta [(\mathbf{l}_\theta^*)^2] \,.$$

C. Prove the formulas (16) and (17) on page 21 of the Chapter 3 notes and interpret these formulas geometrically.

Solution: A. Note that for any $a \in R^{k-1}$ we have

$$\begin{aligned}
E_\theta [l_1^* \dot{l}_2^T a] &= E_\theta \left\{ (\dot{l}_1 - I_{12} I_{22}^{-1} \dot{l}_2) \dot{l}_2^T a \right\} \\
&= \left\{ E_\theta \left\{ \dot{l}_1 \dot{l}_2^T \right\} - I_{12} I_{22}^{-1} E_\theta \left\{ \dot{l}_2 \dot{l}_2^T \right\} \right\} a \\
&= \{ I_{12} - I_{12} \} a = 0 \,.
\end{aligned}$$

Thus l_1^* is orthogonal to $[\dot{l}_2]$ in $L_2(P_\theta)$.
B. Note that for any $c \in R^{k-1}$ we have

$$\begin{aligned}
& E_\theta (\dot{l}_1 - c' \dot{l}_2)^2 \\
&= E_\theta (\dot{l}_1 - I_{12} I_{22}^{-1} \dot{l}_2 + I_{12} I_{22}^{-1} \dot{l}_2 - c' \dot{l}_2)^2 \\
&= E_\theta (\dot{l}_1 - I_{12} I_{22}^{-1} \dot{l}_2)^2 + E_\theta ((I_{12} I_{22}^{-1} - c') \dot{l}_2)^2 \\
&= I_{11} - I_{12} I_{22}^{-1} I_{21} + E_\theta ((I_{12} I_{22}^{-1} - c') \dot{l}_2)^2 \\
&\geq I_{11 \cdot 2}
\end{aligned}$$

with equality if and only if $c' = I_{12} I_{22}^{-1}$. Here the second equality uses the orthogonality proved in A.
C. Formula (16) says that

$$\tilde{l}_1 = I_{11}^{-1} \dot{l}_1 - I_{11}^{-1} I_{12} \tilde{l}_2 \,. \tag{0.1}$$

One way to derive this is as indicated on page 21: since $\tilde{l} = I^{-1} \dot{l}$ we have

$$\tilde{l}_1 = I^{11} \dot{l}_1 + I^{12} \dot{l}_2 \qquad \text{and} \qquad \tilde{l}_2 = I^{21} \dot{l}_1 + I^{22} \dot{l}_2 \,.$$

Hence it follows that

$$\begin{aligned}
& \tilde{l}_1 + I_{11}^{-1} I_{12} \tilde{l}_2 \\
&= I^{11} \dot{l}_1 + I^{12} \dot{l}_2 + I_{11}^{-1} I_{12} (I^{21} \dot{l}_1 + I^{22} \dot{l}_2) \\
&= I_{11}^{-1} \left\{ (I_{11} I^{11} + I_{12} I^{21}) \dot{l}_1 + (I_{11} I^{12} + I_{12} I^{22}) \dot{l}_2 \right\} \\
&= I_{11}^{-1} \left\{ Ident \cdot \dot{l}_1 + 0 \cdot \dot{l}_2 \right\} \\
&= I_{11}^{-1} \dot{l}_1 \,.
\end{aligned}$$

Rearranging yields (0.1). Note that this indentity decomposes the efficient influence function $\tilde{l}_1$ in the larger model with both θ_1 and θ_2 unknown into its projection onto the efficient influence function in the sub-model when θ_2 is known, namely $I_{11}^{-1} \dot{l}_1$, and a term which is orthogonal to $[\dot{l}_1]$. Formula (17) follows immediately from (16) in view of orthogonality of the two terms:

$$
\begin{aligned}
I_{11\cdot 2}^{-1} &= E[\widetilde{l}_1 \widetilde{l}_1^T] = E[I_{11}^{-1}\dot{\mathbf{l}}_1\dot{\mathbf{l}}_1^T I_{11}^{-1}] + I_{11}^{-1} I_{12} E[\widetilde{\mathbf{l}}_2 \widetilde{\mathbf{l}}_2^T] I_{21} I_{11}^{-1} \\
&= I_{11}^{-1} + I_{11}^{-1} I_{12} I_{22\cdot 1}^{-1} I_{21} I_{11}^{-1}
\end{aligned}
$$

4. Suppose that $(T|Z) \sim \text{Weibull}(\lambda^{-1} e^{-\gamma Z}, \beta)$, and $Z \sim G_\eta$ on R with density g_η with respect to some dominating measure μ. Thus the conditional cumulative hazard function $\Lambda(t|z)$ is given by

$$\Lambda_{\gamma,\lambda,\beta}(t|z) = (\lambda e^{\gamma Z} t)^\beta = \lambda^\beta e^{\beta\gamma Z} t^\beta$$

and hence

$$\lambda_{\gamma,\lambda,\beta}(t|z) = \lambda^\beta e^{\beta\gamma Z} \beta t^{\beta - 1}\,.$$

(Recall that $\lambda(t) = f(t)/(1 - F(t))$ and

$$\Lambda(t) \equiv \int_0^t \lambda(s) ds = \int_0^t (1 - F(s))^{-1} dF(s) = -\log(1 - F(t))$$

if F is continuous.) Thus it makes sense to re-parametrize by defining $\theta_1 \equiv \beta\gamma$ (this is the parameter of interest since it reflects the effect of the covariate Z), $\theta_2 \equiv \lambda^\beta$, and $\theta_3 \equiv \beta$. This yields

$$\lambda_\theta(t|z) = \theta_3 \theta_2 \exp(\theta_1 z) t^{\theta_3 - 1}$$

You may assume that

$$a(z) \equiv (\partial/\partial\eta) \log g_\eta(z)$$

exists and $E\{a^2(Z)\} < \infty$. Thus Z is a "covariate" or "predictor variable", θ_1 is a "regression parameter" which affects the intensity of the (conditionally) Weibull variable T, and $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)$ where $\theta_4 \equiv \eta$.
(a) Derive the joint density $p_\theta(t, z)$ of (T, Z) for the re-parametrized model.
(b) Find the information matrix for θ. What does the structure of this matrix say about the effect of $\eta = \theta_4$ being known or unknown about the estimation of $\theta_1, \theta_2, \theta_3$?
(c) Find the information and information bound for θ_1 if the parameters θ_2 and θ_3 are known?
(d) What is the information bound for θ_1 if just θ_3 is known to be equal to 1?
(e) Find the efficient score function and the efficient influence function for estimation of θ_1 when θ_3 is known.
(f) Find the information $I_{11\cdot(2,3)}$ and information bound for θ_1 if the parameters θ_2 and θ_3 are unknown. (Here both θ_2 and θ_3 are in "the second block".)
(g) Find the efficient score function and the efficient influence function for estimation of θ_1 when θ_2 and θ_3 are unknown.
(h) Specialize the calculations in (d) - (g) to the case when $Z \sim \text{Bernoulli}(\theta_4)$ and compare the information bounds.

Solution: (a) Integrating $\lambda_\theta(t|z)$ with respect to t gives

$$\Lambda_\theta(t|z) = \theta_2 \exp(\theta_1 z) t^{\theta_3}\,,$$

and hence the conditional survival function $1 - F_\theta(t|z)$ is given by

$$1 - F_\theta(t|z) = \exp(-\Lambda_\theta(t|z)) = \exp(-\theta_2 \exp(\theta_1 z) t^{\theta_3})\,. \tag{0.2}$$

It follows that

$$f_\theta(t|z) = \theta_2\theta_3 e^{\theta_1 z} t^{\theta_3 - 1} \exp(-\theta_2 e^{\theta_1 z} t^{\theta_3})\,,$$

and hence that

$$\begin{aligned} p_\theta(y,z) &= f_\theta(y|z) g_\eta(z) = \theta_2\theta_3 e^{\theta_1 z} t^{\theta_3 - 1} \exp(-\theta_2 e^{\theta_1 z} t^{\theta_3}) g_\eta(z) \\ &= = \theta_2\theta_3 e^{\theta_1 z} t^{\theta_3 - 1} \exp(-\theta_2 e^{\theta_1 z} t^{\theta_3}) g_{\theta_4}(z)\,. \end{aligned}$$

(b) We first calculate the scores for θ. Note that the random variable $W \equiv \theta_2 \exp(\theta_1 Z) Y^{\theta_3}$ has, conditionally on Z, a standard Exponential(1) distribution:

$$P_\theta(W > w|Z) = P_\theta(\theta_2 \exp(\theta_1 Z) Y^{\theta_3} > w|Z) = e^{-w}$$

by (0.2). We calculate

$$\begin{aligned} l(\theta|Y,Z) &= \log p_\theta(Y,Z) \\ &= \log\theta_2 + \log\theta_3 + \theta_1 Z + (\theta_3 - 1)\log Y - \theta_2 e^{\theta_1 Z} Y^{\theta_3} + \log g_{\theta_4}(Z)\,, \\ \dot{\mathrm{l}}_1(Y,Z) &= Z - Z\theta_2 e^{\theta_1 Z} Y^{\theta_3} = Z(1 - W)\,, \\ \dot{\mathrm{l}}_2(Y,Z) &= \frac{1}{\theta_2} - \frac{\theta_2 e^{\theta_1 Z} Y^{\theta_3}}{\theta_2} = \frac{1}{\theta_2}(1 - W)\,, \\ \dot{\mathrm{l}}_3(Y,Z) &= \frac{1}{\theta_3} + \log Y - \theta_2 e^{\theta_1 Z} Y^{\theta_3} \log Y \\ &= \frac{1}{\theta_3} + \log Y\{1 - \theta_2 e^{\theta_1 Z} Y^{\theta_3}\} \\ &= \frac{1}{\theta_3}\left\{1 + \log\frac{\theta_2 e^{\theta_1 Z} Y^{\theta_3}}{\theta_2 e^{\theta_1 Z}}\{1 - W\}\right\} \\ &= \frac{1}{\theta_3}\left\{1 + \{\log W - \log(\theta_2 e^{\theta_1 Z})\}\{1 - W\}\right\} \\ &= \frac{1}{\theta_3}\left\{[1 - (W-1)\log W] + (W-1)\log(\theta_2 e^{\theta_1 Z})\right\} \\ \dot{\mathrm{l}}_4(Y,Z) &= a(Z) = a(Z,\eta)\,. \end{aligned}$$

Moreover,

$$\begin{aligned} \ddot{\mathrm{l}}_{13}(Y,Z) &= -Z\theta_2 e^{\theta_1 Z} Y^{\theta_3} \log Y = -Z\frac{1}{\theta_3}\theta_2 e^{\theta_1 Z} Y^{\theta_3} \log\left(\frac{\theta_2 e^{\theta_1 Z} Y^{\theta_3}}{\theta_2 e^{\theta_1 Z}}\right) \\ &= -\frac{Z}{\theta_3} W\{\log W - \log(\theta_2 e^{\theta_1 Z})\} \\ &= -\frac{z}{\theta_3} W\{\log W - \log(\theta_2) - \theta_1 Z\} \end{aligned}$$

$$
\begin{aligned}
\ddot{\mathbf{l}}_{23}(Y,Z) &= -e^{\theta_1 Z}Y^{\theta_3}\log Y = -\frac{1}{\theta_2\theta_3}\theta_2 e^{\theta_1 Z}Y^{\theta_3}\log\left(\frac{\theta_2 e^{\theta_1 Z}Y^{\theta_3}}{\theta_2 e^{\theta_1 Z}}\right) \\
&= -\frac{1}{\theta_2\theta_3}W\{\log W - \log(\theta_2 e^{\theta_1 Z})\} \\
&= -\frac{1}{\theta_2\theta_3}W\{\log W - \log(\theta_2) - \theta_1 Z\}\,, \\
\ddot{\mathbf{l}}_{33}(Y,Z) &= -\frac{1}{\theta_3^2}\{1 + W[\log W - \log(\theta_2 e^{\theta_1 Z}]^2\}\,.
\end{aligned}
$$

Thus we calculate easily:

$$
\begin{aligned}
I_{11}(\theta) &= E_\theta(\dot{\mathbf{l}}_1(Y,Z)^2) = E_\theta\{E[Z^2(1-W)^2|Z]\} \\
&= E\{Z^2 E[(1-W)^2|Z]\} = E(Z^2)\,, \\
I_{22}(\theta) &= E_\theta(\dot{\mathbf{l}}_2(Y,Z)^2) = E_\theta\{E[\theta_2^{-2}(1-W)^2|Z]\} = \theta_2^{-2}\,, \\
I_{33}(\theta) &= \theta_3^{-2}\left\{1 + E[W(\log W)^2] - 2E(W\log W)\{\log\theta_2 + \theta_1 E(Z)\}\right. \\
&\qquad \left. + E\{(\log\theta_2 + \theta_1 Z)^2\}\right\} \\
&= \theta_3^{-2}\left\{1 + B^2 - 2A\{\log\theta_2 + \theta_1 E(Z)\} + E\{(\log\theta_2 + \theta_1 Z)^2\}\right\} \\
I_{12}(\theta) &= E_\theta(\dot{\mathbf{l}}_1(Y,Z)\dot{\mathbf{l}}_2(Y,Z)) = E_\theta\{E[Z\theta_2^{-1}(1-W)^2|Z]\} = \theta_2^{-1}E(Z)\,, \\
I_{13}(\theta) &= -E_\theta\{\ddot{\mathbf{l}}_{13}(Y,Z)\} \\
&= \theta_3^{-1}\{E(Z)[A - \log\theta_2] - \theta_1 E(Z^2)\}\,, \\
I_{23}(\theta) &= -E_\theta\{\ddot{\mathbf{l}}_{23}(Y,Z)\} \\
&= (\theta_2\theta_3)^{-1}\{A - \log\theta_2 - \theta_1 E(Z)\}
\end{aligned}
$$

where

$$
\begin{aligned}
A &\equiv E\{W\log W\} = \int_0^\infty (w\log w)\exp(-w)dw = 1-\gamma\,, \\
B^2 &\equiv E\{W(\log W)^2\} = \pi^2/6 + (1-\gamma)^2 - 1\,.
\end{aligned}
$$

Note that since $\dot{\mathbf{l}}_4(y,z) = a(z)$ is just a function of Z, it follows easily that for $j = 1,2,3$ we also have

$$
\begin{aligned}
I_{j4}(\theta) &= E_\theta\{\dot{\mathbf{l}}_j(Y,Z)\dot{\mathbf{l}}_4(Y,Z)\} \\
&= E\{g_j(W,Z)a(Z)\} = E\{E[g_j(W,Z)a(Z)|Z]\} \\
&= E\{a(Z)E[g_j(W,Z)|Z]\} = E\{a(Z)\cdot 0\} = 0\,,
\end{aligned}
$$

Because of this orthogonality, the information bounds for $(\theta_1,\theta_2,\theta_3)$ are the same when $\theta_4 = \eta$ is unknown as when it is known.
(c) If θ_2 and θ_3 are known, then the information bound for estimation of θ_1 is just $I_{11}^{-1}(\theta) = 1/E(Z^2)$. It follows that the information matrix for θ is of the following form:

$$
I(\theta) = \begin{pmatrix}
E(Z^2) & \theta_2^{-1}E(Z) & \theta_3^{-1}C & 0 \\
\theta_2^{-1}E(Z) & \theta_2^{-2} & (\theta_2\theta_3)^{-1}D & 0 \\
\theta_3^{-1}C & (\theta_2\theta_3)^{-1}D & \theta_3^{-2}E & 0 \\
0 & 0 & 0 & Ea^2(Z)
\end{pmatrix}
$$

where

$$
\begin{aligned}
C &= E(Z)(A - \log\theta_2) - \theta_1 E(Z^2) \\
D &= A - \log\theta_2 - \theta_1 E(Z) \\
E &= 1 + B^2 - 2A(\log\theta_2 + \theta_1 E(Z)) + E(\log\theta_2 + \theta_1 Z)^2 .
\end{aligned}
$$

(d) If $\theta_3 = 1$ is known, then the information bound for θ_1 is $I_{11\cdot 2}^{-1}$ where

$$
\begin{aligned}
I_{11\cdot 2}(\theta) &= I_{11} - I_{12} I_{22}^{-1} I_{21} \\
&= E(Z^2) - (E(Z)/\theta_2)^2 \theta_2^2 = E(Z^2) - (EZ)^2 = Var(Z) .
\end{aligned}
$$

Thus $I_{11\cdot 2}^{-1} = 1/Var(Z)$.
(e) When θ_3 is known, the efficient score function and the efficient influence function for estimation of θ_1 are given by

$$
\begin{aligned}
\dot{\mathrm{l}}_1^*(Y,Z) &= \dot{\mathrm{l}}_1 - I_{12} I_{22}^{-1} \dot{\mathrm{l}}_2 \\
&= Z(1-W) - \theta_2^{-1} E(Z) \theta_2^2 \frac{1}{\theta_2}(1-W) \\
&= Z(1-W) - E(Z)(1-W) = (Z - E(Z))(1-W) ,
\end{aligned}
$$

and

$$
\begin{aligned}
\widetilde{\mathrm{l}}_1(Y,Z) &= I_{11\cdot 2}^{-1} \dot{\mathrm{l}}_1^*(Y,Z) \\
&= \frac{1}{Var(Z)}(Z - E(Z))(1-W) .
\end{aligned}
$$

(f) When both the parameters θ_2 and θ_3 are unknown, the information $I_{11\cdot(2,3)}$ is given by

$$
\begin{aligned}
I_{1\cdot(2,3)} &\equiv I_{11\cdot 2} \qquad \text{where the "second block" contains both } \theta_2, \theta_3 \\
&= I_{11} - I_{12} I_{22}^{-1} I_{21} \qquad (0.3)
\end{aligned}
$$

where

$$
\begin{aligned}
I_{12} &= (\theta_2^{-1} E(Z), \theta_3^{-1} C) , \\
I_{22}^{-1} &= \begin{pmatrix} \theta_2^2 E & -\theta_2\theta_3 D \\ -\theta_2\theta_3 D & \theta_3^2 \end{pmatrix} \frac{1}{E - D^2} .
\end{aligned}
$$

Thus the second term in (0.3) is

$$
\left\{ [E(Z)]^2 E - 2E(Z) C D + C^2 \right\} / (E - D^2) . \qquad (0.4)
$$

Now the denominator is

$$
\begin{aligned}
E - D^2 &= 1 + B^2 - 2A(\log\theta_2 + \theta_1 E(Z)) + E(\log\theta_2 + \theta_1 Z)^2 \\
&\qquad - (A - \log\theta_2 - \theta_1 E(Z))^2 \\
&= 1 + B^2 - 2A(\log\theta_2 + \theta_1 E(Z)) + E(\log\theta_2 + \theta_1 Z)^2 \\
&\qquad - [A^2 - 2A(\log\theta_2 + \theta_1 E(Z)) + (\log\theta_2 + \theta_1 E(Z))^2 \\
&= 1 + B^2 - A^2 + Var[\log\theta_2 + \theta_1 Z] \\
&= \pi^2/6 + \theta_1^2 Var(Z) ,
\end{aligned}
$$

and, upon noting that

$$\begin{aligned} C - E(Z)D &= E(Z)(A - \log\theta_2) - \theta_1 E(Z^2) - \{E(Z)(A - \log\theta_2) - \theta_1[E(Z)]^2\} \\ &= -\theta_1 Var(Z)\,, \end{aligned}$$

it follows that the numerator of (0.4) is

$$\begin{aligned} C^2 - 2E(Z)CD + [E(Z)]^2 E &= C^2 - 2E(Z)CD + [E(Z)]^2 D^2 + [E(Z)]^2(E - D^2) \\ &= (C - E(Z)D)^2 + [E(Z)]^2\{\pi^2/6 + \theta_1^2 Var(Z)\} \\ &= \theta_1^2[Var(Z)]^2 + [E(Z)]^2\{\pi^2/6 + \theta_1^2 Var(Z)\}\,. \end{aligned}$$

It follows that the information for θ_1 when θ_2 and θ_3 are unknown is equal to

$$\begin{aligned} I_{11\cdot(2,3)} &= E(Z^2) - \frac{\theta_1^2[Var(Z)]^2 + [E(Z)]^2\{\pi^2/6 + \theta_1^2 Var(Z)\}}{\pi^2/6 + \theta_1^2 Var(Z)} \\ &= \frac{\pi^2/6}{\pi^2/6 + \theta_1^2 Var(Z)} Var(Z) \le Var(Z) \le E(Z^2) \end{aligned}$$

with equality in the first inequality if and only if $\theta_1 = 0$. Note that the information decreases as θ_1 increases, and it converges to $\pi^2/(6\theta_1^2)$ as $Var(Z) \to \infty$.
(g) When θ_2 and θ_3 are unknown the efficient score function for θ_1 is, with the "second block" containing both θ_2 and θ_3,

$$\begin{aligned} \mathbf{l}_1^* &= \dot{\mathbf{l}}_1 - I_{12} I_{22}^{-1} \dot{\mathbf{l}}_2 \\ &= \dot{\mathbf{l}}_1 - (\theta_2(E(Z)E - CD), \theta_3(C - DE(Z))\dot{\mathbf{l}}_2/(E - D^2) \\ &= Z(1 - W) - \frac{E(Z)E - CD}{E - D^2}(1 - W) \\ &\quad + \frac{\theta_1 Var(Z)}{\pi^2/6 + \theta_1^2 Var(Z)}\left\{[1 - (W - 1)\log W] + (W - 1)\log(\theta_2 e^{\theta_1 Z})\right\} \\ &= \left\{Z - \frac{E(Z)E - CD + \log(\theta_2 e^{\theta_1 Z})}{\pi^2/6 + \theta_1^2 Var(Z)}\right\}(1 - W) \\ &\quad + \frac{\theta_1^2 Var(Z)}{\pi^2/6 + \theta_1^2 Var(Z)}\{1 - (W - 1)\log W\}\,. \end{aligned}$$

(h) When $Z \sim$ Bernoulli(η), then

$$\begin{aligned} &I_{11} = E(Z^2) = \eta = \theta_4\,, \\ &I_{11\cdot 2} = Var(Z) = \eta(1 - \eta) = \theta_4(1 - \theta_4)\,, \\ &I_{11\cdot(2,3)} = \frac{\pi^2/6}{\pi^2/6 + \theta_1^2 Var(Z)} Var(Z) \\ &\qquad\quad = \frac{\pi^2/6}{\pi^2/6 + \theta_1^2 \eta(1 - \eta)}\eta(1 - \eta)\,. \end{aligned}$$

The corresponding information bounds are given by the reciprocals of these quantities. See the following figures for comparisons of the information and information bounds.

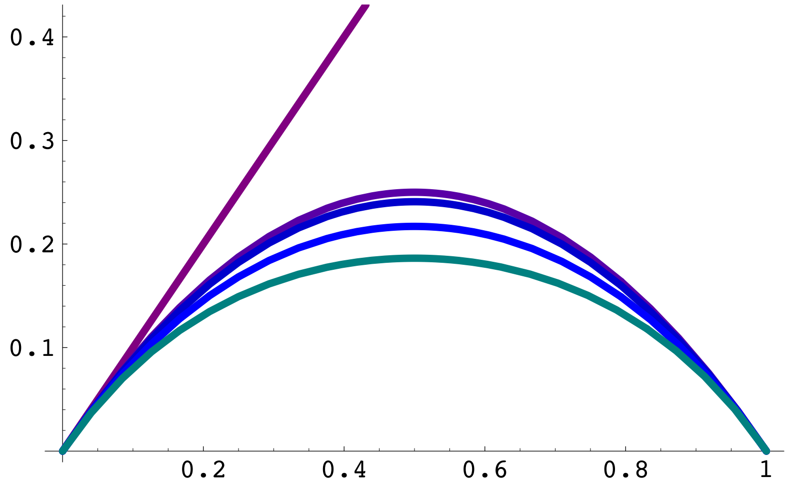

Figure 1: Plots of I_{11}, $I_{11\cdot 2}$, and $I_{11\cdot(2,3)}$ as a function of $\eta = \theta_4$, and for $\theta_1 = .5,\ 1.0, 1.5$

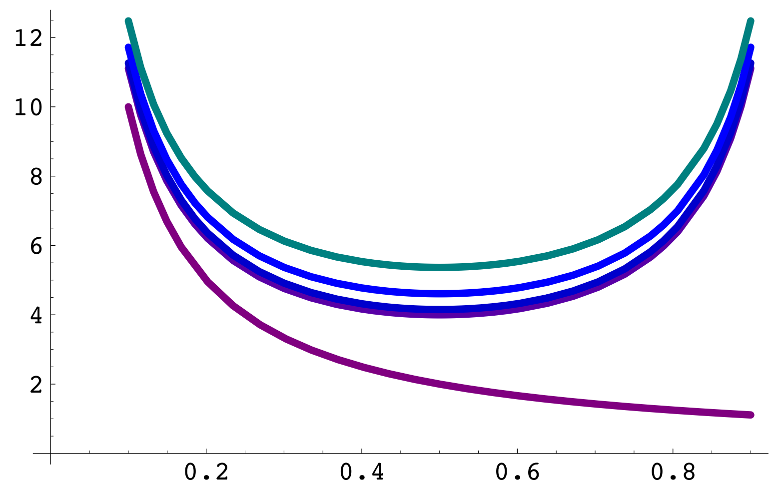

Figure 2: Plots of I_{11}^{-1}, $I_{11\cdot 2}^{-1}$, and $I_{11\cdot(2,3)}^{-1}$ as a function of $\eta = \theta_4$, , and for $\theta_1 = .5,\ 1.0, 1.5$

5. Suppose that $X \sim F_\theta = \text{exponential}(\theta)$ with density $f_\theta(x) = \theta e^{-\theta x} 1_{(0,\infty)}(x)$ and $Y \sim G_\eta$ independent of X with densities $\{g_\eta : \eta \in R^+\}$, a regular parametric model on $(0, \infty)$. Consider the following three scenarios for observation of X or functions of X:
(a) Uncensored: we observe X and Y.
(b) Right-censored: we observe $T(X,Y) = (X \wedge Y, 1\{X \le Y\} \equiv (\min\{X,Y\}, 1\{X \le Y\}) \equiv (Z, \Delta)$.
(c) Interval-censored (case 1): we observe $S(X,Y) = (Y, 1\{X \le Y\}) \equiv (Y, \Delta)$.
(i) Find the joint density of (X,Y) and joint distributions of $T(X,Y)$ and $S(X,Y)$.
(ii) Find the scores for θ and η in each of the three scenarios (a), (b), and (c). (Let $(\partial/\partial\eta) \log g_\eta(y) \equiv a(y)$ with $a \in L_2^0(G_\eta)$.)
(iii) Compute and compare $I_{X,Y}(\theta)$, $I_{T(X,Y)}(\theta)$, and $I_{S(X,Y)}(\theta)$. Make the comparisons in general and then explicitly by making one or more choices of the family $\{g_\eta\}$.

Solution: (i) In case (a) when we observe X and Y the joint density of X, Y

is simply $f_\theta(x)g_\eta(y) = \theta\exp(-\theta x)g_\eta(y)$. In case (b) the joint density $p(z,\delta) = p(z,\delta;\theta,\eta)$ (with respect to Lebesgue measure on $(0,\infty)$ times counting measure on $\{0,1\}$) is given by

$$p(z,\delta) = \{(1-G_\eta(z))f_\theta(z)\}^\delta\{(1-F_\theta(z))g_\eta(z)\}^{1-\delta}.$$

In case (c) the joint density $p(y,\delta) = p(y,\delta;\theta,\eta)$ of $S(X,Y) = (Y,\Delta)$ given by

$$p(y,\delta) = F_\theta(y)^\delta(1-F_\theta(y))^{1-\delta}g_\eta(y).$$

(ii) In case (a),

$$\log p_{X,Y}(x,y;\theta,\eta) = \log f_\theta(x) + \log g_\eta(y) = \log\theta - \theta x + \log g_\eta(y),$$

and hence the scores for θ and η are

$$\begin{aligned}\dot{l}_\theta(x,y) &= \theta^{-1} - x,\\ \dot{l}_\eta(x,y) &= a(y).\end{aligned}$$

In case (b) we find that

$$\begin{aligned}\log p(z,\delta;\theta,\eta) &= \delta(\log f_\theta(z) + \log(1-G_\eta(z))\} + (1-\delta)\{\log g_\eta(z) + \log(1-F_\theta(z))\}\\ &= \delta\log f_\theta(z) + (1-\delta)\log(1-F_\theta(z)) + (1-\delta)g_\eta(z) + \delta(1-G_\eta(z)).\end{aligned}$$

Thus the scores for θ and η are given by

$$\begin{aligned}\dot{l}_\theta(z,\delta) &= \delta(\theta^{-1} - z) + (1-\delta)(-z) = \theta^{-1}\delta - z,\\ \dot{l}_\eta(z,\delta) &= (1-\delta)a(z) + \delta(1-G_\eta(z))^{-1}\int_z^\infty a(y)dG_\eta(y).\end{aligned}$$

In case (c),

$$\log p(y,\delta;\theta,\eta) = \delta F_\theta(y) + (1-\delta)(1-F_\theta(y)) + \log g_\eta(y).$$

Thus the scores for θ and η are given by

$$\begin{aligned}\dot{l}_\theta(y,\delta) &= \left\{\frac{\delta}{F_\theta(y)}\frac{\partial}{\partial\theta}F_\theta(y) + \frac{(1-\delta)}{1-F_\theta(y)}(-\frac{\partial}{\partial\theta}F_\theta(y)\right\}\\ &= \left\{\frac{\delta}{F_\theta(y)} - \frac{(1-\delta)}{1-F_\theta(y)}\right\}\frac{\partial}{\partial\theta}F_\theta(y)\\ &= \left\{\frac{\delta}{F_\theta(y)} - \frac{(1-\delta)}{1-F_\theta(y)}\right\}(y\exp(-\theta y)\\ &= \{\delta - F_\theta(y)\}\frac{y(1-F_\theta(y))}{F_\theta(y)(1-F_\theta(y))},\\ \dot{l}_\eta(y,\delta) &= a(y).\end{aligned}$$

(iii) In case (a), the information matrix for (θ,η) is given by

$$I_{X,Y}(\theta,\eta) = \begin{pmatrix}\theta^{-2} & 0\\ 0 & Ea^2(Y)\end{pmatrix},$$

and hence the information for θ is simply θ^{-2}.
(b) In case (b),

$$\begin{aligned} I_{11}(\theta,\eta) &= E_{\theta,\eta}\dot{l}_\theta^2(Z,\Delta) \\ &= E_{\theta,\eta}(\theta^{-1}\Delta - Z)^2. \end{aligned}$$

But we can also calculate

$$\ddot{l}_{\theta,\theta}(z,\delta) = -\theta^{-2}\delta,$$

and hence

$$\begin{aligned} I_{11}(\theta,\eta) &= -E_{\theta,\eta}\ddot{l}_{\theta,\theta}(Z,\Delta) = \theta^{-2}P_{\theta,\eta}(\Delta=1) && (0.5) \\ &= \theta^{-2}\int_0^\infty F_\theta dG_\eta = \theta^{-2}E_\eta g(\theta Y) \le \theta^{-2} && (0.6) \end{aligned}$$

where $g(v) \equiv 1 - e^{-v}$ where the inequality is strict if $P_\eta(Y < \infty) > 0$. Note that

$$\ddot{l}_{\theta,\eta}(z,\delta) = 0,$$

and hence $I_{12}(\theta,\eta) = I_{21}(\theta,\eta) = 0$. Thus we conclude that the information for θ is simply $I_{11}(\theta,\eta) = \theta^{-2}P_{\theta,\eta}(\Delta = 1)$ as calculated in (0.6). When $Y \sim$ Exponential(η) this yields

$$\begin{aligned} I_{11\cdot 2}(\theta,\eta) &= \theta^{-2}\int_0^\infty (1-\exp(-\theta y)\eta\exp(-\eta y)dy \\ &= \theta^{-2}\left\{1 - \eta\int_0^\infty \exp(-(\theta+\eta)y)\right\} \\ &= \theta^{-2}\left\{1 - \frac{\eta}{\theta+\eta}\right\} \\ &= \theta^{-2}\frac{\theta}{\theta+\eta} = \theta^{-2}\frac{1}{1+r} \quad \text{with } r \equiv \eta/\theta. \end{aligned}$$

(c) In case (c), since $(\Delta|Y) \sim$ Bernoulli$(F_\theta(Y))$ we calculate conditionally on Y to find that

$$\begin{aligned} I_{11}(\theta,\eta) &= E_{\theta,\eta}\dot{l}_\theta(Y,\Delta)^2 \\ &= E_{\theta,\eta}\{F_\theta(Y)(1-F_\theta(Y))\}\frac{Y^2(1-F_\theta(Y))^2}{F_\theta(Y)^2(1-F_\theta(Y))^2} \\ &= \theta^{-2}E_{\theta,\eta}\frac{(\theta Y)^2(1-F_\theta(Y))}{F_\theta(Y)} \\ &= \theta^{-2}E_\eta h(\theta Y) \end{aligned}$$

where $h(v) \equiv v^2e^{-v}/(1-e^{-v})$ is a bounded function vanishing at 0 and ∞ and $\|h\|_\infty \le .65$; see Figure yy.

Again by computing conditionally we see that

$$\begin{aligned} I_{12}(\theta,\eta) &= E_{\theta,\eta}\dot{l}_\theta(Y,\Delta)\dot{l}_\eta(Y,\Delta) \\ &= E\left\{E\left\{(\Delta - F_\theta(Y))\frac{Ya(Y)}{F_\theta(Y)}|Y\right\}\right\} \\ &= E\left\{\frac{Ya(Y)}{F_\theta(Y)}E\left\{(\Delta - F_\theta(Y))|Y\right\}\right\} \\ &= 0. \end{aligned}$$

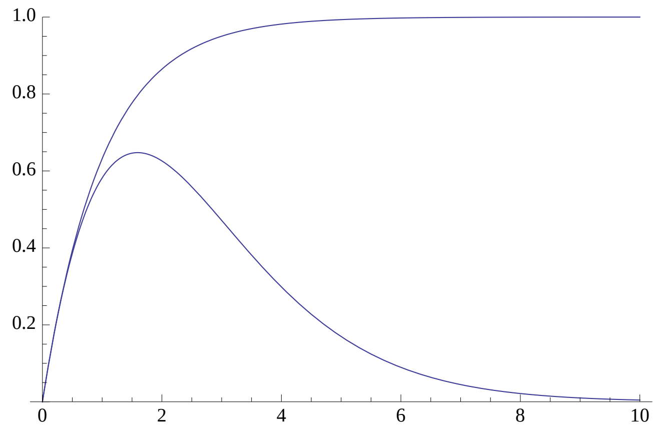

Figure 3: The functions $g(v) = 1 - e^{-v}$ and $h(v) = v^2 e^{-v}/(1 - e^{-v})$.

Thus the information for θ based on observation of $S(X, Y) = (Y, \Delta)$ is

$$I_{11}(\theta, \eta) = \theta^{-2} E_\eta h(\theta Y) \le \theta^{-2} E_\eta g(\theta Y) \le \theta^{-2}$$

where $h(v) \equiv v^2 e^{-v}/(1 - e^{-v}) \le 1 - e^{-v} \equiv g(v)$; to see this last inequality note it holds if and only if

$$v^2 e^{-v} \le (1 - e^{-v})^2 = 1 - 2e^{-v} + e^{-2v},$$

or, if and only if

$$(2 + v^2)e^{-v} \le 1 + e^{-2v} \quad \text{or, if and only if} \quad 2 + v^2 \le e^v + e^{-v},$$

or, if and only if

$$1 + \frac{v^2}{2} \le \frac{1}{2}(e^v + e^{-v});$$

and this last inequality is indeed true.

When $Y \sim \text{Exponential}(\eta)$ this becomes

$$I_{11}(\theta, \eta) = \theta^{-2} 2 \frac{\eta}{\theta} \zeta(3, 1 + \eta/\theta) = \theta^{-2} 2r\zeta(3, 1 + r)$$

where $\zeta(s, a) = \sum_{k=0}^{\infty} (k + a)^{-s}$ is the generalized zeta function and (again) $r \equiv \eta/\theta$. Figure 4 shows $I_{X,Y}(\theta)/I_{T(X,Y)}(\theta)$ and $I_{X,Y}(\theta)/I_{S(X,Y)}(\theta)$ when $Y \sim \text{Exponential}(\eta)$ as a function of $r \equiv \eta/\theta$.

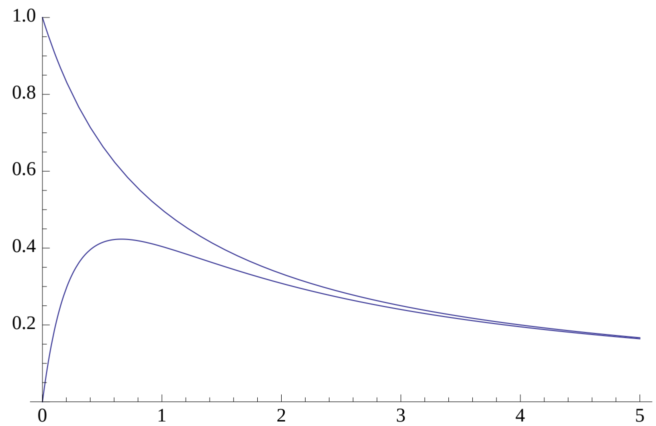

Figure 4: ARE's $I_{X,Y}(\theta)/I_{T(X,Y)}(\theta)$ and $I_{X,Y}(\theta)/I_{S(X,Y)}(\theta)$ as a function of r