Statistics 581, Problem Set 8 Solutions

Wellner; 11/23/2005

1. 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

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

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:

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

2. Suppose that $X \sim \text{Gamma}(\alpha, \beta)$; i.e. X has density p_θ given by

$$p_\theta(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1}\exp(-\beta x)1_{(0,\infty)}(x), \quad \theta = (\alpha, \beta) \in (0,\infty)\times(0,\infty) \equiv \Theta\,.$$

Consider estimation of : A. $q_A(\theta) \equiv E_\theta X$. B. $q_B(\theta) \equiv F_\theta(x_0)$ for a fixed x_0; here $F_\theta(x) \equiv P_\theta(X \le x)$.
(i) Compute $I(\theta) = I(\alpha, \beta)$; compare Lehmann & Casella page 127, Table 6.1
(ii) Compute $q_A(\theta)$, $q_B(\theta)$, $\dot{q}_A(\theta)$, and $\dot{q}_B(\theta)$.
(iii) Find the efficient influence functions for estimation of q_A and q_B.
(iv) Compare the efficient influence functions you find in (iii) with the influence functions ψ_A and ψ_B of the natural nonparametric estimators $\overline{X}_n$ and $\mathbb{F}_n(x_0)$ respectively; in particular, show that $\psi_A \in \dot{\mathcal{P}}$, while $\psi_B \notin \dot{\mathcal{P}}$.

Solution: For the Gamma(α, β) parametrized *my* way:

$$p_\theta(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1}\exp(-\beta x)1_{(0,\infty)}(x).$$

Thus

$$\log p_\theta(x) = (\alpha - 1)\log x + \alpha\log\beta - \log\Gamma(\alpha) - \beta x,$$

and hence

$$\begin{aligned} \dot{l}_\alpha(x) = \log x + \log\beta - \frac{\Gamma'}{\Gamma}(\alpha) &= \log(\beta x) - \psi(\alpha), \\ \dot{l}_\beta(x) &= \frac{\alpha}{\beta} - x \end{aligned}$$

Furthermore,

$$\begin{aligned} \ddot{l}_{\alpha\alpha}(x) &= -\psi'(\alpha), \\ \ddot{l}_{\alpha\beta}(x) &= \frac{1}{\beta} = \ddot{l}_{\beta\alpha}(x), \\ \ddot{l}_{\beta\beta}(x) &= -\frac{\alpha}{\beta^2}. \end{aligned}$$

Hence

$$I(\theta) = \begin{pmatrix} \psi'(\alpha) & -1/\beta \\ -1/\beta & \alpha/\beta^2 \end{pmatrix}.$$

(ii). Now $q_A(\theta) = \alpha/\beta$, and

$$
\begin{aligned}
q_B(\theta) = P_\theta(X \le x_0) &= \int_0^{x_0} \frac{1}{\Gamma(\alpha)} \beta^\alpha x^{\alpha-1} e^{-\beta x}\, dx \\
&= \int_0^{\beta x_0} \frac{1}{\Gamma(\alpha)} t^{\alpha-1} e^{-t}\, dt \\
&\equiv \frac{\Gamma(\alpha, \beta x_0)}{\Gamma(\alpha)}
\end{aligned}
$$

where $\Gamma(\alpha, y)$ is the incomplete gamma function; note that $\Gamma(\alpha, \infty) = \Gamma(\alpha)$. Therefore

$$
\begin{aligned}
\dot{q}_A^T(\theta) = \left(\frac{\partial}{\partial \alpha} q_A, \frac{\partial}{\partial \beta} q_B \right) &= (\frac{1}{\beta}, -\frac{\alpha}{\beta^2}) = \frac{1}{\beta}(1, -\frac{\alpha}{\beta}) \\
&= \mathrm{Cov}_\theta(X - E_\theta(X), \dot{l}_\theta^T(X)),
\end{aligned}
$$

while, with

$$
\psi(\alpha, y) \equiv \frac{\partial}{\partial \alpha} \log \Gamma(\alpha, y) \equiv \Gamma'(\alpha, y)/\Gamma(\alpha, y)\,,
$$

$$
\begin{aligned}
\dot{q}_B^T(\theta) &= \left(\frac{\Gamma'(\alpha, \beta x_0)}{\Gamma(\alpha)} - \frac{\Gamma(\alpha, \beta x_0)\Gamma'(\alpha)}{\Gamma^2(\alpha)}, \frac{(\beta x_0)^{\alpha-1}}{\Gamma(\alpha)} e^{-\beta x_0} \frac{x_0}{\beta} \beta \right) \\
&= \left(\frac{\Gamma(\alpha, \beta x_0)}{\Gamma(\alpha)} \{\psi(\alpha, \beta x_0) - \psi(\alpha)\}, \frac{x_0}{\beta} p_\theta(x_0) \right) \\
&= (q_B(\theta)\{\psi(\alpha, \beta x_0) - \psi(\alpha)\}, \frac{x_0}{\beta} p_\theta(x_0)) \\
&= \mathrm{Cov}_\theta[(1_{[0,x_0]}(X) - F_\theta(x_0)), \dot{l}_\theta^T]\,.
\end{aligned}
$$

(iii). The scores are given by

$$
\dot{l}_\theta(x) = \begin{pmatrix} \dot{l}_\alpha(x) \\ \dot{l}_\beta(x) \end{pmatrix} = \begin{pmatrix} \log(\beta x) - \psi(\alpha) \\ \frac{\alpha}{\beta} - x \end{pmatrix}
$$

and the information matrix is

$$
I(\theta) = \begin{pmatrix} \psi'(\alpha) & -1/\beta \\ -1/\beta & \alpha/\beta^2 \end{pmatrix}.
$$

Thus

$$
I^{-1}(\theta) = \begin{pmatrix} \alpha/\beta^2 & 1/\beta \\ 1/\beta & \psi'(\alpha) \end{pmatrix} \frac{\beta^2}{\alpha\psi'(\alpha) - 1},
$$

and the efficient influence function for estimation of q_A is

$$
\begin{aligned}
\tilde{l}_A &= \dot{q}_A(\theta)^T I^{-1}(\theta) \dot{l}_\theta \\
&= \frac{1}{\beta}(1, -\frac{\alpha}{\beta}) \begin{pmatrix} \alpha/\beta^2 & 1/\beta \\ 1/\beta & \psi'(\alpha) \end{pmatrix} \frac{\beta^2}{\alpha\psi'(\alpha) - 1} \begin{pmatrix} \log(\beta x) - \psi(\alpha) \\ \frac{\alpha}{\beta} - x \end{pmatrix} \\
&= \frac{\beta}{\alpha\psi'(\alpha) - 1} \{0 \cdot (\log(\beta x) - \psi(\alpha)) + (\frac{1}{\beta} - \frac{\alpha}{\beta}\psi'(\alpha))(\frac{\alpha}{\beta} - x)\} \\
&= (x - \frac{\alpha}{\beta}).
\end{aligned}
$$

Note that $X - E_\theta(X) \in [\dot{l}_\theta] = \dot{\mathcal{P}}$; in fact, $X - E_\theta(X) = -\dot{l}_\beta(X)$.
Similarly, $\tilde{l}_B(x) = \dot{q}_B(\theta) I^{-1}(\theta)\dot{l}_\theta(x)$; unfortunately, this does not simplify much, largely due to the fact that $1_{[0,x_0]}(X) - F_\theta(x_0) \notin [\dot{l}_\theta] = \dot{\mathcal{P}}$.
(iii) The information bound for estimation of q_A is

$$\begin{aligned} I^{-1}(P|q_A, \mathcal{P}) &= \dot{q}_A^T I^{-1}(\theta)\dot{q}_A \\ &= \frac{1}{\beta}(1, -\frac{\alpha}{\beta}) \begin{pmatrix} \alpha/\beta^2 & 1/\beta \\ 1/\beta & \psi'(\alpha) \end{pmatrix} \frac{\beta^2}{\alpha\psi'(\alpha) - 1} \begin{pmatrix} 1 \\ -\alpha/\beta \end{pmatrix} \frac{1}{\beta} \\ &= \frac{\alpha}{\beta^2} = Var_\theta(X). \end{aligned}$$

Similarly,

$$I^{-1}(P|q_B, \mathcal{P}) \quad = \quad = \dot{q}_B^T I^{-1}(\theta)\dot{q}_B,$$

which does not simplify appreciably because $1_{[0,x_0]}(X) - F_\theta(x_0) \notin [\dot{l}_\theta] = \dot{\mathcal{P}}$. However, since we know that $\tilde{l}_B = \Pi(1_{[0,x_0]}(x) - F(x_0)|\dot{\mathcal{P}})$, it follows easily that

$$I^{-1}(P|q_B, \mathcal{P}) < E_\theta(1_{[0,x_0]}(X) - F_\theta(x_0))^2 = F_\theta(x_0)(1 - F_\theta(x_0));$$

i.e. it is possible to improve on the natural nonparametric estimator $\mathbb{F}_n(x_0)$ of $q_B(\theta) = F_\theta(x_0)$ when the model holds.

3. Suppose that $(Y|Z) \sim \text{Weibull}(\lambda^{-1}e^{-\gamma Z}, \beta)$, 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 reparametrize 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) Exponential variable Y, and $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)$ where $\theta_4 \equiv \eta$.
(a) Derive the joint density $p_\theta(y, z)$ of (Y, 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

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

$$
\begin{aligned}
&= \frac{1}{\theta_3}\left\{[1-(W-1)\log W] + (W-1)\log(\theta_2 e^{\theta_1 Z})\right\} \\
\dot{\mathbf{l}}_4(Y,Z) &= a(Z) = a(Z,\eta)\,.
\end{aligned}
$$

Moreover,

$$
\begin{aligned}
\ddot{\mathbf{l}}_{13}(Y,Z) &= -Z\theta_2 e^{\theta_1 Z} Y^{\theta_3}\log Y = -Z\frac{1}{\theta_3}\theta_2 e^{\theta_1 Z} Y^{\theta_3}\log\left(\frac{\theta_2 e^{\theta_1 Z} Y^{\theta_3}}{\theta_2 e^{\theta_1 Z}}\right) \\
&= -\frac{Z}{\theta_3} W\{\log W - \log(\theta_2 e^{\theta_1 Z})\} \\
&= -\frac{z}{\theta_3} W\{\log W - \log(\theta_2) - \theta_1 Z\} \\
\ddot{\mathbf{l}}_{23}(Y,Z) &= -e^{\theta_1 Z} Y^{\theta_3}\log Y = -\frac{1}{\theta_2\theta_3}\theta_2 e^{\theta_1 Z} Y^{\theta_3}\log\left(\frac{\theta_2 e^{\theta_1 Z} Y^{\theta_3}}{\theta_2 e^{\theta_1 Z}}\right) \\
&= -\frac{1}{\theta_2\theta_3} W\{\log W - \log(\theta_2 e^{\theta_1 Z})\} \\
&= -\frac{1}{\theta_2\theta_3} W\{\log W - \log(\theta_2) - \theta_1 Z\}\,, \\
\ddot{\mathbf{l}}_{33}(Y,Z) &= -\frac{1}{\theta_3^2}\{1 + W[\log W - \log(\theta_2 e^{\theta_1 Z}]^2\}\,.
\end{aligned}
$$

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

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

and

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

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

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

where

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

Thus the second term in (0.3) is

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

Now the denominator is

$$\begin{aligned}
E-D^2 &= 1+B^2-2A(\log\theta_2+\theta_1E(Z))+E(\log\theta_2+\theta_1Z)^2\\
&\quad -(A-\log\theta_2-\theta_1E(Z))^2\\
&= 1+B^2-2A(\log\theta_2+\theta_1E(Z))+E(\log\theta_2+\theta_1Z)^2\\
&\quad -[A^2-2A(\log\theta_2+\theta_1E(Z))+(\log\theta_2+\theta_1E(Z))^2\\
&= 1+B^2-A^2+Var[\log\theta_2+\theta_1Z]\\
&= \pi^2/6+\theta_1^2Var(Z)\,,
\end{aligned}$$

and, upon noting that

$$\begin{aligned}
C-E(Z)D &= E(Z)(A-\log\theta_2)-\theta_1E(Z^2)-\{E(Z)(A-\log\theta_2)-\theta_1[E(Z)]^2\}\\
&= -\theta_1Var(Z)\,,
\end{aligned}$$

it follows that the numerator of (0.4) is

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

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

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

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,

$$\begin{aligned}
\mathbf{l}_1^* &= \dot{\mathbf{l}}_1-I_{12}I_{22}^{-1}\dot{\mathbf{l}}_2\\
&= \dot{\mathbf{l}}_1-(\theta_2(E(Z)E-CD),\theta_3(C-DE(Z))\dot{\mathbf{l}}_2/(E-D^2)\\
&= Z(1-W)-\frac{E(Z)E-CD}{E-D^2}(1-W)\\
&\quad +\frac{\theta_1Var(Z)}{\pi^2/6+\theta_1^2Var(Z)}\left\{[1-(W-1)\log W]+(W-1)\log(\theta_2e^{\theta_1Z})\right\}\\
&= \left\{Z-\frac{E(Z)E-CD+\log(\theta_2e^{\theta_1Z})}{\pi^2/6+\theta_1^2Var(Z)}\right\}(1-W)\\
&\quad +\frac{\theta_1^2Var(Z)}{\pi^2/6+\theta_1^2Var(Z)}\{1-(W-1)\log W\}\,.
\end{aligned}$$

(h) When $Z \sim$ Bernoulli(η), then

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

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

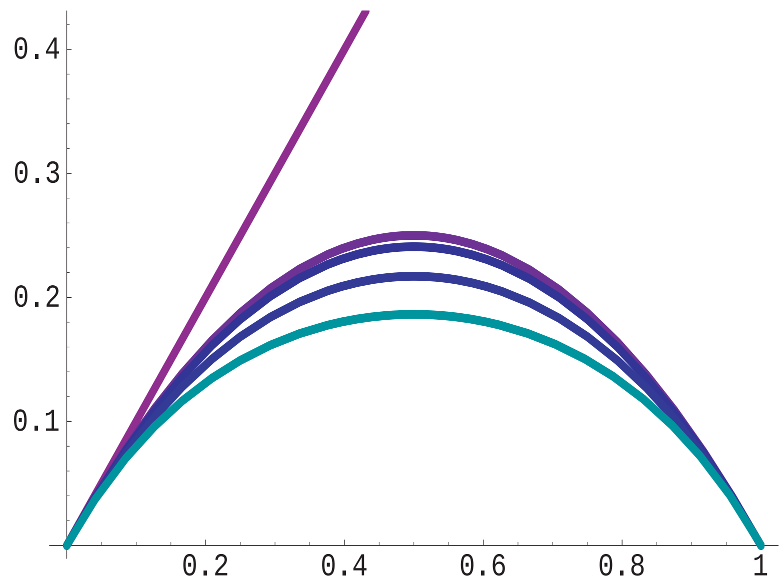

Figure 1: Plots of I_{11}, $I_{11\cdot 2}$, and $I_{11\cdot(2,3)}$ as a function of $\eta = \theta_4$, and for $\theta_1 = .5,\ 1.0, 1.5$

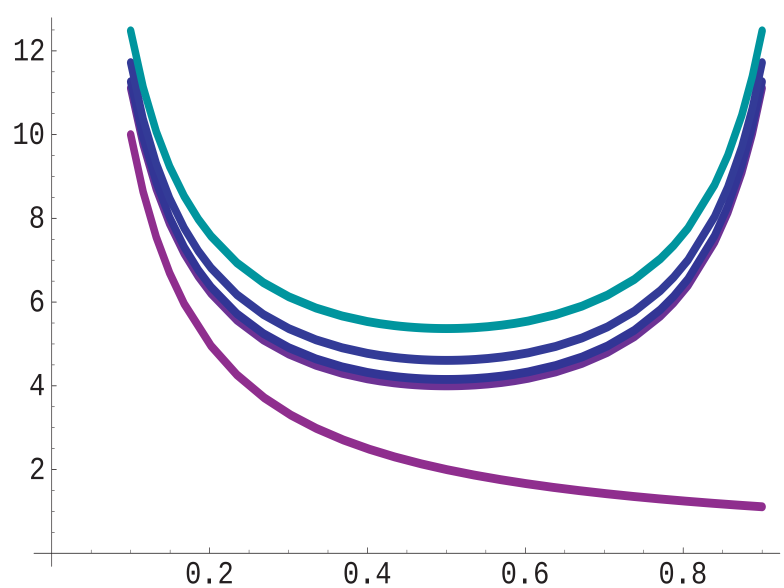

Figure 2: Plots of I_{11}^{-1}, $I_{11\cdot 2}^{-1}$, and $I_{11\cdot(2,3)}^{-1}$ as a function of $\eta = \theta_4$, , and for $\theta_1 = .5,\ 1.0, 1.5$