Statistics 581, Problem Set 8, Revised Solutions

Wellner; 11/22/2000

1. Suppose that $(Y|Z) \sim \text{Poisson}(\lambda e^{\gamma Z})$, and $Z \sim G_\eta$ on R with density g_η with respect to some dominating measure μ. 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", γ is a "regression parameter" which affects the intensity of the (conditionally) Poisson variable Y, and $\theta = (\lambda, \gamma, \eta)$.
(a) Find the information matrix for θ. What does the structure of this matrix say about the effect of η being known or unknown about the estimation of λ and γ?
(b) Find the information and information bound for γ if the parameter λ is known.
(c) Find the efficient score function and the efficient influence function for estimation of γ when λ is known.
(d) Find the information and information bound for γ if the parameter λ is unknown, $I_{\gamma\gamma\cdot\lambda}$.
(e) Find the efficient score function and the efficient influence function for estimation of γ when λ is unknown.
(f) In the case when $Z \sim \text{Bernoulli}(\eta)$, compute the ratio of the information when λ is unknown, to the information when λ is known as a function of γ and of η.

Solution: (a) The density of $X = (Y, Z)$ is

$$p_\theta(y,z) = f_{\lambda,\gamma}(y|z) g_\eta(z) = e^{-\lambda e^{\gamma z}} \frac{(\lambda e^{\gamma z})^y}{y!} g_\eta(z)\,,$$

and hence

$$\log p_\theta(y,z) = y \log(\lambda e^{\gamma z}) - \lambda e^{\gamma z} - \log y! + \log g_\eta(z)\,.$$

From this we calculate the scores $\dot{l}_\lambda$, $\dot{l}_\gamma$, and $\dot{l}_\eta$:

$$\begin{aligned}
\dot{l}_\lambda(y,z) &= \frac{y}{\lambda} - e^{\gamma z} = \frac{1}{\lambda}(y - \lambda e^{\gamma z})\,,\\
\dot{l}_\gamma(y,z) &= yz - \lambda e^{\gamma z} z = z(y - \lambda e^{\gamma z})\,,\\
\dot{l}_\eta(y,z) &= a(z)\,.
\end{aligned}$$

This leads to calculating the entries of the information matrix as follows:

$$\begin{aligned}
I_{\lambda,\lambda} &= E_\theta(\dot{l}_\lambda(X)^2) = \lambda^{-2} E[(Y - \lambda e^{\gamma Z})^2]\\
&= \lambda^{-2} E[E[(Y - \lambda e^{\gamma Z})^2 | Z]]\\
&= \lambda^{-2} E[\lambda e^{\gamma Z}] = E[e^{\gamma Z}]/\lambda\,,\\
I_{\gamma,\gamma} &= E_\theta(\dot{l}_\gamma(X)^2) = E[Z^2 (Y - \lambda e^{\gamma Z})^2]\\
&= E[E[Z^2(Y - \lambda e^{\gamma Z})^2 | Z]]\\
&= E[Z^2 \lambda e^{\gamma Z}] = \lambda E[Z^2 e^{\gamma Z}]\,,\\
I_{\eta,\eta} &= E_\theta a^2(Z)\,,
\end{aligned}$$

$$
\begin{aligned}
I_{\lambda,\gamma} &= E_\theta(\dot{l}_\lambda \dot{l}_\gamma(X)) = \lambda^{-1}E_\theta(Z(Y-\lambda e^{\gamma Z})^2) \\
&= \lambda^{-1}E[E[Z(Y-\lambda e^{\gamma Z})^2|Z]] \\
&= \lambda^{-1}E[Z\lambda e^{\gamma Z}]\,, \\
I_{\lambda,\eta} &= E_\theta(\dot{l}_\lambda(X)\dot{l}_\eta(X)) = \lambda^{-1}E[(Y-\lambda e^{\gamma Z})a(Z)] \\
&= E[E[a(Z)(Y-\lambda e^{\gamma Z})|Z]] \\
&= E[a(Z)\cdot 0] = 0\,, \\
I_{\gamma,\eta} &= E_\theta(\dot{l}_\gamma(X)\dot{l}_\eta(X)) = E[Z(Y-\lambda e^{\gamma Z})a(Z)] \\
&= E[E[a(Z)Z(Y-\lambda e^{\gamma Z})|Z]] \\
&= E[Za(Z)\cdot 0] = 0\,.
\end{aligned}
$$

Thus the information matrix $I(\theta) = I(\lambda,\gamma,\eta)$ is given by:

$$
I(\theta) = \begin{pmatrix} \lambda^{-1}E(e^{\gamma Z}) & E(Ze^{\gamma Z}) & 0 \\ E(Ze^{\gamma Z}) & \lambda E(Z^2e^{\gamma Z}) & 0 \\ 0 & 0 & E_\theta a^2(Z) \end{pmatrix}.
$$

(b) If λ is known, the information for γ is $I_{\gamma,\gamma} = \lambda E(Z^2e^{\gamma Z})$, and the information bound is $1/I_{\gamma,\gamma} = 1/\{\lambda E(Z^2e^{\gamma Z})\}$.
(c) When λ is known, the efficient score function for γ is just the score function $\dot{l}_\gamma$, and the efficient influence function is $\tilde{l}_\gamma = \dot{l}_\gamma/I_{\gamma,\gamma}$.
(d) When λ is unknown, the (efficient) information for γ is

$$
\begin{aligned}
I_{\gamma\gamma\cdot\lambda} &= I_{\gamma,\gamma} - I_{\gamma,\lambda}I_{\lambda,\lambda}^{-1}I_{\lambda,\gamma} \\
&= \lambda E(Z^2e^{\gamma Z}) - \frac{[E(Ze^{\gamma Z})]^2}{E(e^{\gamma Z})/\lambda} \\
&= \lambda\left\{E(Z^2e^{\gamma Z}) - \left(\frac{E(Ze^{\gamma Z})}{E(e^{\gamma Z})}\right)^2 E(e^{\gamma Z})\right\} \\
&= \lambda E(e^{\gamma Z})Var_{G_\gamma}(Z)\,,
\end{aligned}
$$

where G_γ is the $\gamma-$tilted distribution corresponding to G given by

$$
G_\gamma(A) = \frac{\int_A e^{\gamma z}dG(z)}{\int e^{\gamma z}dG(z)}\,.
$$

(e) When λ is unknown, the efficient score function for γ is

$$
\begin{aligned}
\dot{l}_\gamma^*(x) &= \dot{l}_\gamma(x) - I_{\gamma\lambda}I_{\lambda\lambda}^{-1}\dot{l}_\lambda(x) \\
&= z(y-\lambda e^{\theta z}) - \frac{\lambda E(Ze^{\gamma Z})}{E(e^{\gamma Z})}\lambda^{-1}(y-\lambda e^{\gamma z}) \\
&= \left(z - \frac{E(Ze^{\gamma Z})}{E(e^{\gamma Z})}\right)(y-\lambda e^{\gamma z})\,.
\end{aligned}
$$

Note that $E[l_\gamma^*(X)^2] = I_{\gamma,\gamma\cdot\lambda}$ which we computed in (d). Thus the efficient influence function for γ is $\tilde{l}_\gamma(x) = l_\gamma^*(x)/I_{\gamma\gamma\cdot\lambda}$.

(f) When $Z \sim$ Bernoulli(η), the ratio of the information when λ is unknown to the information when λ is known is

$$\begin{aligned}
\frac{I_{\gamma,\gamma\cdot\lambda}}{I_{\gamma,\gamma}} &= \frac{\lambda Var_{G_\gamma}(Z)}{\lambda E(Z^2 e^{\gamma Z})} \\
&= \frac{\eta e^\gamma - \left(\frac{\eta e^\gamma}{\eta e^\gamma + (1-\eta)}\right)^2}{\eta e^\gamma} \\
&= \frac{1-\eta}{1-\eta+\eta e^\gamma}\,.
\end{aligned}$$

See Figure 1 for a plot of this as a function of η for various values of γ.

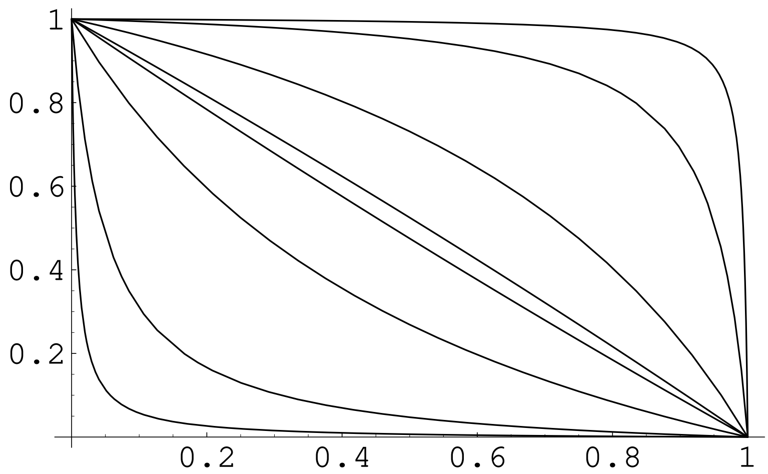

Figure 1: Ratio of Information for γ with λ known and unknown.

2. For the same set-up as in problem 1, but with the distribution G of Z assumed to be known (so η is known), consider the two parameters $q_1(\theta) = E_\theta(Y) = \nu_1(P_\theta)$ and $q_2(\theta) = P_\theta(Y \le y_0) = \nu_2(P_\theta)$ for a fixed integer y_0.
(a) Find the information bounds for estimation of $q_1(\theta)$ and $q_2(\theta)$.
(b) Find the efficient influence functions $\tilde{l}_{\nu_1}$ and $\tilde{l}_{\nu_2}$ for estimation of q_1 and q_2.
(c) If $(Y_1, Z_1), \ldots, (Y_n, Z_n)$ are i.i.d. as (Y, Z), find the influence functions ψ_1 and ψ_2 and the asymptotic variances $V_1^2(\theta)$ and $V_2^2(\theta)$ of the natural empirical estimators of q_1 and q_2 respectively.
(d) Show that $\Pi(\psi_i|\dot{\mathcal{P}}) = \tilde{l}_{\nu_i}$ for $i = 1, 2$.

Solution: My solution to this problem was incomplete at best – misleading at worst, especially since the complete solution for $q_1(\theta)$ is quite pretty. The following solution is due to Fadoua Balabdaoui and Saonli Basu:
We first give the solution for $q_1(\theta) = \nu_1(P_\theta)$:
(a:1) First, with $\theta = (\lambda, \gamma)$,

$$\begin{aligned} q_1(\theta) &= E_\theta(Y) = \nu_1(P_\theta) \\ &= E_\theta[E[Y|Z]] = E_\theta[\lambda e^{\theta Z}] = \lambda E[e^{\gamma Z}]\,. \end{aligned}$$

Thus it follows that

$$\dot{q}_1(\theta) = \begin{pmatrix} E[e^{\gamma Z}] \\ \lambda E[Z e^{\gamma Z}] \end{pmatrix} .$$

To compute the information bound for estimation of $q_1(\theta)$, we first compute $I^{-1}(\theta)$:

$$I^{-1}(\theta) = \frac{1}{D} \begin{pmatrix} \lambda E(Z^2 e^{\gamma Z}) & -E(Z e^{\gamma Z}) \\ -E(Z e^{\gamma Z}) & \lambda^{-1} E(e^{\gamma Z}) \end{pmatrix}$$

where $D \equiv \det(I(\theta)) = E(Z^2 e^{\gamma Z}) E(e^{\gamma Z}) - [E(Z e^{\gamma Z})]^2$. Hence it follows that

$$\begin{aligned} I^{-1}(\theta)\dot{q}(\theta) &= \frac{1}{D} \begin{pmatrix} \lambda E(Z^2 e^{\gamma Z}) & -E(Z e^{\gamma Z}) \\ -E(Z e^{\gamma Z}) & \lambda^{-1} E(e^{\gamma Z}) \end{pmatrix} \begin{pmatrix} E(e^{\gamma Z}) \\ \lambda E(Z e^{\gamma Z}) \end{pmatrix} \\ &= \frac{1}{D} \begin{pmatrix} \lambda E(Z^2 e^{\gamma Z}) E(e^{\gamma Z}) - \lambda [E(Z e^{\gamma Z})]^2 \\ -E(Z e^{\gamma Z}) E(e^{\gamma Z}) + E(e^{\gamma Z}) E(Z e^{\gamma Z}) \end{pmatrix} \\ &= \frac{1}{D} \begin{pmatrix} \lambda D \\ 0 \end{pmatrix} = \begin{pmatrix} \lambda \\ 0 \end{pmatrix}, \end{aligned} \tag{0.1}$$

so

$$\dot{q}_1^T(\theta) I^{-1}(\theta) \dot{q}_1(\theta) = (E(e^{\gamma Z}), \lambda E(Z e^{\gamma Z})) \begin{pmatrix} \lambda \\ 0 \end{pmatrix} = \lambda E(e^{\gamma Z})\,. \tag{0.2}$$

(b:1) By using (0.1), the efficient influence function for estimation of q_1 is

$$\begin{aligned} \tilde{l}_{\nu_1}(x) &= \dot{q}_1(\theta)^T I(\theta)^{-1} \dot{l}_\theta(x) \\ &= (\lambda, 0) \begin{pmatrix} \dot{l}_\lambda(x) \\ \dot{l}_\gamma(x) \end{pmatrix} \\ &= \lambda \dot{l}_\lambda(x) = y - \lambda e^{\gamma z}\,. \end{aligned}$$

Note that

$$E_\theta(\tilde{l}_{\nu_1}(X)^2) = E_\theta\{Var_\theta(Y|Z)\} = \lambda E(e^{\gamma Z}) = \dot{q}_1^T(\theta) I^{-1}(\theta) \dot{q}_1(\theta)$$

from (a:1).
(c:1) the natural empirical estimator of $q_1(\theta) = E_\theta(Y)$ is $\overline{Y}_n$ with influence function $\psi_1(x) = \psi_1(y,z) = y - E_\theta(Y)$. Note that

$$\begin{aligned}
\psi_1(y,z) &= y - E_\theta(Y) = y - \lambda E(e^{\gamma Z}) \\
&= y - \lambda e^{\gamma z} + \lambda e^{\gamma z} - \lambda E(e^{\gamma Z}) \\
&= \tilde{l}_{\nu_1}(x) + E_\theta(Y|Z=z) - E_\theta(Y)\,.
\end{aligned}$$

(d:1) In the previous display we have expressed ψ_1 as the sum of $\tilde{l}_{\nu_1} \in \dot{\mathcal{P}}$ (since $\tilde{l}_{\nu_1} = \lambda \dot{l}_\lambda$) and $E_\theta(Y|Z) - E_\theta(Y) = \lambda e^{\gamma Z} - \lambda E(e^{\gamma Z})$. Note that the second part of ψ_1 is orthonogal to $\dot{\mathcal{P}}$:

$$\begin{aligned}
E\{(E(Y|Z) - E(Y))\dot{l}_\lambda(X)\} &= E\{(e^{\gamma Z} - E(e^{\gamma Z}))E\{Y - E(Y|Z)|Z\}\} \\
&= E\{(e^{\gamma Z} - E(e^{\gamma Z}))\cdot 0\} = 0\,,
\end{aligned}$$

$$\begin{aligned}
E\{(E(Y|Z) - E(Y))\dot{l}_\gamma(X)\} &= \lambda E\{Z(e^{\gamma Z} - E(e^{\gamma Z}))E\{Y - E(Y|Z)|Z\}\} \\
&= \lambda E\{Z(e^{\gamma Z} - E(e^{\gamma Z}))\cdot 0\} = 0\,.
\end{aligned}$$

Thus it follows that

$$\Pi(\psi_1|\dot{\mathcal{P}}) = \tilde{l}_{\nu_1}(X) = Y - \lambda e^{\gamma Z} = Y - E(Y|Z)\,,$$

and

$$\begin{aligned}
E\psi_1^2(Y,Z) &= Var(Y) = E\{Var(Y|Z)\} + Var(E(Y|Z)) \\
&= E\{\tilde{l}_{\nu_1}(X)^2\} + E\{\lambda^2(e^{\gamma Z} - E(e^{\gamma Z}))^2\} \\
&> E\{\tilde{l}_{\nu_1}(X)^2\} = \lambda E(e^{\gamma Z})\,,
\end{aligned}$$

with strict inequality if the distribution of Z is not degenerate.
(a:2) Now for q_2: we compute

$$q_2(\theta) = P_\theta(Y \le y_0) = \nu_1(P_\theta) = E_\theta[1_{[0,y_0]}(Y)]\,.$$

Thus we compute

$$\dot{q}_2(\theta) = \begin{pmatrix} E_\theta(1_{[0,y_0]}(Y)\dot{l}_\lambda(Y,Z)) \\ E_\theta(1_{[0,y_0]}(Y)\dot{l}_\gamma(Y,Z)) \end{pmatrix} = E\{\psi_2(X)\dot{l}_\theta(X)\}\,.$$

Thus the information bound for estimation of $q_2(\theta)$ is

$$\dot{q}_2(\theta)^T I(\theta)^{-1} \dot{q}_2(\theta)$$

however this does not simplify as in the case of q_1.
(b:2) The efficient influence function for estimation of q_2 is

$$\tilde{l}_{\nu_2}(x) = \dot{q}_2(\theta)^T I(\theta)^{-1} \dot{l}_\theta(x)\,.$$

(c:2) The the natural empirical estimator of $q_2(\theta) = \nu_2(P_\theta)$ is $\nu_2(\mathbb{P}_n) = \mathbb{F}_n(y_0) = n^{-1}\sum_{i=1}^n 1_{[0,y_0]}(Y_i)$, which consistently estimates $P_\theta(Y \le y_0) = q_2(\theta)$, and is linear with influence function

$$\psi_2(x) = 1_{[0,y_0]}(y) - P_\theta(Y \le y_0)\,.$$

The asymptotic variance of $\nu_2(\mathbb{P}_n) = \mathbb{F}_n(y_0)$ is

$$\begin{aligned} V_2^2(\theta) &= E_\theta \psi_2^2(X) = q_2(\theta)(1 - q_2(\theta)) \\ &= E_\theta\{\tilde{l}_{\nu_2}(X)^2\} + E_\theta\{\psi_2(X) - \tilde{l}_{\nu_2}(X)\}^2 \\ &> E_\theta\{\tilde{l}_{\nu_2}(X)^2\} = \dot{q}_2(\theta)^T I(\theta)^{-1}\dot{q}_2(\theta)\,. \end{aligned}$$

(d:2) The influence function ψ_2 for the empirical estimator of $q_2(\theta)$ satisfies $\psi_2 \notin \dot{\mathcal{P}}$. The fact that its projection onto the tangent space coincides with $\tilde{l}_{\nu_2} \equiv \dot{q}_2(\theta)^T I(\theta)^{-1}\dot{l}_\theta$ follows from the same derivation we did in class in the Weibull case.