Statistics 581, Problem Set 6 Solutions

Wellner; 11/15/2006

1. Chapter 2, Exercise 5.3, page 25. [Hint: use the fact that $\mathbb{S}_n(t_j) - \mathbb{S}_n(t_{j-1}) = n^{-1/2}\sum_{i=[nt_{j-1}]+1}^{[nt_j]} X_i$, $j = 1, \ldots, t_k$ with $t_0 \equiv 0$ are independent random variables.]

   Solution: Note that

$$\begin{pmatrix} \mathbb{S}_n(t_1) \\ \mathbb{S}_n(t_2) \\ \vdots \\ \mathbb{S}_n(t_k) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} \mathbb{S}_n(t_1) \\ \mathbb{S}_n(t_2) - \mathbb{S}_n(t_1) \\ \vdots \\ \mathbb{S}_n(t_k) - \mathbb{S}_n(t_{k-1}) \end{pmatrix}.$$

   where the components of the vector on the right side are independent and

$$\begin{aligned} \mathbb{S}_n(t_j) - \mathbb{S}_n(t_{j-1}) &= n^{-1/2} \sum_{i=[nt_{j-1}]+1}^{[nt_j]} X_i \\ &= \sqrt{\frac{[nt_j] - [nt_{j-1}]}{n}} \frac{1}{\sqrt{[nt_j] - [nt_{j-1}]}} \sum_{i=[nt_{j-1}]+1}^{[nt_j]} X_i \\ &\to_d \mathbb{S}(t_j) - \mathbb{S}(t_{j-1}) \sim \sqrt{t_j - t_{j-1}} N(0,1) = N(0, t_j - t_{j-1}). \end{aligned}$$

   Thus it follows that

$$\begin{pmatrix} \mathbb{S}_n(t_1) \\ \mathbb{S}_n(t_2) - \mathbb{S}_n(t_1) \\ \vdots \\ \mathbb{S}_n(t_k) - \mathbb{S}_n(t_{k-1}) \end{pmatrix} \to_d \begin{pmatrix} \mathbb{S}(t_1) \\ \mathbb{S}(t_2) - \mathbb{S}(t_1) \\ \vdots \\ \mathbb{S}(t_k) - \mathbb{S}(t_{k-1}) \end{pmatrix}$$

   where the coordinates of the vector on the right side are independent. (This justifies the notation, since Brownian motion has independent increments.) Then the continuous mapping (or Mann-Wald) theorem yields

$$\begin{aligned} \begin{pmatrix} \mathbb{S}_n(t_1) \\ \mathbb{S}_n(t_2) \\ \vdots \\ \mathbb{S}_n(t_k) \end{pmatrix} &= \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} \mathbb{S}_n(t_1) \\ \mathbb{S}_n(t_2) - \mathbb{S}_n(t_1) \\ \vdots \\ \mathbb{S}_n(t_k) - \mathbb{S}_n(t_{k-1}) \end{pmatrix} \\ &\to_d \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} \mathbb{S}(t_1) \\ \mathbb{S}(t_2) - \mathbb{S}(t_1) \\ \vdots \\ \mathbb{S}(t_k) - \mathbb{S}(t_{k-1}) \end{pmatrix} \\ &= \begin{pmatrix} \mathbb{S}(t_1) \\ \mathbb{S}(t_2) \\ \vdots \\ \mathbb{S}(t_k) \end{pmatrix} \sim N_k(0, (t_j \wedge t_{j'})_{j,j'=1}^k). \end{aligned}$$

2. Consider a function $T : \mathcal{F} \to \mathbb{R}$ where $\mathcal{F}$ is some (sub) class of distribution functions F (examples include the mean, $T(F) = \mu(F) = \int x dF(x)$, the variance $T(F) = \sigma^2(F) = \int (x - \int y dF(y))^2 dF(x)$, the median $T(F) = F^{-1}(1/2)$, linear combinations of order statistics $T(F) = \int_0^1 F^{-1}(u)w(u)du$, the *mean residual life function* at $x > 0$ $T(F) \equiv e(x, F) \equiv \int_{(x,\infty)}(1 - F(u))du/(1 - F(x)) = E(X - x|X > x)$, and so forth). [The mean residual life function gives the mean life conditional on surviving beyond x.] The "principle of substitution" says that $T(F)$ can be estimated by $T(\widehat{F}_n)$. for some estimator $\widehat{F}_n$ of F. If T is sufficiently "smooth", then frequently the empirical distribution function $\mathbb{F}_n$ can be taken as the estimator $\hat{F}_n$ of F.

   Give a treatment of consistency and asymptotic normality of the estimator $e(x, \mathbb{F}_n)$ of $e(x, F)$ based on our results from sections 2.4 and 2.6. You may assume that with $X \sim F$ on $(0, \infty)$ we have $E_F X < \infty$, $E_F X^2 < \infty$, and $1 - F(x) > 0$ (as well as any other additional assumptions you need).

   Solution: First note that

$$e(x, \mathbb{F}_n) = \frac{\int_{(x,\infty)}(1 - \mathbb{F}_n(u))du}{1 - \mathbb{F}_n(x)} \tag{0.1}$$

   where we know that

$$1 - \mathbb{F}_n(x) \to 1 - F(x) \tag{0.2}$$

   by the strong law of large numbers (or the Glivenko-Cantelli theorem), and where

$$\begin{aligned}
\int_{(x,\infty)}(1 - \mathbb{F}_n(u))du &= \int_{(x,\infty)} n^{-1}\sum_{i=1}^n 1_{[X_i > u]}du \\
&= n^{-1}\sum_{i=1}^n \int_{(x,\infty)} 1_{[X_i > u]}du \\
&= n^{-1}\sum_{i=1}^n (X_i - x)1_{[X_i > x]} \equiv n^{-1}\sum_{i=1}^n Y_i
\end{aligned}$$

   where the Y_i's are i.i.d. with $Y_i \geq 0$ and

$$\begin{aligned}
EY_i &= E(X_i - x)1\{X_i > x\} = \int_{(x,\infty)}(y - x)dF(y) \\
&= \int_{(x,\infty)} \left(\int_x^y du\right) dF(y) \\
&= \int_{(x,\infty)}\int_{(x,\infty)} 1\{u < y\}dudF(y) \\
&= \int_{(x,\infty)}\int_{(x,\infty)} 1\{u < y\}dF(y)du \\
&= \int_{(x,\infty)} (1 - F(u))du.
\end{aligned}$$

   Thus

$$\int_{(x,\infty)}(1 - \mathbb{F}_n(u))du = n^{-1}\sum_{i=1}^n Y_i \to_{a.s.} EY_1 = \int_{(x,\infty)}(1 - F(u))du \tag{0.3}$$

by the strong law of large numbers. It follows from (0.1), (0.3), (0.2), and the continuous mapping theorem that $e(x, \mathbb{F}_n) \rightarrow_{a.s.} e(x, F)$ for any x for which $1 - F(x) > 0$.

To establish asymptotic normality of $e(x, \mathbb{F}_n)$ we first establish joint asymptotic normality of

$$\begin{aligned}
&\begin{pmatrix} \sqrt{n}\left(\int_{(x,\infty)}(1-\mathbb{F}_n(u))du - \int_{(x,\infty)}(1-F(u))du\right) \\ \sqrt{n}\,(1-\mathbb{F}_n(x) - (1-F(x))) \end{pmatrix} \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\begin{pmatrix} Y_i - EY_i \\ Z_i - EZ_i \end{pmatrix} \equiv \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\begin{pmatrix} Y_i^c \\ Z_i^c \end{pmatrix} \\
&\equiv \sqrt{n}\begin{pmatrix} R_n - r \\ S_n - s \end{pmatrix} \qquad (0.4)
\end{aligned}$$

where $Y_i \equiv (X_i - x)1_{[X_i > x]}$ as before and $Z_i = 1_{[X_i > x]}$ with $EZ_i = 1 - F(x)$, and then use the delta - method. Note that $r = \int_{(x,\infty)}(1 - F(u))du$ and $s = 1 - F(x) \equiv \overline{F}(x)$.

Now the random vectors $(Y_i^c, Z_i^c)'$ are i.i.d. with mean 0 and $E|(Y^c, Z^c)|^2 = E\{(Y^c)^2 + (Z^c)^2\} < \infty$ since $E(X^2) < \infty$, and with covariance matrix

$$\begin{aligned}
&E\begin{pmatrix} (Y^c)^2 & Y^c Z^c \\ Z^c Y^c & (Z^c)^2 \end{pmatrix} \\
&= \begin{pmatrix} E(X-x)^2 1\{X > x\} - (E(X-x)1\{X > x\})^2 & F(x)\overline{F}(x)e(x) \\ F(x)\overline{F}(x)e(x) & F(x)\overline{F}(x)) \end{pmatrix} \equiv \Sigma.
\end{aligned}$$

Thus, using the notation of (0.4), the multivariate CLT yields

$$\sqrt{n}\begin{pmatrix} R_n - r \\ S_n - s \end{pmatrix} \rightarrow_d \begin{pmatrix} \mathbb{R} \\ \mathbb{S} \end{pmatrix} \sim N_2(0, \Sigma).$$

Now we apply the delta method with $g(u, v) = u/v$ so that $\dot{g}(u, v) = v^{-1}(1, -u/v)$:

$$\begin{aligned}
\sqrt{n}(e(x, \mathbb{F}_n) - e(x, F)) &= \sqrt{n}(g(R_n, S_n) - g(r, s)) \\
&\rightarrow_d \dot{g}(r, s)\begin{pmatrix} \mathbb{R} \\ \mathbb{S} \end{pmatrix} \\
&= \frac{1}{\overline{F}(x)}(\mathbb{R} - e(x)\mathbb{S}) \sim N(0, V^2)
\end{aligned}$$

where, using the notation $u_+ \equiv u1\{u > 0\}$,

$$\begin{aligned}
V^2 &= \frac{1}{\overline{F}^2(x)}\left\{E(X-x)_+^2 - (\overline{F}(x)e(x))^2 - 2e(x)F(x)\overline{F}(x)e(x) + e^2(x)F(x)\overline{F}(x)\right\} \\
&= \frac{1}{\overline{F}(x)}\left\{\frac{E(X-x)_+^2}{\overline{F}(x)} - e^2(x)\right\} \\
&= \frac{1}{\overline{F}(x)} Var(X - x | X > x).
\end{aligned}$$

3. 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

$$l(\mu; X) \equiv \log \frac{dP_{\mu,\sigma}}{dP_{0,\sigma}}(X)$$

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

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

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

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

(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 $\delta = \mu^2/\sigma^2$.

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

$$P(X \geq x) = \exp(-\lambda x), \qquad x > 0.$$

(a) Show that the r-th moment of X, $\mu_r \equiv \mu_r(\lambda)$ is given by

$$\mu_r(\lambda) = EX^r = \frac{\Gamma(r+1)}{\lambda^r}.$$

(b) Use the moment calculation in (a) to show that

$$\frac{\mu_r(\lambda)}{\mu_{r+1}(\lambda)} = \frac{\lambda}{r+1}$$

and hence that the family of estimators $\{\hat{\lambda}_n^{(k)}\}_{k\geq 0}$ given by

$$\hat{\lambda}_n^{(k)} \equiv (k+1)\frac{\overline{X^k}_n}{\overline{X^{k+1}}_n} \equiv (k+1)\frac{n^{-1}\sum_1^n X_i^k}{n^{-1}\sum_1^n X_i^{k+1}}$$

are all consistent estimators of λ: $\hat{\lambda}_n^{(k)} \to_p \lambda$ for each $k = 0, 1, 2, \ldots$.
(c) Show that

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

and compute $\sigma_k^2(\lambda)$ explicitly as a function of k and λ.
(c) What is the asymptotic relative efficiency of $\hat{\lambda}_n^{(k)}$ to $\widehat{\lambda}_n \equiv \hat{\lambda}_n^{(0)} = 1/\overline{X}_n$ for $k > 1$?

Solution:
(a) We compute

$$\begin{aligned} E(X^r) &= \int_0^\infty x^r \lambda e^{-\lambda x} dx = \lambda^{-r} \int_0^\infty (\lambda x)^r e^{-\lambda x} \lambda dx \\ &= \lambda^{-r} \int_0^\infty y^{(r+1)-1} e^{-y} dy = \lambda^r \Gamma(r+1). \end{aligned}$$

(b) It follows from (a) that

$$\frac{\mu_r(\lambda)}{\mu_{r+1}(\lambda)} = \frac{\lambda}{r+1}$$

and hence

$$\begin{aligned} \hat{\lambda}_n^{(k)} &\equiv (k+1)\frac{\overline{X^k}_n}{\overline{X^{k+1}}_n} \equiv (k+1)\frac{n^{-1}\sum_1^n X_i^k}{n^{-1}\sum_1^n X_i^{k+1}} \\ &\to_p (k+1)\frac{\mu_k(\lambda)}{\mu_{k+1}(\lambda)} = \lambda. \end{aligned}$$

(c) Now by the multivariate CLT it follows that

$$\sqrt{n}\begin{pmatrix} \overline{X^k}_n - \mu_k \\ \overline{X^{k+1}}_n - \mu_{k+1} \end{pmatrix} \to_d \underline{Z} \sim N_2(0, \Sigma)$$

where

$$\begin{aligned} \Sigma &= \begin{pmatrix} \frac{\Gamma(2k+1)-\Gamma(k+1)^2}{\lambda^{2k}} & \frac{\Gamma(2k+2)-\Gamma(k+1)\Gamma(k+2)}{\lambda^{2k+1}} \\ \frac{\Gamma(2k+2)-\Gamma(k+1)\Gamma(k+2)}{\lambda^{2k+1}} & \frac{\Gamma(2k+3)-(\Gamma(k+2)^2}{\lambda^{2k+2}} \end{pmatrix} \\ &= \frac{1}{\lambda^{2k}}\begin{pmatrix} \Gamma(2k+1)-\Gamma(k+1)^2 & \frac{\Gamma(2k+2)-\Gamma(k+1)\Gamma(k+2)}{\lambda^1} \\ \frac{\Gamma(2k+2)-\Gamma(k+1)\Gamma(k+2)}{\lambda^1} & \frac{\Gamma(2k+3)-(\Gamma(k+2)^2}{\lambda^2} \end{pmatrix}. \end{aligned}$$

Thus by the delta method with $g(u,v) = u/v$, so that $\dot{g}(u,v) = v^{-1}(1, -u/v)$

$$\begin{aligned}
\sqrt{n}(\hat{\lambda}_n^{(k)} - \lambda) &= (k+1)\sqrt{n}(g(\overline{X^k}_n, \overline{X^{k+1}}_n) - g(\mu_k(\lambda), \mu_{k+1}(\lambda)) \\
&\to_d (k+1)\dot{g}(\mu_k(\lambda), \mu_{k+1}(\lambda))\underline{Z} \\
&= \frac{k+1}{\mu_{k+1}(\lambda)}(Z_1 - \frac{\mu_k}{\mu_{k+1}} Z_2 \\
&= \frac{1}{\mu_{k+1}}((k+1)Z_1 - \lambda Z_2) \\
&\sim \frac{1}{\mu_{k+1}} N(0, \lambda^{-2k} C_k) = N(0, \frac{1}{\lambda^{2k}\mu_{k+1}^2} C_k) \\
&= N(0, \lambda^2 \frac{C_k}{\Gamma(k+2)^2}) \equiv N(0, \lambda^2 D_k)
\end{aligned}$$

where

$$C_k = (k+1)^2 \left\{\Gamma(2k+1) - \Gamma(k+1)^2\right\} - 2(k+1)\{\Gamma(2k+2) - \Gamma(k+1)\Gamma(k+2)\} + \Gamma(2k+3) - \Gamma(k+2)^2.$$

and (after a bit of algebra)

$$D_k = \frac{\Gamma(2k+1)}{\Gamma(k+1)^2}\left\{1 - 2\frac{2k+1}{k+1} + \frac{(2k+2)(2k+1)}{(k+1)^2}\right\}.$$

(c) When $k = 0$, we compute $D_k = 1$. Thus the asymptotic relative efficiency of $\hat{\lambda}_n^{(k)}$ with respect to $\hat{\lambda}_n^{(0)}$ is D_0/D_k. These estimators become inefficient relative to the mean very rapidly as k increases, as is shown by the following plot of the relative efficiency.

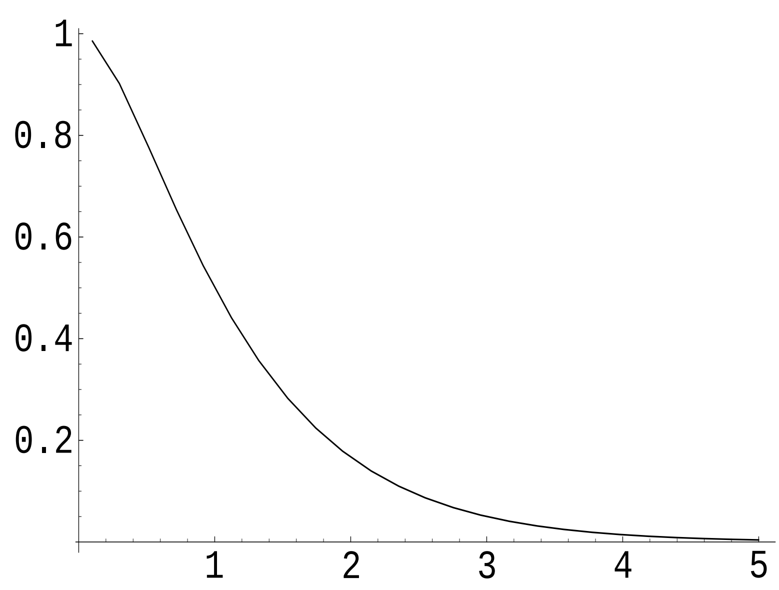

Figure 1: Asymptotic relative efficiency of $\hat{\lambda}_n^{(k)}$ with respect to $\hat{\lambda}_n^{(0)}$

5. Problem 4, page 132, Ferguson, AFCILST, modified slightly.
Suppose that $X_1, \ldots, X_n$ are independent random variables with $X_i \sim$

Poisson($\exp(\theta z_i)$) for $i = 1, \ldots, n$ where $z_1, \ldots, z_n$ are known real numbers.
(a) Find the Cramér - Rao bound for the variance of an unbiased estimator of θ based on $X_1, \ldots, X_n$.
(b) Find the Cramér - Rao bound of an unbiased estimator of $q(\theta) = P_\theta(X < 3)$.
(c) Repeat (a) and (b) for the case in which we observe (X_i, Z_i) for $i = 1, \ldots, n$ where $(X_i|Z_i) \sim$ Poiss($\exp(\theta Z_i)$), and $Z_i \sim G$ on R with density g where g is unknown. Compare the bounds for this model with those computed in (a) and (b).

Solution: (a) First, the joint density (with respect to the $n-$fold product of counting measure on $\{0, 1, \ldots\}$) of $X_1, \ldots, X_n$ is given by

$$p_\theta(x_1, \ldots, x_n) = \prod_{i=1}^n \exp(-e^{\theta z_i}) \frac{(e^{\theta z_i})^{x_i}}{x_i!},$$

so

$$\log p_\theta(x_1, \ldots, x_n) = \sum_{i=1}^n \{x_i \theta z_i - e^{\theta z_i} - \log(x_i!)\},$$

and

$$\dot{\mathbf{l}}_\theta(x_1, \ldots, x_n) = \sum_{i=1}^n \{x_i z_i - z_i e^{\theta z_i}\} = \sum_{i=1}^n z_i \{x_i - e^{\theta z_i}\}.$$

It follows that the information for θ is given by

$$I(\theta) = Var(\dot{\mathbf{l}}_\theta(X_1, \ldots, X_n)) = \sum_{i=1}^n z_i^2 e^{\theta z_i}.$$

Thus the Cramér - Rao lower bound for unbiased estimators T of θ is given by

$$Var_\theta(T(X_1, \ldots, X_n)) \geq \frac{1}{I(\theta)} = \frac{1}{\sum_{i=1}^n z_i^2 e^{\theta z_i}}.$$

(b) For estimation of

$$q(\theta) = n^{-1} \sum_{i=1}^n P_\theta(X_i < 3) = n^{-1} \sum_{i=1}^n \int 1_{[x_i \leq 2]} p_\theta(x_i) d\nu(x_i)$$

we compute

$$\begin{aligned} \dot{q}(\theta) &= n^{-1} \sum_{i=1}^n \int 1_{[x_i \leq 2]} \dot{\mathbf{l}}_\theta(x_i) p_\theta(x_i) d\nu(x_i) \\ &= n^{-1} \sum_{i=1}^n \int 1_{[x_i \leq 2]} z_i (x_i - e^{\theta z_i}) \exp(-e^{\theta z_i}) \frac{(e^{\theta z_i})^{x_i}}{x_i!} d\nu(x_i), \end{aligned}$$

and the Cramér-Rao lower bound for unbiased estimators of $q(\theta)$ is given by

$$Var_\theta(T) \geq \frac{\dot{q}(\theta)^2}{I(\theta)}.$$

where $I(\theta)$ is as in (a).
(c) When the Z_i's are i.i.d. with density g we have

$$p_\theta(x,z) = \exp(-e^{\theta z})\frac{(e^{\theta z})^x}{x!}g(z),$$

so

$$\log p_\theta(x,z) = x\theta z - e^{\theta z} - \log(x!) + \log g(z),$$

and the score for θ for a sample of size 1 is

$$\dot{\mathbf{l}}(x,z) = z(x - e^{\theta z}).$$

Thus the information for θ is given by

$$\begin{aligned} I(\theta) &= E\dot{\mathbf{l}}^2(X,Z) = E\{Z^2(X-e^{\theta Z})^2\} \\ &= E\{E[Z^2(X-e^{\theta Z})^2|Z]\} \\ &= E\{Z^2E[(X-e^{\theta Z})^2|Z]\} \\ &= E\{Z^2e^{\theta Z}\}. \end{aligned}$$

For estimation of $q(\theta) = P_\theta(X \le 2) = E_\theta 1_{[X\le 2]}$ we compute

$$\dot{q}(\theta) = E_\theta\{1_{[X\le 2]}\dot{\mathbf{l}}_\theta(X,Z)\},$$

and the resulting information bound for unbiased estimators of $q(\theta)$ is given by

$$Var_\theta(T) \ge \frac{\dot{q}^2(\theta)}{nI(\theta)}.$$

Remark: It is of interest to consider this same example, but with a "baseline mean parameter" λ in addition, so that $E(X|Z) = \lambda\exp(\theta Z)$ (with $(X|Z) \sim Poisson(E(X|Z))$ again). What is the information for θ in the presence of the nuisance parameter λ?