Spring 2014

1. For known values $x_{i,1}, x_{i,2}$, $i = 1, \ldots, n$, let

$$Z_i = \beta_1 x_{i,1} + \epsilon_i$$

and

$$Y_i = \beta_1 x_{i,1} + \beta_2 x_{i,2} + \epsilon_i, \qquad i = 1, \ldots, n,$$

where ϵ_i, $i = 1, 2, \ldots, n$ are independent normal random variables with mean 0 and variance 1.

a) Given the data $\boldsymbol{Z} = (Z_1, \ldots, Z_n)$, compute the maximum likelihood estimate of β_1 and show that it achieves the Cramer-Rao lower bound. Throughout this part and the following, make explicit any non-degeneracy assumptions that may need to be made.

Solution. Note that $Z_i \sim \mathcal{N}(\beta_1 x_{i,1}, 1)$. Hence, the likelihood is

$$\mathcal{L}(\beta_1; \boldsymbol{z}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{(z_i - \beta_1 x_{i,1})^2}{2}} = (2\pi)^{-\frac{n}{2}} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}(z_i - \beta_1 x_{i,1})^2\right),$$

and so, the log-likelihood is

$$\log \mathcal{L}(\beta_1; \boldsymbol{z}) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{n}(z_i - \beta_1 x_{i,1})^2.$$

Taking the derivative with respect to β_1, we get

$$\frac{d}{d\beta_1}\log \mathcal{L}(\beta_1; \boldsymbol{z}) = \sum_{i=1}^{n} x_{i,1}(z_i - \beta_1 x_{i,1}) = \sum_{i=1}^{n} x_{i,1} z_i - \beta_1 \sum_{i=1}^{n} x_{i,1}^2.$$

Setting the above equation to zero and solving for β_1 gives us

$$\boxed{\hat{\beta}_1 = \frac{\sum x_{i,1} z_i}{\sum x_{i,1}^2},}$$

assuming that $\sum x_{i,1}^2 \neq 0$. We see that this is indeed the MLE since the second derivative

$$\frac{d^2}{d\beta_1^2}\mathcal{L}(\beta_1; \boldsymbol{z}) = -\sum_{i=1}^{n} x_{i,1}^2$$

is always negative. Note that $\hat{\beta}_1$ is unbiased:

$$\mathbb{E}\hat{\beta}_1 = \frac{\sum x_{i,1}\mathbb{E}z_i}{\sum x_{i,1}^2} = \frac{\sum x_{i,1}^2 \beta_1}{\sum x_{i,1}^2} = \beta_1.$$

Now, the variance of $\hat{\beta}_1$ is

$$\operatorname{Var}\left(\hat{\beta}_1\right) = \frac{1}{\left(\sum x_{i,1}^2\right)^2}\operatorname{Var}\left(\sum x_{i,1} z_i\right) = \frac{1}{\left(\sum x_{i,1}^2\right)^2}\sum_{i=1}^{n}\operatorname{Var}(x_{i,1} z_i) = \frac{1}{\left(\sum x_{i,1}^2\right)^2}\sum_{i=1}^{n} x_{i,1}^2 = \frac{1}{\sum x_{i,1}^2}.$$

On the other hand, the Fisher information is

$$\mathcal{I}(\beta_1) = -\mathbb{E}_{\beta_1}\left[-\sum_{i=1}^{n} x_{i,1}^2\right] = \sum_{i=1}^{n} x_{i,1}^2.$$

Hence, the Cramer-Rao lower bound is achieved by $\hat{\beta}_1$. $\square$

b) Based on $\boldsymbol{Y} = (Y_1, \ldots, Y_n)$, compute the Cramer-Rao lower bound for the estimation of (β_1, β_2), and in particular, compute a variance lower bound for the estimation of β_1 in the presence of unknown β_2.

Solution. Note that $Y_i \sim \mathcal{N}(\beta_1 x_{i,1} + \beta_2 x_{i,2}, 1)$. Similarly as in part (a), we see that the likelihood function is

$$\mathcal{L}(\beta_1, \beta_2; \boldsymbol{y}) = (2\pi)^{-\frac{n}{2}} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}(y_i - \beta_1 x_{i,1} - \beta_2 x_{i,2})^2\right),$$

and so, the log-likelihood is

$$\log \mathcal{L}(\beta_1, \beta_2; \boldsymbol{y}) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{n}(y_i - \beta_1 x_{i,1} - \beta_2 x_{i,2})^2.$$

Since the first partial derivatives are

$$\frac{\partial}{\partial \beta_1}\log \mathcal{L}(\beta_1, \beta_2; \boldsymbol{y}) = \sum_{i=1}^{n} x_{i,1}(y_i - \beta_1 x_{i,1} - \beta_2 x_{i,2}), \text{ and}$$
$$\frac{\partial}{\partial \beta_2}\log \mathcal{L}(\beta_1, \beta_2; \boldsymbol{y}) = \sum_{i=1}^{n} x_{i,2}(y_i - \beta_1 x_{i,1} - \beta_2 x_{i,2}),$$

the second partial derivatives are

$$\frac{\partial^2}{\partial \beta_1^2}\log \mathcal{L} = -\sum_{i=1}^{n} x_{i,1}^2, \quad \frac{\partial^2}{\partial \beta_1 \beta_2}\log \mathcal{L} = -\sum_{i=1}^{n} x_{i,1}x_{i,2}, \text{ and } \frac{\partial^2}{\partial \beta_2^2}\log \mathcal{L} = -\sum_{i=1}^{n} x_{i,2}^2.$$

Hence, the Cramer-Rao lower bound is

$$\begin{aligned}\mathcal{I}^{-1}(\beta_1, \beta_2) &= \begin{pmatrix} \sum x_{i,1}^2 & \sum x_{i,1}x_{i,2} \\ \sum x_{i,1}x_{i,2} & \sum x_{i,2}^2 \end{pmatrix}^{-1} \\ &= \frac{1}{\sum x_{i,1}^2 \sum x_{i,2}^2 - \left(\sum x_{i,1}x_{i,2}\right)^2}\begin{pmatrix} \sum x_{i,2}^2 & -\sum x_{i,1}x_{i,2} \\ -\sum x_{i,1}x_{i,2} & \sum x_{i,1}^2 \end{pmatrix}.\end{aligned}$$

In particular, the variance lower bound for the estimation of β_1 in the presence of unknown β_2 is

$$\boxed{\frac{\sum x_{i,2}^2}{\sum x_{i,1}^2 \sum x_{i,2}^2 - \left(\sum x_{i,1}x_{i,2}\right)^2}.}$$

c) Compare the variance lower bound in (a), which is the same as the one for the model for Y_i where β_2 is known to be equal to zero, to the one in (b), where β_2 is unknown, and show the latter one is always at least as large as the former.

Solution. Since $\sum x_{i,1}^2 \sum x_{i,2}^2 \geq \sum x_{i,1}^2 \sum x_{i,2}^2 - \left(\sum x_{i,1}x_{i,2}\right)^2 \geq 0$, where the last inequality follows by Cauchy-Schwarz, it follows that

$$\frac{\sum x_{i,2}^2}{\sum x_{i,1}^2 \sum x_{i,2}^2 - \left(\sum x_{i,1}x_{i,2}\right)^2} \geq \frac{1}{\sum x_{i,1}^2},$$

and so, the variance lower bound in part (b), where β_2 is unknown, is greater than or equal to the variance lower bound in part (a), where β_2 is known to be zero. □

2. Suppose we observe the pair (X, Y), where X has a Poisson(λ) distribution and Y has a Bernoulli($\lambda/(1+\lambda)$) distribution, that is,

$$P_\lambda(X = j) = \frac{\lambda^j e^{-\lambda}}{j!}, j = 0, 1, 2, \ldots,$$

and

$$P_\lambda(Y = 1) = \frac{\lambda}{1+\lambda} = 1 - P_\lambda(Y = 0),$$

with X and Y independent, and $\lambda \in (0, \infty)$ unknown.

a) Find a one-dimensional sufficient statistic for λ based on (X, Y).
The joint density is, for $j = 0, 1, 2, \ldots$ and $k = 0, 1$,

$$P_\lambda(X = x, Y = y) = \frac{\lambda^x e^{-\lambda}}{x!}\left(\frac{\lambda}{1+\lambda}\right)^y \left(\frac{1}{1+\lambda}\right)^{1-y} = \frac{\lambda^{x+y}e^{-\lambda}}{x!(1+\lambda)} = \left(\frac{e^{-\lambda}}{1+\lambda}\lambda^{x+y}\right)\cdot\frac{1}{x!},$$

and so, by the Neyman-Fisher Factorization Theorem, $T(X, Y) = X + Y$ is a sufficient statistic for λ. Since the joint pmf is

$$\frac{1}{x!}\frac{e^{-\lambda}}{1+\lambda}\exp\left((x+y)\log\lambda\right),$$

and since $\{\log\lambda : \lambda > 0\}$ contains an open set in $\mathbb{R}$, we see that $\boxed{T(X, Y) = X + Y}$ is complete and sufficient.

b) Is there a UMVUE of λ? If so, find it.
Solution. Note that

$$P(Y = 1 \mid X + Y = k) = \frac{P(Y = 1, X + Y = k)}{P(X + Y = k)} = \frac{P(Y = 1, X = k-1)}{P(Y = 1, X = k-1) + P(Y = 0, X = k)}$$
$$= \frac{\left(\frac{e^{-\lambda}}{1+\lambda}\lambda^k\right)\frac{1}{(k-1)!}}{\left(\frac{e^{-\lambda}}{1+\lambda}\lambda^k\right)\frac{1}{(k-1)!} + \left(\frac{e^{-\lambda}}{1+\lambda}\lambda^k\right)\frac{1}{k!}} = \frac{k}{k+1},$$

and so,

$$\mathbb{E}(Y \mid X + Y) = P(Y = 1 \mid X + Y) = \frac{X+Y}{X+Y+1}.$$

Since $X = X + Y - Y$ is an unbiased estimator of λ and $X + Y$ is a complete and sufficient statistic, we see that

$$\mathbb{E}(X \mid X + Y) = \mathbb{E}(X + Y \mid X + Y) - \mathbb{E}(Y \mid X + Y) = X + Y - \frac{X+Y}{X+Y+1}$$
$$= \boxed{\frac{(X+Y)^2}{(X+Y)+1}}$$

is a UMVUE of λ.

c) Is there a UMVUE of $\lambda/(1 + \lambda)$? If so, find it.
Solution. Note that Y is an unbiased estimator of $\lambda/(1 + \lambda)$, and so, as we saw in part (b),

$$\mathbb{E}(Y \mid X + Y) = \boxed{\frac{X+Y}{(X+Y)+1}}$$

is a UMVUE of $\frac{\lambda}{1+\lambda}$.