1. *Suppose* $X_1, X_2, \ldots, X_m$ *and* $Y_1, Y_2, \ldots, Y_n$ *are independent random samples, respectively, from* $N(\mu_1, \sigma^2)$ *and* $N(\mu_2, \sigma^2)$, *where* $-\infty < \mu_1, \mu_2 < \infty$, $\sigma^2 > 0$.

   (a) *Does this model belong to the exponential family of distributions? Justify.*

   (b) *Find the minimal sufficient statistics for the unknown parameters. Is it complete?*

   (c) *Find the MLE and UMVUE of* σ^2.

Solution:

(a) Let $\boldsymbol{\theta} = (\mu_1, \mu_2, \sigma^2)$. The joint probability density function of $X_1, X_2, \ldots, X_m$ and $Y_1, Y_2, \ldots, Y_n$ is

$$f(\mathbf{x}, \mathbf{y}|\boldsymbol{\theta}) = \frac{1}{(\sqrt{2\pi})^{m+n}(\sigma^2)^{(m+n)/2}} exp\Big(-\frac{\sum_{i=1}^m (x_i - \mu_1)^2 + \sum_{j=1}^n (y_j - \mu_2)^2}{2\sigma^2}\Big), \text{ for } \mathbf{x} \in \mathbb{R}^m, \mathbf{y} \in \mathbb{R}^n.$$

The pdf can be written as

$$f(\mathbf{x}, \mathbf{y}|\boldsymbol{\theta}) = c(\boldsymbol{\theta}) h(\mathbf{x}, \mathbf{y}) exp\Big(\sum_{i=1}^3 w_i(\boldsymbol{\theta}) t_i(\mathbf{x}, \mathbf{y})\Big),$$

where

$$c(\boldsymbol{\theta}) = \frac{1}{(\sigma^2)^{(m+n)/2}} exp\Big(-\frac{m\mu_1^2}{\sigma^2} - \frac{n\mu_2}{2\sigma^2}\Big),$$

$$w_1(\boldsymbol{\theta}) = -\frac{1}{2\sigma^2} \qquad t_1(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^m x_i^2 + \sum_{j=1}^n y_j^2$$

$$w_2(\boldsymbol{\theta}) = \frac{\mu_1}{\sigma^2} \qquad t_2(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^m x_i$$

$$w_3(\boldsymbol{\theta}) = \frac{\mu_2}{\sigma^2} \qquad t_3(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^n y_j$$

$$h(\mathbf{x}, \mathbf{y}) = \frac{1}{(\sqrt{2\pi})^{m+n}}.$$

The distribution of $X_1, X_2, \ldots, X_m, Y_1, Y_2, \ldots, Y_n$ belongs to a 3-variate exponential family.

(b) From (a) and using the Factorization Theorem, $(\sum_{i=1}^m X_i^2 + \sum_{j=1}^n Y_j^2, \sum_{i=1}^m X_i, \sum_{j=1}^n Y_j)$ is sufficient for (μ_1, μ_2, σ^2). Next, we show that it is also a minimal sufficient statistic.
Let $(\mathbf{x}_1, \mathbf{y}_1)$ and $(\mathbf{x}_2, \mathbf{y}_2)$ denote two sample points. Then, the ratio

$$\frac{f(\mathbf{x}_1, \mathbf{y}_1|\boldsymbol{\theta})}{f(\mathbf{x}_2, \mathbf{y}_2|\boldsymbol{\theta})} = \frac{exp\Big(-[\sum_{i=1}^m x_{i,1}^2 + \sum_{j=1}^n y_{j,1}^2 - 2\mu_1 \sum_{i=1}^m x_{i,1} - 2\mu_2 \sum_{j=1}^n y_{j,1}]/2\sigma^2\Big)}{exp\Big(-[\sum_{i=1}^m x_{i,2}^2 + \sum_{j=1}^n y_{j,2}^2 - 2\mu_1 \sum_{i=1}^m x_{i,2} - 2\mu_2 \sum_{j=1}^n y_{j,2}]/2\sigma^2\Big)}$$

is independent of μ_1, μ_2, σ^2 if and only if $\sum_{i=1}^m x_{i,1}^2 + \sum_{j=1}^n y_{j,1}^2 = \sum_{i=1}^m x_{i,2}^2 + \sum_{j=1}^n y_{j,2}^2$, $\sum_{i=1}^m x_{i,1} = \sum_{i=1}^m x_{i,2}$ and $\sum_{j=1}^n y_{j,1} = \sum_{j=1}^n y_{j,2}$.

Hence, $(\sum_{i=1}^m X_i^2 + \sum_{j=1}^n Y_j^2, \sum_{i=1}^m X_i, \sum_{j=1}^n Y_j)$ is a minimal sufficient statistic for (μ_1, μ_2, σ^2).

The statistic $(\sum_{i=1}^m X_i^2 + \sum_{j=1}^n Y_j^2, \sum_{i=1}^m X_i, \sum_{j=1}^n Y_j)$ is complete if $\{(w_1(\boldsymbol{\theta}), w_2(\boldsymbol{\theta}), w_3(\boldsymbol{\theta})) : \boldsymbol{\theta} \in \boldsymbol{\Theta}\}$ contains an open set in $\mathbb{R}^3$. Here, $\boldsymbol{\Theta} = \mathbb{R}^2 \times (0, \infty)$. Hence, the statistic is also a complete sufficient statistic.

(c) The MLE for σ^2.

From (a), the log-likelihood for estimating $\boldsymbol{\theta}$ can be written as

$$L(\boldsymbol{\theta}|\mathbf{x}, \mathbf{y}) = log(f(\boldsymbol{\theta}|\mathbf{x}, \mathbf{y})) = C - \frac{\sum_{i=1}^m (x_i - \mu_1)^2 + \sum_{j=1}^n (y_j - \mu_2)^2}{2\sigma^2} - \frac{n+m}{2} log(\sigma^2),$$

where C is independent of $\boldsymbol{\theta}$. The partial derivatives, with respect to μ_1, μ_2 and σ^2 are

$$\frac{\partial L(\boldsymbol{\theta}|\mathbf{x}, \mathbf{y})}{\partial \mu_1} = \sum_{i=1}^m \frac{x_i - \mu_1}{\sigma^2}$$
$$\frac{\partial L(\boldsymbol{\theta}|\mathbf{x}, \mathbf{y})}{\partial \mu_2} = \sum_{j=1}^n \frac{y_j - \mu_2}{\sigma^2}$$
$$\frac{\partial L(\boldsymbol{\theta}|\mathbf{x}, \mathbf{y})}{\partial \sigma^2} = -\frac{n+m}{2\sigma^2} + \frac{1}{2\sigma^4} (\sum_{i=1}^m (x_i - \mu_1)^2 + \sum_{j=1}^n (y_j - \mu_2)^2).$$

Setting these partial derivatives to 0 and solving the equations yield the following solutions

$$\hat{\mu}_1 = \sum_{i=1}^m \frac{x_i}{m} = \bar{x}_m \qquad \hat{\mu}_2 = \sum_{j=1}^n \frac{y_j}{n} = \bar{y}_n$$
$$\hat{\sigma}^2 = \frac{1}{n+m} \Big(\sum_{i=1}^m (x_i - \bar{x}_m)^2 + \sum_{j=1}^n (y_j - \bar{y}_n)^2 \Big)$$

Next, to show that the $\hat{\sigma}^2$ is the MLE of σ^2.

For $\mu_1 \neq \bar{x}_m$, $\sum_{i=1}^m (x_i - \mu_1)^2 > \sum_{i=1}^m (x_i - \bar{x}_m)^2$. Similarly, for $\mu_2 \neq \bar{y}_n$, $\sum_{j=1}^m (y_j - \mu_2)^2 > \sum_{j=1}^n (y_j - \bar{y}_n)^2$. Hence, for any value of σ^2,

$$\frac{1}{(\sigma^2)^{(m+n)/2}} exp\Big(- \frac{\sum_{i=1}^m (x_i - \mu_1)^2 + \sum_{j=1}^n (y_j - \mu_2)^2}{2\sigma^2} \Big) \leq$$
$$\frac{1}{(\sigma^2)^{(m+n)/2}} exp\Big(- \frac{\sum_{i=1}^m (x_i - \bar{x}_m)^2 + \sum_{j=1}^n (y_j - \bar{y}_n)^2}{2\sigma^2} \Big).$$

From the above, we only need to show that $\frac{1}{(\sigma^2)^{(m+n)/2}} exp\Big(- (\sum_{i=1}^m (x_i - \bar{x}_m)^2 + \sum_{j=1}^n (y_j - \bar{y}_n)^2)/2\sigma^2 \Big)$ attains its maximum at $\hat{\sigma}^2$. Let

$$log(g(\sigma^2|\mathbf{x}, \mathbf{y})) = -\frac{\sum_{i=1}^m (x_i - \bar{x}_m)^2 + \sum_{j=1}^n (y_j - \bar{y}_n)^2}{2\sigma^2} - \frac{n+m}{2} log(\sigma^2).$$

Then, setting the derivative of this function with respect to σ^2 to 0, yields the unique solution $\hat{\sigma}^2$. Also,

$$\left.\frac{d^2 log(g(\sigma^2|\mathbf{x},\mathbf{y}))}{d(\sigma^2)^2}\right|_{\sigma^2=\hat{\sigma}^2} < 0.$$

Therefore, the MLE of σ^2 is $\hat{\sigma}^2$.
From (b) the statistic $(\sum_{i=1}^m X_i^2 + \sum_{j=1}^n Y_j^2, \sum_{i=1}^m X_i, \sum_{j=1}^n Y_j)$ is complete. To find the UMVUE for σ^2 we only need to look for an unbiased estimator for σ^2 based on the statistic. $\hat{\sigma}^2$ is the MLE for σ^2 and is based on $(\sum_{i=1}^m X_i^2 + \sum_{j=1}^n Y_j^2, \sum_{i=1}^m X_i, \sum_{j=1}^n Y_j)$. We check for unbiasedness.

$$E(\hat{\sigma}^2) = \frac{n+m-2}{n+m}$$

Hence, $(n+m)\hat{\sigma}^2/(n+m-2)$ is the UMVUE for σ^2.

□

2. (a) *Let U and V be two (jointly distributed) statistics such that U has finite variance. Show that*

$$Var(U) = Var(E(U|V)) + E(Var(U|V)).$$

(b) *Suppose $(X_1, X_2, \ldots, X_n)$ has probability distribution $P_\theta, \theta \in \Theta$. Let $\delta(X_1, X_2, \ldots, X_n)$ be an estimator of θ with finite variance. Suppose that T is sufficient for θ, and let $\delta^\star(t) = E(\delta(X_1, X_2, \ldots, X_n)|T = t)$, be the conditional expectation of $\delta(X_1, X_2, \ldots, X_n)$ given $T = t$. Then, arguing as in (a), and without applying Jensen's inequality, prove that*

$$E(\delta^\star(T) - \theta)^2 \le E(\delta(X_1, X_2, \ldots, X_n) - \theta)^2,$$

with strict inequality unless $\delta = \delta^\star$ (i.e., δ is already a function of T).

Solution:

(a)

$$\begin{aligned} Var(U) = E(U^2) - [E(U)]^2 &= E(E(U^2|V)) - E([E(U|V)]^2) + E([E(U|V)]^2) - [E(U)]^2 \\ &= E(Var(U|V)) + E([E(U|V)]^2) - [E(E(U|V))]^2 \\ &= E(Var(U|V)) + Var([E(U|V)]). \end{aligned}$$

(b) $E(\delta(X_1, X_2, \ldots, X_n)) = \theta$. In (a), put $U = \delta(X_1, X_2, \ldots, X_n)$ and $V = T$. Then,

$$\begin{aligned} E(\delta(X_1, X_2, \ldots, X_n) - \theta)^2 &= Var(E(\delta(X_1, X_2, \ldots, X_n)|T)) + E(Var(\delta(X_1, X_2, \ldots, X_n)|T)) \\ &\ge Var(E(\delta(X_1, X_2, \ldots, X_n)|T)), \end{aligned} \tag{1}$$

because $E(Var(\delta(X_1, X_2, \ldots, X_n)|T)) \ge 0$. As

$$Var(E(\delta(X_1, X_2, \ldots, X_n)|T)) = Var(\delta^\star(T)) = E(\delta^\star(T) - \theta)^2$$

we get

$$E(\delta(X_1, X_2, \ldots, X_n) - \theta)^2 \ge E(\delta^\star(T) - \theta)^2.$$

The equality exists if $E(Var(\delta(X_1, X_2, \ldots, X_n)|T)) = 0$, i.e.

$$\begin{aligned} &E(E[(\delta(X_1, X_2, \ldots, X_n) - \delta^\star(T))^2|T]) = 0 \\ \implies &\delta(X_1, X_2, \ldots, X_n) - E(\delta(X_1, X_2, \ldots, X_n)|T) = 0, \end{aligned}$$

i.e. δ is a function of T.

$\square$

3. *Suppose $X_1 \sim$ Binomial(n_1, p) which is independent of $X_2 \sim$ Binomial(n_2, p), where n_1 and n_2 are fixed and $0 < p < 1$.*

   (a) *What is the conditional distribution of X_1 given $X_1 + X_2 = k$?*

   (b) *Using (a) show that $X_1 + X_2$ sufficient for p.*

Solution: The distribution of X_i, $i = 1, 2$ is

$$P(X_i = x) = \binom{n_i}{x} p^x (1-p)^{n_i - x}, x = 0, 1, \ldots, n_i.$$

The distribution of $X_1 + X_2$ is

$$P(X_1+X_2 = k) = \sum_{j=0}^{k} P(X_1 = j)P(X_2 = k-j) = \sum_{j=0}^{k} \binom{n_1}{j}\binom{n_2}{k-j} p^k (1-p)^{n_1+n_2-k}, k = 0, 1, \ldots, n_1+n_2.$$

(a) The conditional distribution of X_1 given $X_1 + X_2 = k$ is

$$\begin{aligned} P(X_1 = x | X_1 + X_2 = k) &= \frac{P(X_1 = x, X_2 = k - x)}{P(X_1 + X_2 = k)} \\ &= \frac{\binom{n_1}{x}\binom{n_2}{k-x} p^k (1-p)^{n_1+n_2-k}}{\sum_{i=0}^{k} \binom{n_1}{i}\binom{n_2}{k-i} p^k (1-p)^{n_1+n_2-k}} \\ &= \frac{\binom{n_1}{x}\binom{n_2}{k-x}}{\sum_{i=0}^{k} \binom{n_1}{i}\binom{n_2}{k-i}}, x = 0, 1 \ldots, k. \end{aligned}$$

(b) The ratio of the joint distribution of X_1 and X_2, and the distribution of $X_1 + X_2$ is

$$\frac{P(X_1 = x_1, X_2 = x_2)}{P(X_1 + X_2 = t)} = \frac{\binom{n_1}{x_1}\binom{n_2}{x_2}}{\sum_{i=0}^{t} \binom{n_1}{i}\binom{n_2}{t-i}}, t = x_1 + x_2,$$

for all $x_i = 0, 1, \ldots, n_i$, $i = 1, 2$. As this ratio is independent of p, $X_1 + X_2$ is sufficient for p.

$\square$

4. *Let $X \sim$ Poisson(λ), $\lambda > 0$, and let $Y = 1$ when $X > 0$, and 0 otherwise.*

   (a) *Find the Fisher information on λ (say, $I^{(X)}(\lambda)$ and $I^{(Y)}(\lambda)$, respectively) contained in X and Y.*

   (b) *Compare $I^{(X)}(\lambda)$ and $I^{(Y)}(\lambda)$.*

Solution:

(a) The pmf of X is

$$f_\lambda(x|\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}, x = 0, 1, \ldots.$$

Y is a bernoulli random variable, with $P(Y = 1) = 1 - e^{-\lambda}$. The pmf of Y is

$$g_\lambda(y|\lambda) = (1 - e^{-\lambda})^y e^{-\lambda(1-y)}, y = 0, 1.$$

As

$$\frac{\partial log(f_\lambda(x|\lambda))}{\partial\lambda} = -1 + \frac{x}{\lambda},$$

$$I^{(X)}(\lambda) = E_\lambda\Big[-1 + \frac{X}{\lambda}\Big]^2 = \frac{1}{\lambda}.$$

As

$$\frac{\partial log(g_\lambda(y|\lambda))}{\partial\lambda} = -1 + \frac{y}{(1 - e^{-\lambda})},$$

$$I^{(Y)}(\lambda) = E_\lambda\Big[-1 + \frac{Y}{(1 - e^{-\lambda})}\Big]^2 = \frac{1}{e^\lambda - 1}.$$

(b) From (a) and that $e^\lambda - 1 > \lambda$ for all $\lambda > 0$, $I^{(Y)}(\lambda) < I^{(X)}(\lambda)$ for all $\lambda > 0$.

□