1. For observations $Y_1, \cdots, Y_n$, consider the linear model

$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i, i = 1, \cdots, n,$$

where x_i is the value of a co-variate corresponding to Y_i and ϵ_i are i.i.d. errors having the $N(0, \sigma^2)$ distribution. Here β_0, β_1 and $\sigma^2 > 0$ are unknown parameters and x_i are treated as known constants. Also $-\infty < \beta_0 < \infty$, $-\infty < \beta_1 < \infty$.

(a) Show that the distribution of $Y_1, \cdots, Y_n$ belongs to k-variate exponential family. Find k.

(b) Use properties of exponential family to set up equations and solve them to find MLE of $(\beta_0, \beta_1, \sigma^2)$.

Solution:

(a) Let $\boldsymbol{\theta} = (\beta_0, \beta_1, \sigma^2)$. The probability density function of $Y_1, \cdots, Y_n$ can be written as

$$\begin{aligned} f(y_1, \cdots, y_n|\boldsymbol{\theta}) &= \frac{1}{(2\pi)^{n/2}\sigma^n} exp\big(- \frac{\sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2}{2\sigma^2}\big), \text{ for } y_1, \cdots, y_n \in \mathbb{R} \\ &= exp\big(\sum_{i=1}^3 \eta_i(\boldsymbol{\theta}) t_i(y_1, \cdots, y_n)\big) c(\boldsymbol{\theta}) h(y_1, \cdots, y_n), \qquad (1) \end{aligned}$$

where

$$\begin{aligned} \eta_1(\boldsymbol{\theta}) &= -\frac{1}{2\sigma^2}, \qquad t_1(y_1, \cdots, y_n) = \sum_{i=1}^n y_i^2, \\ \eta_2(\boldsymbol{\theta}) &= \frac{\beta_0}{\sigma^2}, \qquad t_2(y_1, \cdots, y_n) = \sum_{i=1}^n y_i, \\ \eta_3(\boldsymbol{\theta}) &= \frac{\beta_1}{\sigma^2}, \qquad t_3(y_1, \cdots, y_n) = \sum_{i=1}^n x_i y_i, \\ c(\boldsymbol{\theta}) &= \frac{1}{\sigma^n} exp\big(- \frac{n\beta_0^2}{2\sigma^2} - \frac{2\beta_0\beta_1 \sum_{i=1}^n x_i}{2\sigma^2} - \frac{\beta_1^2 \sum_{i=1}^n x_i^2}{2\sigma^2}\big) \text{ and} \\ h(y_1, \cdots, y_n) &= \frac{1}{(2\pi)^{n/2}}. \end{aligned}$$

The distribution of $Y_1, \cdots, Y_n$ belongs to 3-variate exponential family.

(b) Using (a), the log-likelihood is

$$\begin{aligned} log f(\boldsymbol{\theta}|y_1, \cdots, y_n) = &\sum_{i=1}^3 \eta_i(\boldsymbol{\theta}) t_i(y_1, \cdots, y_n) - \frac{n\beta_0^2}{2\sigma^2} - \frac{2\beta_0\beta_1 \sum_{i=1}^n x_i}{2\sigma^2} \\ &- \frac{\beta_1^2 \sum_{i=1}^n x_i^2}{2\sigma^2} - \frac{n}{2} log(\sigma^2) + H(y_1, \cdots, y_n), \end{aligned}$$

where $H(y_1,\cdots,y_n)=log(h(y_1,\cdots,y_n))$ is independent of $\boldsymbol{\theta}$.
Taking the first order partial derivatives of the log-likelihood with respect to β_0, β_1 and σ^2, and equating them to 0, we get the following equations (under the assumption $\sigma^2>0$):

$$\sum_{i=1}^{n} y_i x_i - \beta_1 \sum_{i=1}^{n} x_i^2 - \beta_0 \sum_{i=1}^{n} x_i = 0. \tag{2}$$

$$\sum_{i=1}^{n} y_i - \beta_0 - \beta_1 \sum_{i=1}^{n} x_i = 0. \tag{3}$$

$$\frac{\sum_{i=1}^n y_i^2}{2(\sigma^2)^2} - \frac{\beta_0\sum_{i=1}^n y_i}{(\sigma^2)^2} - \frac{\beta_1\sum_{i=1}^n y_i x_i}{(\sigma^2)^2} + \frac{n\beta_0^2}{2(\sigma^2)^2} + \frac{\beta_0\beta_1\sum_{i=1}^n x_i}{(\sigma^2)^2} + \frac{\beta_1^2\sum_{i=1}^n x_i^2}{2(\sigma^2)^2} - \frac{n}{2(\sigma^2)} = 0. \tag{4}$$

Define

$$S_{xy} = \sum_{i=1}^{n}(x_i-\bar{x}_n)(y_i-\bar{y}_n) \text{ and } S_{xx} = \sum_{i=1}^{n}(x_i-\bar{x}_n)^2.$$

Let $\hat{\beta}_0$, $\hat{\beta}_1$ and $\hat{\sigma}^2$ be the solutions to $(2)-(4)$. Then,

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}, \hat{\beta}_0 = \bar{y}_n - \hat{\beta}_1\bar{x}_n \text{ and } \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{\beta}_0-\hat{\beta}_1x_i)^2.$$

Next, to show that the $(\hat{\beta}_1,\hat{\beta}_0,\hat{\sigma}^2)$ is the MLE of $(\beta_0,\beta_1,\sigma^2)$.
Define $S(\beta_0,\beta_1)=\sum_{i=1}^n(y_i-\beta_0-\beta_1x_i)^2$. The function $S(\beta_0,\beta_1)$ is quadractic in β_0 and β_1. Taking the first order partial derivative of $S(\beta_0,\beta_1)$ with respect β_0 and β_1 and equating them to 0 yields the solutions $\hat{\beta}_0$ and $\hat{\beta}_1$. The Hessian matrix

$$\begin{bmatrix} 2n & 2\sum_{i=1}^n x_i \\ 2\sum_{i=1}^n x_i & 2\sum_{i=1}^n x_i^2 \end{bmatrix}$$

is positive definitive. Therefore, $S(\beta_0,\beta_1)$ attains its global minimun at $(\beta_0,\beta_1)=(\hat{\beta}_0,\hat{\beta}_1)$. Hence, for any value of σ^2,

$$\frac{1}{(\sigma^2)^{n/2}}exp\Big(-\frac{\sum_{i=1}^n(y_i-\beta_0-\beta_1x_i)^2}{2\sigma^2}\Big) \le \frac{1}{(\sigma^2)^{n/2}}exp\Big(-\frac{\sum_{i=1}^n(y_i-\hat{\beta}_0-\hat{\beta}_1x_i)^2}{2\sigma^2}\Big).$$

From the above, we only need to show that $\frac{1}{(\sigma^2)^{n/2}}exp\Big(-\frac{\sum_{i=1}^n(y_i-\hat{\beta}_0-\hat{\beta}_1x_i)^2}{2\sigma^2}\Big)$ attains its maximum at $\hat{\sigma}^2$. Let

$$log(g(\sigma^2|y_1,\ldots,y_n)) = \Big(-\frac{\sum_{i=1}^n(y_i-\hat{\beta}_0-\hat{\beta}_1x_i)^2}{2\sigma^2}\Big) - \frac{n}{2}log(\sigma^2).$$

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

$$\frac{d^2log(g(\sigma^2|y_1,\ldots,y_n))}{d(\sigma^2)^2}\Big|_{\sigma^2=\hat{\sigma}^2} < 0.$$

Hence, the MLE of σ^2 is $\hat{\sigma}^2$.
The MLEs of β_1, β_0 and σ^2 are $\hat{\beta}_1$, $\hat{\beta}_0$ and $\hat{\sigma}^2$, respectively.

$\square$

2. Consider a family of regular models with density $f(x|\theta)$ such that $f(x|\theta) > 0$ for all $\theta \in \Theta$ and for all $x \in \chi$. Suppose $T(X)$ is sufficient for this family of distributions indexed by θ. Show that if $T(x) = T(y)$ for two sample points x and y then $\frac{f(x|\theta)}{f(y|\theta)}$ is free of θ.

   Solution: As $T(X)$ is a sufficient statistic, using the Factorization Theorem, there exist functions $g(\cdot|\theta)$ and $h(x)$, such that

   $$f(x|\theta) = g(T(x)|\theta)h(x), \text{ for all sample points } x \text{ and for all } \theta \in \Theta.$$

   Then, for the two sample points x, y with $T(x) = T(y)$,

   $$\frac{f(x|\theta)}{f(y|\theta)} = \frac{g(T(x)|\theta)h(x)}{g(T(y)|\theta)h(y)} = \frac{h(x)}{h(y)}$$

   is free of θ. $\square$

3. Roll a balanced six-faced die and let N denote the number of dots that show up. Having observed $N = n$, perform n Bernoulli(θ) trials, getting X successes.

   (a) Find a minimal sufficient statistic for θ.

   (b) Show that minimal sufficient statistic is not complete in this case.

   Solution: The pmf of (X, N) is

   $$f(x, n|\theta) = \frac{\binom{n}{x}\theta^x(1-\theta)^{n-x}}{6}, \text{ for } x = 0, \cdots, n \text{ and } n = 1, \cdots, 6. \tag{5}$$

   (a) The ratio $(0 < \theta < 1)$

   $$\frac{f(x, n_1|\theta)}{f(y, n_2|\theta)} = \frac{\binom{n_1}{x}\theta^x(1-\theta)^{n_1-x}}{\binom{n_2}{y}\theta^y(1-\theta)^{n_2-y}}, \tag{6}$$

   is a constant as a function of θ if and only if $x = y$ and $n_1 = n_2$. Hence, the minimal sufficient statistic for θ is (X, N).

   (b) N is an ancilliary statistic. The minimal sufficient statistic (X, N) and N are not independent. Define $g(X, N) = N - 7/2$. Then,

   $$E(g(X, N)) = E(N - \frac{7}{2}) = 0 \text{ for all } \theta, \text{ but } P(N - \frac{7}{2} = 0) = 0 \neq 1.$$

   Hence, the minimal sufficient statistic (X, N) is not complete.

$\square$

4. Suppose $X_1, X_2, \cdots, X_n$ are i.i.d. observations from Exponential(λ) (with density proportional to $exp(-\lambda x)$), where $n \geq 3$ and $0 < \lambda$.

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

   (b) Find the UMVUE of λ.

   (c) Does the UMVUE attain the C-R lower bound?

Solution: The joint probability density function of $X_1, \cdots, X_n$ is

$$f(x_1, \cdots, x_n|\lambda) = exp\big(-\lambda\sum_{i=1}^{n} x_i\big)\lambda^n, \text{ for } x_i > 0, i = 1, \cdots n.$$

(a) The distribution of $X_1, \cdots, X_n$ belongs to an exponential family. The joint distribution can be written as

$$f(x_1, \cdots, x_n|\lambda) = exp\big(\eta(\lambda)t(x_1, \cdots, x_n)\big)c(\lambda)I(x_i > 0, i = 1, \cdots, n), \tag{7}$$

where

$$\eta(\lambda) = \lambda, t(x_1, \cdots, x_n) = -\sum_{i=1}^{n} x_i, c(\lambda) = \lambda^n, \text{ and } h(x_1, \cdots, x_n) = I(x_i > 0, i = 1, \cdots, n).$$

(b) Using the Factorization Theorem, $\sum_{i=1}^{n} X_i$ is a sufficient statistic for λ. $T = \sum_{i=1}^{n} X_i$ follows a Gamma distribution, with pdf as

$$g_\lambda(t|\lambda) = \frac{1}{\Gamma(n)}\lambda^n t^{n-1} exp(-t\lambda), \text{ for } t > 0.$$

$$E_\lambda\big(\frac{1}{T}\big) = \int_0^\infty \frac{1}{t}\frac{1}{\Gamma(n)}\lambda^n t^{n-1} exp(-t\lambda)dt = \frac{\lambda}{n-1}.$$

Hence, an unbiased estimator for λ is $(n-1)/(\sum_{i=1}^{n} X_i)$. Using the Rao-Blackwell Theorem, $(n-1)/(\sum_{i=1}^{n} X_i)$ is also the UMVUE.

(c) Using (a), the C-R lower bound for an unbiased estimator of λ is

$$\frac{1}{-E_\lambda\Big(\frac{\partial^2}{\partial\lambda^2}\big(-\lambda\sum_{i=1}^{n} x_i + nlog(\lambda)\big)\Big)} = \frac{\lambda^2}{n}. \tag{8}$$

The variance of the UMVUE, $(n-1)/(\sum_{i=1}^{n} X_i)$, is $\lambda^2/(n-2)$. The UMVUE does not attain the C-R lower bound.

□

5. Consider a random sample $X_1, X_2, \ldots, X_n$ from $U(0, \theta)$, where $\theta > 0$.

   (a) Construct a 95% confidence interval for θ which has the form: $[X_{(n)}, X_{(n)}/c]$ for some constant c.

   (b) If the confidence interval constructed from observed data is the interval $[1, 12.5]$, how will you interpret it?

Solution: The random sample $X_1, \cdots, X_n$ is from $U(0, \theta)$, $\theta > 0$.

(a) The pdf of $X_{(n)}/\theta$ is

$$f(x) = nx^{n-1}, \text{ for } 0 \le x \le 1. \tag{9}$$

We need to find c, such that $P[X_{(n)} \le \theta \le \frac{X_{(n)}}{c}] = 0.95$.

$$P[X_{(n)} \le \theta \le \frac{X_{(n)}}{c}] = P[c \le \frac{X_{(n)}}{\theta} \le 1] = 1 - c^n.$$

Choosing $c = 0.05^{1/n}$, we get $P[X_{(n)} \le \theta \le \frac{X_{(n)}}{c}] = 0.95$.

(b) While $X_{(n)} \le \theta \le \frac{X_{(n)}}{c}$ is an interval estimator for θ, $[1, 12.5]$ is the interval estimate for θ. The observed interval $[1, 12.5]$ contains the true value of θ with 95% confidence.

$\square$