Bios 661: 1 – 5; Bios 673: 2 – 6.

1. C&B 6.22
2. C&B 7.2(a)
3. C&B 7.6
4. Suppose that $Y_x \sim N(x\mu, x^3\sigma^2)$, $x = 1, 2, \ldots, n$, and assume that $\{Y_1, Y_2, \ldots, Y_n\}$ constitute a set of n mutually independent random variables, and that σ^2 is a known positive constant.

   (a) Derive the method of moments estimator $\hat{\mu}_1$;

   Solution: Method of moment matching in non-iid data can be done by matching the sample moment to its own expectation. In this question, since

$$E\left(\frac{1}{n}\sum_{x=1}^{n} Y_x\right) = \frac{1}{n}\sum_{x=1}^{n} x\mu = \frac{(n+1)}{2}\mu,$$

   we may claim

$$\hat{\mu}_1 = \frac{2}{n(n+1)}\sum_{x=1}^{n} Y_x,$$

   by matching the first sample moment $n^{-1}\sum_{x=1}^{n} Y_x$ to its own expectation.

   (b) Derive the maximum likelihood estimator $\hat{\mu}_2$.

   Solution: The log-likelihood function:

$$\ell(\mu|\boldsymbol{y}) \propto -\frac{1}{2\sigma^2}\sum_{x=1}^{n} x^{-3}(y_x - x\mu)^2.$$

   Thus, set

$$\frac{\partial}{\partial\mu}\ell(\mu|\boldsymbol{y}) = \frac{1}{\sigma^2}\sum_{x=1}^{n} x^{-3}(y_x - x\mu)x = 0.$$

   We have

$$\hat{\mu}_2 = \frac{\sum_{x=1}^{n} x^{-2}Y_x}{\sum_{x=1}^{n} x^{-1}}.$$

(c) Determine the exact distribution of $\hat{\mu}_1$ and $\hat{\mu}_2$.

Solution:

$$E(\hat{\mu}_1) = \frac{2}{n(n+1)}\sum_{x=1}^{n} E(Y_x) = \frac{2}{n(n+1)}\sum_{x=1}^{n} x\mu = \mu.$$

$$Var(\hat{\mu}_1) = \frac{4}{n^2(n+1)^2}\sum_{x=1}^{n} Var(Y_x) = \frac{4}{n^2(n+1)^2}\sum_{x=1}^{n} x^3\sigma^2 = \sigma^2.$$

$$E(\hat{\mu}_2) = \frac{\sum_{x=1}^n x^{-2}E(Y_x)}{\sum_{x=1}^n x^{-1}} = \frac{\sum_{x=1}^n x^{-2}x\mu}{\sum_{x=1}^n x^{-1}} = \mu.$$

$$Var(\hat{\mu}_2) = \frac{\sum_{x=1}^n x^{-4}Var(Y_x)}{(\sum_{x=1}^n x^{-1})^2} = \frac{\sum_{x=1}^n x^{-4}x^3\sigma^2}{(\sum_{x=1}^n x^{-1})^2} = \frac{\sigma^2}{\sum_{x=1}^n x^{-1}}.$$

Both $\hat{\mu}_1$ and $\hat{\mu}_2$ are linear combination of normal variables and hence are normal variables. We can claim $\hat{\mu}_1 \sim N(\mu, \sigma^2)$ and $\hat{\mu}_2 \sim N(\mu, \sigma^2/\sum_{x=1}^n x^{-1})$.

(d) Which one has a smaller variance?

Solution: Both estimator are unbiased, but since $\sum_{x=1}^n x^{-1} > 1$, one would prefer $\hat{\mu}_2$ because of smaller variance.

5. Suppose that the random variables $Y_1, \ldots, Y_n$, $n > 2$ are independent and normally distributed with $E(Y_i) = \theta x_i$, where $x_1, \ldots, x_n$ are known non-zero constants, and $\text{var}(Y_i) = \sigma^2$. Both $\theta \in \mathbb{R}$ and $\sigma^2 \in (0, \infty)$ are unknown.

   (a) Find a two-dimensional sufficient statistic for (θ, σ^2).

   Solution: One can write the joint pdf as

$$\begin{aligned} f(\boldsymbol{y}|\theta, \sigma^2) &= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left\{-\frac{\sum_{i=1}^n (y_i - \theta x_i)^2}{2\sigma^2}\right\} \\ &= \left(2\pi\sigma^2\right)^{-n/2} \exp\left(-\frac{\theta^2}{2\sigma^2}\sum_{i=1}^n x_i^2\right) \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n y_i^2 + \frac{\theta}{\sigma^2}\sum_{i=1}^n y_i x_i\right). \end{aligned}$$

   Hence, one can claim $(\sum_{i=1}^n Y_i^2, \sum_{i=1}^n Y_i x_i)$ are a two-dimensional sufficient statistic.

(b) With σ^2 fixed (in some sense as known), find the MLE $\hat{\theta}$ of θ and show that it is an unbiased estimator of θ.

Solution: The likelihood function $L(\theta)$ is proportional to

$$\exp\left\{-\frac{\sum_{i=1}^n (y_i - \theta x_i)^2}{2\sigma^2}\right\}.$$

Maximizing the function above is equivalent to minimizing the function

$$Q(\theta) = \sum_{i=1}^n (y_i - \theta x_i)^2.$$

One can easily see that

$$\frac{\partial Q(\theta)}{\partial \theta} = 2\sum_{i=1}^n (y_i - \theta x_i)(-x_i).$$

Setting $\partial Q(\theta)/\partial\theta = 0$, we can get $\hat{\theta} = \sum_{i=1}^n x_i Y_i / \sum_{i=1}^n x_i^2$. The second derivative of $Q(\theta)$ is

$$\frac{\partial^2 Q(\theta)}{\partial \theta^2} = 2\sum_{i=1}^n x_i^2,$$

which is apparently positive, so $\hat{\theta}$ is a minimizer of $Q(\theta)$.

(c) Find the distribution of the MLE $\hat{\theta}$.

Solution: One can see $\hat{\theta}$ is a linear combination of Y_i, $i = 1, \ldots, n$, so it follows a normal distribution with mean

$$E(\hat{\theta}) = \sum_{i=1}^n x_i E(Y_i) / \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i^2 \theta / \sum_{i=1}^n x_i^2 = \theta,$$

and variance

$$\text{var}(\hat{\theta}) = \sum_{i=1}^n x_i^2 \sigma^2 / (\sum_{i=1}^n x_i^2)^2 = \sigma^2 / \sum_{i=1}^n x_i^2.$$

(d) With θ fixed at the MLE $\hat{\theta}$ in (b), find the MLE $\hat{\sigma}^2$ of σ^2.

Solution: With θ fixed, the likelihood function can be written as

$$L(\sigma^2) = \left(2\pi\sigma^2\right)^{-n/2} \exp\left(-\frac{Q(\theta)}{2\sigma^2}\right),$$

with log-likelihood function

$$\ell(\sigma^2) = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{Q(\theta)}{2\sigma^2}.$$

Taking the first derivative

$$\frac{\partial}{\partial\sigma^2}\ell(\sigma^2) = -\frac{n}{2}\frac{2\pi}{2\pi\sigma^2} - Q(\theta)\frac{1}{2\sigma^4}(-1),$$

and set it as zero, one have MLE of σ^2 as

$$\hat{\sigma}^2 = \frac{1}{n}Q(\theta) = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \theta x_i)^2.$$

(e) When both θ and σ^2 are unknown, the MLE of (θ, σ^2) are $\hat{\theta}$ in (b) and $\hat{\sigma}_e^2 = n^{-1}\sum_{i=1}^{n}(Y_i - \hat{\theta}x_i)^2$, which is $\hat{\sigma}^2$ in (d) but with θ replaced by $\hat{\theta}$. Show that $\hat{\theta}$ and $\hat{\sigma}_e^2$ are independent.

Solution: Let $R_i = Y_i - \hat{\theta}x_i$. One can show that

$$\begin{aligned}
\mathrm{Cov}(\hat{\theta}, R_i) &= \mathrm{Cov}(\hat{\theta}, Y_i - \hat{\theta}x_i) \\
&= \mathrm{Cov}(\hat{\theta}, Y_i) - \mathrm{var}(\hat{\theta})x_i \\
&= \frac{x_i\sigma^2}{\sum_{i=1}^{n} x_i} - \frac{x_i\sigma^2}{\sum_{i=1}^{n} x_i} \\
&= 0.
\end{aligned}$$

Since both Y_i and R_i are normally distributed, one can claim Y_i and R_i are independent for $i = 1, \ldots, n$. So as Y_i and $\sigma_e^2 = n^{-1}\sum_{i=1}^{n} R_i^2$.

6. Suppose $X_1, \ldots, X_n$ are iid with pdf $f(x|\theta) = h(x)c(\theta)\exp\{\theta t(x)\}$ (an exponential family with $w(\theta) = \theta$).

(a) Show that $E\{T(X)\} = -c'(\theta)/c(\theta)$.

Solution: The moment generating function of $T(X)$ is

$$\begin{aligned}
M_{T(x)}(s) = E[\exp\{st(x)\}] &= \int \exp\{st(x)\}h(x)c(\theta)\exp\{\theta t(x)\}dx \\
&= \frac{c(\theta)}{c(s+\theta)}\int h(x)c(s+\theta)\exp\{(s+\theta)t(x)\}dx \\
&= \frac{c(\theta)}{c(s+\theta)}.
\end{aligned}$$

Using the moment generating function, one can derive the expectation as

$$E\{T(x)\} = \frac{d}{ds}M_{T(x)}(s)\,|_{s=0} = c(\theta)(-1)c(\theta)^{-2}c'(\theta) = -\frac{c'(\theta)}{c(\theta)}.$$

Similarly,

$$E\{T(x)^2\} = \frac{d^2}{ds^2}M_{T(x)}(s)\,|_{s=0} = 2\frac{c'(\theta)^2}{c(\theta)^2} - \frac{c''(\theta)}{c(\theta)},$$

and

$$\text{var}\{T(x)\} = E\{T(x)^2\} - E\{T(x)\}^2 = \frac{c'(\theta)^2}{c(\theta)^2} - \frac{c''(\theta)}{c(\theta)}.$$

Another way of deriving this result starts with

$$1 = \int f(x)dx = \int h(x)c(\theta)\exp\{\theta t(x)\}dx.$$

Differentiating both sides with respect to θ, one has

$$\begin{aligned}0 &= \int h(x)c'(\theta)\exp\{\theta t(x)\}dx + \int t(x)h(x)c(\theta)\exp\{\theta t(x)\}dx \\ &= \frac{c'(\theta)}{c(\theta)} + E\{T(x)\}.\end{aligned}$$

It apparently leads to the result. Notice that we use the assumption that differentiation and integral are exchangeable, which stands for exponential family.

(b) Show that the MLE for θ has to satisfy $E\{\sum_{i=1}^n T(X_i)\} = \sum_{i=1}^n T(X_i)$. That is, the equation holds when one plugs in $\theta = \hat{\theta}$ in the left hand side of the equation.

Solution: The log-likelihood function can be written as

$$\ell(\theta) = \sum_{i=1}^n \log\{h(x_i)\} + n\log\{c(\theta)\} + \theta\sum_{i=1}^n t(x_i).$$

Taking partial derivative to the log-likelihood function, one has

$$n\frac{c'(\theta)}{c(\theta)} + \sum_{i=1}^n t(x_i) = 0.$$

Since $-c'(\theta)/c(\theta) = E\{T(x_i)\}$ for each $i = 1, \ldots, n$, one can conclude the MLE for θ has to satisfy

$$E\left\{\sum_{i=1}^n T(x_i)\right\} = \sum_{i=1}^n T(x_i).$$

The second derivative of the log-likelihood function equals

$$n\frac{c''(\theta)c(\theta) - c'(\theta)^2}{c(\theta)^2},$$

which is negative for all θ since $c''(\theta)/c(\theta) - c'(\theta)^2/c(\theta)^2 = -\mathrm{var}\{T(x)\} < 0$.

(c) Suppose $f(x|\beta) = \beta^{-1}\exp(-x/\beta)$ (exponential distribution with mean β). Use the result in (b) to find the MLE for β.

Solution: Let $\theta = \beta^{-1}$, $c(\theta) = \theta$, and $t(x) = -x$. By the result in (b), the MLE of θ has to satisfy

$$E\left\{\sum_{i=1}^n T(X_i)\right\} = -n\theta^{-1} = -\sum_{i=1}^n X_i.$$

One has the MLE for θ as $\hat{\theta} = n/\sum_{i=1}^n X_i = \bar{X}^{-1}$. The MLE for β equals $\bar{X}$.

7. (Bios 673 class material, C&B 7.17) The same result holds for the general exponential family

$$f(x|\theta) = h(x)c(\theta)\exp\left\{\sum_{j=1}^k w_j(\theta)t_j(x)\right\},$$

where $\theta = (\theta_1, \ldots, \theta_k)'$. Suppose that the two-dimensional vectors $(X_1, Y_1), \ldots, (X_n, Y_n)$ follows a bivariate normal distribution

$$N\left(\left(\begin{array}{c}\mu_x\\ \mu_y\end{array}\right), \left(\begin{array}{cc}\sigma_x^2 & \rho\sigma_x\sigma_y\\ \rho\sigma_x\sigma_y & \sigma_y^2\end{array}\right)\right).$$

Find the MLE for $\theta = (\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)'$ using the result in the previous question (You may find it is much easier than using partial derivatives).

(a) Show that the joint pdf of (X, Y) can be written as

$$f(x, y|\theta) = h(x, y)c(\theta)\exp\left\{\sum_{j=1}^k w_j(\theta)t_j(x, y)\right\},$$

where $\theta = (\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)'$ and

$$t_1(x, y) = x^2,\ \ t_2(x, y) = x,\ \ t_3(x, y) = y^2,\ \ t_4(x, y) = y,\ \ t_5(x, y) = xy.$$

Solution: The pdf for the pair of normal random variables is

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \times$$
$$\exp\left[-\frac{1}{2(1-\rho^2)}\left\{\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right) + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right\}\right].$$

It is not difficult to see that the pdf belongs to an exponential family and

$$t_1(x,y) = x^2,\ \ t_2(x,y) = x,\ \ t_3(x,y) = y^2,\ \ t_4(x,y) = y,\ \ t_5(x,y) = xy.$$

(b) Find the MLE using the constraints $E\{\sum_{i=1}^n T_j(X_i,Y_i)\} = \sum_{i=1}^n T_j(X_i,Y_i)$.

Solution: By the result we have, the MLE has to satisfy

$$E\left(\sum_{i=1}^n X_i^2\right) = \sum_{i=1}^n X_i^2,$$
$$E\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n X_i,$$
$$E\left(\sum_{i=1}^n Y_i^2\right) = \sum_{i=1}^n Y_i^2,$$
$$E\left(\sum_{i=1}^n Y_i\right) = \sum_{i=1}^n Y_i,$$
$$E\left(\sum_{i=1}^n X_iY_i\right) = \sum_{i=1}^n X_iY_i.$$

One can get $\hat{\mu}_x = \bar{X}$, $\hat{\sigma}_x^2 = n^{-1}\sum_{i=1}^n (X_i-\bar{X})^2$, $\hat{\mu}_y = \bar{Y}$, $\hat{\sigma}_y^2 = n^{-1}\sum_{i=1}^n (Y_i-\bar{Y})^2$, and

$$\hat{\rho} = \frac{\sum_{i=1}^n (X_i-\bar{X})(Y_i-\bar{Y})}{\sqrt{\sum_{i=1}^n (X_i-\bar{X})^2 \sum_{i=1}^n (Y_i-\bar{Y})^2}}.$$