Mathematical Statistics I
Assignment 5 - Solutions

1. Note that $X_1, X_2, \ldots, X_n$ are identically distributed and therefore have the same conditional expected value, so we have, by linearity of conditional expectation,

$$E\left(X_1 + 2X_2 + 3X_3 \mid \sum_i X_i\right) = 6E\left(X_1 \mid \sum_i X_i\right)$$

Hence it suffices to find the conditional distribution of X_1 given $\sum_{i=1}^n X_i = x$. Assuming $x \geq x_1$ (otherwise the following probability is 0) we have,

$$\begin{aligned} P\left(X_1 = x_1 \mid \sum_{i=1}^n X_i = x\right) &= \frac{P\left(X_1 = x_1, \sum_{i=1}^n X_i = x\right)}{P\left(\sum_{i=1}^n X_i = x\right)} \\ &= \frac{P\left(X_1 = x_1, \sum_{i=2}^n X_i = x - x_1\right)}{P\left(\sum_{i=1}^n X_i = x\right)} \\ &= \frac{P\left(X_1 = x_1\right)\left(\sum_{i=2}^n X_i = x - x_1\right)}{P\left(\sum_{i=1}^n X_i = x\right)} \\ &= \binom{x}{x_1}\left(\frac{1}{n}\right)^{x_1}\left(1 - \frac{1}{n}\right)^{x - x_1}, \end{aligned}$$

where the last equality comes from the fact that $\sum_{i=1}^n X_i \sim P(n\theta)$ and $\sum_{i=2}^n X_i \sim P((n-1)\theta)$. Thus, the conditional distribution is binomial$(x, 1/n)$ and therefore $E\left[X_1 \mid \sum_{i=1}^n X_i = x\right] = x/n$. Finally we have,

$$E\left(X_1 + 2X_2 + 3X_3 \mid \sum_i X_i\right) = 6\sum_i X_i/n = 6\overline{X}.$$

2. Let $Y_i = 1$ if the ith observation is less than 0 and $Y_i = 0$ otherwise. Then $Y_1, \ldots, Y_n$ are the actual observations. Let $p = P\left(Y_i = 1\right) = \Phi(-\mu)$, where Φ is the cumulative distribution function of $N(0,1)$, $\ell(p)$ be the likelihood function in p, and $T = \sum_{i=1}^n Y_i$. Then

$$\frac{\partial \log \ell(p)}{\partial p} = \frac{T}{p} - \frac{n-T}{1-p}$$

The likelihood equation has a unique solution T/n. Hence the MLE of p is T/n. Then, the MLE of μ is $-\Phi^{-1}(T/n)$.

3. Let $Y_1 = X_1 + X_2$ and $Y_2 = X_2$, therefore $X_1 = Y_1 - Y_2$, $X_2 = Y_2$ and with the Jacobian determinant being 1, the joint probability density function is,

$$f_{Y_1,Y_2}(y_1, y_2) = f_{X_1,X_2}(y_1 - y_2, y_2) = f_{X_1}(y_1 - y_2) f_{X_2}(y_2) = \frac{1}{\theta^2} e^{-y_1/\theta}, \quad 0 < y_2 < y_1.$$

Considering Y_2 separately, the pdf is,

$$f_{Y_2}(y_2) = \frac{1}{\theta} e^{-y_2/\theta}, \quad y_2 > 0,\, \theta > 0.$$

Therefore $Y_2 \sim \Gamma(1, \theta)$ and we have $E\left(Y_2\right) = \theta$ and $Var\left(Y_2\right) = \theta^2$. We showed that Y_2 is indeed an unbiased estimator of θ with variance θ^2. The probability density function of $Y_1 = X_1 + X_2$ is,

$$f_{Y_1}\left(y_1\right) = \int_0^{y_1} f_{Y_1,Y_2}(y_1, y_2)\, dy_2 = \frac{y_1}{\theta^2} e^{-y_1/\theta}, \quad y_1 > 0.$$

Hence, the conditional probability density function is,

$$f_{Y_2|Y_1=y_1}\left(y_2, y_1\right) = \frac{f_{Y_1,Y_2}(y_1, y_2)}{f_{Y_1}(y_1)} = \frac{1}{y_1}, \quad 0 < y_2 < y_1.$$

Now we have,

$$\varphi(y_1) = E\left(Y_2 \mid Y_1 = y_1\right) = \int_0^{y_1} y_2 f_{Y_2|Y_1=y_1}\left(y_2, y_1\right)\, dy_2 = y_1/2.$$

Finally we have,

$$\mathrm{Var}(\varphi\left(Y_1\right)) = \frac{1}{4}\mathrm{Var}(Y_1) = \frac{1}{4}\left(\mathrm{Var}(X_1) + \mathrm{Var}(X_2)\right) = \theta^2/2.$$

4. The marginal pdf of Y is given by,

$$f_Y(y) = \int_0^y f_{X,Y}(x, y)\, dx = \frac{2}{\theta} e^{-y/\theta}(1 - e^{-y/\theta}), \quad y > 0.$$

Thus we have,

$$\begin{aligned}
E(Y) &= \int_0^\infty y \frac{2}{\theta} e^{-\frac{y}{\theta}} \left(1 - e^{-\frac{y}{\theta}}\right) dy \\
&= 2\int_0^\infty y \frac{1}{\theta} e^{-\frac{y}{\theta}}\, dy - \int_0^\infty y \frac{2}{\theta} e^{-\frac{2y}{\theta}}\, dy \\
&= 2E\left(Y_1\right) - E\left(Y_2\right) \\
&= 2\theta - \frac{\theta}{2} = \frac{3\theta}{2}.
\end{aligned}$$

where $Y_1 \sim \Gamma(1, \theta)$ and $Y_1 \sim \Gamma(1, \theta/2)$ and we also have,

$$\begin{aligned}
\mathrm{Var}(Y) &= \int_0^\infty y^2 \frac{2}{\theta} e^{-\frac{y}{\theta}} \left(1 - e^{-\frac{y}{\theta}}\right) dy - \left(\frac{3}{2}\theta\right)^2 \\
&= 2\int_0^\infty y^2 \frac{1}{\theta} e^{-\frac{y}{\theta}}\, dy - \int_0^\infty y^2 \frac{2}{\theta} e^{-\frac{2y}{\theta}}\, dy - \frac{9}{4}\theta^2 \\
&= 2\left[\mathrm{var}\left(Y_1\right) + \theta^2\right] - \left[\mathrm{var}\left(Y_2\right) + \frac{\theta^2}{4}\right] - \frac{9}{4}\theta^2 \\
&= 4\theta^2 - \frac{\theta^2}{2} - \frac{9}{4}\theta^2 = \frac{5\theta^2}{4}.
\end{aligned}$$

The marginal pdf of X is given by,

$$f_X(x) = \int_x^\infty f_{X,Y}(x, y) dy = \frac{2}{\theta} e^{-\frac{2}{\theta}x}, \quad 0 < x < \infty$$

Thus, $X \sim \Gamma(1, \theta/2)$ and the conditional pdf of Y given X is as follows,

$$f_{Y|X=x}(y,x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{1}{\theta} e^{\frac{x-y}{\theta}}, \quad 0 < x < y < \infty$$

Thus we have,

$$E(Y|X=x) = \int_x^\infty y \frac{1}{\theta} e^{\frac{x-y}{\theta}}\, dy = x + \theta.$$

Notice that

$$E(X+\theta) = E(E(Y|X)) = E(Y) = \frac{3\theta}{2},$$

and

$$\mathrm{Var}(X+\theta) = \mathrm{Var}(X) = \frac{\theta^2}{4}$$

which, in accordance with the theory, is less than $\mathrm{Var}(Y)$.

5. Let $\ell(\theta)$ be the likelihood function and

$$h(\theta) = \frac{\partial \log \ell(\theta)}{\partial \theta} = n\left(\frac{\overline{X}}{\theta} - 1 - \frac{1}{e^\theta - 1}\right).$$

Obviously, $\lim_{\theta\to\infty} h(\theta) = -n$ and since $\lim_{\theta\to 0} \theta/\left(e^\theta - 1\right) = 1$,

$$\lim_{\theta\to 0} h(\theta) = n \lim_{\theta\to 0} \frac{1}{\theta}\left(\overline{X} - \frac{\theta}{e^\theta - 1}\right) - n = \infty$$

when $\overline{X} > 1$. Note that h is continuous. Hence by intermediate value theorem, when $\bar{X} > 1$, $h(\theta) = 0$ has at least one solution. Note that

$$h'(\theta) = n\left[-\frac{\overline{X}}{\theta^2} + \frac{e^\theta}{\left(e^\theta - 1\right)^2}\right] = n\left[\frac{-\overline{X}\left(e^\theta - 1\right)^2/e^\theta + \theta^2}{\theta^2\left(e^\theta - 1\right)^2/e^\theta}\right].$$

The taylor expansion of e^θ is $\sum_{n\geq 0}\frac{\theta^n}{n!}$ hence,

$$e^\theta + e^{-\theta} = \sum_{n\geq 0}\frac{\theta^n}{n!} + \sum_{n\geq 0}(-1)^n\frac{\theta^n}{n!} = 2\sum_{n\geq 0}\frac{\theta^{2n}}{(2n)!} = 2 + \theta^2 + 2\sum_{n\geq 2}\frac{\theta^{2n}}{(2n)!}.$$

Hence $\left(e^\theta - 1\right)^2/e^\theta = \left(e^\theta - 1\right)\left(1 - e^{-\theta}\right) = e^\theta + e^{-\theta} - 2 > \theta^2$ and for $\overline{X} > 1$, $\overline{X}\left(e^\theta - 1\right)^2/e^\theta > \theta^2$. Therefore $h'(\theta) < 0$, so $h(\theta) = 0$ has a unique solution and $\log \ell(\theta)$ is convex. Therefore, the unique solution is the MLE of θ.