MATH 556 - ASSIGNMENT 4 – SOLUTIONS

1. *Consider the three-level hierarchical model*

LEVEL 3 :	$\theta = (\theta_1, \theta_2) \in \mathbb{R}^+ \times \mathbb{R}^+$	*Fixed*
LEVEL 2 :	$X \sim Gamma(\theta_1, \theta_2)$	
LEVEL 1 :	$Y_1, \ldots, Y_n \vert X = x \sim Poisson(x)$	$Y_1, \ldots, Y_n$ *independent given* X

(a) *Find the (marginal) joint pmf of* $Y_1, \ldots, Y_n$.

We have by direct calculation, for $(y_1, \ldots, y_n) \in \{\mathbb{Z}^+\}^n$,

$$
\begin{aligned}
f_{Y_1,\ldots,Y_n}(y_1, \ldots, y_n) &= \int_0^\infty \prod_{i=1}^n f_{Y|X}(y_i|x) f_X(x)\, dx \\
&= \int_0^\infty \prod_{i=1}^n \frac{e^{-x} x^{y_i}}{y_i!} \frac{\theta_2^{\theta_1}}{\Gamma(\theta_1)} x^{\theta_1 - 1} \exp\{-\theta_2 x\}\, dx \\
&= \frac{\theta_2^{\theta_1}}{\Gamma(\theta_1)} \frac{1}{\prod\limits_{i=1}^n y_i!} \int_0^\infty x^{s+\theta_1-1} \exp\{-(n+\theta_2)x\}\, dx \qquad s = \sum_{i=1}^n y_i \\
&= \frac{\theta_2^{\theta_1}}{\Gamma(\theta_1)} \frac{1}{\prod\limits_{i=1}^n y_i!} \frac{\Gamma(s+\theta_1)}{(n+\theta_2)^{s+\theta_1}}
\end{aligned}
$$

4 Marks

(b) *Find the marginal pmf of* Y_1.

Setting $n = 1$ in the above formula, we have that

$$
f_{Y_1}(y_1) = \frac{\theta_2^{\theta_1}}{\Gamma(\theta_1)} \frac{1}{y_1!} \frac{\Gamma(y_1+\theta_1)}{(1+\theta_2)^{y_1+\theta_1}} = \frac{\Gamma(y_1+\theta_1)}{\Gamma(\theta_1) y_1!} \left(\frac{1}{1+\theta_1}\right)^{y_1} \left(\frac{\theta_2}{1+\theta_1}\right)^{\theta_1} \qquad y_1 \in \mathbb{Z}^+
$$

and zero otherwise.

2 Marks

This is in fact a Negative Binomial distribution.

(c) *Find the correlation between* Y_1 *and* Y_2.

This is most easily computed using iterated expectation: we have

$$
\mathbb{E}_{Y_1}[Y_1] = \mathbb{E}_X\left[\mathbb{E}_{Y_1|X}[Y_1|X]\right] = \mathbb{E}_X[X] = \frac{\theta_1}{\theta_2}
$$

from the formula sheet. Clearly $\mathbb{E}_{Y_1}[Y_1] = \mathbb{E}_{Y_2}[Y_2]$. Also

$$
\mathbb{E}_{Y_1}[Y_1^2] = \mathbb{E}_X\left[\mathbb{E}_{Y_1|X}[Y_1^2|X]\right] = \mathbb{E}_X\left[\mathrm{Var}_{Y_1|X}[Y_1|X] + \{\mathbb{E}_{Y_1|X}[Y_1|X]\}^2\right] = \mathbb{E}_X[X + X^2]
$$

by properties of the Poisson distribution. Thus

$$
\mathrm{Var}_{Y_1}[Y_1] = \mathbb{E}_{Y_1}[Y_1^2] - \{\mathbb{E}_{Y_1}[Y_1]\}^2 = \mathbb{E}_X[X] + \mathbb{E}_X\left[X^2\right] - \{\mathbb{E}_X[X]\}^2 = \frac{\theta_1}{\theta_2} + \frac{\theta_1}{\theta_2^2} = \frac{\theta_1(1+\theta_2)}{\theta_2^2}
$$

Finally,

$$\mathbb{E}_{Y_1,Y_2}[Y_1Y_2] = \mathbb{E}_X\left[\mathbb{E}_{Y_1,Y_2|X}[Y_1Y_2|X]\right] = \mathbb{E}_X\left[\mathbb{E}_{Y_1|X}[Y_1|X]\mathbb{E}_{Y_2|X}[Y_2|X]\right]$$

by conditional independence. As before, $\mathbb{E}_{Y_1|X}[Y_1|X] = \mathbb{E}_{Y_2|X}[Y_2|X] = X$. Thus

$$\mathbb{E}_{Y_1,Y_2}[Y_1Y_2] = \mathbb{E}_X\left[X^2\right] = \text{Var}_X[X] + \{\mathbb{E}_X[X]\}^2 = \frac{\theta_1}{\theta_2^2} + \frac{\theta_1^2}{\theta_2^2} = \frac{\theta_1(1+\theta_1)}{\theta_2^2}$$

and hence

$$\text{Cov}_{Y_1,Y_2}[Y_1,Y_2] = \frac{\theta_1(1+\theta_1)}{\theta_2^2} - \frac{\theta_1^2}{\theta_2^2} = \frac{\theta_1}{\theta_2^2}$$

and

$$\text{Corr}_{Y_1,Y_2}[Y_1,Y_2] = \frac{\text{Cov}_{Y_1,Y_2}[Y_1,Y_2]}{\text{Var}_{Y_1}[Y_1]} = \frac{1}{1+\theta_2}.$$

4 Marks

2. *For $n \geq 1$ random variables $X_1, \ldots, X_n$, the order statistics, $Y_1, \ldots, Y_n$, are defined by*

$$Y_i = X_{(i)} - \text{"the } i\text{th smallest value in } X_1, \ldots, X_n\text{"}$$

for $i = 1, \ldots, n$. For example

$$Y_1 = X_{(1)} = \min\{X_1, \ldots, X_n\} \qquad Y_n = X_{(n)} = \max\{X_1, \ldots, X_n\}.$$

For $X_1, \ldots, X_n$ independently distributed from continuous distribution with pdf f_X, the joint pdf of order statistics $Y_1, \ldots, Y_n$ can be shown to be

$$f_{Y_1,\ldots,Y_n}(y_1, \ldots, y_n) = n! f_X(y_1) \ldots f_X(y_n) \qquad y_1 < \ldots < y_n$$

and zero otherwise.

(a) *Suppose X_1, X_2, X_3 are independent random variables having an $Exponential(1)$ distribution. Find the distribution of the second order statistic, Y_2, that is, the second smallest of X_1, X_2, X_3.*

From first principles, we have

$$f_{Y_1,Y_2,Y_3}(y_1,y_2,y_3) = 3!\exp\{-(y_1+y_2+y_3)\} \qquad 0 < y_1 < y_2 < y_3 < \infty$$

so, for $y_2 > 0$,

$$\begin{aligned} f_{Y_2}(y_2) &= 6\int_0^{y_2}\int_{y_2}^{\infty} \exp\{-(y_1+y_2+y_3)\}\, dy_3\, dy_1 \\ &= 6\exp\{-y_2\}\int_0^{y_2} \exp\{-y_1\}\exp\{-y_2\}\, dy_1 \\ &= 6\exp\{-2y_2\}\int_0^{y_2} \exp\{-y_1\}\, dy_1 \\ &= 6\exp\{-2y_2\}(1-\exp\{-y_2\}). \end{aligned}$$

Alternatively, using the general result from lectures, for $Y_j = X_{(j)}$, we have

$$\begin{aligned} f_{Y_j}(y_j) &= \frac{n!}{(j-1)!(n-j)!}\{F_X(y_j)\}^{j-1} f_X(y_j)\{1-F_X(y_j)\}^{n-j} \qquad y_j > 0 \\ &= \frac{3!}{(2-1)!(3-2)!}\{1-\exp\{-y_2\}\}^{2-1}\exp\{-y_2\}\{\exp\{-y_2\})\}^{3-2} \qquad j=2 \\ &= 6\{1-\exp\{-y_2\})\}\exp\{-2y_2\} \end{aligned}$$

as before.

5 Marks

(b) *Suppose $X_1, \ldots, X_n$ are independent continuous random variables with cdf F_X*

$$F_X(x) = 1 - x^{-1} \qquad x \geq 1$$

and zero otherwise.

Show that $Z_n = \min\{X_1, \ldots, X_n\}$ has a ***degenerate*** *distribution in the limit as $n \longrightarrow \infty$, that is, that*

$$\lim_{n \longrightarrow \infty} P_{Z_n}[Z_n = c] = 1$$

for some c to be identified, but that there exists a sequence of real values $\{\alpha_n\}$ such that $U_n = Z_n^{\alpha_n}$ has distribution F_X for each n.

Hint: for the first part, having identified c, show that

$$P_{Z_n}[Z_n < c] + P_{Z_n}[Z_n > c] \longrightarrow 0$$

as $n \longrightarrow \infty$.

For the first part, it is evident that we should inspect $c = 1$ as the degenerate limit. Then

$$P_{Z_n}[Z_n < 1] = 0$$

by definition of F_X, and

$$P_{Z_n}[Z_n > 1] = 1 - F_{Z_n}(1) = \{1 - F_X(1)\}^n = 0$$

by the result in lectures that for the minimum order statistic.

$$F_{Z_n}(z) = 1 - \{1 - F_X(z)\}^n = 1 - \frac{1}{z^n} \qquad z \geq 1.$$

Now for $u \geq 1$,

$$F_{U_n}(u) = P_{U_n}[U_n \leq u] = P_{Z_n}[Z_n^{\alpha_n} \leq u] = P_{Z_n}[Z_n \leq u^{1/\alpha_n}] = F_{Z_n}(u^{1/\alpha_n})$$

so

$$F_{U_n}(u) = 1 - \frac{1}{u^{n/\alpha_n}}.$$

Thus choosing $\alpha_n = n$ yields that

$$F_{U_n}(u) = 1 - \frac{1}{u} = F_X(u)$$

as required.

5 Marks