Queen's University
Department of Mathematics and Statistics

MTHE/STAT 353
Homework 3 Solutions, Winter 2022

- For each question, your solution should start on a fresh page. You can write your solution using one of the following three formats:
  (1) Use your own paper.
  (2) Use a tablet, such as an ipad.
  (3) Use document creation software, such as Word or LaTeX.
- Write your name and student number on the first page of each solution, and number your solution.
- For each question, photograph or scan each page of your solution (unless your solution has been typed up and is already in electronic format), and combine the separate pages into a single file. Then upload each file (one for each question), into the appropriate box in Crowdmark.

Instructions for submitting your solutions to Crowdmark are also here.

Total Marks : 30

Student Number Name

1. (6 marks) Let X_1, X_2, X_3 be independent, identically distributed continuous random variables. Find the probability that the second largest value (i.e., the median) is closer to the smallest value than to the largest value, when the common distribution of the X_i is

(a) (2 marks) the Uniform$(0,1)$ distribution (a symmetry argument should suffice here);

(b) (4 marks) the Exponential(λ) distribution.

Solution:

(a) For the Uniform$(0,1)$ distribution, by symmetry the median should be equally likely to be closer to the smallest value or to the largest value. So the probability of this is 1/2.

(b) For the Exponential(λ) distribution we integrate the joint density of $(X_{(1)}, X_{(2)}, X_{(3)})$ over the set of points (x_1, x_2, x_3) in 3-space satisfying $x_2 - x_1 < x_3 - x_2$ and $0 < x_1 < x_2 < x_3 < \infty$. The joint density of $(X_{(1)}, X_{(2)}, X_{(3)})$ is given by

$$f_{123}(x_1, x_2, x_3) = \begin{cases} 6\lambda^3 e^{-\lambda(x_1+x_2+x_3)} & \text{for } 0 < x_1 < x_2 < x_3 < \infty \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, the desired probability is

$$\begin{aligned} P(X_{(2)} - X_{(1)} < X_{(3)} - X_{(2)}) &= \int_0^\infty \int_{x_1}^\infty \int_{2x_2-x_1}^\infty 6\lambda^3 e^{-\lambda(x_1+x_2+x_3)} dx_3 dx_2 dx_1 \\ &= \int_0^\infty \int_{x_1}^\infty 6\lambda^2 e^{-\lambda(x_1+x_2)} \left[-e^{-\lambda x_3} \right]_{2x_2-x_1}^\infty dx_2 dx_1 \\ &= \int_0^\infty \int_{x_1}^\infty 6\lambda^2 e^{-\lambda(x_1+x_2)} e^{-\lambda(2x_2-x_1)} dx_2 dx_1 \\ &= \int_0^\infty 2\lambda \int_{x_1}^\infty 3\lambda e^{-3\lambda x_2} dx_2 dx_1 \\ &= \int_0^\infty 2\lambda \left[-e^{-3\lambda x_2} \right]_{x_1}^\infty dx_1 \\ &= \int_0^\infty 2\lambda e^{-3\lambda x_1} dx_1 \\ &= \frac{2}{3} \int_0^\infty 3\lambda e^{-3\lambda x_1} dx_1 \\ &= \frac{2}{3}. \end{aligned}$$

Student Number　　　　　　　　　　　　Name

2. (6 marks) Let $X_1, \ldots, X_n$ be a sequence of independent Uniform$(0,1)$ random variables, with $X_{(1)}, \ldots, X_{(n)}$ denoting their order statistics. Let A_n denote the expected area of the triangle formed by the vertices $(X_{(n-2)}, 0)$, $(X_{(n-1)}, X_{(n-1)})$, and $(X_{(n)}, 0)$. Find A_n (in terms of n) and show that $nA_n \to 1$ as $n \to \infty$.

Solution: In terms of $X_{(n-2)}$, $X_{(n-1)}$, and $X_{(n)}$ the expected area of the triangle is given by

$$A_n = E[(X_{(n)} - X_{(n-2)})X_{(n-1)}/2].$$

The joint density of $(X_{(n-2)}, X_{(n-1)}, X_{(n)})$ is

$$f_{n-2,n-1,n}(x,y,z) = n!\frac{F(x)^{n-3}}{(n-3)!}f(x)f(y)f(z),$$

where F and f are, respectively, the distribution function and the probability density function of the Uniform$(0,1)$ distribution. Therefore, we have

$$f_{n-2,n-1,n}(x,y,z) = \begin{cases} n(n-1)(n-2)x^{n-3} & \text{for } 0 < x < y < z < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Therefore,

$$\begin{aligned}
A_n &= \frac{n(n-1)(n-2)}{2}\int_0^1\int_0^z\int_0^y (z-x)yx^{n-3}dxdydz \\
&= \frac{n(n-1)(n-2)}{2}\int_0^1\int_0^z\left[\frac{zyx^{n-2}}{n-2} - \frac{yx^{n-1}}{n-1}\right]_0^y dydz \\
&= \frac{n(n-1)(n-2)}{2}\int_0^1\int_0^z\left(\frac{zy^{n-1}}{n-2} - \frac{y^n}{n-1}\right)dydz \\
&= \frac{n(n-1)(n-2)}{2}\int_0^1\left[\frac{zy^n}{n(n-2)} - \frac{y^{n+1}}{(n+1)(n-1)}\right]_0^z dz \\
&= \frac{n(n-1)(n-2)}{2}\int_0^1\left(\frac{z^{n+1}}{n(n-2)} - \frac{z^{n+1}}{(n+1)(n-1)}\right)dz \\
&= \frac{n-1}{2}\int_0^1 z^{n+1}dz - \frac{n(n-2)}{2(n+1)}\int_0^1 z^{n+1}dz \\
&= \frac{n-1}{2(n+2)} - \frac{n(n-2)}{2(n+1)(n+2)} \\
&= \frac{(n-1)(n+1) - n(n-2)}{2(n+1)(n+2)} = \frac{2n-1}{2(n+1)(n+2)}.
\end{aligned}$$

Then

$$nA_n = \frac{2n^2 - n}{2n^2 + 6n + 4} = \frac{2 - 1/n}{2 + 6/n + 4/n^2} \to \frac{2}{2} = 1 \quad \text{as } n \to \infty.$$

Student Number Name

3. (6 marks) Let $X_1, \ldots, X_n$ be mutually independent Uniform(0,1) random variables. Find the probability that the interval $(\min(X_1, \ldots, X_n), \max(X_1, \ldots, X_n))$ contains the value 1/2 and find the smallest n such that this probability is at least 0.95.

Solution: For a given n, the probability we want to compute is $P(X_{(1)} < 1/2, X_{(n)} > 1/2)$. The joint pdf of $(X_{(1)}, X_{(n)})$ is given by

$$f_{1n}(x_1, x_n) = \begin{cases} n(n-1)(x_n - x_1)^{n-2} & \text{for } 0 < x_1 < x_n < 1 \\ 0 & \text{otherwise} \end{cases}$$

Integrating this over the region of interest gives

$$\begin{aligned} P(X_{(1)} \leq 1/2, X_{(n)} \geq 1/2) &= \int_0^{1/2} \int_{1/2}^1 n(n-1)(x_n - x_1)^{n-2} dx_n dx_1 \\ &= n \int_0^{1/2} \left[(x_n - x_1)^{n-1} \Big|_{1/2}^{1} \right] dx_1 \\ &= n \int_0^{1/2} \left[(1 - x_1)^{n-1} - \left(\frac{1}{2} - x_1\right)^{n-1} \right] dx_1 \\ &= -(1 - x_1)^n \Big|_0^{1/2} + \left(\frac{1}{2} - x_1\right)^n \Big|_0^{1/2} \\ &= 1 - \left(\frac{1}{2}\right)^n - \left(\frac{1}{2}\right)^n \\ &= 1 - \left(\frac{1}{2}\right)^{n-1}. \end{aligned}$$

Setting this equal to 0.95 gives

$$0.95 = 1 - \left(\frac{1}{2}\right)^{n-1} \Leftrightarrow 0.05 = \left(\frac{1}{2}\right)^{n-1} \Leftrightarrow \ln(.05) = (n-1)\ln(1/2) \Leftrightarrow n = \frac{\ln(.05)}{\ln(.5)} + 1.$$

This gives $n = \frac{\ln 20}{\ln 2} + 1 = 4.322 + 1 = 5.322$. Thus, the smallest n for which the probability is at least 0.95 is $n = 6$.

Student Number Name

4. (6 marks) Let $X_1, \ldots, X_n$ be a sequence of independent Uniform$(0,1)$ random variables, with $X_{(1)}, \ldots, X_{(n)}$ denoting their order statistics. For fixed k let $g_n(x)$ denote the probability density function of $nX_{(k)}$. Find $g_n(x)$ and show that

$$\lim_{n\to\infty} g_n(x) = \begin{cases} \frac{x^{k-1}}{(k-1)!}e^{-x} & \text{for } x \geq 0 \\ 0 & \text{for } x < 0. \end{cases}$$

(note that as n increases the sample size also increases).

Solution: The probability density function of $X_{(k)}$ is

$$\begin{aligned} f_k(x_k) &= \frac{n!}{(k-1)!(n-k)!} f(x_k)F(x_k)^{k-1}(1-F(x_k))^{n-k} \\ &= \begin{cases} \frac{n!}{(k-1)!(n-k)!} x_k^{k-1}(1-x_k)^{n-k} & \text{for } 0 < x_k < 1 \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

Let $X = nX_{(k)}$. By the change of variable formula, the probability density function of X is

$$\begin{aligned} g_n(x) &= \frac{1}{n} f_k(x/n) \\ &= \begin{cases} \frac{(n-1)!}{(k-1)!(n-k)!}(x/n)^{k-1}(1-x/n)^{n-k} & \text{for } 0 < x < n \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

To more easily see the behaviour of $g_n(x)$ for large n we write it as

$$g_n(x) = \left[\left(\frac{n-1}{n}\right) \times \ldots \left(\frac{n-(k-1)}{n}\right)(1-x/n)^{-k}\right] \frac{x^{k-1}}{(k-1)!}\left(1-\frac{x}{n}\right)^n.$$

The term in square brackets above has k factors, each of which goes to 1 as $n \to \infty$. Also,

$$\left(1-\frac{x}{n}\right)^n \to e^{-x} \quad \text{as } n \to \infty.$$

Finally, the sample space of $nX_{(k)}$, which is $(0,n)$ converges to $(0,\infty)$ as $n \to \infty$. Therefore,

$$\lim_{n\to\infty} g_n(x) = \begin{cases} \frac{x^{k-1}}{(k-1)!}e^{-x} & \text{for } x > 0 \\ 0 & \text{for } x \leq 0, \end{cases}$$

Student Number Name

5. (6 marks) Let $X_1, X_2, \ldots$ be a sequence of independent random variables with the exponential distribution with mean 1, and let $X_{(n)} = \max(X_1, \ldots, X_n)$. For $x > 0$, show that

$$\lim_{n\to\infty} P(X_{(n)} - \ln n \leq x) = \exp(-e^{-x}).$$

Solution: We have

$$\begin{aligned}
P(X_{(n)} - \ln n \leq x) &= P(X_{(n)} \leq x + \ln n) \\
&= P(\max(X_1, \ldots, X_n) \leq x + \ln n) \\
&= P(X_1 \leq x + \ln n, \ldots, X_n \leq x + \ln n) \\
&= P(X_1 \leq x + \ln n) \times \ldots \times P(X_n \leq x + \ln n) \\
&= (1 - e^{-(x+\ln n)})^n \\
&= \left(1 - \frac{e^{-x}}{n}\right)^n.
\end{aligned}$$

Therefore,

$$\lim_{n\to\infty} P(X_{(n)} - \ln n \leq x) = \lim_{n\to\infty} \left(1 - \frac{e^{-x}}{n}\right)^n = \exp\{-e^{-x}\},$$

using the well known fact from calculus that $(1 + x/n)^n \to e^x$ as $n \to \infty$.