Resit Probability Theory (WBMA046-05)

7 July 2023, 11:45-13:45

READ THE INSTRUCTIONS BELOW CAREFULLY BEFORE STARTING THE EXAM.

- This exam contains 7 pages (including this cover page) and **5** exercises.
- Write your name and student number at the top of EACH page (including this cover page).
- Your answers should be written in this booklet. Preferably, avoid handing in extra paper.
  If you do not have enough space to write your solution below the exercise, you can use the back of the sheet. If this is still not enough space, you can add one (or more) extra sheet(s) to the exam. In this case, indicate it clearly in the booklet, and write your name and student number on the top of that (these) sheet(s).
- Do not write on the table below.
- Sticking to the rules above is worth 10 points.
- It is absolutely NOT allowed to use calculator, phone, smartwatch, books, lecture notes or any other aids.
- Always give a short proof of your answer or a calculation to justify it, or clearly state the facts from the lecture notes you are using (unless it is stated explicitly in the question this is not needed.).

Exercise	Points	Score
1	15	
2	15	
3	20	
4	10	
5	30	
Follow the rules (name on each page...)	10	
Total	100	

Exercise 1 (a:4, b:5, c:6 pts)
You have a beautiful flower collection consisting of 12 tulips (T), 2 roses (R) and 1 daisy (D). You can distinguish between the different types of flowers but not between flowers of the same type. That means, for instance, all tulips look the same to you. You would like to display your flowers side by side on 3 shelves. Each shelf has space for 5 flowers. The shelves themselves are distinguishable and the positions from 1 to 5 on each shelf matter.

	shelf 1:	D	R	R	T	T
Example:	shelf 2:	T	T	T	T	T
	shelf 3:	T	T	T	T	T

a. In how many ways can you display your flowers?

The number of possibilities is:

b. You do not want any shelf to only contain tulips. In how many ways can you display your flowers now?

The number of possibilities is:

c. You want each shelf to be symmetric. That means if for a shelf $i \in \{1,2,3\}$ at position $j \in \{1,\dots,5\}$ there is a flower of type $X \in \{D,R,T\}$, there has to be a flower of the same type at position $6-j$ as well. How many ways are there now?

The number of possibilities is:

Remark: Give your answer (a number or a product of numbers and binomial coefficients) directly in the box above. You are not required to provide any explanation in this exercise. For each question you get either zero or the full points.

Solution:

a) $15 \times \binom{14}{2}$

b) $3 \times 5 \times 5^2$

c) Case: roses are on the same shelf: $3 \times 6 = 18$
Case: roses are on different shelves: 3
Together: 21

Exercise 2 (15 pts)
Let $a > 0$ and $\lambda > 0$. A continuous random variable X has *Pareto distribution* with positive parameters a and λ if its pdf is

$$f(x) = \mathbf{1}(x > a)\lambda a^{\lambda} x^{-\lambda-1}.$$

Let $x_0 > a$. Let Y be a random variable with distribution given by the conditional distribution of X under the condition $X > x_0$, i.e. $F_Y(y) = \mathbb{P}(X \leq y \mid X > x_0)$ for any $y \in \mathbb{R}$.

Find the pdf of Y.
Hint: You might want to compute $1 - F_X(x)$ and $1 - F_Y(y)$ first.

Solution: For $y \leq x_0$, $F_Y(y) = 0$. Thus $\boxed{f_Y(y) = 0 \text{ for all } y \leq x_0}$.
Let $y > x_0$. Following the hint we compute

$$1 - F_X(y) = \int_y^{\infty} f_X(t)\,\mathrm{d}t = \int_y^{\infty} \lambda a^{\lambda} t^{-\lambda-1}\,\mathrm{d}t = \left[-a^{\lambda} t^{-\lambda}\right]_{t=y}^{\infty} = a^{\lambda} y^{-\lambda}$$

and

$$1 - F_Y(y) = \mathbb{P}(X \geq y | X > x_0) = \frac{\mathbb{P}(X \geq y \,\&\, X > x_0)}{\mathbb{P}(X > x_0)} = \frac{\mathbb{P}(X \geq y)}{\mathbb{P}(X > x_0)} = \frac{1 - F_X(y)}{1 - F_X(x_0)} = \frac{a^{\lambda} y^{-\lambda}}{a^{\lambda} x_0^{-\lambda}} = \left(\frac{y}{x_0}\right)^{-\lambda}.$$

Thus

$$f_Y(y) = \frac{\partial F_Y(y)}{\partial y} = \frac{\partial}{\partial y}\left[1 - \left(\frac{y}{x_0}\right)^{-\lambda}\right] = \boxed{\lambda x_0^{\lambda} y^{-\lambda-1} \text{ for all } y > x_0}.$$

Remark: We observe that Y has Pareto distribution with parameters x_0 and λ.

Exercise 3 (a:5, b:5, c:5, d:5 pts)

a) Let X be a random variable with expectation $\mathbb{E}[X] = \mu \in \mathbb{R}$ and variance $\mathbb{V}\mathrm{ar}(X) < \infty$. State Chebyshev's inequality for X.

Let $(X_i)_{i\in\mathbb{N}}$ be a sequence of i.i.d. random variables with expectation $\mathbb{E}[X_1] = \mu$ and $\mathbb{E}[X_1^2] < \infty$.
Let $S_n := \binom{n}{2}^{-1} \sum_{1\leq i<j\leq n} X_i X_j$.

b) Compute $\mathbb{E}[S_n]$.

c) Show that $\mathbb{E}[S_n^2] = \binom{n}{2}^{-2}\left(\binom{n}{2}\mathbb{E}[X_1^2]^2 + \binom{n}{2}2(n-2)\mu^2\mathbb{E}[X_1^2] + \binom{n}{2}\binom{n-2}{2}\mu^4\right)$.

d) Show that $S_n \underset{\mathbb{P}}{\overset{n\to\infty}{\longrightarrow}} \mu^2$.

Solution:

a) For any $a > 0$,

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq a) \leq \frac{\mathbb{V}\mathrm{ar}(X)}{a^2}.$$

b) By linearity of the expectation and the i.i.d. property of the random variables we get $\mathbb{E}[S_n] = \mathbb{E}[X_1X_2] = \mathbb{E}[X_1]\mathbb{E}[X_2] = \boxed{\mu^2}$.

c)

$$\begin{aligned}\mathbb{E}[S_n^2] &= \binom{n}{2}^{-2}\mathbb{E}\Bigg[\sum_{i<j} X_i^2X_j^2 + \sum_{\substack{i<j,k<l\\ |\{i,j,k,l\}|=3}} X_iX_jX_kX_l + \sum_{\substack{i<j,k<l\\ |\{i,j,k,l\}|=4}} X_iX_jX_kX_l\Bigg]\\ &= \binom{n}{2}^{-2}\left(\binom{n}{2}\mathbb{E}[X_1^2]^2 + \binom{n}{2}2(n-2)\mu^2\mathbb{E}[X_1^2] + \binom{n}{2}\binom{n-2}{2}\mu^4\right)\end{aligned}$$

d)

$$\mathbb{E}[S_n^2] = \underbrace{\binom{n}{2}^{-1}\mathbb{E}[X_1^2]^2}_{\overset{n\to\infty}{\longrightarrow}0} + \underbrace{2\binom{n}{2}^{-1}(n-2)\mu^2\mathbb{E}[X_1^2]}_{\overset{n\to\infty}{\longrightarrow}0} + \underbrace{\binom{n}{2}^{-1}\binom{n-2}{2}\mu^4}_{\overset{n\to\infty}{\longrightarrow}\mu^4} \overset{n\to\infty}{\longrightarrow} \mu^4$$

From this, Chebyshev's inequality yields for any $\varepsilon > 0$

$$\mathbb{P}(|S_n - \mu^2| \geq \varepsilon) \leq \frac{\mathbb{V}\mathrm{ar}(S_n)}{\varepsilon^2} = \varepsilon^{-2}(\mathbb{E}[S_n^2] - \mu^4) \overset{n\to\infty}{\longrightarrow} 0.$$

Exercise 4 (10 pts) Let X, Y, Z be independent and uniformly distributed over $(-1, 1)$. Compute $\mathbb{P}(X \geq YZ)$
Solution: Since

$$f_{X,Y,Z}(x,y,z) = f_X(x)f_Y(y)f_Z(z) = \frac{1}{8}, \qquad -1 \leq x \leq 1, \quad -1 \leq y \leq 1, \quad -1 \leq z \leq 1,$$

we have

$$\begin{aligned}
\mathbb{P}(X \geq YZ) &= \iiint_{x \geq yz} f_{X,Y,Z}(x,y,z)dx\,dy\,dz \\
&= \int_{-1}^{1}\int_{-1}^{1}\int_{yz}^{1} \frac{1}{8}dx\,dy\,dz \\
&= \int_{-1}^{1}\int_{-1}^{1} \frac{1}{8} - \frac{yz}{8}dy\,dz \\
&= \int_{-1}^{1} \frac{1}{4}dz \\
&= \frac{1}{2}.
\end{aligned}$$

ALTERNATIVE SOLUTION: Note that X and $-X$ have the same distribution. Thus $\mathbb{P}(X \geq YZ) = \mathbb{P}(-X \geq YZ)$ and

$$\mathbb{P}(X \geq YZ) = \frac{\mathbb{P}(X \geq YZ) + \mathbb{P}(-X \geq YZ)}{2} = \frac{1 - \mathbb{P}(X = YZ)}{2} = \frac{1}{2}.$$

Exercise 5 (a:5, b:5, c:5, d:5, e:5, f:5 pts)
The joint density of X and Y is given by

$$f(x,y) = C(1+|x|)^{-4}(1+|y|)\mathbf{1}(-x > y > x).$$

a) Find the value of the constant C.

b) Find the density function f_X of X.
(*Hint: When $f_X(x) \neq 0$, it is of the form $\sum_{k\geq 2} a_k(1+|x|)^{-k}$ with $a_k \in \mathbb{R}$ and $a_2 > 0$.*)

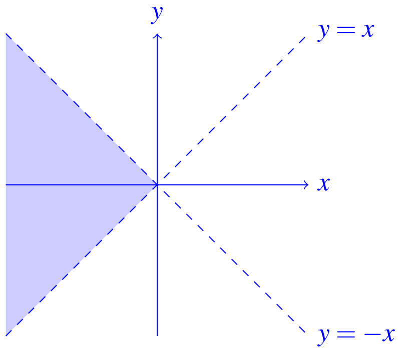
c) Find the density function f_Y of Y.
(*Hint: When $f_Y(y) \neq 0$, it is of the form $\sum_{k\geq 2} b_k(1+|y|)^{-k}$ with $b_k \in \mathbb{R}$ and $b_2 > 0$.*)

d) Is the expectation of X well defined? If yes find it.
(*Hint: You might want to compute $\mathbb{E}[1-X]$ first.*)

e) Is the expectation of Y well defined? If yes find it.
(*Hint: You might want to compute $\mathbb{E}[(1+Y)\mathbf{1}(y \geq 0)]$ and $\mathbb{E}[(1-Y)\mathbf{1}(y \leq 0)]$ first.*)

f) Determine if X and Y are independent.

Remark 1: For the questions b, c, d and e, you are allowed to write your answered with the constant C non-evaluated.
Remark 2: You are allowed to answer the questions in a different order, for example answering b before a.
Remark 3: We recall that expectations are allowed to be infinite.
Solution: *Remark*: Drawing the domain $\{(x,y) : -x > y > x\}$ can be helpful.



We will first compute the marginals f_X and f_Y because integrating any of the two will lead us to finding C.

b) For $x > 0$, one has $\mathbf{1}(-x > y > x)$ for any $y \in \mathbb{R}$, thus $\boxed{f_X(x) = 0 \text{ for } x > 0}$. Now, let $x \leq 0$.

$$\begin{aligned} f_X(x) &= C\int_x^{|x|}(1+|x|)^{-4}(1+|y|)\,dy \\ &= 2C(1+|x|)^{-4}\int_0^{|x|}(1+y)\,dy \\ &= 2C(1+|x|)^{-4}\frac{(1+|x|)^2-1}{2} \\ &= \boxed{C(1+|x|)^{-2} - C(1+|x|)^{-4} \text{ for } x \leq 0}. \end{aligned}$$

c) First we consider $y \geq 0$.

$$\begin{aligned} f_Y(y) &= C\int_{-\infty}^{-y}(1+|x|)^{-4}(1+|y|)\,dx \\ &= C(1+|y|)\int_{y}^{\infty}(1+x)^{-4}\,dx \\ &= C(1+|y|)\frac{(1+y)^{-3}}{3} \\ &= \frac{C}{3}(1+|y|)^{-2}. \end{aligned}$$

Moreover, since $f(x,-y) = f(x,y)$ for all $x,y \in \mathbb{R}$, we have that $f_Y(-y) = f_Y(y)$ for all $y \in \mathbb{R}$. Thus, we can conclude that $\boxed{f_Y(y) = \frac{C}{3}(1+|y|)^{-2} \text{ for all } y \in \mathbb{R}}$.

a) To find C, one can use any of the following equalities: $1 = \iint f$, $1 = \int f_X$ or $1 = \int f_Y$. In this solutions we will use the last one.

$$1 = \int_{-\infty}^{\infty} f_Y(y)\,dy = \frac{C}{3}\int_{-\infty}^{\infty}(1+|y|)^{-2}\,dy = \frac{2C}{3}\int_{0}^{\infty}(1+y)^{-2}\,dy = \frac{2C}{3}\int_{1}^{\infty}y^{-2}\,dy = \frac{2C}{3}.$$

Thus $\boxed{C = \frac{3}{2}}$.

d)

$$\begin{aligned} \mathbb{E}[1-X] &= \int(1-x)f_X(x)\,dx \\ &= \int_{-\infty}^{0}(1-x)[C(1+|x|)^{-2} - C(1+|x|)^{-4}]\,dx \\ &= \int_{0}^{\infty}(1+x)[C(1+x)^{-2} - C(1+x)^{-4}]\,dx \\ &= \int_{1}^{\infty}x[Cx^{-2} - Cx^{-4}]\,dx \\ &= \int_{1}^{\infty}[Cx^{-1} - Cx^{-3}]\,dx = \infty. \end{aligned}$$

Hence, because of the linearity of the expectation $\mathbb{E}[X] = 1 - \mathbb{E}[1-X] = 1 - \infty = \boxed{-\infty}$.

e)

$$\mathbb{E}[(1+Y)\mathbf{1}(y \geq 0)] = \int_{0}^{\infty}(1+y)f_Y(y)\,dy = \frac{C}{3}\int_{0}^{\infty}(1+y)\times(1+|y|)^{-2}\,dy = \frac{C}{3}\int_{1}^{\infty}y^{-1}\,dy = \infty.$$

From this, we deduce that

$$\mathbb{E}[Y\mathbf{1}(y \geq 0)] = \mathbb{E}[(1+Y)\mathbf{1}(y \geq 0)] - \mathbb{E}[\mathbf{1}(y \geq 0)] = \mathbb{E}[(1+Y)\mathbf{1}(y \geq 0)] - \mathbb{P}[y \geq 0] = \infty - \frac{1}{2} = \infty.$$

$$\mathbb{E}[(1-Y)\mathbf{1}(y \leq 0)] = \int_{-\infty}^{0}(1-y)f_Y(y)\,dy = \frac{C}{3}\int_{-\infty}^{0}(1+|y|)\times(1+|y|)^{-2}\,dy = \frac{C}{3}\int_{1}^{\infty}y^{-1}\,dy = \infty.$$

From this, we deduce that

$$\mathbb{E}[Y\mathbf{1}(y \leq 0)] = -\mathbb{E}[(1-Y)\mathbf{1}(y \leq 0)] + \mathbb{E}[\mathbf{1}(y \leq 0)] = -\mathbb{E}[(1-Y)\mathbf{1}(y \leq 0)] + \mathbb{P}[y \leq 0] = -\infty + \frac{1}{2} = -\infty.$$

If $\mathbb{E}[Y]$ would be well defined, we would have $\mathbb{E}[X] = \mathbb{E}[Y\mathbf{1}(y \leq 0)] + \mathbb{E}[Y\mathbf{1}(y \geq 0)] = -\infty + \infty$. But the last sum is undefined. Hence $\boxed{\mathbb{E}[Y] \text{ is undefined}}$.

f) They are $\boxed{\text{not independent}}$ because $f \neq f_X f_Y$.