Final Exam (2h)

The exam is open book.

Exercise 1 [Random vectors]
Let (X, Y) have joint density

$$f(x,y) = \begin{cases} \frac{2x+4y}{3} & \text{if } (x,y) \in [0,1]^2 \\ 0 & \text{otherwise.} \end{cases}$$

1. Find the marginal densities f_X and f_Y of X and Y.
2. Are X and Y independent ? Justify briefly.
3. Prove that
$$\mathbb{E}[X|Y] = \frac{2+6Y}{3+12Y}.$$

Answers :

1. Let us compute, for a fixed $x \in [0,1]$,
$$f_X(x) = \int_{y=0}^{1} \frac{2x+4y}{3} dy = \frac{2x}{3} + \int_{y=0}^{1} \frac{4}{3} y dy = \frac{2x}{3} + \frac{2}{3}.$$
Similarly, for a fixed $y \in [0,1]$
$$f_Y(y) = \int_{x=0}^{1} \frac{2x+4y}{3} dx = \frac{4y}{3} + \int_{x=0}^{1} \frac{2}{3} x dx = \frac{4y}{3} + \frac{1}{3}.$$
2. If X, Y were independent, then one would have $f(x,y) = f_X(x) f_Y(y)$, this is not the case.
3. According to the course formula (p.37),
$$\begin{aligned} \mathbb{E}[X|Y] &= \frac{\int_x x f(x,Y) dx}{f_Y(Y)} \\ &= \frac{\int_x x(2x+4Y)/3 dx}{4Y/3+1/3} \\ &= \frac{\int_x x(2x+4Y) dx}{4Y+1} \\ &= \frac{2\int_{x=0}^{1} x^2 dx + 4Y \int_{x=0}^{1} x dx}{4Y+1} \\ &= \frac{2/3+4Y/2}{4Y+1} = \frac{2+6Y}{3+12Y}. \end{aligned}$$

Exercise 2 [Random variables]
Let X have density $f_X(x) = \begin{cases} \frac{3}{x^4} & \text{if } x \geq 1, \\ 0 & \text{otherwise.} \end{cases}$.

1. Compute $\mathbb{E}[X^2]$. Does X belong to L^2 ?
2. Compute the cumulative distribution function F_X of X.
3. Prove that $1 - \frac{1}{X^3}$ have the uniform distribution over $[0,1]$.
(Hint : you can use the CDF, or the "change of variable" method.)

Answers :

1.
$$\mathbb{E}[X^2] = \int_1^{+\infty} x^2 \frac{3}{x^4} dx = -3/x\big|_1^\infty = 3 < +\infty,$$

therefore $X \in L^2$.

2. Let $t \geq 1$,
$$F_X(t) = \mathbb{P}(X \leq t) = \int_1^t \frac{3}{x^4} dx = -1/x^3\big|_1^t = 1 - \frac{1}{t^3}.$$

3. We first observe that $X \in [1, +\infty)$, so $1/X^3 \in [0,1]$, and $1 - 1/X^3 \in [0,1]$ also.
1st method : CDF. Let us compute its CDF : for any $0 \leq t \leq 1$
$$\begin{aligned}\mathbb{P}(1 - 1/X^3 \leq t) &= \mathbb{P}(-1/X^3 \leq t - 1)\\ &= \mathbb{P}(X^3 \leq 1/(1-t))\\ &= \mathbb{P}\left(X \leq (1/(1-t))^{1/3}\right)\\ &= F_X\left((1/(1-t))^{1/3}\right)\\ &= 1 - \frac{1}{((1/(1-t))^{1/3})^3} = 1 - \frac{1}{1/(1-t)} = 1 - (1-t) = t.\end{aligned}$$

Now, $F(t) = t$ is the CDF of the uniform distribution in $[0,1]$.
2d method : "change of variable". Let ϕ be any continuous and bounded function, and let us compute
$$\mathbb{E}[\phi(1 - 1/X^3)] = \int_1^{+\infty} \phi(1 - 1/x^3) 3/x^4 dx.$$

We put $u = 1 - 1/x^3$, then $x = \frac{1}{(1-u)^{1/3}}$, and $\frac{du}{dx} = \frac{3}{x^4}$. The boundaries become : $x = 1 \Leftrightarrow u = 0$ and $x = +\infty \Leftrightarrow u = 1$. Thus,
$$\mathbb{E}[\phi(1 - 1/X^3)] = \int_{u=0}^1 \phi(u) du = \int_u \phi(u) \mathbf{1}_{[0,1]}(u) du,$$

which proves that $1 - 1/X^3$ is uniform (Theorem 2.7).

Exercise 3 [Gaussian Vectors]
Let $\binom{X}{Y}$ be a gaussian vector with mean $\mu = \binom{0}{0}$ and with covariance matrix
$$C = \begin{pmatrix} 1 & -1 \\ -1 & 3 \end{pmatrix}.$$

1. Briefly explain why $\binom{aY}{X-aY}$ is a gaussian vector.
2. Let a be a fixed constant. Compute $\mathrm{Cov}(Y, X - aY)$, and deduce the value of a such that Y and $X - aY$ are independent.
3. Compute $\mathbb{E}[X|Y]$.
(Hint : Write $\mathbb{E}[X|Y] = \mathbb{E}[aY + (X - aY)|Y]$, where a is the answer to the previous question.)

Correction :

1. The vector $\binom{aY}{X-aY}$ is the image of $\binom{X}{Y}$ by a linear transformation :
$$\begin{pmatrix} aY \\ X - aY \end{pmatrix} = \begin{pmatrix} 0 & a \\ 1 & -a \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}.$$

Thus, it is a gaussian vector (see Proposition 4.5).

2.
$$\mathrm{Cov}(Y, X - aY) = \mathrm{Cov}(Y, X) - a\mathrm{Cov}(Y, Y) = -1 - 3a.$$

Therefore we have $\mathrm{Cov}(Y, X - aY) = 0$ for $a = -1/3$. Now, Y and $X - aY = X + Y/3$ are the two components of a gaussian vector with a null covariance. Therefore they are independent (Proposition 4.4).

3. Since $X = -Y/3 + (X + Y/3)$, we have by linearity of the conditional expectation

$$\begin{array}{cccc} \mathbb{E}[X|Y] & = & \mathbb{E}[-Y/3|Y] & + \quad \mathbb{E}[X+Y/3|Y] \\ & = & -Y/3 & + \quad \mathbb{E}[X+Y/3] \\ & & \text{('taking out what is known')} & \text{(independence)} \end{array}$$

Finally, $\mathbb{E}[X+Y/3] = \mathbb{E}[X] + \mathbb{E}[Y/3] = 0 + 0$ since $\binom{aY}{X-aY}$ has mean $\binom{0}{0}$. We get

$$\mathbb{E}[X|Y] = -Y/3.$$

Exercise 4 [A kind of Law of Large Numbers]
Let $(X_k)_{k\geq 1}$ be a sequence of independent random variables such that for all $k \geq 1$,

$$X_k = \begin{cases} \sqrt{k} & \text{with probability } \frac{1}{\sqrt{k}} \\ 0 & \text{with probability } 1 - \frac{1}{\sqrt{k}}. \end{cases}$$

We set $S_n = X_1 + \cdots + X_n$.

1. Compute, for every k, $\mathbb{E}[X_k], \mathbb{E}[S_n], \mathrm{Var}[X_k], \mathrm{Var}[S_n]$.

The goal of the remainder of the exercise is to prove that

$$\frac{S_n}{n} \overset{\text{(prob.)}}{\to} 1.$$

Yet, the law of large numbers cannot be applied here, since X_k's have not the same distribution.

2. Check that for every n we have $\mathrm{Var}[S_n] \leq n\sqrt{n}$.
3. Using the Chebychev's inequality (page 15), prove that $\frac{S_n}{n}$ converges to 1 in probability.

Answers :

1. We have

$$\mathbb{E}[X_k] = \sqrt{k} \times \frac{1}{\sqrt{k}} + 0 \times \left(1 - \frac{1}{\sqrt{k}}\right) = 1,$$
$$\mathbb{E}[S_n] = \mathbb{E}[X_1] + \ldots \mathbb{E}[X_n] = n \times 1 = n.$$

Besides,

$$\mathbb{E}[X_k^2] = (\sqrt{k})^2 \times \frac{1}{\sqrt{k}} = \sqrt{k}.$$

We get $\mathrm{Var}[X_k] = \mathbb{E}[X_k^2] - \mathbb{E}[X_k]^2 = \sqrt{k} - 1$.
Since X_k's are independent,

$$\begin{aligned} \mathrm{Var}[S_n] &= \mathrm{Var}[X_1] + \cdots + \mathrm{Var}[X_n] \\ &= \sqrt{1} - 1 + \sqrt{2} - 1 + \cdots + \sqrt{n} - 1 = \sqrt{1} + \sqrt{2} + \cdots + \sqrt{n} - n. \end{aligned}$$

2. We have

$$\begin{aligned} \mathrm{Var}[S_n] &= \sqrt{1} + \sqrt{2} + \cdots + \sqrt{n} - n \\ &\leq \sqrt{1} + \sqrt{2} + \cdots + \sqrt{n} \\ &\leq \sqrt{n} + \sqrt{n} + \cdots + \sqrt{n} \\ &= n\sqrt{n}. \end{aligned}$$

3. We need to prove that for every $\varepsilon > 0$,

$$\mathbb{P}\left(\left|\frac{S_n}{n} - 1\right| \geq \varepsilon\right) \overset{n\to+\infty}{\to} 0.$$

Using the Chebychev's inequality

$$\begin{aligned}
\mathbb{P}\left(\left|\frac{S_n}{n}-1\right|\geq\varepsilon\right) &= \mathbb{P}\left(\left|\frac{S_n}{n}-\mathbb{E}\left[\frac{S_n}{n}\right]\right|\geq\varepsilon\right) \\
&\leq \frac{\operatorname{Var}(S_n/n)}{\varepsilon^2} \\
&= \frac{\operatorname{Var}(S_n)/n^2}{\varepsilon^2} \qquad \text{(see equation (\$) page 14)} \\
&\leq \frac{\sqrt{n}/n}{\varepsilon^2} \qquad \text{(previous question)} \\
&\overset{n\to+\infty}{\longrightarrow} 0.
\end{aligned}$$