Stochastic Methods for Engineering

Answers to LN Exercises

2.5.1. i) We have

$$\rho(aX+b, cY+d) = \frac{\mathrm{Cov}(aX+b, cY+d)}{\mathbb{V}[aX+b]^{1/2}\mathbb{V}[cY+d]^{1/2}}.$$

Now,

$$\mathrm{Cov}(aX+b, cY+d) = \mathbb{E}[a(X-\mathbb{E}[X])c(Y-\mathbb{E}[Y])] = ac\,\mathrm{Cov}(X,Y),$$

and

$$\mathbb{V}[aX+b] = \mathrm{Cov}(aX+b, aX+b) = a^2\mathbb{V}[X],$$

so

$$\rho(aX+b, cX+d) = \frac{ac\,\mathrm{Cov}(X,Y)}{|a||c|\mathbb{V}[X]^{1/2}\mathbb{V}[Y]^{1/2}} = \mathrm{sgn}(ac)\rho(X,Y).$$

ii) If $Y = aX+b$ we have

$$\rho(X,Y) = \rho(X, aX+b) = \mathrm{sgn}(a)\rho(X,X) = \mathrm{sgn}(a) = \pm 1,$$

being $\rho(X,X) = 1$. Vice versa, assume $\rho(X,Y) = \pm 1$, that is

$$\frac{\mathrm{Cov}(X,Y)}{\mathbb{V}[X]^{1/2}\mathbb{V}[Y]^{1/2}} = \pm 1.$$

Since

$$|\mathrm{Cov}(X,Y)| = |\mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])]| \overset{CS}{\leqslant} \mathbb{E}[(X-\mathbb{E}[X])^2]^{1/2}\mathbb{E}[(Y-\mathbb{E}[Y])^2]^{1/2} = \mathbb{V}[X]^{1/2}\mathbb{V}[Y]^{1/2},$$

and equality holds iff one or both $X-\mathbb{E}[X]$, $Y-\mathbb{E}[Y]$ vanish or if they are linearly dependent. In the first case, one or both X and Y should be constant, but being X, Y non costant by assumption, it must be they are linearly inedependent, so for example

$$Y - \mathbb{E}[X] = \lambda(X - \mathbb{E}[X]).$$

From this $Y = aX+b$ for suitable $a, b \in \mathbb{R}$. □

2.5.2. On $(\Omega, \mathscr{F}, \mathbb{P}) = ([-1,1], \mathscr{B}_{\mathbb{R}}, \frac{1}{2}\lambda_1)$ take

$$Y(\omega) = \omega,$$

and $X = Y^2$. In this way $\rho(X, Y^2) = \rho(Y^2, Y^2) = 1$. Clearly $\mathbb{E}[Y] = 0$ and

$$\mathrm{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \mathbb{E}[XY] = \mathbb{E}[Y^3] = 0,$$

from which $\rho(X,Y) = 0$. □

2.5.3. i) $\emptyset, X \in \mathcal{S}$: indeed, $\mu_X(\emptyset) = 0 = \mu_Y(\emptyset)$ and $\mu_X(\mathbb{R}) = 1 = \mu_Y(\mathbb{R})$. If $E \in \mathcal{S}$ then $\mu_X(E^c) = 1 - \mu_X(E) = 1 - \mu_Y(E) = \mu_Y(E^c)$, so als $E^c \in \mathcal{S}$. Finally, if $(E_n) \subset \mathcal{S}$. We can always transform into a disjoint union

$$\bigcup_n E_n = \bigsqcup_n F_n,$$

setting $F_1 = E_1$ and, for $n \geqslant 2$, $F_n = F_{n-1} \cup (E_n \setminus F_{n-1})$. Therefore

$$\mu_X\left(\bigcup_n E_n\right) = \mu_X\left(\bigsqcup_n F_n\right) = \sum_n \mu_X(F_n) = \sum_n \mu_Y(F_n) = \mu_Y\left(\bigsqcup_n F_n\right) = \mu_Y\left(\bigcup_n E_n\right).$$

This proves that $\mathcal{S}$ is a $\sigma-$algebra.

ii) Since $\mathcal{S} \supset$ intervals, it contains the $\sigma-$algebra generated by intervals, which is $\mathscr{B}_{\mathbb{R}}$, that is $\mathcal{S} \supset \mathscr{B}_{\mathbb{R}}$. Therefore, $\mu_X(E) = \mu_Y(E)$ for every $E \in \mathscr{B}_{\mathbb{R}}$. □

3.4.1. We have

$$\mathbb{P}(a \leqslant X \leqslant b) = \mathbb{P}\left(\{X \leqslant b\} \setminus \{X < a\}\right),$$

and since $\{X < a\} \subset \{X \leqslant b\}$ (here $a < b$) and $\mathbb{P}$ is a probability measure, subtractivity holds and

$$\mathbb{P}\left(\{X \leqslant b\} \setminus \{X < a\}\right) = \mathbb{P}\left(\{X \leqslant b\}\right) - \mathbb{P}\left(\{X < a\}\right) = F_X(b) - \mathbb{P}\left(\{X < a\}\right).$$

From the continuity from below of $\mathbb{P}$,

$$\mathbb{P}\left(\{X < a\}\right) = \lim_{x \to a-} \mathbb{P}(X \leqslant x) = \lim_{x \to a-} F_(x),$$

so

$$\mathbb{P}(a \leqslant X \leqslant b) = F_X(b) - \lim_{x \to a-} F_X(x).$$

Similarly,

$$\mathbb{P}(a < X < b) = \mathbb{P}(X < b) - \mathbb{P}(X \leqslant a) = \lim_{x \to b-} F_X(x) - F_X(a),$$

and

$$\mathbb{P}(X \geqslant b) = 1 - \mathbb{P}(X < b) = 1 - \lim_{x \to b-} F_X(x)$$

and

$$\mathbb{P}(X = a) = \mathbb{P}\left(\{X \leqslant a\} \setminus \{X < a\}\right) = F_X(a) - \lim_{x \to a-} F_X(x). \quad \square$$

3.4.5. Let X be a.c. with density f_X. We notice that $X^2 \geqslant 0$ $\mathbb{P}-$a.s., so $F_{X^2}(x) = \mathbb{P}(X^2 \leqslant x) = 0$ for $x < 0$. For $x \geqslant 0$ we have

$$F_{X^2}(x) = \mathbb{P}(X^2 \leqslant x) = \mathbb{P}(-\sqrt{x} \leqslant X \leqslant \sqrt{x}) = \int_{-\sqrt{x}}^{\sqrt{x}} f_X(y)\, dy,$$

so, for $x > 0$,

$$\partial_x F_{X^2}(x) = \partial_x\left(\int_{-\sqrt{x}}^{\sqrt{x}} f_X(y)\, dy\right) = f_X(\sqrt{x})\frac{1}{2\sqrt{x}} - f(-\sqrt{x})\left(-\frac{1}{2\sqrt{x}}\right) = \frac{1}{2\sqrt{x}}\left(f_X(\sqrt{x}) + f_X(-\sqrt{x})\right).$$

Therefore,

$$\exists \partial_x F_{X^2}(x) = \begin{cases} 0, & x < 0, \\ \frac{1}{2\sqrt{x}}\left(f_X(\sqrt{x}) + f_X(-\sqrt{x})\right), & x > 0 \end{cases} =: f_{X^2}(x).$$

This says that X^2 is a.c. and f_{X^2} is its density. □

3.4.6. i) We have

$$F_Y(y) = \mathbb{P}(Y \leqslant y) = \mathbb{P}(\sqrt{|X|} \leqslant y) = \begin{cases} 0, & y < 0, \\ \\ \mathbb{P}(|X| \leqslant y^2) = \int_{-y^2}^{y^2} e^{-\frac{(x-m)^2}{2\sigma^2}} \frac{dx}{\sqrt{2\pi\sigma^2}}, & y > 0, \end{cases}$$

By the fundamental thm of integral calculus,

$$\exists \partial_y F_Y(y) = \begin{cases} 0, & y < 0, \\ \\ \frac{2y}{\sqrt{2\pi\sigma^2}} \left(e^{-\frac{(y^2-m)^2}{2\sigma^2}} + e^{-\frac{(y^2-m)^2}{2\sigma^2}} \right) & y \geqslant 0, \end{cases} =: f_Y(y).$$

ii) In this case $Y = \Phi(X)$ where $\Phi(x) = \frac{1}{1+e^{-x}}$ is bijective. Indeed,

$$y = \Phi(x), \overset{0<y<1}{\iff} 1 + e^{-x} = \frac{1}{y}, \iff e^{-x} = \frac{1}{y} - 1 = \frac{1-y}{y}, \iff x = -\log\frac{1-y}{y} = \Phi^{-1}(y).$$

Therefore

$$\begin{aligned} f_Y(y) &= f_X(\Phi^{-1}(y))|(\Phi^{-1})'(y)| = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\log\frac{1-y}{y}+m)^2}{2\sigma^2}} \left| -\frac{y}{1-y}\left(-\frac{1}{y^2}\right)\right| \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \frac{1}{y(1-y)} e^{-\frac{(\log\frac{1-y}{y}+m)^2}{2\sigma^2}} 1_{]0,1[}(y). \quad \square \end{aligned}$$

3.4.7. Let $Y = [X]$ with $X \sim \exp(\lambda)$. We remind that $F_X(x) := (1 - e^{-\lambda x})1_{[0,+\infty[}(x)$. Being $Y = [X]$ we have that $\mathbb{P}(Y < 0) = 0$, so $F_Y(y) = 0$ for every $y < 0$. Let $y \geqslant 0$. We notice that

$$F_Y(y) = \mathbb{P}(Y \leqslant y) = \mathbb{P}([X] \leqslant y) = \mathbb{P}([X] \leqslant [y]) = \mathbb{P}(X \leqslant [y]) = F_X([y]) = 1 - e^{-\lambda[y]}.$$

Therefore,

$$F_Y(y) = \left(1 - e^{-\lambda[y]}\right) 1_{y \geqslant 0}. \quad \square$$

3.4.8. Without loss of generality we may assume $x_0 = 0$, so $P_0 = (0, y_0)$. If $\theta \in]-\pi/2, \pi/2[$ is the angle made by r_θ and the $y-$axis, the abscissa of the intersection with the $x-$axis is $x = -y_0 \tan\theta$. Assuming $\Theta \sim U(]-\pi/2, \pi/2[)$ we have that $X = -y_0 \tan\Theta = \Phi(\Theta)$ and since Φ is invertible on $]-\pi/2, \pi/2[$ with inverse

$$\Phi^{-1}(x) = \arctan\frac{x}{y_0},$$

we get

$$f_X(x) = f_\Theta(\Phi^{-1}(x))|(\Phi^{-1})'(x)| = \frac{1}{\pi} 1_{]-\pi/2,\pi/2[}\left(\arctan\frac{x}{y_0}\right) \frac{1}{1+\left(\frac{x}{y_0}\right)^2} \frac{1}{y_0} = \frac{1}{\pi} \frac{y_0}{y_0^2 + x^2}. \quad \square$$

3.4.9. i) If $x^* \neq y^*$ say $x^* < y^*$ then, being $F_X \uparrow$, we have that $F_X(x^*) \leqslant F_X(y^*-)$. Therefore

$$]F_X(x^*-), F_X(x^*)[\cap]F_X(y^*-), F_X(y^*)[= \emptyset.$$

ii) For every I_{x^*} there is at least a rational $q_{x^*} \in I_{x^*}$. Since the (I_{x^*}) are disjoint, the correspondence $x^* \longmapsto I_{x^*} \longmapsto q^*_{x^*}$ is injective, so theset of discontinuity points $\{x^*\}$ is at most countable. □

4.3.1. See slides.

4.3.2. i) To be a probability density, $f_{X,Y}$ must be $\geqslant 0$ a.e. $(x, y) \in \mathbb{R}^2$ (this needs $c \geqslant 0$) and such that $\int_{\mathbb{R}^2} f_{X,Y} = 1$. Since

$$\int_{\mathbb{R}^2} f_{X,Y} = \int_0^1 \int_0^2 c\left(x^2 + \frac{xy}{2}\right)\, dy\, dx = c\int_0^1 2x^2 + \frac{x}{2}\left[\frac{y^2}{2}\right]_{y=0}^{y=2} dx = c\left(2\left[\frac{x^3}{3}\right]_{x=0}^{x=1} + [x^2]_{x=0}^{x=1}\right) = c\frac{5}{3},$$

from which $c = \frac{3}{5}$.

ii) The marginal density f_X is obteined by the formula

$$f_X(x) = \int_{\mathbb{R}} f_{X,Y}(x, y)\, dy = \frac{3}{5}1_{[0,1]}(x)\int_0^2 x^2+\frac{xy}{2}\, dy = \frac{3}{5}1_{[0,1]}(x)\left(2x^2 + \frac{x}{2}\left[\frac{y^2}{2}\right]_{y=0}^{y=2}\right) = \frac{6}{5}(x^2+x)1_{[0,1]}(x).$$

iii) We have

$$\mathbb{P}(X > Y) = \int_{x\geqslant y} f_{X,Y}(x, y)\, dxdy = \int_0^1 \int_0^x \frac{3}{5}(x^2 + \frac{xy}{2})\, dy\, dx = \frac{3}{5}\int_0^1 \frac{5}{4}x^3\, dx = \frac{3}{4}\left[\frac{x^4}{4}\right]_{x=0}^{x=1} = \frac{3}{16}. \quad \square$$

4.3.3. Let $f_{X,Y}(x, y) := e^{-(x+y)}1_{[0,+\infty[^2}(x, y)$. Since $f_{X,Y} \equiv 0$ on $\mathbb{R}^2\backslash[0, +\infty[^2$, we have that $\mathbb{P}((X, Y) \in \mathbb{R}^2_+) = 1$. We also notice that

$$\mathbb{P}(Y = 0) = \int_{x\in\mathbb{R}, y=0} f_{X,Y}(x, 0)\, dxdy = 0,$$

and, similarly $\mathbb{P}(X = 0) = 0$. Therefore $\mathbb{P}(Y \leqslant 0) = \mathbb{P}(X \leqslant 0) = 0$, so

$$F_{X/Y}(u) = \mathbb{P}(X/Y \leqslant u) = 0,\ \forall u \leqslant 0.$$

For $u > 0$, being $\mathbb{P}(Y > 0) = 1$ we have

$$\begin{aligned}
F_{X/Y}(u) &= \mathbb{P}(X/Y \leqslant u) = \mathbb{P}(X \leqslant uY) = \int_{x\leqslant uy} f_{X,Y}(x, y)\, dxdy = \int_{x\leqslant uy} e^{-(x+y)}1_{[0,+\infty[^2}(x, y)\, dxdy \\
&= \int_0^{+\infty}\int_{x/u}^{+\infty} e^{-(x+y)}dy\, dx = \int_0^{+\infty} e^{-x}\int_{x/u}^{+\infty} e^{-y}dy\, dx = \int_0^{+\infty} e^{-x}\,[-e^{-y}]_{x/u}^{+\infty}\, dx \\
&= \int_0^{+\infty} e^{-x}e^{-x/u}\, dx = \int_0^{+\infty} e^{-\left(1+\frac{1}{u}\right)x}\, dx = \left[-\frac{e^{-\left(1+\frac{1}{u}\right)x}}{1+\frac{1}{u}}\right]_{x=0}^{x=+\infty} = \frac{u}{u+1}.
\end{aligned}$$

Therefore,

$$F_{X/Y}(u) = \frac{u}{u+1}1_{[0,+\infty[}(u).$$

From this we see that $F_{X/Y}$ is a.e. differentiable with

$$\partial_u F_{X/Y}(u) = f_{X/Y}(u) = \frac{1}{(u+1)^2}1_{u\geqslant 0}. \quad \square$$

4.3.4. Clearly, $f \geqslant 0$. To be a probability density it must verify

$$1 = \int_{\mathbb{R}^3} f(x,y)\,dxdy = \int_{B(0,R]} c(R-\sqrt{x^2+y^2})\,dxdy \overset{pol.coords}{=} \int_{0\leqslant\rho\leqslant R,\ 0\leqslant\theta\leqslant 2\pi} c(R-\rho)\rho\,d\rho d\theta$$

$$= c2\pi\int_0^R (R-\rho)\rho\,d\rho = 2\pi c\left(R\frac{R^2}{2}-\frac{R^3}{3}\right) = \frac{2\pi R^3}{6}c$$

from which $c = \frac{3}{\pi R^3}$. Now, for $0 \leqslant a \leqslant R$ we have

$$\mathbb{P}((X,Y)\in B(0,a]) = \int_{B(0,a]} f(x,y)\,dxdy = \frac{6}{R^3}\int_0^a (R-\rho)\rho\,d\rho = \frac{6}{R^3}\left(R\frac{a^2}{2}-\frac{a^3}{3}\right) = \frac{a^2}{R^3}(3R-2a).$$

The distance from the centre of the target is $D := \sqrt{X^2+Y^2}$. The problem asks to compute the distribution of D. Starting from the cdf

$$F_D(a) = \mathbb{P}\left((X,Y)\in B(0,a]\right) = \begin{cases} 0, & a<0, \\ \frac{a^2}{R^3}(3R-2a), & 0\leqslant a\leqslant R, \\ 1, & a\geqslant R. \end{cases}$$

Clearly, F_D is differentiable for $a \neq 0, R$, so D is absolutely continuous with density

$$f_D(a) = \partial_a F_D(a) = \begin{cases} 0, & a<0,\ a>R \\ \frac{2a}{R^3}(3R-2a)+\frac{a^2}{R^3}(-2) = \frac{2a}{R^3}(3R-a), & 0<a<R, \end{cases} = \frac{6a}{R^3}(R-a)1_{[0,R]}(a).$$

Finally,

$$\mathbb{E}[D] = \int_{\mathbb{R}} a f_D(a)\,da = \int_0^R \frac{6a^2}{R^3}(R-a)\,da = \frac{6}{R^3}\left(R\frac{R^3}{3}-\frac{R^4}{4}\right) = \frac{R}{2}. \qquad \square$$

4.3.9. i) We notice that

$$x^2-2\rho xy+y^2 = \underbrace{\begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix}}_{=:M}\begin{pmatrix} x \\ y \end{pmatrix}\cdot\begin{pmatrix} x \\ y \end{pmatrix}$$

where M is a 2×2 matrix with $\det M = 1-\rho^2$. Therefore

$$\frac{1}{1-\rho^2}M = C^{-1},\quad C = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix},$$

with C positive definite (provided $\rho^2 \leqslant 1$) and symmetric matrix. Therefore, if $v = \binom{x}{y}$, we have

$$f(v) = ce^{-\frac{1}{2}C^{-1}v\cdot v},$$

so f is a Gaussian density with $c = \frac{1}{\sqrt{(2\pi)^2(1-\rho^2)}}$.

ii) Let $(X,Z) = (X, \frac{Y-\rho X}{\sqrt{1-\rho^2}}) = \Psi(X,Y)$. Notice that Ψ is a linear transformation of $\mathbb{R}^2$ into itself, and

$$\Psi'(x,y) = \begin{bmatrix} 1 & 0 \\ -\frac{\rho}{\sqrt{1-\rho^2}} & \frac{1}{\sqrt{1-\rho^2}} \end{bmatrix}, \implies |\det \Psi'| = \frac{1}{\sqrt{1-\rho^2}}, \implies |\det(\Psi^{-1})'| = \sqrt{1-\rho^2}.$$

Moreover

$$(x,z) = \Psi(x,y), \iff \begin{cases} x = x, \\ z = \frac{y-\rho x}{\sqrt{1-\rho^2}}, \end{cases} \iff \begin{cases} x = x, \\ y = \sqrt{1-\rho^2} z + \rho x, \end{cases}$$

Therefore, if $v = (x,z)$,

$$f_{X,Z}(x,z) = f_{X,Y}(x, \sqrt{1-\rho^2} z + \rho x)\sqrt{1-\rho^2} = \frac{1}{\sqrt{(2\pi)^2}} e^{-\frac{1}{2} C^{-1} T v \cdot T v}$$

where

$$T = \begin{bmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{bmatrix}$$

We notice that $TT^\top = C$, so

$$T^\top C^{-1} T = T^\top (TT^\top)^{-1} T = T^\top (T^\top)^{-1} T^{-1} T = \mathbb{I}_2,$$

from which

$$f_{X,Z}(x,z) = \frac{1}{\sqrt{(2\pi)^2}} e^{-\frac{1}{2}(x^2+z^2)}.$$

This shows that $(X,Z) \sim \mathcal{N}(0, \mathbb{I}_2)$. Clearly $f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$, that is $Z \sim \mathcal{N}(0,1)$.

iii) We have

$$\begin{aligned} \mathbb{P}(X>0, Y>0) &= \mathbb{P}\left(X>0,\ \sqrt{1-\rho^2} Z + \rho X > 0\right) = \int_{x>0,\ z>-\frac{\rho}{\sqrt{1-\rho^2}} x} e^{-\frac{x^2+z^2}{2}} \frac{dxdz}{\sqrt{(2\pi)^2}} \\ &= \frac{1}{2\pi} \int_0^{+\infty} e^{-\frac{x^2}{2}} \int_{-\frac{\rho}{\sqrt{1-\rho^2}} x}^{+\infty} e^{-\frac{z^2}{2}}\, dz\, dx = \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} e^{-\frac{x^2}{2}} \left(1 - \Phi\left(-\frac{\rho}{\sqrt{1-\rho^2}} x\right)\right) dx \\ &= \frac{1}{2} - \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} e^{-\frac{x^2}{2}} \Phi\left(-\frac{\rho}{\sqrt{1-\rho^2}} x\right) dx. \end{aligned}$$

where $\Phi(u) := \int_{-\infty}^u e^{-\frac{z^2}{2}} \frac{dz}{\sqrt{2\pi}}$ is the cdf of the standard Gassian. Let, for brevity, $r := -\frac{\rho}{\sqrt{1-\rho^2}}$. We notice that $\Phi'(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$. Therefore

$$\frac{1}{\sqrt{2\pi}} \int_0^{+\infty} e^{-\frac{x^2}{2}} \Phi\left(-\frac{\rho}{\sqrt{1-\rho^2}} x\right) dx = \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} e^{-\frac{x^2}{2}} \Phi(rx)\, dx =: I(r).$$

where $r=-\frac{\rho}{\sqrt{1-\rho^2}}$. We notice that, differentiating under integral sign, being $\Phi'(u)=e^{-\frac{u^2}{2}}\frac{1}{\sqrt{2\pi}}$,

$$\partial_r I(r)=\int_0^{+\infty}e^{-\frac{x^2}{2}}\frac{1}{\sqrt{2\pi}}e^{-\frac{r^2x^2}{2}}x\,\frac{dx}{\sqrt{2\pi}}=\frac{1}{2\pi}\int_0^{+\infty}\underbrace{xe^{-\frac{(1+r^2)x^2}{2}}}_{=\partial_x-\frac{1}{1+r^2}e^{-(1+r^2)\frac{x^2}{2}}}\,dx=\frac{1}{(2\pi)(1+r^2)}.$$

From this,

$$I(r)=\frac{1}{2\pi}\arctan r+c,$$

and since $I(0)=\frac{1}{\sqrt{2\pi}}\int_0^{+\infty}e^{-\frac{x^2}{2}}\Phi(0)\,dx=\frac{1}{4}$, we get

$$I(r)=\frac{1}{2\pi}\arctan r+\frac{1}{4}.$$

Therefore

$$\mathbb{P}(X>0,Y>0)=\frac{1}{2}-\left(\frac{1}{2\pi}\arctan\left(-\frac{\rho}{\sqrt{1-\rho^2}}\right)+\frac{1}{4}\right)=\frac{1}{4}+\frac{1}{2\pi}\arctan\frac{\rho}{\sqrt{1-\rho^2}}.\qquad\square$$

5.3.1. For a standard Bernoulli r.v., $\mathbb{P}(X=0)=p$, $\mathbb{P}(X=1)=1-p$ we have

$$\phi_X(\xi)=\mathbb{E}[e^{i\xi X}]=e^{i\xi 0}\mathbb{P}(X=0)+e^{i\xi 1}\mathbb{P}(X=1)=p+(1-p)e^{i\xi}.$$

For a binomial r.v. $X\in\{0,\dots,n\}$, $\mathbb{P}(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$ (with $p\in[0,1]$), we have

$$\begin{aligned}\phi_X(\xi)&=\mathbb{E}[e^{i\xi X}]=\sum_{k=0}^n e^{i\xi k}\mathbb{P}(X=k)=\sum_{k=0}^n\binom{n}{k}p^k(1-p)^{n-k}e^{i\xi k}=\sum_{k=0}^n\binom{n}{k}(e^{i\xi}p)^k(1-p)^{n-k}\\&=\left(1-p+pe^{i\xi}\right)^n.\end{aligned}$$

For a Poisson r.v. $X\in\mathbb{N}$, $\mathbb{P}(X=k)=e^{-\lambda}\frac{\lambda^k}{k!}$, so

$$\phi_X(\xi)=\mathbb{E}[e^{i\xi X}]=\sum_{k=0}^\infty e^{i\xi k}e^{-\lambda}\frac{\lambda^k}{k!}=e^{-\lambda}\sum_{k=0}^\infty\frac{(\lambda e^{i\xi})^k}{k!}=e^{-\lambda}e^{\lambda e^{i\xi}}=e^{\lambda(e^{i\xi}-1)}.\qquad\square$$

5.3.2. Since $f_X\in L^1(\mathbb{R})$, we have

$$\phi_X(\xi)=\widehat{f_X}(-\xi)=\frac{1}{\pi}\widehat{\frac{a}{a^2+(\sharp-m)^2}}(-\xi)=e^{i\xi m}\frac{a}{\pi}\widehat{\frac{1}{a^2+\sharp^2}}(-\xi)=e^{i\xi m-a|\xi|}.\qquad\square$$

5.3.3. If $f_X(x)=\frac{\lambda^{\alpha-1}}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x}1_{[0,+\infty[}(x)$, then

$$\begin{aligned}\phi_X(\xi)&=\widehat{f_X}(-\xi)=\frac{\lambda^{\alpha-1}}{\Gamma(\alpha)}\int_0^{+\infty}x^{\alpha-1}e^{-\lambda x}e^{i\xi x}\,dx=\frac{\lambda^{\alpha-1}}{\Gamma(\alpha)}\int_0^{+\infty}x^{\alpha-1}e^{-(\lambda-i\xi)x}\,dx\\&=\frac{\lambda^{\alpha-1}}{\Gamma(\alpha)}\Gamma(\alpha)(\lambda-i\xi)^{-\alpha}=\frac{1}{\lambda}\left(1-\frac{i\xi}{\lambda}\right)^{-\alpha}.\end{aligned}$$

5.3.6. Since $\mu(E) = \mathbb{E}[Y 1_E(X)]$, that is

$$\int_{\mathbb{R}} 1_E \, d\mu = \mathbb{E}[Y 1_E(X)].$$

By linearity,

$$\int_{\mathbb{R}} s \, d\mu = \mathbb{E}[Y s(X)],$$

for every simple function s. By standard approximation, we get

$$\int_{\mathbb{R}} \varphi \, d\mu = \mathbb{E}[Y \varphi(X)], \; \forall \varphi \in L^\infty.$$

In particular, setting $\varphi(x) = e^{i\xi x}$ we get

$$\int_{\mathbb{R}} e^{i\xi x} \, d\mu(x) = \mathbb{E}[Y e^{i\xi X}] = 0, \; \forall \xi \in \mathbb{R},$$

so $\widehat{\mu}(-\xi) \equiv 0$. From injectivity it follows that $\mu = 0$. The second part ($Y = 0$) is not true unless $Y = f(X)$.

5.3.7. i) We have

$$|\widehat{\mu}(\xi)| = \left| \int_{\mathbb{R}} e^{-i\xi x} \, d\mu(x) \right| \leqslant \int_{\mathbb{R}} |e^{-i\xi x}| \, d\mu = 1.$$

ii) We have

$$\overline{\widehat{\mu}(\xi)} = \overline{\int_{\mathbb{R}} e^{-i\xi x} \, d\mu(x)} = \int_{\mathbb{R}} \overline{e^{-i\xi x}} \, d\mu(x) = \int_{\mathbb{R}} e^{i\xi x} \, d\mu(x) = \widehat{\mu}(-\xi).$$

iii) We have

$$\sum_{j,k} \widehat{\mu}(\xi_j - \xi_k) z_j \overline{z_k} = \int_{\mathbb{R}} \sum_{j,k} e^{-i(\xi_j - \xi_k)x} z_j \overline{z_k} \, d\mu(x) = \int_{\mathbb{R}} \underbrace{\sum_j e^{-i\xi_j x} z_j}_{=:w(x)} \underbrace{\sum_k e^{i\xi_k x} \overline{z_k}}_{=:\overline{w(x)}} \, d\mu(x) \geqslant 0.$$

iv) We have

$$\widehat{\mu}(\xi) = \int_{\mathbb{R}} e^{-i\xi x} \, d\mu(x) = \int_{\mathbb{R}} f(\xi, x) \, d\mu(x).$$

We notice that $f(\sharp, x) \in \mathscr{C}(\mathbb{R})$ for μ–a.e. $x \in \mathbb{R}$, and $|f(\xi, x)| = 1 \in L^1(\mathbb{R}, \mu)$. By continuity of integrals depending on parameters, $\widehat{\mu} \in \mathscr{C}(\mathbb{R})$. □

5.3.8. Let $\phi_X(\xi) = \mathbb{E}[e^{i\xi X}] = \widehat{f_X}(-\xi)$. Then

$$|\phi_X(\xi)|^2 = \phi_X(\xi)\overline{\phi_X(\xi)} = \widehat{f_X}(-\xi)\overline{\widehat{f_X}(-\xi)} = \widehat{f_X(-\sharp)}(\xi)\widehat{f_X}(\xi) = \widehat{f_X(-\sharp) * f_X}(\xi) = \widehat{(f_X(-\sharp) * f_X)\,(-\sharp)}(-\xi).$$

Now, let

$$f(y) := (f_X(-\sharp) * f_X)(-y) = \int_{\mathbb{R}} f_X(-x) f_X(-y - x) \, dx = \int_{\mathbb{R}} f_X(x) f_X(-y + x) \, dx.$$

By Young theorem, being $f_X \in L^1$, f is well defined and L^1. Clearly, $f \geqslant 0$ (being $f_X \geqslant 0$) and

$$\begin{aligned}\int_{\mathbb{R}} f(y)\,dy &= \int_{\mathbb{R}}\int_{\mathbb{R}} f_X(x)f_X(-y+x)\,dx\,dy = \int_{\mathbb{R}}\int_{\mathbb{R}} f_X(x)f_X(-y+x)\,dy\,dx \\ &= \int_{\mathbb{R}} f_X(x) \underbrace{\int_{\mathbb{R}} f_X(-y+x)\,dy}_{=1}\,dx = \int_{\mathbb{R}} f_X(x)\,dx = 1,\end{aligned}$$

so f is a probability density. As well known, there is Y r.v. such that $f_Y = f$ and the conclusion follows. □

5.3.9. If $d\mu_X = f_X(x)\,dx$, then

$$\widehat{\mu_X}(\xi) = \widehat{f_X}(-\xi).$$

Known that $\widehat{\mu_X} \in L^1(\mathbb{R})$, inversion formula would apply and

$$\widehat{f_X}(\xi) = \widehat{\mu_X}(-\xi), \implies f_X(x) = \frac{1}{2\pi}\widehat{\widehat{f_X}}(-x) = \frac{1}{2\pi}\widehat{\widehat{\mu_X}(-\sharp)}(-x).$$

This formula yileds the possible f_X. This is the starting point. Let

$$f_X(x) := \frac{1}{2\pi}\widehat{\widehat{\mu_X}(-\sharp)}(-x).$$

The goal is to check that $d\mu_X = f_X(x)\,dx$. We notice that, by the duality Lemma, if $\psi, \widehat{\psi} \in L^1$ we have

$$\begin{aligned}\int_{\mathbb{R}} \psi(x)f_X(x)\,dx &= \int_{\mathbb{R}} \psi(x)\frac{1}{2\pi}\widehat{\widehat{\mu_X}(-\sharp)}(-x)\,dx = \frac{1}{2\pi}\int_{\mathbb{R}} \psi(-x)\widehat{\widehat{\mu_X}(-\sharp)}(x)\,dx \\ &= \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{\psi(-\sharp)}(\xi)\widehat{\mu_X}(-\xi)\,d\xi = \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{\psi}(-\xi)\widehat{\mu_X}(\xi)\,d\xi \\ &\overset{duality}{=} \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{\widehat{\psi(-\sharp)}}(x)\,d\mu_X(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{\widehat{\psi}}(-x)\,d\mu_X(x) \\ &\overset{inversion}{=} \int_{\mathbb{R}} \psi(x)\,d\mu_X(x).\end{aligned}$$

Therefore,

$$\int_{\mathbb{R}} \psi(x)f_X(x)\,dx = \int_{\mathbb{R}} \psi(x)\,d\mu_X(x),\ \forall\psi\ :\ \psi, \widehat{\psi} \in L^1(\mathbb{R}).$$

In particular, this holds for every $\psi \in \mathcal{S}(\mathbb{R})$, and by standard approximation arguments, for every Borel function. □

6.4.1. We notice that

$$\begin{aligned}\mathbb{P}(X > x, Y > y) &= 1 - \mathbb{P}(\{X \leqslant x\} \cup \{Y \leqslant y\}) = 1 - (\mathbb{P}(X \leqslant x) + \mathbb{P}(Y \leqslant y) - \mathbb{P}(X \leqslant x, Y \leqslant y)) \\ &= 1 - \left(F_X(x) + F_Y(y) - F_{X,Y}(x,y)\right).\end{aligned}$$

On the other side,

$$\mathbb{P}(X > x)\mathbb{P}(Y > y) = (1 - \mathbb{P}(X \leqslant x))\,(1 - \mathbb{P}(Y \leqslant y)) = 1 - F_X(x) - F_Y(y) + F_X(x)F_Y(y),$$

and by the assumption

$$1 - \big(F_X(x) + F_Y(y) - F_{X,Y}(x, y)\big) = 1 - F_X(x) - F_Y(y) + F_X(x)F_Y(y),$$

from which

$$F_{X,Y}(x, y) = F_X(x)F_Y(y),\ \forall (x, y) \in \mathbb{R}^2,$$

which is equivalent to the independence of X and Y. □

6.4.3. i) We have $f_{-Y}(y) = f_Y(-y) = e^{-2|-y|} = e^{-2|y|} = f_Y(y)$. The characteristic function is

$$\begin{aligned}\phi_{-Y}(\xi) &= \phi_Y(\xi) = \int_{\mathbb{R}} e^{i\xi y} e^{-2|y|}\,dy = \int_{-\infty}^{0} e^{(i\xi+2)y}\,dy + \int_{0}^{+\infty} e^{(i\xi-2)y}\,dy \\ &= \left[\frac{e^{(i\xi+2)y}}{i\xi+2}\right]_{y=-\infty}^{y=0} + \left[\frac{e^{(i\xi-2)y}}{i\xi-2}\right]_{y=0}^{y=+\infty} = \frac{1}{i\xi+2} - \frac{1}{i\xi-2} = \frac{4}{4+\xi^2}.\end{aligned}$$

ii) Since $Z = X - Y$ and X, Y are independent,

$$\phi_Z(\xi) = \phi_{X-Y}(\xi) = \phi_X(\xi)\phi_{-Y}(\xi) = \left(\frac{4}{4+\xi^2}\right)^2.$$

As we can see, Z is not a Laplace random variable. □

6.4.5. Let $X \in [0, a]$ and $Y \in [0, b]$. Then, the area of the triangle is $T_{X,Y} = \frac{1}{2}XY$. So, being X, Y independent (so $f_{X,Y}(x, y) = f_X(x)f_Y(y) = \frac{1}{a}1_{[0,a]}(x)\frac{1}{b}1_{[0,b]}(y)$), the required probability is

$$\begin{aligned}\mathbb{P}\left(T_{X,Y} \geqslant \frac{1}{4}ab\right) &= \mathbb{P}\left(XY \geqslant \frac{ab}{2}\right) = \int_{xy \geqslant ab/2} \frac{1}{ab} 1_{[0,a]}(x)1_{[0,b]}(y)\,dxdy \\ &= \frac{1}{ab}\int_{0\leqslant x\leqslant a,\ 0\leqslant y\leqslant b,\ xy\geqslant ab/2} 1\,dxdy \\ &= \frac{1}{ab}\int_{a/2}^{a}\int_{ab/2x}^{b} 1\,dy\,dx = \frac{1}{ab}\int_{a/2}^{a}\left(b - \frac{ab}{2x}\right)dx = \int_{a/2}^{a}\left(\frac{1}{a} - \frac{1}{2x}\right)dx \\ &= \frac{1}{2}\left(1 - [\log x]_{x=a/2}^{x=a}\right) = \frac{1+\log 2}{2}.\end{aligned}$$ □

6.4.8. i) Let $S_n := \min(T_1, \dots, T_{n-1})$. Let

$$F_{S_n}(t) = \mathbb{P}(\min(T_1, \dots, T_{n-1}) \leqslant t).$$

Since $0 \leqslant T_j \leqslant 1$ for every j, $F_{S_n}(t) = 0$ if $t < 0$ and $F_{S_n}(t) = 1$ if $t \geqslant 1$. For $0 \leqslant t < 1$,

$$F_{S_n}(t) = 1 - \mathbb{P}(\min(T_1, \dots, T_{n-1}) > t) = 1 - \prod_{k=1}^{n-1}\underbrace{\mathbb{P}(T_k > t)}_{=1-t} = 1 - (1-t)^{n-1}.$$

In conclusion

$$F_{S_n}(t) = \begin{cases} 0, & t < 0, \\ 1 - (1-t)^{n-1}, & 0 \leqslant t < 1, \\ 1, & t \geqslant 1. \end{cases}$$

ii) Since $S_n := \min(T_1, \dots, T_{n-1})$ depends on $T_1, \dots, T_{n-1}$, which are independent of T_n, we have that S_n and T_n are independent, and the conclusion follows.

iii) Let A_n be the event "a new record is set in the $n-$th race. We can write this as

$$A_n = \{T_n < \min(T_1, \dots, T_n)\} = \{T_n < S_n\}.$$

Therefore

$$\mathbb{P}(A_n) = \int_{t<s} f_{T_n, S_n}(t,s)\, dtds \overset{indep}{=} \int_{t<s} f_{T_n}(t) f_{S_n}(s)\, dtds.$$

We notice that $f_{T_n}(t) = 1_{[0,1]}(t)$ while

$$f_{S_n}(s) = 1_{[0,1]}(n-1)(1-s)^{n-2},$$

so

$$\begin{aligned} \mathbb{P}(A_n) &= \int_{0\leqslant t<s\leqslant 1} (n-1)(1-s)^{n-2}\, dtds = \int_0^1 \int_t^1 (n-1)(1-s)^{n-2}\, ds\, dt \\ &= \int_0^1 \left[(1-s)^{n-1}\right]_{s=t}^{s=1}\, dt = \int_0^1 -(1-t)^{n-1}\, dt = \left[\frac{(1-t)^n}{n}\right]_{t=0}^{t=1} = \frac{1}{n}. \end{aligned}$$

iv) The event "a record remains unbroken" is $A := A_1 \cap \bigcap_{n\geqslant 2} A_n^c$. By independence,

$$\mathbb{P}(A) = \mathbb{P}(A_1) \prod_{n\geqslant 2} \mathbb{P}(A_n^c) = \prod_{n\geqslant 2}\left(1 - \frac{1}{n}\right) = e^{\sum_{n\geqslant 2} \log\left(1-\frac{1}{n}\right)},$$

and recalling of the inequality $\log(1+x) \leqslant x$, we have

$$\mathbb{P}(A) \leqslant e^{-\sum_{n\geqslant 2}\frac{1}{n}} = e^{-\infty} = 0. \quad \square$$

7.3.3. Let X be such that $\mathbb{P}(X = n) = \frac{\lambda_1^n}{n!} e^{-\lambda_1}$ and Y with $\mathbb{P}(Y = m) = \frac{\lambda_2^m}{m!} e^{-\lambda_2}$. To compute the conditional expectation, since $X + Y$ is a discrete random variable, $\sigma(X + Y)$ is generated by the events $\{X + Y = k\}$ who form a partition of the sample space Ω. Therefore

$$\mathbb{E}[X \mid X+Y] = \sum_{k=0}^{\infty} \frac{1}{\mathbb{P}(X+Y=k)} \mathbb{E}[X 1_{X+Y=k}] 1_{X+Y=k}.$$

Now,

$$\begin{aligned} \mathbb{P}(X+Y=k) &= \sum_{j=0}^{k} \mathbb{P}(X+Y=k, Y=j) = \sum_{j=0}^{k} \mathbb{P}(X=k-j, Y=j) \overset{indep}{=} \sum_{j=0}^{k} \mathbb{P}(X=k-j)(Y=j) \\ &= \sum_{j=0}^{k} \frac{\lambda_1^{k-j}}{(k-j)!} e^{-\lambda_1} \frac{\lambda_2^j}{j!} e^{-\lambda_2} = e^{-(\lambda_1+\lambda_2)} \frac{1}{k!} \sum_{j=0}^{k} \binom{k}{j} \lambda_1^{k-j} \lambda_2^j = e^{-(\lambda_1+\lambda_2)} \frac{(\lambda_1+\lambda_2)^k}{k!}. \end{aligned}$$

Similarly

$$\begin{aligned}\mathbb{E}[X1_{X+Y=k}] &= \sum_{j=0}^{k}\mathbb{E}[X1_{X=k-j}1_{Y=j}] = \sum_{j=0}^{k}(k-j)\frac{\lambda_1^{k-j}}{(k-j)!}e^{-\lambda_1}\frac{\lambda_2^j}{j!}e^{-\lambda_2}\\ &= \lambda_1 e^{-(\lambda_1+\lambda_2)}\sum_{j=0}^{k-1}\frac{\lambda_1^{k-1-j}\lambda_2^j}{(k-1-j)!j!} = \lambda_1 e^{-(\lambda_1+\lambda_2)}\frac{(\lambda_1+\lambda_2)^{k-1}}{(k-1)!}.\end{aligned}$$

Therefore

$$\mathbb{E}[X \mid X+Y] = \sum_{k=0}^{\infty}\frac{1}{e^{-(\lambda_1+\lambda_2)}\frac{(\lambda_1+\lambda_2)^k}{k!}}\lambda_1 e^{-(\lambda_1+\lambda_2)}\frac{(\lambda_1+\lambda_2)^{k-1}}{(k-1)!}1_{X+Y=k} = \frac{\lambda_1}{\lambda_1+\lambda_2}\sum_k k1_{X+Y=k}. \quad \square$$

7.3.4. We start noticing that

$$\mathbb{E}[X \mid Y=y] = \mathbb{E}[(X-m_X)+m_X \mid Y-m_Y = y-mY] = m_X + \mathbb{E}[X-m_X \mid Y-m_Y = y-mY],$$

so we are reduced to the case $m_X = m_Y = 0$. We have

$$\mathbb{E}[X \mid Y=y] = \int_{\mathbb{R}} x\frac{f_{X,Y}(x,y)}{f_Y(y)}\,dx = \frac{1}{\sqrt{2\pi\frac{\det C}{c_{22}}}}e^{\frac{y^2}{2c_{22}}}\int_{\mathbb{R}} xe^{-\frac{1}{2}C^{-1}\binom{x}{y}\cdot\binom{x}{y}}dx$$

Now,

$$C^{-1} = \frac{1}{\det C}\begin{bmatrix} c_{22} & -c_{12}\\ -c_{12} & c_{11}\end{bmatrix}$$

so,

$$C^{-1}\binom{x}{y}\cdot\binom{x}{y} = \frac{1}{\det C}\left(c_{22}x^2 + c_{11}y^2 - 2c_{12}xy\right).$$

Therefore,

$$\int_{\mathbb{R}} xe^{-\frac{1}{2}C^{-1}\binom{x}{y}\cdot\binom{x}{y}}\,dx = \int_{\mathbb{R}} xe^{-\frac{x^2-2x\frac{c_{12}}{c_{22}}y+\left(\frac{c_{12}}{c_{22}}y\right)^2}{2\det C/c_{22}}}\,dx\ e^{-\frac{c_{11}y^2}{2\det C}}e^{\frac{c_{12}^2y^2}{2c_{22}\det C}}$$

from which

$$\begin{aligned}\mathbb{E}[X \mid Y=y] &= \frac{1}{\sqrt{2\pi\frac{\det C}{c_{22}}}}\exp\left(\frac{1}{2}\frac{\det C - c_{11}c_{22}+c_{12}^2}{c_{22}\det C}y^2\right)\int_{\mathbb{R}} xe^{-\frac{\left(x-\frac{c_{12}}{c_{22}}y\right)^2}{2\det C/c_{22}}}\,dx\\ &= \frac{c_{12}}{c_{22}}y.\end{aligned}$$

Returning to the conditional expectation we get

$$\mathbb{E}[X \mid Y=y] = m_X + \frac{c_{12}}{c_{22}}(y-m_Y),$$

from which we finally have

$$\mathbb{E}[X \mid Y] = m_X + \frac{c_{12}}{c_{22}}(Y-m_Y). \quad \square$$

7.3.5. Let $(Z,W) := (X-Y, X+Y) = \Psi(X,Y)$. Then

$$\mathbb{E}[X-Y \mid X+Y=w] = \int_{\mathbb{R}} z f_{Z|W}(z \mid w)\, dz = \frac{1}{f_W(w)} \int_{\mathbb{R}} z f_{Z,W}(z,w)\, dz,$$

if $f_W(w) \neq 0$, and 0 elsewhere. Now,

$$f_{Z,W}(z,w) = f_{X,Y}(\Psi^{-1}(z,w))|\det(\Psi^{-1})'(z,w)|.$$

We have

$$\Psi \ : \ \begin{cases} z = x-y, \\ \\ w = x+y, \end{cases} \iff \Psi^{-1} \ : \ \begin{cases} x = \frac{z+w}{2}, \\ \\ y = \frac{w-z}{2}, \end{cases}$$

and

$$|\det(\Psi^{-1})'(z,w)| = \left| \det \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \right| = \frac{1}{2}.$$

Therefore,

$$f_{Z,W}(z,w) = \frac{1}{2} f_{X,Y}\left(\frac{w+z}{2}, \frac{w-z}{2}\right) = \frac{1}{2} f\left(\frac{w+z}{2}\right) f\left(\frac{w-z}{2}\right),$$

so

$$I := \int_{\mathbb{R}} z f_{Z,W}(z,w)\, dz = \frac{1}{2}\int_{\mathbb{R}} z f\left(\frac{w+z}{2}\right) f\left(\frac{w-z}{2}\right)\, dz = -\frac{1}{2}\int_{\mathbb{R}} z f\left(\frac{w-z}{2}\right) f\left(\frac{w+z}{2}\right)\, dz = -I$$

from which $I = 0$. We conclude that

$$\mathbb{E}[X-Y \mid X+Y=w] = 0,$$

from which $\mathbb{E}[X-Y \mid X+Y] = 0$. □

7.3.6. Let $Z := \mathbb{E}[X \mid Y]$. We have to check that

$$\mathbb{E}[X \mid Z] = Z, \quad \iff \quad \mathbb{E}[X 1_F] = \mathbb{E}[Z 1_F],\ \forall F \in \sigma(Z).$$

Let $F \in \sigma(Z)$. Since $Z = \mathbb{E}[X \mid Y]$ is, in particular, $Y-$measurable, $F \in \sigma(Y)$. So,

$$\mathbb{E}[X1_F] = \mathbb{E}[\mathbb{E}[X \mid Y]1_F] = \mathbb{E}[Z1_F],$$

and since $Z \in \sigma(Z)$ we conclude that $\mathbb{E}[X \mid Z] = Z$, from which the conclusion follows. □

7.3.7. Let $(X_n) \uparrow$, $X_n \geqslant 0$. Let $X := \lim_n X_n$, the limit existing because $X_n \uparrow$ with n. Moreover, being $(X_n) \subset L(\Omega)$, also $X \in L(\Omega)$ and since $X_n \geqslant 0$ we have $X \geqslant 0$. Let

$$Y_n := \mathbb{E}[X_n \mid G], \ \ Y := \mathbb{E}[X \mid \mathscr{G}].$$

By the monotonicity of the conditional expectation, $Y_n = \mathbb{E}[X_n \mid \mathscr{G}] \leqslant \mathbb{E}[X_{n+1} \mid \mathscr{F}] = Y_{n+1}$ so, in particular,

$$\exists \lim_n Y_n =: Z.$$

Let's check that $Z = \mathbb{E}[X \mid \mathscr{G}]$. First, since $Y_n \in \mathscr{G}$, also $Z \in \mathscr{G}$. Moreover, for $G \in \mathscr{G}$, we have

$$\mathbb{E}[X1_G] \overset{monot.conv.}{=} \lim_n \mathbb{E}[X_n 1_G] = \lim_n \mathbb{E}[Y_n 1_G] \overset{monot.conv.}{=} \mathbb{E}[Z1_G],$$

so $Z = \mathbb{E}[X \mid G]$ and from this it follows that

$$\lim_n \mathbb{E}[X_n \mid \mathscr{G}] = \mathbb{E}[X \mid \mathscr{G}],\ \mathbb{P}-a.s. \quad \square$$

7.3.8. We notice that

$$|\mathbb{E}[X_n \mid \mathcal{G}] - \mathbb{E}[X \mid \mathcal{G}]| = |\mathbb{E}[X_n - X \mid \mathcal{G}]| \leqslant \mathbb{E}\left[|X_n - X| \mid \mathcal{G}\right].$$

Let $Z_n := \sup_{k\geqslant n} |X_n - X|$. Clearly

$$|X_n - X| \leqslant Z_n, \implies \mathbb{E}\left[|X_n - X| \mid \mathcal{G}\right] \leqslant \mathbb{E}\left[Z_n \mid \mathcal{G}\right]$$

and $Z_n \downarrow 0$, so $0 \leqslant Z_1 - Z_n \uparrow Z_1 - \lim_n Z_n = Z_1$ because, by i), $Z_n \longrightarrow 0$. Therefore, by 7.3.7.,

$$\lim_n \mathbb{E}[Z_1 - Z_n \mid \mathcal{G}] = \mathbb{E}[Z_1 \mid \mathcal{G}], \ \mathbb{P} - a.s.$$

We also notice since $|X_n| \leqslant Y \in L^1$, also $|X| \leqslant Y$ and $Z_n \leqslant \sup_{k\geqslant n}(|X_n| + |X|) \leqslant 2Y \in L^1$. Therefore

$$\mathbb{E}[Z_1 - Z_n \mid \mathcal{G}] = \mathbb{E}[Z_1 \mid \mathcal{G}] - \mathbb{E}[Z_n \mid \mathcal{G}],$$

and from this

$$\lim_n \mathbb{E}[Z_n \mid \mathcal{G}] = 0, \ \mathbb{P} - a.s.$$

from which the conclusion follows. □

8.5.2. Let $\varepsilon > 0$. We have

$$\mathbb{P}(|X_n| \geqslant \varepsilon) = \mathbb{P}(X_n \geqslant \varepsilon) = \int_\varepsilon^{+\infty} ne^{-nx}\,dx = [-e^{-nx}]_{x=\varepsilon}^{x=+\infty} = e^{-n\varepsilon} \longrightarrow 0.$$

Is also $X_n \xrightarrow{L^1} 0$? We have

$$\begin{aligned}\mathbb{E}[|X_n - 0|] &= \mathbb{E}[|X_n|] = \int_{\mathbb{R}} |x|ne^{-nx}1_{[0,+\infty[}(x)\,dx = \int_0^{+\infty} xne^{-nx}\,dx \\ &= [-xe^{-nx}]_{x=0}^{x=+\infty} + \int_0^{+\infty} e^{-nx}\,dx = \frac{1}{n}[-e^{-nx}]_{x=0}^{x=+\infty} = \frac{1}{n} \longrightarrow 0.\end{aligned}$$

Is also $X_n \xrightarrow{a.s.} 0$? This happens iff

$$\mathbb{P}\left(\limsup_n \{|X_n| \geqslant \varepsilon\right) = 0.$$

We notice that

$$\mathbb{P}(|X_n| \geqslant \varepsilon) = e^{-n\varepsilon}, \implies \sum_n \mathbb{P}(|X_n| \geqslant \varepsilon) = \sum_n e^{-n\varepsilon} = \sum_n (e^{-\varepsilon})^n \overset{e^{-\varepsilon}<1}{=} \frac{1}{1 - e^{-\varepsilon}} < +\infty.$$

By the first Borel-Cantelli lemma it follows that

$$\mathbb{P}\left(\limsup_n \{|X_n| \geqslant \varepsilon\right) = 0,$$

so $X_n \xrightarrow{a.s.} 0$. □

8.5.3. We have $\mathbb{P}(X_n = 0) = 1 - \frac{1}{n}$, $\mathbb{P}(X_n = 1) = \frac{1}{n}$. We check $X_n \xrightarrow{L^1} 0$. Indeed

$$\mathbb{E}[|X_n - 0|] = \mathbb{E}[|X_n|] \overset{X_n\geqslant 0}{=} \mathbb{E}[X_n] = 0 \cdot \mathbb{P}(X_n = 0) + 1 \cdot \mathbb{P}(X_n = 1) = \frac{1}{n} \longrightarrow 0.$$

About a.s. limit, if $X_n \xrightarrow{a.s.} X$ then, by L^1 convergence necessarily $X = 0$. This happens iff

$$\forall \varepsilon > 0,\ \mathbb{P}\left(\limsup\{|X_n| \geqslant \varepsilon\}\right) = 0.$$

Now, since $|X_n| \geqslant \varepsilon$ iff $X_n \geqslant \varepsilon$ and, for $0 < \varepsilon \leqslant 1$,

$$\mathbb{P}(X_n \geqslant \varepsilon) = \mathbb{P}(X_n = 1) = \frac{1}{n},$$

being the X_n independent the events $E_n = \{X_n \geqslant \varepsilon\}$ are independent. Moreover, since

$$\sum_n \mathbb{P}(E_n) = \sum_n \frac{1}{n} = +\infty,$$

the second Borel-Cantelli lemma applies: we conclude that

$$\mathbb{P}\left(\limsup\{|X_n| \geqslant \varepsilon\}\right) = 1,$$

thus (X_n) is almost never convergent.

Since $X_n \xrightarrow{L^1} 0$, this implies that $X_n \xrightarrow{\mathbb{P}} 0$ and $X_n \xrightarrow{d} 0$. □

8.5.4. Since $\sum_n \varepsilon_n < +\infty$, we notice that if

$$(\star)\ \ \exists N,\ :\ |X_n| \leqslant \varepsilon_n,\ \forall n \geqslant N,$$

then

$$\sum_n |X_n| \leqslant \underbrace{\sum_{n=0}^{N-1} |X_n|}_{finite\ sum} + \sum_{n=N}^{\infty} \varepsilon_n < +\infty,$$

that is $\sum_n |X_n|$ converges, hence also $\sum_n X_n$ converges. Therefore $(\star)$ is a sufficient condition for convergence. We notice that

$$p := \mathbb{P}\left(\exists N,\ :\ |X_n| \leqslant \varepsilon_n,\ \forall n \geqslant N\right) = \mathbb{P}\left(\bigcup_N \bigcap_{n \geqslant N} \{|X_n| \leqslant \varepsilon_n\}\right),$$

so the conclusion holds if $p = 1$ or, equivalently,

$$0 = \mathbb{P}\left(\bigcap_N \bigcup_{n \geqslant N} \{|X_n| > \varepsilon_n\}\right) = \mathbb{P}\left(\limsup_n\{|X_n| > \varepsilon_n\}\right).$$

According to first Borel-Cantelli Lemma, this happens if $\sum_n \mathbb{P}(\{|X_n| > \varepsilon_n)$ is finite, this being true because, by assumption, $\mathbb{P}(|X_n| > \varepsilon_n) \leqslant \mathbb{P}(|X_n| \geqslant \varepsilon_n) \leqslant \varepsilon_n$ and $\sum_n \varepsilon_n < +\infty$. □

8.5.5. By assumption

$$\mathbb{P}(|X_n - X| \geqslant \varepsilon) \longrightarrow 0,\ \ \mathbb{P}(|Y_n - Y| \geqslant \varepsilon) \longrightarrow 0,\ \forall \varepsilon > 0.$$

Now, since

$$|(X_n + Y_n) - (X + Y)| = |(X_n - X) + (Y_n - Y)| \leqslant |X_n - X| + |Y_n - Y|,$$

it is clear that

$$|(X_n + Y_n) - (X + Y)| \geqslant \varepsilon,\ \implies\ |X_n - X| \geqslant \frac{\varepsilon}{2}\ \vee\ |Y_n - Y| \geqslant \frac{\varepsilon}{2},$$

otherwise, if $|X_n - X|, |Y_n - Y| < \frac{\varepsilon}{2}$ we would have $|(X_n + Y_n) - (X + Y)| < \varepsilon$. This means that

$$\begin{aligned}\mathbb{P}\left(|(X_n + Y_n) - (X + Y)| \geqslant \varepsilon\right) &\leqslant \mathbb{P}\left(\{|X_n - X| \geqslant \frac{\varepsilon}{2}\} \cup \{|Y_n - Y| \geqslant \frac{\varepsilon}{2}\}\right) \\ &\leqslant \mathbb{P}\left(\{|X_n - X| \geqslant \frac{\varepsilon}{2}\}\right) + \mathbb{P}\left(\{|Y_n - Y| \geqslant \frac{\varepsilon}{2}\}\right) \longrightarrow 0.\end{aligned}$$

8.5.6. i) Let ϕ_{X_n} be the characteristic function of X_n. Since $f_{X_n} \in L^1(\mathbb{R})$ we have

$$\phi_{X_n}(\xi) = \mathbb{E}[e^{i\xi X_n}] = \int_{\mathbb{R}} f_{X_n}(x) e^{i\xi x}\, dx = \widehat{f_{X_n}}(-\xi).$$

Now, we remind that

$$\widehat{\frac{1}{a^2 + \sharp^2}}(\xi) = \frac{\pi}{a} e^{-a|\xi|},$$

so

$$\widehat{f_{X_n}}(\xi) = \frac{n}{n^2\pi} \widehat{\frac{1}{\left(\frac{1}{n}\right)^2 + \sharp^2}}(\xi) = \frac{1}{n\pi} \frac{\pi}{1/n} e^{-\frac{|\xi|}{n}} = e^{-\frac{|\xi|}{n}}.$$

From this,

$$\phi_{X_n}(\xi) = e^{-\frac{|\xi|}{n}} \longrightarrow e^0 = 1\phi_0(\xi), \implies X_n \stackrel{d}{\longrightarrow} 0.$$

ii) We have

$$\begin{aligned}\mathbb{P}\left(|X_n - 0| \geqslant \varepsilon\right) &= \mathbb{P}\left(|X_n| \geqslant \varepsilon\right) = \int_{|x|\geqslant\varepsilon} f_{X_n}(x)\, dx = \frac{1}{\pi} \int_{|x|\geqslant\varepsilon} \frac{n}{1 + n^2x^2}\, dx = \frac{2}{\pi} \int_{\varepsilon}^{+\infty} \frac{n}{1 + (nx)^2}\, dx \\ &= \frac{2}{\pi} \left[\arctan(nx)\right]_{x=\varepsilon}^{x=+\infty} = \frac{2}{\pi}\left(\frac{\pi}{2} - \arctan(n\varepsilon)\right) \longrightarrow \frac{2}{\pi}\left(\frac{\pi}{2} - \frac{\pi}{2}\right) = 0, \quad n \to +\infty.\end{aligned}$$

We conclude that $X_n \stackrel{\mathbb{P}}{\longrightarrow} 0$.

iii) Since a.s. convergence implies convergence in probability, if $X_n \stackrel{a.s.}{\longrightarrow} X$, then, necessarily, $X = 0$, that is $X_n \stackrel{a.s.}{\longrightarrow} 0$. Now, this happens iff

$$\mathbb{P}\left(\limsup_n \{|X_n| \geqslant \varepsilon\}\right) = 0, \ \forall \varepsilon > 0.$$

Notice that

$$\mathbb{P}(|X_n| \geqslant \varepsilon) = \frac{2}{\pi}\left(\frac{\pi}{2} - \arctan(n\varepsilon)\right) \longrightarrow 0.$$

Now, the question is: is this enough to make $\sum_n \mathbb{P}(|X_n| \geqslant \varepsilon)$ convergent? If yes, the conclusion would follow by the first Borel-Cantelli Lemma. To discuss this, we need to establish the asymptotic behavior of $\arctan(n\varepsilon)$ when $n \sim +\infty$, for $\varepsilon > 0$ fixed. To this aim we recall the remarkable identity

$$\arctan x + \arctan \frac{1}{x} = \frac{\pi}{2}, \ \forall x > 0,$$

so

$$\frac{\pi}{2} - \arctan(n\varepsilon) = \arctan \frac{1}{n\varepsilon} \sim \frac{1}{n\varepsilon},$$

being $\arctan y \sim y$ when $y \to 0$. Therefore,

$$\sum_n \mathbb{P}(|X_n| \geqslant \varepsilon) \sim \sum_n \frac{1}{n\varepsilon} = +\infty.$$

Since the (X_n) are independent, and $\sum_n \mathbb{P}(|X_n| \geqslant \varepsilon) = +\infty$, we can apply the second Borel-Cantelli Lemma and conclude that

$$\mathbb{P}\left(\limsup\{|X_n| \geqslant \varepsilon\}\right) = 1,$$

that is (X_n) is almost never convergent. □

8.5.7. Let $U_n \sim U([0,1])$ be i.i.d. and let $X_n := \min(U_1, \dots, U_n)$. Then

$$F_{X_n}(x) = \mathbb{P}(\min(U_1, \dots, U_n) \leqslant x) = \begin{cases} 0, & x < 0, \\ \\ 1, & x \geqslant 1. \end{cases}$$

For $0 \leqslant x \leqslant 1$ we have

$$\begin{aligned} F_{X_n}(x) &= \mathbb{P}(\min(U_1, \dots, U_n) \leqslant x) \\ &= \mathbb{P}\left(\{U_1 \leqslant x\} \sqcup \{U_1 > x, U_2 \leqslant x\} \sqcup \dots \{U_1, U_2, \dots, U_{n-1} > x, U_n \leqslant x\}\right) \\ &= \mathbb{P}(U_1 \leqslant x) + \mathbb{P}(U_1 > x,\ U_2 \leqslant x) + \cdots \mathbb{P}(U_1 > x, \dots, U_{n-1} > x, U_n \leqslant x). \end{aligned}$$

Now, by the independence,

$$\mathbb{P}(U_1 > x, \dots, U_{k-1} > x, U_k \leqslant x) = \mathbb{P}(U_1 > x) \cdots \mathbb{P}(U_{k-1} > x)\mathbb{P}(U_k \leqslant x) = x(1-x)^{k-1}.$$

Therefore,

$$F_{X_n}(x) = x + x(1-x) + x(1-x)^2 + \cdots + x(1-x)^{n-1} = x\sum_{k=0}^{n-1}(1-x)^k = x\frac{1-(1-x)^n}{1-(1-x)} = 1-(1-x)^n.$$

We conclude that

$$F_{X_n}(x) = \begin{cases} 0, & x < 0, \\ 1-(1-x)^n, & 0 \leqslant x \leqslant 1, \\ 1, & x > 1. \end{cases}$$

ii) We notice that

$$F_{nX_n}(x) = \mathbb{P}(nX_n \leqslant x) = \mathbb{P}(X_n \leqslant \frac{x}{n}) = F_{X_n}\left(\frac{x}{n}\right) = \begin{cases} 0, & x < 0, \\ 1-\left(1-\frac{x}{n}\right)^n, & 0 \leqslant x \leqslant n, \\ 1, & x > n. \end{cases}$$

From this we notice that

$$F_{nX_n}(x) \longrightarrow \begin{cases} 0, & x < 0, \\ 1-e^{-x}, & x \geqslant 0, \end{cases} =: F_Y(x),$$

where $Y \sim \exp(1)$, and $nX_n \xrightarrow{d} Y$.

Let's discuss convergence in probability: since this is stronger than convergence in distribution, the unique possibility is $nX_n \xrightarrow{\mathbb{P}} Y$, that is

$$\mathbb{P}(|nX_n - Y| \geqslant \varepsilon) \longrightarrow 0.$$

Let $\alpha > 0$ be fixed and let $n \geqslant \alpha$ (that is $n \geqslant [\alpha] + 1$). We notice that

$$\begin{aligned}\{|nX_n - Y| \geqslant \varepsilon\} &= \{Y \leqslant \alpha,\ nX_n \leqslant Y - \varepsilon \ \vee\ nX_n \geqslant Y + \varepsilon\} \sqcup \{Y > \alpha,\ nX_n \leqslant Y - \varepsilon \ \vee\ nX_n \geqslant Y + \varepsilon\} \\ &\supset \{Y \leqslant \alpha,\ nX_n \geqslant \alpha + \varepsilon\} \sqcup \{Y > \alpha,\ nX_n \leqslant \alpha - \varepsilon\}\end{aligned}$$

so

$$\mathbb{P}(|nX_n - Y| \geqslant \varepsilon) \geqslant \mathbb{P}(Y \leqslant \alpha,\ nX_n \geqslant \alpha + \varepsilon) + \mathbb{P}(Y > \alpha,\ nX_n \leqslant \alpha - \varepsilon).$$

Now,

$$\mathbb{P}(Y \leqslant \alpha,\, nX_n \geqslant \alpha+\varepsilon) = \mathbb{P}(\{Y \leqslant \alpha\}\setminus\{nX_n \leqslant \alpha+\varepsilon\}) \geqslant \mathbb{P}(Y \leqslant \alpha) - \mathbb{P}(nX_n \leqslant \alpha+\varepsilon) = 1 - e^{-\alpha} - \left(1 - \frac{\alpha + \varepsilon}{n}\right)^n,$$

and, similarly,

$$\mathbb{P}(Y > \alpha,\ nX_n \leqslant \alpha - \varepsilon) \geqslant \mathbb{P}(nX_n \leqslant \alpha - \varepsilon) - \mathbb{P}(Y < \alpha) = \left(1 - \frac{\alpha - \varepsilon}{n}\right)^n - (1 - e^{-\alpha}).$$

Therefore

$$\mathbb{P}(|nX_n - Y| \geqslant \varepsilon) \geqslant \left(1 - \frac{\alpha - \varepsilon}{n}\right)^n - \left(1 - \frac{\alpha + \varepsilon}{n}\right)^n \longrightarrow e^{-(\alpha-\varepsilon)} - e^{-(\alpha+\varepsilon)} = 2e^{-\alpha}\cosh\varepsilon > 0.$$

From this it follows that $\lim_n \mathbb{P}(|nX_n - Y| \geqslant \varepsilon)$ cannot be $= 0$, so (X_n) cannot converge in probability to Y.

8.5.8. i) We have

$$F_{M_n}(x) = \mathbb{P}(\max(X_1, \dots, X_n) \leqslant x) = \mathbb{P}(\{X_1 \leqslant x\} \cap \cdots \{X_n \leqslant x\}) = \prod_{k=1}^{n} \mathbb{P}(X_k \leqslant x).$$

Since we know that

$$\mathbb{P}(X_k > x) = \frac{1}{\sqrt{x}},\ x \geqslant 1,\quad F_{X_k}(x) = \mathbb{P}(X_k \leqslant x) = 1 - \mathbb{P}(X_k > x) = 1 - \frac{1}{\sqrt{x}},\ x \geqslant 1,$$

in particular $F_{X_k}(1) = 0$ and since $F_{X_k} \nearrow$, this means that $F_{X_k}(x) = 0$ for every $x \leqslant 1$. Therefore, $F_{M_n}(x) = 0$ for $x \leqslant 1$, while

$$F_{M_n}(x) = \prod_{k=1}^{n}\left(1 - \frac{1}{\sqrt{x}}\right) = \left(1 - \frac{1}{\sqrt{x}}\right)^n,\ x \geqslant 1.$$

ii) Clearly, $\lim_n F_{M_n}(x) = 0$ for every $x \leqslant 1$. For $x > 1$ fixed, since $1 - \frac{1}{\sqrt{x}} < 1$,

$$\lim_{n\to\infty} F_{M_n}(x) = \lim_{n\to\infty}\left(1 - \frac{1}{\sqrt{x}}\right)^n = 0.$$

Therefore,

$$\lim_{n\to+\infty} F_{M_n}(x) = 0,\ \forall x \in \mathbb{R}.$$

From this it follows that (M_n) cannot be convergent in distribution to any M. Indeed, if this would happen, we would have $F_{M_n}(x) \longrightarrow F_M(x)$ at every x continuity point of M. Since the number of discontinuity points of M is at most countable (see exercise 3.4.9), it means that $F_M(x) = 0$ apart, at most, for a countable set of x. This implies $F_M \equiv 0$. Indeed, if $F_M(x_0) > 0$ for some x_0, then by monotonicity, $F_M(x) \geqslant F_M(x_0) > 0$ for every $x > x_0$. Therefore F_M shuld be discontinuous at every $x > x_0$ (otherwise, $F_M(x) = 0$), thus the set of discontinuity points of F_M would contain $[x_0, +\infty[$, that is, it would be uncountable.

9.4.1. We have

$$\phi_{\overline{X_n}}(\xi) = \mathbb{E}\left[e^{i\xi\overline{X_n}}\right] = \mathbb{E}\left[e^{i\frac{\xi}{n}\sum_{k=1}^n X_k}\right] = \mathbb{E}\left[\prod_k e^{i\frac{\xi}{n}X_k}\right] \stackrel{indep}{=} \prod_k \mathbb{E}[e^{i\frac{\xi}{n}X_k}] = \prod_k \phi_{X_k}\left(\frac{\xi}{n}\right).$$

Since $X_k \in L^1(\Omega)$,

$$\phi_{X_k}(\eta) = \phi_{X_k}(0) + \partial_\eta \psi_{X_k}(0)\eta + o(\eta) = 1 + i\eta\mathbb{E}[X_k] + o(\eta) = 1 + i\eta m + o(\eta),$$

so

$$\phi_{X_k}\left(\frac{\xi}{n}\right) = 1 + i\frac{\xi}{n}m + o\left(\frac{\xi}{n}\right).$$

Therefore

$$\phi_{\overline{X_n}}(\xi) = \prod_k \left(1 + i\frac{\xi}{n}m + o\left(\frac{\xi}{n}\right)\right) = \left(1 + i\frac{\xi}{n}m + o\left(\frac{\xi}{n}\right)\right)^n \longrightarrow e^{i\xi m},\ \forall \xi \in \mathbb{R}.$$

The conclusion now follows by the continuity theorem. □

9.4.2. We notice that

$$\frac{1}{\sqrt{n}}|(X_1, \dots, X_n)| = \sqrt{\frac{1}{n}\sum_{k=1}^n X_k^2}$$

so, the conclusion is equivalent to

$$\mathbb{P}\left(\frac{1-\varepsilon}{\sqrt{3}} \leqslant \sqrt{\frac{1}{n}\sum_{k=1}^n X_k^2} \leqslant \frac{1+\varepsilon}{\sqrt{3}}\right) = \mathbb{P}\left(\frac{(1-\varepsilon)^2}{3} \leqslant \frac{1}{n}\sum_{k=1}^n X_k^2 \leqslant \frac{(1+\varepsilon)^2}{3}\right) \longrightarrow 1.$$

This is equivalent to

$$\frac{1}{n}\sum_{k=1}^n X_k^2 \xrightarrow{\mathbb{P}} \frac{1}{3}.$$

Since the X_k are i.i.d. (with $f_{X_k}(x) = \frac{1}{2}1_{[-1,1]}(x)$), the same holds for the X_k^2. Moreover,

$$\mathbb{E}[X_k^2] = \int_{-1}^1 x^2\frac{dx}{2} = \int_0^1 x^2\,dx = \left[\frac{x^3}{3}\right]_{x=0}^{x=1} = \frac{1}{3},$$

and,

$$\mathbb{V}[X_k^2] = \mathbb{E}[X_k^4] - \mathbb{E}[X_k]^4 = \int_{-1}^1 x^4\,\frac{dx}{2} = \int_0^1 x^4\,dx = \left[\frac{x^5}{5}\right]_{x=0}^{x=1} = \frac{1}{5}.$$

Therefore, the Chebishev theorem applies and the conclusion follows. □

9.4.3. i) We notice that

$$\mathbb{E}[X_k] = k \times \frac{1}{2k\log k} - k \times \frac{1}{2k\log k} + 0 \times \left(1 - \frac{1}{2k\log k}\right) = 0,$$

and

$$\mathbb{V}[X_k] = \mathbb{E}[X_k^2] - \mathbb{E}[X_k]^2 = 2k^2 \times \frac{1}{2k\log k} = \frac{k}{\log k}.$$

By the Chebishev bound,

$$\mathbb{P}\left(|\overline{X_n}| \geqslant \varepsilon\right) \leqslant \frac{1}{\varepsilon^2 n^2}\mathbb{V}\left[\sum_{k=1}^{n} X_k\right].$$

By the independence,

$$\mathbb{V}\left[\sum_{k=1}^{n} X_k\right] = \sum_{k=1}^{n}\mathbb{V}\left[X_k\right] = \sum_{k=2}^{n}\frac{k}{\log k},$$

so

$$\lim_{n\to+\infty}\mathbb{P}\left(|\overline{X_n}| \geqslant \varepsilon\right) \leqslant \frac{1}{\varepsilon^2}\lim_{n}\frac{1}{n^2}\sum_{k=2}^{n}\frac{k}{\log k}.$$

We notice that, if $f(k) := \frac{k}{\log k}$, $f'(k) = \frac{\log k - 1}{(\log k)^2} \geqslant 0$ for $k \geqslant 1$. In particular, $\frac{k}{\log k}\uparrow$, so

$$\frac{1}{n^2}\sum_{k=1}^{n}\frac{k}{\log k} \leqslant \frac{1}{n^2}\sum_{k=1}^{n}\frac{n}{\log n} = \frac{1}{n^2}\frac{n}{\log n}\cdot n = \frac{1}{\log n} \longrightarrow 0,$$

so

$$\lim_{n\to+\infty}\mathbb{P}\left(|\overline{X_n}| \geqslant \varepsilon\right) = 0,$$

this ensuring that $\overline{X_n} \xrightarrow{\mathbb{P}} 0$.

ii) If $\overline{X_n}$ converges $\mathbb{P}$–a.s., necessarily $\overline{X_n} \xrightarrow{a.s.} 0$. We notice that

$$\overline{X_n} = \frac{1}{n}\sum_{k=1}^{n} X_k = \frac{1}{n}\sum_{k=1}^{n-1} X_k \pm 1 = \frac{n-1}{n}\overline{X_{n-1}} \pm 1 = \frac{n-1}{n}\overline{X_{n-1}} + \frac{X_n}{n} =: \frac{n-1}{n}\overline{X_{n-1}} + Y_n.$$

If $\overline{X_n} \xrightarrow{a.s.} 0$, then

$$Y_n = \overline{X_n} - \frac{n-1}{n}\overline{X_{n-1}} \longrightarrow 0,\ \mathbb{P} - a.s.$$

This happens iff

$$\mathbb{P}(\limsup_n |Y_n| \geqslant \varepsilon) = 0.$$

However, since the Y_n are independent, and

$$\mathbb{P}(|Y_n| \geqslant \varepsilon) = \mathbb{P}(X_n = \pm n) = \frac{1}{n\log n}, \implies \sum_n \mathbb{P}(|Y_n| \geqslant \varepsilon) = \sum_n \frac{1}{n\log n} = +\infty,$$

by the second Borel-Cantelli lemma we obtain

$$\mathbb{P}(\limsup_n |Y_n| \geqslant \varepsilon) = 1.$$

We get a contradiction, The conclusion is that $\overline{X_n}$ cannot be convergent with probability 1. □

9.4.4. We notice that, setting $Y_k := X_k X_{k+1}$, $Y_k \in L^1(\Omega)$ (indeed, $\mathbb{E}[|Y_k|] = \mathbb{E}[|X_k||X_{k+1}|] = \mathbb{E}[|X_k|]\mathbb{E}[|X_{k+1}|] < +\infty$), and moreover

$$\mathbb{E}[Y_k] = \mathbb{E}[X_k X_{k+1}] = \mathbb{E}[X_k]\mathbb{E}[X_{k+1}] = m^2.$$

We also notice that, if μ is the common law of the X_k, then

$$F_{Y_k}(y) = \mathbb{P}(X_k X_{k+1} \leqslant u) = \int_{xy \leqslant u} \mu(dx)\mu(dy)$$

is independent of k, that is the Y_k are identicaslly distributed. However, Y_k is not independent of Y_{k+1} because both variables depend on X_{k+1}. However, $Y_1, Y_3, \dots, Y_{2k+1}, \dots$ are independent and, similarly, $Y_2, Y_4, \dots, Y_{2k}, \dots$ areindependent too. The L^1 SLLN applies, so if $n := 2N + 1$ we have

$$\begin{aligned} \overline{Y_n} &= \frac{1}{2N+1} \sum_{k=1}^{2N+1} Y_k = \frac{1}{2N+1} \left(\sum_{j=0}^{N} Y_{2j+1} + \sum_{j=1}^{N} Y_{2j} \right) \\ &= \frac{N-1}{2N+1} \frac{1}{N+1} \sum_{j=0}^{N} Y_{2j+1} + \frac{N}{2N+1} \frac{1}{N} \sum_{j=1}^{N} Y_{2j} \xrightarrow{a.s.} \frac{1}{2} m^2 + \frac{1}{2} m^2 = m^2, \end{aligned}$$

and, similarly

$$\overline{Y_{2N}} \xrightarrow{a.s.} m^2.$$

We conclude that $\overline{Y_n} \xrightarrow{a.s.} m^2$. □

9.4.5. As suggested, let X_k be independent random variables uniformly distributed on $[0, 1]$. In this way

$$\mathbb{E}[\varphi(X_1, \dots, X_n)] = \int_{\mathbb{R}^n} \varphi(x_1, \dots, x_n) \, d\mu_{X_1,\dots,X_n}(x_1, \dots, x_n) = \int_{[0,1]^n} \varphi(x_1, \dots, x_n) \, dx_1 \cdots dx_n.$$

Therefore,

$$\int_{[0,1]^n} \frac{x_1^2 + \cdots + x_n^2}{x_1 + \cdots + x_n} \, dx_1 \cdots dx_n = \mathbb{E}\left[\frac{X_1^2 + \cdots + X_n^2}{X_1 + \cdots + X_n} \right].$$

Now, since $(X_k), (X_k^2) \subset L^1(\Omega)$ are i.i.d., the L^1−SLLN applies and

$$\frac{1}{n} \sum_{k=1}^{n} X_k \xrightarrow{a.s.} \mathbb{E}[X_1] = \frac{1}{2},$$

while

$$\frac{1}{n} \sum_{k=1}^{n} X_k^2 \xrightarrow{a.s.} \mathbb{E}[X_1^2] = \frac{1}{3}.$$

Therefore,

$$\frac{X_1^2 + \cdots + X_n^2}{X_1 + \cdots + X_n} = \frac{\frac{1}{n} \sum_{k=1}^{n} X_k^2}{\frac{1}{n} \sum_{k=1}^{n} X_k} \xrightarrow{a.s.} \frac{1/3}{1/2} = \frac{2}{3}.$$

Therefore,

$$\lim_n \int_{[0,1]^n} \frac{x_1^2+\cdots+x_n^2}{x_1+\cdots+x_n}\,dx_1\cdots dx_n = \lim_n \mathbb{E}\left[\frac{X_1^2+\cdots+X_n^2}{X_1+\cdots+X_n}\right] \overset{(*)}{=} \mathbb{E}\left[\lim_n \frac{X_1^2+\cdots+X_n^2}{X_1+\cdots+X_n}\right] = \frac{2}{3},$$

provided $(*)$ applies. To carry $\lim_n$ inside the $\mathbb{E}$, we invoke the Lebesgue dominated convergence theorem: we already know that

$$\lim_n \frac{X_1^2+\cdots+X_n^2}{X_1+\cdots+X_n} = \frac{2}{3},\ \mathbb{P}-a.s..$$

We need to dominate (independently of n) the ratio. To this aim, we just notice that, since $X_k \in [0,1]$, then $0 \leqslant X_k^2 \leqslant X_k \leqslant 1$, so

$$X_1^2+\cdots+X_n^2 \leqslant X_1+\cdots+X_n,\ \Longrightarrow\ \frac{X_1^2+\cdots+X_n^2}{X_1+\cdots+X_n} \leqslant 1 := Z \in L^1(\Omega),\ \mathbb{P}-a.s.$$

The conclusion now follows. □

9.4.6. i) No: indeed,

$$\mathbb{E}[|X_k|]\int_{\mathbb{R}} |x| f_{X_k}(x)\,dx = \int_{\mathbb{R}} |x|\frac{1}{\pi}\frac{a}{a^2+x^2}\,dx = +\infty.$$

ii) Let $\overline{X_n} = \frac{1}{n}\sum_{k=1}^n X_k$. Then

$$\phi_{\overline{X_n}}(\xi) = \mathbb{E}\left[e^{i\frac{\xi}{n}\sum_{k=1}^n X_k}\right] = \mathbb{E}\left[\prod_{k=1}^n e^{i\frac{\xi}{n}X_k}\right] \overset{indep}{=} \prod_{k=1}^n \mathbb{E}\left[e^{i\frac{\xi}{n}X_k}\right] = \prod_{k=1}^n \phi_{X_k}\left(\frac{\xi}{n}\right).$$

Notice that

$$\phi_{X_k}(\xi) = \widehat{\frac{1}{\pi}\frac{a}{a^2+\sharp^2}}(-\xi) = \frac{a}{\pi}\frac{\pi}{a}e^{-a|-\xi|} = e^{-a|\xi|},$$

so

$$\phi_{\overline{X_n}} = \prod_{k=1}^n e^{-a|\frac{\xi}{n}} = e^{-a|\xi|}.$$

iii) Since $\phi_{\overline{X_n}}(\xi) \equiv e^{-a|\xi|} \longrightarrow e^{-a|\xi|}$, $\forall \xi \in \mathbb{R}$, by the continuity theorem $\overline{X_n} \xrightarrow{d} X$ where X has the same distribution of X_k.

Let's analyze the convergence in probability. Since this is stronger than convergence in distribution, if true, necessarily $\overline{X_n} \xrightarrow{\mathbb{P}} X$ where X is still a Cauchy distribution of same type of the X_k. We notice that

$$\{|\overline{X_n}-X| \geqslant \varepsilon\} \subset \left\{|\overline{X_n}| \geqslant \frac{\varepsilon}{2}\right\} \cup \left\{|X| \geqslant \frac{\varepsilon}{2}\right\}$$

9.4.7. We notice that

$$\left(\frac{e^{-n}}{X_1\cdot X_2\cdots X_n}\right)^{1/\sqrt{n}} = e^{\frac{1}{\sqrt{n}}\left(-n-\sum_{k=1}^n \log X_k\right)} = e^{-\frac{1}{\sqrt{n}}\sum_{k=1}^n(\log X_k+1)}$$

so, if $0 < a < b$,

$$a \leqslant \left(\frac{e^{-n}}{X_1 \cdots X_n}\right)^{1/\sqrt{n}} \leqslant b, \quad \Longleftrightarrow \quad -\log b \leqslant \frac{1}{\sqrt{n}} \sum_{k=1}^{n} (\log X_k + 1) \leqslant -\log a$$

Now, if $Y_k := \log X_k + 1$ we have that

$$\mathbb{E}[Y_k] = 1 + \mathbb{E}[\log X_k] = 1 + \int_0^1 \log x \, dx = 1 + [x \log x]_{x=0}^{x=1} - \int_0^1 1 \, dx = 0,$$

and

$$\mathbb{V}[Y_k] = \mathbb{E}[Y_k^2] - \mathbb{E}[Y_k]^2 = \mathbb{E}\left[(1 + \log X_k)^2\right] = 1 + 2 \underbrace{\mathbb{E}[\log X_k]}_{=-1} + \mathbb{E}[\log^2 X_k] = \mathbb{E}[\log^2 X_k] - 1.$$

We have

$$\mathbb{E}[\log^2 X_k] = \int_0^1 \log^2 x \, dx = \left[x \log^2 x\right]_{x=0}^{x=1} - \int_0^1 x 2 \log x \cdot \frac{1}{x} \, dx = 0 - 2 \int_0^1 \log x \, dx = 2,$$

from which $\mathbb{V}[Y_k] = 2 - 1 = 1$. From the CLT

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} Y_k \xrightarrow{d} \mathcal{N}(0, 1),$$

and since the limit distribution is absolutely continuous, in particular

$$\mathbb{P}\left(a \leqslant \left(\frac{e^{-n}}{X_1 \cdots X_n}\right)^{1/\sqrt{n}} \leqslant b\right) = \mathbb{P}\left(-\log b \leqslant \frac{1}{\sqrt{n}} \sum_{k=1}^{n} Y_k \leqslant -\log a\right) \longrightarrow \int_{-\log b}^{-\log a} e^{-\frac{x^2}{2}} \frac{dx}{\sqrt{2\pi}}. \quad \square$$

9.4.8. We notice that

$$Y_n := \frac{\sum_{k=1}^{n} X_k}{\sqrt{\sum_{k=1}^{n} X_k^2}} = \frac{1}{\sqrt{\frac{1}{n} \sum_{k=1}^{n} X_k^2}} \frac{1}{\sqrt{n}} \sum_{k=1}^{n} X_k.$$

By the assumptions, $(X_k^2) \subset L^1(\Omega)$ are i.i.d., hence, by the L^1-SLLN we have

$$\frac{1}{n} \sum_{k=1}^{n} X_k^2 \xrightarrow{a.s.} \mathbb{E}[X_1^2] =: \sigma^2.$$

Therefore,

$$Z_n := \frac{\sigma}{\sqrt{\frac{1}{n} \sum_{k=1}^{n} X_k^2}} \xrightarrow{a.s.} 1,$$

and

$$Y_n = Z_n \frac{1}{\sigma \sqrt{n}} \sum_{k=1}^{n} X_k.$$

Now, by the CLT, being $\mathbb{E}[X_k] = 0$, we have

$$\frac{1}{\sigma \sqrt{n}} \sum_{k=1}^{n} X_k \xrightarrow{d} \mathcal{N}(0, 1),$$

We claim that also

$$Y_n \xrightarrow{d} \mathcal{N}(0,1).$$

We need a general fact:

$$Z_n \xrightarrow{a.s.} 1, \;\; Y_n \xrightarrow{d} Y, \;\; \Longrightarrow \;\; Z_n Y_n \xrightarrow{d} Y.$$

Indeed, we have

$$\phi_{Z_n Y_n}(\xi) = \mathbb{E}[e^{i\xi Z_n Y_n}] = \mathbb{E}[e^{i\xi Y_n} e^{i\xi(Z_n-1)Y_n}] = \mathbb{E}[e^{i\xi Y_n}] + \mathbb{E}\left[e^{i\xi Y_n}\left(e^{i\xi(Z_n-1)Y_n} - 1\right)\right]$$

and since, by assumption, $\mathbb{E}[e^{i\xi Y_n}] = \phi_{Y_n}(\xi) \longrightarrow \phi_Y(\xi)$, we need to prove that the last term goes to 0. We have

$$\begin{aligned}\left|\mathbb{E}\left[e^{i\xi Y_n}\left(e^{i\xi(Z_n-1)Y_n} - 1\right)\right]\right| &\leqslant \mathbb{E}\left[\left|e^{i\xi(Z_n-1)Y_n} - 1\right|\right] \\ &= \mathbb{E}\left[\left|e^{i\xi(Z_n-1)Y_n} - 1\right| 1_{|Y_n|<K}\right] + \mathbb{E}\left[\left|e^{i\xi(Z_n-1)Y_n} - 1\right| 1_{|Y_n|\geqslant K}\right]\end{aligned}$$

Here K is fixed independently of n (we will see son how). The first expectation goes to 0 because of the dominated convergence: since $Z_n \xrightarrow{a.s.} 1$, and $|Y_n| < K$, $(Z_n - 1)Y_n \longrightarrow 0$, and everything is controlled by $1 \in L^1(\Omega)$. About the second expectation we have

$$\mathbb{E}\left[\left|e^{i\xi(Z_n-1)Y_n} - 1\right| 1_{|Y_n|\geqslant K}\right] \leqslant 2\mathbb{P}(|Y_n| \geqslant K),$$

and since $\mathbb{P}(|Y_n| \geqslant K) \longrightarrow \mathbb{P}(|Y| \geqslant K)$, we can say that $\mathbb{P}(|Y_n| \geqslant K) \leqslant \mathbb{P}(|Y| \geqslant K) + \varepsilon \leqslant 2\varepsilon$ for n large, and K large enough. Therefore

$$\lim_n \left|\mathbb{E}\left[e^{i\xi Y_n}\left(e^{i\xi(Z_n-1)Y_n} - 1\right)\right]\right| \leqslant 2\varepsilon,$$

and since $\varepsilon > 0$ can be take arbitrarily, we have the conclusion.