Seminar 6

Exercise 1

There are 7 red and 5 white balls in a box. Two balls are randomly drawn from the box.

a. Find the expectation and variance of the number of red balls drawn.
b. Would the answer change if the balls were drawn as follows: the first ball is drawn, put back, and then the second ball is drawn?

Solution

Let N be the number of red balls drawn. Let $X_i = 1$ if the i-th ball drawn is red, and $X_i = 0$ otherwise. $N = X_1 + X_2$. In both cases we have $\mathbb{E}[X_1] = \mathbb{P}(X_1 = 1) = 7/12 = \mathbb{E}[X_2] = 7/12$ and thus $\mathbb{E}[N] = 7/6$. The computation of the variance however depends on the joint law of (X_1, X_2) and we need to separate the two cases.

a. **Without replacement:** In this case $\mathbb{P}(N = 2) = \mathbb{P}(X_1 = 1, X_2 = 1) = \binom{7}{2}/\binom{12}{2}$; $\mathbb{P}(N = 0) = \binom{5}{2}/\binom{12}{2}$. Thus we can compute

$$\mathbb{E}[N^2] = 4\mathbb{P}(N = 2) + 1(1 - \mathbb{P}(N = 2) - \mathbb{P}(N = 0)) = 1 + 3\frac{\binom{7}{2}}{\binom{12}{2}} - \frac{\binom{5}{2}}{\binom{12}{2}}$$

$$\mathrm{Var}[N] = \mathbb{E}[N^2] - \mathbb{E}[N]^2 = \frac{175}{396} \approx 0.442$$

b. **With replacement:** Now X_1, X_2 are independent, $N \sim \text{Binomial}(2, p = 7/12)$. Thus $\mathrm{Var}(N) = 2p(1-p) = 35/72 \approx 0.486$. As intuitive, the variance increased.

Exercise 2

Let $X \sim \mathcal{N}(0, \sigma^2)$. Calculate $\mathbb{E}[X^n]$ for $n \in \mathbb{N}$.

Solution

- If n is odd, then $\mathbb{E}[X^n] = 0$ due to the symmetry of the normal distribution about zero.
- If $n = 2k$ is even, we can reduce to the case $\sigma = 1$ considering $Z = X/\sigma$, since then $\mathbb{E}[X^{2k}] = \sigma^{2k}\mathbb{E}[Z^{2k}]$. An integration by parts shows a crucial property of standard normal random variables

$$\mathbb{E}[Zf(Z)] = \mathbb{E}[f'(Z)]$$

Let $f(Z) = Z^{n-1}$. Then $\mathbb{E}[Z^n] = (n-1)\mathbb{E}[Z^{n-2}]$. Applying this formula repeatedly: $\mathbb{E}[Z^{2k}] = (2k-1)\mathbb{E}[Z^{2k-2}] = (2k-1)(2k-3)\cdots 1 \cdot \mathbb{E}[Z^0] = (2k-1)!!$. Then $\mathbb{E}[X^{2k}] = \sigma^{2k}(2k-1)!!$.

Exercise 3

The random variable X has a Poisson distribution with parameter λ_1, the random variable Y is exponentially distributed with parameter λ_2, and X and Y are independent. Find the expectation and variance of the random variables $X + Y$, XY.

> **Solution**
>
> $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y] = \lambda_1 + 1/\lambda_2$, and due to independence
>
> $$\begin{aligned} \mathrm{Var}(X+Y) &= \mathrm{Var}(X) + \mathrm{Var}(Y) = \lambda_1 + 1/\lambda_2^2 \\ \mathbb{E}[XY] &= \mathbb{E}[X]\mathbb{E}[Y] = \lambda_1/\lambda_2 \\ \mathrm{Var}(XY) &= \mathbb{E}[X^2]\mathbb{E}[Y^2] - (\mathbb{E}[XY])^2 = (\lambda_1 + \lambda_1^2)(2\lambda_2^{-2}) - (\lambda_1/\lambda_2)^2 \end{aligned}$$

Exercise 4

Let the joint probability density of the random variables ξ, η be $p_{\xi,\eta}(x,y) = C\exp(-x^2 + 2xy - 2y^2)$. Find the constant C and $\mathrm{Cov}(\xi, \eta)$.

> **Solution**
>
> Complete the square in the exponent: $-x^2 + 2xy - 2y^2 = -(x^2 - 2xy + y^2) - y^2 = -((x-y)^2 + y^2)$. Via the substitution $z = x - y$ in the integral, we then get $\int p_{\xi,\eta}(x,y)dx\,dy = C\pi$. So $C = 1/\pi$.
> From the same computation we see that $\zeta := \xi - \eta$ and η are independent, so $\mathrm{Cov}(\xi,\eta) = \mathrm{Cov}(\zeta,\eta) - \mathrm{Cov}(\eta,\eta) = 0 + 1/2$.

Exercise 5

Give an example of dependent random variables with zero covariance.

> **Solution**
>
> Let X be any symmetric (non-constant) random variable, meaning that X and $-X$ have the same distribution, e.g. $X \sim \mathcal{N}(0,1)$. Take $Y = X^{2n}$, then X and Y are dependent, $\mathbb{E}[X] = 0$ and $\mathbb{E}[XY] = \mathbb{E}[X^{2n+1}] = 0$.

Exercise 6

Calculate $\mathbb{E}[\xi^2]$

a. If

$$F_\xi(x) = \begin{cases} 0, & \text{for } x < -1, \\ 1/3, & \text{for } -1 \leq x < 0, \\ 1 - \frac{1}{2}e^{-x}, & \text{for } x \geq 0. \end{cases}$$

b. If $F_\xi(x) = \frac{1}{2} + \frac{1}{\pi}\arctan(x)$, for $x \geq 0$?

Solution

For $a \in \mathbb{R}$ we have that

$$\mathbb{E}[f(\xi)] = f(a) + \int_a^\infty f'(t)\mathbb{P}(X > t)dt - \int_{-\infty}^a f'(t)\mathbb{P}(X \leq t)dt \tag{1}$$

a. We can solve the problem in two different ways.

- Using Equation 1, with $a = -1$

$$\mathbb{E}\left[\xi^2\right] = (-1)^2 + \int_{-1}^0 2t(1-1/3)dt + \int_0^\infty 2t\tfrac{1}{2}e^{-t}dt = 1 - 2/3 + 1 = 4/3$$

- Using the fact that the distribution of ξ is $\mu_\xi = \frac{1}{3}\delta_{-1} + \frac{1}{6}\delta_0 + \frac{1}{2}e^{-x}\mathbf{1}_{x\geq 0}dx$ so that

$$\mathbb{E}\left[\xi^2\right] = \int x^2 d\mu_\xi(x) = \tfrac{1}{3}(-1)^2 + \tfrac{1}{6}0^2 + \tfrac{1}{2}\int_0^\infty t^2 e^{-t} = 1/3 + 0 + 2/2 = 4/3$$

b. This is the Cauchy distribution. The distribution of ξ admits a density: $p_\xi(x) = F_\xi'(x) = \frac{1}{\pi(1+x^2)}$. Thus $x^2 p_\xi(x)$ is not integrable and $\mathbb{E}[\xi^2] = +\infty$. We can also check that Equation 1 (say with $a = 0$) gives the same result since

$$\mathbb{E}\left[\xi^2\right] = 0^2 + 2\int_0^\infty 2t(1/2 - \arctan(t))dt = +\infty$$

Exercise 7

Let ξ be distributed according to the Cauchy law with density $\frac{1}{\pi}\frac{1}{1+x^2}$. Find the quantile $q_{2/3}$ for $|\xi|$.

Solution

We need to solve $\mathbb{P}(|\xi| \leq q) = 2/3$. This yields

$$\frac{2}{3} = \int_{-q}^q \frac{1}{\pi}\frac{1}{1+x^2}dx = \frac{2}{\pi}\arctan(q)$$

Therefore $q_{2/3} = \tan(\pi/3) = \sqrt{3}$.

Exercise 8

$n = 100$ letters are randomly placed into n envelopes which already have addresses written on them. Find the expectation and variance of the number of letters that ended up in the correct envelopes.

Solution

Here we are counting fixed points in a random permutation. Let X be the number of letters in correct envelopes. $X = \sum_{i=1}^n X_i$, where X_i is the indicator that the i-th letter went into its own envelope. Since there are $(n-1)!$ permutations that fix the point i

$$\mathbb{E}[X_i] = \mathbb{E}[X_i^2] = \mathbb{P}(X_i = 1) = 1/n$$

Similarly, for $i \neq j$, there are $(n-2)!$ permutations that fix i, j

$$\mathbb{E}[X_i X_j] = \mathbb{P}(X_i = 1, X_j = 1) = \frac{1}{n(n-1)}$$

Therefore

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{i=1}^{n} \mathbb{E}[X_i] = n \cdot (1/n) = 1 \\
\mathbb{E}[X^2] &= \sum_{i,j} \mathbb{E}[X_i X_j] = \sum_{i} \mathbb{E}[X_i^2] + \sum_{i \neq j} \mathbb{E}[X_i X_j] = n \cdot (1/n) + n(n-1) \cdot 1/(n(n-1)) = 2 \\
\text{Var}(X) &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = 2 - 1 - 1
\end{aligned}$$

The answer does not depend on $n \geq 2$.

Exercise 9*

Let the random vector $\xi = (\xi_1, \dots, \xi_n)$ have a uniform distribution on the n-dimensional sphere of radius 1. Find $\text{Var}(\xi_j)$.

Solution

It holds $\mathbb{E}[\xi_i] = 0$ by symmetry. Morever each ξ_i has the same distribution and $\sum_j \xi_j^2 = 1$. Therefore $\mathbb{E}[\xi_i^2] = \text{Var}(\xi_j) = 1/n$.

Exercise 10

Around a round table sit n men and m women. Find the expectation and variance of the number of pairs of neighbors of the type MW (man-woman).

Solution

Let $X_i = 1$ if the pair at seats $(i, i+1)$ (sums are understood $\pmod{n+m}$) is of type MW, and $X_i = 0$ otherwise. Let $X = \sum_i X_i$ be the number of pairs of neighbors of MW. We have

$$\begin{aligned}
\mathbb{E}[X] &= \sum_i \mathbb{E}[X_i] = (m+n)\mathbb{P}(X_1 = 1) = (m+n)\frac{2nm}{(n+m)(n+m-1)} = \frac{2nm}{n+m-1} \\
\mathbb{E}[X^2] &= \sum_{i,j} \mathbb{E}[X_i X_j] = \sum_{i,j\,:\,i=j} \mathbb{E}[X_i X_j] + \sum_{i,j\,:\,|i-j|=1} \mathbb{E}[X_i X_j] + \sum_{i,j\,:\,|i-j|>1} \mathbb{E}[X_i X_j] = \mathbb{E}[X] + 2(n+m)\mathbb{P}(X_1 = 1, X \\
&= \frac{2nm}{n+m-1} + 2(n+m)\frac{nm(n-1) + mn(m-1)}{(m+n)(m+n-1)(m+n-2)} + ((n+m)^2 - 3(n+m))\frac{4nm(n-1)(m}{(m+n-1)(m+} \\
\text{Var}(X) &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{4nm(n-1)(m-1)}{(n+m-1)^2(n+m-2)}
\end{aligned}$$

Indeed

- To compute $\mathbb{P}(X_1 = 1)$, we can have MW or WM, each with probability $nm/((n+m)(n+m-1))$.
- To compute $\mathbb{P}(X_1 = 1, X_2 = 1)$ we can have MWM or WMW.
- To compute $\mathbb{P}(X_1 = 1, X_3 = 1)$, there are four arrangements, MWMW, MWWM, WMMW, WMWM.

Exercise 11

Let η and $\xi_0, \xi_1, \dots$ be independent random variables taking values $0, 1, 2, \dots$, where the ξ_j have identical distributions. Consider the random variable $\beta = \sum_{j=1}^{\eta} \xi_j$. Prove the following relation between generating functions: $Q_\beta = Q_\eta \circ Q_\xi$.

Solution

Conditioning on η

$$Q_\beta(s) := \mathbb{E}[s^\beta] = \sum_{k=0}^{\infty} \mathbb{E}[s^\beta | \eta = k]\mathbb{P}(\eta = k) = \sum_{k=0}^{\infty} \mathbb{E}[s^{\sum_{j=1}^{k} \xi_j}]\mathbb{P}(\eta = k)$$

where in the last step we used that $\sum_{j=1}^{k} \xi_j$ and η are independent. Since the (ξ_j) are i.i.d., the expectation of the product $\prod s^{\xi_j}$ factorizes to get

$$Q_\beta(s) = \sum_{k=0}^{\infty} \prod_{j=1}^{k} \mathbb{E}[s^{\xi_j}]\mathbb{P}(\eta = k) = \sum_{k=0}^{\infty} \mathbb{P}(\eta = k)\mathbb{E}[s^\xi]^k = Q_\eta(Q_\xi(s))$$

Exercise 12

Given an infinite i.i.d. sequence of indicators $\{\xi_i\}, i = 0, 1 \dots$ with parameter $p = 1/3$, and a random variable β, find $\mathbb{E}(\alpha)$ for the discrete random variable $\alpha = \sum_{k=1}^{\beta} \xi_k$.

Solution

We can use the previous exercise to get $Q_\alpha = Q_\beta \circ Q_\xi$. In particular $\mathbb{E}[\alpha] = Q'_\alpha(1) = Q'_\beta(Q_\xi(1))Q'_\xi(1) = Q'_\beta(1)Q'_\xi(1) = \mathbb{E}[\beta]\mathbb{E}[\xi]$.

Exercise 13

Two lecturers teach probability theory. The first one already knows that there are n students in his class. The other, however, has not yet held the first class, so he considers the number N of his students to be a random variable with an expectation equal to n.

An exam is planned, for which the probability of passing is $p \in (0, 1)$ for each student (independent of other students and of N). In which class is the expected number of students who will pass the exam greater? In which class is the variance of the number of students who will pass the exam greater?

Solution

Let Y be the number of students who passed in the first class, and X in the second.
$Y \sim \text{Binomial}(n, p)$. Thus $\mathbb{E}[Y] = np$ and $\text{Var}(Y) = np(1-p)$.
For the second class, the number of students N is a random variable with $\mathbb{E}[N] = n$. We however know that $\mathbb{P}(X = k|N = m) = \binom{m}{k} p^k (1-p)^{m-k}$. Namely X is binomial *conditionally to* N. From the previous exercise $\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|N]] = np$. The expected number of students who pass is the same in both classes.
The variance can be computed as:

$$\text{Var}(X) = \mathbb{E}[\text{Var}(X|N)] + \text{Var}(\mathbb{E}[X|N])$$

where $\text{Var}(X|N) = Np(1-p)$, thus the first term on the r.h.s. is $np(1-p)$. $\mathbb{E}[X|N] = Np$, thus the second term is $p^2 \text{Var}(N)$. In particular $\text{Var}(X) = \text{Var}[Y] + p^2 \text{Var}[N]$. The variance is greater in the second class.

Exercise 14

Let the random variable ξ satisfy $\xi \in L_1(\Omega, \mathbb{P})$ and

a. $\xi = 0, 1, \ldots$. Prove that $\mathbb{E}[\xi] = \sum_{n=1}^{\infty} \mathbb{P}(\xi \geq n)$.
b. $\xi \geq 0$. Prove that $\mathbb{E}[\xi] = \int_0^{\infty} \mathbb{P}(\xi \geq x)\, dx$.

Solution

In general we have, for $X \geq 0$, $\xi = \int_0^{\infty} \mathbf{1}_{x<\xi} dx$, there via Fubini

$$\mathbb{E}[X] = \int_0^{\infty} \mathbb{P}(\xi > x) dx$$

Exercise 15

You have one hour for a nap, but you are waiting for two messages and do not turn off your phone. The messages will wake you up. Assuming the arrival times of the messages are independent and uniformly distributed over this hour, how much time on average will you have left to sleep after they arrive?

Solution

Let the hour be the interval $[0, 1]$. The message arrival times $T_1, T_2 \sim \text{Uniform}([0, 1])$ are independent. You will wake up last time at $M = \max(T_1, T_2)$. The remaining sleep time is $1 - M$. Thus

$$\mathbb{E}[1 - M] = \int_0^1 \mathbb{P}(M \leq t) dt = \int_0^1 \mathbb{P}(T_1 \leq t)\mathbb{P}(T_2 \leq t) dt = \int_0^1 t^2 dt = 1/3$$

or 20 minutes.

Exercise 16

Let X be a random variable. Prove the following statements.

a. If $X \in L^1$, then the median m minimizes the function $\phi_1(r) = \mathbb{E}[|X - r|]$.
b. If $X \in L^2$, then the expectation $\mathbb{E}[X]$ minimizes the function $\phi_2(r) = \mathbb{E}[|X - r|^2]$.
c. Use a. and Jensen's inequality to prove that $|m - \mathbb{E}[X]|^2 \leq \text{Var}[X]$.

> **Solution**
>
> a. Let m be a median and $r > m$. Then $\mathbb{E}[|X - r|] - \mathbb{E}[|X - m|] = \mathbb{E}[|X - r| - |X - m|]$. The integrand is equal to $r - m$ for $X \leq m$, $m - r$ for $X > r$, and $m + r - 2X$ for $m < X \leq r$. Thus
>
> $$\phi_1(r) - \phi(m) = \mathbb{E}[|X - r|] - \mathbb{E}[|X - m|] \geq (r - m)\mathbb{P}(X \leq m) + (m - r)\mathbb{P}(X > r) + (m - r)\mathbb{P}(m < X \leq r) = (r - m$$
>
> If $r < m$, the same computation applied to $-X$ yields a similar result.
>
> b. Let $\mu := \mathbb{E}[X]$. Then
>
> $$\phi_2(r) = \mathbb{E}[(X - r)^2] = \mathbb{E}[(X - \mu)^2] + 2\mathbb{E}[(X - \mu)](r - \mu) + (r - \mu)^2 = \phi_2(\mu) + (r - \mu)^2$$
>
> Thus m is the unique minimizer of ϕ_2.
>
> c. $|\mathbb{E}X - m| \leq \mathbb{E}[|X - m|] \leq \mathbb{E}[|X - \mathbb{E}[X]|]$ from point a. Thus
>
> $$|\mathbb{E}[X] - m|^2 \leq \mathbb{E}[|X - \mathbb{E}[X]|]^2 \leq \mathbb{E}[|X - \mathbb{E}[X]|^2] = \text{Var}(X).$$

Exercise 17*

[In the context of this problem, the Monte Carlo method was first applied in history] A needle of unit length is randomly thrown onto a strip of infinite length and unit width on the plane. What is the probability that the needle will intersect at least one of the lines forming the strip? *Hint: Replace the problem with the following more general question: instead of a needle, consider an arbitrary Lipschitz curve of length ℓ. Find the expected value of the number of its intersections with an infinite lattice formed by parallel lines with a step of 1. Start with a curve in the form of a segment.*

> **Solution**
>
> With probability 1, the needle cannot intersect more than one line. So the probability of intersection is just the expected value of the number of intersections.
> Consider a segment of length ℓ, partition it in finitely many intervals, and let X_i be the number of intersections of the interval i, with the vertical lines on the plane. The number of intersections $X = \sum_i X_i$ satisfies
>
> $$\mathbb{E}[X] = \sum_i \mathbb{E}[X_i] \tag{2}$$
>
> therefore $\mathbb{E}[X]$ is just a linear function of ℓ. We need to find out the constant of proportionality. Consider now a finite union of segments. We can reason in the same way, and the number of intersections will be proportional to the total length of the segments.
> As we take a piecewise linear curve to approximate a smooth curve, for instance a circle, we see that the number of intersections of the approximations converges a.s. to the number of intersection of the curve (Which for instance for a circle is bounded). Thus for a smooth curve $\mathcal{C}$ the number of intersections $X_{\mathcal{C}}$ satisfies
>
> $$\mathbb{E}[X_{\mathcal{C}}] = A\ell(\mathcal{C})$$
>
> where A is a constant independent of the curve $\mathcal{C}$ and $\ell(\mathcal{C})$ is the length of the curve. To determine A, we see that a circle of diameter 1 has a.s. 2 intersections with the vertical lines. Therefore $2 = A\pi$. Thus for a smooth (or rectifiable) curve
>
> $$\mathbb{E}[X_{\mathcal{C}}] = 2\ell(\mathcal{C})/\pi$$

Exercise 18*

$S^1 = \mathbb{R}/\mathbb{Z}$ is a circle of length 1. Let $I_0 = [0, 1/n] \subset S^1$ and $I_k = \frac{k}{n} + I_0 = [k/n, (k+1)/n]$, $k = 1, \dots, n-1$. Then $|I_0| = 1/n$ and $| \cup_{i=0}^{n-1} I_k| = 1$. Thus, we have constructed $ n $ shifts of the set I_0 such that their union has full measure.

However, in the general case, we have a different picture. For $E \subset S^1$ and $x \in S^1$, let $x + E := \{x + y, y \in E\}$. Let the set E be measurable and $n \geq 1$. Prove that there exist $x_1, \dots, x_n$ such that $| \cup_{i=1}^{n} (x_i + E)| \geq 1 - (1 - |E|)^n$.

Solution

Choose $x_1, \dots, x_n$ independently with a uniform distribution on S^1. By Fubini's theorem:

$$\begin{aligned}
\mathbb{E}[| \cup_{i=1}^{n} (x_i + E)|] &= \int_{S^1} \mathbb{E}[\mathbf{1}_{\cup_i (x_i+E)}(y)]dy \\
&= \int_{S^1} (1 - \mathbb{P}(\cap_i \{y \notin x_i + E\}))\, dy = 1 - \int_{S^1} \mathbb{P}(\cap_i \{x_i \notin y - E\})dy \\
&= 1 - \int_{S^1} \prod_i \mathbb{P}(x_i \notin y - E)dy = 1 - (1 - |E|)^n
\end{aligned}$$

Since the average value of the measure of the union is $1 - (1 - |E|)^n$, there must exist at least one specific placement $(x_1, \dots, x_n)$, for which the measure of the union is not less than this average value.

Exercise 19*

Find the expectation and variance of $\max(X, Y)$, where X, Y are the random variables from exercise 3.

Solution

Since X, Y are independent

$$\mathbb{E}[\max(X, Y)] = \int_0^\infty 1 - \mathbb{P}(\max(X, Y) \leq t)dt = \int_0^\infty 1 - \mathbb{P}(X \leq t)\mathbb{P}(Y \leq t)dt$$

For the variance, we can compute it is $\mathrm{Var}(\max(X, Y)) = \mathbb{E}[\max(X, Y)^2] - \mathbb{E}[\max(X, Y)]^2$. So we are left to compute

$$\mathbb{E}[\max(X, Y)^2] = \int_0^\infty 2t(1 - \mathbb{P}(X \leq t)\mathbb{P}(Y \leq t))dt$$

Exercise 20*

Let $f \in C^1(\mathbb{R})$ (or just absolutely continuous), ξ be a random variable and $f(\xi) \in L_1$. Prove that $\forall a \in \mathbb{R}$

$$\mathbb{E}[f(\xi)] = f(a) + \int_a^\infty f'(x)\mathbb{P}(\xi \geq x)\, dx - \int_{-\infty}^a f'(x)\mathbb{P}(\xi \leq x)\, dx.$$

Solution

We can restrict to the case $a = 0$ and $f(0) = 0$, by considering the function $f(\cdot + a) - f(a)$ otherwise. Let us also consider the case $\xi \geq 0$, the general case being similar. Then by Fubini

$$f(\xi) = \int_0^{\xi} f'(t)dt = \int_0^{\infty} f'(t)\mathbf{1}_{0\leq t\leq \xi}dt$$
$$\mathbb{E}[f(\xi)] = \int_0^{\infty} f'(t)\mathbb{E}[\mathbf{1}_{0\leq t\leq \xi}]dt$$

which is indeed the stated formula.

Exercise 21*

Let E be an ordered measurable space. Let $X\colon \Omega \to E$ be a random variable. Let $f, g\colon E \to \mathbb{R}$ be measurable monotone functions. Prove that

$$\mathbb{E}[f(X)\,g(X)] \geq \mathbb{E}[f(X)]\mathbb{E}[g(X)].$$

Solution

Take X, Y i.i.d.. Then, since $(f(y) - f(x))(g(y) - g(x))$ is pointwise non-negative

$$0 \leq \mathbb{E}[(f(Y) - f(X))(g(Y) - g(X))] = \mathbb{E}[f(X)g(X) + f(Y)g(Y)] - \mathbb{E}[f(X)g(Y) + f(Y)g(X)]$$

Since X, Y are i.i.d., this gives the wanted inequality.

Exercise 22*

Let the random variables ξ, η satisfy $\mathbb{E}\xi = \mathbb{E}\eta = 0$, and $\operatorname{Var}[\xi] = \operatorname{Var}[\eta] = 1$ and have correlation coefficient ρ. Prove that

$$\mathbb{E}\max(\xi^2, \eta^2) \leq 1 + \sqrt{1 - \rho^2}.$$

Solution

For $a, b \in \mathbb{R}$ it holds $\max(a, b) = \frac{a+b+|a-b|}{2}$. Thus:

$$\mathbb{E}\max(\xi^2, \eta^2) = \frac{1}{2}\mathbb{E}[\xi^2 + \eta^2] + \frac{1}{2}\mathbb{E}[|\xi^2 - \eta^2|] = \frac{1}{2}\operatorname{Var}[\xi] + \frac{1}{2}\operatorname{Var}[\eta] + \frac{1}{2}\mathbb{E}[|\xi + \eta||\xi - \eta|]$$

Moreover we have $\mathbb{E}[\xi\eta] = \rho\sqrt{\operatorname{Var}[\xi]\operatorname{Var}[\eta]}$, and

$$\mathbb{E}[|\xi + \eta||\xi - \eta|] \leq \mathbb{E}[(\xi + \eta)^2]^{1/2}\mathbb{E}[(\xi - \eta)^2]^{1/2}$$

Putting all together

$$\mathbb{E}\max(\xi^2, \eta^2) \leq \frac{\operatorname{Var}[\xi] + \operatorname{Var}[\eta]}{2} + \frac{1}{2}\sqrt{(\operatorname{Var}[\xi] + \operatorname{Var}[\eta])^2 - 4\rho^2\operatorname{Var}[\xi]\operatorname{Var}[\eta]}$$

Exercise 23*

The random variables $\xi_1, \xi_2, \ldots$ are iid, $\xi_j \sim \mathrm{Uniform}([0,1])$. Let ν be a random variable equal to the minimum k for which $\sum_{i=1}^k \xi_i \geq 1$. Find $\mathbb{E}[\nu]$.

> **Solution**
>
> For $S_n = \sum_{i=1}^n \xi_i$, it holds $\{\nu > n\} = \{S_n < 1\}$. Thus
>
> $$\mathbb{E}[\nu] = \sum_{n=0}^{\infty} \mathbb{P}(\nu > n) = \sum_{n=0}^{\infty} \mathbb{P}(S_n < 1)$$
>
> The density of S_n on $[0,1]$ is $x^{n-1}/(n-1)!$. Therefore
>
> $$\mathbb{E}[\nu] = \int_0^1 \sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!} dx = e$$

Additional Exercises

Exercise 24*

Let ξ be a random variable and $f \colon \mathbb{R} \to \mathbb{R}$ be absolutely continuous and such that $\mathbb{E}[|f(\xi)|] < \infty$. Prove that for $a \in \mathbb{R}$

$$\mathbb{E}[f(\xi)] = f(a) + \int_a^{\infty} f'(t)\mathbb{P}(\xi > t)dt - \int_{-\infty}^{a} f'(t)\mathbb{P}(\xi \leq t)dt$$

In particular if f admits a limit at $-\infty$

$$\mathbb{E}[f(\xi)] = f(-\infty) + \int_{-\infty}^{\infty} f'(t)\mathbb{P}(\xi > t)dt \tag{3}$$

> **Solution**
>
> Since f is absolutely continuous, for ω such that $\xi(\omega) \geq a$ (notice that at least one of the integral terms vanishes for each ω, depending on $\xi(\omega) > a$ or $\xi(\omega) < a$)
>
> $$f(\xi) = f(a) + \int_a^{\infty} \mathbf{1}_{\xi > t} f'(t)dt - \int_{-\infty}^{a} \mathbf{1}_{\xi \leq t} f'(t)dt$$
>
> Taking expectation inside the integrals, we get the wanted formula.

Exercise 25*

Let ξ be a random variable and $\varepsilon \in (0,1]$. Define

$$\begin{aligned}\varphi(t) &:= -\log \mathbb{P}(\xi > t) \in [0, +\infty] \\ \psi(t) &:= -\log \mathbb{P}(\xi \geq t) \in [0, +\infty]\end{aligned}$$

Prove that

$$\mathbb{E}[e^{(1-\varepsilon)\psi(\xi)}] \le \varepsilon^{-1} \le \mathbb{E}[e^{(1-\varepsilon)\varphi(\xi)}]$$

In particular there is equality if $\mathbb{P}(\xi = t) = 0$ for all t.

Solution

First take an absolutely continuous, non-decreasing function $\chi \ge \varphi$. Using Equation 3

$$\mathbb{E}[e^{(1-\varepsilon)\chi(\xi)}] = 1 + \int_0^\infty (1-\varepsilon)e^{(1-\varepsilon)\chi(t)}\chi'(t)\mathbb{P}(\xi > t)dt \ge 1 + \int_0^\infty (1-\varepsilon)e^{(1-\varepsilon)\chi(t)}\chi'(t)e^{-\chi(t)}dt$$
$$= 1 + (1-\varepsilon)\int_0^\infty e^{-\varepsilon\chi(t)}\chi'(t)dt = 1 + (1-\varepsilon)/\varepsilon = 1/\varepsilon$$

And similarly if we consider $\chi \le \psi$. Therefore

$$\sup_{\chi \le \psi, \chi \text{a.c.}} \mathbb{E}[e^{(1-\varepsilon)\chi(\xi)}] \le \varepsilon^{-1} \le \sup_{\chi \ge \varphi, \chi \text{a.c.}} \mathbb{E}[e^{(1-\varepsilon)\chi(\xi)}]$$

Now the point is that we can approximate φ with smooth functions from above, and ψ from below. E.g. take η_n with a smooth density supported in $[0, 1/n]$ and independent of ξ. Then set

$$\varphi_n(t) := -\log \mathbb{P}(\xi + \eta_n > t) \in [\varphi(t - 1/n), \psi(t)], \qquad \psi_n(t) := -\log \mathbb{P}(\xi - \eta_n \ge t) \in [\varphi(t), \psi(t + 1/n)]$$

and take the limit $n \to \infty$.