1. Let $\text{cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X) \cdot \mathbb{E}(Y)$ denote the covariance of two random variables $X, Y$.

(a) Show that $\text{cov}(X,Y) = \mathbb{E}[(X - \mu_X)(Y - \mu_Y)]$.

(b) Show that $\sigma^2(X + Y) = \sigma^2(X) + \sigma^2(Y) + 2 \text{cov}(X,Y)$.

(c) Let $X \sim \text{Unif}[-1,1]$. Show that the random variables $X$ and $X^2$ are uncorrelated, but not independent.

(d) Let $p(X,Y)$ be the correlation coefficient of $X, Y$. Show that $|p(X,Y)| \leq 1$.

*Hint:* Imitate the proof of the Cauchy-Schwarz inequality from linear algebra to show that $|\mathbb{E}(X_1X_2)|^2 \leq \mathbb{E}(X_1)^2 \cdot \mathbb{E}(X_2^2)$ for any two RV’s $X_1, X_2$. The claim follows from this by a short calculation.

**Solution:**

(a)

$$
\mathbb{E}[(X - \mu_X)(Y - \mu_Y)] = \mathbb{E}(XY) - \mu_X \mathbb{E}(Y) - \mu_Y \mathbb{E}(X) + \mu_X \mu_Y \\
= \mathbb{E}(XY) - \mu_X \mu_Y = \text{cov}(X,Y).
$$

(b)

$$
\sigma^2(X + Y) = \mathbb{E}(X + Y - \mu_X + Y - \mu_Y)^2 \\
= \mathbb{E}(X - \mu_X)^2 + \mathbb{E}(Y - \mu_Y)^2 + 2 \mathbb{E}(X - \mu_X)(Y - \mu_Y) \\
= \sigma^2(X) + \sigma^2(Y) + 2 \text{cov}(X,Y).
$$

(c) The density function of $X$ is

$$
f(x) = \begin{cases} 
\frac{1}{2} & \text{if } -1 \leq x \leq 1, \\
0 & \text{otherwise}.
\end{cases}
$$

The function $f(x)$ is even, that is, $f(x) = f(-x)$, so

$$
\mathbb{E}(X) = \int_{-\infty}^{\infty} xf(x) \, dx = 0 \quad \text{and} \quad \mathbb{E}(X^3) = \int_{-\infty}^{\infty} x^3 f(x) \, dx = 0.
$$

Therefore

$$
\text{Cov}(X, X^2) = E(X^3) - \mathbb{E}(X)\mathbb{E}(X^2) = 0,
$$

so $X$ and $X^2$ are uncorrelated. By definition $X$ and $X^2$ are independent if

$$
\mathbb{P}(X \in A, X^2 \in B) = \mathbb{P}(X \in A)\mathbb{P}(X^2 \in B)
$$

for all (reasonable) sets $A, B \subset \mathbb{R}$, in particular for intervals. For example let $A = B = [0,1/4]$, then

$$
\mathbb{P}(X \in A, X^2 \in A) = \mathbb{P}(X \in A, X \in [-1/2,1/2]) = \mathbb{P}(X \in A) = \frac{1}{8}.
$$

On the other hand side

$$
\mathbb{P}(X \in A)\mathbb{P}(X^2 \in A) = \mathbb{P}(X \in A)\mathbb{P}(X \in [-1/2,1/2]) = \frac{1}{8} \cdot \frac{1}{2} = \frac{1}{16}.
$$

Therefore

$$
\mathbb{P}(X \in A, X^2 \in A) \neq \mathbb{P}(X \in A)\mathbb{P}(X^2 \in A),
$$
so $X$ and $X^2$ are not independent.

(d) We first show the hint. Set $Z = X_1 - \frac{\mathbb{E}(X_1 X_2)}{\mathbb{E}X_2^2} X_2$ (if $\mathbb{E}X_2^2 = 0$, then $X_2 = 0$, and the statement is trivial). Then $X_1 = Z + \frac{\mathbb{E}(X_1 X_2)}{\mathbb{E}X_2^2} X_2$ and

$$
\mathbb{E}(Z X_2) = \mathbb{E}(X_1 X_2) - \frac{\mathbb{E}(X_1 X_2)}{\mathbb{E}X_2^2} \mathbb{E}X_2^2 = 0.
$$

Therefore,

$$
\mathbb{E}X_1^2 = \mathbb{E}\left(Z + \frac{\mathbb{E}(X_1 X_2)}{\mathbb{E}X_2^2} X_2 \right)^2
= \mathbb{E}Z^2 + \left(\frac{\mathbb{E}(X_1 X_2)}{\mathbb{E}X_2^2}\right)^2 \mathbb{E}X_2^2 + 2 \frac{\mathbb{E}(X_1 X_2)}{\mathbb{E}X_2^2} \mathbb{E}(Z X_2)
= \mathbb{E}Z^2 + \left(\frac{\mathbb{E}(X_1 X_2)}{\mathbb{E}X_2^2}\right)^2 \mathbb{E}X_2^2 \geq \left(\mathbb{E}(X_1 X_2)\right)^2 \mathbb{E}X_2^2
$$

since $\mathbb{E}Z^2 \geq 0$. Rearranging this inequality gives $|\mathbb{E}(X_1 X_2)|^2 \leq \mathbb{E}X_1^2 \mathbb{E}X_2^2$.

Now, since the covariance is $\text{cov}(X, Y) = \mathbb{E}[(X - \mu_X)(Y - \mu_Y)]$ we conclude from the hint that

$$
\text{cov}(X, Y)^2 \leq \mathbb{E}(X - \mu_X)^2 \cdot \mathbb{E}(Y - \mu_Y)^2 = \sigma^2(X) \sigma^2(Y).
$$

This immediately implies that $|p(X, Y)| \leq 1$

2. Suppose that $X, Y$ are discrete random variables with joint p.m.f. as shown below. Calculate $\text{cov}(X, Y)$ and $p(X, Y)$.

<table>
<thead>
<tr>
<th>$X \downarrow Y \rightarrow$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1/15</td>
<td>1/15</td>
<td>2/15</td>
<td>1/15</td>
</tr>
<tr>
<td>2</td>
<td>1/10</td>
<td>1/10</td>
<td>1/5</td>
<td>1/10</td>
</tr>
<tr>
<td>3</td>
<td>1/30</td>
<td>1/30</td>
<td>0</td>
<td>1/10</td>
</tr>
</tbody>
</table>

**Solution:** The marginal probability mass functions are

$$
p_X(1) = 1/3, \ p_X(2) = 1/2, \ p_X(3) = 1/6
$$

and

$$
p_Y(0) = p_Y(1) = 1/5, \ p_Y(2) = 1/3, \ p_Y(3) = 4/15.
$$

Thus

$$
\mathbb{E}(X) = \frac{11}{6} \quad \text{and} \quad \mathbb{E}(Y) = \frac{5}{3}.
$$

From the table we have

$$
\mathbb{E}(XY) = 1 \cdot \frac{1}{15} + 2 \cdot \frac{2}{15} + 3 \cdot \frac{1}{15} + 2 \cdot \frac{1}{10} + 4 \cdot \frac{1}{5} + 6 \cdot \frac{1}{10} + 3 \cdot \frac{1}{30} + 9 \cdot \frac{1}{10} = \frac{47}{15}.
$$

Thus

$$
\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \frac{47}{15} - \frac{55}{18} = \frac{7}{90}.
$$
From the marginal probability mass functions

\[ E(X^2) = \frac{23}{6} \quad \text{and} \quad E(Y^2) = \frac{59}{15}, \]

so

\[ \text{Var}(X) = \frac{17}{36} \quad \text{and} \quad \text{Var}(Y) = \frac{52}{45}. \]

We have

\[ \text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \text{Var}(Y)}} \approx 0.1053. \]

3. Let \( X \) be the number of rolls of a fair die until I see the first six. Next, I choose a sample, with replacement, of size \( X \) from an urn with 5 red and 4 green balls. Let \( Y \) be the number of green balls in my sample.

(a) Compute the conditional probability distribution \( p_{Y \mid X}(y \mid x) \).

(b) Compute \( E[Y \mid X] \).

(c) Use your answer from part (b) to compute \( E[Y] \).

Solution:

(a) Given \( X = x \), the distribution of \( Y \) is Bin\((x, 4/9)\), so

\[ p_{Y \mid X}(y \mid x) = \begin{cases} \binom{x}{y} \left(\frac{4}{9}\right)^y \left(\frac{5}{9}\right)^{x-y} & y \in \{0, \ldots, x\} \\ 0 & \text{else} \end{cases} \]

(b) Since \( E\text{Bin}(x, 4/9) = \frac{4}{9}x \), we have

\[ E(Y \mid X = x) = \frac{4}{9}x. \]

Thus

\[ E(Y \mid X) = \frac{4}{9}X. \]

(c) Using that \( X \sim \text{Geom}(1/6) \) we have

\[ E(Y) = E(E(Y \mid X)) = E((4/9)X) = \frac{4}{9} \cdot 6 = \frac{8}{3}. \]

4. Assume that the random variables \( X, Y \) have joint density function

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

(a) What is the conditional probability density function of \( X \) given that \( Y = 1 \)?

(b) Find the conditional expectation \( E(X \mid Y = 1) \).

Solution: (a) If \( 0 \leq y \leq 2 \) then we have

\[ f_Y(y) = \int_{-\infty}^{\infty} f(x, y) \, dx = \frac{y + 1}{4}. \]
Therefore
\[ f_{X|Y}(x \mid 1) = \frac{f(x, 1)}{f_Y(1)} = \begin{cases} \frac{2x+1}{2} & \text{if } 0 \leq x \leq 1, \\ 0 & \text{otherwise.} \end{cases} \]

(b) We have
\[ \mathbb{E}(X \mid Y = 1) = \int_{-\infty}^{\infty} xf_{X|Y}(x \mid 1) \, dx = \int_{0}^{1} x \frac{2x+1}{2} \, dx = \left( \frac{x^3}{3} + \frac{x^2}{4} \right) \big|_{0}^{1} = \frac{7}{12}. \]

5. Suppose a laser pointer is located at the origin of your coordinate system, and is pointing toward the vertical line \( L = \{(x, y) \text{ s.t. } x = 1\} \). Suppose that the angle \( X \) between the laser beam and the \( x \) axis is a uniform random variable. Calculate the p.d.f. of the \( Y \) coordinate of the point on \( L \) which the beam points at. **Hint:** Draw a picture and argue that this means that \( X \sim \text{Unif}[-\pi/2, \pi/2] \) and \( Y = \tan X \).

**Solution:**
The distance between the laser pointer and the line \( L \) is 1, and therefore the tangens of the angle of the beam with the \( x \) axis is exactly the \( Y \) coordinate of the point on \( L \) the laser pointer is pointing at. The angles between the pointer and the \( x \) axis can be in the interval \([-\pi/2, \pi/2]\) in order for the beam to point towards \( L \) (and not away from it). Therefore, \( X \sim \text{Unif}[-\pi/2, \pi/2] \), and \( Y = \tan X \).

We now first compute the c.d.f. of \( Y \), using the fact that tan is a strictly increasing function:
\[ F_Y(b) = \mathbb{P}(Y \leq b) = \mathbb{P}(\tan X \leq b) \]
\[ = \mathbb{P}(X \leq \arctan b) = \int_{-\pi/2}^{\arctan b} \frac{1}{\pi} \, dx = \frac{\arctan b + \frac{\pi}{2}}{\pi}. \]

We get the p.d.f. by differentiating:
\[ f_Y(y) = F_Y'(y) = \frac{1}{\pi} \frac{d}{dy} \arctan y = \frac{1}{\pi} \frac{1}{1 + y^2}. \]
This distribution is called the Cauchy distribution.

6. Let \( X_1, X_2 \sim \text{Exp}(\lambda) \) be independent. Calculate the p.d.f. of \( X_1 + X_2 \).
Solution:
By the convolution formula
\[ f_{X_1+X_2}(x) = \int_{-\infty}^{\infty} f_{X_1}(x_1) f_{X_2}(x-x_1) \, dx_1 \]
\[ = \int_0^{\infty} \lambda e^{-\lambda x_1} \lambda e^{-\lambda(x-x_1)} \, dx_1 \]
\[ = \lambda^2 e^{-\lambda x} \int_0^x 1 \, dx_1 \]
\[ = \lambda^2 xe^{-\lambda x}. \]
To find the correct boundaries of integration in the second step, we noticed that
\[ f_{X_1}(x_1) = 0 \text{ unless } x_1 \geq 0 \text{ and } f_{X_2}(x-x_1) = 0 \text{ unless } x-x_1 \geq 0 \iff x_1 \leq x. \]The random variable with such a p.d.f. is called the \( \Gamma(2, \lambda) \) RV.

7. Let \( X \sim \text{Poisson}(\lambda) \).
(a) Calculate the moment generating function of \( X \).
(b) Let \( Y \sim \text{Poisson}(\mu) \) be independent of \( X \). Show that \( X + Y \sim \text{Poisson}(\lambda + \mu) \).

**Hint:** What is the m.g.f. of \( X + Y \)? Remember that the m.g.f. uniquely determines the p.m.f.

**Solution:**
(a) We have
\[ M_X(t) = E(e^{tX}) \]
\[ = \sum_{k=0}^{\infty} e^{tk} P(X = k) \]
\[ = \sum_{k=0}^{\infty} e^{tk} \frac{\lambda^k}{k!} e^{-\lambda} \]
\[ = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^t)^k}{k!} = e^{\lambda(e^t-1)}. \]
(b) Since \( X \) and \( Y \) are independent, we have
\[ M_{X+Y}(t) = M_X(t)M_Y(t) = e^{\lambda(e^t-1)}e^{\mu(e^t-1)} = e^{(\lambda+\mu)(e^t-1)}. \]
This is the m.g.f. of a \( \text{Poisson}(\lambda + \mu) \) RV, and since the m.g.f. determines the distribution, we have proven that \( X + Y \sim \text{Poisson}(\lambda + \mu) \)

8*. Let \( X \) be a continuous random variable with p.d.f. \( f(x) \) and \( g : \mathbb{R} \to \mathbb{R} \) be a strictly increasing function. Compute the p.d.f. of \( g(X) \).
Solution:
We compute the c.d.f., using that $g$ is strictly increasing.

$$F_{g(X)}(b) = P(g(X) \leq b) = P(X \leq g^{-1}(b)) = F_X(g^{-1}(b)).$$

Here, $g^{-1}(b)$ is the inverse function of $g$ (e.g. $g^{-1}(b) = \sqrt{b}$ if $g(x) = x^2$, or $g^{-1}(b) = \arctan b$ if $g(x) = \tan x$). Using the chain rule

$$\frac{d}{dx} F_{g(X)}(x) = F_X'(g^{-1}(x)) \cdot \frac{d}{dx} g^{-1}(x)$$

$$= f_X(g^{-1}(x)) \cdot \frac{1}{g'(g^{-1}(x))},$$

where we used a theorem about the derivative of the inverse function.