1) In a bowl there are 3 balls numbered form 1 to 3. You extract two of them without reinsertion. Let X_1 the result of the first extraction and X_2 the result of the second one. Compute:

a) the joint probability mass function $p(x_1, x_2)$ of X_1 and X_2.
There are six possible outcomes: $\{(1,2),(1,3),(2,3),(2,1),(3,2),(3,1)\}$ *and they have all the same probability. So we have*

$$\begin{aligned} &p(1,2) = p(1,3) = p(2,3) = p(2,1) = p(3,2) = p(3,1) = \frac{1}{6} \\ &p(1,1) = p(2,2) = p(3,3) = 0 \end{aligned}$$

b) the probability mass function of $Y = X_1+X_2$ and the expected value $E(X_1+X_2)$.
There are 3 possible values for Y*:* 3*,* 4 *and* 5*. Each of them can be obtained in two ways. So that we have*

$$p_Y(3) = p_Y(4) = p_Y(5) = \frac{1}{3}$$

The expected value is

$$E(X_1 + X_2) = \frac{1}{3}(3+4+5) = 4$$

c) the marginal $p_{X_1}(x_1)$ with respect to X_1 of $p(x_1, x_2)$ and the conditional probability mass function $p_{X_1|X_2}(x_1|x_2)$.
The marginal $p_{X_1}(x_1)$ *is just the probability of getting* x_1 *at the first extraction so that:*

$$p_{X_1}(1) = p_{X_1}(2) = p_{X_1}(3) = \frac{1}{3}$$

The conditional probability mass function is given by:

$$p_{X_1|X_2}(x_1|x_2) = \frac{p(x_1,x_2)}{p_{X_2}(x_2)}$$

where

$$p_{X_2}(1) = p_{X_2}(2) = p_{X_2}(3) = \frac{1}{3}$$

like for X_1*. So we get*

$$\begin{aligned} p_{X_1|X_2}(1|2) =& p_{X_1|X_2}(1|3) = p_{X_1|X_2}(2|3) = p_{X_1|X_2}(2|1) = \\ &= p_{X_1|X_2}(3|2) == p_{X_1|X_2}(3|1) = \frac{1}{2} \\ p_{X_1|X_2}(1|1) =& p_{X_1|X_2}(2|2) = p_{X_1|X_2}(3|3) = 0 \end{aligned}$$

d) $\operatorname{cov}(X, Y)$. Are X_1 and X_2 independent? Why?

We have that

$$E(X_1) = E(X_2) = \frac{1}{3}(1 + 2 + 3) = 2$$

and

$$E(X_1 X_2) = \frac{1}{6}(1 \cdot 2 + 1 \cdot 3 + 2 \cdot 3 + 2 \cdot 1 + 3 \cdot 1 + 3 \cdot 2) = \frac{11}{3}$$

so that

$$\operatorname{cov}(X, Y) = \frac{11}{3} - 4 = -\frac{1}{3}.$$

The covariance is not 0 so they cannot be independent.

2) Let X be an exponential rv with parameter $\lambda = 2$ and Y be a uniform rv between 0 and 1. Assume that X and Y are independent. Compute:

a) the joint probability distribution function $f(x,y)$ of X and Y.
The variable are independent so that the joint probability distribution function $f(x,y)$ is the product of the marginal:

$$f(x,y) = \begin{cases} 2e^{-2x} & x > 0 \text{ and } 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

b) the probability that $X > 3$ and $Y > 0.5$
We have that $P(\{X > 3\} \cap \{Y > 0.5\}) = P(X > 3)PY > 0.5) = e^{-2\cdot 3} \cdot 0.5 = e^{-6}/2$.

c) the probability that X is larger than Y, that is $P(X > Y)$.

$$P(X > Y) = \int\int_{x>y} f(x,y)dxdy = \int_0^1 \int_y^\infty 2e^{-2x} dxdy =$$
$$= \int_0^1 e^{-2y} dy = \frac{1}{2}(1 - e^{-2})$$

(Bonus) Assume now that X and Y are not independent but X is still exponential with parameter $\lambda = 2$ while the conditional probability of Y given X is $f_{Y|X}(y|x) = 1/x$ for $0 < y < x$ and $f_{Y|X}(y|x) = 0$ otherwise. Compute

d) the joint probability distribution function $f(x,y)$ of X and Y.
We have

$$f(x,y) = f_{Y|X}(y|x)f_X(x) = \begin{cases} \frac{2}{x}e^{-2x} & x > 0 \text{ and } 0 < y < x \\ 0 & \text{otherwise} \end{cases}$$

e) the covariance of X and Y, $\mathrm{cov}(X,Y)$.
(**Hint**: remember that:

$$\int_0^\infty \lambda x e^{-\lambda x} = \frac{1}{\lambda} \qquad \int_0^\infty \lambda x^2 e^{-\lambda x} = \frac{2}{\lambda^2}$$

).

$$E(X) = \int_0^\infty \int_0^x 2e^{-2x} dydx = \int_0^\infty 2xe^{-2x} dx = \frac{1}{2}$$
$$E(Y) = \int_0^\infty \int_0^x \frac{2y}{x} e^{-2x} dydx = \int_0^\infty xe^{-2x} dx = \frac{1}{4}$$

while

$$E(XY) = \int_0^\infty \int_0^x 2ye^{-2x} dydx = \int_0^\infty x^2 e^{-2x} dx = \frac{1}{4}$$

so that

$$\mathrm{cov}(X,Y) = \frac{1}{4} - \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{4}.$$

3) You reach campus driving while a friend of yours use public transportation. The time you need to reach campus is described by a rv X with $E(X) = 1h$ and standard deviation $\sigma_X = 0.3h$ while the time needed by your friend is a rv Y with $E(Y) = 1.2h$ and standard deviation $\sigma_Y = 0.5h$. In a semester you both go to campus 100 times. Moreover the times needed to reach campus on different days are independent. Let T_x be the total time you spend driving during a semester.

   a) What is the (approximate) distribution of T_X? Compute $P(T_X > 110)$.
   T_X is (approximatively) a normal random variable with parameters $\mu = 100$ and $\sigma^2 = 100 \cdot 0.3^2$. So that we have

$$P(T_X > 110) = 1 - \Phi\left(\frac{110 - 100}{10 \cdot 0.3}\right) = 1 - \Phi(3.33) = 0.0004.$$

Call $\bar{X}$ the average time (over the semester) you spend driving to campus and $\bar{Y}$ the average time your friend spend in the public transportation system.

   b) compute $P(\bar{X} > 1.05)$ and $P(1.1 < \bar{Y} < 1.3)$.

$$\bar{X} \simeq N\left(1, \frac{0.3^2}{100}\right) \qquad \bar{Y} \simeq N\left(1.2, \frac{0.5^2}{100}\right)$$

so that

$$P(\bar{X} > 1.05) = 1 - \Phi\left(\frac{10(1.05 - 1)}{0.3}\right) = 1 - \Phi(1.67) = 0.0475$$

$$P(1.1 < \bar{Y} < 1.3) = \Phi\left(\frac{10(1.3 - 1.2)}{0.5}\right) - \Phi\left(\frac{10(1.1 - 1.2)}{0.5}\right) =$$
$$= \Phi(2) - \Phi(-2) = 0.9544$$

   c) compute $P(\bar{X} > \bar{Y})$.
   Let $Z = \bar{X} - \bar{Y}$. Then

$$Z \simeq N\left(-0.2, \frac{0.3^2 + 0.5^2}{100}\right)$$

so that

$$P(\bar{X} > \bar{Y}) = P(\bar{X} - \bar{Y} > 0) = 1 - \Phi\left(\frac{-0.2 \cdot 10}{\sqrt{0.3^2 + 0.5^2}}\right) = 1 - \Phi(3.42) = 0.0003$$