Machine Learning:
Assignment Sheet #2

Due on February 22, 2022 at 10:00

Group HB
Henri Sota, Enis Mustafaj

Problem 2.1

We consider the chance experiment of a fair coin that is tossed four times.

a) Give the sample space.

The sample space is:

$$\Omega = \{(HHHH), (HHHT), (HHTH), (HTHH), (THHH), (HHTT), (HTHT), (THHT),\\ (THTH), (TTHH), (HTTH), (HTTT), (THTT), (TTHT), (TTTH), (TTTT)\}$$

b) Define the random variable X to be the number of heads, Y as the number of tails and Z as $Z = |X - Y|$.
Random variables X, Y, Z are defined as following:

$$\begin{array}{llll}
X(TTTT) = 0, & X(TTTH) = 1, & X(TTHT) = 1, & X(THTT) = 1 \\
X(HTTT) = 1, & X(HHTT) = 2, & X(HTHT) = 2, & X(THHT) = 2 \\
X(THTH) = 2, & X(TTHH) = 2, & X(HTTH) = 2, & X(HHHT) = 3 \\
X(HHTH) = 3, & X(HTHH) = 3, & X(THHH) = 3, & X(HHHH) = 4
\end{array}$$

$$\begin{array}{llll}
Y(TTTT) = 4, & Y(TTTH) = 3, & Y(TTHT) = 3, & Y(THTT) = 3 \\
Y(HTTT) = 3, & Y(HHTT) = 2, & Y(HTHT) = 2, & Y(THHT) = 2 \\
Y(THTH) = 2, & Y(TTHH) = 2, & Y(HTTH) = 2, & Y(HHHT) = 1 \\
Y(HHTH) = 1, & Y(HTHH) = 1, & Y(THHH) = 1, & Y(HHHH) = 0
\end{array}$$

$$\begin{array}{llll}
Z(TTTT) = 4, & Z(TTTH) = 2, & Z(TTHT) = 2, & Z(THTT) = 2 \\
Z(HTTT) = 2, & Z(HHTT) = 0, & Z(HTHT) = 0, & Z(THHT) = 0 \\
Z(THTH) = 0, & Z(TTHH) = 0, & Z(HTTH) = 0, & Z(HHHT) = 2 \\
Z(HHTH) = 2, & Z(HTHH) = 2, & Z(THHH) = 2, & Z(HHHH) = 4
\end{array}$$

c) Compute the joint PMF of (X, Z).
The PMF of variables X and Z is is defined as following:

$$p(0, 4) = p(\{(TTTT)\} \cap \{(TTTT), (HHHH)\}) = p(\{(TTTT)\}) = \frac{1}{16}$$

$$p(1, 1) = p(\{(TTTH), (TTHT), (THTT), (HTTT)\} \cap \{(TTTH), (TTHT), (THTT),\\ (HTTT), (HHHT), (HHTH), (HTHH), (THHH)\}) = p(\{(TTTH), (TTHT), (THTT), (HTTT)\}) = \frac{1}{4}$$

$$p(2, 0) = p(\{(HHTT), (HTHT), (THHT), (THTH), (TTHH), (HTTH)\} \cap \{(HHTT), (HTHT), (THHT),\\ (THTH), (TTHH), (HTTH)\}) = p(\{(HHTT), (HTHT), (THHT), (THTH), (TTHH), (HTTH)\}) = \frac{3}{8}$$

$$p(3, 1) = p(\{(HHHT), (HHTH), (HTHH), (THHH)\} \cap \{(TTTH), (TTHT), (THTT),\\ (HTTT), (HHHT), (HHTH), (HTHH), (THHH)\}) = p(\{(HHHT), (HHTH), (HTHH), (THHH)\}) = \frac{1}{4}$$

$$p(4, 4) = p(\{(HHHH)\} \cap \{(TTTT), (HHHH)\}) = p(\{(HHHH)\}) = \frac{1}{16}$$

PMF for the other values of X and Z is $p(X, Z) = P(\emptyset) = 0$

d) Compute the expectation $E(Y)$. The expectation $E(Y)$ is computed as following:

$$E(Y) = \sum_{y \in R_y} y \cdot P(Y = y) = 0 \cdot \frac{1}{16} + 1 \cdot \frac{1}{4} + 2 \cdot \frac{5}{16} + 3 \cdot \frac{4}{16} + 4 \cdot \frac{1}{16} = \frac{30}{16} = 1.875$$

Problem 2.2

We consider the chance experiment of a fair dice that is rolled two times. Let X_1 be the outcome of the first roll, while X_2 is the outcome of the second roll.

a) Describe the chance experiment by its sample space, and define the random variables.

The sample space is:

$$\begin{aligned}\Omega = \{&11, 12, 13, 14, 15, 16\ 21, 22, 23, 24, 25, 26\ 31, 32, 33, 34, 35, 36\\ &41, 42, 43, 44, 45, 46\ 51, 52, 53, 54, 55, 56\ 61, 62, 63, 64, 65, 66\}\end{aligned}$$

Definition of X_1

$$\begin{array}{llllll}
X_1(11) = 1 & X_1(12) = 1 & X_1(13) = 1 & X_1(14) = 1 & X_1(15) = 1 & X_1(16) = 1 \\
X_1(21) = 2 & X_1(22) = 2 & X_1(23) = 2 & X_1(24) = 2 & X_1(25) = 2 & X_1(26) = 2 \\
X_1(31) = 3 & X_1(32) = 3 & X_1(33) = 3 & X_1(34) = 3 & X_1(35) = 3 & X_1(36) = 3 \\
X_1(41) = 4 & X_1(42) = 4 & X_1(43) = 4 & X_1(44) = 4 & X_1(45) = 4 & X_1(46) = 4 \\
X_1(51) = 5 & X_1(52) = 5 & X_1(53) = 5 & X_1(54) = 5 & X_1(55) = 5 & X_1(56) = 5 \\
X_1(61) = 6 & X_1(62) = 6 & X_1(63) = 6 & X_1(64) = 6 & X_1(65) = 6 & X_1(66) = 6
\end{array}$$

Definition of X_2

$$\begin{array}{llllll}
X_2(11) = 1 & X_2(12) = 2 & X_2(13) = 3 & X_2(14) = 4 & X_2(15) = 5 & X_2(16) = 6 \\
X_2(21) = 1 & X_2(22) = 2 & X_2(23) = 3 & X_2(24) = 4 & X_2(25) = 5 & X_2(26) = 6 \\
X_2(31) = 1 & X_2(32) = 2 & X_2(33) = 3 & X_2(34) = 4 & X_2(35) = 5 & X_2(36) = 6 \\
X_2(41) = 1 & X_2(42) = 2 & X_2(43) = 3 & X_2(44) = 4 & X_2(45) = 5 & X_2(46) = 6 \\
X_2(51) = 1 & X_2(52) = 2 & X_2(53) = 3 & X_2(54) = 4 & X_2(55) = 5 & X_2(56) = 6 \\
X_2(61) = 1 & X_2(62) = 2 & X_2(63) = 3 & X_2(64) = 4 & X_2(65) = 5 & X_2(66) = 6
\end{array}$$

b) Compute the conditional expectation $E[X_1 X_2 | X_2 = x_2]$.

$$\begin{aligned}E[X_1 + X_2 | X_2 = x_2] &= E[X_1 | X_2 = x_2] + E[X_2 | X_2 = x_2]\\ &= E[X_1] + x_2\\ &= 3.5 + x_2\end{aligned}$$

$E[X_1 | X_2 = x_2] = E[X_1]$ holds due to the independence between the two variables (they refer to two different rolls of the fair dice)

c) Compute the conditional expectation $E[X_1 X_2 | X_2 = x_2]$.

$$\begin{aligned}E[X_1 X_2 | X_2 = x_2] &= x_2 E[X_1 | X_2 = x_2]\\ &= x_2 E[X_1]\\ &= 3.5 \cdot x_2\end{aligned}$$

$E[X_1 | X_2 = x_2] = E[X_1]$ holds due to the independence between the two variables (they refer to two different rolls of the fair dice)

d) Compute the conditional variance $\text{Var}(X_1^2 X_2 | X_2 = x_2)$.

By definition of variance, it can computed as:

$$\begin{aligned}
\text{Var}(X_1^2 X_2 | X_2 = x_2) &= E[(X_1^2 X_2 - E[X_1^2 X_2 | X_2 = x_2])^2 | X_2 = x_2] \\
&= E\left[X_1^4 X_2^2 - 2X_1^2 X_2 E[X_1^2 X_2 | X_2 = x_2] + E[X_1^2 X_2 | X_2 = x_2]^2 | X_2 = x_2\right] \\
&= x_2^2 E[X_1^4 | X_2 = x_2] - 2E[X_1^2 X_2 x_2 E[X_1^2 | X_2 = x_2] | X_2 = x_2] \\
&\quad + E[x_2^2 E[X_1^2 | X_2 = x_2]^2 | X_2 = x_2] \\
&= x_2^2 E[X_1^4] - 2E[X_1^2 X_2 x_2 E[X_1^2]] | X_2 = x_2] + E[x_2^2 E[X_1^2]^2 | X_2 = x_2] \\
&= x_2^2 E[X_1^4] - 2x_2^2 E[X_1^2 E[X_1^2]] + x_2^2 E[E[X_1^2]^2] \\
&= x_2^2 E[X_1^4] - 2x_2^2 E[X_1^2] E[X_1^2] + x_2^2 E[X_1^2]^2
\end{aligned}$$

Substituting the expected values for X_1^4 and X_1^2

$$E_[X_1^2] = \frac{91}{6} \quad E[X_1^4] = \frac{2275}{6}$$

$$\begin{aligned}
\text{Var}(X_1^2 X_2 | X_2 &= x_2^2 E[X_1^4] - 2x_2^2 E[X_1^2] E[X_1^2] + x_2^2 E[X_1^2] \\
&= x_2^2 \left(\frac{2275}{6} - 2\frac{91}{6}\frac{91}{6} + \left(\frac{91}{6}\right)^2 \right) \\
&= x_2^2 \frac{2275 \cdot 6 - 91^2}{36} \\
&= x_2^2 \frac{5369}{36} \\
&\approx 149.14 \cdot x_2^2
\end{aligned}$$

Problem 2.3

Let (X, Y) be random variables describing the two coordinates of points that are uniformly distributed in a triangle that is bound $-1 \leq x \leq 1, y \geq 0$ and the two lines $y = 1 + x$ and $y = 1 - x$.
Firstly, we find the joint probability density function PDF by the property:

$$\int_x \int_y p(x,y) \cdot dy \cdot dx = 1$$

We set the $p(x, y)$ to be a constant c.

$$\int_{-1}^{0} \int_{0}^{1+x} c \cdot dy \cdot dx + \int_{0}^{1} \int_{0}^{1-x} c \cdot dy \cdot dx = 1$$

$$\int_{-1}^{0} c \cdot (1+x) \cdot dx + \int_{0}^{1} c \cdot (1-x) \cdot dx = 1$$

$$c - \frac{c}{2} + c - \frac{c}{2} = 1$$

$$c = 1 => p(x,y) = 1$$

a) Find $P(X \geq -0.5)$.

$$P(X \geq -0.5) = 1 - P(X < -0.5)$$
$$P(X \geq -0.5) = 1 - \int_{-1}^{-0.5} \int_{0}^{1+x} p(x,y) \cdot dy \cdot dx$$
$$P(X \geq -0.5) = 1 - \int_{-1}^{-0.5} \int_{0}^{1+x} 1 \cdot dy \cdot dx$$
$$P(X \geq -0.5) = 1 - \int_{-1}^{-0.5} (1+x) \cdot dx$$
$$P(X \geq -0.5) = 1 - (\frac{x^2}{2} + x)\Big|_{-1}^{-0.5}$$
$$P(X \geq -0.5) = 1 - (\frac{1}{8} - \frac{1}{2}) - (\frac{1}{2} - 1) = \frac{7}{8}$$

b) Find $P(Y \geq 0.5)$.
For each value of Y, the value of X goes from $y-1$ to $1-y$

$$P(Y \geq 0.5) = \int_{0.5}^{1} \int_{y-1}^{1-y} p(x,y) \cdot dx \cdot dy$$
$$P(Y \geq 0.5) = \int_{0.5}^{1} \int_{y-1}^{1-y} 1 \cdot dx \cdot dy$$
$$P(Y \geq 0.5) = \int_{0.5}^{1} (2 - 2 \cdot y) \cdot dy$$
$$P(Y \geq 0.5) = 2 \cdot y - y^2\Big|_{0.5}^{1}$$
$$P(Y \geq 0.5) = (2-1) - (1 - \frac{1}{4}) = \frac{1}{4}$$

c) Find the marginal densities and expectations of X and Y.

$$p_X(x) = \int_Y p(x,y) \cdot dy$$

Since the values of Y varies from 0 to $1 - |x|$:

$$p_X(x) = \int_{0}^{1-|x|} p(x,y) \cdot dy$$
$$p_X(x) = 1 - |x|$$

The marginal density of Y:

$$p_Y(y) = \int_X p(x,y) \cdot dx$$

The values of X varies from $y-1$ to $1-y$, so the marginal density ca be written as:

$$p_Y(y) = \int_{y-1}^{1-y} 1 \cdot dx = 2 - 2 \cdot y$$

The expectation value of X:

$$E[X] = \int_X \int_Y x \cdot p(x,y) dy \cdot dx$$

$$E[X] = \int_{-1}^{1}\int_{0}^{1-|x|} xdy \cdot dx$$

$$E[X] = \int_{-1}^{1} x \cdot (1-|x|) \cdot dx$$

Sine we have absolute value of x, we can split the integral into 2 intervals:

$$E[X] = \int_{-1}^{0} x \cdot (1+x) \cdot dx + \int_{0}^{1} x \cdot (1-x) \cdot dx$$

$$E[X] = (\frac{x^3}{3} + \frac{x^2}{2})\Big|_{-1}^{0} + (-\frac{x^3}{3} + \frac{x^2}{2})\Big|_{0}^{1} = -(\frac{-1}{3} + \frac{1}{2}) + (\frac{-1}{3} + \frac{1}{2}) = 0$$

The expectation value of Y:

$$E[Y] = \int_{Y}\int_{X} y \cdot p(x,y)dx \cdot dy$$

$$E[Y] = \int_{0}^{1}\int_{y-1}^{1-y} y \cdot dx \cdot dy$$

$$E[Y] = \int_{0}^{1} 2 \cdot y - 2 \cdot y^2 \cdot dy$$

$$E[Y] = (y^2 - \frac{2 \cdot y^3}{3})\Big|_{0}^{1} = (1 - \frac{2}{3}) = \frac{1}{3}$$

Programming Problem 2.1

a) Generate a random set of two-dimensional points and plot them as a scatter plot.

```
size = 100
points = np.random.rand(size, 2)

fig = plt.figure()
ax = fig.add_subplot(1, 1, 1)

ax.scatter(points[:, 0], points[:, 1])
plt.show()
```

b) Plot the density of the 2D Gaussian distribution $\mathcal{N}(\mu, \mathbf{\Sigma})$ with $\mu = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $\mathbf{\Sigma} = \begin{pmatrix} 1 & 0.25 \\ 0.25 & 1 \end{pmatrix}$.

```
def two_dimensional_gaussian(mean, cov, pos):
    cov_det = np.linalg.det(cov)
    cov_inv = np.linalg.inv(cov)

    N = np.sqrt((2 * np.pi) ** 2 * cov_det)

    fac = np.einsum('...k,kl,...l->...', pos - mean, cov_inv, pos - mean)

    return np.exp(-fac / 2) / N

sample = 100

X = np.linspace(-3, 3, sample)
Y = np.linspace(-2, 4, sample)
X, Y = np.meshgrid(X, Y)
```

```

pos = np.empty(X.shape + (2,))
pos[:, :, 0] = X
pos[:, :, 1] = Y

mean = np.array([0, 1])
cov = np.array([[1, 0.25], [0.25, 1]])

dst = two_dimensional_gaussian(mean, cov, pos)

fig = plt.figure(figsize=(15, 15))
ax = fig.add_subplot(projection='3d')

ax.plot_surface(X, Y, dst, linewidth=1, cmap=cm.jet)

ax.set_zlim(-0.05, 0.2)
ax.set_zticks(np.linspace(0, 0.2, 5))
ax.view_init(20, -20)

plt.show()
```