Exercise 2

Berkay Günes & Niklas Knöll

1 Coin-tossing game

(a) The mean is given by:

$$E[X] = 1 * \frac{1}{4} + 2 * \frac{1}{4} + (-1) * \frac{1}{2} = 0.25€$$

The Variance is given by:

$$Var[X] = (1 - 0.25)^2 * \frac{1}{4} + (2 - 0.25)^2 * \frac{1}{4} + (-1 - 0.25)^2 * \frac{1}{2} = 1.69$$

(b) The fair price is $0.25€$, because the expected value of the game is $0.25€$. With that cost the game is fair, because the new expected value will be $0€$.

2 Expectations and variances

(a)

$$
\begin{aligned}
Var[aX+b] &= E[(aX+b)^2] - E[aX+b]^2 \\
&= E[a^2X^2 + 2abX + b^2] - (aE[X]+b)^2 \\
&= a^2E[X^2] + 2abE[X] + b^2 - a^2E[X]^2 - 2abE[X] - b^2 \\
&= a^2E[X^2] - a^2E[X]^2 \\
&= a^2(E[X^2] - E[X]^2) \\
&= a^2Var[X]
\end{aligned}
$$

(b)

$$
\begin{aligned}
E[X] &= \sum_x xP(X=x) \\
&= \sum_x x \sum_y P(X=x|Y=y)P(Y=y) \\
&= \sum_x \sum_y xP(X=x|Y=y)P(Y=y) \\
&= \sum_y \sum_x xP(X=x|Y=y)P(Y=y) \\
&= \sum_y E[X|Y=y]P(Y=y) \\
&= E_Y[E_X[X|Y]]
\end{aligned}
$$

(c)

$$
\begin{aligned}
Var[X] &= E[X^2] - E[X]^2 \\
&= E[E[X^2|Y]] - E[E[X|Y]]^2 \\
&= E[E(X^2|Y) - E(X|Y)^2] + E[E(X|Y)^2] - E[E(X|Y)]^2 \\
&= E[Var(X|Y)] + Var(E(X|Y))
\end{aligned}
$$

3 Covariance and correlation

We will look at $\mathrm{Var}\left(\frac{X}{\sigma_Y} \pm \frac{Y}{\sigma_Y}\right)$. This is always non-negative, because a variance has to be non-negative. From the lecture we already know $\mathrm{Var}(aX) = a^2\mathrm{Var}(X)$ and $\mathrm{Cov}(aX, bY) = ab\mathrm{Cov}(X, Y)$ and
$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X, Y)$. With this we get:

$$
\begin{aligned}
&\mathrm{Var}\left(\frac{X}{\sigma_Y} \pm \frac{Y}{\sigma_Y}\right) \\
=&\mathrm{Var}\left(\frac{X}{\sigma_X}\right) + \mathrm{Var}\left(\frac{Y}{\sigma_Y}\right) \pm 2\mathrm{Cov}\left(\frac{X}{\sigma_X}, \frac{Y}{\sigma_Y}\right) \\
=&\frac{\mathrm{Var}(X)}{\sigma_X^2} + \frac{\mathrm{Var}(Y)}{\sigma_Y^2} \pm 2\frac{\mathrm{Cov}(X, Y)}{\sigma_X\sigma_Y} \\
=&2 \pm 2\rho(X, Y) \geq 0
\end{aligned}
$$

This can be easily rearranged as:

$$-1 \leq \rho(X, Y) \leq +1$$

4 Correlation Between CO_2 levels and Earth's surface temperature

In [ ]:

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
```

In [ ]:

```
time=pd.read_csv('monthly_in_situ_co2_mlo.csv',skiprows=59,usecols=[3],de
co2=pd.read_csv('monthly_in_situ_co2_mlo.csv',skiprows=59,usecols=[4],del
co2=co2.to_numpy()
time=time.to_numpy()

co2_annual=[]
k=0
count=0
for idx,i in enumerate(time):
    if co2[idx]>0:
        k+=co2[idx]
        count+=1
    if (idx%12==0) & (idx!=0):
        co2_annual.append(k/count)
        k=0
        count=0


yr=np.linspace(1958,2022,65)
temp=pd.read_csv('temp.csv',skiprows=79,usecols=[13],delimiter=",")
yr_t=pd.read_csv('temp.csv',skiprows=79,usecols=[0],delimiter=",")
yr_t=yr_t.to_numpy()
temp=temp.to_numpy()
temp=temp[:-1]
yr_t=yr_t[:-1]
for idx,i in enumerate(temp):
    temp[idx]=float(i)
```

In [ ]:

```
#variances
def var(X):
    return np.sum((np.mean(X)-X)**2)/(len(X))
temp_var=var(temp)
co2_var=var(co2_annual)
#covariance
def cov(X,Y):
    return np.mean(X*Y)-np.mean(X)*np.mean(Y)
co2_annual=np.asarray(co2_annual)
corr=cov(co2_annual,temp)/(co2_var**0.5*temp_var**0.5)
print(corr)
```

```
0.961713608600864
```

The correlation coefficient between annual CO_2 levels and temperature deviation is

$$\rho = \frac{\mathrm{COV}(T, CO_2)}{\sigma_T \sigma_{CO_2}} = 0.96$$

In [ ]:

```
##temp vs co2 levels
plt.scatter(temp,co2_annual,s=4)
plt.xlabel("temperature deviation")
plt.ylabel(r"CO$_2$ [ppm]")
plt.title("Visualization of the correlation")
# plt.legend()
```

Out[ ]:

```
Text(0.5, 1.0, 'Visualization of the correlation')
```