# Advanced Statistics, homework 2

## 3.8 Exercises

1. Suppose we play a game where we start with $c$ dollars. On each play of the game you either double or halve your money, with equal probability. What is your expected fortune after $n$ trials?

**Solution:**

$$X_0 = c$$

We have only 2 cases, either it gets double or takes half. So, probability for both cases = $\frac{1}{2}$.

$$X_i = 2c \ (if\ it\ wins),\ X_i = \frac{1}{2}\ c\ (if\ it\ looses).$$

Expectation for X, when it's discrete

$$\mathbf{E(X)} = \sum_x \boldsymbol{x} * \boldsymbol{P}(\boldsymbol{X})$$

So, the expectation:

$$E(X) = \frac{1}{2} * 2c + \frac{1}{2} * \frac{1}{2}c = 1c + \frac{1}{4}c = \frac{5}{4}\ c$$

After n trials, by induction:

$$E(X_n) = (\frac{5}{4})^n c$$

**So, the answer:**

$$E(X_n) = (\frac{5}{4})^n c$$

4. A particle starts at the origin of the real line and moves along the line in jumps of one unit. For each jump the probability is $p$ that the particle will jump one unit to the left and the probability is $1-p$ that the particle will jump one unit to the right. Let $X_n$ be the position of the particle after $n$ units. Find $\mathbb{E}(X_n)$ and $\mathbb{V}(X_n)$. (This is known as a **random walk**.)

**Solution:**

Because it moves along the line:

$$X_i = 1\ (if\ it\ jumps\ right), \qquad X_i = -1\ (if\ it\ jumps\ left)$$

$$P(X_i = 1\ ) = 1 - p, \qquad P(X_i = -1\ ) = p$$

As in the previous task:

$$\mathbf{E(X)} = \sum\nolimits_x \boldsymbol{x * P(X)}$$

$$E(X_i) = 1 * (1-p) + (-1) * p = 1 - p - p = 1 - 2p$$

$$E(X_n) = n * E(X_i) = \boldsymbol{n(1-2p)}$$

$$\boldsymbol{Var(X_i) = E\left(X_i^2\ \right) - \left(E(X_i)\right)^2}$$

$$E\left(X_i^2\right) = 1^2 * (1-p) + (-1)^2 * p = 1 - p + p = 1$$

$$\left(E(X_i)\right)^2 = (1-2p)^2 = 1 - 4p + 4p^2$$

$$Var(X_i) = E(X_i^2\ ) - \left(E(X_i)\right)^2 = 1 - 1 + 4p - 4p^2 = 4p(1-p)$$

$$Var(X_n) = n * Var(X_i) = \boldsymbol{4np(1-p)}$$

**So, the answer:**

$$E(X_n) = n * E(X_i) = n(1-2p)$$

$$Var(X_n) = n * Var(X_i) = 4np(1-p)$$

5. A fair coin is tossed until a head is obtained. What is the expected number of tosses that will be required?

**Solution:**

$$X \sim Geom(p)$$

$$P(X = H) = \frac{1}{2} = 0.5\ , because\ we\ have\ either\ Tails\ or\ Heads.$$

As we know, probability in Geometric Distribution:

$$P(X = k) = p(1-p)^{k-1}$$

So, the expectation, also as we know:

$$E(X) = \sum_{k=1}^{\infty} k * P(X = k) = \sum_{k=1}^{\infty} k * p * (1-p)^{k-1} = \frac{1}{p} = \frac{1}{0.5} = 2$$

**So, the answer:**

$$E(X) = 2$$

## EX: 11 (Computer experiment)

$$Y_i = 1, \qquad Y_i = -1$$

$$P(Y_i = 1) = P(Y_i = -1) = \frac{1}{2}$$

$$E(Y_i) = \frac{1}{2} * 1 + \frac{1}{2} * (-1) = 0$$

$$E\left(Y_i^2\right) = 1^2 * \frac{1}{2} + (-1)^2 * \frac{1}{2} = 1$$

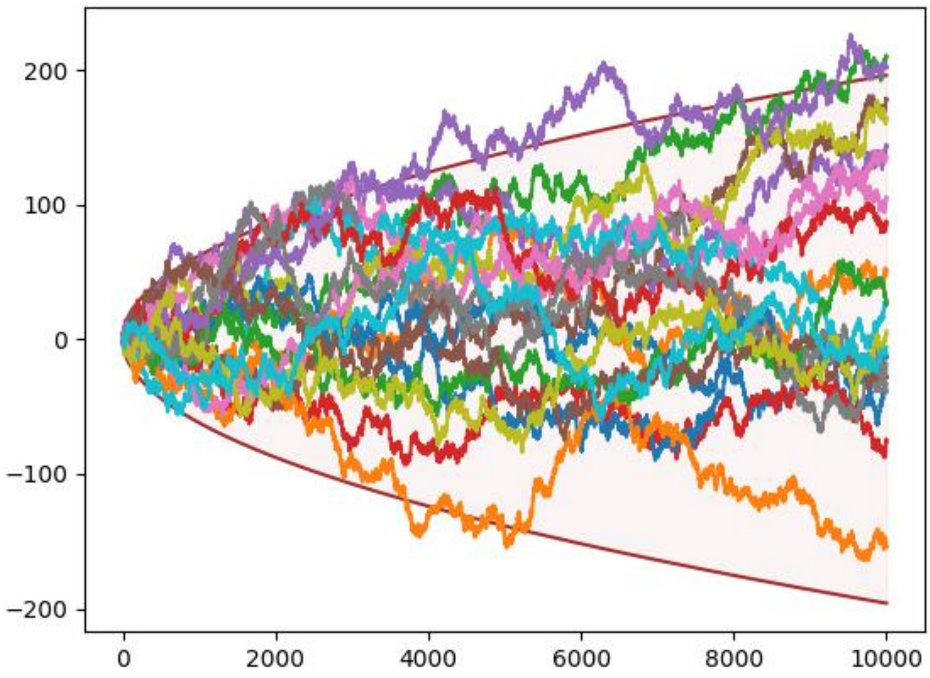
$$Var(Y_i) = E\left(Y_i^2\right) - \left(E(Y_i)\right)^2 = 1 - 0 = 1$$

$$E(X_n) = \sum_{i=1}^{n} E(Y_i) = 0$$

$$Var(X_n) = \sum_{i=1}^{n} Var(Y_i) = n$$

```
import numpy as np
from scipy.stats import norm, bernoulli
import matplotlib.pyplot as plt

n = 10000
b = 20

Y = 2 * bernoulli.rvs(p = 1/2, loc = 0, size = (b, n)) - 1
X = np.cumsum(Y, axis = 1)

# print(Y)
# print(X)

arr = np.arange(1, n + 1)
# print(arr)
Z = norm.ppf(0.975)
```

```
# print(Z)
plt.plot(arr, Z * np.sqrt(arr), color='brown')
plt.plot(arr, -Z * np.sqrt(arr), color='brown')
plt.fill_between(arr, Z * np.sqrt(arr), -Z * np.sqrt(arr), color='brown',
alpha=0.05)

for b in range(b):
    plt.plot(arr, X[b])

plt.show()
```




2. Let $X \sim \text{Poisson}(\lambda)$. Use Chebyshev's inequality to show that $\mathbb{P}(X \geq 2\lambda) \leq 1/\lambda$.

**Solution:**

$$E(X) = \lambda, \;\; Var(X) = \lambda$$

By Chebyshev's:

$$P(|X - E(X)| \geq t) \leq \frac{E(X)}{t^2}$$

$$P(|X-\lambda| \geq \lambda) \leq \frac{\lambda}{\lambda^2} = \frac{1}{\lambda}$$

6. Suppose that the height of men has mean 68 inches and standard deviation 2.6 inches. We draw 100 men at random. Find (approximately) the probability that the average height of men in our sample will be at least 68 inches.

**Given:**

$E(X) = 68,\ \sigma = 2.6 \quad n = 100$

**Solution:**

In addition to $Z_n \rightsquigarrow N(0,1)$, there are several forms of notation to denote the fact that the distribution of $Z_n$ is converging to a Normal. They all mean the same thing. Here they are:

$$\begin{aligned} Z_n &\approx N(0,1) \\ \overline{X}_n &\approx N\left(\mu, \frac{\sigma^2}{n}\right) \\ \overline{X}_n - \mu &\approx N\left(0, \frac{\sigma^2}{n}\right) \\ \sqrt{n}(\overline{X}_n - \mu) &\approx N\left(0, \sigma^2\right) \\ \frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma} &\approx N(0,1). \end{aligned}$$

Getting from this we have:

$$P(X \geq 68) = 1 - P(x < 68) = 1 - P\left(\frac{\sqrt{100}(68-68)}{2.6} < 0\right) = 1 - \Phi(0)$$
$$= 1 - 0.5 = 0.5$$

**So, the answer:**

$$P(X \geq 68) = 0.5$$

8. Suppose we have a computer program consisting of $n = 100$ pages of code. Let $X_i$ be the number of errors on the $i^{th}$ page of code. Suppose that the $X_i's$ are Poisson with mean 1 and that they are independent. Let $Y = \sum_{i=1}^{n} X_i$ be the total number of errors. Use the central limit theorem to approximate $\mathbb{P}(Y < 90)$.

**Solution:**

$$X_i \sim Posson(1)$$

$$E(X) = Var(X) = 1$$

$$E(X_n) = E(X_1 + \cdots + E(X_n) = 100 \qquad Var(X) = 100$$

$$Y \sim Normal(100, 100)$$

$$P(Y < 90) = P\left(Z < \frac{90 - 100}{\sqrt{100}}\right) = P(Z < -1) = 0.159$$

**So, the answer:**

$$P(Y < 90) = 0.159$$