

Practice sheet 2

Exercise 1

A random variable X is with Binomial distribution with parameters $n \in \mathbb{N}^*$ (size) and $p \in [0, 1]$ (probability) if the outcomes of X are in the set $\{0, 1, 2, \dots, n\}$ and for any $k \in \{0, 1, 2, \dots, n\}$:

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k} = \frac{n!}{k!(n-k)!} p^k (1 - p)^{n-k}$$

For example, we flip a coin $n = 10$ times, and X can be the number of Head outcomes. If p is the probability of observing Head, then $1 - p$ is observing a Tail, X is a random variable with Binomial distribution with parameter size $n = 10$, and probability p .

`stats.binom()`, from `scipy` library, is the Python command that can define a Binomial random variable. It has two arguments: the size n and the probability p . For example, `stats.binom(10, .4)` means that we have defined a Binomial random variable with $n = 10$ and $p = .4$.

Let's consider the following Python code

```
[1]: import numpy as np
```

```
[2]: from scipy import stats
```

```
[3]: X= stats.binom(15,.6)
```

```
[4]: (.4)**15
```

```
[4]: 1.073741824000001e-06
```

```
[5]: X.pmf(2)
```

```
[5]: 0.00025367150592000057
```

```
[6]: X.pmf(6)
```

```
[6]: 0.06121410527231985
```

```
[7]: X.pmf(1)
```

```
[7]: 2.415919104000003e-05
```

```
[8]: X.pmf(3)
```

```
[8]: 0.0016488647884800034
```

```
[9]: X.cdf(5)
```

```
[9]: 0.03383330288435202
```

1. Compute the following probabilities by using the previous code:

1. $\mathbb{P}(X \leq 2)$?
2. $\mathbb{P}(X > 6)$?
3. $\mathbb{P}(X = 3)$?

2. A multiple-choice test has 15 questions, each of which has five choices. An unprepared student taking the test answers each of the questions completely randomly by choosing an arbitrary answer from the five provided. Suppose X denotes the number of answers that the student gets right. Compute using a Python code the following probabilities:

3. All the students' answers are wrong.
4. All the students' answers are correct.
5. At least the student answers correctly to 7 questions.
6. We consider two grading systems:
 - System A, the student gets +2 for each correct answer and 0 for each wrong answer.
 - System B, the student receives +2 for each correct answer and -1 for each wrong answer.

Compare the probability of getting at least 15 between the grading systems? What's your conclusion?

Exercise 2

Let's consider the random variable X that represents the rank of the occurrence of the first success. If the trials are independent and the probability of success on each trial is p , the probability distribution of X is called then the **geometric** distribution.

the probability that the k th trial (out of k trials) is the first success is

$$\mathbb{P}(X = k) = (1 - p)^{k-1}p, \quad \text{for all } k \in \mathbb{N}^*$$

It's known that

$$\mathbb{E}(X) = \frac{1}{p}, \quad \mathbb{P}(X \geq k) = (1 - p)^{k-1}$$

1. Complete the following Python code:

```
[10]: import numpy as np
      from scipy import stats
```

```
[11]: X= stats.geom(.....)
```

```
[12]: X.pmf(3)
```

```
[12]: .....
```

```
[13]: X.cdf(4)
```

```
[13]: .....
```

```
[14]: X.pmf(2)
```

```
[14]: 0.24
```

```
[15]: X.mean()
```

```
[15]: 1.6666666666666667
```

```
[16]: X.cdf(0)
```

```
[16]: .....
```

```
[17]: X.pmf(-1)
```

```
[17]: .....
```

2. You play a game of chance that you can either win or lose (there are no other possibilities) until you lose. Your probability of losing is $p=0.57$.

1. What is the probability that it takes five games until you lose?
2. What is the probability that it takes at least three games until you lose?

Exercise 3

Complete the following the Python code

```
[18]: import math
      from scipy import stats
```

```
[19]: X= stats.norm(1.4,2)
```

```
[20]: X.mean()
```

```
[20]: .....
```

```
[21]: X.var()
```

[21]:

```
[22]: X.cdf(.....)
```

[22]: 0.5

```
[23]: X.cdf(0)
```

[23]:

```
[24]: 1/math.sqrt(2*math.pi*4)
```

[24]: 0.19947114020071635

```
[25]: X.pdf(1.4)
```

[25]:

Exercise 4

Let's consider the following Python

```
[26]: import math
      from scipy import stats
```

```
[27]: X = stats.norm(3,2)
```

```
[28]: X.cdf(0)
```

[28]: 0.06680720126885807

```
[29]: X.cdf(4)
```

[29]: 0.6914624612740131

```
[30]: X.cdf(-2)
```

[30]: 0.006209665325776132

1. What's the probability distribution of the random variables introduced in the code above?
2. What's $\mathbb{E}(X)$ and $\mathbb{E}(X^2)$?
3. Compute the probability $\mathbb{P}(0 \leq X \leq 4)$.
4. Compute the probability $\mathbb{P}(|X| < 2)$
5. Compute the probability $\mathbb{P}(X < -4)$