Exercises # 2: Random variables

Léo Neufcourt

September 16, 2019

Please do all problems. Due in class on September 13, 2019. Python codes must be sent by email before September 13, 2019, 12:40pm.

Exercise 1. *A player shoots on a target with radius 10 inches, composed of concentric rings delimited by circles with radii 1, 2, ..., 10 cm, and respectively numbered from 10 to 1. We suppose that the player reaches the the target all the time, and that the probability to hit ring number k is proportional to the area of this ring. Let X be the random variable which associates to each throw the number of the target.*

(1) What is the law (i.e. here p.m.f.) of X?

(2) The player wins $\$k$ *if he hits ring number* k *for* k *between* 10 *and* 6*, but loses* $\$2$ *if he reaches the peripheral rings numbered* 1 *to* 5*. Is this game favorable to the player?*

Hint: *(1) Compute the area of each ring. (2) Introduce the random variable equal to the gain of the player and compute its expectation...*

Solution 1.

(1) *We denote* A_k *the area of the ring* k*, and* A *the total area. By the equiprobability hypotheses we have* $\mathbb{P}(X = k) = \frac{A_k}{A}$*, for* $k = 1, 2, \ldots 10$*. For ring* k *the external radius is* $11 - k$ *and the interior radius is* $10 - k$ *(draw a picture). We thus have, in square inches,*

$$A_k = \pi((11 - k)^2 - (10 - k)^2) = \pi(21 - 2k)$$

so that $\mathbb{P}(X = k) = \frac{21-2k}{100}$.

(2) *Denote* Y *the random variable equal to the (signed) gain of the player.* Y *takes values in* $\{-2, 6, 7, 8, 9, 10\}$, *hence the event* $\{Y = -2\}$ *is equal to the event* $\{X\ leq5\}$, *hence*

$$\mathbb{P}(Y = -2) = \mathbb{P}(X = 1) + \mathbb{P}(X = 2) + \mathbb{P}(X = 3) + \mathbb{P}(X = 4) + \mathbb{P}(X = 5) = \frac{75}{100}$$

by the result of the previous question. The expectation of Y *is*

$$\mathbb{E}[Y] = -2 \times \frac{75}{100} + \frac{6 \times 9 + 7 \times 7 + 8 \times 5 + 9 \times 3 + 10 \times 1}{100} = \frac{3}{10},$$

which is positive: the game is favorable to the player, since he will win in average \$0.3 *per game.*

Exercise 2. *A restaurant has seventy guests every evening. The chef knows that, in average, two customers every five order a* crème brûlée. *He thinks that if he prepares thirty crème brûlées, then he will be able to satisfy all customers with seventy percent probability.*

(1) Is the chef right?

(2) How many crème brûlées (at least) does the chef need to bake to satisfy all customers with ninety percent probability?

Hint: *Introduce the random variable equal to the number of crème brûlées ordered on a given evening.*

Solution 2.

(1) *Let* X *be the random variables counting the number of crème brûlées ordered on a given evening.* X *counts the number of successes of a repetition of* $n = 70$ *independent Bernouilli trials, each of which has a success probability of* $p = 0.4$, *hence has a binomial distribution with parameters* n *and* p. *All requests will be satisfied if (and only if)* $X\ leq30$, *and one easily obtains from hand calculations or a calculator* $\mathbb{P}(X \leq 30) \approx 0.73 \geq 0.7$: *the chef is right.*

(2) *We are looking for the smallest integer* k *such that* $\mathbb{P}(X\ leqk) \geq 0.9$. *hand calculations or a calculator provide* $\mathbb{P}(X \geq 32) \approx 0.86$, *and* $\mathbb{P}(X \geq 33) \approx 0.91$: *it is sufficient to bake* 33 *crème brûlées to satisfy all customers with* 90% *probability. He should at least bake one more for his crew, though.*

Exercise 3. *Let X and Y be two Gaussian random variables (with values in $\mathbb{R}$, with respective means μ and ν, respective standard deviations σ and τ, and correlation ρ. What is $\mathbb{E}[X|Y=y], y \in \mathbb{R}$?*

Solution 3. *See textbooks for this classical result. Alternatively, let us decompose $X =: \tilde{X} + \rho\frac{\sigma}{\tau}Y$, i.e. define $\tilde{X} := X - \rho\frac{\sigma}{\tau}Y$. The key is to see that $\tilde{X}$ is independent from Y, while the other term is a (orthogonal) projection of X on Y. First, $\tilde{X}$ is a Gaussian as the sum of two Gaussian random variables. Second,*

$$\mathbb{E}[Y\tilde{X}] = E[YX] - \rho\frac{\sigma}{\tau}E[Y^2] = \rho\sigma\tau - \rho\frac{\sigma}{\tau}\tau^2 = 0,$$

*so that $\tilde{X}$ and Y are two uncorrelated **Gaussian** random variables hence independent. It follows that*

$$\mathbb{E}[X|Y=y] = \mathbb{E}[\tilde{X}|Y=y] + \rho\frac{\sigma}{\tau}\mathbb{E}[Y|Y=y] = \mathbb{E}[\tilde{X}] + \rho\frac{\sigma}{\tau}y = \rho\frac{\sigma}{\tau}y.$$

Exercise 4 (Python)**.** *Write a function* `sample_discrete(n, q)` *that takes as input an integer* `n` *and a* pandas *series* `q`*, with length* `d = len(q)`*, index* `x = list(q.index)` *and values* `v = q.values`*, and returns a* numpy *array containing* `n` *independent samples from the discrete distribution which gives probability* `v[i]` *to the outcome* `x[i]`*,* $\mathtt{i} = 1, 2, \dots, d$*.*

Solution 4.

```
import numpy as np
import pandas as pd
import math

def sample_discrete(n, q):
    elements = q.index.tolist()
    probabilities = q.values.tolist()
    x = np.random.choice(elements, n, p = probabilities)
    return x

### Example
sample_discrete(
    15, pd.Series(
        [.3, .4, .2, .1],
        index = ['a', 3, 'potato', math.exp(1)]
    )
)
```