Lisbon School of Economics & Management
Universidade de Lisboa

STATISTICS I

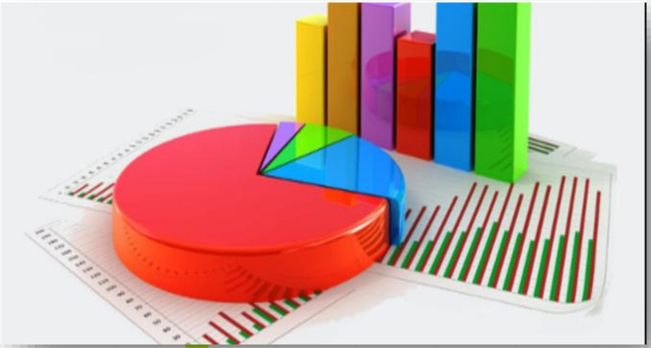
Bachelor's degrees in Economics and Finance

2nd Year/2nd Semester

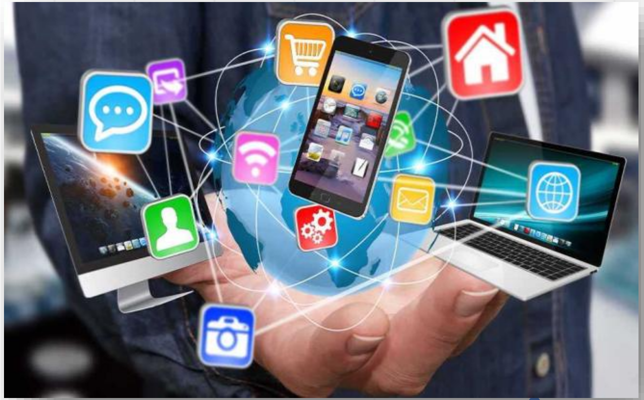
2025/2026

Pratical Class Nº 10

Professor: Elisabete Fernandes
E-mail: efernandes@iseg.ulisboa.pt



https://doity.com.br/estatistica-aplicada-a-nutricao



https://basiccode.com.br/produto/informatica-basica/

1. Basic Probability Theory.
2. Univariate random variables.
3. Expected Values.
4. Multivariate random variables (random vectors).
5. Expected Values of Functions of Random Vectors.
6. Special Random Variables and Repeated Sampling Distributions.

Bibliography:

- Miller & Miller, John E. , Freund's Mathematical Statistics with applications , 8th Edition, Pearson Education, [MM], 2013
- P. Newbold, W. Carlson, B. Thorne, , Statistics for Business and Economics , 8th Edition, Pearson Education, [N], 2012

1

Discrete Distributions: Exercises

Binomial Distribution and Hypergeometric Distribution

Chapter 6

1. A box contains 10 balls, of which 3 are red, 2 are yellow, and 5 are blue. Five balls are randomly selected with replacement.

   (a) Calculate the probability that less than 2 of the selected balls are red.

   (b) Assume now that the five balls are randomly selected without replacement. Compute the previous probability.

Exercise 1 a)

10 balls $\begin{cases} 3 \text{ red} \\ 2 \text{ yellow} \\ 5 \text{ blue} \end{cases}$ ⟶ extract 5 (with replacement)

Binomial Distribution

a)

$X \equiv$ # of red balls selected (in 5 possible)

$p = P(\text{red ball}) = \frac{3}{10} = 0.3$

$X \sim Bin(5, 0.3)$ (n, p)

$f_X(x) = \binom{5}{x} 0.3^x \times 0.7^{5-x} \quad (x = 0, 1, 2, 3, 4, 5)$

$P(X < 2) = f_X(0) + f_X(1) = \binom{5}{0} 0.3^0 \times 0.7^5 + \binom{5}{1} 0.3^1 \times 0.7^4 \simeq 0.528$

Binomial Distribution

- The *Binomial random variable* is defined as the number of successes in n trials, each of which has the probability of success p.

- *The Binomial random variable:* $X =$ number of successes in n trials. One can show that the probability function is given by

$$f_X(x) = \binom{n}{x} \times p^x (1-p)^{n-x}, \; x = 0, 1, 2, \cdots, n$$

where

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

is the number of x combinations from a set with n elements and $k! = k \times (k-1) \times \ldots \times 2 \times 1.$

Adriana, 2025

Properties:

1. $E(X) = np,$
2. $Var(X) = np(1-p)$
3. $M_X(t) = [(1-p) + pe^t]^n$

Exercise 1 b)

b)

Hypergeometric Distribution ← Without replacement

$X \equiv$ # of red balls selected (in 5 possible)

$\theta = P(\text{red ball}) = \frac{3}{10} = 0.3$

$X \sim \text{Hypergeometric}(10, 3, 5)$ $\quad f_X(x) = \frac{\binom{7}{5-x}\binom{3}{x}}{\binom{10}{5}} \quad (x = 1, 2, 3, 4, 5)$

10 ↑ N balls; 3 ↑ M red balls; 5 ↑ n draws

$P(X < 2) = f_X(0) + f_X(1) = \frac{\binom{7}{5}\binom{3}{0}}{\binom{10}{5}} + \frac{\binom{7}{4}\binom{3}{1}}{\binom{10}{5}} = 0.5$

Hypergeometric Distribution

Consider a finite population of size N that contains exactly M objects with a specific feature. The hyper-geometric distribution is a discrete probability distribution that describes the probability of k successes in n draws (to get k objects with the referred feature), without replacement.

Adriana, 2025

$$X \sim Hypergeometric(N, M, n)$$

$$P(X = k) = \begin{cases} \frac{\binom{N-M}{n-k}\binom{M}{k}}{\binom{N}{n}}, & k = \max\{0, n-(N-M)\}, \cdots, \min\{n, M\} \\ 0, & \text{otherwise} \end{cases}$$

Properties: If $X \sim Hypergeometric(N, M, n)$, then

1. $E(X) = n \times \frac{M}{N}$
2. $Var(X) = n\frac{M}{N}(1 - \frac{M}{N})\frac{N-n}{N-1}$
3. There is no closed form solution for $M_X(t)$

2. Prove that if $X_1 \sim B(n_1, p)$ and $X_2 \sim B(n_2, p)$ and X_1 and X_2 are independent random variables, then $X_1 + X_2 \sim B(n_1 + n_2, p)$

(**Hint:** Recall that if $X \sim B(n, p)$, then $M_X(t) = [(1-p) + pe^t]^n$.)

Exercise 2

$X \sim B(n,p)$, then $M_X(t) = [(1-p) + pe^t]^n$

$X_1 \sim \text{Bin}(n_1, p)$ $\quad m_{X_1}(t) = [(1-p) + pe^t]^{n_1}$

$X_2 \sim \text{Bin}(n_2, p)$ $\quad m_{X_2}(t) \; [(1-p) + pe^t]^{n_2} =$

$X_1 \perp X_2$

$$m_{X_1+X_2}(t) = E(e^{t(X_1+X_2)}) = E(e^{tX_1} e^{tX_2}) =$$

$$= E(e^{tX_1}) E(e^{tX_2}) = m_{X_1}(t)\, m_{X_2}(t)$$

because $X_1 \perp X_2$

$$= [(1-p) + pe^t]^{n_1} \times [(1-p) + pe^t]^{n_2} =$$

$$= [(1-p) + pe^t]^{n_1 + n_2}$$

→ M. G. F. of a $\text{Bin}(n_1 + n_2, p)$ random variable

Therefore $X_1 + X_2 \sim \text{Bin}(n_1 + n_2, p)$, QED

Moment generating function:

$$M_X(t) = E(e^{tX})$$

3. Let $X \sim B(n, p)$ and $X^* \sim B(n, 1-p)$, show that $P\left(n - X^* = x\right) = P\left(X = x\right)$.

Exercise 3

$X \sim Bin(n, p)$ $\qquad f_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$

$X^* \sim Bin(n, 1-p)$ $\qquad f_{X^*}(x^*) = \binom{n}{x^*} (1-p)^{x^*} p^{n-x^*}$

By definition

$$P(n - X^* = x) = P(X^* = n - x) = f_{X^*}(n-x)$$

$$= \binom{n}{n-x} (1-p)^{n-x} p^{n-(n-x)} =$$

$$= \binom{n}{n-x} (1-p)^{n-x} p^{x} =$$

$$\overset{(*)}{=} \binom{n}{x} p^{x} (1-p)^{n-x} = f_X(x) = P(X = x), \quad QED$$

(*) Auxiliary calculation:

$$\binom{n}{n-x} = \frac{n!}{(n-x)![n-(n-x)]!} = \frac{n!}{(n-x)!\,x!} = \frac{n}{x!(n-x!)} = \binom{n}{x}$$

4. Just prior to jury selection for O. J. Simpson's murder trial in 1995, a poll found that about 20% of the adult population believed Simpson was innocent (after much of the physical evidence in the case had been revealed to the public). Ignore the fact that this 20% is an estimate based on a subsample from the population; for illustration, take it as the true percentage of people who thought Simpson was innocent prior to jury selection. Assume that the 12 jurors were selected randomly and independently from the population

   (a) Find the probability that the jury had at least one member who believed in Simpson's innocence.
   (b) Find the probability that the jury had at least two members who believed in Simpson's innocence.
   (c) What is the expected value and the variance of the number of jurors, in a sample of 12 jurors, who believed in Simpson's innocence?

Exercise 4 a)

$X \equiv$ # of jurors that believe in Simpson's innocence

$X \sim Bin(12, 0.2)$ $\qquad f_X(x) = \binom{12}{x} 0.2^{x} \times 0.8^{12-x} \quad (x = 0, 1, 2, \dots, 12)$

a)

$P(X \geq 1) = 1 - P(X < 1) = 1 - f_X(0) = 1 - \binom{12}{0} 0.2^{0} \times 0.8^{12} \simeq 0.9313$

Exercise 4 b)

b)

$$P(X \geq 2) = 1 - P(X < 2) = 1 - P(X \leq 1) = 1 - (f_X(0) + f_X(1)) =$$

$$= 1 - \left(\binom{12}{0} 0.2^{0} \times 0.8^{12} + \binom{12}{1} 0.2^{1} \times 0.8^{11} \right) \simeq 0.7551$$

Exercise 4 c)

c) $X \sim Bin(n, p) \Rightarrow E(X) = np$; $Var(X) = np(1-p)$

$E(X) = 12 \times 0.2 = 2.4$

$Var(X) = 12 \times 0.2 \times 0.8 = 1.92$

Discrete Distributions: Exercises

2

Geometric Distribution and Negative Binomial Distribution

Chapter 6

5. A student takes a multiple choice test with 20 questions, each with 4 choices and only one is correct.

(a) Assume that the student blindly guesses and gets one question correct. Find the probability that the reader has to read no more than 4 questions until we get the right one.

(b) Assume that the student blindly guesses and gets two question correct. Find the probability that the reader has to read 10 questions until he gets the second question that is correct.

Exercise 5 a)

$n = 20$ questions

p = Prob(correct answer in one question) = 0.25

a)

$X \equiv$ # of questions until the first correct answer

$X \sim Geo(0.25)$ $\qquad f_X(x) = 0.75^{x-1} \times 0.25 \quad (x = 1, 2, \dots)$

$$P(X \leq 4) = \sum_{x=1}^{4} f_X(x) = 0.25 \sum_{x=1}^{4} 0.75^{x-1} = 0.25\,(0.75^0 + 0.75^1 + 0.75^2 + 0.75^3) \simeq 0.683$$

Note: This answer assumes that the reader can check any question more than one time.
The geometric distribution assumes unlimited possible trials (but here we have 20)

Geometric Distribution

Adriana, 2025

- If X is the random variable that counts the number of trials until a success! and the 1^{st} success occurs on the x^{th} trial (the first $x-1$ trials are failures), then X follows a geometric distribution with a probability of success p.

$$X \sim Geo(p)$$

$$P(X = x) = (1-p)^{x-1}p, \quad x = 1, 2, 3, \cdots$$

Exercise 5 b)

b)

$Y \equiv$ # of questions until the second correct answer

$Y \sim NB(2, 0.25)$ $\quad f_Y(y) = \binom{y-1}{1} 0.25^2 \times 0.75^{y-2}, \quad y = 2, 3, \dots$

$f_Y(10) = \binom{9}{1} 0.25^2 \times 0.75^8 \simeq 0.0563$

Negative Binomial Distribution

Adriana, 2025

- If the k^{th} success is to occur on the x^{th} trial, there must be $k-1$ successes on the first $x-1$ trials, and the probability for this is

$$P(X=x)=\binom{x-1}{k-1}p^k(1-p)^{x-k},\quad x=k,k+1,k+2,\cdots$$

where X follows a negative binomial distribution with parameters k and p

$$X \sim NB(k,p)$$

Properties:

1. $E(X)=\frac{k}{p}$
2. $Var(X)=\frac{k}{p}\left(\frac{1}{p}-1\right)$
3. $M_X(t)=\left(\frac{pe^t}{1-e^t(1-p)}\right)^k$

Discrete Distributions: Exercises

3

Poisson Distribution

Chapter 6

6. Past Experience indicates that an average number of 6 customers per hour stop for petrol at a petrol station. Assuming that the number of customers that stop for petrol at a petrol station is a Poisson random variable:

   (a) What is the probability of having 3 customers stopping in any hour?

   (b) What is the probability of having 3 customers or less stopping in any hour?

   (c) What is the expected value, and standard deviation of the distribution.

Poisson Distribution

The Poisson random variable is a discrete rv that describes the number of occurrences within a randomly chosen unit of time or space. For example, within a minute, hour, day, kilometer.

The Poisson probability function is a discrete function defined for non-negative integers. If X is a Poisson random variable with parameter λ, we write $X \sim Poisson(\lambda)$. The Poisson distribution with parameter $\lambda > 0$, it is defined by

$$f_X(x) = P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}, x = 0, 1, 2, ..$$

Properties:

1. $E(X) = Var(X) = \lambda$.
2. $M_X(t) = e^{\lambda(e^t - 1)}$.
3. If $X_i \sim Poisson(\lambda_i)$ and the X_i are independent random variables, then $\sum_{i=1}^{n} X_i \sim Poisson\left(\sum_{i=1}^{n} \lambda_i\right)$.

Adriana, 2025

Exercise 6 a)

$X \equiv$ customers in one hour $\quad E(X) = 6$

$X \sim Poi(6) \qquad f_X(x) = \dfrac{e^{-6} 6^x}{x!} \quad (x = 0, 1, 2, \dots)$

a)

$$f_X(3) = \frac{e^{-6} 6^3}{3!} \simeq 0.089$$

Exercise 6 b)

b)

$$F_X(3) = \sum_{x=0}^{3} f_X(x) = \sum_{x=0}^{3} \frac{e^{-6} 6^x}{x!} = e^{-6}\left(\frac{6^0}{0!} + \frac{6^1}{1!} + \frac{6^2}{2!} + \frac{6^3}{3!}\right)$$

$$= e^{-6}\left(\frac{6^0}{0!} + \frac{6^1}{1!} + \frac{6^2}{2!} + \frac{6^3}{3!}\right) \simeq 0.1512$$

Exercise 6 b)

e)

$X \sim Poi(\lambda) \Rightarrow E(X) = Var(X) = \lambda$

$\sigma_X = \sqrt{Var(X)} = \sqrt{\lambda}$

$E(X) = Var(X) = 6$

$\sigma_X = \sqrt{6}$

7. The average number of trucks arriving on any one day at a truck depot in a certain city is known to be 2.

   (a) If we assume a Poisson distribution, what is the probability that on a given day fewer than 2 trucks will arrive at this depot?

   (b) Assume that in a set of 10 random days the number of trucks arriving at the truck depot is independent in each day. Compute the probability of having fewer than 2 trucks arriving at the truck depot each day, for 5 days.

Exercise 7 a)

$X \equiv$ # of trucks arriving at the depot in one day

$E(X) = 2$

a)

$X \sim Poi(2)$ $\qquad f_X(x) = \dfrac{e^{-2}\,2^x}{x!} \quad x = 0, 1, 2, \dots$

$P(X<2) = \underbrace{P(X \leqslant 1)}_{F_X(1)} = f_X(0) + f_X(1) =$

$= \dfrac{e^{-2}\,2^0}{0!} + \dfrac{e^{-2}\,2^1}{1!} \simeq 0.41$

Exercise 7 b)

b)

$Y \equiv$ # of days (in 10 possible days) with fewer than 2 trucks arriving at the depot

$n = 10$

$p = P(X < 2) = F_X(1) \simeq 0.41$

$Y \sim Bin(10, 0.41)$ $\quad f_Y(y) = \binom{10}{y} 0.41^y (1-0.41)^{10-y} \quad (y = 0, 1, \cdots, 10)$

$f_W(5) = \binom{10}{5} 0.41^5 \, 0.59^5 \simeq 0.209$

Thanks!

Questions?