Lisbon School of Economics & Management
Universidade de Lisboa

STATISTICS I

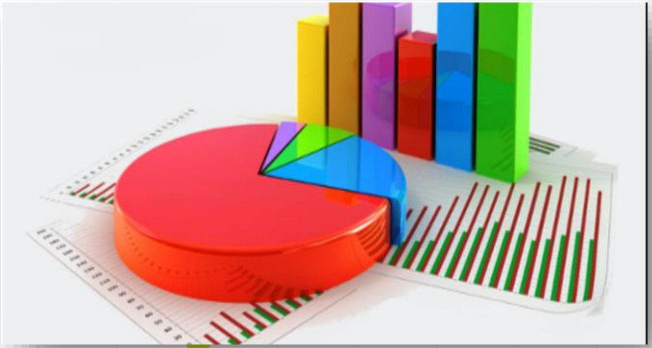
Economics / Finance/ Management

2nd Year/2nd Semester

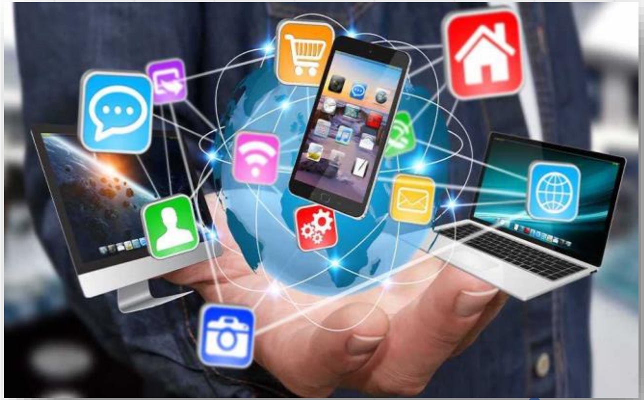
2024/2025

LESSON 8

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



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



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

Roadmap:

- Probability
- Random variable and two dimensional random variables:
  - Distribution
  - Joint distribution
  - Marginal distribution
  - Conditional distribution functions
- Expectations and parameters for a random variable and two dimensional random variables
- Discrete Distributions
- Continuous Distributions

Bibliography: Miller & Miller, John E. (2014) Freund's Mathematical Statistics with applications, 8th Edition, Pearson Education, [MM]

Two-Dimensional Discrete Random Variables: Exercises

1

Expectations and Parameters for two Dimensional Random Variables

Chapter 5

1. If X and Y have the joint probability distribution

$$f(x, y) = 1/4, \text{ for all } (x, y) \in \{(-3, -5), (-1, -1), (1, 1), (3, 5)\}.$$

Compute the $E(Z^2)$, where $Z = XY - 2X$.

Exercise 1

$f(x,y) = 1/4$, for all $(x,y) \in \underbrace{\{(-3,-5),(-1,-1),(1,1),(3,5)\}}_{D_{X,Y} = \{(x,y):\, x \in D_X \,\wedge\, y \in D_Y \,\wedge\, f_{X,Y}(x,y) > 0\}}$.

$$E[g(X,Y)] = \sum_{(x,y)\in D_{X,Y}} g(x,y)\, f_{X,Y}(x,y)$$

$$E(Z^2) = E[(\underbrace{XY-2X}_{g(X,Y)})^2] = \sum_{(x,y)\in D_{X,Y}} (xy-2x)^2 f_{X,Y}(x,y) =$$

$$= \frac{((-3)(-5)-2(-3))^2 + ((-1)(-1)-2(-1))^2 + (1\times 1 - 2\times 1)^2 + (3\times 5 - 2\times 3)^2}{4} =$$

$$= \frac{532}{4} = 133$$

3. If X and Y have the joint probability distribution

X/Y	-1	0	1
0	0	1/6	1/12
1	1/4	0	1/2

Show that

(a) $cov(X, Y) = 0$;

(b) the two random variables are not independent.

Exercise 3 a)

X and Y have the joint probability distribution:

X/Y	-1	0	1	$f_X(x)$
0	0	1/6	1/12	3/12
1	1/4	0	1/2	3/4
$f_Y(y)$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{7}{12}$	

a)

$$E(X) = \sum_{x \in D_X} x \, f_X(x) = 0 \times \frac{3}{12} + 1 \times \frac{3}{4} = \frac{3}{4}$$

$$E(Y) = \sum_{y \in D_Y} y \, f_Y(y) = -1 \times \frac{1}{4} + 0 \times \frac{1}{6} + 1 \times \frac{7}{12} = -\frac{1}{4} + \frac{7}{12} = -\frac{3}{12} + \frac{7}{12} = \frac{4}{12} = \frac{2}{6} = \frac{1}{3}$$

$$E(XY) = \sum_{(x,y) \in D_{X,Y}} \sum x y \, f_{X,Y}(x,y) = 0 \times 0 \times \frac{1}{6} + 0 \times 1 \times \frac{1}{12} + 1 \times (-1) \times \frac{1}{4} + 1 \times 1 \times \frac{1}{2} =$$

$$= -\frac{1}{4} + \frac{1}{2} = -\frac{1}{4} + \frac{2}{4} = \frac{1}{4}$$

$$\mathrm{cov}(X,Y) = E(XY) - E(X)E(Y) = \frac{1}{4} - \frac{1}{3} \times \frac{3}{4} = \frac{1}{4} - \frac{1}{4} = 0, \quad \text{QED}$$

Exercise 3 b)

b)

$$f_{X,Y}(0,-1) = 0 \neq f_X(0)\, f_X(-1) = \frac{3}{12} \times \frac{1}{4} \Rightarrow X \not\perp Y, \ QED$$

5. Let X_1,X_2, and X_3 be independent random variables with means 4, 9, and 3 and the variances 3, 7, and 5.

   (a) Find the means and the variances of $Y = 2X_1 - 3X_2 + 4X_3$ and $Z = X_1 + 2X_2 - X_3$.

   (b) Repeat (a) and (b) dropping the assumption of independence and using instead the information that $cov(X_1, X_2) = 1$, $cov(X_2, X_3) = -2$, and $cov(X_1, X_3) = -3$.

Exercise 5 a)

$\epsilon(X_1) = 4 \qquad Var(X_1) = 3$
$\epsilon(X_2) = 9 \qquad Var(X_2) = 7$
$\epsilon(X_3) = 3 \qquad Var(X_3) = 5$

X_1, X_2 and X_3 are independent from each other

Note 1: $\epsilon\left(\sum_{i=1}^{m} a_i \epsilon(X_i)\right) = \sum_{i=1}^{m} a_i \epsilon(X_i)$

Note 2: If $X_1, \ldots, X_n$ are random variables and $a_1, \ldots, a_n$ are constants and $Y = \sum_{i=1}^{n} a_i X_i$, then

$$Var(Y) = \sum_{i=1}^{n} a_i^2 Var(X_i) + 2 \underbrace{\sum_{i=1}^{n} \sum_{j=1, j<i}^{n} a_i a_j Cov(X_i, X_j)}_{=0, \text{ if } X_i, X_j \text{ are independent}}.$$

With $m = 3$ we have: $Var(Y) = \sum_{i=1}^{3} a_i^2 Var(X_i) + 2 \sum_{i=1}^{3} \sum_{j=1, j<i}^{3} a_i a_j cov(X_i, X_j)$

a) Since $m = 3$ and $X_i \perp X_j$ $(i, j = 1, 2, 3;\ i \neq j)$ we have

$$Var\left(\sum_{i=1}^{3} a_i X_i\right) = \sum_{i=1}^{3} a_i^2 Var(X_i)$$

$Y = 2X_1 - 3X_2 + 4X_3$

$\epsilon(Y) = \epsilon(2X_1 - 3X_2 + 4X_3) = 2\epsilon(X_1) - 3\epsilon(X_2) + 4\epsilon(X_3) =$

$= 2 \times 4 - 3 \times 9 + 4 \times 3 = 8 - 27 + 12 = -7$

$Var(Y) = Var(2X_1 - 3X_2 + 4X_3) = 2^2 Var(X_1) + (-3)^2 Var(X_2) + 4^2 Var(X_3) =$

$= 4 \times 3 + 9 \times 7 + 16 \times 5 = 155$

Exercise 5 a)

$$Z = X_1 + 2X_2 - X_3$$

$$E(Z) = E(X_1 + 2X_2 - X_3) = E(X_1) + 2E(X_2) - E(X_3) = 4 + 2 * 9 - 3 = 19$$

$$Var(Z) = Var(X_1 + 2X_2 - X_3) = Var(X_1) + 2^2 Var(X_2) + (-1)^2 Var(X_3) =$$

$$= 3 + 4 * 7 + 5 = 36$$

Exercise 5 b)

$\mathrm{cov}(X_1, X_2) = 1 = \mathrm{Cov}(X_2, X_1)$
$\mathrm{cov}(X_2, X_3) = -2 = \mathrm{Cov}(X_3, X_2)$
$\mathrm{cov}(X_1, X_3) = -3 = \mathrm{cov}(X_3, X_1)$

With $n = 3$ and not assuming independence we have:

$$\mathrm{Var}\left(\sum_{i=1}^{3} a_i X_i\right) = \sum_{i=1}^{3} a_i^2 \,\mathrm{Var}(X_i) + 2 \sum_{i=1}^{3} \sum_{j=1,\, j<i}^{3} a_i a_j \,\mathrm{Cov}(X_i, X_j)$$

The expected values, $E(Y) = -7$ and $E(Z) = 19$, remain unchanged.

$$\mathrm{Var}(Y) = \mathrm{Var}(2X_1 - 3X_2 + 4X_3) = 2^2\,\mathrm{Var}(X_1) + (-3)^2\,\mathrm{Var}(X_2) + 4^2\,\mathrm{Var}(X_3) + 2\big((-3 \times 2)\underbrace{\mathrm{Cov}(X_2, X_1)}_{1} + (4 \times 2)\underbrace{\mathrm{cov}(X_3, X_1)}_{-3} + (4 \times (-3))\big)\underbrace{\mathrm{Cov}(X_3, X_2)}_{-2} =$$

$$= 4 \times 3 + 9 \times 7 + 16 \times 5 + 2(-6 + 8(-3) - 12(-2)) =$$

$$= 143$$

$$\mathrm{Var}(Z) = \mathrm{Var}(X_1 + 2X_2 - X_3) =$$

$$= \mathrm{Var}(X_1) + 2^2\,\mathrm{Var}(X_2) + (-1)^2\,\mathrm{Var}(X_3) + 2\big((1 \times 2)\underbrace{\mathrm{cov}(X_2, X_1)}_{1} + (-1 \times 1)\underbrace{\mathrm{Cov}(X_3, X_1)}_{-3} + (-1 \times 2)\underbrace{\mathrm{Cov}(X_3, X_2)}_{-2}\big) =$$

$$= 3 + 4 \times 7 + 5 + 2(2 + 3 + 4) = 54$$

7. Let (X, Y) be a discrete random vector with joint probability function given by:

$$f(x,y) = \frac{x+y}{21}, \quad x = 1, 2; \; y = 1, 2, 3$$

(a) Compute the means and variances of X and Y

(b) Using $E(XY)$ analyze the independence of the two random variables and compute the correlation coefficient.

(c) Compute $E(X|Y = 1)$

Exercise 7 a)

$f_{X,Y}(x,y)$:

$x \backslash y$	1	2	3	$f_X(x)$
1	$\frac{2}{21}$	$\frac{3}{21}$	$\frac{4}{21}$	$\frac{9}{21}$
2	$\frac{3}{21}$	$\frac{4}{21}$	$\frac{5}{21}$	$\frac{12}{21}$
$f_Y(y)$	$\frac{5}{21}$	$\frac{7}{21}$	$\frac{9}{21}$	

$$E(X) = \sum_{x \in D_X} x \, f_X(x) = \frac{9}{21} + 2 \times \frac{12}{21} = \frac{33}{21} = \frac{11}{7}$$

$$E(Y) = \sum_{y \in D_Y} y \, f_Y(y) = \frac{5}{21} + 2 \times \frac{7}{21} + 3 \times \frac{9}{21} = \frac{46}{21}$$

$$E(X^2) = \sum_{x \in D_X} x^2 f_X(x) = \frac{9}{21} + 2^2 \times \frac{12}{21} = \frac{57}{21}$$

$$E(Y^2) = \sum_{y \in D_Y} y^2 f_Y(y) = \frac{5}{21} + 2^2 \times \frac{7}{21} + 3^2 \times \frac{9}{21} = \frac{114}{21}$$

$$Var(X) = E(X^2) - E(X)^2 = \frac{57}{21} - \left(\frac{33}{21}\right)^2 = \cdots = \frac{12}{49}$$

$$Var(Y) = E(Y^2) - E(Y)^2 = \frac{114}{21} - \left(\frac{46}{21}\right)^2 = \cdots = \frac{278}{441}$$

Exercise 7 b)

$$E(XY) = \sum_{(x,y)\in D_{X,Y}} xy\, f_{X,Y}(x,y) =$$

$$= (1\times1)\frac{2}{21} + (1\times2)\frac{3}{21} + (1\times3)\frac{4}{21} + (2\times1)\frac{3}{21} + (2\times2)\frac{4}{21} + (2\times3)\frac{5}{21} =$$

$$= \frac{2}{21} + \frac{6}{21} + \frac{12}{21} + \frac{6}{21} + \frac{16}{21} + \frac{30}{21} = \frac{72}{21} = \frac{24}{7} \neq E(X)E(Y) = \frac{11}{7}\times\frac{46}{21} = \frac{506}{147}$$

Therefore X and Y are not independent.

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sigma_X\, \sigma_Y} = \frac{E(XY) - E(X)E(Y)}{\sqrt{Var(X)\,Var(Y)}} =$$

$$= \frac{\frac{24}{7} - \frac{506}{147}}{\sqrt{\frac{12}{49}\times\frac{278}{441}}} \simeq -0.035 \rightarrow$$

This makes sense because, as we already know, X and Y are not independent (if they were we would have $cov(X,Y) = \rho_{X,Y} = 0$).

Exercise 7 c)

$$E(X \mid Y=1) = \sum_{x \in D_X} x \, f_{X|Y=1}(x) \, dx =$$

$$= \sum_{x=1}^{2} x \frac{f_{X,Y}(x,1)}{f_Y(1)} = 1 \frac{f_{X,Y}(1,1)}{f_Y(1)} + 2 \frac{f_{X,Y}(2,2)}{f_Y(1)} =$$

$$= \frac{\frac{2}{21}}{\frac{5}{21}} + 2 \frac{\frac{3}{21}}{\frac{5}{21}} = \frac{2}{5} + \frac{6}{5} = \frac{8}{5}$$

10. Let (X, Y) be a two dimensional random variable, such that its set of discontinuities is $D_{X,Y} = \{(0,1),(0,2),(1,0),(1,1),(1,2)\}$ and its probability function is

$$f_{X,Y}(x,y) = \begin{cases} \frac{x+y}{a}, & (x,y) \in D_{X,Y} \\ 0, & \text{otherwise} \end{cases}.$$

(a) Find a and represent $f_{X,Y}$ by filling a suitable table.

(b) Compute $E(X)$ and $E(Y)$.

(c) Calculate $Cov(2X, 3Y)$ and $Var(X+Y)$.

(d) Characterize the distribution of $E(Y|X)$.

Exercise 10 a)

$f_{X,Y}(x,y)$:

$X \backslash Y$	0	1	2	$f_X(x)$
0	0	$\frac{1}{a} = \frac{1}{9}$	$\frac{2}{a} = \frac{2}{9}$	$\frac{3}{a} = \frac{3}{9}$
1	$\frac{1}{a} = \frac{1}{9}$	$\frac{2}{a} = \frac{2}{9}$	$\frac{3}{a} = \frac{3}{9}$	$\frac{6}{a} = \frac{6}{9}$
$f_Y(y)$	$\frac{1}{a} = \frac{1}{9}$	$\frac{3}{a} = \frac{3}{9}$	$\frac{5}{a} = \frac{5}{9}$	

$$\sum_{(x,y) \in D_{X,Y}} f_{X,Y}(x,y) = 1 \Leftrightarrow \frac{1}{a} + \frac{2}{a} + \frac{1}{a} + \frac{2}{a} + \frac{3}{a} = \frac{9}{a} = 1 \Leftrightarrow a = 9$$

Exercise 10 b)

b)

$$f_X(x) = \begin{cases} \frac{3}{9} = \frac{1}{3} & (x = 0) \\ \frac{6}{9} = \frac{2}{3} & (x = 1) \\ 0 & (\text{elsewhere}) \end{cases}$$

$$f_Y(y) = \begin{cases} \frac{1}{9} & (y = 0) \\ \frac{3}{9} = \frac{1}{3} & (y = 1) \\ \frac{5}{9} & (y = 2) \\ 0 & (\text{elsewhere}) \end{cases}$$

$$E(X) = \sum_{x \in D_X} x \, f_X(x) = 0 \times \frac{1}{3} + 1 \times \frac{2}{3} = \frac{2}{3}$$

$$E(Y) = \sum_{y \in D_Y} y \, f_Y(y) = 0 \times \frac{1}{9} + 1 \times \frac{3}{9} + 2 \times \frac{5}{9} = \frac{3}{9} + \frac{10}{9} = \frac{13}{9}$$

Exercise 10 c)

$$\mathrm{cov}(2X, 3Y) = 6\,\mathrm{cov}(X,Y) = 6 \times \left(-\frac{2}{27}\right) = -\frac{12}{27} = -\frac{4}{9}$$

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{cov}(X,Y) =$$
$$= \frac{2}{9} + \frac{38}{81} - 2\left(\frac{2}{27}\right) = \cdots = \frac{44}{81} \rightarrow$$

This is the correct value of Var(X+Y). The one in the solutions is wrong.

Cálculos auxiliares:

$$E(XY) = \sum_{(x,y)\in D_{X,Y}} x y\, f_{X,Y}(x,y) = (1\times 1)\frac{2}{9} + (1\times 2)\frac{3}{9} = \frac{2}{9} + \frac{6}{9} = \frac{8}{9}$$

$$\mathrm{cov}(X,Y) = E(XY) - E(X)E(Y) = \frac{8}{9} - \frac{2}{3}\times\frac{13}{9} = \frac{8}{9} - \frac{26}{27} = \frac{24}{27} - \frac{26}{27} = -\frac{2}{27}$$

$$\mathrm{Var}(X) = E(X^2) - E(X)^2 = \frac{2}{3} - \left(\frac{2}{3}\right)^2 = \frac{2}{3} - \frac{4}{9} = \frac{6}{9} - \frac{4}{9} = \frac{2}{9}$$

$$E(X^2) = \sum_{x\in D_X} x^2 f_X(x) = 0^2\times\frac{1}{3} + 1^2\times\frac{2}{3} = \frac{2}{3}$$

$$\mathrm{Var}(Y) = E(Y^2) - E(Y)^2 = \frac{23}{9} - \left(\frac{13}{9}\right)^2 = \frac{207}{81} - \frac{169}{81} = \frac{38}{81}$$

$$E(Y^2) = \sum_{y\in D_Y} y^2 f_Y(y) = 0^2\times\frac{1}{9} + 1^2\times\frac{3}{9} + 2^2\times\frac{5}{9} = \frac{3}{9} + \frac{20}{9} = \frac{23}{9}$$

Exercise 10 d)

Let $Z = E(Y|X)$

$$E(Y|X=x) = \sum_{y \in D_Y} y\, f_{Y|X=x}(y) = \sum_{y=0}^{2} y \frac{f_{X,Y}(x,y)}{f_X(x)} \qquad (x = 0, 1)$$

Therefore:

$$E(Y|X=0) = \frac{1 \times \frac{1}{9} + 2 \times \frac{2}{9}}{\frac{3}{9}} = \frac{5}{3} \quad \text{so, } f_Z\left(\tfrac{5}{3}\right) = f_X(0) = \tfrac{1}{3}$$

$$E(Y|X=1) = \frac{0 \times \frac{1}{9} + 1 \times \frac{2}{9} + 2 \times \frac{3}{9}}{\frac{6}{9}} = \frac{8}{6} = \frac{4}{3} \quad \text{so, } f_Z\left(\tfrac{4}{3}\right) = f_X(1) = \tfrac{2}{3}$$

$$D_Z = \left\{ \tfrac{4}{3}, \tfrac{5}{3} \right\}$$

$$f_Z(z) = \begin{cases} \frac{1}{3} & \left(z = \frac{5}{3}\right) \\ \frac{2}{3} & \left(z = \frac{4}{3}\right) \\ 0 & (\text{elsewhere}) \end{cases}$$

Thanks!

Questions?