STAT 511 su 2020 Lec 14 slides

Covariance and correlation and bivariate Normal distribution

Karl B. Gregory

University of South Carolina

These slides are an instructional aid; their sole purpose is to display, during the lecture, definitions, plots, results, etc. which take too much time to write by hand on the blackboard. They are not intended to explain or expound on any material.

Covariance

The covariance between two rvs X and Y is defined as

$$\mathrm{Cov}(X, Y) = \mathbb{E}(X - \mu_X)(Y - \mu_Y) =: \sigma_{XY},$$

where $\mu_X = \mathbb{E}X$ and $\mu_Y = \mathbb{E}Y$.

Useful expression: $\mathrm{Cov}(X, Y) = \mathbb{E}XY - \mathbb{E}X\mathbb{E}Y$

Exercise: Derive the useful expression for computing covariances.

Correlation

The correlation between two rvs X and Y is defined as

$$\operatorname{corr}(X, Y) = \frac{\operatorname{Cov}(X, Y)}{\sqrt{\operatorname{Var} X}\sqrt{\operatorname{Var} Y}} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} =: \rho_{XY},$$

where $\sigma_X = \sqrt{\operatorname{Var} X}$ and $\sigma_Y = \sqrt{\operatorname{Var} Y}$.

We will later show that $\operatorname{corr}(X, Y) \in [-1, 1]$ for any rvs X and Y.

Exercise: Let (X, Y) be a pair of rvs with joint pdf given by

$$f(x,y) = \frac{1}{8}(x+y) \cdot \mathbf{1}(0 < x < 2, 0 < y < 2).$$

1. Find $\text{Cov}(X, Y)$.
2. Find $\text{corr}(X, Y)$.

Covariance and correlation of linearly transformed rvs

For any two rvs X and Y and constants $a, b, c, d \in \mathbb{R}$, we have

$$\begin{aligned} \mathrm{Cov}(aX + b, cY + d) &= ac \cdot \mathrm{Cov}(X, Y) \\ \mathrm{corr}(aX + b, cY + d) &= \mathrm{sign}(ac) \cdot \mathrm{corr}(X, Y). \end{aligned}$$

Exercise: Prove the result.

Sample space for rolling two dice, tabulated as $(\text{roll 1}, \text{roll 2})$:

$$\mathcal{S} = \left\{ \begin{array}{llllll} (1,1), & (1,2), & (1,3), & (1,4), & (1,5), & (1,6), \\ (2,1), & (2,2), & (2,3), & (2,4), & (2,5), & (2,6), \\ (3,1), & (3,2), & (3,3), & (3,4), & (3,5), & (3,6), \\ (4,1), & (4,2), & (4,3), & (4,4), & (4,5), & (4,6), \\ (5,1), & (5,2), & (5,3), & (5,4), & (5,5), & (5,6), \\ (6,1), & (6,2), & (6,3), & (6,4), & (6,5), & (6,6) \end{array} \right\}$$

Exercise: Let X = max and Y = min of rolls.

1. Find $\text{Cov}(X, Y)$.
2. Find $\text{corr}(X, Y)$.

Theorem (independence implies covariance equal to zero)

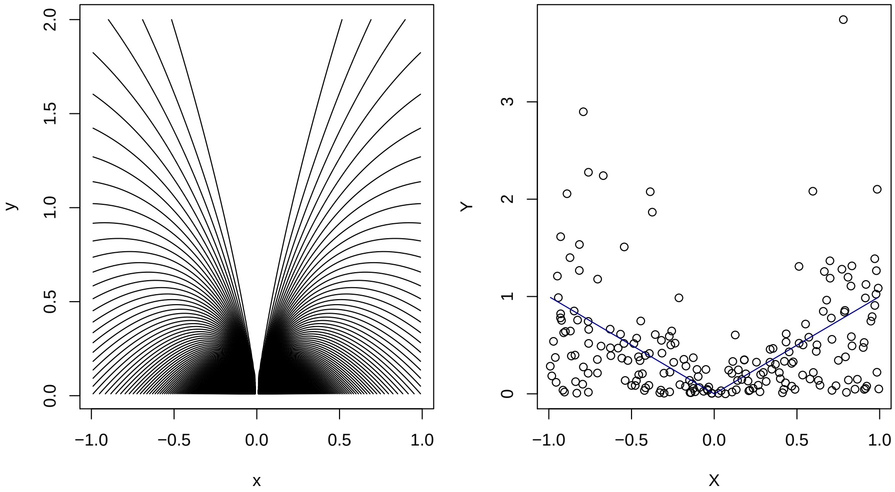
If X and Y are independent then $\text{Cov}(X, Y) = 0$.

If $\text{Cov}(X, Y) = 0$, it does not mean that X and Y are independent!

Exercise: Let (X, Y) be a pair of rvs with joint pdf given by

$$f(x, y) = \frac{1}{2|x|} e^{-y/|x|} \mathbf{1}(x \in (-1, 1) \setminus \{0\}, y > 0).$$

1. Check whether X and Y are independent.
2. Compute $\text{Cov}(X, Y)$.


0.0
0.5
1.0
1.5
2.0
y
−1.0
−0.5
0.0
0.5
1.0
x
0
1
2
3
Y
−1.0
−0.5
0.0
0.5
1.0
X

Theorem (Variance of linear combination of random variables)

For any $a, b \in \mathbb{R}$ *we have*

$$\operatorname{Var}(aX + bY) = a^2 \operatorname{Var} X + b^2 \operatorname{Var} Y + 2ab \operatorname{Cov}(X, Y).$$

Exercise: Prove the above.

Exercise: Suppose $\operatorname{Var} X = 2$, $\operatorname{Var} Y = 3$ and $\operatorname{Cov}(X, Y) = -3/2$.

Find $\operatorname{Var}(3X - Y)$.

Exercise: Let (X, Y) be a pair of rvs with joint pdf given by

$$f(x, y) = \frac{1}{8}(x + y) \cdot \mathbf{1}(0 < x < 2, 0 < y < 2).$$

1. Find $\text{Cov}(2X, Y)$
2. Find $\text{corr}(2X, Y)$.
3. Find $\text{Var}(X - Y)$
4. Find $\text{Var}(3X + Y/2)$.

Theorem (Variance of linear combination of random variables)

Let $X_1, \dots, X_n$ be random variables and let $a_1, \dots, a_n \in \mathbb{R}$. Then

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

Exercise: Prove the above.

Theorem (Variance of mean of independent rvs with same variance)

Let $X_1, \dots, X_n$ *be independent rvs all with variance* σ^2 *and let* $\bar{X} = n^{-1} \sum_{i=1}^{n} X_i$.

Then

$$\mathrm{Var}(\bar{X}) = \frac{\sigma^2}{n}.$$

Exercise: Prove the above.

Exercise: Let $Y_1, \dots, Y_n$ be independent rvs such that

$$Y_i \sim \text{Normal}(\mu, \sigma_i^2), \quad i = 1, \dots, n$$

and let

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \quad \text{and} \quad \tilde{Y} = \frac{\sum_{i=1}^{n} \sigma_i^{-2} Y_i}{\sum_{j=1}^{n} \sigma_j^{-2}}.$$

1. Find $\mathbb{E}\bar{Y}$.
2. Find $\text{Var}\,\bar{Y}$.
3. Find $\mathbb{E}\tilde{Y}$.
4. Find $\text{Var}\,\tilde{Y}$.

Exercise: Let $Z_1, \dots, Z_n$ have unit variance and suppose

$$\text{corr}(Z_i, Z_j) = \rho \in (-1, 1) \text{ for } i \neq j.$$

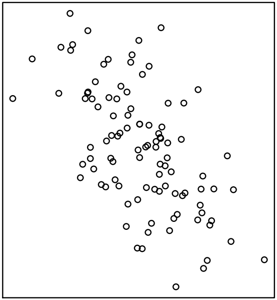
Find $\text{Var}\, \bar{Z}$, where $\bar{Z} = n^{-1} \sum_{i=1}^{n} Z_i$.

Bivariate Normal distribution

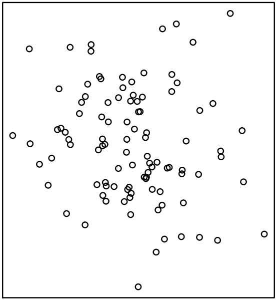
The rvs (X, Y) have the bivariate Normal distribution if they have joint pdf

$$f(x, y; \mu_X, \mu_Y, \sigma_X, \sigma_Y, \rho) = \frac{1}{2\pi} \frac{1}{\sigma_X \sigma_Y \sqrt{1-\rho^2}}$$
$$\times \exp\left[-\frac{1}{2(1-\rho^2)} \left(\left[\frac{X-\mu_X}{\sigma_X} \right] - 2\rho \left[\frac{X-\mu_X}{\sigma_X} \right] \left[\frac{Y-\mu_Y}{\sigma_Y} \right] + \left[\frac{Y-\mu_Y}{\sigma_Y} \right]^2 \right) \right].$$

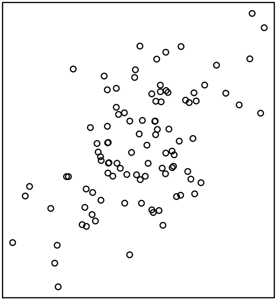
- μ_X is mean of X.
- μ_Y is mean of Y.
- σ_X^2 is variance of X.
- σ_Y^2 is variance of Y.
- ρ is $\text{corr}(X, Y)$.



rho = −0.5



rho = 0



rho = 0.5

Theorem (Cauchy-Schwarz Inequality)

For any rvs X and Y

$$|\mathbb{E}XY| \leq \mathbb{E}|XY| \leq \sqrt{\mathbb{E}X^2}\sqrt{\mathbb{E}Y^2}.$$

Exercise: Prove the above.

Exercise: Use the CS inequality to prove $\text{corr}(X, Y) \in [-1, 1]$ for any rvs X, Y.