Lecture 11

Law of the unconscious statistician (LOTUS) - reminder

Let (X, Y) be a bivariate random vector and $g : \mathbb{R}^2 \to \mathbb{R}$, then

- discrete case:

$$\mathbb{E}(g(X, Y)) = \sum_y \sum_x g(x, y) p_{X,Y}(x, y),$$

- continuous case:

$$\mathbb{E}(g(X, Y)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{X,Y}(x, y),$$

If X and Y are **independent**, then

- $$\mathbb{E}(XY) = \mathbb{E}X\mathbb{E}Y$$
- for any function h_1 and h_2:
$$\mathbb{E}(h_1(X)h_2(Y)) = \mathbb{E}(h_1(X))\mathbb{E}(h_2(Y)),$$
- $$Var(X + Y) = Var(X) + Var(Y).$$

Numerical characteristics of multivariate distributions

Definition (Independence)

Random variables $X_1, \ldots, X_n$ are said to be independent if and only if

- discrete case:

$$\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n) = \mathbb{P}(X = x_1) \cdot \ldots \cdot \mathbb{P}(X_n = x_n)$$

- continuous case:

$$f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = f_{X_1}(x_1) \cdot \ldots \cdot f_{X_n}(x_n)$$

Definition

The expected value of a random vector $X = (X_1, \ldots, X_n)$ is a vector

$$\mathbb{E}(X) = (\mathbb{E}(X_1), \ldots \mathbb{E}(X_n)),$$

the coordinates are the expected values of the respective random variables.

Covariance

Definition

Let X, Y be random variables such that $\mathbb{E}(X^2) < \infty$ and $\mathbb{E}(Y^2) < \infty$. **The covariance** of X and Y is defined as

$$Cov(X, Y) = \mathbb{E}\left(X - \mathbb{E}(X)\right)\left(Y - \mathbb{E}(Y)\right).$$

Properties of covariance:

1. $Cov(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$,
2. $Cov(X, Y) = Cov(Y, X)$,
3. $Cov(X, X) = Var(X)$,
4. $Cov(aX + b, cY + d) = acCov(X, Y)$ for constants a, b, c, d,
5. $Cov(X_1 + X_2, Y_1 + Y_2) = Cov(X_1, Y_1) + Cov(X_1, Y_2) + Cov(X_2, Y_1) + Cov(X_2, Y_2)$,
6. Schwarz inequality:

$$|Cov(X, Y)| \leq \sqrt{Var(X)}\sqrt{Var(Y)}.$$

Remark

Let X, Y be such that $\mathbb{E}(X^2) \leq \infty$ and $\mathbb{E}(Y^2) < \infty$, then

$$\begin{aligned} Var(X+Y) = Cov(X+Y, X+Y) &= Cov(X,X) + 2Cov(X,Y) + Cov(Y,Y) \\ &= Var(X) + 2Cov(X,Y) + Var(Y). \end{aligned}$$

Thus, for $X_1, \ldots, X_n$ such that $\mathbb{E}(X_i^2) < \infty$, $i = 1, 2, \ldots, n$, we have

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

Definition

Random variables X and Y are **uncorrelated**, if $Cov(X, Y) = 0$.

Remark

Let X, Y *be random variables such that* $\mathbb{E}(X^2) < \infty$ *and* $\mathbb{E}(Y^2) < \infty$.

1. *If* X *and* Y *are uncorrelated, then*

$$Var(X + Y) = Var(X) + Var(Y).$$

2. *If* X *and* Y *are independent then*

$$Cov(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = 0,$$

thus they are also uncorrelated.

However, **the converse** is not true: $Cov(X, Y) = 0$ **does not imply independence**, see two following examples:

Example

Let X has uniform distribution on the set $\{-2, -1, 0, 1, 2\}$. Let $Y = X^2$. Show that $Cov(X, Y) = 0$ but X and Y are not independent.

Solution

$$p_X(-2) = p_X(-1) = p_X(0) = p_X(1) = p_X(2) = \frac{1}{5},$$

$$Cov(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \mathbb{E}(X^3) - \mathbb{E}(X)\mathbb{E}(X^2),$$

$$\mathbb{E}(X) = (-2)\frac{1}{5} + (-1)\frac{1}{5} + 1\frac{1}{5} + 2\frac{1}{5} = 0,$$

$$\mathbb{E}(X^3) = (-2)^3\frac{1}{5} + (-1)^3\frac{1}{5} + \frac{1}{5} + 2^3\frac{1}{5} = 0,$$

therefore $Cov(X, Y) = 0$.

Example

Let (X, Y) be uniformly distributed on $D = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$. Determine $Cov(X, Y)$.

Solution

$(X, Y) \sim \mathcal{U}(D)$, *to* $f_{X,Y}(x, y) = \frac{1}{\pi}\mathbb{I}_D(x, y)$,

$$Cov(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y).$$

The symmetry about the origin of the joint distribution leads to

$$\mathbb{E}(X) = \mathbb{E}(Y) = \mathbb{E}(XY) = 0.$$

Therefore $Cov(X, Y) = 0$*. However* X *and* Y *are not independent (check the shape of the suport of the joint distribution of* (X, Y)*).*

Remark

The fact that the random variables are uncorrelated ($Cov(X, Y) = 0$*) does not indicate their independence.*

Examples - covariance in calculations

Example

n people throw their hats in a box and then pick a hat at random. Let X be the number of people that pick their own hat. Determine the variance of X.

Solution

$$X = X_1 + \ldots + X_n, \; \textit{where} \;\; X_i = \begin{cases} 1, & \textit{the "i"th person gets his own hat,} \\ 0, & \textit{otherwise.} \end{cases}$$

$\implies X_i \sim B\left(\frac{1}{n}\right),\; i = 1, \ldots, n.$

$$\mathbb{E}(X) = \mathbb{E}(X_1) + \ldots + \mathbb{E}(X_n) = n \cdot \frac{1}{n} = 1.$$

$$\mathrm{Var}(X) = \mathrm{Var}(X_1 + \ldots + X_n) = Var(X_1) + \ldots + Var(X_n) + 2 \sum_{1 \le i < j \le n} \mathrm{Cov}(X_i, X_j).$$

$\mathrm{Var}(X_i) = \frac{1}{n}(1 - \frac{1}{n}),\; i = 1, \ldots, n.$

Solution (cd)

For $i \neq j$,

$$Cov(X_i, X_j) = \mathbb{E}(X_i X_j) - \mathbb{E}(X_i)\mathbb{E}(X_j) = \mathbb{P}(X_i = 1, X_j = 1) - \mathbb{P}(X_i = 1)\mathbb{P}(X_j = 1)$$

$$= \mathbb{P}(X_i = 1)\mathbb{P}(X_j = 1 | X_i = 1) - \mathbb{P}(X_i = 1)\mathbb{P}(X_j = 1) = \frac{1}{n}\frac{1}{n-1} - \frac{1}{n^2}.$$

Then

$$\mathrm{Var}(X) = n\frac{1}{n}(1 - \frac{1}{n}) + 2\binom{n}{2}\left(\frac{1}{n(n-1)} - \frac{1}{n^2}\right) = 1.$$

Example

Flip a fair coin three times. Let X be the number of heads in the first two flips and let Y be the number of heads in the last two flips (there is overlap on the middle flip). Compute $Cov(X, Y)$.

Solution

We will present two ways of solving the task: the first method uses the joint probability distribution of (X, Y) *and the definition of covariance:*

$X\setminus Y$	0	1	2
0	1/8	1/8	0
1	1/8	2/8	1/8
2	0	1/8	1/8

From the marginal distributions: $\mathbb{E}(X) = \mathbb{E}(Y) = 1$, *and*

$$\mathbb{E}(XY) = 1 \cdot 1 \cdot \frac{2}{8} + 1 \cdot 2 \cdot \frac{1}{8} + 2 \cdot 1 \cdot \frac{1}{8} + 2 \cdot 2 \cdot \frac{1}{8} = \frac{10}{8},$$

hence $Cov(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \frac{10}{8} - 1 = \frac{1}{4}$.

Solution (cont'd)

Now we will solve the same task using the properties of covariance. Let

$$X_i = \begin{cases} 1, & \text{if } H \text{ comes up in the } i-th \text{ toss}, \\ 0, & \text{otherwise}, \end{cases} \qquad i = 1, 2, 3.$$

Thus $X = X_1 + X_2$, $Y = X_2 + X_3$. *Obviously, random variables* X_1, X_2, X_3 *are independent (since the different tosses are independent) and* $X_i \sim B(\frac{1}{2})$, $i = 1, 2, 3$. *So, using the property 5 of covariance, we get*

$$\begin{aligned} Cov(X, Y) &= Cov(X_1 + X_2, X_2 + X_3) \\ &= Cov(X_1, X_2) + Cov(X_1, X_3) + Cov(X_2, X_2) + Cov(X_2, X_3) \\ &= Cov(X_2, X_2) = Var(X_2) = \frac{1}{4}. \end{aligned}$$

Correlation

Definition

The correlation coefficient of two random variables X and Y having nonzero variances is defined as

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}.$$

Remark

The correlation coefficient removes the scale from the covariance. This characteristic is also dimensionless.

Properties of $\rho_{X,Y}$:

1. $|\rho_{X,Y}| \leq 1$,
2. $\rho_{X,Y} = 0 \iff$ X and Y are uncorrelated,
3. $|\rho_{X,Y}| = 1 \iff \exists a, b \in \mathbb{R} : \mathbb{P}(Y = aX + b) = 1.$

Definition

The covariance matrix of a random vector $X = (X_1, X_2)$ is defined as

$$C_X = \begin{bmatrix} Cov(X_1, X_1) & Cov(X_1, X_2) \\ Cov(X_2, X_1) & Cov(X_2, X_2) \end{bmatrix}.$$

Properties of the covariance matrix:

1. C_X is symmetric and positive definite,
2. $C_{aX} = a^2 C_X \quad \forall a \in \mathbb{R}$,
3. $\forall A \in \mathbb{R}^{m \times 2} \quad C_{AX} = ACA^T$.

Example

Determine the covariance matrix of (X, Y) and the correlation coefficient $\rho_{X,Y}$, and the covariance matrix of the random vector $(U, V) = (2X + Y, X - Y)$ if the joint pmf is of the form:

$X \backslash Y$	0	1
-1	1/2	1/4
0	1/8	1/8

Solution

$$\mathbb{E}(X) = (-1)\frac{1}{2} + (-1)\frac{1}{4} = -\frac{3}{4},\ \mathbb{E}(X^2) = \frac{3}{4},\ \mathbb{E}(Y) = \frac{3}{8},\ \mathbb{E}(Y^2) = \frac{3}{8},$$

$$\mathbb{E}(XY) = (-1)1\frac{1}{4} = -\frac{1}{4}.$$

$$C_{X,Y} = \begin{array}{|c|c|}\hline Var(X) & Cov(X,Y) \\ \hline Cov(X,Y) & Var(Y) \\ \hline \end{array} = \begin{array}{|c|c|}\hline 3/16 & 5/32 \\ \hline 5/32 & 15/64 \\ \hline \end{array},\quad \rho_{X,Y} = \frac{5/32}{\sqrt{3/16} \cdot \sqrt{15/64}}$$

$$C_{U,V} = AC_{X,Y}A^T,\quad where\ \ A = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}$$