Introduction to Probability Theory
7th week: Multivariate random variables.
The variance and covariance of random variables.

Definition. *The* **expectation** *(or the* **mean***) of a random variable X with atoms $\{x_1, x_2, \dots\}$ is defined as*

$$\mathbb{E}X = \sum_i x_i \mathbb{P}\left(X = x_i\right).$$

Remark: If $X : \Omega \to \mathbb{R}$ is a random variable and $g : \mathbb{R} \to \mathbb{R}$ is a (measurable) function, then $g(X) : \Omega \to \mathbb{R}$ is a new random variable.

Theorem (Law of the Unconscious Statistician). *If X is a discrete random variable and $g : \mathbb{R} \to \mathbb{R}$ is a (measurable) function, then*

$$\mathbb{E}(g(X)) = \sum_x g(x) \mathbb{P}\left(X = x\right).$$

Definition. *The* **variance** *of a random variable X is defined as*

$$\operatorname{Var} X = \mathbb{E}((X - \mathbb{E}X)^2) \quad (\textit{a useful formula: } \operatorname{Var} X = \mathbb{E}(X^2) - (\mathbb{E}X)^2).$$

Remark: For some random variables X the expectation does not exists and then the variance of such X is not defined. It might also happen that $\mathbb{E}X$ exists but the variance $\operatorname{Var} X$ does not, since the sum on the right-hand side of the formula does not (absolutely) converge.

Definition. *By a k-***dimensional random variable** *we mean a (measurable) function*

$$\bar{X} = (X_1, \dots, X_k) : \Omega \to \mathbb{R}^k,$$

where Ω is the sample space of a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Remark: On this course we will be studying mainly two dimensional random variables.

Definition. *A 2-dimensional random variable (X, Y) is* **discrete***, if for some countable set $A = \{(x_1, y_1), (x_2, y_2), \dots\}$ we have*

$$\sum_x \sum_y \mathbb{P}\left(X = x, Y = y\right) = 1.$$

For a discrete random variable (X, Y) the probabilities $p_{xy} = \mathbb{P}\left(X = x, Y = y\right)$ determine the **(joint) probability mass function (pmf)** which describes the **(joint) distribution of** (X, Y).

Definition. *The* **(joint) cumulative distribution function (CDF)** *of a 2-dimensional random variable (X, Y) is defined as the function $F_{X,Y} : \mathbb{R} \to \mathbb{R}$ such that*

$$F_{X,Y}(x, y) = \mathbb{P}\left(X \leqslant x, Y \leqslant y\right).$$

Theorem (Law of the Unconscious Statistician for two dimensions). *If (X, Y) is a 2-dimensional discrete random variable and $g : \mathbb{R}^2 \to \mathbb{R}$ is a (measurable) function, then*

$$\mathbb{E}(g(X, Y)) = \sum_x \sum_y g(x, y) \mathbb{P}\left(X = x, Y = y\right).$$

Definition. *The* **cumulative distribution function (CDF)** *of a 2-dimensional random variable (X, Y) is a function $F_{X,Y} : \mathbb{R} \to \mathbb{R}$ defined as*

$$F_{X,Y}(x, y) = \mathbb{P}\left(X \leqslant x, Y \leqslant y\right).$$

Definition. *The* **covariance** $\operatorname{Cov}(X, Y)$ *of random variables X and Y is defined as*

$$\operatorname{Cov} X = \mathbb{E}[(X - \mathbb{E}X)(Y - \mathbb{E}Y)] = \mathbb{E}(XY) - (\mathbb{E}X)(\mathbb{E}Y),$$

while the **correlation coefficient** $\rho(X, Y)$ *is given by*

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sqrt{\operatorname{Var} X \operatorname{Var} Y}}$$

Remark: For some random variables X and Y the covariance $\mathrm{Cov}(X,Y)$ and the correlation coefficient $\rho(X,Y)$ do not exist. However, if $\rho(X,Y)$ exists, then $|\rho(X,Y)| \leqslant 1$.

DEFINITION. *Let (X,Y) be a 2-dimensional random variable. The* **marginal distribution of** X *is given by*

$$\mathbb{P}(X = x_0) = \sum_y \mathbb{P}(X = x_0, Y = y)$$

while for **marginal distribution of** Y *we have*

$$\mathbb{P}(Y = y_0) = \sum_x \mathbb{P}(X = x, Y = y_0) .$$

If for every pair (x_0, y_0) we have

$$\mathbb{P}(X = x_0, Y = y_0) = \mathbb{P}(X = x_0)\,\mathbb{P}(Y = y_0) ,$$

then the random variables X and Y are **independent**.

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

$$\mathbb{E}XY = \mathbb{E}X\mathbb{E}Y .$$

In particular, for independent random variables Y and Y we have $\mathrm{Cov}(X,Y) = 0$ *and* $\rho(X,Y) = 0$ *(provided both* $\mathrm{Cov}(X,Y)$ *and* $\rho(X,Y)$ *exist).*

Remark: It is generally not true that if $\mathbb{E}XY = \mathbb{E}X\mathbb{E}Y$, then X and Y are independent.

Theorem. *For any real numbers $a, b, c \in \mathbb{R}$ and any functions $f, g : \mathbb{R}^2 \to \mathbb{R}$ we have*

$$\mathbb{E}(af(X,Y) + bg(X,Y) + c) = a\mathbb{E}f(X,Y) + b\mathbb{E}g(X,Y) + c.$$

Moreover,

$$\mathrm{Var}(aX + b) = a^2 \,\mathrm{Var}\, X .$$

If X **and** Y **are independent,** *then we have also*

$$\mathrm{Var}(X + Y) = \mathrm{Var}\, X + \mathrm{Var}\, Y .$$

PROBLEMS WE SHALL DO IN CLASS ON APRIL 18TH

Problem 1. The pmf of X is given by

k	0	1	2	3
$\mathbb{P}(X = k)$	0.4	0.3	0.2	0.1

a) Find $\mathbb{E}X$ and $\mathrm{Var}\, X$.
b) Find the pmf of $Y = |2 - X|$ and use it to compute $\mathbb{E}Y$.
c) Find $\mathbb{E}|2 - X|$ using *the Law of the Unconscious Statistician.*

Problem 2. The pmf of X is the following

i	-2	-1	0	1	2	7
$\mathbb{P}(X = i)$	0.18	0.16	0.28	0.23	0.14	0.01

Find the pmf of $Y = \mathrm{sgn}(X)$ and calculate $\mathbb{E}Y$ i $\mathrm{Var}\, Y$.

Problem 3. The joint pmf of (X,Y) is given by

$$\mathbb{P}(X = 0, Y = -1) = 1/2,\ \mathbb{P}(X = 1, Y = 1) = \mathbb{P}(X = -1, Y = 1) = 1/4 .$$

Find both marginal distributions and compute $\rho(X,Y)$. Are X and Y independent?

Problem 4. We toss a die n times. Let X_n denote the number of ones, and let Y_n stand for the number of twos. Are these two random variables independent?

Problem 5. The first urn contains 5 white balls and 3 black balls, in the second one we have 2 white balls and 2 black balls. We choose one ball from the first urn and one ball from the second. Let X denote the number of white balls we selected and Y the number of black balls we chose.

(i) Find the joint pmf of (X, Y).
(ii) Are X i Y independent?
(iii) Find $\mathbb{E}X$ i $\mathbb{E}Y$ using the marginal distributions of X and Y.
(iv) Find $\mathbb{E}X$ representing X as $X = X_1 + X_2$, where X_i is the number of white balls selected from the ith urn.
(v) What is the pmf of $Z = X + Y$?
(vi) Find $\rho(X, Y)$. Can one do it in a shorter way, not using the joint pmf of (X, Y) (but using (v))?

Homework assignment for Quiz 5, April 18th

Problem 1. The pmf of X is given by

$$\mathbb{P}(X = 8) = \frac{1}{8} \quad \text{and} \quad \mathbb{P}(X = i) = \frac{1}{2^i}, \text{ for } i = 1, 2, 3.$$

Find the pmfs of $Y = (-2)^X$ and $Z = 4X + 1$. Compute $\mathbb{E}Y$ and $\mathbb{E}Z$ in two ways: directly from their pmfs and using *the Law of the Unconscious Statistician.*

Problem 2. The pmf of X is given by

k	-1	0	1	3
$\mathbb{P}(X = k)$	$\frac{2}{10}$	$\frac{5}{10}$	$\frac{3}{20}$	$\frac{3}{20}$

(a) Find the pmf of $Y = |2 - X|$ and compute $\mathbb{E}Y$ and $\operatorname{Var} Y$.
(b) Find the pmf of $Z = \sin(\pi X/2)$, and compute $\mathbb{E}Z$ and $\operatorname{Var} Z$.

Problem 3. There are 37 numbers on the French roulette wheel: 18 of them are red, 18 are black, and one is green. Find the expectation and the variance of a win if the player:
(a) bets on token on red (then, if there will be red, he will win another token, if not, he will lose the token he bet);
(b) bets one token on 28 (then, in the case of 28, he will get another 35 tokens, otherwise he loses the token he bet).

Problem 4. We toss a die once. The first player wins 2, if the result is even, and loses 1 (i.e. wins -1) if it is odd. The second player wins 2 if the result is at least 4, and loses 1 in all other cases. Let X_i denote the win of the ith player.
(a) Find the joint pmf of (X_1, X_2).
(b) Find the marginal ditributions of (X_1, X_2) and check if X_1 i X_2 are independent.
(c) Compute $\rho(X_1, X_2)$.
(d) Find $\mathbb{E}X_1^2 X_2^3$.

Problem 5. We select three balls from an urn which contains 2 white and 3 black balls. The first player wins 6 if all three balls are black and loses 1 in all other cases. The second player wins 1 for every black ball (white balls does not affect his win). By X we denote the win of the first player, and Y the gain of the second player.
(a) Find the joint pmf of (X, Y).
(b) Find the marginal ditributions of (X, Y) and check if X and Y are independent.
(c) Find $\rho(X, Y)$.