3 WEEK

Introduction to the 3rd week

Now, we will consider **systems of random variables.** It means that for a particular random experiment, we will consider **several random variables that are associated with this experiment**. And we'll study **the relationship between these random variables.**

Example:

Let's return to example with casino

- Alice bets \$100 on red
- *At the same time* Bob bets \$100 on black

These're 2 different variables which values are defined by the result of the same random experiment "spinning of the wheel". We also see that these random variables are related to each other. For example, if we know that Alice won \$100 then it means that the outcome if our experiment gives us right number, and it means that Bob cannot win and he'll lose \$100 (he's payout is negative 100)

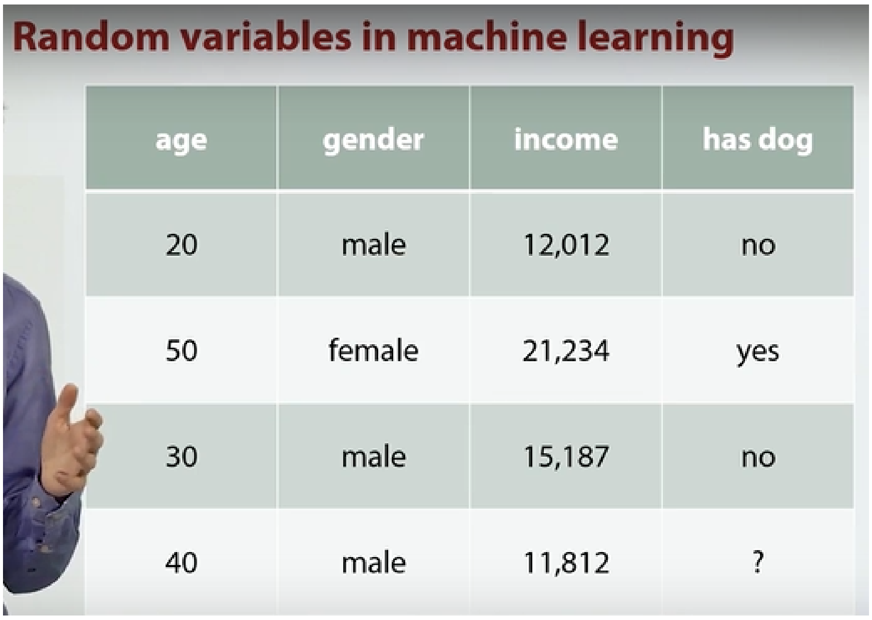
- Claudia bets $100 on 32

We have new random variable that is related to Alice payout. If Alice wins (we have red) and it increase chances of Claudia to win because 32 is red. There're some relationship between variables. But this relationship isn't deterministic. We cannot answer exactly what will be Caludia's payout even if we know Alice failed.

- Dan bets $100 on red on different roulette

Two roulette are independent to each other, we can use any information Alice payout to say smth new about Dan's payout. So, these random variables are independent to each other.

In machine learning


Random variables in machine learning

age	gender	income	has dog
20	male	12,012	no
50	female	21,234	yes
30	male	15,187	no
40	male	11,812	?

We can assume that these columns are random variables, because, for example, if we look at the 1st row of this table, we think about some particular person that gave these answers to some questions, and we understand that this particular person, was chosen randomly from some population. And the same thing holds for every row of this table. We can think about each row as a realization of some system of random variables. And we can try to find some relations between these random variables. For example, if we see a person with large age, we probably expect taht their income will be also large compared to some person with low age. Just because very young persons usually don't have large incomes.

In any case, we'll think about this variables as a kind of random variables. Because we don't know exactly what is the income of a particular person with age 21.

Linear transformations of random variables

What do we mean under a system of random variable?

Let's consider a random experiment. We know that to define a random experiment, we have to specify a probability space that is associated with it

Omega - set of elementary outcomes(w)

P(w) - probability of outcome, w belongs Omega

Let's define two random variables (or mathematically, two functions from omega to space of all real numbers)

X: Omega ⟶ R

Y: Omega ⟶ R

(X, Y) - system of random variables

What does it mean that it's a system? → This two variables can interact with each other. That means that if we know smth about X, we can say smth about Y, and wise versa.

Let's discuss how we can **create new random variable** from existing random variables:

We want to define Y by using variable X.

Transformations of random variables:

1 way:

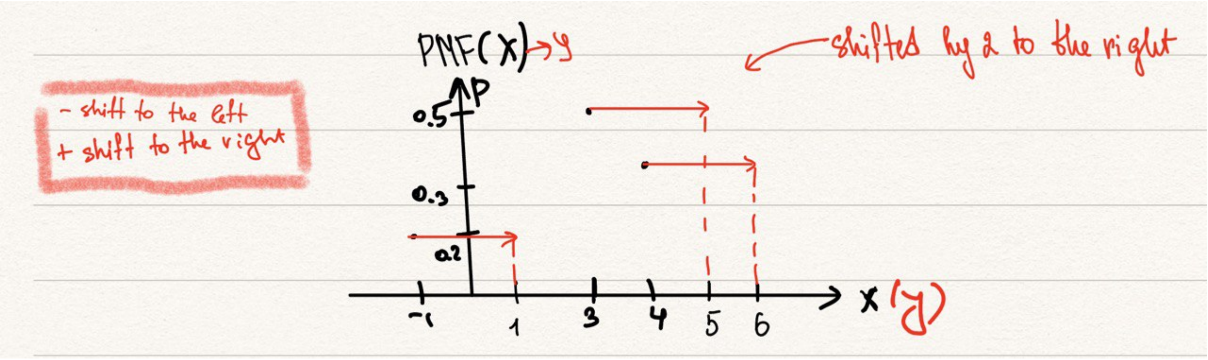
X - random variable

C - constant (for example, c = 2)

Y = X + C

P	0.2	0.5	0.3
X	-1	3	4
Y	1	5	6

It may be useful to think about PMF that are associated with this transformation (Y = X + C)



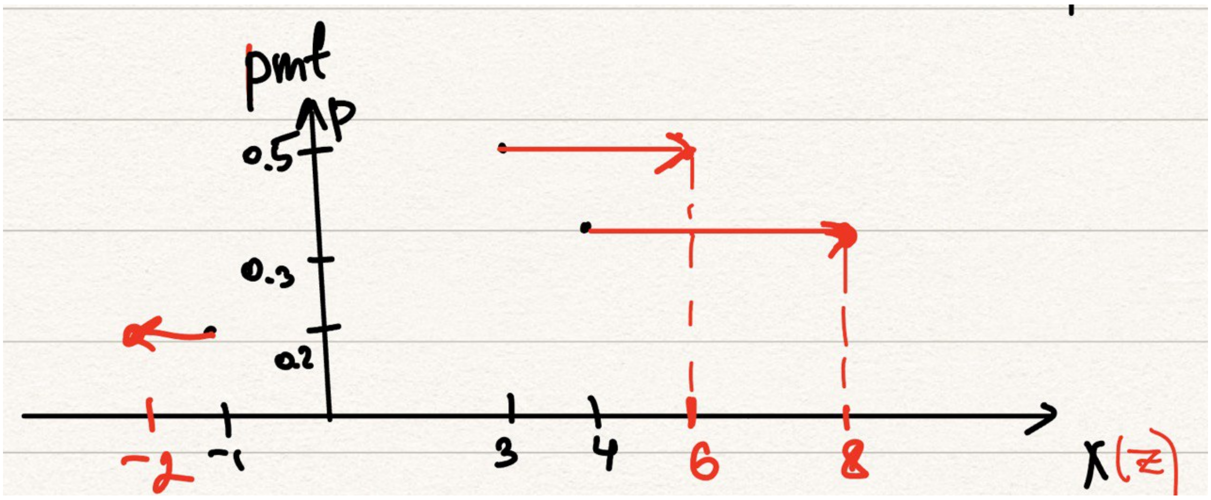
* on PMF we shift to the right if + C and shift to the right if -C

2 way:

C - constant (for example, c = 2)

Z = c * X

P	0.2	0.5	0.3
X	-1	3	4
Z	-2	6	8

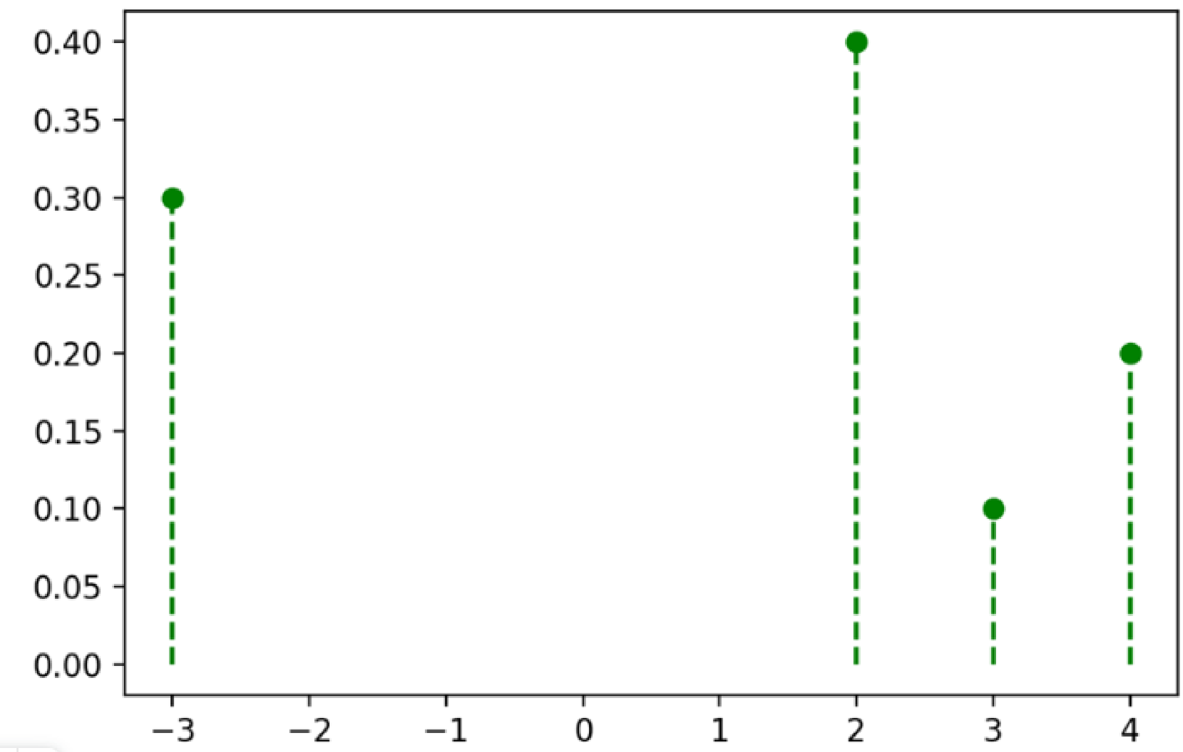


Multiplying → stretch whole pic

PMF of linear transformations

Question 1

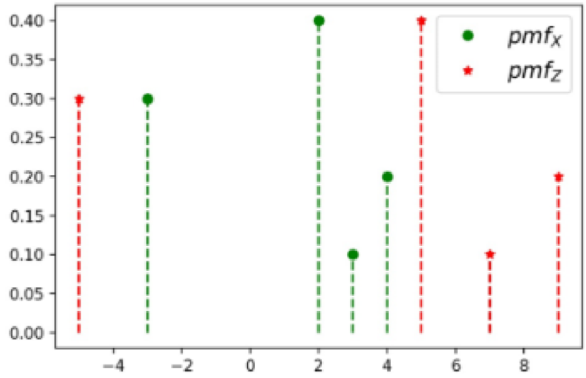
Discrete random variable X has PMF with the following graph:



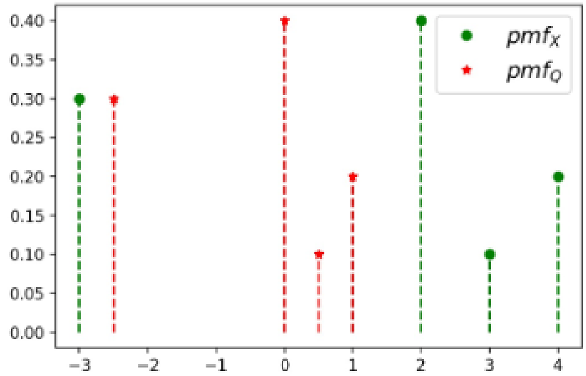
Provided the PMF graphs below, choose the correct sequence in which these random variables are listed:

1. 2X
2. 2X+1

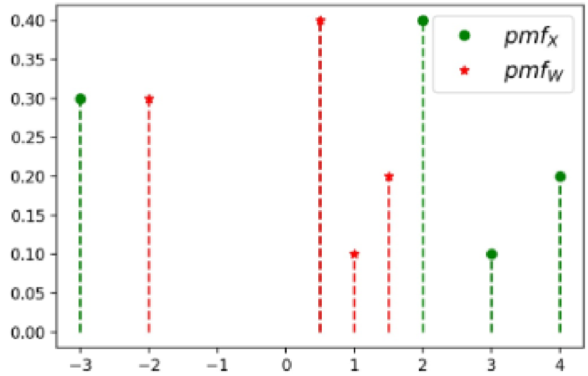
3. X/2 − 1
4. (X-1)/2

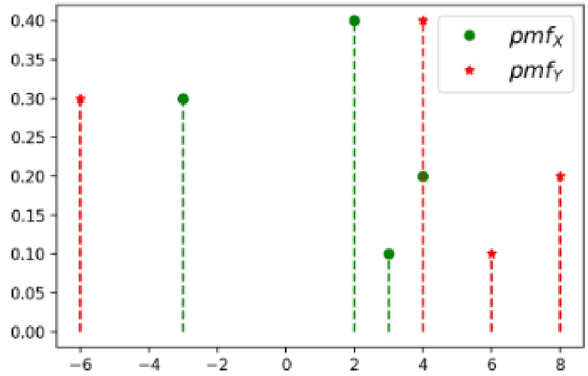


0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
−4
−2
0
2
4
6
8
pmf_X
pmf_Z



0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
−3
−2
−1
0
1
2
3
4
pmf_X
pmf_Q

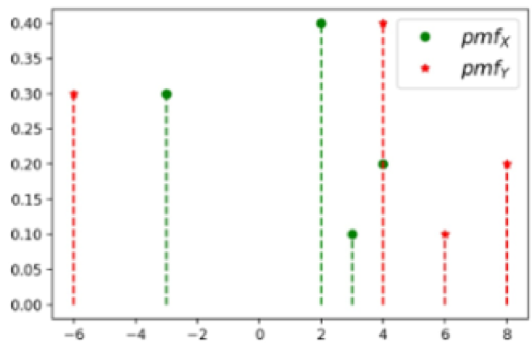




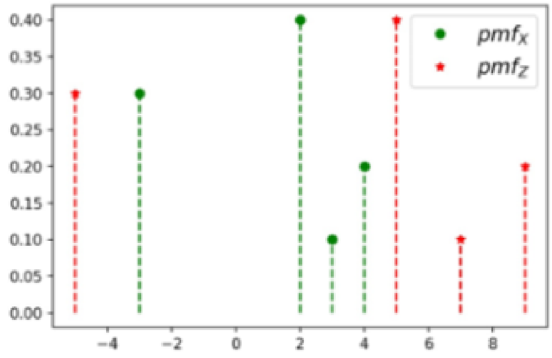
Let’s discuss how change Expected value by using this transformation

- **Y, Z, Q, W**
- Z, Y, Q, W
- Y, Z, W, Q
- W,Q,Z,Y
- Q,W,Y, Z

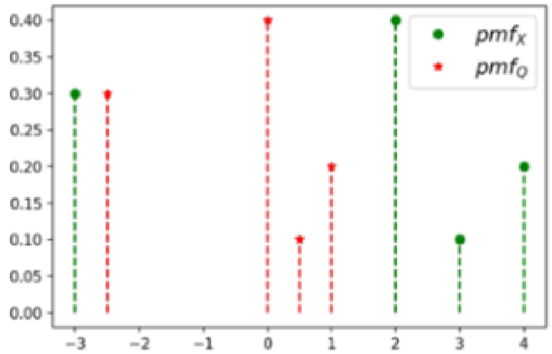
1. 2X



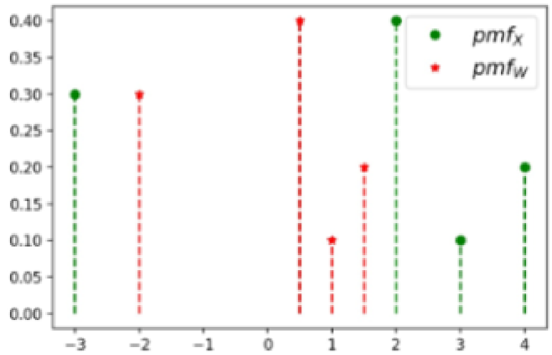
2. 2X+1



3. X/2 − 1



4. (X-1)/2



Linearity of expected value

Let's discuss some properties of **expected value**

$$\boxed{\mathbb{E}X = \sum_{i=1}^{n} p_i \cdot x_i}$$

p	p_1	p_2	p_3	$\cdots$	p_n
X	x_1	x_2	x_3	$\cdots$	x_n

1) $Z = X + c$ → 1st property

$$\mathbb{E}Z = \mathbb{E}(X+c) = \sum_{i=1}^{n} (x_i + c)p_i = \sum_{i=1}^{n} (x_i p_i + c p_i) = \sum_{i=1}^{n} x_i p_i + \sum_{i=1}^{n} c p_i$$

const

$$= \sum_{i=1}^{n} x_i p_i + c \underbrace{\boxed{\sum_{i=1}^{n} p_i}}_{=1} = (\mathbb{E}X) + c$$

(because it's sum of all possible probabilities)

p	p_1	p_2	p_3	$\cdots$	p_n
X	x_1	x_2	x_3	$\cdots$	x_n
Z	x_1+c	x_2+c	x_3+c	$\cdots$	x_n+c

→ 2nd property

2) $$\mathbb{E}(cX) = \sum_{i=1}^{n} c x_i p_i = c \sum_{i=1}^{n} x_i p_i = c \cdot \mathbb{E}X$$

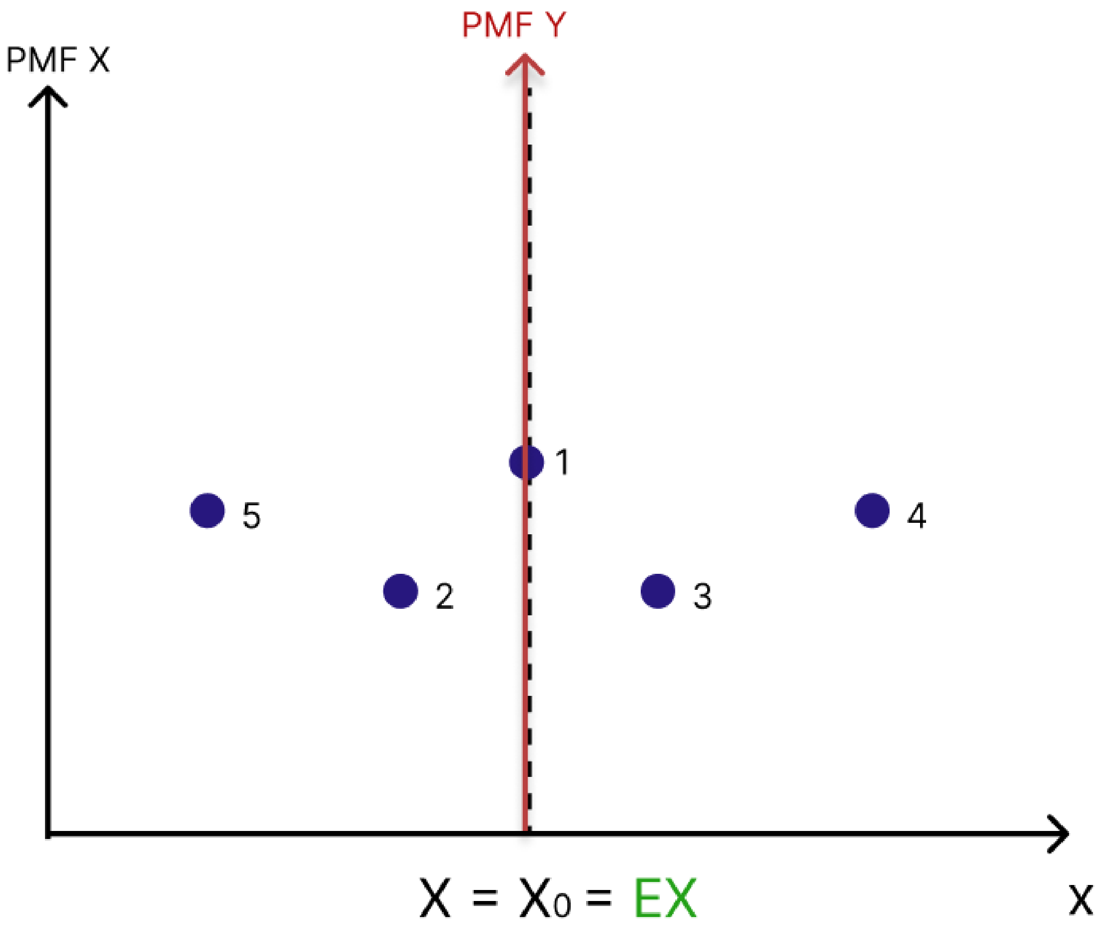
p	p_1	p_2	$\cdots$	p_n
X	x_1	x_2	$\cdots$	x_n
cX	cx_1	cx_2	$\cdots$	cx_n

Итого properties of Expected Values:

1. **E (X + c) = EX + c**
2. **E (cX) = c*EX**

Symmetric distributions and their expected values

Let's discuss symmetric distributions. It means that PMF have some vertical axis of symmetry. Let's make a graph:



Points 2 and 3 or 4 and 5 have the same probability, same deviation. If we try to find Expected value these dots cancel each other.

The symmetric with respectful to vertical line. In this case, it occurs that we can find EX of corresponding value without any calculations

EX = x0 (at least if it exists ←→ have finite number of values)

Let us consider the new variable Y = X - x0

PMF for Y is related to PMF of X. In fact, we just shift to this graph to the left.

Then we can introduce new vertical axis which is PMF of Y PMF of Y is symmetric with respect to vertical axis. It means that this function is even. The symmetry of this distribution means that the distribution of Y is the same as the distribution of negative Y. It means that they have the same Expected Value.

<u>Prof:</u>

EY = E(-Y) = E((-1) * Y) = (-1) * EY ⇒ EY = 0

We can see that the EY is a number that doesn't change when we multiply it by negative one. The only possible value for EY is to be 0.

EY = E(X - x0) = EX - x0 (1st property of expected value) = 0

EX - x0 =0

EX = x0.

Functions of random variables is new random variable

Now we can discuss a new way to create new random variable from existing. Actually, we can apply any function, just a function from a set of real numbers to set of real numbers

$$X : \Omega \to \mathbb{R}$$
$$f : \mathbb{R} \to \mathbb{R}$$
$$Z = f(x)$$

p	p_1	p_2	$\dots$	p_n
X	x_1	x_2	$\dots$	x_n
Z	$f(x_1)$	$f(x_2)$	$\dots$	$f(x_n)$

$$\Omega \xrightarrow{X} \mathbb{R} \xrightarrow{f} \mathbb{R}$$

$$Z = f \circ X \qquad Z(\omega) = f(X(\omega))$$

$$\mathbb{E}Z = \mathbb{E}f(X) = \sum_{i=1}^{n} f(x_i) \cdot p_i$$

Z is composition of two functions: x and f

Properties of variance

Variance is a measure how far from mean and how often this variable deviates from the Expected value.

$$\text{VarX} = \text{E}\,(\text{X} - \text{EX})^2$$

- we need this square to kill the sign of this expression. So positive deviance and negative deviance do not cancel each other.

Let us discuss how Variance behaves with transformations of random variable.

1. Z = X + c, c - const

VarZ = E (X + c - E (X + c))^2 = E (X + c - ((EX) + c))^2 = E (X + c - EX - c)^2 = E (X - EX)^2 = VarX

VarX + c = VarZ = VarX

So if we add some const to a random variable, its variance doesn't change. This is actually a quite natural because adding a const will shift our PMF to the left or to the tight, but doesn't change the deviation of values with respect to its Expected value.

2. Y = c * X, c - const

VarY = E (cX - EcX)^2 = E(cX - c*EX)^2 = E(c(X - EX))^2 = E (c)^2 (X - EX)^2 = (c^2) E(X - EX)^2 = (c^2) * VarX

VarY * c= (c^2) * VarX

Sum of random variables. Expected value and variance

Now, let us consider a pair of random variables defined on the same probability space. We can consider the sum of random variables.

1. How **Expected value** behaves with the sum:

X: Omega → R

Y: Omega → R

Z = X + Y

EZ = ?

EZ = E(X + Y) = EX + EY

Proof:

Let $\Omega = \{w_1, \dots, w_N\}$

$$\mathbb{E}X = \sum_{i=1}^{N} P(w_i)\, X(w_i)$$

$$\mathbb{E}Y = \sum_{i=1}^{N} P(w_i) \cdot Y(w_i)$$

$$\mathbb{E}Z = \sum_{i=1}^{N} P(w_i)\,(X(w_i) + Y(w_i)) = \sum_{i=1}^{N} \left(p(w_i)(X(w_i)) + P(w_i)Y(w_i)\right)$$

$$= \underbrace{\sum_{i=1}^{N} p(w_i) X(w_i)}_{\mathbb{E}X} + \underbrace{\sum_{i=1}^{N} p(w_i)\, Y(w_i)}_{\mathbb{E}Y}$$

2. How **Variance** behaves with the sum:

X - random variable

The only property we demand from X is indeed random. It means that it has nonzero or, better say, positive variance.

VarX = v >0

Y = -X

VarY = Var(-X) = Var((-1)*X) = (-1)^2 VarX = VarX = v

Z = X + Y ⇒ X - X

VarZ = Var(X - X) = Var(0) = 0 ⇒ Z actually is constant zero. Because Variance of any constant is equal to zero

So we can see that Variance of Z doesn't equal to sum of Variance of X and Y:

VarX + VarY = v + v = 2v >0

Var(X + Y) ≠ VarX + VarY

*But for some special case of independent variables this can be true

Expectation and variance properties practice

Question 1

The examiner gives grades (from 2 to 5) on the exam using the following approach: he tosses a coin 3 times, counts the number of heads and adds two. What would be the average grade for the exam in a class? Enter the exact value below with two decimal places or as an irreducible fraction. (e.g. 0.12 or 13/28):

Экзаменатор выставляет оценки (от 2 до 5) на экзамене следующим образом: он подбрасывает монету 3 раза, считает количество орлов и прибавляет две. Какова будет средняя оценка за экзамен в классе? Введите ниже точное значение с двумя знаками после запятой или в виде несократимой дроби. (например, 0,12 или 13/28):

Solution:

Давайте разберемся, как работает система оценивания и как посчитать среднюю оценку.

Вероятности выпадения орлов:

- **0 орлов (3 решки):** Вероятность - 1/2 *1/2* 1/2 = 1/8
- **1 орел:** Вероятность - 3 *1/2* 1/2 * 1/2 = 3/8 (три варианта, где может выпасть орел)
- **2 орла:** Вероятность - 3 *1/2* 1/2 * 1/2 = 3/8
- **3 орла:** Вероятность - 1/2 *1/2* 1/2 = 1/8

Оценки:

- 0 орлов + 2 = 2
- 1 орел + 2 = 3
- 2 орла + 2 = 4
- 3 орла + 2 = 5

Средняя оценка:

Для расчета средней оценки нужно умножить каждую оценку на ее вероятность, а затем сложить полученные значения:

- (2 *1/8) + (3* 3/8) + (4 *3/8) + (5* 1/8) = 2/8 + 9/8 + 12/8 + 5/8 = 28/8 = 3.5

Ответ: Средняя оценка за экзамен в классе будет 3.5. 15:44

Use properties of Expectation

The correct answer is: 7/2

Question 2

Two discrete random variables X and Y take values (0, 2, 5, 5, 7, 10, 11) and (3, 5, 8, 8, 10, 13, 14). All values have equal probability. How are their variances related? Try to use properties of variance to answer without precisely calculating the variances:

- VarY=3VarX
- VarY=VarX+9
- **VarY=VarX**

Solution:

VarX = E(X - EX) ^2

X	0	2	5	5	7	10	11
ДИСПХ	15,90476						
Y	3	5	8	8	10	13	14
ДИСПY	15,90476						

Expectation properties skill test

Question 1

Let X and Y be random variables on the same probability space, $\mathbb{E}X = 1, \mathbb{E}Y = 7$. Find expectation of $Z = X + 2Y - 5$. Enter the value below:

Solution:

Z = 1 + 14 - 5 = 10

The correct answer is: 10

Question 2

Find expectation of Binomial random variable. Use that it's a sum of n independent Bernoulli random variables with probability distribution:

Найдите математическое ожидание биномиальной случайной величины. Используйте это как сумму n независимых случайных величин Бернулли с распределением вероятностей:

- n^2p
- $n(1-p)$
- $np+n$
- np ✔
- $np(1-p)/2$
- $np(1-p)$

Solution:

The expected value of $Y = \sum_{i=1}^{n} X_i$ is

$$E(Y) = E(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} E(X_i) = np$$

where the last equality follows because $E(X_i) = p$ which is a constant and $\sum_{i=1}^{n} c = nc$ for any constant c.

Joint probability distribution

To describe interactions between random variables, we have to introduce a new notion.

Experimental: two tossing a fair coin

Let's introduce three random variables which are associated with this experiment:

- X = {1 if 1st tossing gives head, 0 otherwise} = [1st tossing gives head] = {HT, HH}
- Y = 1 - X
- Z = [2d tossing gives head] = {TH, HH}

Probability distributions:

P	1/2	1/2
X	1	0

P	1/2	1/2
Y	1	0

P	1/2	1/2
Z	1	0

No intersection: P(X = 0 ∩ Y = 0) = 0 as when X = 0 ⟶ Y = 1 - X ⟶ Y = 1

P(X = 0 ∩ Z =0) = P(TT) = 1/4

This shows us that despite the similarity between their probability distributions, these variables are distinct.

To capture this distinction, we have to introduce the notion of joint probability.

Definition of Joint probability distribution:

Def: X, Y - random variables

support $X = \{x_1, x_2, \dots, x_m\}$

support $Y = \{y_1, y_2, \dots, y_n\}$

$$P(X = x_i \cap Y = y_j) = p_{ij} \quad , i = 1, \dots, m \; ; j = 1, \dots, n$$

$$p_{ij} \geq 0 \quad , \quad \sum_{i=1}^{m} \sum_{i=1}^{n} p_{ij} = 1$$

What is the probability that X equals xi and Y equals yi?

Probability distribution:

Y \ X	0	1	
0	0	1/4	
1	1/4	0	

Z \ X	0	1	
0	1/4	1/4	
1	1/4	1/4	

> X = {1 if 1st tossing gives head, 0 otherwise} = [1st tossing gives head] = {HT, HH}
>
> Y = 1 - X

- P(X = 0 ∩ Y = 0) = 0 as when X = 0 → Y = 1 - X → Y = 1
- P(X = 0 ∩ Y = 1) = 1/4
- P(X = 1 ∩ Y = 1) = 0
- P(X = 1 ∩ Y = 0) = 1/4

—

All possible outcomes: {HH, TT, HT, TH}

> X = {1 if 1st tossing gives head, 0 otherwise} = [1st tossing gives head] = {HT, HH}
>
> Z = [2d tossing gives head] = {TH, HH}

- P(X = 0 ∩ Y = 0) = 1/4 = {TT}
- P(X = 1 ∩ Y = 1) = 1/4 = {HH}
- P(X = 1 ∩ Y = 0) = 1/4 = {HT}
- P(X = 0 ∩ Y = 1) = 1/4 = {TH}

Marginal distribution

Joint distribution of two variables gives us information about the distribution of both of them. **But what if we are interested only in one variable?**

Marginal distribution - is the distribution of one variable (Если нам известны joint probabilities of two events we have just sum their probability and we gey marginal distribution).

X, Y $\quad \operatorname{supp} X = \{x_1, \dots, x_m\}$

$\operatorname{supp} Y = \{y_1, \dots, y_n\}$

$p_{ij} = P(X = x_i \cap Y = y_j) \quad i = 1, \dots, m; \; j = 1, \dots, n$

$$P(X = x_i) = p_{i1} + \dots + p_{in} = \sum_{i=1}^{n} p_{ij}$$ (сложить)

Proof: $P(X = x_i) = P((X = x_i \cap Y = y_1) \cup \dots \cup (X = x_i \cap Y = y_n)) = P(X = x_i \cap Y = y_1) + \dots + P(X = x_i \cap Y = y_n) = p_{i1} + \dots + p_{in} = \sum_{j=1}^{n} p_{ij}$ ■

Y \ X	0	1	Marginal distribution
0	0	1/4	1/4
1	1/4	0	1/4
Marginal distribution	1/4	1/4	

Z \ X	0	1	Marginal distribution
0	1/4	1/4	1/2
1	1/4	1/4	1/2
Marginal distribution	1/2	1/2	

Marginal distribution X and Y:

- P(X = 0) = P(X = 0 ∩ Y = 0) U P(X = 1 ∩ Y = 0) = 0 + 1/4 = 1/4
- P(X = 1) = P(X = 1 ∩ Y = 0) U P(X = 1 ∩ Y = 1) = 0 + 1/4 = 1/4
- P(Y = 0) = P(X = 0 ∩ Y = 0) U P(X = 1 ∩ Y = 0) = 0 + 1/4 = 1/4
- P(Y = 1) = P(X = 0 ∩ Y = 1) U P(X = 1 ∩ Y = 1) = 0 + 1/4 = 1/4

So we can see that having the same marginal distribution we have different joint distributions. So it means that in general, if you know marginal distribution you cannot extract joint distribution of two variables because you don't know how they interact with each other. However, if you know that two variables are independent of each other. If you smth about the value of one variable, you don't know anything new about the value of the other value. Then you can extract joint distribution just by looking at these marginal distributions. We will discuss it later.

Joint distribution skill test

Question 1

Consider 2 random variables X and Y with joint distribution given by the table:

	Y=0	Y=1	Y=2
X=0	1/4	1/6	1/12
X=1	1/16	1/8	5/16

Find $P(\{X=1\} \cap \{Y>0\})$. Enter the exact value below with two decimal places or as an irreducible fraction. (e.g. 0.12 or 13/28):

Solution:

1/8 + 5/16 = 2/16 + 5/16 = 7/16

The correct answer is: 7/16

Question 2

Consider 2 random variables X and Y with joint distribution given by the table:

	Y=0	Y=1	Y=2
X=0	1/4	1/6	1/12
X=1	1/16	1/8	5/16

Find $P(\{Y=2\} \mid \{X=1\})$. Enter the exact value below with two decimal places or as an irreducible fraction. (e.g. 0.12 or 13/28):

*Express conditional probabilities through unconditional. Then find unconditional probabilities using joint PMF. Be accurate.

Solution:

B = {X=1} : {1/16, 1/8, 5/16} = 1/16 + 1/8 + 5/16 = 8/16 = 1/2

A = {Y=2}

P(A ∩ B) = 5/16

P(A|B) = P(A ∩ B) / P(B) = 5/16 / 1/2 = 5/16 ∗ 2 = 5/8

The correct answer is: 5/8

Independent random variables

If you have two random variables defined on the same probability space, these random variables can interact with each other. It means that if you know smth about the value of the one random variable, then you have some new information about the value of another random variable.

Definition of random independent variables:

X, Y - random variables

suppX = {x1, …., xm}

suppY = {y1, …., ym}

1. P (Y = yj | X = xi) = P(Y = yi) for all possible i and j ⇒ **X and Y are independent**
2. P(Y = yj ∩ X = xi) = P(Y = yj) * (X = xi) or all possible i and j ⇒ **X and Y are independent**

Z \ X	0	1	Marginal distribution
0	1/4	1/4	1/2
1	1/4	1/4	1/2
Marginal distribution	1/2	1/2	

So, in this case, if know that two **random variables are independent**, than we can find joint probability by multiplication of the corresponding values of marginal probabilities.

Y \ X	0	1	Marginal distribution
0	0	1/4	1/4
1	1/4	0	1/4
Marginal distribution	1/4	1/4	

But int his case, we see that these **two variables are not independent**. For example, because joint probability of P(X = 0 ∩ Y = 0) = 0 is not equal to multiplication of two marginal probabilities: 1/4 ∗ 1/4 ≠ 0

Another example of non-independent random variables

Now consider a little bit more complicated example of dependent random variables.

Random experiment: 3 coin tossing of fair coin

X - # of H for 1st and 2d tossings

Y - # of H for 2d and 3d tossings

<u>Is it true that X and Y are independent variable or not?</u>

To answer this questions let us construct joint distribution of these two random variables

Y \ X	0	1	2	Marginal distribution
0	1/8	1/8	0	1/4
1	1/8	1/4	1/8	1/2
2	0	1/8	1/8	1/4
Marginal distribution	1/4	1/2	1/4	

All possible outcomes: 2^3 = 8

P(X = 0 ∩ Y = 0) = P({TTT}) = 1/8

P(X = 1 ∩ Y = 0) = P({HTT}) = 1/8

P(X = 1 ∩ Y = 1) = P({HTH, THT}) = 2/8 = 1/4

1/4 ∗ 1/4 ≠ 1/8 ⇒ **They are not independent** but thay are not in deterministic dependence between other. The dependence between them is a kind of probabilistic. It gives us some information about probabilities of y depending on the value of x vice versa.

Joint PMF

Question 1

Two discrete random variables are independent if and only if the rows of their joint PMF are pairwise collinear as vectors (values in one row are multiples of corresponding values in another row with the same factor, which only depends on the chosen pair). Is that true?

Две дискретные случайные величины независимы тогда и только тогда, когда строки их совместной PMF попарно коллинеарны как векторы (значения в одной строке кратны соответствующим значениям в другой строке с тем же множителем, который зависит только от выбранной пары). Это правда?

- no
- yes

The correct answer is: yes

Question 2

A fair dice rolls. Let X be the number on the dice and Y be 0 for even X and 1 for odd. Find the joint PMF of X and Y:

	X=1	X=2	X=3	X=4	X=5	X=6
Y=0	1	0	1	0	1	0
Y=1	0	1	0	1	0	1

✓

	X=1	X=2	X=3	X=4	X=5	X=6
Y=0	0	1/6	0	1/6	0	1/6
Y=1	1/6	0	1/6	0	1/6	0

	X=1	X=2	X=3	X=4	X=5	X=6
Y=0	1/6	0	1/6	0	1/6	0
Y=1	0	1/6	0	1/6	0	1/6

Expected value of product of independent random variables

Previously we discussed that expected value of sum of two random variables is equal to sum of expected values. But, for example, for product(*) of expected value , it's not true. However, if two random variables are independent then a very natural relation between expected value of their product and product of expected values takes place.

X, Y - random variables

X and Y are independent

E(XY) = (EX) * (EY)

<u>Prof</u>: $\operatorname{supp} X = \{x_1, \ldots, x_m\}$ $\quad P(X=x_i)=p_i$

$\operatorname{supp} Y = \{y_1, \ldots, y_n\}$ $\quad P(Y=y_j)=q_j$

$$P(X=x_i \cap Y=y_j) \underset{\uparrow \text{independ.}}{=} P(X=x_i)\cdot P(Y=y_j) = p_i * q_j$$

This was a joint distribution due independence X and Y

$$\mathbb{E}(X \cdot Y) = \sum_{i=1}^{m} \sum_{j=1}^{n} P(X = x_i \cap Y = y_j) \cdot x_i * y_j =$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{n} p_i * q_j \cdot x_i y_j = \sum_{i=1}^{m} \left(p_i \cdot x_j \sum_{j=1}^{n} q_j \cdot y_j\right) =$$

const

$$= \left(\sum_{i=1}^{m} p_i x_i\right) \cdot \left(\sum_{j=1}^{n} q_j y_j\right) = \mathbb{E}X \cdot \mathbb{E}Y \blacksquare$$

We see that the crucial point in this prof is the fact that we can express this probability as a product of two probabilities. This allows us to move these terms out of this sum and replace the double sum with product of these two sums. If this condition is violated(dependence), this prof doesn't work.

Expectation properties skill test

Question 1

There are 60 white balls in a box and 40 black balls. Twenty random balls are taken. Find expected value of the number of black balls among them.

В коробке 60 белых шаров и 40 черных шаров. Берется двадцать случайных шаров. Найдите математическое ожидание количества черных шаров среди них.

Solution:

x	1 (black)	0 (white)
P(X = x)	2/5	3/5

Sum of Expected values:

20 * (2/5) + 0 ∗ 3/5 = 8

The correct answer is: 8

Question 2

Two dice are thrown and a random variable X is taken as a sum of the values on both dice plus a product of those values. What's EX? Enter the exact value below with two decimal places or as an irreducible fraction. (e.g. 0.12 or 13/28):

Брошены две игральные кости, и случайная величина X принимается как сумма значений на обоих кубиках плюс произведение этих значений. Что такое EX? Введите ниже точное значение с двумя знаками после запятой или в виде несократимой дроби. (например, 0,12 или 13/28):

Solution:

Possible	Summation	Multiplication	Result	Probability	EX
(1, 1)	2	1	3	**1/36**	3/36
(1, 2) (2, 1)	3	2	5	2/36	10/36
(1, 3) (3, 1)	4	3	7	2/36	14/36
(2, 2)	4	4	8	1/36	8/36
(1, 4) (4, 1)	5	4	9	2/36	18/36
(2, 3)(3, 2)	5	6	11	2/36	22/36
(1, 5) (5,1)	6	5	11	2/36	22/36
(2, 4)(4, 2)	6	8	14	2/36	28/36
(3, 3)	6	9	15	1/36	15/36
(1, 6) (6,1)	7	6	13	2/36	26/36
(2, 5) (5,2)	7	10	17	2/36	34/36
(3, 4) (4, 3)	7	12	19	2/36	38/36
(2,6)(6,2)	8	12	20	2/36	40/36
(3, 5) (5, 3)	8	15	23	2/36	46/36
(4, 4)	8	16	24	1/36	24/36
(3, 6) (6,3)	9	18	27	2/36	54/36
(4, 5) (5, 4)	9	20	29	2/36	58/36
(4, 6) (6,4)	10	24	34	2/36	68/36
(5, 5)	10	25	35	1/36	35/36
(6, 5) (5, 6)	11	30	41	2/36	82/36
(6,6)	12	36	48	1/36	48/36

	3
	10
	14
	8
	18
	22
	22
	28
	15
	26
	34
	38
	40
	46
	24
	54
	58
	68
	35
	82
	48
Sum	693

(693)/36 = 77/4 = 19.25

The correct answer is: 77/4

Question 3

Each of the three players throws a dice and then they sum the values. Afterwards, 2 is raised to the obtained degree. What's the expected value of the obtained random variable?

Каждый из трех игроков бросает кубик, а затем суммирует значения. После этого 2 возводится в полученную степень. Каково ожидаемое значение полученной случайной величины?

Solution:

Note that we can write:

$$E[X] = \frac{1}{6^3} \sum_{i=1}^{6} \sum_{j=1}^{6} \sum_{k=1}^{6} 2^{i+j+k}$$

$$E[X] = \frac{1}{6^3} \sum_{i=1}^{6} 2^i \sum_{j=1}^{6} 2^j \sum_{k=1}^{6} 2^k$$

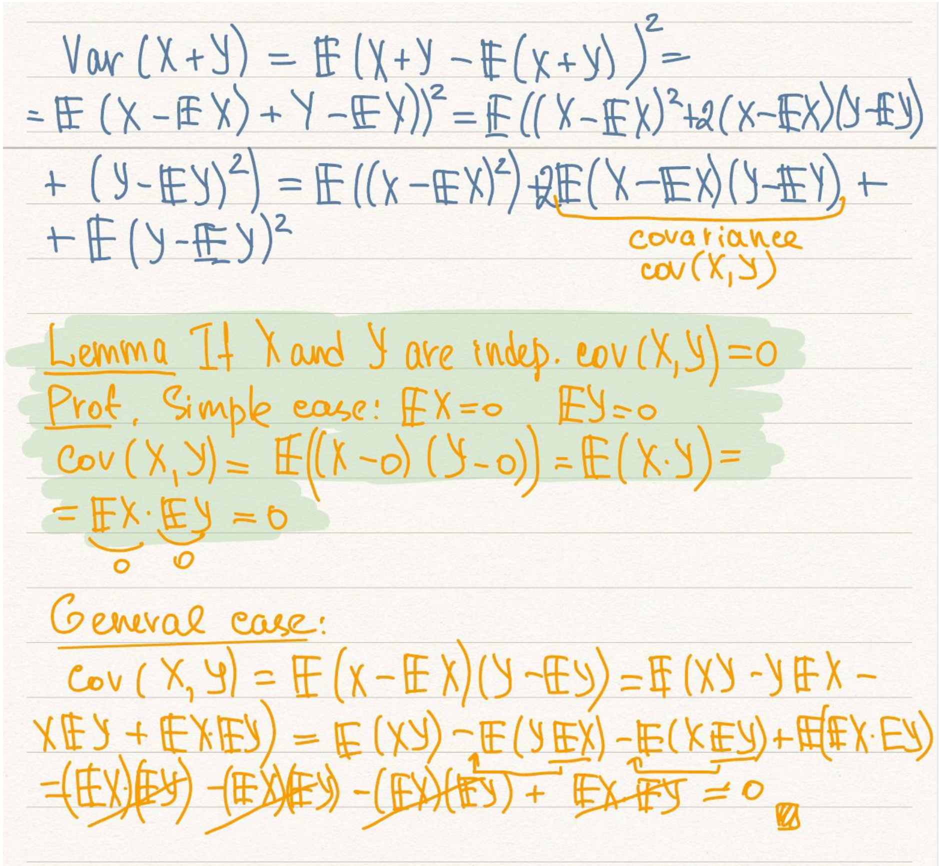
If $f(n) = \sum_{k=1}^{n} 2^k$ then $f(n) = 2^{n+1} - 2$, proof left as an exercise to the reader (hint: induction).

$$E[X] = \frac{f(6)^3}{6^3} = \left(\frac{2^7 - 2}{6}\right)^3 = 21^3 = 9261$$

The correct answer is: 9261

Variance of sum of random variables. Covariance

To proceed, we have to discuss variance of sum of two random variables. Previously, we found that variance of sum of two random variables isn’t necessarily equal to sum of variances. However, it appears that if two random variables are independent, it’s true that variance of sum is equal to sum of variances. To prove it, first we have to prove an additional lemma and also introduce a notion of covariance of two random variables.



$$Var(X+Y) = \mathbb{E}(X+Y-\mathbb{E}(X+Y))^2 =$$
$$= \mathbb{E}(X-\mathbb{E}X) + Y - \mathbb{E}Y))^2 = \mathbb{E}((X-\mathbb{E}X)^2 + 2(X-\mathbb{E}X)(Y-\mathbb{E}Y)$$
$$+ (Y-\mathbb{E}Y)^2) = \mathbb{E}((X-\mathbb{E}X)^2) + 2\underbrace{\mathbb{E}(X-\mathbb{E}X)(Y-\mathbb{E}Y)}_{\substack{\text{covariance} \\ cov(X,Y)}} +$$
$$+ \mathbb{E}(Y-\mathbb{E}Y)^2$$

Lemma If X and Y are indep. $cov(X,Y)=0$

Prof. Simple case: $\mathbb{E}X=0 \quad \mathbb{E}Y=0$

$$Cov(X,Y) = \mathbb{E}((X-0)(Y-0)) = \mathbb{E}(X\cdot Y) =$$
$$= \underbrace{\mathbb{E}X}_{0}\cdot\underbrace{\mathbb{E}Y}_{0} = 0$$

General case:

$$Cov(X,Y) = \mathbb{E}(X-\mathbb{E}X)(Y-\mathbb{E}Y) = \mathbb{E}(XY - Y\mathbb{E}X -$$
$$X\mathbb{E}Y + \mathbb{E}X\mathbb{E}Y) = \mathbb{E}(XY) - \mathbb{E}(Y\mathbb{E}X) - \mathbb{E}(X\mathbb{E}Y) + \mathbb{E}(\mathbb{E}X\cdot\mathbb{E}Y)$$
$$= \cancel{(\mathbb{E}X)(\mathbb{E}Y)} - \cancel{(\mathbb{E}X)(\mathbb{E}Y)} - \cancel{(\mathbb{E}X)(\mathbb{E}Y)} + \cancel{\mathbb{E}X\,\mathbb{E}Y} = 0 \quad \blacksquare$$

Var(X + Y) = E((X - EX)^2) + 2* E(X - EX)(Y - EY) + E((Y - EY)^2)

Covariance measures some kind association between X and Y. It's strongly dependence between X and Y.

Covariance = E(X - EX)(Y - EY)

If X and Y are independent ⇒ Cov(X, Y) = 0

The crucial idea in **Cov(X, Y) = 0 is that we have inequality due to independence of X and Y.**

So if X and Y are independent Cov(X, Y) = 0 we can see that Var(X + Y) = E((X - EX)^2) + ~~2* E(X - EX)(Y - EY)~~ + E((Y - EY)^2) **⇒ Var(X + Y) = E((X - EX)^2) + E((Y - EY)^2) = Var(X) + Var(Y)**

Variance properties practice

Question 1

Let X be the sum of values on 6 dice rolled. Find VarX. Enter the exact value below with two decimal places or as an irreducible fraction. (e.g. 0.12 or 13/28):

Solution:

- E(X) = (1 + 2 + 3 + 4 + 5 + 6) / 6 = 3.5
- E(X^2) = (1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2) / 6 = 15.17
- Var(X) = E(X^2) - (E(X))^2 = 15.17 - 3.5^2 = 2.92

*Var(X) = 6**2.92 = 17.52

The correct answer is: 35/2 = 17.52

Variance of Binomial random variable

Question 1

Find variance of Binomial random variable. Use that it's a sum of n independent Bernoulli random variables with PMF:

- a. $np(1-p)$ ✔
- b. $n^2(1-p)^2$
- c. n^2p
- d. np^2

The correct answer is: np(1–p)

Properties of covariance

Covariance measures some property of association between two random variables. Let us discuss some properties of covariance.

1. **Var(X + Y) = E((X - EX)^2) + 2* E(X - EX)(Y - EY) + E((Y - EY)^2)**
2. **If X and Y are independent ⇒ Cov(X, Y) = 0 . The inverse isn't true**

Does it follow that if covariance equals to zero, then X and Y are independent? Does it true that inverse holds?

Y \ X	-1	0	1	Marginal distribution
1	1/3	0	1/3	2/3
-2	0	1/3	0	1/3
Marginal distribution	1/3	1/3	1/3	

It's clear that X and Y are not independent to each other. For example, 1/3 ∗ 2/3 ≠ 1/3

But what can we say about covariance between X and Y?

> 💡 Expected value = p * marginal distribution

EX = -1 ∗ 1/3 + 0∗1/3 +1∗1/3 = -1/3 +1/3 = 0

EY = 1∗ 2/3 -2∗1/3 = 0

This is a symmetric distribution

Cov(X, Y) = E(XY) = (-1) * 1 ∗ 1/3 + (-2) * 0 ∗ 1/3 + 1∗ 1 ∗1/3 = 0

X and Y are not independent with Cov(X, Y) = 0

3. Cov(X, Y) = Cov(Y, X). If the swap the product doesn't change.
4. Cov(X + c, Y) = Cov(X, Y). Holds for both variables.
5. Cov(c*X=, Y) = c * Cov(X, Y)

Prof: E(cX - EcX) (Y- EY) = E(cX - cEX) (Y- EY) = Ec(X - EX) (Y- EY) = c* E(X - EX) (Y - EY)

6. Y = kX + b, k and b are costants

Cov(X, Y) =Cov(X, kX + b) =Cov(X, kX) = k * Cov(X, X) = k * VarX

> 💡 Cov(X, X) = VarX

In a sense, covariance measure some kind of linear relationship between two variables. If we know that the knowledge that on variable is large gives us dome information like the other variable is also large and vice versa. Then they have non-zero covariance.

A little bit better tool to measure this kind of linear relationship is correlation.

Covariance for a dice roll

Question 1

A fair dice rolls. Let X be the number on the dice and Y be 0 for even X and 1 for odd. Find covariance of X and Y. Enter the exact value below with two decimal places or as an irreducible fraction. (e.g. 0.12 or 13/28):

Бросается справедливая игра в кости. Пусть X — число на игральной кости, Y — 0 для четного X и 1 для нечетного. Найдите ковариацию X и Y

Solution:

Y \ X	1	2	3	4	5	6	Marginal distribution
0	0	1/6	0	1/6	0	1/6	1/2
1	1/6	0	1/6	0	1/6	0	1/2
Marginal distribution	1/6	1/6	1/6	1/6	1/6	1/6	

X and Y are not independent

EX = 1/6 + 2/6 + 3/6 + 4/6 + 5/6 + 6/6 = 21/6 = 7/2

EY = 1/2

X	Y	XY	Probability
1	1	1	1/6
2	0	0	1/6
3	1	3	1/6
4	0	0	1/6
5	1	5	1/6
6	0	0	1/6

E[XY] = (1/6)(1 + 3 + 5) = 3/2

This is a symmetric distribution

Cov(X, Y) = E(X - EX)(Y - EY) = E[XY] - E[X]E[Y] = (3/2) - (3.5)(0.5) = **-1/4**

The correct answer is: -1/4

Correlation of two random variables

Covariance gives us some information about relation between two variables. There is another coefficient that can be used for similar objective but has a little bit different and, in some cases, a little bit better properties. This coefficient is called correlation.

Actually, correlation strictly related to covariance

> 💡 Corr(X, Y) = Cov(X, Y) / sqrt(VarX * VarY)

Let us consider a case of perfect linear relationship:

Y = kX + b

corr(X, Y) = corr(X, kX + b) = cov(X, kX + b) /sqrt(VarX * kX + b) = k*VarX / sqrt(VarX * VarkX) = k*VarX / sqrt(k^2 (VarX)^2) = k*~~VarX~~ / |k|*~~VarX~~ = k / |k|

= {1 if k>0; -1 if k < 0}

So we see that if there is perfectly linear relationship between two random variable. And correlation between these two variables is equal either 1 or -1 depending on the sign of coefficient k. This makes correlation a perfect measure of linear relationship between two random variables.

Properties of correlation:

1. Corr(k*X, Y) = Corr(X, Y) if k > 0. So scaling random variables doesn't effect on correlation
2. Corr(X, Y) belongs to [-1, 1] ⇒ Y = kX + b. k influence on the sign of correlation
3. Corr(X, Y) = 0 <⇒ Cv(X, Y) = 0 uncorrelated

uncorrelated (if two variables are independent, they are uncorrelated but inverse is not true)

Correlation quiz

Question 1

Select all correct statements:

- **if X and Y are independent then corr(X,Y) = 0**
- if corr(X,Y) = 0 then X and Y are independent
- **correlation takes values on the segment [-1, 1]**
- **the closer absolute value of correlation of random variables to 1, the more intense is linear dependence between them**
- corr(X, -X) is always equal to 0

Week three skill test

Question 1

There are 2 yes/no questions in a questionnaire. Let X and Y be random variables that correspond to indicators of positive answer for the questions. Joint probability distribution of X and Y is given by the table:

	X=0	X=1
Y=0	0.8	0.1
Y=1	0.05	0.05

find EX. Enter the exact value below with two decimal places or as an irreducible fraction. (e.g. 0.12 or 13/28):

2 вопроса да/нет в анкете. Пусть X и Y — случайные величины, соответствующие показателям положительного и положительного результата.

Solution:

- P(X=0) = 0.8 + 0.05 = 0.85
- P(X=1) = 0.1 + 0.05 = 0.15

*EX = (0 ** 0.85) + (1 ∗ 0.15) = 0.15 = 3/20

The correct answer is: 3/20

Question 2

Under the same conditions, find EY. Enter the exact value below with two decimal places or as an irreducible fraction.

Solution:

- P(Y=0) = 0.8 + 0.1 = 0.9
- P(Y=1) = 0.05 + 0.05 = 0.1

*EX = (0 ** 0.9) + (1 ∗ 0.1) = 0.1 = 1/10

The correct answer is: 1/10

Question 3

Under the same conditions, find cov(X,Y). Enter the exact value below with two decimal places or as an irreducible fraction.

Solution:

Cov(X, Y) = E(X - EX)(Y - EY) = E[XY] - E[X]E[Y] =

E[XY]

X	Y	XY	Probability
0	0	0	0.8
0	1	0	0.05
1	0	0	0.1
1	1	1	0.05

E[XY] = (0 * *0.8) + (0 ** 0.05) + (0 * *0.1) + (1 ** 0.05) = 0.05
Cov(X, Y) = E[XY] - E[X]E[Y] = 0.05 - (0.15)(0.1) = 0.035 = 7/200

The correct answer is: 7/200

Question 4

1. **SD(X):**

* Var(X) = E[X^2] - (E[X])^2 *E[X^2] = (0^2* 0.85) + (1^2 ∗ 0.15) = 0.15
 * Var(X) = 0.15 - (0.15)^2 = 0.1275
 * SD(X) = √Var(X) = √0.1275 ≈ 0.357

2.
SD(Y):

* Var(Y) = E[Y^2] - (E[Y])^2 *E[Y^2] = (0^2* 0.9) + (1^2 ∗ 0.1) = 0.1
 * Var(Y) = 0.1 - (0.1)^2 = 0.09
 * SD(Y) = √Var(Y) = √0.09 = 0.3

Corr(X, Y) = Cov(X, Y) / (SD(X) *SD(Y)) = 0.035 / (0.357* 0.3) ≈ 0.327

The correct answer is: 0.32

Systems of random variables. Highlights

- One can consider several random variables associated with the same random experiment.
- Such random variables can either be independent of each other or they can "interact" in some way.
- For example, random variables can be positively correlated. It means that if we know that one variable obtained large value, the other most likely will also be large, and vice versa.
- To describe system of random variables, one can use *joint distribution,* i.e. a table of all possible combinations of values and corresponding probabilities.