Математические и статистические методы в психологии
Семинар 6. Дискретные случайные величины: дисперсия и ее свойства, ковариация и корреляция (решения)
А. А. Макаров, А. А. Тамбовцева, Н. А. Василёнок

Задача 1. В 1963 году Стэнли Милгрем начал серию экспериментов, исследовавших готовность человека подчиниться авторитету. В эксперименте участвовало трое: экспериментатор, испытуемый (*«учитель»*), и актер, игравший роль другого испытуемого (*«ученик»*), который был привязан к креслу с электродами. Учитель давал ученику задание на запоминание слов. Как только ученик ошибался, экспериментатор давал указание учителю «наказывать» ученика ударом тока (который, конечно, не был настоящим, о чем учитель, конечно, не знал). С каждой допущенной ошибкой экспериментатор требовал все большего разряда тока.

Милгрем выяснил, что примерно 65% испытуемых соглашались наказывать ученика током, тогда как 35% – отказывались. Последующие эксперименты, проведенные в других странах, показали схожие результаты.

В эксперименте, аналогичном эксперименту Милгрема, принимают участие 5 испытуемых. Примем за успех ситуацию, в которой испытуемый отказывается от «наказания» ученика. Рассчитайте вероятность, с которой:

(a) Ровно 3 испытуемых откажутся от «наказания» ученика;
(b) Менее 2 испытуемых согласятся наказывать «ученика».
(c) Сколько, в среднем, испытуемых откажутся от «наказания» ученика?
(d) Охарактеризуйте разброс результатов относительно среднего значения, которые можно ожидать в этом эксперименте.

Решение. В эксперименте принимают участие 5 испытуемых, каждый из которых может либо подчиниться авторитету и согласиться с предложенным наказанием, либо отказаться от предложения наказать «ученика» током. Нас интересует случайная величина X, которая описывает количество *отказавшихся* от подчинения авторитету. Минимальное значение этой случайной величины – 0 (все согласились), а максимальное – 5 (все отказались). Каждый испытуемый принимает решение *независимо от другого*, и, судя по результатам предыдущих исследований, вероятность оказаться – вероятность успеха – у всех участников эксперимента постоянна и составляет $p = 0.35$. Следовательно, случайная величина X имеет биномиальное распределение с параметрами $X \sim Bin(0.35, 5)$.

1. Воспользуемся формулой вероятности значения случайной величины, распределенной биномиально:

$$P(X = k) = C_n^k \times p^k \times q^{n-k}$$

$$P(X = 3) = C_5^3 \times 0.35^3 \times 0.65^2 = 10 \times 0.0429 \times 0.4654 = 0.18$$

2. Рассмотрим следующие ситуации, подходящие под условие: (а) ни один испытуемый не согласится с предложенным наказанием и (б) только один испытуемый согласится, тогда как 4 не согласятся. Следовательно, нас интересует сумма вероятностей $P(X = 5)$ и $P(X = 4)$[1]

$$P(X = 4) = C_5^4 \times 0.35^4 \times 0.65^1 = 5 \times 0.015 \times 0.65 = 5 \times 0.00975 = 0.04877$$

$$P(X = 5) = C_5^0 \times 0.35^5 \times 0.65^0 = 0.005$$

$$P(X \geq 4) = P(X = 4) + P(X = 5) = 0.054$$

[1] Так как за *успех* мы обозначили несогласие испытуемого с экспериментатором. Мы могли бы переобозначить вероятности и тогда бы считали сумму $P(X = 0 \cup X = 1) = C_5^0 \times 0.65^0 \times 0.35^5 + C_5^1 \times 0.65^1 \times 0.35^4$, что, как не сложно заметить, то же самое. Главное здесь – не запутаться в том, какие вероятности в какие степени мы возводим.

3. В этом пункте нас спрашивают про среднее значение – математическое ожидание случайной величины X. Мы знаем, что математическое ожидание биномиального распределения рассчитывается как произведение параметров n и p.

$$E(x) = n \times p = 5 \times 0.35 = 1.75$$

4. В этом пункте нас спрашивают про дисперсию случайной величины X. Мы знаем, что для биномиального распределения она рассчитывается как $D(x) = n \times p \times q$.

$$D(x) = 5 \times 0.35 \times 0.65 = 1.1375$$

Задача 2. Случайная величина X задана рядом распределения:

X	-3	0	1	2
P	0.1	0.2	0.4	**0.3**

Рассчитайте дисперсию и стандартное отклонение:
(a) Случайной величины X;
(b) Случайной величины $Y = 3X - 2$.

Решение. Сначала восстановим пропущенную вероятность:

$$P(X = 2) = 1 - (0.1 + 0.2 + 0.4) = 0.3$$

1. Дисперсия дискретной случайной величины расчитывается как

$$D(X) = E(X^2) - [E(X)]^2$$

Рассчитаем математическое ожидание:

$$E(X) = (-3) \times 0.1 + 1 \times 0.4 + 2 \times 0.3 = -0.3 + 0.4 + 0.6 = 0.7$$

$$[E(X)]^2 = 0.7^2 = 0.49$$

Рассчитаем $E(X^2)$, возводя каждое из значений случайной величины X в квадрат:

$$E(X^2) = (-3)^2 \times 0.1 + 1^2 \times 0.4 + 2^2 \times 0.3 = 9 \times 0.1 + 1 \times 0.4 + 4 \times 0.3 = 0.9 + 0.4 + 1.2 = 2.5$$

Наконец, мы можем рассчитать дисперсию:

$$D(X) = 2.5 - 0.49 = 2.01$$

Стандартное отклонение равно корню из дисперсии: $D(X) = \sqrt{2.01} = 1.42$

2. Воспользуемся свойствами дисперсии:

$$D(Y) = D(3X - 2) = D(3X) + 0 = 9D(X) = 9 \times 2.01 = 18.09$$

Задача 3. Случайные величины X и Y независимы. Известно, что $E(X) = -1$, $E(Y) = 3$, $D(X) = 2$, $D(Y) = 5$. Найдите математическое ожидание и дисперсию случайной величины W:
(a) $W = 3X + 2Y$;
(b) $W = 3Y + 6$;
(c) $W = X + 2Y - 5$;
(d) $W = -4X - Y - 3$.

1. Воспользуемся свойствами математического ожидания:

$$E(W) = E(3X + 2Y) = E(3X) + E(2Y) = 3E(X) + 2E(Y) = 3 \times (-1) + 2 \times 3 = 3$$

Воспользуемся свойствами дисперсии:

$$D(W) = D(3X) + E(2Y) = 9D(X) + 4E(Y) = 9 \times 2 + 4 \times 5 = 38$$

2. Воспользуемся свойствами математического ожидания:

$$E(W) = E(3Y + 6) = E(3Y) + E(6) = 3E(Y) + 6 = 3 \times 3 + 6 = 15$$

Воспользуемся свойствами дисперсии:

$$D(W) = D(3Y) + 0 = 9D(Y) = 9 \times 5 = 45$$

3. Воспользуемся свойствами математического ожидания:

$$E(W) = E(X + 2Y - 5) = E(X) + 2E(Y) - 5 = (-1) + 6 - 5 = 0$$

Воспользуемся свойствами дисперсии:

$$D(W) = D(X) + 4D(Y) + 0 = 2 + 4 \times 5 = 22$$

4. Воспользуемся свойствами математического ожидания:

$$E(W) = E(-4X - Y - 3) = -4E(X) - E(Y) - 3 = (-4) \times (-1) - 3 - 3 = 4 - 6 = -2$$

Воспользуемся свойствами дисперсии:

$$D(W) = D(-4X) + D(-Y) + 0 = (-4)^2 D(X) + (-1)^2 D(Y) = 16 \times 2 + 5 = 37$$

Задача 4. Совместное распределение двух случайных величин X и Y задано таблицей:

X / Y	-2	0	2
0.2	0.03	0.05	**0.12**
0.6	0.15	0.3	0.35

(a) Выпишите маргинальные распределения случайных величин X и Y.
(b) Рассчитайте математические ожидания случайных величин X и Y.
(c) Рассчитайте дисперсии случайных величин X и Y.
(d) Являются ли случайные величины X и Y независимыми?
(e) Выпишите ряд распределения вероятностей для случайной величины $X \times Y$.
(f) Рассчитайте $Cov(X, Y)$.
(g) Рассчитайте $Corr(X, Y)$. Проинтерпретируйте результаты.

Решение. Сначала восстановим пропущенную вероятность:

$$P(X = 0.2 \cap Y = 2) = 1 - (0.15 + 0.3 + 0.35) - (0.03 + 0.05) = 1 - 0.8 - 0.08 = 0.12$$

1. Маргинальное распределение случайной величины X можно найти, просуммировав вероятности по строчкам:

X	0.2	0.6
P	0.2	0.8

Маргинальное распределение случайной величины Y можно найти, просуммировав вероятности по колонкам:

Y	-2	0	2
P	0.18	0.35	0.47

2. Найдем математические ожидания, используя маргинальные распределения:

$$E(X) = 0.2 \times 0.2 + 0.6 \times 0.8 = 0.04 + 0.48 = 0.52$$

$$E(Y) = (-2) \times 0.18 + 2 \times 0.47 = -0.36 + 0.94 = 0.58$$

3. Найдем дисперсию случайной величины X. Сначала найдем математическое ожидание случайной величины X^2, возведя каждое из значений случайной величины X в квадрат и умножив на соответствующую вероятность:

$$E(X^2) = (0.2)^2 \times 0.2 + (0.6)^2 \times 0.8 = 0.008 + 0.288 = 0.296$$

$$D(X) = 0.296 - (0.52)^2 = 0.296 - 0.2704 = 0.0256$$

Найдем дисперсию случайной величины Y:

$$E(Y^2) = (-2)^2 \times 0.18 + 2^2 \times 0.47 = 4 \times (0.18 + 0.47) = 2.6$$

$$D(Y) = 2.6 - (0.58)^2 = 2.6 - 0.3364 = 2.2636$$

4. Проверим случайные величины X и Y на независимость. Рассмотрим $P(X = 0.2 \cap Y = (-2)) = 0.03$. Произведение маргинальных вероятностей равно $P(X = 0.2) \times P(Y = (-2)) = 0.2 \times 0.18 = 0.36$. Видно, что

$$P(X = 0.2 \cap Y = (-2)) \neq P(X = 0.2) \times P(Y = (-2))$$

. Следовательно, случайные величины являются зависимыми.

5. Выпишем ряд распределения случайной величины $X \times Y$. Значения этой случайной величины получаются последовательным попарным перемножением каждого значения случайной величины X на каждое значение случайной величины Y. Вероятности этой случайной величины мы получаем из таблицы совместного распределения (дана в условии): например, чтобы найти $P(XY = 1.2)$ нам нужно найти вероятность пересечения тех значений X и Y, произведение которых даст $XY = 1.2$: $P(X = 0.6 \cap Y = 2) = 0.35$.

XY	-1.2	-0.4	0	0.4	1.2
P	0.15	0.03	0.35	0.12	0.35

6. Вспомним, как расчитывается ковариация – мера линейной связи двух случайных величин:

$$Cov(X, Y) = E(XY) - E(X)E(Y)$$

Нам нужно найти $E(XY)$, и сделать мы это можем, основываясь на таблице распределения случайной величины XY из предыдущего пункта:

$$E(XY) = (-1.2) \times 0.15 + (-0.4) \times 0.03 + 0.4 \times 0.12 + 1.2 \times 0.35 =$$

$$-0.18 - 0.012 + 0.048 + 0.42 = 0.276$$

$$Cov(X, Y) = 0.276 - 0.52 \times 0.58 = 0.276 - 0.3016 = -0.0256$$

7. Чтобы найти корреляцию между случайными величинами X и Y, нам нужно посчитать стандатрные отклонения каждой из величин.

$$SD(X) = \sqrt{D(X)} = \sqrt{0.0256} = 0.16$$

$$SD(Y) = \sqrt{D(Y)} = \sqrt{2.2636} = 1.5$$

$$Corr(X,Y) = \frac{Cov(X,Y)}{SD(X) \times SD(Y)} = \frac{-0.0256}{0.16 \times 1.5} = -0.107$$

Полученное значение корреляции говорит о *слабой обратной линейной взаимосвязи* между случайными величинами.