Random variables. Expected value and variance

Probability theory

Anton Afanasev

Higher School of Economics

DSBA 221
September 30, 2023

Seminar Overview

Quiz

1. Three shooters fired their guns, with two bullets hitting the target. Find the probability that the third shooter has hit the target if the probabilities of hitting the target by the first, second and third shooters are 0.6, 0.5, and 0.4, respectively.
2. If a coin with probability of tail equal to $2/3$ is tossed 5 times, find the probability of at least 4 heads.

Random variables

Definition

Random variable is any measurable function $X : \Omega \to \mathbb{R}$

Example

Let X be a number of heads in two independent fair coin tosses.

- Sample space:
$$\Omega = \{HH, HT, TH, TT\}.$$
- Values of X:
$$X(HH) = 2, \quad X(HT) = 1, \quad X(TH) = 1, \quad X(TT) = 0.$$

- Measurable $\iff$ distribution function exists.

Cumulative distribution function

Definition

Cumulative distribution function c.d.f. (or distribution function) F_X of random variable X:

$$\forall x \in \mathbb{R}: \quad F_X(x) = \mathsf{P}(X \leq x) = \mathsf{P}\left(\{\omega \in \Omega : X(\omega) \leq x\}\right).$$

- Properties of c.d.f.:

1. $F_X(x) \in [0, 1]$,
2. $F_X(x)$ is monotonically non-decreasing,
3. $F_X(x)$ is right-continuous,
4. $\lim\limits_{x \to +\infty} F_X(x) = 1$,
5. $\lim\limits_{x \to -\infty} F_X(x) = 0$.

Probability mass function

- Discrete case:

$$X: \quad \Omega = \{\omega_1, \dots, \omega_n\} \longrightarrow \{x_1, \dots, x_k\}, \quad 1 \leq k \leq n.$$

- Probability mass function p.m.f. of X:

$$\forall i \in \overline{1,k}: \quad \mathsf{P}(X = x_i) = \sum_{j:\ X(\omega_j) = x_j}^{n} \mathsf{P}(\omega_j).$$

Example

Let X be a number of heads in two independent fair coin tosses.

- P.m.f. of X:

x_i	0	1	2
$\mathsf{P}(X = x_i)$	$\mathsf{P}(TT) = 1/4$	$\mathsf{P}(TH) + \mathsf{P}(HT) = 1/2$	$\mathsf{P}(HH) = 1/4$

P.m.f. of the example with 2 fair coin tosses

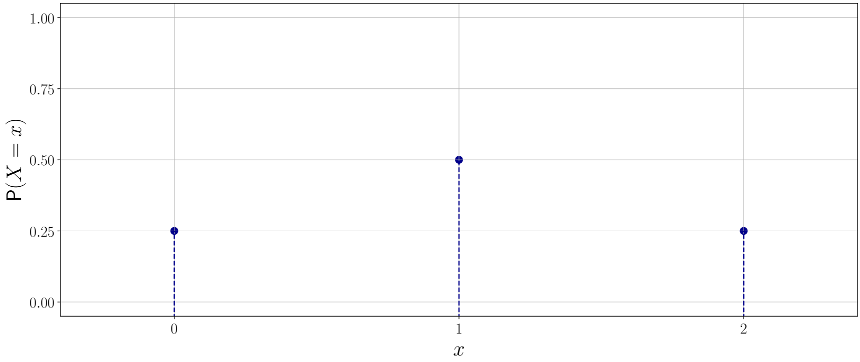

Figure: Probability mass function of X.

C.d.f. of the example with 2 fair coin tosses

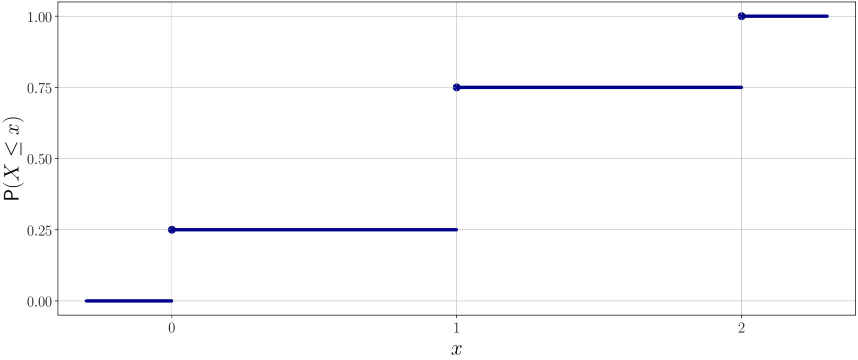

Figure: Cumulative distribution function of X.

Expected value
Discrete

- Discrete case:

$$X: \quad \Omega = \{\omega_1, \ldots, \omega_n\} \longrightarrow \{x_1, \ldots, x_k\}, \quad 1 \leq k \leq n.$$

- Expected value of X:

$$\mathsf{E}(X) = \sum_{j=1}^{n} X(\omega_j) \cdot \mathsf{P}(\omega_j) \overset{\text{linearity}}{=} \sum_{i=1}^{k} x_i \cdot \mathsf{P}(X = x_i).$$

- Applying arbitrary function $f(X)$.
  Law of the unconscious statistician (LOTUS):

$$\mathsf{E}(f(X)) = \sum_{j=1}^{n} f(X(\omega_j)) \cdot \mathsf{P}(\omega_j) \overset{\text{linearity}}{=} \sum_{i=1}^{k} f(x_i) \cdot \mathsf{P}(X = x_i).$$

Expected value
Discrete

Example

$$\mathsf{E}\left(X^2\right) = \sum_{i=1}^{k} x_i^2 \cdot \mathsf{P}(X = x_i),$$

$$\mathsf{E}(\sin X) = \sum_{i=1}^{k} \sin x_i \cdot \mathsf{P}(X = x_i).$$

- $\mathsf{E}(X)$ is linear:

$$\mathsf{E}(c_1 X_1 + c_2 X_2) = c_1 \mathsf{E}(X_1) + c_2 \mathsf{E}(X_2).$$

- The product is separable only in case of factors' independence:

$$\mathsf{E}(X_1 X_2) \overset{\text{independence}}{=} \mathsf{E}(X_1) \cdot \mathsf{E}(X_2).$$

Variance and standard deviation

- Variance of X:

$$\mathsf{V}(X) = \mathsf{E}\left[(X - \mathsf{E}(X))^2\right] = \mathsf{E}\left(X^2\right) - \mathsf{E}\left(X\right)^2.$$

- $\mathsf{V}(X)$ is quadratic:

$$\mathsf{V}(cX) = c^2\mathsf{V}(X).$$

- The sum is separable only in case of terms' independence:

$$\mathsf{V}(X_1 + X_2) \overset{\text{independence}}{=} \mathsf{V}(X_1) + \mathsf{V}(X_2).$$

- Standard deviation of X:

$$\sigma(X) = \sqrt{\mathsf{V}(X)}.$$

St. dev. is often used instead of variance due to its reasonable unit of measurement: $[X] = [\mathsf{E}(X)] = [\sigma(X)]$.

Variance and standard deviation

- Some identities:

$$\mathsf{V}(c) = 0,$$
$$\mathsf{V}(aX + b) = a^2\mathsf{V}(X),$$
$$\sigma(aX + b) = |a|\sigma(X),$$

where a, b and c are constants.

Problem 1

A number is chosen at random from set $S = \{-1, 0, 1\}$. Let X be the number chosen. Find the expected value, variance, and standard deviation of X.

Problem 2

A random variable X has the following distribution:

X	0	1	2	4
P_X	$1/3$	$1/3$	$1/6$	$1/6$

Find $\mathsf{E}(X)$, $\mathsf{E}\left(X(X+1)\right)$, $\mathsf{V}(X)$ and $\sigma(X)$.

Problem 3

X is a random variable with $\mathsf{E}(X) = 100$ and $\mathsf{V}(X) = 15$. Find

1. $\mathsf{E}\left(X^2\right)$,
2. $\mathsf{E}\left(3X + 10\right)$,
3. $\mathsf{E}\left(-X\right)$,
4. $\mathsf{V}\left(-X\right)$,
5. $\sigma\left(-X\right)$.

Problem 4

A coin is tossed three times. Let X be the number of heads that turn up. Find $\mathsf{V}(X)$ and $\sigma(X)$.

- Bernoulli distribution of $X_1 \quad \Longleftrightarrow \quad X_1 \sim \text{Bernoulli}(p)$.
  X_1 shows the number of successes in one Bernoulli trial with probability of success p.

$$\mathsf{P}(X_1 = k) = \begin{cases} p, & k = 1, \\ 1 - p, & k = 0. \end{cases} \qquad \mathsf{E}(X_1) = p, \quad \mathsf{V}(X_1) = p(1-p).$$

- Binomial distribution of $X \quad \Longleftrightarrow \quad X \sim \text{Bin}(n, p)$.
  X shows the number of successes in n Bernoulli trials with probability of success p.

$$\mathsf{P}(X = k) = C_n^k p^k (1-p)^{n-k}. \qquad \mathsf{E}(X) = np, \quad \mathsf{V}(X) = np(1-p).$$

Problem 5

A box contains 10 white balls and 2 black balls. 6 balls are selected at random. Random variable X is equal to number of black balls in the 6 selected.

1. Find the distribution of random variable X.
2. Find expected value $\mathsf{E}(X)$.
3. Find expected value of X, given that you were told that $X > 0$ $\mathsf{E}\left(X \mid X > 0\right)$.

- Geometric distribution of $G \iff G \sim \mathrm{Geom}(p)$.
  G shows the number of Bernoulli trials needed to get one success, while probability of success is p.

$$\mathsf{P}(G = k) = (1-p)^{k-1}p. \qquad \mathsf{E}(G) = \frac{1}{p}, \quad \mathsf{V}(G) = \frac{1-p}{p^2}.$$

- Hypergeometric distribution of $H \iff H \sim \mathrm{Hypergeom}(N, K, n)$.
  H shows the number of successes in n draws without replacement from a population with size N, which contains K “successful” objects.

$$\mathsf{P}(H = k) = \frac{C_K^k C_{N-K}^{n-k}}{C_N^n}, \quad \mathsf{E}(H) = \frac{nK}{N}, \quad \mathsf{V}(H) = \frac{nK}{N} \cdot \frac{N-K}{N} \cdot \frac{N-n}{N-1}.$$

That's all Folks