Mathematics for Machine Learning

Vazgen Mikayelyan

September 5, 2020

FAST DISCOVERING THE FUTURE

Trials

Definition

A fixed number of repetitions of the same experiment can be thought of as one compound experiment. In this case, the individual repetitions are called **trials.**

Trials

Definition

A fixed number of repetitions of the same experiment can be thought of as one compound experiment. In this case, the individual repetitions are called **trials.**

Definition

If the outcomes of any group of the trials do not affect the probabilities of the outcomes of any other trial, then the trials are called **independent trials.**

Trials

Bernoulli Trials

Suppose there are only two possible outcomes in the sample space of each trial classified as "Success" and "Failure". Such trials are called **Bernoulli trials.** If moreover, the probability p of Success remains the same for each trial throughout the entire experiment, then such trials are called **repeated** Bernoulli trials.

Trials

Bernoulli Trials

Suppose there are only two possible outcomes in the sample space of each trial classified as "Success" and "Failure". Such trials are called **Bernoulli trials.** If moreover, the probability p of Success remains the same for each trial throughout the entire experiment, then such trials are called **repeated** Bernoulli trials.

Binomial Experiment

Let n be some fixed positive integer. An experiment consisting of n repeated Bernoulli trials is called a **Binomial Experiment.**

Examples of Binomial experiments

Examples of Binomial experiments

- Tossing a fair coin 10 times:
  Success = Heads, Failure = Tails
  $p = 0.5$

Examples of Binomial experiments

- Tossing a fair coin 10 times:
  Success = Heads, Failure = Tails
  p = 0.5
- Rolling a fair die 5 times:
  Success = opened number is even,
  Failure = the other numbers
  $p = \frac{3}{6} = \frac{1}{2}.$

Binomial Probabilities

Suppose a Binomial Experiment has n trials. Let X be the number of successes in a sequence of n trials. Then for $0 \le k \le n$,

$$p_X(k) = \mathbb{P}(X = k) =$$

Binomial Probabilities

Suppose a Binomial Experiment has n trials. Let X be the number of successes in a sequence of n trials. Then for $0 \leq k \leq n$,

$$p_X(k) = \mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

We denote this fact as $X \sim B(n, p)$.

Binomial Probabilities

Suppose a Binomial Experiment has n trials. Let X be the number of successes in a sequence of n trials. Then for $0 \leq k \leq n$,

$$p_X(k) = \mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

We denote this fact as $X \sim B(n, p)$.

Example

We are tossing a fair coin 100 times. The success is "Heads".
1) What is the probability of being successful exactly 5 times?
2) What is the probability that we will be successful at least once?

Example

An examine consists of 16 multiple-choice questions. Each question provides 4 answers, where only 1 is correct. A student, who is absolutely unfamiliar with the subject, chooses 16 answers randomly.

1) What is the probability that the student will answer correctly at least once?

2) The exam is passed if at least 3 questions are answered correctly. What is the probability that the student above will pass the exam?

Example

An examine consists of 16 multiple-choice questions. Each question provides 4 answers, where only 1 is correct. A student, who is absolutely unfamiliar with the subject, chooses 16 answers randomly.
1) What is the probability that the student will answer correctly at least once?
2) The exam is passed if at least 3 questions are answered correctly. What is the probability that the student above will pass the exam?

Example

A communication system consists of 5 components, each of which will, independently, function with probability p. The total system will be able to operate effectively if at least one-half of its components function. Find the probability that a 5-component system will be effective.

Multinomial Probabilities

Suppose an experiment has n independent trials and r possible outcomes:

$$A_1, A_2, \ldots, A_r.$$

If A is the event that in n trials the event A_i will occur k_i times for all $i = 1, 2, \ldots, r$ and $k_1 + k_2 + \ldots + k_r = n$, then

$$\mathbb{P}(A) = \frac{n!}{k_1! k_2! \ldots k_r!} p_1^{k_1} p_2^{k_2} \ldots p_r^{k_r},$$

where $\mathbb{P}(A_i) = p_i$, for all $i = 1, 2, \ldots, r$.

Expected Value

Definition

The expectation, or the expected value, of discrete r.v. X, denoted by $\mathbb{E}[X]$, is defined by

$$\mathbb{E}[X] = \sum_{x:p(x)>0} xp(x)$$

Expected Value

Definition

The expectation, or the expected value, of discrete r.v. X, denoted by $\mathbb{E}[X]$, is defined by

$$\mathbb{E}[X] = \sum_{x:p(x)>0} xp(x)$$

Proposition

If X is a discrete random variable that takes on one of the values $x_i, i \geq 1$, with respective probabilities $p(x_i)$, then,

$$\mathbb{E}[X] = \sum_n x_n p(x_n).$$

moreover, for any real-valued function g,

$$\mathbb{E}[g(X)] = \sum_n g(x_n)p(x_n).$$

Expected Value

Definition

The expectation, or the expected value, of continuous r.v. X*, denoted by* $\mathbb{E}[X]$*, is defined by*

$$\mathbb{E}[X] = \int_{-\infty}^{+\infty} x f_X(x)\, dx.$$

Expected Value

Definition

The expectation, or the expected value, of continuous r.v. X*, denoted by* $\mathbb{E}[X]$*, is defined by*

$$\mathbb{E}[X] = \int_{-\infty}^{+\infty} x f_X(x)\, dx.$$

Proposition

If X *is a continuous random variable, then for any real-valued continuous function* g*,*

$$\mathbb{E}[g(X)] = \int_{-\infty}^{+\infty} g(x) f_X(x)\, dx.$$

Expected Value

In words, the expected value of discrete random variable X is a weighted average of the possible values that X can take on, each value being weighted by the probability that X assumes it.

Expected Value

In words, the expected value of discrete random variable X is a weighted average of the possible values that X can take on, each value being weighted by the probability that X assumes it.

Example

Find $E[X]$ where X is the outcome when we roll a fair die.

Properties

1. If $X(\omega) = I_A(\omega)$ is the indicator function for the event A, then $E[X] = \mathbb{P}(A)$.

Expected Value

Properties

1. If $X(\omega) = I_A(\omega)$ is the indicator function for the event A, then $E[X] = \mathbb{P}(A)$.
2. If a and b are constants, then

$$E[aX + b] = aE[X] + b.$$

Expected Value

Properties

1. If $X(\omega) = I_A(\omega)$ is the indicator function for the event A, then $E[X] = \mathbb{P}(A)$.
2. If a and b are constants, then

$$E[aX + b] = aE[X] + b.$$

3. $E[X + Y] = E[X] + E[Y]$

Expected Value

Properties

1. If $X(\omega) = I_A(\omega)$ is the indicator function for the event A, then $E[X] = \mathbb{P}(A)$.
2. If a and b are constants, then

$$E[aX + b] = aE[X] + b.$$

3. $E[X + Y] = E[X] + E[Y]$
4. $E[X_1 + \ldots + X_n] = E[X_1] + \ldots + E[X_n]$

Expected Value

Properties

1. If $X(\omega) = I_A(\omega)$ is the indicator function for the event A, then $E[X] = \mathbb{P}(A)$.
2. If a and b are constants, then
$$E[aX + b] = aE[X] + b.$$
3. $E[X + Y] = E[X] + E[Y]$
4. $E[X_1 + \ldots + X_n] = E[X_1] + \ldots + E[X_n]$
5. If $X \geq Y$, then $E[X] \geq E[Y]$.

Expected Value

Definition

Random variables X, Y *are called independent, if for all* $x, y \in \mathbb{R}$ *holds*

$$\mathbb{P}\left(X \leq x \cap Y \leq y\right) = F_X\left(x\right) F_Y\left(y\right).$$

Expected Value

Definition

Random variables X, Y are called independent, if for all $x, y \in \mathbb{R}$ holds

$$\mathbb{P}\left(X \leq x \cap Y \leq y\right) = F_X\left(x\right) F_Y\left(y\right).$$

Proposition

If X and Y are independent RVs, then $E\left[XY\right] = E\left[X\right] E\left[Y\right]$.

Variance

Definition

If X is a RV, then the **variance** *of X, denoted by $Var(X)$, is defined by*

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

Variance

Definition

If X is a RV, then the **variance** *of X, denoted by $Var(X)$, is defined by*

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

Definition

The square root of the $Var\left(X\right)$ is called the standard deviation of X.

Variance

Definition

If X is a RV, then the **variance** *of X, denoted by $Var(X)$, is defined by*

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

Definition

The square root of the $Var(X)$ is called the standard deviation of X.

Properties

1. $Var(X) = \mathbb{E}\left[X^2\right] - (\mathbb{E}[X])^2$

Variance

Definition

If X is a RV, then the **variance** *of X, denoted by $Var(X)$, is defined by*

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

Definition

The square root of the $Var\left(X\right)$ is called the standard deviation of X.

Properties

1. $Var\left(X\right) = \mathbb{E}\left[X^2\right] - \left(\mathbb{E}\left[X\right]\right)^2$
2. $Var\left(X\right) \geq 0$

Variance

Definition

If X is a RV, then the **variance** *of X, denoted by $Var(X)$, is defined by*

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

Definition

The square root of the $Var\left(X\right)$ is called the standard deviation of X.

Properties

1. $Var\left(X\right) = \mathbb{E}\left[X^2\right] - \left(\mathbb{E}\left[X\right]\right)^2$
2. $Var\left(X\right) \geq 0$
3. If X and Y are independent random variables, then

$$Var\left(X + Y\right) = Var\left(X\right) + Var\left(Y\right).$$

Variance

Example

Calculate $Var(X)$ if X represents the outcome when a fair die is rolled.

Variance

Example

Calculate $Var(X)$ if X represents the outcome when a fair die is rolled.

Properties of Binomial Random Variables

If X is a binomial RV with parameters n and p, then

$$\mathbb{E}[X] = np, Var\left(X\right) = np\left(1-p\right).$$

Expected Value and Variance

$\mathbb{E}[X] =$

Expected Value and Variance

$$\mathbb{E}[X] = \sum_{k=1}^{n} k C_n^k p^k \left(1 - p\right)^{n-k}$$

Expected Value and Variance

$$\mathbb{E}[X] = \sum_{k=1}^{n} kC_n^k p^k (1-p)^{n-k} = np \sum_{k=1}^{n} C_{n-1}^{k-1} p^{k-1} (1-p)^{n-k} =$$

Expected Value and Variance

$$\mathbb{E}[X] = \sum_{k=1}^{n} k C_n^k p^k (1-p)^{n-k} = np \sum_{k=1}^{n} C_{n-1}^{k-1} p^{k-1} (1-p)^{n-k} = np$$

Expected Value and Variance

$$\mathbb{E}[X] = \sum_{k=1}^{n} k C_n^k p^k (1-p)^{n-k} = np \sum_{k=1}^{n} C_{n-1}^{k-1} p^{k-1} (1-p)^{n-k} = np$$

$$\mathbb{E}[X^2] - \mathbb{E}[X] =$$

Expected Value and Variance

$$\mathbb{E}[X] = \sum_{k=1}^{n} k C_n^k p^k (1-p)^{n-k} = np \sum_{k=1}^{n} C_{n-1}^{k-1} p^{k-1} (1-p)^{n-k} = np$$

$$\mathbb{E}[X^2] - \mathbb{E}[X] = \sum_{k=1}^{n} k(k-1) C_n^k p^k (1-p)^{n-k}$$

Expected Value and Variance

$$\mathbb{E}[X] = \sum_{k=1}^{n} k C_n^k p^k (1-p)^{n-k} = np \sum_{k=1}^{n} C_{n-1}^{k-1} p^{k-1} (1-p)^{n-k} = np$$

$$\mathbb{E}[X^2] - \mathbb{E}[X] = \sum_{k=1}^{n} k(k-1) C_n^k p^k (1-p)^{n-k}$$

$$= n(n-1) p^2 \sum_{k=2}^{n} C_{n-2}^{k-2} p^{k-2} (1-p)^{n-k} = n(n-1) p^2.$$

Expected Value and Variance

$$\mathbb{E}[X] = \sum_{k=1}^{n} kC_n^k p^k (1-p)^{n-k} = np \sum_{k=1}^{n} C_{n-1}^{k-1} p^{k-1} (1-p)^{n-k} = np$$

$$\mathbb{E}[X^2] - \mathbb{E}[X] = \sum_{k=1}^{n} k(k-1) C_n^k p^k (1-p)^{n-k}$$

$$= n(n-1) p^2 \sum_{k=2}^{n} C_{n-2}^{k-2} p^{k-2} (1-p)^{n-k} = n(n-1) p^2.$$

$Var(X) =$

Expected Value and Variance

$$\mathbb{E}[X] = \sum_{k=1}^{n} k C_n^k p^k (1-p)^{n-k} = np \sum_{k=1}^{n} C_{n-1}^{k-1} p^{k-1} (1-p)^{n-k} = np$$

$$\mathbb{E}[X^2] - \mathbb{E}[X] = \sum_{k=1}^{n} k(k-1) C_n^k p^k (1-p)^{n-k}$$

$$= n(n-1)p^2 \sum_{k=2}^{n} C_{n-2}^{k-2} p^{k-2} (1-p)^{n-k} = n(n-1)p^2.$$

$$Var(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 =$$

Expected Value and Variance

$$\mathbb{E}[X] = \sum_{k=1}^{n} kC_n^k p^k (1-p)^{n-k} = np \sum_{k=1}^{n} C_{n-1}^{k-1} p^{k-1} (1-p)^{n-k} = np$$

$$\mathbb{E}[X^2] - \mathbb{E}[X] = \sum_{k=1}^{n} k(k-1) C_n^k p^k (1-p)^{n-k}$$

$$= n(n-1)p^2 \sum_{k=2}^{n} C_{n-2}^{k-2} p^{k-2} (1-p)^{n-k} = n(n-1)p^2.$$

$$Var(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = np + n(n-1)p^2 - (np)^2 = np(1-p).$$

Markov's Inequality

If X is a random variable, then for any value $\varepsilon > 0$

$$\mathbb{P}\left(|X| \geq \varepsilon\right) \leq \frac{E\left[|X|\right]}{\varepsilon}.$$

Markov's Inequality

If X is a random variable, then for any value $\varepsilon > 0$

$$\mathbb{P}\left(|X| \geq \varepsilon\right) \leq \frac{E\left[|X|\right]}{\varepsilon}.$$

Theorem

If X is a random variable, then for any value $\varepsilon > 0$ and $r > 0$

$$\mathbb{P}\left(|X| \geq \varepsilon\right) \leq \frac{E\left[|X|^r\right]}{\varepsilon^r}.$$

Expected Value and Variance

Markov's Inequality

If X is a random variable, then for any value $\varepsilon > 0$

$$\mathbb{P}(|X| \geq \varepsilon) \leq \frac{E[|X|]}{\varepsilon}.$$

Theorem

If X is a random variable, then for any value $\varepsilon > 0$ and $r > 0$

$$\mathbb{P}(|X| \geq \varepsilon) \leq \frac{E[|X|^r]}{\varepsilon^r}.$$

Chebyshev's inequality

If X is a random variable, then for any value $\varepsilon > 0$

$$\mathbb{P}(|X - E[X]| \geq \varepsilon) \leq \frac{Var(X)}{\varepsilon^2}.$$

Poisson Distribution

Definition

A random variable X that takes on one of the values $0, 1, 2, \ldots$ is said to be a Poisson random variable with parameter $\lambda > 0$ if the PMF of X is representable as

$$\mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \ldots.$$

The fact that X is a Poisson RV with parameter λ is denoted as $X \sim Po(\lambda)$ or $X \sim Poiss(\lambda)$.

Poisson Distribution

Definition

A random variable X that takes on one of the values $0, 1, 2, \ldots$ is said to be a Poisson random variable with parameter $\lambda > 0$ if the PMF of X is representable as

$$\mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \ldots.$$

The fact that X is a Poisson RV with parameter λ is denoted as $X \sim Po(\lambda)$ or $X \sim Poiss(\lambda)$.

The Poisson random variable may be used as an approximation for a binomial random variable with parameters (n, p) when n is large and p is small enough so that $n \cdot p$ is of moderate size.

$$\text{If } X \sim B(n, p), \text{ then } P(X = k) \approx \frac{\lambda^k}{k!} e^{-\lambda}, \text{ where } \lambda = np.$$

The following random variables can be considered as Poisson RVs because of their closeness to the binomial.

The following random variables can be considered as Poisson RVs because of their closeness to the binomial.

1. The number of typos on a page of a book.

The following random variables can be considered as Poisson RVs because of their closeness to the binomial.

1. The number of typos on a page of a book.
2. The number of people in a community who survive to age 100.

The following random variables can be considered as Poisson RVs because of their closeness to the binomial.

1. The number of typos on a page of a book.
2. The number of people in a community who survive to age 100.
3. The number of telephone calls per hour received by an office.

The following random variables can be considered as Poisson RVs because of their closeness to the binomial.

1. The number of typos on a page of a book.
2. The number of people in a community who survive to age 100.
3. The number of telephone calls per hour received by an office.
4. The number of car accidents per year in front of YSU.

The following random variables can be considered as Poisson RVs because of their closeness to the binomial.

1. The number of typos on a page of a book.
2. The number of people in a community who survive to age 100.
3. The number of telephone calls per hour received by an office.
4. The number of car accidents per year in front of YSU.
5. The number of broken bottles in a delivery.

The following random variables can be considered as Poisson RVs because of their closeness to the binomial.

1. The number of typos on a page of a book.
2. The number of people in a community who survive to age 100.
3. The number of telephone calls per hour received by an office.
4. The number of car accidents per year in front of YSU.
5. The number of broken bottles in a delivery.

Example

Assume that the number of university professor in Armenia is 400 and the probability of reaching age 100 is 0.005*. Then let's assume that there are only* $n = 400$ *professors. Find the probability that at least 3 professors will reach age 100.*

The following random variables can be considered as Poisson RVs because of their closeness to the binomial.

1. The number of typos on a page of a book.
2. The number of people in a community who survive to age 100.
3. The number of telephone calls per hour received by an office.
4. The number of car accidents per year in front of YSU.
5. The number of broken bottles in a delivery.

Example

Assume that the number of university professor in Armenia is 400 and the probability of reaching age 100 is 0.005. Then let's assume that there are only $n = 400$ professors. Find the probability that at least 3 professors will reach age 100.

Example

Let X be a Poisson random variable with parameter λ. Prove that

$$\mathbb{E}\left[X\right] = \lambda, \quad Var\left(X\right) = \lambda.$$

Expected Value and Variance

$$\mathbb{E}[X] =$$

Expected Value and Variance

$$\mathbb{E}[X] = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} =$$

Expected Value and Variance

$$\mathbb{E}\left[X\right] = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} =$$

Expected Value and Variance

$$\mathbb{E}[X] = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda.$$

Expected Value and Variance

$$\mathbb{E}\left[X\right] = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda.$$

$$\mathbb{E}\left[X^2\right] - \mathbb{E}\left[X\right] =$$

Expected Value and Variance

$$\mathbb{E}\left[X\right] = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda.$$

$$\mathbb{E}\left[X^2\right] - \mathbb{E}\left[X\right] = \sum_{k=0}^{\infty} k\left(k-1\right) \frac{\lambda^k}{k!} e^{-\lambda} =$$

Expected Value and Variance

$$\mathbb{E}\left[X\right] = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda.$$

$$\mathbb{E}\left[X^2\right] - \mathbb{E}\left[X\right] = \sum_{k=0}^{\infty} k\left(k-1\right) \frac{\lambda^k}{k!} e^{-\lambda} = \lambda^2 e^{-\lambda} \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} =$$

Expected Value and Variance

$$\mathbb{E}[X] = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda.$$

$$\mathbb{E}\left[X^2\right] - \mathbb{E}[X] = \sum_{k=0}^{\infty} k(k-1) \frac{\lambda^k}{k!} e^{-\lambda} = \lambda^2 e^{-\lambda} \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} = \lambda^2.$$

Expected Value and Variance

$$\mathbb{E}\left[X\right] = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda.$$

$$\mathbb{E}\left[X^2\right] - \mathbb{E}\left[X\right] = \sum_{k=0}^{\infty} k\left(k-1\right) \frac{\lambda^k}{k!} e^{-\lambda} = \lambda^2 e^{-\lambda} \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} = \lambda^2.$$

$$Var\left(X\right) =$$

Expected Value and Variance

$$\mathbb{E}\left[X\right] = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda.$$

$$\mathbb{E}\left[X^2\right] - \mathbb{E}\left[X\right] = \sum_{k=0}^{\infty} k\left(k-1\right) \frac{\lambda^k}{k!} e^{-\lambda} = \lambda^2 e^{-\lambda} \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} = \lambda^2.$$

$$Var\left(X\right) = \mathbb{E}[X^2] - \left(\mathbb{E}[X]\right)^2 =$$

Expected Value and Variance

$$\mathbb{E}\left[X\right] = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda.$$

$$\mathbb{E}\left[X^2\right] - \mathbb{E}\left[X\right] = \sum_{k=0}^{\infty} k\left(k-1\right) \frac{\lambda^k}{k!} e^{-\lambda} = \lambda^2 e^{-\lambda} \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} = \lambda^2.$$

$$Var\left(X\right) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \lambda + \lambda^2 - \lambda^2 = \lambda.$$