American University of Beirut

STAT 210: Elementary Statistics for Sciences

2022–2023 Fall

Siamak Taati

Chapter 5

Discrete random variables

Part 2

Discrete random variables

Discrete random variables

Example (Flipping a biased coin twice)

$$\boxed{\Omega := \{\mathtt{HH}, \mathtt{HT}, \mathtt{TH}, \mathtt{TT}\}}$$

$$\boxed{\mathbb{P}(\mathtt{HH}) = 0.3025 \qquad \mathbb{P}(\mathtt{HT}) = 0.2475 \qquad \mathbb{P}(\mathtt{TH}) = 0.2475 \qquad \mathbb{P}(\mathtt{TT}) = 0.2025}$$

$N := \#$ of heads

outcome	HH	HT	TH	TT
N	2	1	1	0

Discrete random variables

Example (Flipping a biased coin twice)

$$\boxed{\Omega := \{\mathtt{HH}, \mathtt{HT}, \mathtt{TH}, \mathtt{TT}\}}$$

$$\boxed{\mathbb{P}(\mathtt{HH}) = 0.3025 \qquad \mathbb{P}(\mathtt{HT}) = 0.2475 \qquad \mathbb{P}(\mathtt{TH}) = 0.2475 \qquad \mathbb{P}(\mathtt{TT}) = 0.2025}$$

$N :=$ # of heads

outcome	HH	HT	TH	TT
N	2	1	1	0

Q What is the mean value of N?

Discrete random variables

Example (Flipping a biased coin twice)

$$\boxed{\Omega := \{\mathtt{HH}, \mathtt{HT}, \mathtt{TH}, \mathtt{TT}\}}$$

$$\boxed{\mathbb{P}(\mathtt{HH}) = 0.3025 \qquad \mathbb{P}(\mathtt{HT}) = 0.2475 \qquad \mathbb{P}(\mathtt{TH}) = 0.2475 \qquad \mathbb{P}(\mathtt{TT}) = 0.2025}$$

$N := \#$ of heads

outcome	HH	HT	TH	TT
N	2	1	1	0

Q What is the mean value of N?

A1 We have

$$\begin{aligned}\mathbb{E}[N] &= \mathbb{P}(\mathtt{HH}) \times 2 + \mathbb{P}(\mathtt{HT}) \times 1 + \mathbb{P}(\mathtt{TH}) \times 1 + \mathbb{P}(\mathtt{TT}) \times 0 \\ &= 0.3025 \times 2 + 0.2475 \times 1 + 0.2475 \times 1 + 0.2025 \times 0 \\ &= \boxed{1.1}\end{aligned}$$

Discrete random variables

Example (Flipping a biased coin twice)

$$\boxed{\Omega := \{\mathtt{HH}, \mathtt{HT}, \mathtt{TH}, \mathtt{TT}\}}$$

$\mathbb{P}(\mathtt{HH}) = 0.3025$ $\mathbb{P}(\mathtt{HT}) = 0.2475$ $\mathbb{P}(\mathtt{TH}) = 0.2475$ $\mathbb{P}(\mathtt{TT}) = 0.2025$

$N := \#$ of heads

outcome	HH	HT	TH	TT
N	2	1	1	0

Q What is the mean value of N?

A2 Using the distribution of N, we have

$$\begin{aligned}
\mathbb{E}[N] &= \mathbb{P}(N=0) \times 0 + \mathbb{P}(N=1) \times 1 + \mathbb{P}(N=2) \times 2 \\
&= 0.2025 \times 0 + 0.4950 \times 1 + 0.3025 \times 2 \\
&= \boxed{1.1}
\end{aligned}$$

k	$\mathbb{P}(N=k)$
0	0.2025
1	0.4950
2	0.3025

Discrete random variables

Example (Rolling two dice)

$$\boxed{\Omega := \{⚀⚀, ⚀⚁, ⚀⚂, \ldots, ⚅⚅\}}$$

$$\boxed{\mathbb{P}(⚀⚀) = \mathbb{P}(⚀⚁) = \mathbb{P}(⚀⚂) = \cdots = \mathbb{P}(⚅⚅) = 1/36}$$

$X :=$ sum of the numbers on the dice

Discrete random variables

Example (Rolling two dice)

$$\boxed{\Omega := \{\text{⚀⚀}, \text{⚀⚁}, \text{⚀⚂}, \dots, \text{⚅⚅}\}}$$

$$\boxed{\mathbb{P}(\text{⚀⚀}) = \mathbb{P}(\text{⚀⚁}) = \mathbb{P}(\text{⚀⚂}) = \cdots = \mathbb{P}(\text{⚅⚅}) = {}^{1}\!/_{36}}$$

$$X := \text{sum of the numbers on the dice}$$

Q What is the mean value of X?

Discrete random variables

Example (Rolling two dice)

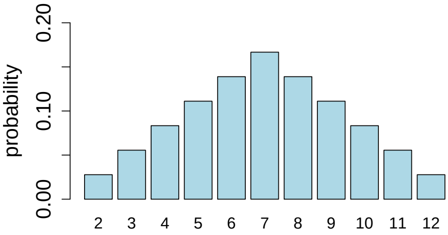
$$\boxed{\Omega := \{⚀⚀, ⚀⚁, ⚀⚂, \ldots, ⚅⚅\}}$$

$$\boxed{\mathbb{P}(⚀⚀) = \mathbb{P}(⚀⚁) = \mathbb{P}(⚀⚂) = \cdots = \mathbb{P}(⚅⚅) = 1/36}$$

$$X := \text{sum of the numbers on the dice}$$

Q What is the mean value of X?

A1 Since the distribution of X is symmetric around 7, $\mathbb{E}[X] = \boxed{7}$.



Discrete random variables

Example (Rolling two dice)

$$\boxed{\Omega := \{⚀⚀, ⚀⚁, ⚀⚂, \ldots, ⚅⚅\}}$$

$$\boxed{\mathbb{P}(⚀⚀) = \mathbb{P}(⚀⚁) = \mathbb{P}(⚀⚂) = \cdots = \mathbb{P}(⚅⚅) = 1/36}$$

$$X := \text{sum of the numbers on the dice}$$

Q What is the mean value of X?

A2 We have

$$\mathbb{E}[X] = \mathbb{P}(⚀⚀) \times 2 + \mathbb{P}(⚀⚁) \times 3 + \mathbb{P}(⚀⚂) \times 4 + \cdots + \mathbb{P}(⚅⚅) \times 12$$

$$= \frac{\overbrace{2 + 3 + 4 + \cdots + 12}^{\text{values for each of the 36 possible outcomes}}}{36}$$

$$= \frac{2 + 2 \times 3 + 3 \times 4 + 4 \times 5 + 5 \times 6 + 6 \times 7 + 5 \times 8 + 4 \times 9 + 3 \times 10 + 2 \times 11 + 12}{36}$$

$$= \boxed{7}$$

Discrete random variables

Example (Rolling two dice)

$$\boxed{\Omega := \{⚀⚀, ⚀⚁, ⚀⚂, \ldots, ⚅⚅\}}$$

$$\boxed{\mathbb{P}(⚀⚀) = \mathbb{P}(⚀⚁) = \mathbb{P}(⚀⚂) = \cdots = \mathbb{P}(⚅⚅) = 1/36}$$

$X :=$ sum of the numbers on the dice

Q What is the mean value of X?

A3 Using the distribution of X, we have

$$\begin{aligned}\mathbb{E}[X] &= \mathbb{P}(X=2)\times 2 + \mathbb{P}(X=3)\times 3 + \mathbb{P}(X=4)\times 4 + \cdots \\ &\qquad + \mathbb{P}(X=12)\times 12 \\ &= \frac{1}{36}\times 2 + \frac{2}{36}\times 3 + \frac{3}{36}\times 4 + \cdots + \frac{1}{36}\times 12 \\ &= \boxed{7}\end{aligned}$$

a	$\mathbb{P}(X=a)$
2	1/36
3	2/36
4	3/36
5	4/36
6	5/36
7	6/36
8	5/36
9	4/36
10	3/36
11	2/36
12	1/36

Discrete random variables

Example (Rolling two dice)

$$\boxed{\Omega := \{\text{⚀⚀}, \text{⚀⚁}, \text{⚀⚂}, \ldots, \text{⚅⚅}\}}$$

$$\boxed{\mathbb{P}(\text{⚀⚀}) = \mathbb{P}(\text{⚀⚁}) = \mathbb{P}(\text{⚀⚂}) = \cdots = \mathbb{P}(\text{⚅⚅}) = 1/36}$$

$$X := \text{sum of the numbers on the dice}$$

Q What is the mean value of X?

A4 Note that

- The mean of the number on the 1st die is 3.5.
- The mean of the number on the 2nd die is 3.5.

Therefore,

- The mean of the sum of the two numbers is $3.5 + 3.5 = \boxed{7}$.

Discrete random variables

Example (Rolling two dice)

$$\boxed{\Omega := \{⚀⚀, ⚀⚁, ⚀⚂, \ldots, ⚅⚅\}}$$

$$\boxed{\mathbb{P}(⚀⚀) = \mathbb{P}(⚀⚁) = \mathbb{P}(⚀⚂) = \cdots = \mathbb{P}(⚅⚅) = 1/36}$$

$X :=$ sum of the numbers on the dice

Q What is the **mean** value of X?

$\boxed{\text{A4}}$ More precisely,

$$X = X_1 + X_2$$

$X_1 :=$ number on the 1st die

$X_2 :=$ number on the 2nd die

Now, $\mathbb{E}[X_1] = \mathbb{E}[X_2] = 3.5$, hence

$$\mathbb{E}[X] = \mathbb{E}[X_1] + \mathbb{E}[X_2] = 3.5 + 3.5 = \boxed{7}$$

Expected value of discrete RVs

The expected value (a.k.a. the mean) of a discrete RV X is the average value of X over all the possible outcomes of the experiment weighted by their probabilities, and is denoted by $\mathbb{E}[X]$ or μ.

It is interpreted as the (approximate) mean of the values of X in many many repeated experiments.

Expected value of discrete RVs

The expected value (a.k.a. the mean) of a discrete RV X is the average value of X over all the possible outcomes of the experiment weighted by their probabilities, and is denoted by $\mathbb{E}[X]$ or μ.

It is interpreted as the (approximate) mean of the values of X in many many repeated experiments.

As in the case of the mean of a variable in a data set, the expected value of a RV X can be calculated in two ways:

Expected value of discrete RVs

The expected value (a.k.a. the mean) of a discrete RV X is the average value of X over all the possible outcomes of the experiment weighted by their probabilities, and is denoted by $\mathbb{E}[X]$ or μ.

It is interpreted as the (approximate) mean of the values of X in many many repeated experiments.

As in the case of the mean of a variable in a data set, the expected value of a RV X can be calculated in two ways:

1. If the experiment has possible outcomes $\omega_1, \omega_2, \omega_3, \ldots$, then

$$\boxed{\mathbb{E}[X] = \mathbb{P}(\omega_1) \cdot X(\omega_1) + \mathbb{P}(\omega_2) \cdot X(\omega_2) + \mathbb{P}(\omega_3) \cdot X(\omega_3) + \cdots}$$

[similar to using the raw data]

Expected value of discrete RVs

The expected value (a.k.a. the mean) of a discrete RV X is the average value of X over all the possible outcomes of the experiment weighted by their probabilities, and is denoted by $\mathbb{E}[X]$ or μ.

It is interpreted as the (approximate) mean of the values of X in many many repeated experiments.

As in the case of the mean of a variable in a data set, the expected value of a RV X can be calculated in two ways:

1. If the experiment has possible outcomes $\omega_1, \omega_2, \omega_3, \ldots$, then

$$\boxed{\mathbb{E}[X] = \mathbb{P}(\omega_1) \cdot X(\omega_1) + \mathbb{P}(\omega_2) \cdot X(\omega_2) + \mathbb{P}(\omega_3) \cdot X(\omega_3) + \cdots}$$

[similar to using the raw data]

2. If X has possible values $a_1, a_2, a_3, \ldots$, then

$$\boxed{\mathbb{E}[X] = \mathbb{P}(X = a_1) \cdot a_1 + \mathbb{P}(X = a_2) \cdot a_2 + \mathbb{P}(X = a_3) \cdot a_3 + \cdots}$$

[similar to using the relative frequency table]

Expected value of discrete RVs

The expected value (a.k.a. the mean) of a discrete RV X is the average value of X over all the possible outcomes of the experiment weighted by their probabilities, and is denoted by $\mathbb{E}[X]$ or μ.

It is interpreted as the (approximate) mean of the values of X in many many repeated experiments.

Remark

If the experiment involves picking a member of a population uniformly at random (i.e., all members being equally likely to be picked), then the expected value of a RV X is the same as its population mean.

Expected value of discrete RVs

The expected value (a.k.a. the mean) of a discrete RV X is the average value of X over all the possible outcomes of the experiment weighted by their probabilities, and is denoted by $\mathbb{E}[X]$ or μ.

It is interpreted as the (approximate) mean of the values of X in many many repeated experiments.

Remark

If the experiment involves picking a member of a population uniformly at random (i.e., all members being equally likely to be picked), then the expected value of a RV X is the same as its population mean.

Example

The expected weight of a bird chosen uniformly at random from the population of all birds within a species is the same as the population mean of the weight for that species.

Standard deviation of discrete RVs

The variance of a discrete RV X is the expected value of the square of the deviation of X from its mean, and is denoted by $\mathbb{V}\mathrm{ar}[X]$ or σ^2.

The standard deviation of a X is simply the square root of its variance, and is denoted by $\mathbb{SD}[X]$ or σ.

Standard deviation of discrete RVs

The variance of a discrete RV X is the expected value of the square of the deviation of X from its mean, and is denoted by $\mathbb{V}\mathrm{ar}[X]$ or σ^2.

The standard deviation of a X is simply the square root of its variance, and is denoted by $\mathbb{SD}[X]$ or σ.

In other words,

$$\mathbb{V}\mathrm{ar}[X] := \mathbb{E}[(X-\mu)^2]$$
$$\mathbb{SD}[X] := \sqrt{\mathbb{V}\mathrm{ar}[X]}$$

Standard deviation of discrete RVs

The variance of a discrete RV X is the expected value of the square of the deviation of X from its mean, and is denoted by $\mathbb{V}\mathrm{ar}[X]$ or σ^2.

The standard deviation of a X is simply the square root of its variance, and is denoted by $\mathbb{SD}[X]$ or σ.

In other words,

$$\mathbb{V}\mathrm{ar}[X] := \mathbb{E}[(X-\mu)^2] = \mathbb{E}[X^2] - \mu^2$$
$$\mathbb{SD}[X] := \sqrt{\mathbb{V}\mathrm{ar}[X]}$$

Standard deviation of discrete RVs

The variance of a discrete RV X is the expected value of the square of the deviation of X from its mean, and is denoted by $\mathbb{V}\mathrm{ar}[X]$ or σ^2.

The standard deviation of a X is simply the square root of its variance, and is denoted by $\mathbb{SD}[X]$ or σ.

In other words,

$$\mathbb{V}\mathrm{ar}[X] := \mathbb{E}[(X-\mu)^2] = \mathbb{E}[X^2] - \mu^2$$
$$\mathbb{SD}[X] := \sqrt{\mathbb{V}\mathrm{ar}[X]}$$

As before, the expected values of $(X-\mu)^2$ and X^2 can be calculated in two different ways, either by summing over the possible outcomes of the experiment, or by summing over the possible values of X.

Standard deviation of discrete RVs

The variance of a discrete RV X is the expected value of the square of the deviation of X from its mean, and is denoted by $\mathbb{V}\mathrm{ar}[X]$ or σ^2.

The standard deviation of a X is simply the square root of its variance, and is denoted by $\mathbb{SD}[X]$ or σ.

Remark

If the experiments involves picking a member of a population uniformly at random (i.e., all members being equally likely to be picked), then the variance of X is the same as its population variance.

Standard deviation of discrete RVs

The variance of a discrete RV X is the expected value of the square of the deviation of X from its mean, and is denoted by $\mathbb{V}\mathrm{ar}[X]$ or σ^2.

The standard deviation of a X is simply the square root of its variance, and is denoted by $\mathbb{SD}[X]$ or σ.

Remark

If the experiments involves picking a member of a population uniformly at random (i.e., all members being equally likely to be picked), then the variance of X is the same as its population variance.

Example

The variance of the weight of a bird chosen uniformly at random from the population of all birds within a species is the same as the population variance of the weight for that species.

Discrete random variables

Some important types of discrete RVs:

- Binomial RVs
- Poisson RVs

Discrete random variables

Some important types of discrete RVs:

- Bernoulli RVs (we will discuss if time permits)
- Binomial RVs
- Poisson RVs

Discrete random variables

Some important types of discrete RVs:

- Bernoulli RVs (we will discuss if time permits)
- Binomial RVs
- Poisson RVs
- Geometric RVs (not covered in this course)
- Hypergeometric RVs (not covered in this course)
- ...

Discrete random variables

Some important types of discrete RVs:

- Bernoulli RVs (we will discuss if time permits)
- Binomial RVs
- Poisson RVs
- Geometric RVs (not covered in this course)
- Hypergeometric RVs (not covered in this course)
- ...

These are identified by their distributions.

Discrete random variables

Some important types of discrete RVs:

- Bernoulli RVs (we will discuss if time permits)
- Binomial RVs
- Poisson RVs
- Geometric RVs (not covered in this course)
- Hypergeometric RVs (not covered in this course)
- ...

These are identified by their distributions.

Binomial RVs

Example (Gene with two alleles)

A certain type of gene in humans has two alleles (versions) A and B. Previous studies suggest that around 70% of the people have allele A and the remaining 30% have allele B. A random sample of 10 individuals are selected.

$$X := \# \text{ of sampled individuals with allele } A$$

Binomial RVs

Example (Gene with two alleles)

A certain type of gene in humans has two alleles (versions) A and B. Previous studies suggest that around 70% of the people have allele A and the remaining 30% have allele B. A random sample of 10 individuals are selected.

$$X := \# \text{ of sampled individuals with allele } A$$

Q1 What is the distribution of X? (e.g., what is $\mathbb{P}(X = 7)$?)

Binomial RVs

Example (Gene with two alleles)

A certain type of gene in humans has two alleles (versions) A and B. Previous studies suggest that around 70% of the people have allele A and the remaining 30% have allele B. A random sample of 10 individuals are selected.

$$X := \# \text{ of sampled individuals with allele } A$$

Q1 What is the distribution of X? (e.g., what is $\mathbb{P}(X = 7)$?)

Q2 What is the expected value of X?

Binomial RVs

Example (Gene with two alleles)

A certain type of gene in humans has two alleles (versions) A and B. Previous studies suggest that around 70% of the people have allele A and the remaining 30% have allele B. A random sample of 10 individuals are selected.

$$X := \# \text{ of sampled individuals with allele } A$$

Q1 What is the distribution of X? (e.g., what is $\mathbb{P}(X = 7)$?)

Q2 What is the expected value of X?

Q3 What is the standard deviation of X?

Binomial RVs

Example (Flipping a biased coin)

A coin has a bias in favor of heads:

- 70% of the times the coin comes up heads.
- 30% of the times the coin comes up tails.

We flip the coin 10 times in a row.

$$X := \# \text{ of heads}$$

Binomial RVs

Example (Flipping a biased coin)

A coin has a bias in favor of heads:

- 70% of the times the coin comes up heads.
- 30% of the times the coin comes up tails.

We flip the coin 10 times in a row.

$$X := \# \text{ of heads}$$

Q1 What is the distribution of X? (e.g., what is $\mathbb{P}(X = 7)$?)

Binomial RVs

Example (Flipping a biased coin)

A coin has a bias in favor of heads:

- 70% of the times the coin comes up heads.
- 30% of the times the coin comes up tails.

We flip the coin 10 times in a row.

$$X := \# \text{ of heads}$$

(Q1) What is the distribution of X? (e.g., what is $\mathbb{P}(X = 7)$?)

(Q2) What is the expected value of X?

Binomial RVs

Example (Flipping a biased coin)

A coin has a bias in favor of heads:

- 70% of the times the coin comes up heads.
- 30% of the times the coin comes up tails.

We flip the coin 10 times in a row.

$$X := \# \text{ of heads}$$

Q1 What is the distribution of X? (e.g., what is $\mathbb{P}(X = 7)$?)

Q2 What is the expected value of X?

Q3 What is the standard deviation of X?

Binomial RVs

Example (Flipping a biased coin)

A coin has a bias in favor of heads:

- 70% of the times the coin comes up heads.
- 30% of the times the coin comes up tails.

We flip the coin 10 times in a row.

$$X := \# \text{ of heads}$$

Q1 What is the distribution of X? (e.g., what is $\mathbb{P}(X = 7)$?)

Q2 What is the expected value of X?

Q3 What is the standard deviation of X?

Do you notice something familiar??!

Binomial RVs

Consider an experiment involving n trials (or subexperiments), each with two possible results: success or failure. Suppose that

- Each trial is successful with probability p.
- The trials are independent of one another.

The random variable

$$X := \# \text{ of successes}$$

is called a binomial RV with parameters n and p.

The distribution of X is called the binomial distribution with parameters n and p.

Binomial RVs

Consider an experiment involving n trials (or subexperiments), each with two possible results: success or failure. Suppose that

- Each trial is successful with probability p.
- The trials are independent of one another.

The random variable

$$X := \# \text{ of successes}$$

is called a binomial RV with parameters n and p.

The distribution of X is called the binomial distribution with parameters n and p.

Clearly, in the last two examples, the two RVs

- $X := \#$ of sampled individuals with allele A
- $X := \#$ of heads

are both binomial RVs with parameters $n = 10$ and $p = 0.7$.

Binomial RVs

Suppose X is a binomial RV with parameters $n = 10$ and $p = 0.7$.

Q1 What is the distribution of X? (e.g., what is $\mathbb{P}(X = 7)$?)

Q2 What is the expected value of X?

Q3 What is the standard deviation of X?

Binomial RVs

Suppose X is a binomial RV with parameters $n = 10$ and $p = 0.7$.

Q1 What is the distribution of X? (e.g., what is $\mathbb{P}(X = 7)$?)

Q2 What is the expected value of X?

Q3 What is the standard deviation of X?

First, check out the following applet: [Kudos to Prof. Matt Bognar]

https://homepage.divms.uiowa.edu/~mbognar/

Binomial RVs

Suppose X is a binomial RV with parameters $n = 10$ and $p = 0.7$.

Q1 What is the distribution of X? (e.g., what is $\mathbb{P}(X = 7)$?)

Q2 What is the expected value of X?

Q3 What is the standard deviation of X?

First, check out the following applet: [Kudos to Prof. Matt Bognar]

https://homepage.divms.uiowa.edu/~mbognar/

The same kind of computation/visualization (and much more) can be done using statistical software such as R:

https://www.r-project.org/

Binomial RVs

Suppose X is a binomial RV with parameters $n = 10$ and $p = 0.7$.

Q1 What is the distribution of X? (e.g., what is $\mathbb{P}(X = 7)$?)

Binomial RVs

Suppose X is a binomial RV with parameters $n = 10$ and $p = 0.7$.

Q1 What is the distribution of X? (e.g., what is $\mathbb{P}(X = 7)$?)

A The possible values of X are $0, 1, 2, \ldots, 10$.

Binomial RVs

Suppose X is a binomial RV with parameters $n = 10$ and $p = 0.7$.

Q1 What is the distribution of X? (e.g., what is $\mathbb{P}(X = 7)$?)

A The possible values of X are $0, 1, 2, \ldots, 10$.
If k is a possible value, then

$$\mathbb{P}(X = k) = \binom{10}{k}(0.7)^k(0.3)^{10-k}$$

Binomial RVs

Suppose X is a binomial RV with parameters $n = 10$ and $p = 0.7$.

Q1 What is the distribution of X? (e.g., what is $\mathbb{P}(X = 7)$?)

A The possible values of X are $0, 1, 2, \ldots, 10$.
If k is a possible value, then

$$\mathbb{P}(X = k) = \binom{10}{k}(0.7)^k(0.3)^{10-k}$$

Binomial coefficient

The notation $\binom{n}{k}$ (or ${}_nC_k$) stands for the number of ways one can select k distinct objects from a collection of n distinguishable objects, disregarding the order of the selection.

It can be calculated as

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Binomial RVs

Suppose X is a binomial RV with parameters $n = 10$ and $p = 0.7$.

Q1 What is the distribution of X? (e.g., what is $\mathbb{P}(X = 7)$?)

A The possible values of X are $0, 1, 2, \ldots, 10$.
If k is a possible value, then

$$\mathbb{P}(X = k) = \binom{10}{k}(0.7)^k(0.3)^{10-k}$$

Binomial coefficient

The notation $\binom{n}{k}$ (or ${}_nC_k$) stands for the number of ways one can select k distinct objects from a collection of n distinguishable objects, disregarding the order of the selection.

It can be calculated as

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Binomial RVs

Suppose X is a binomial RV with parameters $n = 10$ and $p = 0.7$.

Q1 What is the distribution of X? (e.g., what is $\mathbb{P}(X = 7)$?)

A The possible values of X are $0, 1, 2, \ldots, 10$.
If k is a possible value, then

$$\mathbb{P}(X = k) = \binom{10}{k}(0.7)^k(0.3)^{10-k}$$

Binomial coefficient

The notation $\binom{n}{k}$ (or ${}_nC_k$) stands for the number of ways one can select k distinct objects from a collection of n distinguishable objects, disregarding the order of the selection.

It can be calculated as

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Binomial RVs

Suppose X is a binomial RV with parameters $n = 10$ and $p = 0.7$.

Q1 What is the distribution of X? (e.g., what is $\mathbb{P}(X = 7)$?)

A The possible values of X are $0, 1, 2, \ldots, 10$.
If k is a possible value, then

$$\mathbb{P}(X = k) = \binom{10}{k}(0.7)^k(0.3)^{10-k}$$

e.g.,

$$\mathbb{P}(X = 7) = \binom{10}{7}(0.7)^7(0.3)^3$$

Binomial RVs

Suppose X is a binomial RV with parameters $n = 10$ and $p = 0.7$.

Q1 What is the distribution of X? (e.g., what is $\mathbb{P}(X = 7)$?)

A The possible values of X are $0, 1, 2, \ldots, 10$.
If k is a possible value, then

$$\mathbb{P}(X = k) = \binom{10}{k}(0.7)^k(0.3)^{10-k}$$

e.g.,

$$\mathbb{P}(X = 7) = \binom{10}{7}(0.7)^7(0.3)^3 \approx \boxed{0.2668279}$$

Binomial RVs

Suppose X is a binomial RV with parameters $n = 10$ and $p = 0.7$.

Q2 What is the expected value of X?

Binomial RVs

Suppose X is a binomial RV with parameters $n = 10$ and $p = 0.7$.

Q2 What is the expected value of X?

A The expected value is

$$\mu = \mathbb{E}[X] = 10 \times 0.7 = \boxed{7}$$

Binomial RVs

Suppose X is a binomial RV with parameters $n = 10$ and $p = 0.7$.

Q2 What is the expected value of X?

A The expected value is

$$\mu = \mathbb{E}[X] = 10 \times 0.7 = \boxed{7}$$

Binomial RVs

Suppose X is a binomial RV with parameters $n = 10$ and $p = 0.7$.

Q2 What is the expected value of X?

A The expected value is

$$\mu = \mathbb{E}[X] = 10 \times 0.7 = \boxed{7}$$

Binomial RVs

Suppose X is a binomial RV with parameters $n = 10$ and $p = 0.7$.

Q3 What is the standard deviation of X?

Binomial RVs

Suppose X is a binomial RV with parameters $n = 10$ and $p = 0.7$.

Q3 What is the standard deviation of X?

A We have

$$\sigma^2 = \mathbb{V}\mathrm{ar}[X] = 10 \times 0.7 \times 0.3 = \boxed{2.1}$$

Binomial RVs

Suppose X is a binomial RV with parameters $n = 10$ and $p = 0.7$.

Q3 What is the standard deviation of X?

A We have

$$\sigma^2 = \mathbb{V}\text{ar}[X] = 10 \times 0.7 \times 0.3 = \boxed{2.1}$$

Binomial RVs

Suppose X is a binomial RV with parameters $n = 10$ and $p = 0.7$.

Q3 What is the standard deviation of X?

A We have

$$\sigma^2 = \mathbb{V}\text{ar}[X] = 10 \times 0.7 \times 0.3 = \boxed{2.1}$$

Binomial RVs

Suppose X is a binomial RV with parameters $n = 10$ and $p = 0.7$.

Q3 What is the standard deviation of X?

A We have

$$\sigma^2 = \mathbb{V}\mathrm{ar}[X] = 10 \times 0.7 \times 0.3 = \boxed{2.1}$$

Binomial RVs

Suppose X is a binomial RV with parameters $n = 10$ and $p = 0.7$.

Q3 What is the standard deviation of X?

A We have

$$\sigma^2 = \mathbb{V}\text{ar}[X] = 10 \times 0.7 \times 0.3 = \boxed{2.1}$$

Hence, the standard deviation is

$$\sigma = \mathbb{SD}[X] = \sqrt{\mathbb{V}\text{ar}[X]} = \sqrt{2.1} \approx \boxed{1.449138}$$

Binomial RVs

Example (Gene with two alleles)

A certain type of gene in humans has two alleles (versions) A and B. Previous studies suggest that around 70% of the people have allele A and the remaining 30% have allele B. A random sample of 10 individuals are selected.

$$X := \# \text{ of sampled individuals with allele } A$$

Q1 What is the distribution of X? (e.g., what is $\mathbb{P}(X = 7)$?)

Q2 What is the expected value of X?

Q3 What is the standard deviation of X?

Binomial RVs

Example (Gene with two alleles)

A certain type of gene in humans has two alleles (versions) A and B. Previous studies suggest that around 70% of the people have allele A and the remaining 30% have allele B. A random sample of 10 individuals are selected.

$$X := \# \text{ of sampled individuals with allele } A$$

Q1 What is the distribution of X? (e.g., what is $\mathbb{P}(X = 7)$?)

A $\mathbb{P}(X = 7) = \boxed{0.2668279}$.

Q2 What is the expected value of X?

Q3 What is the standard deviation of X?

Binomial RVs

Example (Gene with two alleles)

A certain type of gene in humans has two alleles (versions) A and B. Previous studies suggest that around 70% of the people have allele A and the remaining 30% have allele B. A random sample of 10 individuals are selected.

$$X := \# \text{ of sampled individuals with allele } A$$

Q1 What is the distribution of X? (e.g., what is $\mathbb{P}(X = 7)$?)

A $\mathbb{P}(X = 7) = \boxed{0.2668279}$.

Q2 What is the expected value of X?

A $\mathbb{E}[X] = \boxed{7}$.

Q3 What is the standard deviation of X?

Binomial RVs

Example (Gene with two alleles)

A certain type of gene in humans has two alleles (versions) A and B. Previous studies suggest that around 70% of the people have allele A and the remaining 30% have allele B. A random sample of 10 individuals are selected.

$$X := \# \text{ of sampled individuals with allele } A$$

Q1 What is the distribution of X? (e.g., what is $\mathbb{P}(X = 7)$?)

A $\mathbb{P}(X = 7) = \boxed{0.2668279}$.

Q2 What is the expected value of X?

A $\mathbb{E}[X] = \boxed{7}$.

Q3 What is the standard deviation of X?

A $\mathbb{SD}[X] \approx \boxed{1.449138}$.

Binomial RVs

In general, if X is a binomial RV with parameters n and p, then

- The possible values of X are $0, 1, 2, \ldots, n$.
  If k is a possible value, then

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

- The <u>expected value</u> of X is

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

- The <u>variance</u> of X is

$$\boxed{\sigma^2 = \mathbb{V}\mathrm{ar}[X] = np(1-p)}$$

Binomial RVs

Example (Gene with two alleles)

A certain type of gene in humans has two alleles (versions) A and B. Previous studies suggest that around 70% of the people have allele A and the remaining 30% have allele B. A random sample of 10 individuals are selected.

$$X := \# \text{ of sampled individuals with allele } A$$

Q1 What is the distribution of X? (e.g., what is $\mathbb{P}(X = 7)$?)

A $\mathbb{P}(X = 7) = \boxed{0.2668279}$.

Q2 What is the expected value of X?

A $\mathbb{E}[X] = \boxed{7}$.

Q3 What is the standard deviation of X?

A $\mathbb{SD}[X] \approx \boxed{1.449138}$.

Binomial RVs

Example (Gene with two alleles)

A certain type of gene in humans has two alleles (versions) A and B. Previous studies suggest that around 70% of the people have allele A and the remaining 30% have allele B. A random sample of 10 individuals are selected.

$$X := \# \text{ of sampled individuals with allele } A$$

Q4 What is the probability that at most 1 individual has A?

Binomial RVs

Example (Gene with two alleles)

A certain type of gene in humans has two alleles (versions) A and B. Previous studies suggest that around 70% of the people have allele A and the remaining 30% have allele B. A random sample of 10 individuals are selected.

$$X := \# \text{ of sampled individuals with allele } A$$

Q4 What is the probability that at most 1 individual has A?

A We have

$$\mathbb{P}(X \leq 1)$$

Binomial RVs

Example (Gene with two alleles)

A certain type of gene in humans has two alleles (versions) A and B. Previous studies suggest that around 70% of the people have allele A and the remaining 30% have allele B. A random sample of 10 individuals are selected.

$$X := \# \text{ of sampled individuals with allele } A$$

Q4 What is the probability that at most 1 individual has A?

A We have

$$\mathbb{P}(X \leq 1) = \mathbb{P}(X = 0) + \mathbb{P}(X = 1)$$

Binomial RVs

Example (Gene with two alleles)

A certain type of gene in humans has two alleles (versions) A and B. Previous studies suggest that around 70% of the people have allele A and the remaining 30% have allele B. A random sample of 10 individuals are selected.

$$X := \# \text{ of sampled individuals with allele } A$$

Q4 What is the probability that at most 1 individual has A?

A We have

$$\begin{aligned}\mathbb{P}(X \leq 1) &= \mathbb{P}(X = 0) + \mathbb{P}(X = 1) \\ &= \binom{10}{0}(0.7)^0(0.3)^{10} + \binom{10}{1}(0.7)^1(0.3)^9\end{aligned}$$

Binomial RVs

Example (Gene with two alleles)

A certain type of gene in humans has two alleles (versions) A and B. Previous studies suggest that around 70% of the people have allele A and the remaining 30% have allele B. A random sample of 10 individuals are selected.

$$X := \# \text{ of sampled individuals with allele } A$$

Q4 What is the probability that at most 1 individual has A?

$\boxed{\text{A}}$ We have

$$\begin{aligned}
\mathbb{P}(X \leq 1) &= \mathbb{P}(X = 0) + \mathbb{P}(X = 1) \\
&= \binom{10}{0}(0.7)^0(0.3)^{10} + \binom{10}{1}(0.7)^1(0.3)^9 \\
&= 0.0000059049 + 0.000137781
\end{aligned}$$

Binomial RVs

Example (Gene with two alleles)

A certain type of gene in humans has two alleles (versions) A and B. Previous studies suggest that around 70% of the people have allele A and the remaining 30% have allele B. A random sample of 10 individuals are selected.

$$X := \# \text{ of sampled individuals with allele } A$$

Q4 What is the probability that at most 1 individual has A?

A We have

$$\begin{aligned}
\mathbb{P}(X \le 1) &= \mathbb{P}(X = 0) + \mathbb{P}(X = 1) \\
&= \binom{10}{0}(0.7)^0(0.3)^{10} + \binom{10}{1}(0.7)^1(0.3)^9 \\
&= 0.0000059049 + 0.000137781 \\
&= 0.0001436859 \approx \boxed{1.44 \times 10^{-4}}
\end{aligned}$$

Binomial RVs

Example (Gene with two alleles)

A certain type of gene in humans has two alleles (versions) A and B. Previous studies suggest that around 70% of the people have allele A and the remaining 30% have allele B. A random sample of 10 individuals are selected.

$$X := \# \text{ of sampled individuals with allele } A$$

Q5 What is the probability that at most 8 individual have A?

Binomial RVs

Example (Gene with two alleles)

A certain type of gene in humans has two alleles (versions) A and B. Previous studies suggest that around 70% of the people have allele A and the remaining 30% have allele B. A random sample of 10 individuals are selected.

$$X := \# \text{ of sampled individuals with allele } A$$

Q5 What is the probability that at most 8 individual have A?

A1 We have

$$\mathbb{P}(X \leq 8)$$

Binomial RVs

Example (Gene with two alleles)

A certain type of gene in humans has two alleles (versions) A and B. Previous studies suggest that around 70% of the people have allele A and the remaining 30% have allele B. A random sample of 10 individuals are selected.

$$X := \# \text{ of sampled individuals with allele } A$$

Q5 What is the probability that at most 8 individual have A?

A1 We have

$$\mathbb{P}(X \leq 8) = \mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \cdots + \mathbb{P}(X = 8)$$

Binomial RVs

Example (Gene with two alleles)

A certain type of gene in humans has two alleles (versions) A and B. Previous studies suggest that around 70% of the people have allele A and the remaining 30% have allele B. A random sample of 10 individuals are selected.

$$X := \# \text{ of sampled individuals with allele } A$$

Q5 What is the probability that at most 8 individual have A?

$\boxed{\text{A1}}$ We have

$$\begin{aligned}\mathbb{P}(X \leq 8) &= \mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \cdots + \mathbb{P}(X = 8) \\ &= \binom{10}{0}(0.7)^0(0.3)^{10} + \binom{10}{1}(0.7)^1(0.3)^9 + \cdots + \binom{10}{8}(0.7)^8(0.3)^2\end{aligned}$$

Binomial RVs

Example (Gene with two alleles)

A certain type of gene in humans has two alleles (versions) A and B. Previous studies suggest that around 70% of the people have allele A and the remaining 30% have allele B. A random sample of 10 individuals are selected.

$$X := \# \text{ of sampled individuals with allele } A$$

Q5 What is the probability that at most 8 individual have A?

$\boxed{\text{A1}}$ We have

$$\begin{aligned}
\mathbb{P}(X \le 8) &= \mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \cdots + \mathbb{P}(X = 8) \\
&= \binom{10}{0}(0.7)^0(0.3)^{10} + \binom{10}{1}(0.7)^1(0.3)^9 + \cdots + \binom{10}{8}(0.7)^8(0.3)^2 \\
&= \ldots \qquad \text{(Ugh! Too much calculation!)}
\end{aligned}$$

Binomial RVs

Example (Gene with two alleles)

A certain type of gene in humans has two alleles (versions) A and B. Previous studies suggest that around 70% of the people have allele A and the remaining 30% have allele B. A random sample of 10 individuals are selected.

$$X := \# \text{ of sampled individuals with allele } A$$

Q5 What is the probability that at most 8 individual have A?

A2 We have

$$\mathbb{P}(X \leq 8)$$

Binomial RVs

Example (Gene with two alleles)

A certain type of gene in humans has two alleles (versions) A and B. Previous studies suggest that around 70% of the people have allele A and the remaining 30% have allele B. A random sample of 10 individuals are selected.

$$X := \# \text{ of sampled individuals with allele } A$$

Q5 What is the probability that at most 8 individual have A?

A2 We have

$$\mathbb{P}(X \leq 8) = 1 - \mathbb{P}(X = 9) - \mathbb{P}(X = 10)$$

Binomial RVs

Example (Gene with two alleles)

A certain type of gene in humans has two alleles (versions) A and B. Previous studies suggest that around 70% of the people have allele A and the remaining 30% have allele B. A random sample of 10 individuals are selected.

$$X := \# \text{ of sampled individuals with allele } A$$

Q5 What is the probability that at most 8 individual have A?

$\boxed{\text{A2}}$ We have

$$\begin{aligned}\mathbb{P}(X \leq 8) &= 1 - \mathbb{P}(X = 9) - \mathbb{P}(X = 10) \\ &= 1 - \binom{10}{9}(0.7)^9(0.3)^1 - \binom{10}{10}(0.7)^{10}(0.3)^0\end{aligned}$$

Binomial RVs

Example (Gene with two alleles)

A certain type of gene in humans has two alleles (versions) A and B. Previous studies suggest that around 70% of the people have allele A and the remaining 30% have allele B. A random sample of 10 individuals are selected.

$$X := \# \text{ of sampled individuals with allele } A$$

Q5 What is the probability that at most 8 individual have A?

A2 We have

$$\begin{aligned}
\mathbb{P}(X \leq 8) &= 1 - \mathbb{P}(X = 9) - \mathbb{P}(X = 10) \\
&= 1 - \binom{10}{9}(0.7)^9(0.3)^1 - \binom{10}{10}(0.7)^{10}(0.3)^0 \\
&= 1 - 0.121060821 - 0.0282475249
\end{aligned}$$

Binomial RVs

Example (Gene with two alleles)

A certain type of gene in humans has two alleles (versions) A and B. Previous studies suggest that around 70% of the people have allele A and the remaining 30% have allele B. A random sample of 10 individuals are selected.

$$X := \# \text{ of sampled individuals with allele } A$$

Q5 What is the probability that at most 8 individual have A?

$\boxed{\text{A2}}$ We have

$$\begin{aligned}
\mathbb{P}(X \leq 8) &= 1 - \mathbb{P}(X = 9) - \mathbb{P}(X = 10) \\
&= 1 - \binom{10}{9}(0.7)^9(0.3)^1 - \binom{10}{10}(0.7)^{10}(0.3)^0 \\
&= 1 - 0.121060821 - 0.0282475249 \\
&\approx \boxed{0.850692}
\end{aligned}$$