Lecture 3: Bayes theorem and discrete distributions

MVE055 / MSG810
Mathematical statistics and discrete mathematics

Moritz Schauer

Last updated September 1, 2021, 2021

GU & Chalmers University of Technology

Conditional distribution

If we know some event B occurs, the probability of A given the new information B can be calculated as follows:

Conditional probability

Assume that $\mathsf{P}(B) > 0$. The conditional probability of A given B is defined as

$$\mathsf{P}(A \mid B) = \frac{\mathsf{P}(A \cap B)}{\mathsf{P}(B)}. \tag{0.1}$$

Conditional distribution

If we know some event B occurs, the probability of A given the new information B can be calculated as follows:

Conditional probability

Assume that $\mathsf{P}(B) > 0$. The conditional probability of A given B is defined as

$$\mathsf{P}(A \mid B) = \frac{\mathsf{P}(A \cap B)}{\mathsf{P}(B)}. \tag{0.1}$$

Multiplication rule for probabilities

For events A and B it holds

$$\mathsf{P}(A \cap B) = \boxed{}$$

The multiplication rule is useful to calculate probabilities of multiple events affecting each other.

Conditional distribution

If we know some event B occurs, the probability of A given the new information B can be calculated as follows:

Conditional probability

Assume that $\mathsf{P}(B) > 0$. The conditional probability of A given B is defined as

$$\mathsf{P}(A \mid B) = \frac{\mathsf{P}(A \cap B)}{\mathsf{P}(B)}. \tag{0.1}$$

Multiplication rule for probabilities

For events A and B it holds

$$\mathsf{P}(A \cap B) = \mathsf{P}(B \mid A)\mathsf{P}(A) = \mathsf{P}(A \mid B)\mathsf{P}(B).$$

The multiplication rule is useful to calculate probabilities of multiple events affecting each other.

Conditional distribution

If we know some event B occurs, the probability of A given the new information B can be calculated as follows:

Conditional probability

Assume that $\mathsf{P}(B) > 0$. The conditional probability of A given B is defined as

$$\mathsf{P}(A \mid B) = \frac{\mathsf{P}(A \cap B)}{\mathsf{P}(B)}. \tag{0.1}$$

Multiplication rule for probabilities

For events A and B it holds

$$\mathsf{P}(A \cap B) = \mathsf{P}(B \mid A)\mathsf{P}(A) = \mathsf{P}(A \mid B)\mathsf{P}(B).$$

The multiplication rule is useful to calculate probabilities of multiple events affecting each other.

Bayes formula

Bayes formula

For events A and B

$$\mathsf{P}(A \mid B) = \frac{\mathsf{P}(B \mid A)\mathsf{P}(A)}{\mathsf{P}(B)}$$

Bayes formula

Bayes formula

For events A and B

$$\mathsf{P}(A \mid B) = \frac{\mathsf{P}(B \mid A)\mathsf{P}(A)}{\mathsf{P}(B)}$$

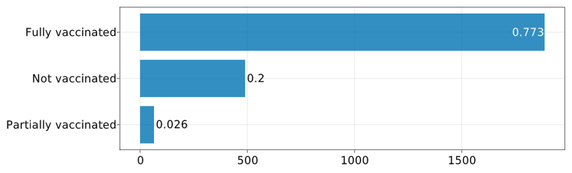
Often it is useful to rewrite the denominator $\mathsf{P}(B)$

$$\mathsf{P}(B) = \mathsf{P}(B \mid A)\mathsf{P}(A) + \mathsf{P}(B \mid A^c)\mathsf{P}(A^c)$$

Base rate fallacy

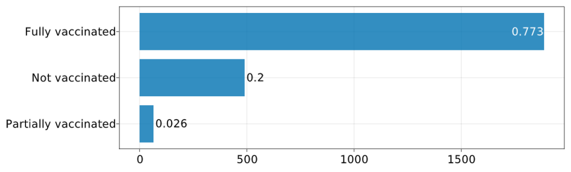
Base rate fallacy

In the fourth wave (July 10 - August 16, 2021) about 2400 (or 0.825 %) people 16 or older in Island have been diagnosed with Covid-19:

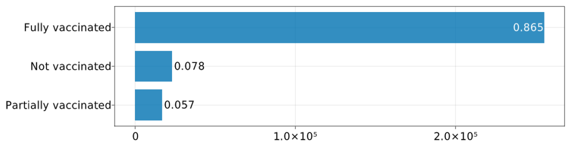


Base rate fallacy

In the fourth wave (July 10 - August 16, 2021) about 2400 (or 0.825 %) people 16 or older in Island have been diagnosed with Covid-19:



But (young) adults in Iceland's population are highly vaccinated



$$\mathrm{P}(\text{diagn} \mid \text{vacc}) = \frac{\mathrm{P}(\text{vacc} \mid \text{diagn})\mathrm{P}(\text{diagn})}{\mathrm{P}(\text{vacc})} = \frac{0.773 \cdot 0.00825}{0.864} = 0.00738$$

Base rate fallacy

$$\mathrm{P}(\mathsf{diagn} \mid \mathsf{vacc}) = \frac{0.773 \cdot 0.00825}{0.864} = 0.00738$$
$$\mathrm{P}(\mathsf{diagn} \mid \mathsf{not\ vacc}) = \frac{0.200 \cdot 0.00825}{0.0783} = 0.0211$$

Base rate fallacy

$$\mathrm{P}(\textsf{diagn} \mid \textsf{vacc}) = \frac{0.773 \cdot 0.00825}{0.864} = 0.00738$$
$$\mathrm{P}(\textsf{diagn} \mid \textsf{not vacc}) = \frac{0.200 \cdot 0.00825}{0.0783} = 0.0211$$

$$\mathrm{P}(\textsf{diagn} \mid \textsf{part. vacc}) = \frac{0.0262 \cdot 0.00825}{0.0570} = 0.00379 \text{ (}\textsf{sic!}\text{)}$$

Independent events

Two events A and B are independent if knowing whether B occured does not change the probability of A

$$\mathsf{P}(A \mid B) = \mathsf{P}(A).$$

Independent events

Two events A and B are independent if $\mathsf{P}(A \cap B) = \mathsf{P}(A)\mathsf{P}(B)$.

Independent events

Two events A and B are independent if knowing whether B occured does not change the probability of A

$$\mathsf{P}(A \mid B) = \mathsf{P}(A).$$

Independent events

Two events A and B are independent if $\mathsf{P}(A \cap B) = \mathsf{P}(A)\mathsf{P}(B)$.

Simple example: Throw a 6-sided die. Are $A = \{5, 6\}$ and $B = \{1, 3, 5\}$ dependent?

Independent events

Two events A and B are independent if knowing whether B occured does not change the probability of A

$$\mathsf{P}(A \mid B) = \mathsf{P}(A).$$

Independent events

Two events A and B are independent if $\mathsf{P}(A \cap B) = \mathsf{P}(A)\mathsf{P}(B)$.

Simple example: Throw a 6-sided die. Are $A = \{5, 6\}$ and $B = \{1, 3, 5\}$ dependent?

$$\mathrm{P}(A)\mathrm{P}(B) = \frac{2}{6}\frac{3}{6} = \frac{1}{6}, \quad \mathrm{P}(A \cap B) = \mathrm{P}(\{5\}) = \frac{1}{6}.$$

If I tell you A happened, that does not change probabilities of B: $\mathrm{P}(B \mid A) = \mathrm{P}(B) = \frac{3}{6}$.

Random variables

Random variables

A random variable is a numeric quantity whose value depends on the outcome of a random experiment.

Random variables

Random variables

A random variable is a numeric quantity whose value depends on the outcome of a random experiment.

Example: X is the number of eyes on a 6-sided die.

We denote random variables with capital letters, often X or Y.

Example:

Pair of dice

Throw a pair of dice, count the total number of eyes, call that random variable X. Consider the **event** that $X = 7$.

Pair of dice

Throw a pair of dice, count the total number of eyes, call that random variable X. Consider the **event** that $X = 7$.

Event? What are the actual ω making our event and sample space? You could take

$$A = \{\epsdice{1}\epsdice{6}, \epsdice{2}\epsdice{5}, \epsdice{3}\epsdice{4}, \epsdice{4}\epsdice{3}, \epsdice{5}\epsdice{2}, \epsdice{6}\epsdice{1}\}, \quad \Omega = \{\epsdice{1}\epsdice{1}, \ldots, \epsdice{6}\epsdice{6}\}$$

Pair of dice

Throw a pair of dice, count the total number of eyes, call that random variable X. Consider the **event** that $X = 7$.

Event? What are the actual ω making our event and sample space? You could take

$$A = \{\epsdice{1}\epsdice{6}, \epsdice{2}\epsdice{5}, \epsdice{3}\epsdice{4}, \epsdice{4}\epsdice{3}, \epsdice{5}\epsdice{2}, \epsdice{6}\epsdice{1}\}, \quad \Omega = \{\epsdice{1}\epsdice{1}, \dots, \epsdice{6}\epsdice{6}\}$$

$$\mathrm{P}(X = 7) = \mathrm{P}(A) = \frac{|A|}{|\Omega|} = \frac{6}{36}$$

Pair of dice

Throw a pair of dice, count the total number of eyes, call that random variable X. Consider the **event** that $X = 7$.

Event? What are the actual ω making our event and sample space? You could take

$$A = \{⚀⚅, ⚁⚄, ⚂⚃, ⚃⚂, ⚄⚁, ⚅⚀\}, \quad \Omega = \{⚀⚀, \ldots, ⚅⚅\}$$

$$\mathrm{P}(X = 7) = \mathrm{P}(A) = \frac{|A|}{|\Omega|} = \frac{6}{36}$$

Value k	2	3	4	5	6	7	8	9	10	11	12
Probability $\mathrm{P}(X{=}k)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Pair of dice

The following holds for $k \in \{2, \dots, 12\}$:

$$\mathrm{P}(X = k) = \frac{6 - |k - 7|}{36}$$

Pair of dice

The following holds for $k \in \{2, \ldots, 12\}$:

$$\mathrm{P}(X = k) = \frac{6 - |k - 7|}{36}$$

Check:

Value k	2	3	4	5	6	7	8	9	10	11	12	other
Probability $\mathrm{P}(X{=}k)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	0

Discrete random variables

Discrete random variables

A random variable is called discrete if it is integer-valued or otherwise has only a finite or countable number of values.

Example: $Y = X/2$ is discrete (but can take non-integers such as $Y = 5.5$ as values.)

Probability mass function

Probability mass function

Define the probability mass function f of a discrete random variable X by

$$f(k) = \mathsf{P}(X = k).$$

Probability mass function

Probability mass function

Define the probability mass function f of a discrete random variable X by

$$f(k) = \mathsf{P}(X = k).$$

Also $f(y) = 0$ for all real y such that $P(X = y) = 0$, okay?

Probability mass function

Probability mass function

Define the probability mass function f of a discrete random variable X by

$$f(k) = \mathsf{P}(X = k).$$

Also $f(y) = 0$ for all real y such that $P(X = y) = 0$, okay?

Sometimes we write f_X to talk about X's own probability mass function.

Sum of two dice

$$f(k) = \begin{cases} \frac{6-|k-7|}{36} & \text{if } k \in \{2, 3, \ldots, 12\} \\ 0 & \text{otherwise} \end{cases}$$

is the probability mass function for the random variable which counts the sum of two dice.

Two coins

Flip two coins... count the number of heads. Call it X.

Two coins

Flip two coins... count the number of heads. Call it X.

$f(0) = \frac{1}{4}, f(1) = \frac{1}{2}$ and $f(2) = \frac{1}{4}$

$f(x) = 0$ otherwise if $x \notin \{0, 1, 2\}$.

Flip two coins... count the number of heads. $f_X(0) = \frac{1}{4}$, $f_X(1) = \frac{1}{2}$ and $f_X(2) = \frac{1}{4}$.

What is $\mathrm{P}(X \in \{1, 2\}) = \mathrm{P}(1 \leqslant X \leqslant 2)$?

Flip two coins... count the number of heads. $f_X(0) = \frac{1}{4}$, $f_X(1) = \frac{1}{2}$ and $f_X(2) = \frac{1}{4}$.

What is $\mathrm{P}(X \in \{1, 2\}) = \mathrm{P}(1 \leqslant X \leqslant 2)$?

$$\mathrm{P}(1 \leqslant X \leqslant 2) = f_X(1) + f_X(2) = \frac{3}{4}$$

Flip two coins... count the number of heads. $f_X(0) = \frac{1}{4}$, $f_X(1) = \frac{1}{2}$ and $f_X(2) = \frac{1}{4}$.

What is $\mathrm{P}(X \in \{1, 2\}) = \mathrm{P}(1 \leqslant X \leqslant 2)$?

$$\mathrm{P}(1 \leqslant X \leqslant 2) = f_X(1) + f_X(2) = \frac{3}{4}$$

Let $Y = X/2$. What is $\mathrm{P}(Y > 0)$?

Flip two coins... count the number of heads. $f_X(0) = \frac{1}{4}$, $f_X(1) = \frac{1}{2}$ and $f_X(2) = \frac{1}{4}$.

What is $\mathrm{P}(X \in \{1, 2\}) = \mathrm{P}(1 \leqslant X \leqslant 2)$?

$$\mathrm{P}(1 \leqslant X \leqslant 2) = f_X(1) + f_X(2) = \frac{3}{4}$$

Let $Y = X/2$. What is $\mathrm{P}(Y > 0)$?

$$\mathrm{P}(Y > 0) = \mathrm{P}(1 \leqslant X \leqslant 2) = f_X(1) + f_X(2) = \frac{3}{4}$$

Rule

For integer valued X

$$\mathsf{P}(m \leqslant X \leqslant n) = \sum_{k=m}^{n} f(k)$$

for any integers m and n.

Probability mass function

Not all functions are probability mass functions. Because they describe probability distributions, some conditions must hold.

$f(k)$ is a probability mass function if and only if

- $f(k) \geqslant 0$ for all k.
- $\sum_{\text{all } k} f(k) = 1$.

If somebody gives you a probability mass function, there is a random variable for it.

Distribution function

Assume X is a discrete random variable. Its distribution function is given by

$$F(x) = \mathsf{P}(X \leqslant x) = \sum_{k \leqslant x} f_X(k),$$

Flip two coins... count the number of heads. Call it X.
$f(0) = \frac{1}{4}, f(1) = \frac{1}{2}$ and $f(2) = \frac{1}{4}$. Find F.

Distribution function

Distribution function

Assume X is a discrete random variable. Its distribution function is given by

$$F(x) = \mathsf{P}(X \leqslant x) = \sum_{k \leqslant x} f_X(k),$$

Flip two coins... count the number of heads. Call it X.
$f(0) = \frac{1}{4}, f(1) = \frac{1}{2}$ and $f(2) = \frac{1}{4}$. Find F.

$$\begin{aligned}
F(0) &= f(0) = \tfrac{1}{4} \\
F(1) &= f(0) + f(1) = \tfrac{1}{4} + \tfrac{1}{2} \\
F(2) &= f(0) + f(1) + f(2) = 1
\end{aligned}$$

Distribution function

What is the probability to throw k times heads in a row with a fair coin?

Distribution function

What is the probability to throw k times heads in a row with a fair coin?

$$f(0) = \frac{1}{2}, \quad f(1) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}, \quad f(2) = \frac{1}{8}, \quad f(k) = \left(\frac{1}{2}\right)^{k+1}$$

Distribution function

What is the probability to throw k times heads in a row with a fair coin?

$$f(0) = \frac{1}{2}, \quad f(1) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}, \quad f(2) = \frac{1}{8}, \quad f(k) = \left(\frac{1}{2}\right)^{k+1}$$

$$P(X > 0) = f(1) + f(2) + f(3) + \ldots = \boxed{}$$

Distribution function

What is the probability to throw k times heads in a row with a fair coin?

$$f(0) = \frac{1}{2}, \quad f(1) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}, \quad f(2) = \frac{1}{8}, \quad f(k) = \left(\frac{1}{2}\right)^{k+1}$$

$$P(X > 0) = f(1) + f(2) + f(3) + ... = 1 - P(X = 0) = 1 - f(0)$$

Distribution function

What is the probability to throw k times heads in a row with a fair coin?

$$f(0) = \frac{1}{2}, \quad f(1) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}, \quad f(2) = \frac{1}{8}, \quad f(k) = \left(\frac{1}{2}\right)^{k+1}$$

$$P(X > 0) = f(1) + f(2) + f(3) + ... = 1 - P(X = 0) = 1 - f(0)$$

Distribution function

For $F(x)$ it holds

- $F(x)$ is increasing
- $F(x) \to 1$ for $x \to \infty$.
- $F(x) \to 0$ for $x \to -\infty$.

Distribution function

Also

- $\mathsf{P}(a < X \leqslant b) = F(b) - F(a)$.
- $\mathsf{P}(X > a) = 1 - F(a)$.
- For integer valued random variables:
  $f(m) = F(m) - F(m-1)$.

We are often interested in the "average" outcome of a random variable.

Expected value

The expected value of a random variable is defined as

$$\mathsf{E}(X) = \sum_{\text{all } k} k f_X(k) \qquad \text{if } X \text{ is discrete,}$$

Recall: the average using fractions

Data set: grades of 24 students

$$5, 5, 6, 5, 6, 6, 6, 5, 5, 7, 6, 7, 5, 5, 5, 6, 6, 6, 5, 6, 5, 7, 6, 7$$

Table:

grade	$x_1 = 7$	$x_2 = 6$	$x_3 = 5$
fraction of students	$p_1 = 4/24$	$p_2 = 10/24$	$p_3 = 10/24$

Recall: the average using fractions

Data set: grades of 24 students

$$5, 5, 6, 5, 6, 6, 6, 5, 5, 7, 6, 7, 5, 5, 5, 6, 6, 6, 5, 6, 5, 7, 6, 7$$

Table:

grade	$x_1 = 7$	$x_2 = 6$	$x_3 = 5$
fraction of students	$p_1 = 4/24$	$p_2 = 10/24$	$p_3 = 10/24$

Average One can write the average in different forms

$$\text{Average} = \frac{5+5+6+\cdots+5+7+6+7}{24}$$

$$= \frac{7\cdot 4 + 6\cdot 10 + 5\cdot 10}{24} = 7\cdot\frac{4}{24} + 6\cdot\frac{10}{24} + 5\cdot\frac{10}{24} = \sum_{i=1}^{3} x_i \cdot p_i$$

Expected value

The expected value of a discrete random variable X with finitely many outcomes can also be written as

$$\mu = \mathrm{E}(X) = \sum_{\text{all } k} x_k \cdot \underbrace{\mathrm{P}(X = x_k)}_{f(x_k)}$$

$$= x_1 \cdot \mathrm{P}(X = x_1) + x_2\, \mathrm{P}(X = x_2) + \cdots + x_n \cdot \mathrm{P}(X = x_n)$$

Expected value

The expected value of a discrete random variable X with finitely many outcomes can also be written as

$$\mu = \mathrm{E}(X) = \sum_{\text{all } k} x_k \cdot \underbrace{\mathrm{P}(X = x_k)}_{f(x_k)}$$

$$= x_1 \cdot \mathrm{P}(X = x_1) + x_2 \, \mathrm{P}(X = x_2) + \cdots + x_n \cdot \mathrm{P}(X = x_n)$$

Here x_i are the n possible outcomes and $P(X = x_i)$ are the probabilities of each outcome.

Expected value

Flip two coins... count the number of heads.

Expected value

Flip two coins... count the number of heads.

$f(0) = \frac{1}{4}, f(1) = \frac{1}{2}$ and $f(2) = \frac{1}{4}$

$\mathsf{E}(X) =$

Expected value

Flip two coins... count the number of heads.

$f(0) = \frac{1}{4}, f(1) = \frac{1}{2}$ and $f(2) = \frac{1}{4}$

$\mathsf{E}(X) =$

Expected value

Flip two coins... count the number of heads.

$f(0) = \frac{1}{4}, f(1) = \frac{1}{2}$ and $f(2) = \frac{1}{4}$

$\mathsf{E}(X) = \boxed{0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} = 1}$

Rules for computing expected values

For the expected value,

- $\mathsf{E}(a) = a$.
- $\mathsf{E}(aX) = a\mathsf{E}(X)$.
- $\mathsf{E}(aX + b) = a\mathsf{E}(X) + b$.
- $\mathsf{E}(X + Y) = \mathsf{E}(X) + \mathsf{E}(Y)$.

Here X and Y are any two random variables and a and b are constants.

Transformations

If we transform the random variables by a function h we have:

Theorem ♡

$$\mathsf{E}(h(X)) = \sum_{\text{all } k} h(k) f(k)$$

Transformations

If we transform the random variables by a function h we have:

Theorem ♡

$$\mathsf{E}(h(X)) = \sum_{\text{all } k} h(k) f(k)$$

Coin example (with $h(x) = x/2$):

$$\mathsf{E}(X/2) = \frac{0}{2} \cdot f_X(0) + \frac{1}{2} \cdot f_X(1) + \frac{2}{2} \cdot f_X(2) = \tfrac{1}{2}$$

Transformations

If we transform the random variables by a function h we have:

Theorem ♡

$$\mathsf{E}(h(X)) = \sum_{\text{all } k} h(k) f(k)$$

Coin example (with $h(x) = x/2$):

$$\mathsf{E}(X/2) = \frac{0}{2} \cdot f_X(0) + \frac{1}{2} \cdot f_X(1) + \frac{2}{2} \cdot f_X(2) = \tfrac{1}{2}$$

$$= (\mathsf{E}(X))/2$$

Common distributions

Bernoulli distribution

The Bernoulli distribution describes a random experiment that can either succeed (with probability p) or fail (with probability $1 - p$.) Suppose we make a random experiment which succeeds with probability p and set

$$X = \begin{cases} 1, & \text{if the experiment succeeds} \\ 0, & \text{in case of failure.} \end{cases}$$

We have $f(1) = p$ and $f(0) = 1 - p$.

Bernoulli distribution

The Bernoulli distribution describes a random experiment that can either succeed (with probability p) or fail (with probability $1 - p$.) Suppose we make a random experiment which succeeds with probability p and set

$$X = \begin{cases} 1, & \text{if the experiment succeeds} \\ 0, & \text{in case of failure.} \end{cases}$$

We have $f(1) = p$ and $f(0) = 1 - p$.

Sometimes useful to write as $f(k) = p^k(1-p)^{1-k}$ for $k \in \{0, 1\}$.

The binomial distribution

Bernoulli distribution

A random variable X is Bernoulli distributed if it has probability mass function $f(1) = p$ and $f(0) = 1 - p$ and $= 0$ otherwise. We write $X \sim \mathrm{Ber}(p)$.

Examples?

The binomial distribution

The binomial distribution describes the probability of having exactly k successes in n independent Bernoulli trials with probability of success p.

If X is Binomial with parameters n and p we write:

$$X \sim \mathrm{Bin}(n, p)$$

The binomial distribution

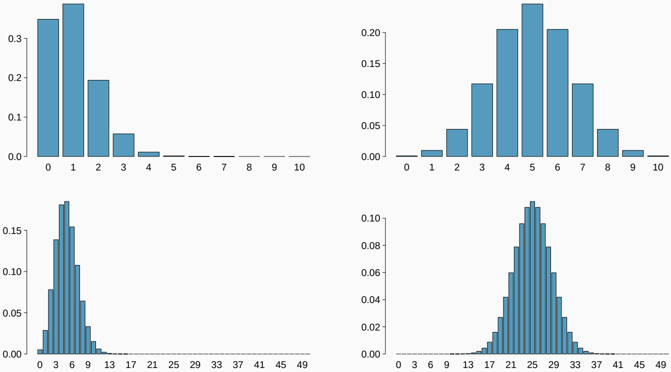
The binomial distribution describes the probability of having exactly k successes in n independent Bernoulli trials with probability of success p.

If X is Binomial with parameters n and p we write:

$$X \sim \mathrm{Bin}(n, p)$$

Ha, the sum of two coins with sides 0 and 1 is $\mathrm{Bin}(2, 0.5)$ distributed.

The binomial distribution



$n = 10$

$n = 50$

$p = 0.1$ $p = 0.5$

The binomial distribution

The binomial distribution describes the probability of having exactly k successes in n independent Bernoulli trials with probability of success p.

If X is Binomial with parameters n and p we write:

$$X \sim \mathrm{Bin}(n, p)$$

Binomial distribution

A random variable X is Binomial distributed with parameters n, p if

$$\mathrm{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \qquad \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Sum of binomial distributed random variables

Sum of binomial distributed random variables.

If $X_1 \sim \mathrm{Bin}(n, p)$ and $X_2 \sim \mathrm{Bin}(m, p)$ are independent, then $X_1 + X_2 \sim \mathrm{Bin}(m + n, p)$.

("Dropping m items, couting the broken ones, dropping n more items, counting the additional broken ones is the same as dropping $m + n$ items...")

Geometric distribution

The experiment consists of a series of independent Bernoulli trials with probability of success equal to p.

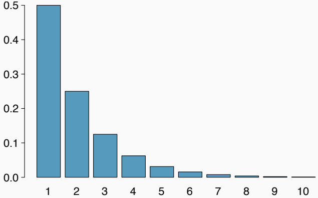
The random variable X denotes the number of trials needed to get the first success.

p is called the parameter of X.

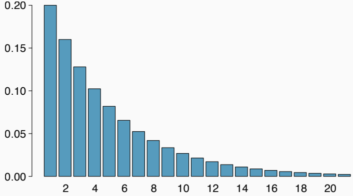
The geometric distribution

The geometric distribution describes the probability distribution of the number of trials needed k to get the first success, for a single event succeeding with probability p. ($k-1$ failures and 1 success.)

$p = 0.5$



$p = 0.2$



The geometric distribution

Geometric distribution

A random variable X is geometrically distributed with parameters p if

$$\mathrm{P}(X = k) = (1 - p)^{k-1} p, \quad k = 1, 2, \ldots$$

We write $X \sim \mathrm{Geom}(p)$.