MTH 463/563

Lectures 8 - 10

Yevgeniy Kovchegov
Oregon State University

Topics:

- Conditional probability.
- Independent and dependent events.
- Bayes' Theorem.
- Introduction to random variables.
- Binomial random variables.
- Expectation of a discrete random variable.

Conditional probability.

Given two events, A and B, in S. If $P(B) \neq 0$, the **conditional probability of A given B** is defined as follows

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

- **Example.** Roll two fair dice.

$\mathcal{S} =$

1, 1	1, 2	1, 3	1, 4	1, 5	1, 6
2, 1	2, 2	2, 3	2, 4	2, 5	2, 6
3, 1	3, 2	3, 3	3, 4	3, 5	3,6
4, 1	4, 2	4, 3	4, 4	4,5	4,6
5, 1	5, 2	5, 3	5,4	5,5	5,6
6, 1	6, 2	6,3	6,4	6,5	6,6

$E = \{$the sum is divisible by three$\}$ and $F = \{$ the sum is $\geq$ 9 $\}$

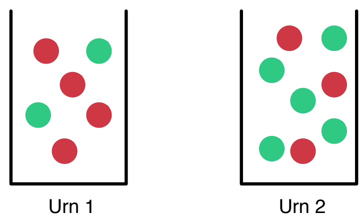
$$P(E) = \frac{|E|}{|\mathcal{S}|} = \frac{12}{36} = \frac{1}{3}, \quad P(F) = \frac{|F|}{|\mathcal{S}|} = \frac{10}{36} = \frac{5}{18}, \quad \text{and} \quad P(E \cap F) = \frac{|E \cap F|}{|\mathcal{S}|} = \frac{5}{36}$$

Then $P(E|F) = \frac{P(E\cap F)}{P(F)} = \frac{5/36}{10/36} = \frac{5}{10} = \frac{1}{2} \neq P(E)$

Conditional probability.

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \Rightarrow P(A \cap B) = P(A|B)P(B)$$

• **Example.** Urn 1 contains 4 red and 2 green marbles, while Urn 2 contains 3 red and 5 green marbles. We pick an urn at random with probability 1/2, and then select a marble from the urn with equal probability for each one. Find the probability that a red marble was selected from the second urn.



Let $U_2 = \{\text{Urn 2 selected}\}$ and $R = \{\text{red marble selected}\}$.

Then $P(U_2) = \frac{1}{2}$ and $P(R|U_2) = \frac{3}{8}$, and

$$P(R \cap U_2) = P(R|U_2)P(U_2) = \frac{3}{16}$$

Independent events. Two events, A and B, are said to be **independent** if

$$P(A \cap B) = P(A)P(B)$$

The above is equivalent to $P(A|B) = P(A)$ whenever $P(B) > 0$.

- **Example.** Roll two fair dice.

$\mathcal{S} =$

1, 1	1, 2	1, 3	1, 4	1, 5	1, 6
2, 1	2, 2	2, 3	2, 4	2, 5	2,6
3, 1	3, 2	3, 3	3, 4	3,5	3,6
4, 1	4, 2	4, 3	4,4	4,5	4,6
5, 1	5, 2	5,3	5,4	5,5	5,6
6, 1	6,2	6,3	6,4	6,5	6,6

$E = \{$the sum is divisible by three$\}$ and $F = \{$the sum is $\geq 8\}$

$$P(E) = \frac{|E|}{|\mathcal{S}|} = \frac{12}{36} = \frac{1}{3}, \quad P(F) = \frac{|F|}{|\mathcal{S}|} = \frac{15}{36} = \frac{5}{12}, \quad \text{and} \quad P(E \cap F) = \frac{|E \cap F|}{|\mathcal{S}|} = \frac{5}{36}$$

Then $P(E \cap F) = \frac{5}{36} = \frac{1}{3} \cdot \frac{5}{12} = P(E)P(F) \Rightarrow E$ and F are independent.

Dependent events. Two events, A and B, are said to be **dependent** if

$$P(A \cap B) \neq P(A)P(B)$$

The above is equivalent to $P(A|B) \neq P(A)$ whenever $P(B) > 0$.

- **Example.** Roll two fair dice.

$\mathcal{S} =$

1, 1	1, 2	1, 3	1, 4	1, 5	1, 6
2, 1	2, 2	2, 3	2, 4	2, 5	2, 6
3, 1	3, 2	3, 3	3, 4	3, 5	3,6
4, 1	4, 2	4, 3	4, 4	4,5	4,6
5, 1	5, 2	5, 3	5,4	5,5	5,6
6, 1	6, 2	6,3	6,4	6,5	6,6

$E = \{$the sum is divisible by three$\}$ and $F = \{$ the sum is $\geq$ 9 $\}$

$$P(E) = \frac{|E|}{|\mathcal{S}|} = \frac{12}{36} = \frac{1}{3}, \quad P(F) = \frac{|F|}{|\mathcal{S}|} = \frac{10}{36} = \frac{5}{18}, \quad \text{and} \quad P(E \cap F) = \frac{|E \cap F|}{|\mathcal{S}|} = \frac{5}{36}$$

Then $P(E \cap F) = \frac{5}{36} \neq \frac{5}{54} = P(E)P(F) \Rightarrow E$ and F are dependent.

Independent and dependent events.

- **Example.** Roll two fair dice.

$\mathcal{S} =$

1, 1	1, 2	1, 3	1, 4	1,5	1,6
2, 1	2, 2	2, 3	2, 4	2,5	2,6
3, 1	3, 2	3, 3	3, 4	3,5	3,6
4, 1	4, 2	4, 3	4, 4	4,5	4,6
5, 1	5, 2	5, 3	5, 4	5,5	5,6
6, 1	6, 2	6, 3	6, 4	6,5	6,6

$A = \{$the first die $\leq 3\}$ and $B = \{$ the second die is $\geq$ 5 $\}$

$$P(A) = \frac{18}{36} = \frac{1}{2}, \quad P(B) = \frac{12}{36} = \frac{1}{3}, \quad \text{and} \quad P(A \cap B) = \frac{|A \cap B|}{|\mathcal{S}|} = \frac{6}{36} = \frac{1}{6}$$

$P(A \cap B) = \frac{1}{6} = \frac{1}{2} \cdot \frac{1}{3} = P(A)P(B) \quad \Rightarrow \quad A$ and B independent.

Also notice $P(A|B) = \frac{1/6}{1/3} = \frac{1}{2} = P(A)$.

Independent and dependent events.

- **Proposition.** If the events E and F are independent, then so are E and $\overline{F}$

- **Proof:** We proved the following proposition

$$P(E) = P(E \cap F) + P(E \cap \overline{F})$$

Since E and F are independent, $P(E \cap F) = P(E)P(F)$ and therefore

$$P(E) = P(E \cap F) + P(E \cap \overline{F}) = P(E)P(F) + P(E \cap \overline{F})$$

So, $P(E) = P(E)P(F) + P(E \cap \overline{F})$ and

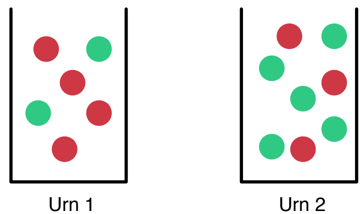
$$P(E \cap \overline{F}) = P(E) - P(E)P(F) = P(E)\big(1 - P(F)\big) = P(E)P(\overline{F})$$

as $1 - P(F) = P(\overline{F})$.

Conditional probability.

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \;\Rightarrow\; P(A \cap B) = P(A|B)P(B)$$

• **Example.** Urn 1 contains 4 red and 2 green marbles, while Urn 2 contains 3 red and 5 green marbles. We pick an urn at random with probability 1/2, and then select a marble from the urn with equal probability for each one. Find the probability that a red marble was selected from the second urn.



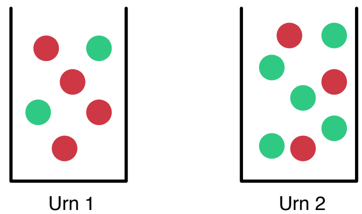
Let $U_2 = \{\text{Urn 2 selected}\}$ and $R = \{\text{red marble selected}\}$.

Then $P(U_2) = \frac{1}{2}$ and $P(R|U_2) = \frac{3}{8}$, and

$$P(R \cap U_2) = P(R|U_2)P(U_2) = \frac{3}{16}$$

Conditional probability.

• **Example.** Urn 1 contains 4 red and 2 green marbles, while Urn 2 contains 3 red and 5 green marbles. We pick an urn at random with probability 1/2, and then select a marble from the urn with equal probability for each one. Find the probability that a red marble was selected.

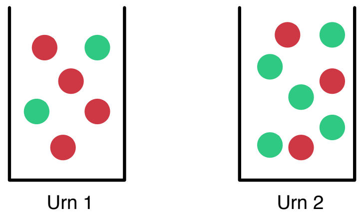


Let $U_1 = \{\text{Urn 1 selected}\}$, $U_2 = \{\text{Urn 2 selected}\}$, and $R = \{\text{red marble selected}\}$. Then $R = (R \cap U_1) \cup (R \cap U_2)$, and

$$P(R) = P(R \cap U_1) + P(R \cap U_2) = P(R|U_1)P(U_1) + P(R|U_2)P(U_2) = \frac{2}{3} \cdot \frac{1}{2} + \frac{3}{8} \cdot \frac{1}{2} = \frac{25}{48}$$

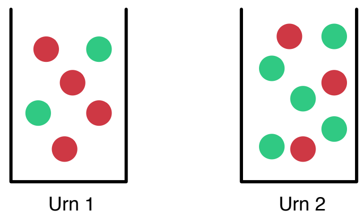
Bayes' Theorem.

• **Example.** Urn 1 contains 4 red and 2 green marbles, while Urn 2 contains 3 red and 5 green marbles. We pick an urn at random with probability 1/2, and then select a marble from the urn with equal probability for each one. We learned that a red marble was selected (but we don't know from what urn). Find the probability that it was taken from Urn 2.



Bayes' Theorem.

• **Example.** Urn 1 contains 4 red and 2 green marbles, while Urn 2 contains 3 red and 5 green marbles. We pick an urn at random with probability 1/2, and then select a marble from the urn with equal probability for each one. We learned that a red marble was selected (but we don't know from what urn). Find the probability that it was taken from Urn 2.

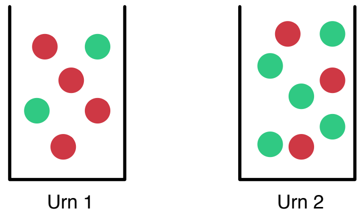


Let $U_1 = \{\text{Urn 1 selected}\}$, $U_2 = \{\text{Urn 2 selected}\}$, and $R = \{\text{red marble selected}\}$. We need to compute $P(U_2|R)$.

$$P(R) = P(R\cap U_1)+P(R\cap U_2) = P(R|U_1)P(U_1)+P(R|U_2)P(U_2) = \frac{2}{3}\cdot\frac{1}{2}+\frac{3}{8}\cdot\frac{1}{2} = \frac{25}{48}$$

$$P(U_2|R) = \frac{P(R\cap U_2)}{P(R)} = \frac{P(R|U_2)P(U_2)}{P(R|U_1)P(U_1)+P(R|U_2)P(U_2)} = \frac{3/16}{25/48} = \frac{9}{25}$$

Bayes' Theorem.



$$P(R) = P(R \cap U_1) + P(R \cap U_2) = P(R|U_1)P(U_1) + P(R|U_2)P(U_2) = \frac{2}{3} \cdot \frac{1}{2} + \frac{3}{8} \cdot \frac{1}{2} = \frac{25}{48}$$

$$P(U_2|R) = \frac{P(R \cap U_2)}{P(R)} = \frac{P(R|U_2)P(U_2)}{P(R|U_1)P(U_1) + P(R|U_2)P(U_2)} = \frac{3/16}{25/48} = \frac{9}{25}$$

Choices

$\swarrow \qquad \searrow$

Urn 1 | Urn 2

$\swarrow \quad \searrow \qquad \swarrow \quad \searrow$

RED	GREEN	RED	GREEN
$\frac{1}{2} \cdot \frac{2}{3}$	$\frac{1}{2} \cdot \frac{1}{3}$	$\frac{1}{2} \cdot \frac{3}{8}$	$\frac{1}{2} \cdot \frac{5}{8}$

Bayes' Theorem.

• **Theorem.** Suppose $B_1,\ B_2,\ \ldots, B_k$ are disjoint events such that

$$B_1 \cup B_2 \cup \cdots \cup B_k = \mathcal{S}$$

Then, for any event A,

• $P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \cdots + P(A|B_k)P(B_k)$

• (Bayes' Theorem) $P(B_j|A) = \frac{P(A|B_j)P(B_j)}{P(A|B_1)P(B_1)+P(A|B_2)P(B_2)+\cdots+P(A|B_k)P(B_k)}$ for any $j = 1, 2, \ldots, k$.

• **Proof.**

$A = A \cap \mathcal{S} = A \cap (B_1 \cup B_2 \cup \cdots \cup B_k) = (A \cap B_1) \cup (A \cap B_2) \cup \cdots \cup (A \cap B_k)$

is a union of disjoint events. Thus

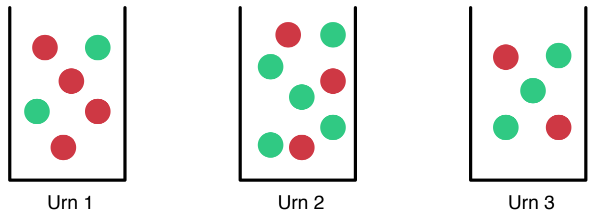
$$P(A) = P(A \cap B_1) + P(A \cap B_2) + \cdots + P(A \cap B_k)$$

$$= P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \cdots + P(A|B_k)P(B_k)$$

Next, $P(B_j|A) = \frac{P(A \cap B_j)}{P(A)} = \frac{P(A|B_j)P(B_j)}{P(A)}$.

Bayes' Theorem.

• **Example.** Let $U_1 = \{\text{Urn 1 selected}\}$, $U_2 = \{\text{Urn 2 selected}\}$, $U_3 = \{\text{Urn 3 selected}\}$, and $R = \{\text{red marble selected}\}$. We need to compute $P(R)$, $P(U_2|R)$, and $P(U_3|R)$.



$$P(R) = P(R|U_1)P(U_1)+P(R|U_2)P(U_2)+P(R|U_3)P(U_3) = \frac{2}{3}\cdot\frac{1}{3}+\frac{3}{8}\cdot\frac{1}{3}+\frac{2}{5}\cdot\frac{1}{3} = \frac{173}{360}$$

$$P(U_2|R) = \frac{P(R|U_2)P(U_2)}{P(R|U_1)P(U_1)+P(R|U_2)P(U_2)+P(R|U_3)P(U_3)} = \frac{1/8}{173/360} = \frac{45}{173}$$

$$P(U_3|R) = \frac{P(R|U_3)P(U_3)}{P(R|U_1)P(U_1)+P(R|U_2)P(U_2)+P(R|U_3)P(U_3)} = \frac{2/15}{173/360} = \frac{48}{173}$$

Examples.

• **Example.** We are given two coins. We know that one of the coins is fair (with probability 1/2 for each outcome), and one is fake, with probability 3/5 for heads and 2/5 for tails. We don't know which one is fair and which one is fake.

We randomly picked a coin (with equal probability for each one of the two), and tossed it once. The result was heads. With what probability can we conclude the coin we picked was the fake coin?

Examples.

- **Solution:** Let $F = \{$ the coin is fake $\}$,

and $H = \{$ coin toss yields heads $\}$.

Then $P(H|F) = \frac{3}{5}$, $P(H|\overline{F}) = \frac{1}{2}$, and $P(F) = P(\overline{F}) = \frac{1}{2}$.

Therefore, by Bayes' Theorem,

$$P(F|H) = \frac{P(H|F) \cdot P(F)}{P(H|F) \cdot P(F) + P(H|\overline{F}) \cdot P(\overline{F})} = \frac{\frac{3}{5} \cdot \frac{1}{2}}{\frac{3}{5} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2}} = \frac{6}{11}$$

is the probability that the coin we picked is fake.

Examples.

• **Example.** We are given three coins. We know that two of the coins are fair (with probability 1/2 for each outcome), and one is fake, with probability $7/8$ for heads and $1/8$ for tails. We don't know which one is fake.

We randomly picked a coin (with equal probability for each one of the three), and tossed it twice. The result was two heads. With what probability can we conclude the coin we picked was the fake coin?

Examples.

- **Solution:** Let $F = \{$ the coin is fake $\}$,

and $D = \{$ double heads $\}$.

Then $P(D|F) = \frac{7}{8} \cdot \frac{7}{8} = \frac{49}{64}$, $P(D|\overline{F}) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$,

$P(F) = \frac{1}{3}$, and $P(\overline{F}) = \frac{2}{3}$.

Therefore, by Bayes' Theorem,

$$P(F|D) = \frac{P(D|F) \cdot P(F)}{P(D|F) \cdot P(F) + P(D|\overline{F}) \cdot P(\overline{F})} = \frac{\frac{49}{64} \cdot \frac{1}{3}}{\frac{49}{64} \cdot \frac{1}{3} + \frac{1}{4} \cdot \frac{2}{3}} = \frac{49}{81} \approx 0.6049$$

is the probability that the coin we picked is fake.

Conditional probability

• **Example.** Let A and B be events of positive probability. Show that $P(A|B) > P(A)$ if and only if $P(B|A) > P(B)$.

Solution: Because $P(A|B) = \frac{P(A\cap B)}{P(B)}$ and $P(B|A) = \frac{P(A\cap B)}{P(A)}$, each of the two inequalities in the example is equivalent to

$$P(A \cap B) > P(A)P(B)$$

Comment: In this case we say that the events A and B are of positive correlation. Think of A being the event that your friend is eating a cookie, and B the event that the same friend is drinking coffee. Then knowing one increases the chances for the other.

Conditional probability

- **Multiplication Rule.**

$$P(E_1 \cap E_2 \cap \ldots \cap E_n) = P(E_1) \cdot P(E_2|E_1) \cdot P(E_3|E_1 \cap E_2) \cdot \ldots \cdot P(E_n|E_1 \cap E_2 \cap \ldots \cap E_{n-1})$$

Proof:

$$\begin{aligned} P(E_1) \cdot P(E_2|E_1) \cdot P(E_3|E_1 \cap E_2) \cdot \ldots \cdot P(E_n|E_1 \cap E_2 \cap \ldots \cap E_{n-1}) \\ = P(E_1) \cdot \frac{P(E_1 \cap E_2)}{P(E_1)} \cdot \frac{P(E_1 \cap E_2 \cap E_3)}{P(E_1 \cap E_2)} \cdot \ldots \cdot \frac{P(E_1 \cap E_2 \cap \ldots \cap E_n)}{P(E_1 \cap E_2 \cap \ldots \cap E_{n-1})} \\ = P(E_1 \cap E_2 \cap \ldots \cap E_n) \end{aligned}$$

Independent events.

- Two events, E and F, are said to be **independent** if

$$P(E \cap F) = P(E)P(F)$$

- Three events, E, F and G, are said to be **independent** if

$$P(E \cap F \cap G) = P(E)P(F)P(G)$$

$$P(E \cap F) = P(E)P(F)$$
$$P(E \cap G) = P(E)P(G)$$
$$P(F \cap G) = P(F)P(G)$$

Note: E is independent of the events formed of F and G:

$$P\Big(E \cap (F \cup G)\Big) = P(E)P(F \cup G) \quad \text{and}$$

$$P\bigg(E \cap \Big((F \cap \overline{G}) \cup (G \cap \overline{F})\Big)\bigg) = P(E)P\Big((F \cap \overline{G}) \cup (G \cap \overline{F})\Big)$$

Independent events.

In general, n events, $E_1, E_2, \ldots, E_n$, are said to be independent if for each k, E_k is independent of the events formed of

$$E_1, \ldots, E_{k-1}, E_{k+1}, \ldots, E_n$$

In other words, consider two non-overlapping subcollections

$$E_{i_1}, \ldots, E_{i_\ell} \quad \text{and} \quad E_{j_1}, \ldots, E_{j_m}$$

If is A an event created from the events in $E_{i_1}, \ldots, E_{i_\ell}$ and if B is an event created from the events in $E_{j_1}, \ldots, E_{j_m}$. Then, A and B are independent.

$$A \in \sigma\Big(E_{i_1}, \ldots, E_{i_\ell}\Big), \quad B \in \sigma\Big(E_{j_1}, \ldots, E_{j_m}\Big) \Rightarrow P(A \cap B) = P(A)P(B).$$

Independent events.

• **Example.** Toss two fair coins. Let E be the event that that the first coin lands heads side up, F be the event that that the second coin lands tails side up, and G be the event that one coin lands heads side up and one lands tails side side up.

Here $\mathcal{S} = \{HH, HT, TH, TT\}$ with equal probability of each outcome, and

$$E = \{HH, HT\} \qquad F = \{HT, TT\} \qquad G = \{HT, TH\}$$

Then E, F, and G are pairwise independent:

$$P(E \cap F) = P(E)P(F)$$
$$P(E \cap G) = P(E)P(G)$$
$$P(F \cap G) = P(F)P(G)$$

However, together, they are dependent:

$$P(E \cap F \cap G) = \frac{1}{4} \neq \frac{1}{8} = P(E)P(F)P(G)$$

and

$$P(G|E \cup F) = \frac{1}{3} \neq \frac{1}{2} = P(G)$$

Conditional probability is a probability.

Theorem. Let F be an event of positive probability, then

- $0 \leq P(E|F) \leq 1$ for all E in $\mathcal{S}$
- $P(\mathcal{S}|F) = 1$
- If $E_1,\ E_2, \ldots$ is a countable collection of disjoint events, then

$$P\left(\bigcup_{j=1}^{\infty} E_j \,\middle|\, F\right) = \sum_{j=1}^{\infty} P(E_j|F)$$

Proof: $\forall E \subseteq \mathcal{S}$, $(E \cap F) \subseteq F$ and $P(E \cap F) \leq P(F)$. Thus

$$0 \leq \ P(E|F) = \frac{P(E \cap F)}{P(F)} \ \leq 1$$

Next,

$$P(\mathcal{S}|F) = \frac{P(\mathcal{S} \cap F)}{P(F)} = \frac{P(F)}{P(F)} = 1$$

Conditional probability is a probability.

Theorem. Let F be an event of positive probability, then

- $0 \leq P(E|F) \leq 1$ for all E in $\mathcal{S}$
- $P(\mathcal{S}|F) = 1$
- If $E_1,\ E_2, \ldots$ is a countable collection of disjoint events, then

$$P\left(\bigcup_{j=1}^{\infty} E_j \,\middle|\, F\right) = \sum_{j=1}^{\infty} P(E_j|F)$$

Proof (continued): $E_1,\ E_2, \ldots$ are disjoint.

$$P\left(\bigcup_{j=1}^{\infty} E_j \,\middle|\, F\right) = \frac{P\left(\left(\bigcup_{j=1}^{\infty} E_j\right) \bigcap F\right)}{P(F)} = \frac{P\left(\bigcup_{j=1}^{\infty} (E_j \cap F)\right)}{P(F)} = \frac{\sum_{j=1}^{\infty} P(E_j \cap F)}{P(F)} = \sum_{j=1}^{\infty} P(E_j|F)$$

Note: Having $E_1 \cap F,\ E_2 \cap F, \ldots$ disjoint is sufficient for the above argument to work.

Conditional probability is a probability.

Theorem. Let F be an event of positive probability, then

- $0 \leq P(E|F) \leq 1$ for all E in $\mathcal{S}$

- $P(\mathcal{S}|F) = 1$

- If $E_1,\ E_2, \ldots$ is a countable collection of disjoint events, then
$$P\left(\bigcup_{j=1}^{\infty} E_j \;\middle|\, F\right) = \sum_{j=1}^{\infty} P(E_j|F)$$

So the axioms of probability apply to $P(\cdot|F)$. Therefore the Inclusion-Exclusion Theorem is valid for $P(\cdot|F)$,

$$P(A \cup B \ |F) = P(A|F) + P(B|F) - P(A \cap B \ |F)$$

as well as all other results about probability functions that we proved.

Bernoulli trials and Bernoulli random variables.

For a given $0 \le p \le 1$, a Bernoulli trial is an experiment with exactly two possible outcomes, success and failure, in which the probability of success is p and probability of failure is $1 - p$.

Here, the sample space $\mathcal{S}$ consists of the two outcomes, success and failure, and

$$P(\text{success}) = p \quad \text{and} \quad P(\text{failure}) = 1 - p$$

• **Example.** Consider tossing a coin such that it will fall heads up with some probability p, and tails up with probability $1 - p$.

Bernoulli random variable X counts the number of successes after one Bernoulli trial, and thus, can equal either 0 or 1, with probabilities $1 - p$ and p.

Here, $P(X = 0) = 1 - p$ and $P(X = 1) = p$.

Bernoulli trials

• **Example.** Consider performing independent Bernoulli trials, each with probability p of success and probability $1-p$ of failure. Let X be a random variable representing the number of successes in **two** Bernoulli trials. Find $P(X=k)$ for $k=0,1,2$.

• **Solution.** Let

$S_1 = \{\text{trial 1} \;\Rightarrow\; \text{success}\}$ and $F_1 = \overline{S}_1 = \{\text{trial 1} \;\Rightarrow\; \text{failure}\}$

$S_2 = \{\text{trial 2} \;\Rightarrow\; \text{success}\}$ and $F_2 = \overline{S}_2 = \{\text{trial 2} \;\Rightarrow\; \text{failure}\}$.

Then $P(S_1) = P(S_2) = p$ and $P(F_1) = P(F_2) = 1-p$, and

$$P(X=0) = P(\text{no successes}) = P(F_1 \cap F_2) = P(F_1)P(F_2) = (1-p)^2$$

$$P(X=1) = P(\text{one success and one failure}) = P(S_1 \cap F_2) + P(F_1 \cap S_2) = 2p(1-p)$$

$$P(X=2) = P(\text{two successes}) = P(S_1 \cap S_2) = P(S_1)P(S_2) = p^2$$

by independence.

Bernoulli trials.

• **Example.** Consider performing independent Bernoulli trials, each with probability p of success and probability $1-p$ of failure. Let X be a random variable representing the number of successes in **three** Bernoulli trials. Find $P(X=k)$ for $k=0,1,2,3$.

• **Solution.** Let

$S_1 = \{\text{trial 1} \;\Rightarrow\; \text{success}\}$ and $F_1 = \overline{S}_1 = \{\text{trial 1} \;\Rightarrow\; \text{failure}\}$

$S_2 = \{\text{trial 2} \;\Rightarrow\; \text{success}\}$ and $F_2 = \overline{S}_2 = \{\text{trial 2} \;\Rightarrow\; \text{failure}\}$

$S_3 = \{\text{trial 3} \;\Rightarrow\; \text{success}\}$ and $F_3 = \overline{S}_3 = \{\text{trial 3} \;\Rightarrow\; \text{failure}\}$.

Then, by independence,

$$P(X=0) = P(F_1 \cap F_2 \cap F_3) = \binom{3}{0}(1-p)^3$$

$$P(X=1) = P(S_1 \cap F_2 \cap F_3) + P(F_1 \cap S_2 \cap F_3) + P(F_1 \cap F_2 \cap S_3) = \binom{3}{1}p(1-p)^2$$

$$P(X=2) = P(S_1 \cap S_2 \cap F_3) + P(S_1 \cap F_2 \cap S_3) + P(F_1 \cap S_2 \cap S_3) = \binom{3}{2}p^2(1-p)$$

$$P(X=3) = P(S_1 \cap S_2 \cap S_3) = \binom{3}{3}p^3$$

Bernoulli trials

• **Example.** Consider performing independent Bernoulli trials, each with probability p of success and probability $1-p$ of failure. Let X be a random variable representing the number of successes in n Bernoulli trials. Find $P(X=k)$ for $k=0,1,\ldots,n$.

• **Solution.** Let

$$\begin{array}{lll} S_1 = \{\text{trial } 1 \ \Rightarrow \ \text{success}\} & \text{and} & F_1 = \overline{S}_1 = \{\text{trial } 1 \ \Rightarrow \ \text{failure}\} \\ S_2 = \{\text{trial } 2 \ \Rightarrow \ \text{success}\} & \text{and} & F_2 = \overline{S}_2 = \{\text{trial } 2 \ \Rightarrow \ \text{failure}\} \\ \vdots & \vdots & \vdots \\ S_n = \{\text{trial } n \ \Rightarrow \ \text{success}\} & \text{and} & F_n = \overline{S}_n = \{\text{trial } n \ \Rightarrow \ \text{failure}\}. \end{array}$$

Then, for each outcome with k successes and $n-k$ failures, its probability

$$P(\underbrace{SFSS\ldots FFS}_{k\ S's\ \text{and}\ n-k\ F's}) = P(S_1)P(F_2)P(S_3)P(S_4)\ldots P(F_{n-2})P(F_{n-1})P(S_n) = p^k(1-p)^{n-k}$$

and $$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k} \quad \text{for each } k=0,1,\ldots,n$$

because there are $\binom{n}{k}$ such outcomes.

Introduction to random variables.

Consider a sample space $\mathcal{S}$ and a probability function P.

- **Definition.** A function from $\mathcal{S}$ to $\mathbb{R}$ is a **random variable**.

- **Example.** Roll two fair dice. Let $X(i,j) = i + j$ for each outcome (i,j) in $\mathcal{S}$. Then X is a random variable representing the sum of the digits on the dice.

$$X : \begin{array}{|c|c|c|c|c|c|}\hline 1,1 & 1,2 & 1,3 & 1,4 & 1,5 & 1,6 \\ \hline 2,1 & 2,2 & 2,3 & 2,4 & 2,5 & 2,6 \\ \hline 3,1 & 3,2 & 3,3 & 3,4 & 3,5 & 3,6 \\ \hline 4,1 & 4,2 & 4,3 & 4,4 & 4,5 & 4,6 \\ \hline 5,1 & 5,2 & 5,3 & 5,4 & 5,5 & 5,6 \\ \hline 6,1 & 6,2 & 6,3 & 6,4 & 6,5 & 6,6 \\ \hline \end{array} \longrightarrow \left\{ \begin{array}{c} 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \\ 11 \\ 12 \end{array} \right\}$$

Introduction to random variables.

- **Example.** Roll two fair dice. Let $X(i,j) = i + j$ for each outcome (i,j) in $\mathcal{S}$. Then X is a random variable representing the sum of the digits on the dice.

$$X : \begin{array}{|c|c|c|c|c|c|} \hline 1,1 & 1,2 & 1,3 & 1,4 & 1,5 & 1,6 \\ \hline 2,1 & 2,2 & 2,3 & 2,4 & 2,5 & 2,6 \\ \hline 3,1 & 3,2 & 3,3 & 3,4 & 3,5 & 3,6 \\ \hline 4,1 & 4,2 & 4,3 & 4,4 & 4,5 & 4,6 \\ \hline 5,1 & 5,2 & 5,3 & 5,4 & 5,5 & 5,6 \\ \hline 6,1 & 6,2 & 6,3 & 6,4 & 6,5 & 6,6 \\ \hline \end{array} \longrightarrow \left\{ \begin{array}{c} 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \\ 11 \\ 12 \end{array} \right\}$$

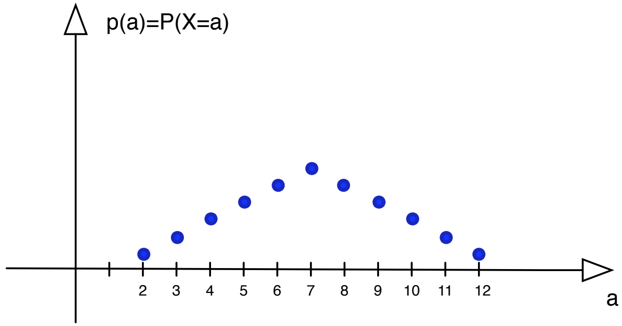
Here, for example, $X(3,1) = 4$ and $X(5,6) = 11$.

We are interested in finding the following probabilities:

$$p(a) = P(X = a) \quad \text{for} \quad a = 2, 3, \ldots, 12$$

Introduction to random variables.

- **Example.** Roll two fair dice. Let $X(i,j) = i + j$ for each outcome (i,j) in $\mathcal{S}$. Then X is a random variable representing the sum of the digits on the dice.



We are interested in finding the following probabilities:

$$p(a) = P(X = a) \quad \text{for} \quad a = 2, 3, \ldots, 12$$

$$p(2) = \tfrac{1}{36},\ p(3) = \tfrac{2}{36},\ p(4) = \tfrac{3}{36},\ p(5) = \tfrac{4}{36},\ p(6) = \tfrac{5}{36},\ p(7) = \tfrac{6}{36}$$

$$p(8) = \tfrac{5}{36},\ p(9) = \tfrac{4}{36},\ p(10) = \tfrac{3}{36},\ p(11) = \tfrac{2}{36},\ p(12) = \tfrac{1}{36}$$

Introduction to random variables.

Let X a discrete random variable. That is X assumes a discrete (countable) number of values.

- **Definition.** Function $p(a) = P(X = a)$ is called the probability mass function (or distribution function).

- **Definition.** Function $F(a) = P(X \leq a)$ is called the cumulative distribution function.

- **Note.** $$\sum_{a:\ p(a)>0} p(a) = 1$$

  In the previous example, $p(2)+p(3)+\cdots+p(12) = 1$.

- **Note.** $0 \leq F(a) \leq 1$

- **Note.** $F(a) = \sum_{x:\ x \leq a} p(x)$

Binomial random variable. Recall the following example.

- **Example.** Consider performing independent Bernoulli trials, each with probability p of success and probability $1-p$ of failure. Let X be a random variable representing the number of successes in n Bernoulli trials. Find $P(X=k)$ for $k=0,1,\ldots,n$.

- **Solution.**

Each outcome with k successes and $n-k$ failures, its probability

$$P(\underbrace{SFSS\ldots FFS}_{k\ S's \text{ and } n-k\ F's}) = p^k(1-p)^{n-k}$$

and $$P(X=k) = \binom{n}{k}p^k(1-p)^{n-k} \quad \text{for each } k=0,1,\ldots,n$$

because there are $\binom{n}{k}$ such outcomes.

- **Definition.** The random variable X in the above example is the binomial random variable with parameters (n,p).

Check: $\displaystyle\sum_{k=0}^{n} p(k) = \sum_{k=0}^{n} \binom{n}{k}p^k(1-p)^{n-k} = \Big(p+(1-p)\Big)^n = 1^n = 1.$

Expectation of a discrete random variable.

• **Definition.** Let X be a discrete random variable with the probability mass function $p(x)$. Then its expected value is

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

• **Example.** Let X be a Bernoulli random variable with parameter p. Then

$$p(1) = P(X = 1) = p \quad \text{and} \quad p(0) = P(X = 0) = 1-p$$

and

$$E[X] = 0 \cdot p(0) + 1 \cdot p(1) = p$$

• **Example.** Roll two fair dice. Let X represent the sum of the digits on the dice. Then

$$E[X] = 2 \cdot p(2) + 3 \cdot p(3) + \cdots + 12 \cdot p(12) = \frac{252}{36} = 7$$

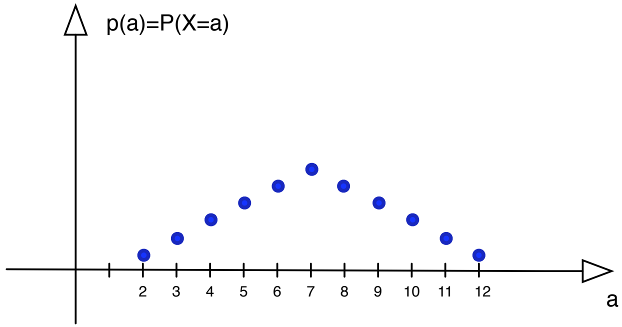
Expectation of a discrete random variable.

• **Definition.** Let X be a discrete random variable with the probability mass function $p(x)$. Then its **expected value** is

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

• **Example.** Roll two fair dice. Let X represent the sum of the digits on the dice. Then

$$E[X] = 2 \cdot p(2) + 3 \cdot p(3) + \cdots + 12 \cdot p(12) = \frac{252}{36} = 7$$



This corresponds to a **center of mass** of $p(a)$.

Expectation of a discrete random variable.

Let X be a binomial random variable with parameters (n, p). Then its probability mass function is known to be

$$p(k) = \binom{n}{k} p^k (1-p)^{n-k} \quad \text{for each } k = 0, 1, \ldots, n$$

• **Definition.** Let X be a discrete random variable with the probability mass function $p(x)$. Then its expected value is

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

• **Example.** Let X be a binomial random variable with parameters (n, p). Then

$$E[X] = \sum_{k=0}^{n} k \cdot p(k) = \sum_{k=0}^{n} k \cdot \binom{n}{k} p^k (1-p)^{n-k} =?$$

Expectation of a discrete random variable.

• **Example.** Let X be a **binomial random variable** with parameters (n, p). Then $E[X] = np$ since

$$E[X] = \sum_{k=0}^{n} k \cdot p(k) = \sum_{k=0}^{n} k \cdot \binom{n}{k} \; p^k (1-p)^{n-k}$$

$$= \sum_{k=1}^{n} k \cdot \frac{n!}{k!(n-k)!} \; p^k (1-p)^{n-k} = \sum_{k=1}^{n} \frac{n!}{(k-1)!(n-k)!} \; p^k (1-p)^{n-k}$$

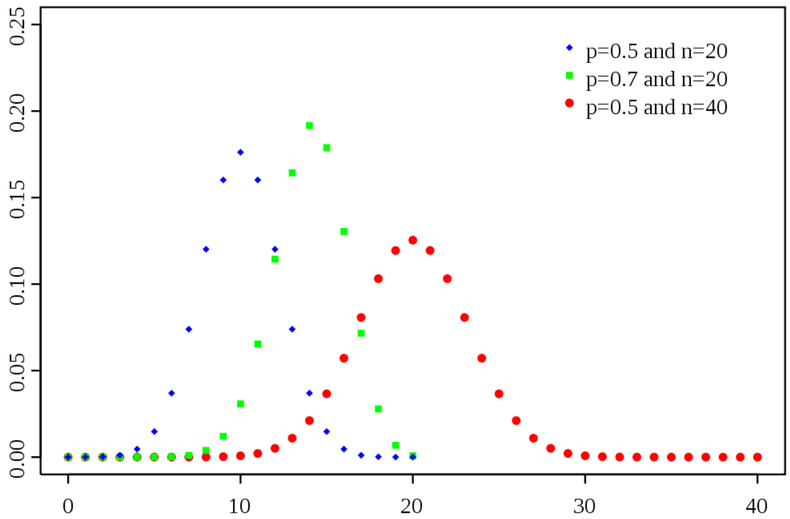
$$= \sum_{j=0}^{n-1} \frac{n!}{j!(n-1-j)!} \; p^{j+1} (1-p)^{n-1-j} = np \cdot \sum_{j=0}^{n-1} \frac{(n-1)!}{j!(n-1-j)!} \; p^j (1-p)^{n-1-j},$$

where the new index $j = k - 1$. Thus

$$E[X] = np \cdot \sum_{j=0}^{n-1} \frac{(n-1)!}{j!(n-1-j)!} \; p^j (1-p)^{n-1-j} = np \cdot \sum_{j=0}^{n-1} \binom{n-1}{j} \; p^j (1-p)^{n-1-j}$$

$$= np \cdot (p + (1-p))^{n-1} = np \qquad \text{by the Binomial theorem}$$

Binomial random variable.



Picture credit: Wikipedia.org

$p(k) = \binom{n}{k} p^k (1-p)^{n-k}$ for each $k = 0, 1, \ldots, n$ and $E[X] = np$