

Discrete Structures. CSCI-150. Fall 2013.

Probability.

Intro.

In any probabilistic problem, you have to identify the source of randomness and uncertainty. Find, what are the events that can occur, and quantitatively describe them.

The walking ant example. Coin tossing example.

First, you need to identify two things:

- (1) The set of all possible outcomes, called **sample space**, denoted by Ω (or S in Rosen).

$$\Omega = \{\text{Heads, Tails}\}$$

- (2) The probability of each individual outcome.

$$P(\text{Heads}) = 1/2, \quad P(\text{Tails}) = 1/2.$$

Another example. Rolling a die once:

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

$$P(i) = 1/6, \quad \text{for } i = 1, \dots, 6$$

There are a few requirements. The **probability** of each outcome must be non-negative, and the sum of the probabilities of all outcomes must be equal to 1:

$$\sum_{u \in \Omega} P(u) = 1$$

An **event** is a subset of Ω .

Consider an event A that happens when you roll a die and get an odd number: $A = \{1, 3, 5\}$. Subsets \emptyset , $\{1\}$, $\{2\}$, $\{4, 5, 6\}$, $\{1, 2, 3, 4, 5, 6\}$ are all valid events too. What do they mean? (describe in English.)

The **probability of an event** is the sum of the probabilities of the outcomes it contains:

$$P(A) = \sum_{u \in A} P(u)$$

$$P(\{1, 3, 5\}) = P(1) + P(3) + P(5) = 1/6 + 1/6 + 1/6 = 1/2$$

So, the probability that the result of a die roll is odd is one half.

Four possible events in coin tossing:

$$P(\emptyset) = 0, \quad P(\{H\}) = 1/2, \quad P(\{T\}) = 1/2, \quad P(\{H, T\}) = 1$$

Conditional probability.

Conditional probability of A , given B . What is the probability that one event, A , happens, given that some other event, B , definitely happens?

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

Example. We roll a die and get x . Given that $x \leq 3$, what is the probability that x is odd?

A : x is odd

B : $x \leq 3$

$$P(A | B) = \frac{P(A \cap B)}{P(B)} = \frac{P(\{1, 3\})}{P(\{1, 2, 3\})} = \frac{1/6 + 1/6}{1/6 + 1/6 + 1/6} = 2/3$$

Example 2. A bit string of length four is generated at random so that each of the 16 bit strings of length four is equally likely. What is the probability that it contains at least two consecutive 0s, given that its first bit is a 0? (We assume that 0 bits and 1 bits are equally likely.)

Note that if you cannot simply list all the elements from the sets $A \cap B$ and B because the sets are too large, apply counting techniques.

What is the probability that it will rain this afternoon, given that it is cloudy this morning?

What is the probability that two rolled dice sum to 10, given that both are odd?

Very good example in LL: 19.1, p.246: The Halting problem (about MIT EECS hockey team).

Independent events and Bernoulli trials.

		Nickel	Dime
A box with four types of coins (or with four compartments).	US	4	12
	Canada	8	24

Take one coin from the box (all choices are equally likely). $P(\text{Nickel} | \text{US}) = \frac{1}{4} = P(\text{Nickel})$.

Two events A and B are independent if $P(A | B) = P(A)$. The occurrence of A does not depend on B in any way.

Since $P(A | B) = \frac{P(A \cap B)}{P(B)}$, A and B are independent if and only if

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

Example. Consider a **biased coin**. When tossing it, with probability p , we get 1, otherwise 0:

$$P(1) = p, \quad P(0) = 1 - p.$$

For a **fair coin**, $p = 1/2$.

When a coin is flipped, the possible outcomes are heads and tails. Each performance of an experiment with two possible outcomes is called a Bernoulli trial, after James Bernoulli. In

general, a possible outcome of a Bernoulli trial is called a success or a failure. Bernoulli trials are mutually independent if the conditional probability of success on any given trial is p , given any information whatsoever about the outcomes of the other trials. We can use 1 to denote a success, and 0 to denote a failure. Therefore, the result of n trials is a bit-string of length n .

Let x_i be the result if the i^{th} trial. A sequence $x_1 \dots x_n$ is a string of 0s and 1s, produced by a sequence of n trials. The trials are independent, so the result of each particular trial does not depend on the history:

$$\begin{aligned} P(x_i = 1 \mid x_1 \dots x_{i-1}) &= P(x_i = 0) = p \\ P(x_i = 0 \mid x_1 \dots x_{i-1}) &= P(x_i = 1) = 1 - p \end{aligned}$$

Consider a particular resulting string of 1s and 0s:

10011101

The probability to generate this string is

$$p(1-p)(1-p)ppp(1-p)p = p^5(1-p)^3$$

For shorter notation, it's common to denote $q = 1 - p$.

In general, when there are k successes and m failures in a given sequence, the probability to generate the sequence is

$$p^k(1-p)^m$$

The probability of exactly k successes in n trials. Binomial distribution.

What is the probability that in a sequence of n Bernoulli trials with probability of success p , you get exactly k successes? (An alternative formulation: after flipping a coin n times, what is the probability to get heads exactly k times?)

The answer is not $p^k(1-p)^{n-k}$, because there is more than one sequence of n trials that results in k successes. Let A be the set of all such sequences.

$$P(A) = \sum_{(x_1 \dots x_n) \in A} P(x_1 \dots x_n) = \sum_{(x_1 \dots x_n) \in A} p^k(1-p)^{n-k} = |A| \cdot p^k(1-p)^{n-k}$$

(Note that the last equality is correct, because all sequences from A have the same probability $p^k(1-p)^{n-k}$.)

We know that there are $\binom{n}{k}$ ways to select k elements from a set of n elements. Therefore, the probability is

$$\sum_{u \in A} p^k(1-p)^{n-k} = \binom{n}{k} p^k(1-p)^{n-k}$$

Probability of a complement

If we know $P(A)$, then

$$P(\bar{A}) \equiv P(\Omega \setminus A) = 1 - P(A)$$

Very useful sometimes, e.g. when the event A is defined by a predicate Q , i.e. $A = \{u \in \Omega \mid Q(u)\}$, and then its complement $\bar{A} = \{u \in \Omega \mid \neg Q(u)\}$.

Random variable

What if we are interested not in subsets of Ω , but in some numerical property of the outcomes. E.g., Ω is a set of people in the room, and we are interested in the height of a randomly selected person $x \in \Omega$.

A **random variable** is a function from the sample space Ω to \mathbb{R} . That is, a random variable assigns a real number to each possible outcome

$$X : \Omega \rightarrow \mathbb{R}$$

Examples:

1. The height of a randomly selected person.
2. The number of successes in n Bernoulli trials.
3. You are playing a game such that on every stage you either win or lose \$1. It continues indefinitely, until you run out of money. Let N_0 be the amount of money you have before the game started. Then the r.v. N_i describes the amount of money you have after the i^{th} stage of the game. Another r.v., Z , is the stage number, when you lose the last dollar.

(Remember that all these random variables are just functions, they map every particular sequence of wins and losses to a number.)

Probability distribution

The **probability distribution** (probability mass function) of a random variable X assigns the probability to each possible value of X :

$$P(X = r)$$

where $r \in X(\Omega)$, i.e. the probability is defined for each r is in the image of the function X – remember that a random variable is nothing more than a real-valued function.

Example:

Let r.v. X be the sum of the numbers that appear when a pair of dice is rolled. What is the probability distribution of X ?

$$\Omega = \{1, 2, 3, 4, 5, 6\}^2 \quad (\text{the set of pairs } (k, m))$$

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Consider one particular value $X = 7$. What is the probability that $P(X = 7)$?

Each pair (k, m) is equally likely, with probability $P((k, m)) = (1/6)^2 = 1/36$, since the rolls are independent, and there are six outcomes with $k + m = 7$:

$$P(X = 7) = \frac{6}{36} = \frac{1}{6}$$

Alternatively,

Let's assume that in the first roll we got k . What's the chance that in the second roll we will have enough to get 7 in total?

(Let R_1 and R_2 be the r.v. corresponding to the individual results of the 1st and 2nd roll respectively.)

Notice that regardless of the outcome of the first roll R_1 , there is always a chance (with probability $1/6$) that in the second roll you get the exactly matching value $R_2 = 7 - R_1$, so

$$P(X = 7) = \dots = P(R_2 = (7 - k) \mid R_1 = k) \stackrel{indep.}{=} P(R_2 = (7 - k)) = \frac{1}{6}$$

So, in other words, the first roll determines the column in the table. And in the second roll, one of the six possibilities results in having $X = 7$ in total.

More detailed (notation is being rather lengthy)

$$\begin{aligned} P(X = 7) &= \sum_{k=1}^6 P(\underbrace{R_1 = k, R_2 = (7 - k)}_{\{u \mid R_1(u)=k\} \cap \{u \mid R_2(u)=7-k\}}) = \sum_{k=1}^6 P(R_2 = (7 - k) \mid R_1 = k) \cdot P(R_1 = k) \\ &= \sum_{k=1}^6 P(R_2 = (7 - k) \mid R_1 = k) \cdot \frac{1}{6} = \\ &= P(R_2 = (7 - k) \mid R_1 = k) \\ &= P(R_2 = (7 - k)) = \frac{1}{6} \end{aligned}$$

Notice the notation: for a joint event $A \cap B$, it's common to write $P(A, B)$, which is equivalent to $P(A \cap B)$.

Binomial distribution

The probability to get k successes in a sequence of n independent Bernoulli trials:

$$b(k; n, p) = P(k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

Expected value

The **expectation** or mean, of the random variable X on the sample space Ω is

$$E(X) = \sum_{u \in \Omega} P(u) \cdot X(u)$$

The expected value of a die roll:

$$\sum_{k=1}^6 \frac{k}{6} = \frac{1}{6} + \frac{2}{6} + \frac{3}{6} + \frac{4}{6} + \frac{5}{6} + \frac{6}{6} = \left(\frac{6 \cdot 7}{2}\right) \cdot \frac{1}{6} = \frac{7}{2} = 3.5$$

Alternatively, we can compute the expected value by enumerating all values of the random variable, instead of enumerating the individual outcomes:

$$E(X) = \sum_{k \in X(\Omega)} P(X = k) \cdot k$$

Geometric distribution

We flip a biased coin, $P(T) = p$, $P(H) = 1 - p$. (Alternatively, we could say that this is a series of independent Bernoulli trials).

What is the expected number of flips until the coin comes up tails?

$$\begin{aligned} P(X = 1) &= p \\ P(X = 2) &= (1 - p)p \\ P(X = 3) &= (1 - p)^2 p \\ &\dots \\ P(X = k) &= (1 - p)^{k-1} p \end{aligned}$$

Thus

$$E(X) = \sum_{k=1}^{\infty} P(X = k) \cdot k = \sum_{k=1}^{\infty} (1 - p)^{k-1} p \cdot k$$

Do we know how to compute such a sum? Essentially, we need a formula for

$$\sum_{k=1}^{\infty} q^{k-1} k$$

We know the sum of an infinite geometric series

$$\sum_{k=1}^{\infty} q^k = q \sum_{k=0}^{\infty} q^k = \frac{q}{1 - q} \quad (\text{when } |q| < 1)$$

Take the derivative with respect to q :

$$\sum_{k=1}^{\infty} q^{k-1} \cdot k = \frac{1(1 - q) - (-1)q}{(1 - q)^2} = \frac{1 - q + q}{(1 - q)^2} = \frac{1}{(1 - q)^2} \quad (\text{when } |q| < 1)$$

Thus

$$E(X) = \sum_{k=1}^{\infty} (1-p)^{k-1} p \cdot k = \left(\sum_{k=1}^{\infty} (1-p)^{k-1} \cdot k \right) p = \frac{1}{(1 - (1-p))^2} \cdot p = \frac{1}{p^2} \cdot p = \frac{1}{p}$$

If $p = 1/2$ (a fair coin), then we expect to wait only two flips. If $p = 1/10$, we expect to wait 10 flips, etc.

We say that this random variable X has a **geometric distribution** with parameter p ,

$$P(X = k) = (1-p)^{k-1} p$$

The expected value of such r.v. is

$$E(X) = \frac{1}{p}$$

Bayes' Theorem

Suppose that A and B are events from a sample space Ω such that $p(A) \neq 0$ and $p(B) \neq 0$.

By the definition of conditional probability, $P(B \cap A) = P(B | A)P(A)$, and $P(B \cap A) = P(A | B)P(B)$. Therefore, the right-hand sides of these two equations are equal and

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

Since $P(A) = P(A \cap B) + P(A \cap \overline{B}) = P(A | B)P(B) + P(A | \overline{B})P(\overline{B})$, we can get an extended version of the same formula

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

Example

Suppose that one person in 100,000 has a particular rare disease for which there is a fairly accurate diagnostic test. This test is correct 99.0% of the time when given to a person selected at random who has the disease; it is correct 99.5% of the time when given to a person selected at random who does not have the disease. Given this information can we find

- (a) the probability that a person who tests positive for the disease has the disease?
- (b) the probability that a person who tests negative for the disease does not have the disease?

Should a person who tests positive be very concerned that he or she has the disease?

To solve, choose A to be the event that a person selected at random has the disease, and B to be the event that a person selected at random is tested positively.

Problem 1.

We randomly select a permutation of $\{1, 2, 3, 4, 5, 6\}$ (all permutations are equally likely). What is the probability that 1 precedes 2?

Problem 2.

What is the conditional probability that exactly four heads appear when a fair coin is flipped five times, given that the first flip came up heads?

Problem 3.

What is the probability that in a series of five independent Bernoulli trials with probability of success p :

- (a) all are successes?
- (b) all are failures?
- (c) only the first is a success?
- (d) there are at least 3 successes?

Problem 4 (Birthdays).

What is the minimum number of people who need to be in a room so that the probability that at least two of them have the same birthday is greater than $1/2$?

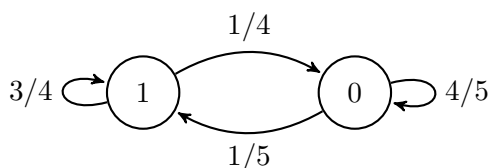
Assume that there are $n = 366$ days in a year, and all birthdays are independent and equally likely.

Problem 5 (Markov chain piano player).

There is a bad piano player called Bob.

If Bob plays a note correctly, then with probability $3/4$, he will play correctly the next note as well. However, if he plays a wrong note, then with probability $4/5$ the next note will be wrong.

$$\begin{aligned}P(X_i = 1 \mid X_{i-1} = 1) &= 3/4 \\P(X_i = 0 \mid X_{i-1} = 1) &= 1/4 \\P(X_i = 1 \mid X_{i-1} = 0) &= 1/5 \\P(X_i = 0 \mid X_{i-1} = 0) &= 4/5\end{aligned}$$



Bob plays a short five note melody, and the first note is guaranteed to be correct.

- (a) What is the probability that he will make no mistakes?

- (b) What is the probability that he will make exactly one mistake?
- (c) What is the probability that he will make at least one mistake?

Problem 6.

Find each of the following probabilities when n independent Bernoulli trials are carried out with probability of success p .

- (a) the probability of no successes
- (b) the probability of at least one success
- (c) the probability of at most one success
- (d) the probability of at least two successes
- (e) the probability of no failures
- (f) the probability of at least one failure
- (g) the probability of at most one failure
- (h) the probability of at least two failures