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 occure, 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).

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

(2) The probability of each individual outcome.

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

Another example. Rolling a die once:

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

$$P(i) = 1/6, \qquad \text{for } i = 1, \ldots, 6$$

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 $\varnothing$, $\{1\}$, $\{2\}$, $\{4, 5, 6\}$, $\{1, 2, 3, 4, 5, 6\}$ 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:

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

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?

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

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

$$A : \quad x \text{ is odd}$$
$$B : \quad x \leq 3$$
$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(\{1,3\})}{P(\{1,2,3\})} = \frac{1/6 + 1/6}{1/6 + 1/6 + 1/6} = 2/3$$

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.

A box with four types of coins (or with four compartments).

	Nickel	Dime
US	4	12
Canada	8	24

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

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

Since $P(A \mid B) = \dfrac{P(A \cap B)}{P(B)}$, 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 ith trial. A sequence $x_1 \ldots x_n$ 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:

$$P(x_i = 1 \mid x_1 \ldots x_{i-1}) = P(x_i = 0) = p$$
$$P(x_i = 0 \mid x_1 \ldots x_{i-1}) = P(x_i = 1) = 1 - p$$

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 sequnce 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 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.

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

(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 elelments. Therefore, the probability is
$$\sum_{u \in A} p^k(1-p)^{n-k} = \binom{n}{k} p^k(1-p)^{n-k}$$

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.

$$P(X_i = 1 \mid X_{i-1} = 1) = 3/4$$
$$P(X_i = 0 \mid X_{i-1} = 1) = 1/4$$
$$P(X_i = 1 \mid X_{i-1} = 0) = 1/5$$
$$P(X_i = 0 \mid X_{i-1} = 0) = 4/5$$
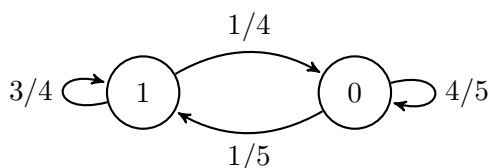
Bob plays a short five note melody, and the first note is guaranteed to be correct.

(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