

1 Probability

1.1 Combinatorial Analysis

The *basic principle of counting*: If an experiment consists of two phases, with n possible outcomes in the first phase, and m possible outcomes in the second phase for each of those outcomes, then the total possible number of outcomes is nm .

There are $n!$ possible linear orderings of n items.

The number of distinct ways to choose k items from n without regard to order is given by

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

For non-negative integers n_1, n_2, \dots, n_r with $\sum n_i = n$, the number distinct of ways to divide n items into non-overlapping groups of sizes n_1, n_2, \dots, n_r is given by

$$\binom{n}{n_1, n_2, \dots, n_r} = \frac{n!}{n_1!n_2! \cdots n_r!}$$

1.2 Axioms of Probability

Some notation and terminology:

- S is the set of all possible outcomes of an experiment, a.k.a. the *sample space*
- An *event* is a subset of S
- For any event A we define A^c to be the complement of A in S
- $S^c = \emptyset$ is the null set
- $A \cap B$ or sometimes AB is the intersection of the sets A and B . If $A \cap B = \emptyset$ then we say that A and B are mutually exclusive

Axioms

If E is an event in a sample space S then the probability that event E occurs as the outcome of a single experiment is denoted by $P(E)$ and satisfies the following axioms

1. $0 \leq P(E) \leq 1$
2. $P(S) = 1$
3. If $\{E_i\}_{i=1}^{\infty}$ is a set of pairwise mutually exclusive events then $P(\cup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} P(E_i)$

Useful Results

$$P(A^c) = 1 - P(A)$$

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

(And the generalization of this, informally: add up all of the probabilities, subtract all of the double counting of intersections, add back all of the double removing of triple intersections, remove all of the restored double counting of quadruple intersections, etc.)

If S is finite and every point in S is assumed to have equal probability of occurring then

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

1.3 Conditional Probability and Independence

Conditional Probability

For events E and F , the conditional probability of E occurs, given that F has occurred is denoted by $P(E|F)$ and is defined by

$$\begin{aligned}P(E|F) &= \frac{P(E \cap F)}{P(F)} \\&= \frac{P(F|E)P(E)}{P(F)} \\&= \frac{P(F|E)P(E)}{P(F|E)P(E) + P(F|E^c)P(E^c)}\end{aligned}$$

where the last line is *Bayes's Formula*

Independence

Two events are said to be *independent* if $P(E \cap F) = P(E)P(F)$. Events that are not independent are said to be *dependent*. This generalizes to sets of events if every the probability of intersection of subsets is equal to the product of the probabilities of the individual events in the subset.

1.4 Random Variables

Random Variables

Given an experiment with sample space S , a *random variable* is a real-valued function, X , defined on S . Because the value of a random variable is determined by the outcome of the experiment, we can assign probabilities to the possible values of the random variable, denoted by $P(X = x)$.

Discrete Random Variables

If the range of X is countable, $\{x_1, x_2, \dots\}$, then X is said to be discrete. In this case we can define the *probability mass function* $p : \mathbb{R} \rightarrow \mathbb{R}$ of X to be

$$p(a) = P(X = a)$$

which has the following properties:

1. $p(x_i) \geq 0$ for $i = 1, 2, \dots$
2. $p(x) = 0$ for any other values of x
3. $\sum_{i=1}^{\infty} p(x_i) = 1$

Expected Value