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

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

Stars and Bars

For any positive integers n and k, the number of k-tuples of positive integers that sum to n is $\binom{n-1}{k-1}$
For any positive integers n and k, the number of k-tuples of non-negative integers that sum to n is $\binom{n+k-1}{k-1}$

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 complent 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 $P(\cup_{i=1}^{\infty}) = \sum_{i=1}^{\infty} P(E_i)$

Useful Results

$$P(A^c) = 1 - P(A)$$
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Informally, the generalization of this formula: add up all of the individual event probabilities, then subtract all of the double counted intersections, then add back all of the twice removed triple intersections, then remove all of the twice restored double counted quadruple intersections, etc.

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

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

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

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

where the last expression 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.

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, $\{x_1, x_2, \ldots\}$, then X is said to be discrete. In this case we can define the *probability mass function* $p : \mathbb{R} \to \mathbb{R}$ of X to be $p(a) = P(X = a)$ which has the following properties:

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

Expected Value

If X is a discrete random variable with probability mass function $p(x)$ then the *expected value* of X, denoted as $E[X]$ is defined by

$$E[X] = \sum_{xp(x)>0} xp(x)$$

Some useful properties, for a random variable X with non-zero outputs $\{x_1, x_2, \ldots\}$:

- If g is any real value function then $E[g(X)] = \sum_{i=1}^{\infty} g(x_i)p(x_i)$
- If a and b are constants, then $E[aX + b] = aE[X] + b$

Variance, Standard Deviation, Covariance, and Correlation

If X is a random variable with *mean* $\mu = E[X]$ then the variance of X, denoted by $\text{Var}(X)$, is defined to be $\text{Var}(X) = E[(X - \mu)^2]$
Some useful formulas that follow from this definition:

- $\text{Var}(X) = E[X^2] - (E[X])^2$
- $\text{Var}(aX + b) = a^2\text{Var}(x)$

The *standard deviation* of X is the square root of the variance: $\sigma = \sqrt{\text{Var}(X)}$
Given random variables, X and Y:

- $\text{Cov}(X, Y) = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y]$ and
  - If X and Y are independent, then $E[XY] = E[X]E[Y]$, so $\text{Cov}(X, Y) = 0$
  - $\text{Cov}(aX, bY) = ab\text{Cov}(X, Y)$
  - $\text{Cov}(aX + bY, cZ) = ac\text{Cov}(X, Z) + bc\text{Cov}(Y, Z)$
- $\text{Cor}(X, Y) = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}$

Units of covariance are the products of the units of X and Y, while correlation is unit-less.

Bernoulli Random Variable

A Bernoulli trial is an experiment that has two outcomes: *success* or *failure*. The sample space is $S = \{\text{success}, \text{failure}\}$ and a *Bernoulli random variable*, X, is defined by $X(\text{success}) = 1$ and $X(\text{failure}) = 0$, along with the probability mass function is given by

$$p(0) = P(X = 0) = 1 - p \qquad\qquad p(0) = P(X = 1) = p$$

where p is the probability of success in a given trial.

Binomial Random Variable

A *binomial random variable* is a random variable X that represents the number of successes that occur in n Bernoulli trials, each with probability of success p.
Properties of Binomial Random Variables:

- $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$
- $E[X] = np$
- $\text{Var}(X) = np(1-p)$

Poisson Random Variable

A *Poisson Random Variable* is a random variable X that takes on one of the values $0, 1, 2, \ldots$, or the number of successes that occur in a given interval, together with a parameter $\lambda > 0$ that describes the expected number of success in that interval.
Properties of Poisson Random Variables:

- $P(X = k) = e^{-\lambda} \dfrac{\lambda^k}{k!}$
- $E[X] = \lambda$
- $\text{Var}(X) = \lambda$

Geometric Random Variable

A *Geometric Random Variable* is a random variable X that takes on one of the values $0, 1, 2, \ldots$, or the number of trials required until a success occurs, when repeatedly performing a Bernoulli trial with probability of success p.
Properties of Geometric Random Variables:

- $P(X = k) = p(1-p)^{k-1}$
- $E[X] = \frac{1}{p}$
- $\text{Var}(X) = \frac{1-p}{p^2}$

Negative Binomial Random Variable

Suppose that independent Bernoulli trials, with proability of success p, are perfomed until r successes occur. A *negative binomial random variable* is a random variable X that takes on one of the values $0, 1, 2, \ldots$, or the number of trials required. Properties of Negative Binomial Random Variables:

- $P(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}$
- $E[X] = \frac{r}{p}$
- $\text{Var}(X) = \frac{r(1-p)}{p^2}$

Hypergeometric Random Variables

A *Hypergeometric Random Variable* is a random variable X that takes on one of the values $0, 1, 2, \ldots$, describing the number of successes in n draws, without replacement, from a finite population of size N, where exactly K members are successes. (In contrast, a binomial random variable can be thought of as describing the probability of k successes with replacement.)
Properties of Hypergeometric Random Variables:

- $P(X = k) = \dfrac{\binom{K}{k}\binom{N-k}{n-k}}{\binom{N}{n}}$
- $E[X] = \frac{nK}{N}$
- $\text{Var}(X) = n\frac{K}{N}\frac{N-K}{N}\frac{N-n}{n-1}$

Note: $\text{Var}(X) \approx n(p)(1-p)$ when N is large relative to n, where $p = K/N$.

Expected Value and Sums of Random Variables

If $X_1, X_2, \ldots, X_n$ are random variables then

$$E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i]$$