

Counting Basics

Sets

- Ways of specifying sets
- Union and intersection
- Universal set and complements
- Empty set and disjoint sets
- Venn diagrams

Counting

Inclusion-exclusion
Multiplication principle
Addition principle
Tree diagrams

Some special cases

Inclusion-Exclusion Principle can be used to count the number of elements in the complement of a set, and the number of elements in a disjoint union.

A special case of the Multiplication principle leads to *Permutations* which count certain *ordered n -tuples* of elements. This then leads to formulas for counting ordered tuples where some of the elements are the same.

Another application of the Multiplication principle leads to *Combinations* which count certain *unordered n -tuples* of elements.

Counting Slogans

$\mathbf{P}(n, m)$ counts the number of ordered subsets of size m selected from a set of n elements.

$\mathbf{C}(n, m)$ counts the number of unordered subsets of size m selected from a set of n elements.

Ordered and unordered partitions

A set with n elements can be partitioned into k subsets of r_1, r_2, \dots, r_k elements (where $r_1 + r_2 + \dots + r_k = n$) and where the subsets are distinguished from one another in the following number of ways:

$$\boxed{\binom{n}{r_1, r_2, \dots, r_k} = \frac{n!}{r_1! r_2! \dots r_k!}}$$

A set of n elements can be partitioned into k **unordered subsets** of r elements each ($kr = n$) in the following number of ways:

$$\boxed{\frac{1}{k!} \binom{n}{r, r, \dots, r} = \frac{n!}{k! \cdot r! \cdot r! \dots r!} = \frac{n!}{k! (r!)^k}}$$

Basic Probability

Suppose we have a set U of all possible ways an experiment could work out, and that S is a subset of U consisting of the outcomes whose occurrence we are interested in. Then the probability that any particular experiment yields an outcome in S is $P(S) = \frac{\mathbf{n}(S)}{\mathbf{n}(U)}$.

Sometimes probabilities are *empirical* and sometimes they can be calculated. Sometimes some probabilities are given to you and you want to calculate others.

A *probability distribution* assigns to each element in a finite set a number between 0 and 1 such that the sum of all the numbers is 1.

Frequency: number of times something occurs; *relative frequency*: proportion of time it occurs

Probability formulae

Most of counting theory applies to enable you to compute probabilities. E.G.

$$P(E) + P(E') = 1 \quad \text{or} \quad \boxed{P(E') = 1 - P(E)}$$

Let E and F be events in a sample space S , then

$$\boxed{P(E \cup F) = P(E) + P(F) - P(E \cap F)}$$

E and F **mutually exclusive** implies that
$$P(E \cup F) = P(E) + P(F)$$

Most of these rules can be visualized with Venn diagrams.

Conditional Probability and Independence

$$P(F|E) = \frac{\mathbf{n}(F \cap E)}{\mathbf{n}(E)} = \frac{P(F \cap E)}{P(E)}$$

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

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

$$P(F|E) = P(F)$$

For independent events, E and F

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

$$P(E \cup F) = P(E) + P(F) - P(E) \cdot P(F)$$

Conditional Probability and Bayes

Tree Diagrams:

Multiply probabilities along a path in a tree and add probabilities for different paths. The probabilities at every node must add up to be 1.

Let E_1 and E_2 be mutually exclusive events ($E_1 \cap E_2 = \emptyset$) whose union is the sample space, i.e. $E_1 \cup E_2 = S$. Let F be an event in S for which $P(F) \neq 0$. Then

Bayes Theorem:

$$P(E_1|F) = \frac{P(E_1 \cap F)}{P(F)} = \frac{P(E_1 \cap F)}{P(E_1 \cap F) + P(E_2 \cap F)} = \frac{P(E_1)P(F|E_1)}{P(E_1)P(F|E_1) + P(E_2)P(F|E_2)}.$$

Conditional Probability

There is a version of Bayes Theorem for more than two mutually exclusive events. Problems along these lines are best dealt with not with the general formula, but with a tree diagram, which always gives the same answer.

Charts, histograms ...

This is all about:

- ▶ Organizing data into meaningful groups and computing frequencies.
- ▶ Drawing histograms and graphs adhering to the equal area principle.
- ▶ Extracting frequencies from histograms.

Mean, median and mode

The **population mean** of m numbers x_1, x_2, \dots, x_m (the data for every member of a population of size m) is denoted by μ and is computed as follows:

$$\mu = \frac{x_1 + x_2 + \cdots + x_m}{m}$$

The **sample mean** of the numbers x_1, x_2, \dots, x_n (data for a sample of size n from the population) is denoted by \bar{x} and is computed similarly:

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

The sample mean is just the population mean of the sample space.

Mean, median and mode

The **population median** is the middle number if you order the data by value. If the number of elements in the data set is even then you need to average two numbers.

The sample median is exactly the same, just for a sample instead of for the whole population.

The **population mode** is number which appears most often. There may be many numbers which are the mode.

The sample mode is exactly the same, just for a sample instead of for the whole population.

Variance and standard deviation

For a set of data $\{x_1, x_2, \dots, x_n\}$ for a population of size n , we define the **population variance**, denoted by σ^2 , to be the average squared distance from the mean, μ :

$$\sigma^2 = \frac{(x_1 - \mu)^2 + (x_2 - \mu)^2 + \dots + (x_n - \mu)^2}{n}$$

For a **sample** $\{x_1, x_2, \dots, x_n\}$ from a larger population, with sample mean \bar{x} , we define the **sample variance**, denoted by s^2 , by

$$s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n - 1}$$

In either case the **standard deviation** is the square root of the variance.

Variance and standard deviation

If a population data set (with n data points) has the values c_1, c_2, \dots, c_m occurring with frequencies f_1, f_2, \dots, f_m (so c_1 occurs f_1 times, etc.), then

$$\sigma^2 = \frac{(c_1 - \mu)^2 f_1 + (c_2 - \mu)^2 f_2 + \cdots + (c_m - \mu)^2 f_m}{n}$$

with a similar formula for s^2 (if the data is a sample).

Random Variables

We perform an experiment, and to each outcome we associate a numerical value (a subject's weight; the time it took until something happened, the number of times we were successful in some number of attempts).

For each possible numerical value that could come up, we can ask “what is the probability that value comes up”?

The resulting table of values and probabilities:

Outcomes	Probability
X	$P(X)$
x_1	p_1
x_2	p_2
\vdots	\vdots
x_n	p_n

is a **Random Variable**

Expected Value of a Random Variable

The **expected value** of a random variable is a measure of the average value the random variable takes, averaged over many repetitions.

If X is a random variable with possible values x_1, x_2, \dots, x_n and corresponding probabilities p_1, p_2, \dots, p_n , then the **expected value of X** , denoted by $\mathbf{E}(X)$, is

$$\mathbf{E}(X) = x_1p_1 + x_2p_2 + \cdots + x_np_n.$$

Outcomes	Probability	Out. \times Prob.
X	$\mathbf{P}(X)$	$X \cdot \mathbf{P}(X)$
x_1	p_1	x_1p_1
x_2	p_2	x_2p_2
\vdots	\vdots	\vdots
x_n	p_n	x_np_n
		Sum = $\mathbf{E}(X)$

Variance of a random variable

If X is a random variable with values x_1, x_2, \dots, x_n , corresponding probabilities p_1, p_2, \dots, p_n , and expected value $\mu = \mathbf{E}(X)$, then

$$\mathbf{Variance} = \sigma^2(X) = p_1(x_1 - \mu)^2 + p_2(x_2 - \mu)^2 + \dots + p_n(x_n - \mu)^2$$

and

$$\mathbf{Standard Deviation} = \sigma(X) = \sqrt{\mathbf{Variance}}.$$

Calculating via a table

$$\text{Variance} = \sigma^2(X) = p_1(x_1 - \mu)^2 + p_2(x_2 - \mu)^2 + \cdots + p_n(x_n - \mu)^2$$

$$\text{Standard Deviation} = \sigma(X) = \sqrt{\text{Variance}}.$$

x_i	p_i	$x_i p_i$	$(x_i - \mu)$	$(x_i - \mu)^2$	$p_i(x_i - \mu)^2$
x_1	p_1	$x_1 p_1$	$(x_1 - \mu)$	$(x_1 - \mu)^2$	$p_1(x_1 - \mu)^2$
x_2	p_2	$x_2 p_2$	$(x_2 - \mu)$	$(x_2 - \mu)^2$	$p_2(x_2 - \mu)^2$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
x_n	p_n	$x_n p_n$	$(x_n - \mu)$	$(x_n - \mu)^2$	$p_n(x_n - \mu)^2$
		Sum = μ			Sum = $\sigma^2(X)$

Bernoulli experiments and binomial distribution

A Bernoulli experiment is some fixed number n of repetitions of independent, identical trials, where in each trial what is measured is either “Success” (which happens with probability p) or “failure” (which happens with probability $q = 1 - p$).

If X is the number of successes then

$$\mathbf{P}(X = k) = \mathbf{C}(n, k)p^k q^{n-k} = \binom{n}{k} p^k q^{n-k}$$

for $k = 0, 1, 2, \dots, n$.

The **expected value** of X is $\mathbf{E}(X) = np$ and the **standard deviation** of X is $\mathbf{\sigma}(X) = \sqrt{npq}$.

X is a **Binomial distribution** with parameters n and p .

Normal distributions

1. All Normal Curves have the same general bell shape.
2. The curve is symmetric with respect to a vertical line that passes through the peak of the curve.
3. The curve is centered at the mean μ which coincides with the median and the mode and is located at the point beneath the peak of the curve.
4. The area under the curve is always 1.
5. The curve is completely determined by the mean μ and the standard deviation σ . For the same mean, μ , a smaller value of σ gives a taller and narrower curve, whereas a larger value of σ gives a flatter curve.

Standard Normal

The standard normal curve has $\mu = 0$ and $\sigma = 1$. For the standard normal curve, you can compute

$$\mathbf{P}(a \leq Z \leq b)$$

either by

$$\mathbf{P}(a \leq Z \leq b) = \mathbf{P}(Z \leq b) - \mathbf{P}(Z \leq a)$$

and consulting a table for these probabilities (table will be provided), or by using the Normalcdf function on your calculator.

For the Normal, there's no difference between $\mathbf{P}(Z \leq b)$ and $\mathbf{P}(Z < b)$.

Standard Normal

For a general normal with mean μ and standard deviation σ , the ***z*-score** of an observation a is

$$z_a = \frac{a - \mu}{\sigma}$$

For a general normal, the probability of seeing a value between a and b is the same as the probability that a standard normal takes a value between z_a and z_b

z-scores allow you to

- ▶ convert non-standard normal distributions to the standard one;
- ▶ to compare two values from different normally distributed data sets;
- ▶ to find percentiles for a normal distribution.

Linear programming

Main points:

Given a collection of inequalities, identify the region of the plane where all the inequalities are satisfied simultaneously. In other words, given a finite set of constraints, find the feasible set (which may be empty).

Find the coordinates of the corners of the feasible set

Given a linear objective function, figure out where it is maximized/minimized on the feasible set — the corners of the feasible set are key.

Game theory

Main points:

Construct payoff matrices for zero-sum and constant sum games.

Find saddle points (if there are any), and figure out if games are strictly determined. Find the values of games, and figure out if they are fair.

Find expected payoffs for mixed strategies.

Use strategy lines to find optimal mixed strategies for players with 2 play options.