List of Definitions, Concepts, and Formulas for Math 125

Definitions and Concepts

Chapter 16

The Law of Averages:

As the number of trials goes up, the difference between the observed number and the expected number goes up, but the difference between the observed percentage and the expected percentage goes down.

These differences are called the chance error: the chance error for the sum of draws (from the 0-1 counting box) and the chance error for the percentage of draws. The chance error for the percentage of draws is just the chance error for the sum expressed as a precentage of the number of draws.

In the case of a fair coin: the chance error for the sum is just the difference between the number of heads and half the number of tosses; the chance error for the percentage is just the difference between the percentage and 50%.

Chance Processes:

A chance process is one that produces a number at random. In this course, only those that produce numbers or results that can be modelled by numbers will be considered.

Box Model: A box model is an analogy between a chance process and drawing from a box with numbers. It will be easy to analyze the sum of the numbers drawn from the box.

Chapter 17

The Expected Value and the Standard Error

The sum of the draws is likely to be around its expected value. give or take its SD or so.

Observed values are rarely more than 2 or 3 SEs away from the expected value.

Using the Normal Curve

For a large number of draws made a random from a box with numbered tickets, the normal curve may be used to find probabilities for the sum of the draws. Just convert the sums of draws to standard units (using the expected value and standard error) and work out areas under the curve, just as in chapter 5. Only here, use the expected value and SE to convert to standard units: not the average and the SD as before.

Chapter 18

Probability Histgograms:

A probability histogram is a new kind of graph. This graph represents chance, not data.

A probability histogram represents chance by area.

The Normal Approximation to Binomial Probabilities:

As long as the number of trials is large, the normmal curve can be used to approximate binomial probabilities. For greater accuracy, find the exact endpoints of the blocks of the histogram by applying $\pm 1/2$. This is called the continuity correction.

The Central Limit Theorem:

When drawing at random with replacement from a box, the probability histogram for the sum will follow the normal curve, even if the contents of the box do not. The histogram must be put into standard units, and the number of draws must be reasonably large.

The expected value pins the center of the probability histogram to the horizontal axis, and the standard error fixes its spread.

Formulas

Chance Processes (those that produce a number as a result of chance)

In many cases the process being studied can be made analagous to the process of drawing numbers from a box. This analogy is called a box model.

Games of Chance

The tickets in the box show the various amounts that can be won or lost on a single play.

The chance of drawing any particular number from the box must equal the chance of winning that amount on a single play. ("Winning" a negative amount is the mathematical equivalent of what most people call losing.)

The number of draws equals the number of plays.

The Sum of Draws

The *sum of draws* from a box is shorthand for the process of drawing tickets at random from a box and then adding up the numbers on the tickets. The formula that applies is

$$\text{observed sum of the draws} = \text{expected sum} + \text{chance error}.$$

The Expected Value for the Sum of the Draws

The expected value for the sum of draws made at random with replacement from a box equals

$$(\text{number of draws}) \times (\text{average of box}).$$

The Chance Error for an Observed Sum of Draws from a Box

The chance error for an observed sum of draws made at random from a box is:

$$\text{observed sum of draws} - \text{expected value for the sum of draws}.$$

The Standard Error for the Sum of the Draws

A sum is likely to be around its expected value, but to be off by a chance error similar in size to the standard error.

When drawing at random with replacement from a box of numbered tickets, the standard error for the sum of the draws is

$$\sqrt{\text{number of draws}} \times (\text{SD of box}).$$

A Shortcut for the SD of a Box with Only Two Kinds of Tickets

When the tickets in the box show only two different numbers, the SD of the box equals

$$\begin{pmatrix}\text{bigger}\\ \text{number}\end{pmatrix} - \begin{pmatrix}\text{smaller}\\ \text{number}\end{pmatrix} \times \sqrt{\begin{matrix}\text{fraction with}\\ \text{bigger number}\end{matrix} \times \begin{matrix}\text{fraction with}\\ \text{smaller number}\end{matrix}}$$

Classifying and Counting

If you have to classify and count the draws, put 0's and 1's on the tickets. Mark 1 on the tickets that count for you, 0 on the others.

Then find the average of the box and the expected value, and the SD (using the shortcut) and the standard error. Now convert any given values of the count to standard units and use the normal curve to approximate any chance being sought.

The Normal Approximation to Binomial Probabilities

If a binomial probability is considered as the sum of repeated draws from a suitable counting box, the normal approximation may be used—provided the number of trials (draws from the box) is suitably large.

The expected value is the product of the number of trials and the average of the counting box. The standard error is the product of the square root of the number of trials and the SD of the counting box (found by the short cut).

Since the sum of the draws is discrete, it is more accurate to correct the endpoints of the intervals by plus or minus one half.

Next convert the endpoints of the given range to standard units using the formula

$$\textbf{standard units} = \frac{\textbf{given value (corrected)} - \textbf{expected value}}{\textbf{standard error}}.$$

The area under the normal curve between the standard units for the corrected endpoints of the given range will be an approximation for the desired chance.