MEASURES OF CENTER

Suppose we have the following test scores. The question now is how do we find the center of this data? By center, we mean an average value.

40, 79, 83, 92, 92

There is one obvious way to find our center and that is by computing the average in the way we normally do, i.e. add up all the scores and divide by the total number. In statistics, we call this average the arithmetic mean or mean.

40, 79, 83, 92, 92

$$mean = \frac{\sum x}{n} = \frac{40 + 79 + 83 + 92 + 92}{5} = 77.2$$

A notation that we should point out is that if we are dealing with a population, then we denote its mean by the Greek letter “mu,” and we denote a sample mean by “x-bar.”

40, 79, 83, 92, 92

$$\text{population mean} = \mu = 77.2$$

$$\text{sample mean} = \bar{x} = 77.2$$

The best thing about the mean is that it is easily incorporated into more advanced statistical procedures.

40, 79, 83, 92, 92

$$\mu = \frac{\sum x}{n} = 77.2$$

The worst thing about the mean is that it is easily affected by extreme values.

40, 79, 83, 92, 92

$$\mu = \frac{\sum x}{n} = 77.2$$

Another way to find the center or average of our data is the median. To find the median, arrange your values in order from lowest to highest, and locate the middle score.

40, 79, 83, 92, 92

median

$$\text{median} = \tilde{x} = 83$$

$$\text{position of the median} = \frac{n+1}{2} = 3$$

If there is no score exactly in the middle, then to find the median we average the two middle scores together.

40, 79, 83, 92

$$median = \frac{79+83}{2} = 81$$

$$\text{median} = \tilde{x} = 81$$

$$\text{position of the median} = \frac{n+1}{2} = 2.5$$

Unlike the mean, the median is not so affected by extreme scores.

40, 79, 83, 92

$$median = \frac{79+83}{2} = 81$$

$$\text{median} = \tilde{x} = 81$$

$$\text{position of the median} = \frac{n+1}{2} = 2.5$$

However, it is harder to incorporate into advanced statistical procedures.

40, 79, 83, 92

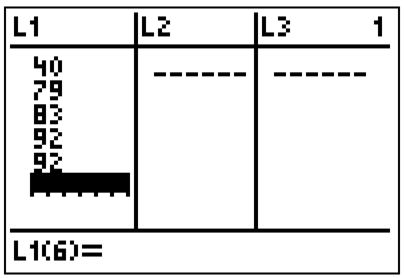
$$median = \frac{79 + 83}{2} = 81$$

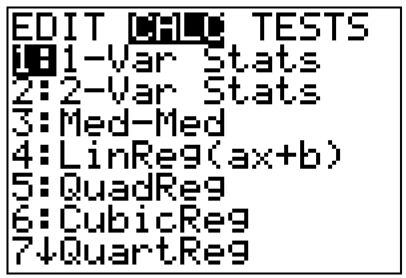
$$\text{median} = \tilde{x} = 81$$

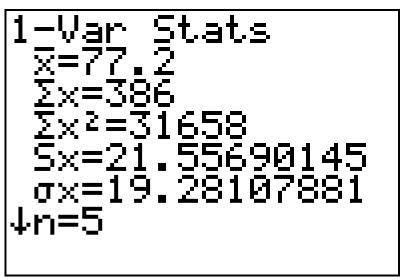
$$\text{position of the median} = \frac{n+1}{2} = 2.5$$

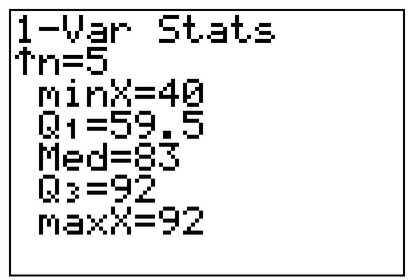
Additionally, both the mean and the median of a list of numbers can easily be found using your calculator.

40, 79, 83, 92, 92









A third type of average is called the mode. This is the data element that occurs most often.

40, 79, 83, 92, 92

$$\text{mode} = 92$$

However, sometimes more than one mode can exist.

40, 40, 79, 83, 92, 92

$$\text{mode} = 40, 92$$

Also, if each data element occurs the same number of times, then no mode exists.

40, 79, 83, 92

no mode

Nonetheless, the mode has one distinct advantage over other measures of central tendency. It can be used with nonnumerical data.

Sample of majors of students:

math	math	math	math
math	math	English	physics
physics	nursing	business	education

The average student majors in math!

Our next task is to find the average value for a probability distribution. As an example, suppose on any given test a certain student was always make either a 70, 80, or 90, and that the probabilities are as shown in the table below.

GRADE	PROBABILITY
70	10%
80	60%
90	30%

Then if this student takes 100 tests, we would expect ten 70s, sixty 80s, and thirty 90s.

GRADE	PROBABILITY
70	10%
80	60%
90	30%

Thus, the student's average grade would be the following:

GRADE	PROBABILITY
70	10%
80	60%
90	30%

$$\mu = \frac{70 \cdot 10 + 80 \cdot 60 + 90 \cdot 30}{100} = 70 \cdot \frac{10}{100} + 80 \cdot \frac{60}{100} + 90 \cdot \frac{30}{100}$$

$$= 70 \cdot P(70) + 80 \cdot P(80) + 90 \cdot P(90) = 82$$

What this illustrates is that we can find the mean of a probability distribution by taking the sum of each value of the random variable times its probability. We also call this the *expected value, E.* This tells us what we expect to happen in the long run.

GRADE	PROBABILITY
70	10%
80	60%
90	30%

$$E = 70 \cdot P(70) + 80 \cdot P(80) + 90 \cdot P(90) = 82$$

For our coin flipping experiment, we have the following expected value.

x = number of heads	$P(x)$
0	1/8
1	3/8
2	3/8
3	1/8

$$E = 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} = \frac{12}{8} = \frac{3}{2} = 1.5$$

This makes good sense. If we flip a fair coin three times and repeat the experiment over and over, then we expect in the long run to come up with heads half the time. In other words, we average 1.5 heads for each run of the experiement.

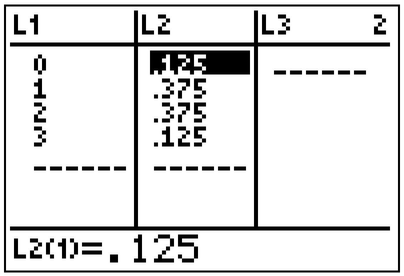
x = number of heads	$P(x)$
0	1/8
1	3/8
2	3/8
3	1/8

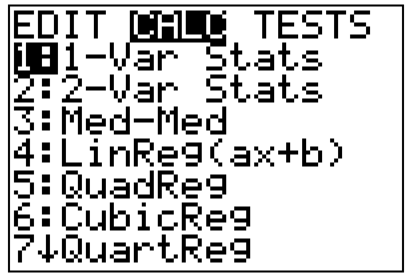
$$E = 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} = \frac{12}{8} = \frac{3}{2} = 1.5$$

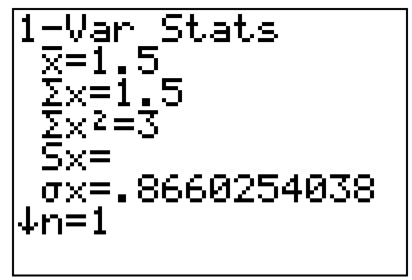
In general, if we have a probability distribution, then the average or expected value of the distribution is given by the formula below.

$$E = \sum \left[x \cdot P(x) \right]$$

This can also be found using tools on your calculator.







$$\mu = \sum \left[x \cdot P(x) \right] = 1.5$$

Suppose you can bet \$5 either in roulette or a dice game, and the probability distributions for each game are as shown below. Which is the better game to play?

ROULETTE

Event	*x*	*P(x)*
lose	-5	37/38
win	175	1/38

DICE

Event	*x*	*P(x)*
lose	-5	251/495
win	5	244/495

Suppose you can bet $5 either in roulette or a dice game, and the probability distributions for each game are as shown below. Which is the better game to play?

ROULETTE

Event	x	$P(x)$
lose	-5	37/38
win	175	1/38

DICE

Event	x	$P(x)$
lose	-5	251/495
win	5	244/495

$$E = (-5)(37/38) + (175)(1/38)$$
$$= -\$0.26$$

$$E = (-5)(251/495) + (5)(244/495)$$
$$= -\$0.08$$

Suppose you can bet $5 either in roulette or a dice game, and the probability distributions for each game are as shown below. Which is the better game to play?

ROULETTE

Event	*x*	*P(x)*
lose	-5	37/38
win	175	1/38

DICE

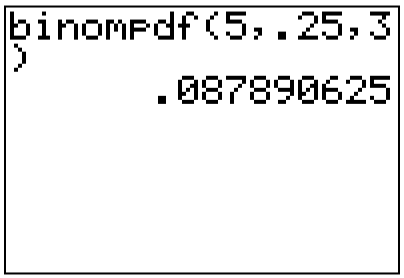
Event	*x*	*P(x)*
lose	-5	251/495
win	5	244/495

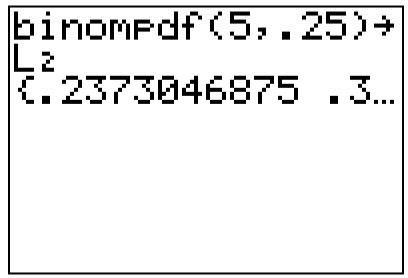
$$E = (-5)(37/38) + (175)(1/38)$$
$$= -\$0.26$$

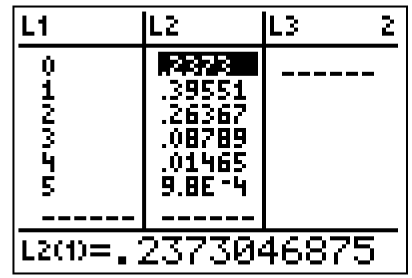
$$E = (-5)(251/495) + (5)(244/495)$$
$$= -\$0.08$$

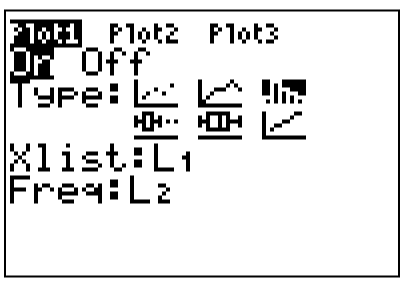
You will lose less in the long run playing dice.

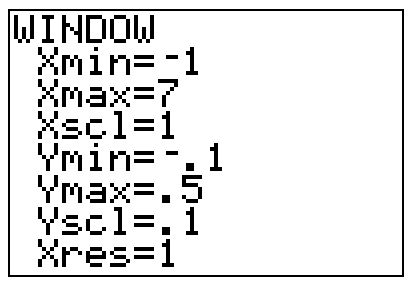
There’s an easy way to figure the *mean* or *expected value* of a binomial distribution. Recall our previous example of a 5-question test with four multiple-choices for each question.

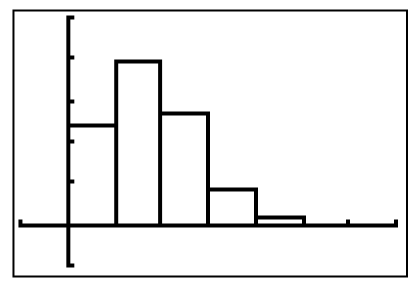












If we took several 5 problem tests of this sort and guessed on each question with probability of success equal to ¼, then in the long run we would expect to get about ¼ of the questions correct on each test. In other words, the *mean* of a binomial distributions is:

$$\mu = np$$

$$(5)(.25) = 1.25$$