RANDOM VARIABLES

A *random variable* is a variable whose numerical values are determined by chance.

A *random variable* is a variable whose values are determined by chance.

A *random variable* is *discrete* if its values do not exist along a continuum. In other words, we think of there being separation between the possible values. Discrete random variables are often the result of a count.

A *random variable* is a variable whose values are determined by chance.

A *random variable* is *discrete* if its values do not exist along a continuum. In other words, we think of there being separation between the possible values. Discrete random variables are often the result of a count.

A *random variable* is *continuous* if its values do exist along a continuum. Thus, between any two values of the variable, other possible values exist.

If you flip a coin three times and let X = *number of heads,* then X is a discrete random variable with possible values of 0, 1, 2, & 3.

If you flip a coin three times and let X = *number of heads,* then X is a discrete random variable with possible values of 0, 1, 2, & 3.

If you let X = *number of eggs a hen lays,* then *that* is also a discrete random variable.

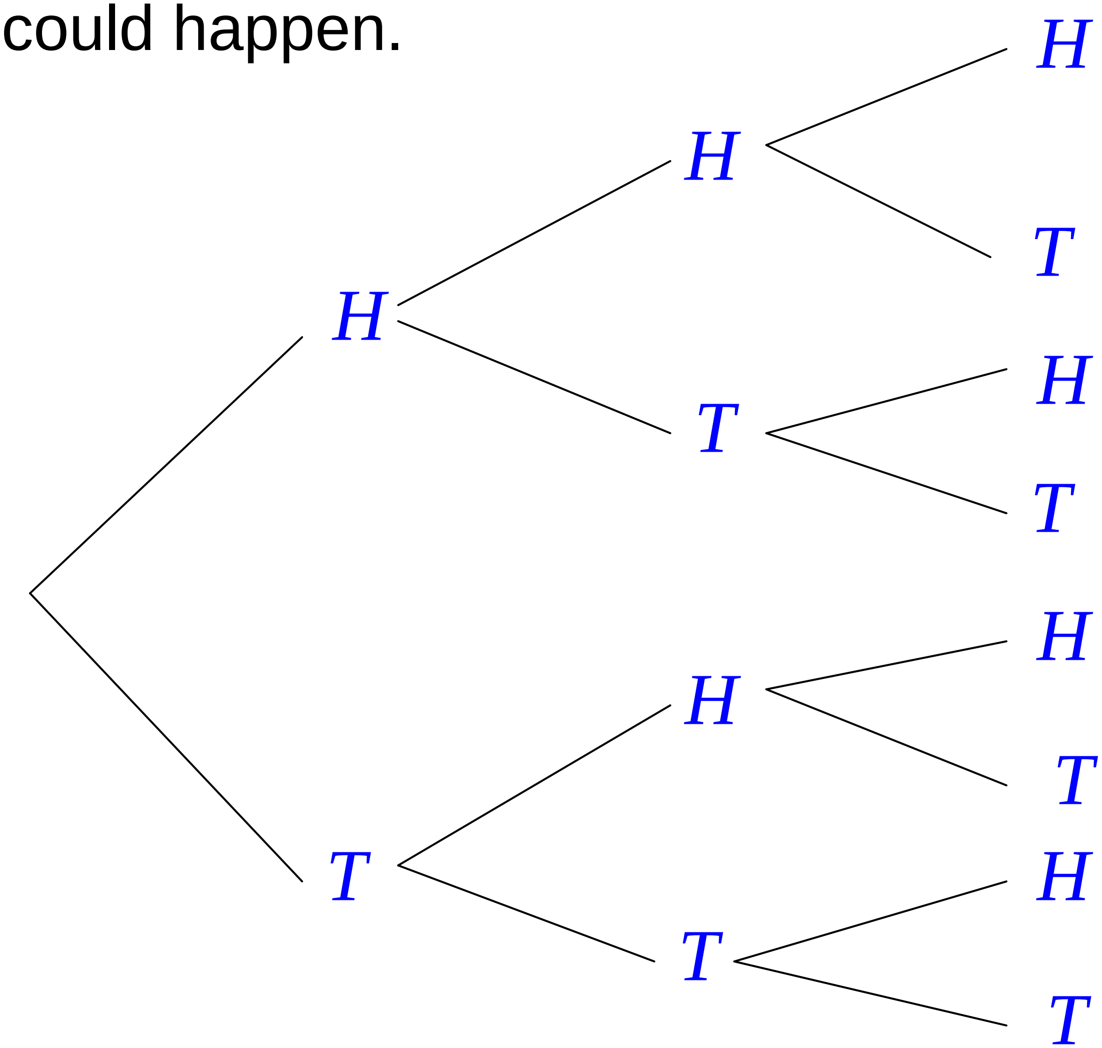
If you flip a coin three times and let X = *number of heads,* then X is a discrete random variable with possible values of 0, 1, 2, & 3.

If you let X = *number of eggs a hen lays,* then *that* is also a discrete random variable.

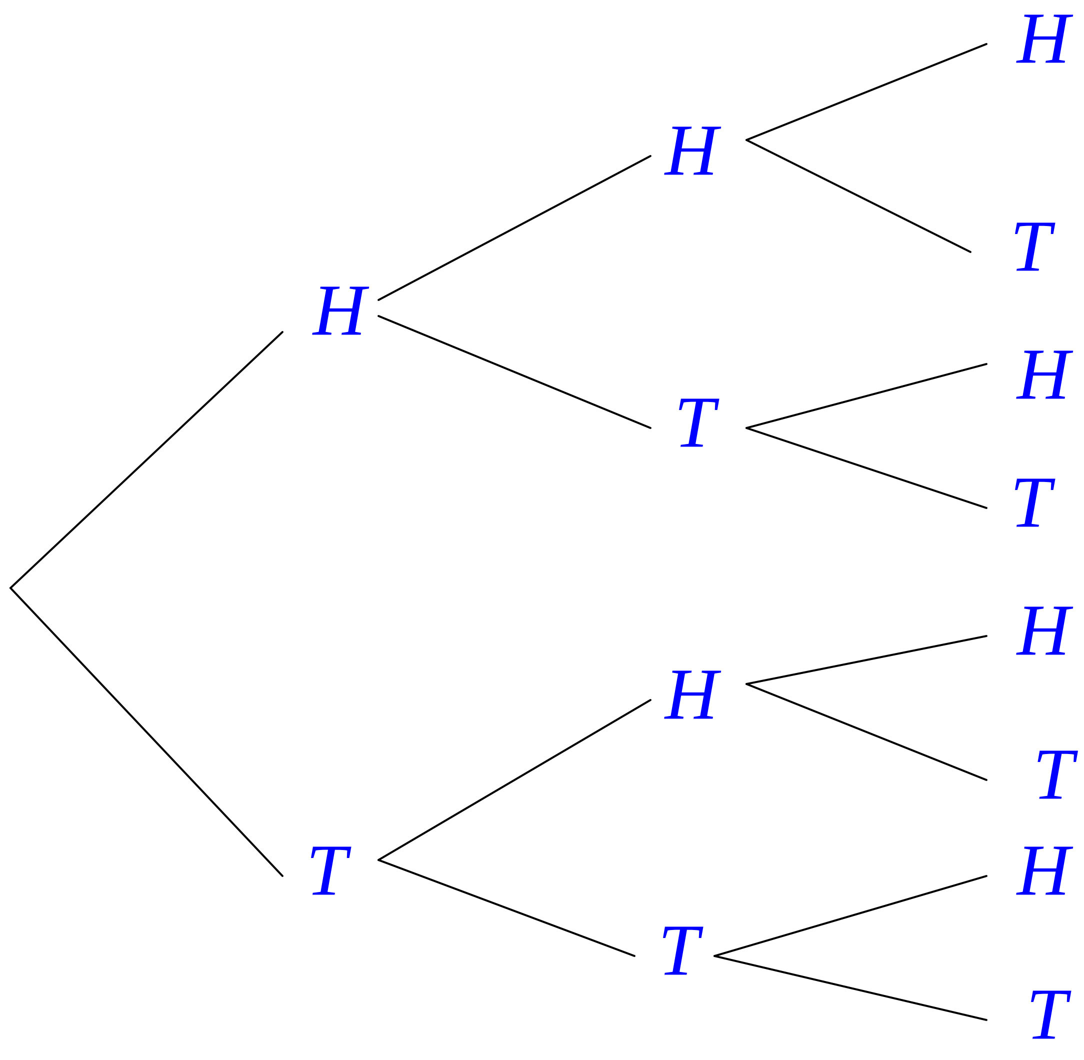
If you let X = *amount of milk a cow produces,* then *that* is a continuous random variable.

Let's do the experiment where we flip a fair coin three times and let X = *number of heads,* and let's consider what could happen.

Let's do the experiment where we flip a fair coin three times and let X = *number of heads,* and let's consider what could happen.



There are eight possible outcomes.



HHH

HHT

HTH

HTT

THH

THT

TTH

TTT

We can summarize the results in the following table.

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

HHH

HHT

HTH

HTT

THH

THT

TTH

TTT

This type of table is called a *probability distribution*.

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

HHH

HHT

HTH

HTT

THH

THT

TTH

TTT

Notice, also, the following:

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

$$1.\quad 0 \le P(x) \le 1$$

$$2.\quad \sum P(x) = 1$$

HHH

HHT

HTH

HTT

THH

THT

TTH

TTT

We can also create a histogram on our calculator for this Probability distribution by completing the following screens:

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

HHH

HHT

HTH

HTT

THH

THT

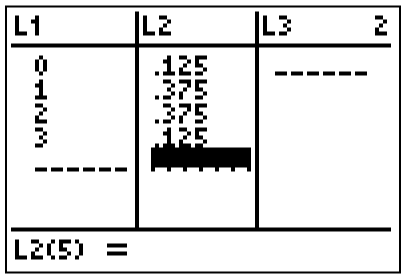
TTH

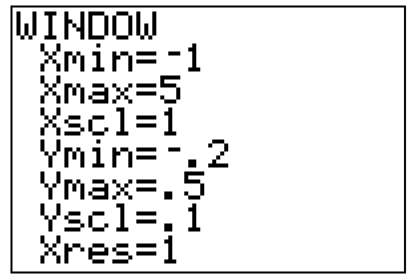
TTT

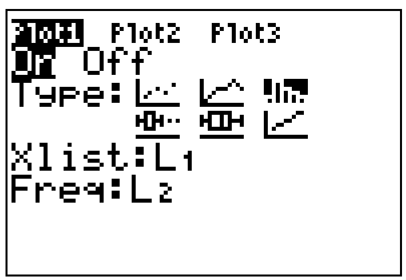
1. $0 \le P(x) \le 1$

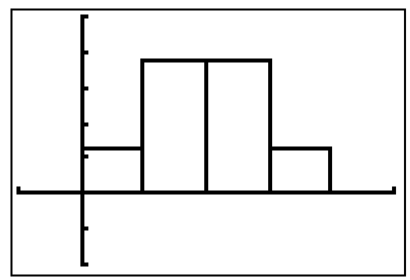
2. $\sum P(x) = 1$

We can also create a histogram on our calculator for this Probability distribution by completing the following screens:









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.

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

At this point, we should now see if we can determine how to calculate variance and standard deviation for a probability distribution. On the one hand, we know that variance is defined by the formula below.

$$\sigma^2 = \frac{\sum (x-\mu)^2}{n}$$

This formula basically computes the average squared deviation from the mean. But on the other hand, if the formula in blue below gives the average value for the distribution, then the modified one in red should give the average squared deviation, i.e. the variance.

$$\sigma^2 = \frac{\sum (x-\mu)^2}{n}$$

$$E = \sum \left[x \cdot P(x) \right] \qquad \sigma^2 = \sum \left[(x-\mu)^2 \cdot P(x) \right]$$

Now we just need to do a little algebra to get this in a better form for computations.

$$\sigma^2 = \sum\left[(x-\mu)^2 \cdot P(x)\right] = \sum\left[(x^2 - 2x\mu + u^2)\cdot P(x)\right]$$

$$= \sum\left[x^2 P(x) - 2x\mu P(x) + \mu^2 P(x)\right]$$

$$= \sum\left[x^2 P(x)\right] - \sum[2x\mu P(x)] + \sum\left[\mu^2 P(x)\right]$$

$$= \sum\left[x^2 P(x)\right] - 2\mu\sum[x \cdot P(x)] + \mu^2\sum[P(x)]$$

$$= \sum\left[x^2 P(x)\right] - 2\mu\cdot\mu + \mu^2 = \sum\left[x^2 P(x)\right] - 2\mu^2 + \mu^2$$

$$= \sum\left[x^2 \cdot P(x)\right] - \mu^2 = \sum\left[x^2 \cdot P(x)\right] - \left(\sum[x \cdot P(x)]\right)^2$$

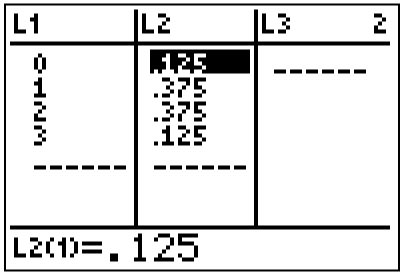
Now lets give it a try using our probability distribution for the coin flipping experiment.

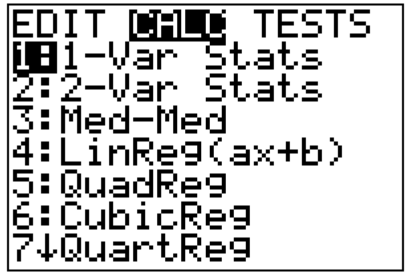
x = number of heads	*P(x)*	x*P(x)	x^2	x^2*P(x)
0	0.125	0	0	0
1	0.375	0.375	1	0.375
2	0.375	0.750	4	1.5
3	0.125	0.375	9	1.125
		1.5		3

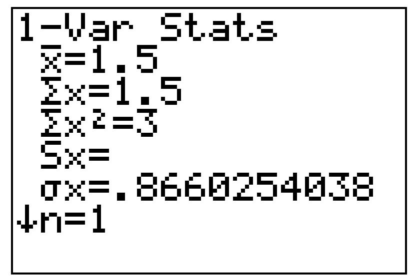
$$\mu = \sum \left[x \cdot P(x) \right] = 1.5$$

$$\sigma = \sqrt{\sum \left[x^2 \cdot P(x) \right] - \mu^2} = \sqrt{.75} \approx .8660254038$$

Of course, there is one other way to do this which now should seem very convenient.







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

$$\sigma = \sqrt{\sum \left[x^2 \cdot P(x) \right] - \mu^2} = \sqrt{.75} \approx .8660254038$$

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.

SOME CRITERIA FOR UNUSUAL RESULTS:

1. The values are more than two standard deviations from the mean.

2. *P(x or more)* ≤ 0.05.

3. *P(x or less)* ≤ 0.05.

Five peapods are produced from parent peas containing genes for both green and yellow pods. The gene for green is dominant, and the probabilities for getting from 0 to 5 green pods as offspring are shown below. Would 1 or fewer green pods be unusual?

x	$P(x)$
0	0.001
1	0.015
2	0.088
3	0.264
4	0.396
5	0.237

Five peapods are produced from parent peas containing genes for both green and yellow pods. The gene for green is dominant, and the probabilities for getting from 0 to 5 green pods as offspring are shown below. Would 1 or fewer green pods be unusual?

x	*P(x)*
0	0.001
1	0.015
2	0.088
3	0.264
4	0.396
5	0.237

```
1-Var Stats
 x̄=3.748251748
 Σx=3.752
 Σx²=15.004
 Sx=
 σx=.9693398892
↓n=1.001
```

Five peapods are produced from parent peas containing genes for both green and yellow pods. The gene for green is dominant, and the probabilities for getting from 0 to 5 green pods as offspring are shown below. Would 1 or fewer green pods be unusual?

x	*P(x)*
0	0.001
1	0.015
2	0.088
3	0.264
4	0.396
5	0.237

```
1-Var Stats
 x̄=3.748251748
 Σx=3.752
 Σx²=15.004
 Sx=
 σx=.9693398892
↓n=1.001
```

$$z=\frac{x-\mu}{\sigma}=\frac{1-3.748251748}{.9693398892}\approx -2.8$$

Five peapods are produced from parent peas containing genes for both green and yellow pods. The gene for green is dominant, and the probabilities for getting from 0 to 5 green pods as offspring are shown below. Would 1 or fewer green pods be unusual?

x	*P(x)*
0	0.001
1	0.015
2	0.088
3	0.264
4	0.396
5	0.237

```
1-Var Stats
 x̄=3.748251748
 Σx=3.752
 Σx²=15.004
 Sx=
 σx=.9693398892
↓n=1.001
```

$$z = \frac{x-\mu}{\sigma} = \frac{1-3.748251748}{.9693398892} \approx -2.8$$

$$P(x \le 1) = 0.001 + 0.015 = 0.016$$

Five peapods are produced from parent peas containing genes for both green and yellow pods. The gene for green is dominant, and the probabilities for getting from 0 to 5 green pods as offspring are shown below. Would 1 or fewer green pods be unusual?

x	*P(x)*
0	0.001
1	0.015
2	0.088
3	0.264
4	0.396
5	0.237

```
1-Var Stats
 x̄=3.748251748
 Σx=3.752
 Σx²=15.004
 Sx=
 σx=.9693398892
↓n=1.001
```

$$z=\frac{x-\mu}{\sigma}=\frac{1-3.748251748}{.9693398892}\approx -2.8$$

$$P(x\leq 1)=0.001+0.015=0.016$$

The result is unusual!