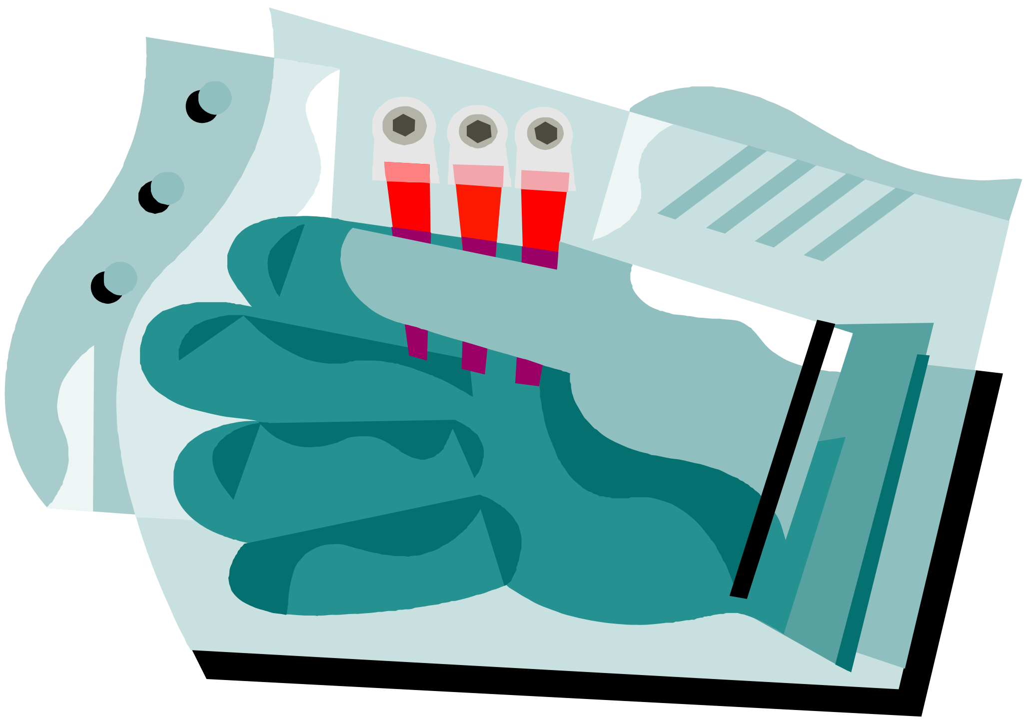
SAMPLING DISTRIBUTIONS



The sampling distribution of a statistic is the distribution of all values of the statistic when all possible samples of the same size *n* are taken from the same population.

The sampling distribution of the mean is the distribution of sample means when all possible samples of the same size *n* are taken from the same population.

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

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

The sampling distribution of the variance is the distribution of sample variances when all possible samples of the same size *n* are taken from the same population.

$$s^2 = \text{sample variance}$$

$$\sigma^2 = \text{population variance}$$

The sampling distribution of the proportion is the distribution of sample proportions when all possible samples of the same size *n* are taken from the same population.

$$\hat{p} = \text{ sample proportion}$$

$$p = \text{ population proportion}$$

As our sample size increases, the values obtained for our sample mean, sample variance, and sample proportion tend to get closer to the actual population mean, variance, and proportion.

As our sample size increases, the values obtained for our sample mean, sample variance, and sample proportion tend to get closer to the actual population mean, variance, and proportion.

In this case, we say that the sample mean, sample variance, and sample proportion *target* the population mean, variance, and proportion.

As our sample size increases, the values obtained for our sample mean, sample variance, and sample proportion tend to get closer to the actual population mean, variance, and proportion.

In this case, we say that the sample mean, sample variance, and sample proportion *target* the population mean, variance, and proportion.

We also say that each of these is an *unbiased estimator* of the corresponding population parameter.

The *standard deviation* is *not* an *unbiased estimator*, but the error is small.

The *standard deviation* is *not* an *unbiased estimator*, but the error is small.

The *median* and the *range* are also *not unbiased estimators*. They do not target the corresponding population parameters as the sample size increases.

Furthermore, as our sample size increases, both the sampling distribution of the sample means and the sampling distribution of the sample proportions tend toward a normal distribution.

Furthermore, as our sample size increases, both the sampling distribution of the sample means and the sampling distribution of the sample proportions tend toward a normal distribution.

On the other hand, the sampling distribution of the sample variances tends toward a distribution that is skewed to the right as the sample size increases.

The Central Limit Theorem: Let X be a random variable with mean μ and standard deviation σ, and consider the sampling distribution of the sample means for all simple random samples of size *n*. We will denote the mean of this distribution by $\mu_{\overline{x}}$ and the standard deviation by $\sigma_{\overline{x}}$. Then,

1. $\mu_{\overline{x}} = \mu$
2. $\sigma_{\overline{x}} = \dfrac{\sigma}{\sqrt{n}}$ (the standard error of the mean)
3. The sampling distribution of the sample means approaches a normal distribution as the sample size increases.

Additionally, we have following two results:

1. If the parent population is normally distributed, then the sampling distribution of the sample means will always be normally distributed, regardless of the sample size *n*.

2. If the parent population is *<u>not</u>* normally distributed, then we do not expect the sampling distribution of the sample means to reasonably approximate a normal distribution unless $n > 30$.

Suppose that weights of men are normally distributed with mean 172 lb and standard deviation 29 lb.

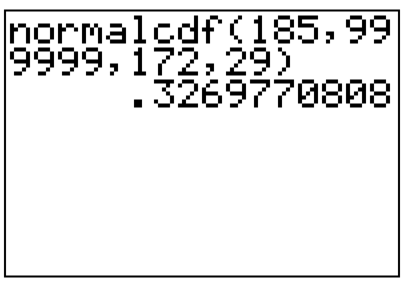
Suppose that weights of men are normally distributed with mean 172 lb and standard deviation 29 lb.

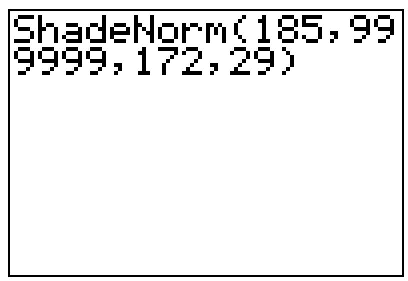
What is the probability that an *individual* man has a weight greater than 185 lb?

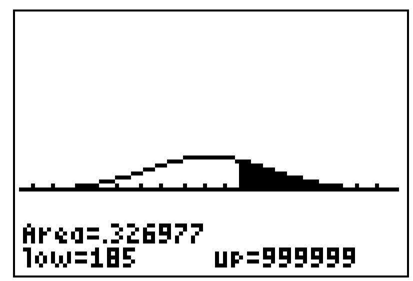
Suppose that weights of men are normally distributed with mean 172 lb and standard deviation 29 lb.

What is the probability that an *individual* man has a weight greater than 185 lb?

$$z = \frac{x - \mu}{\sigma} = \frac{185 - 172}{29} \approx 0.448$$







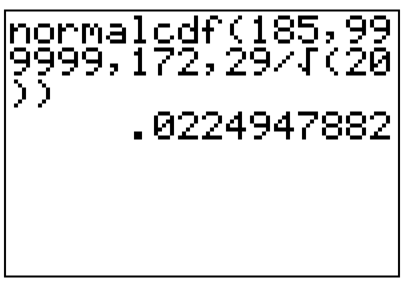
Suppose that weights of men are normally distributed with mean 172 lb and standard deviation 29 lb.

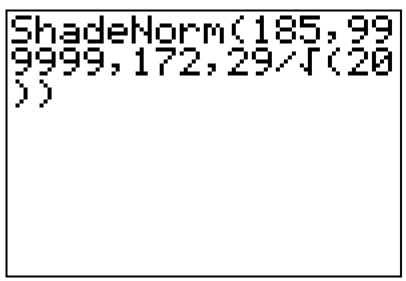
What is the probability that *20 randomly selected men* have a mean weight greater than 185 lb?

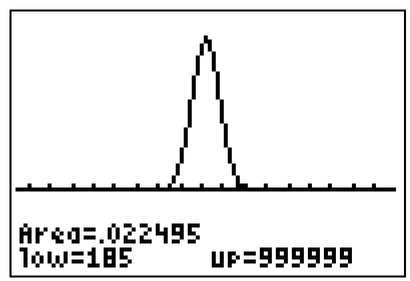
Suppose that weights of men are normally distributed with mean 172 lb and standard deviation 29 lb.

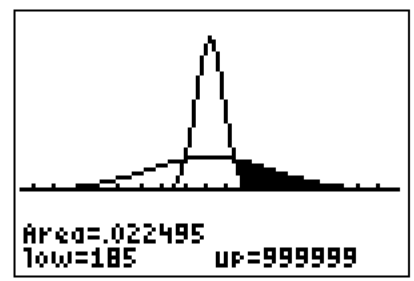
What is the probability that *20 randomly selected men* have a mean weight greater than 185 lb?

$$z = \frac{\overline{x} - \mu_{\overline{x}}}{\sigma_{\overline{x}}} = \frac{\overline{x} - \mu}{\sigma / \sqrt{n}} = \frac{185 - 172}{29 / \sqrt{20}} \approx 2.005$$









A can of Coke should be filled with a mean of 12 oz and a standard deviation of 0.11 oz.

A can of Coke should be filled with a mean of 12 oz and a standard deviation of 0.11 oz.

If a sample of size 36 has a mean fill of 12.04 oz, would this lead you to reject the assumption that the mean fill is 12 oz?

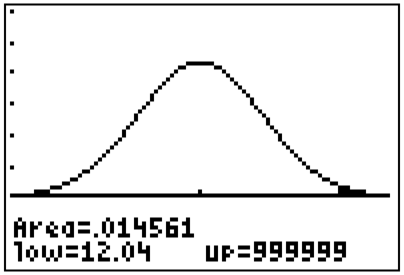
A can of Coke should be filled with a mean of 12 oz and a standard deviation of 0.11 oz.

If a sample of size 36 has a mean fill of 12.04 oz, would this lead you to reject the assumption that the mean fill is 12 oz?

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{12.04 - 12}{0.11 / \sqrt{36}} \approx 2.18$$

```
normalcdf(12.04,
999999,12,.11/6)
          .0145614254
```

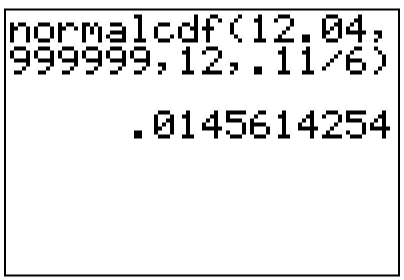
```
ShadeNorm(12.04,
999999,12,.11/6)
```

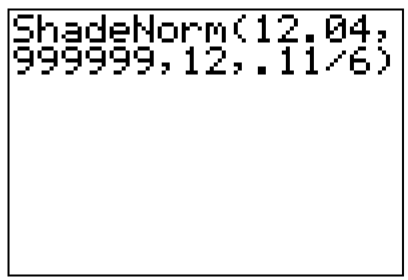


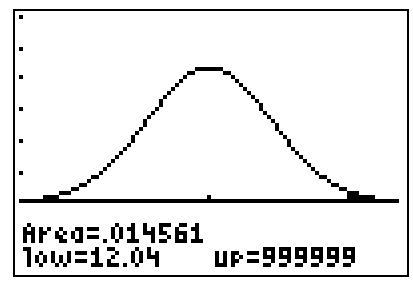
A can of Coke should be filled with a mean of 12 oz and a standard deviation of 0.11 oz.

If a sample of size 36 has a mean fill of 12.04 oz, would this lead you to reject the assumption that the mean fill is 12 oz?

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{12.04 - 12}{0.11 / \sqrt{36}} \approx 2.18$$







Does this difference have practical significance to you?