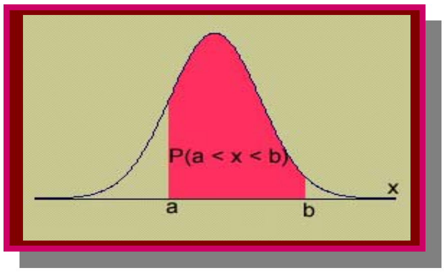
Normal Distribution and Central Limit Theorem

Normal Distribution: (Bell shape distribution)

To find P(a < x < b), we need to find the area under the appropriate normal curve.



The area under the curve is equal to **1.**

$p(a \le x \le b)$ = **area under the curve** between a and b.

There is no probability attached to any single value of x.

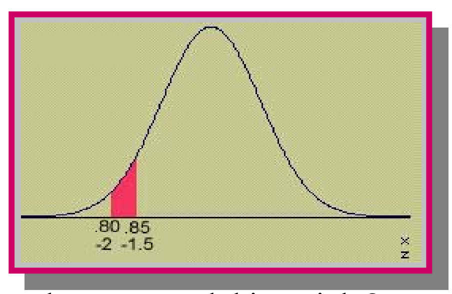
That is, **P(x = a) = 0.**

To simplify the tabulation of these areas, we **standardize** each value of x by expressing it as a ***z*-score**, the number of standard deviations σ it lies from the mean μ, $z = \dfrac{x-\mu}{\sigma}$

Example:

The weights of packages of ground beef are normally distributed with mean 1 pound and standard deviation .10. What is the probability that a randomly selected package weighs between 0.80 and 0.85 pounds?

$$P(0.80 < x < 0.85)$$
$$= P(\frac{0.80-1}{.10} < z < \frac{0.85-1}{.10})$$
$$= P(-2 < z < -1{,}5)$$
$$= P(z < -1.5) - P(z < -2)$$
$$= 0.0668 - 0.0228 = 0.044$$



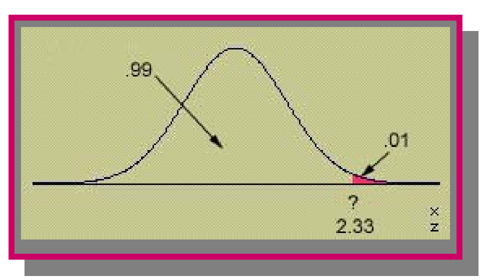
What is the weight of a package such that only 1% of all packages exceed this weight?

$$P(x > ?) = 0.01$$
$$P(z > \frac{?-1}{.1}) = 0.01$$

From Normal Table, we get

$$\frac{?-1}{.1} = 2.33$$

So, $? = 2.33 \times .1 + 1 = 1.233$



Central Limit Theorem:

If random samples of n observations are drawn from a nonnormal population with finite μ and standard deviation σ , then, when n is large, the sampling distribution of **the sample mean** $\bar{x}$ is approximately **normally distributed**, with mean μ and standard $\frac{\sigma}{\sqrt{n}}$

The **Central Limit Theorem** also implies that the **sum** of n measurements is approximately normal with mean $n\mu$ and standard deviation $\sigma\sqrt{n}$

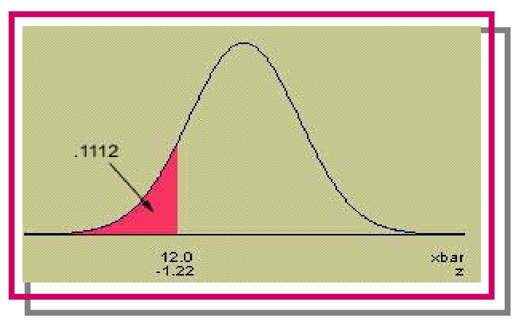
Many statistics that are used for statistical inference are **sums** or **averages** of sample measurements. When n is large ($n \geq 30$) , these statistics will have approximately **normal** distributions.

If the sampling distribution of $\bar{x}$ is normal or approximately normal, *standardize or rescale* the interval of interest in terms of $z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$. Look at normal table and find the probability of the sample mean.

<u>Example:</u>

A soda filling machine is supposed to fill cans of soda with 12 fluid ounces. Suppose that the fills are actually normally distributed with a mean of 12.1 oz and a standard deviation of .2 oz. What is the probability that the average fill for a 6-pack of soda is less than 12 oz?

$$P(\bar{x} < 12)$$
$$= P(\frac{\bar{x} - \mu}{\sigma / \sqrt{n}} < \frac{12 - 12.1}{2 / \sqrt{6}})$$
$$= P(z < -1.22)$$
$$= .1112$$



<u>Example:</u>

The mayor of a small city claims that the average income in his city is $35,000 with a standard deviation of $5000. We take a sample of 64 families, and find that their average income is $30,000. Is his claim correct?

The claim is about population information: $\mu = 35000, \quad \sigma = 5000$.

The sample information are $n = 64, \quad \bar{x} = 30000$.

Can we use sample information to support the population information?

Let us standardize the sample mean $\bar{x} = 30000$: $z = \frac{30000 - 35000}{5000 / \sqrt{64}} = -8$

That is, the sample mean lies on –8 standard deviations below the population mean.

From the Empirical Rule, values more than three standard deviations away from the mean are considered **extremely unlikely.** Therefore, the claim is not correct.