• **Properties** of the normal distribution:
  - continuous bell-shaped probability density curve
  - mean = median = mode
  - unimodal
  - symmetric about the mean
  - the curve gets closer and closer to the $x$-axis as $x \to \infty$ and as $x \to -\infty$
  - the total area under the curve and above the $x$-axis is 1
  - the empirical rule is based on the normal distribution

• The **probability density function** $f$ for a normal random variable is given by
  $$f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma}\right)^2} \text{ for } -\infty < x < \infty,$$
  where $\mu$ is the mean and $\sigma$ is the standard deviation.

• **Notation**: $X \sim N(\mu, \sigma)$ means random variable $X$ has a normal distribution with mean $\mu$ and standard deviation $\sigma$.

• If a normal random variable has mean 0 and standard deviation 1, then it is called a **standard normal** random variable.

• If $X \sim N(\mu, \sigma)$, then $Z = \frac{X-\mu}{\sigma} \sim N(0, 1)$.

• Plot of a **standard normal probability density curve**
To plot a standard normal curve using TI-83:

\[ Y = , \quad \text{2nd Vars 1} , \quad Y_1 = \text{normalpdf}(X, 0, 1) \]

- the empirical rule is based on probabilities associated with a standard normal rv Z

\[
P(-1 < Z < 1) = .6826894809 \\
P(-2 < Z < 2) = .954499876 \\
P(-3 < Z < 3) = .9973000656
\]

- Area under the curve gives probabilities. \( P(a \leq X \leq b) \) is shaded in the plot below.

Find probabilities using TI-83 \text{normalcdf} command.

**Example.** Suppose \( X \sim N(\mu = 7.25, \sigma = 1.47) \).

(i) Find \( P(5 \leq X < 8) \).

\[
\text{2nd Vars (DISTR) 2 normalcdf(5, 8, 7.25, 1.47) ENTER} \quad .6321131201
\]

Therefore, \( P(5 \leq X < 8) = .632 \).

(ii) Find \( P(X > 8) \).

\[
\text{2nd Vars (DISTR) 2 normalcdf(8, 10^{99}, 7.25, 1.47) ENTER} \quad .30499542345
\]

Note that \( 10^{99} \) is playing the part of infinity.

Therefore \( P(X > 8) = .305 \).
(iii) Find \( P(X < 6) \).

\[
\text{normalcdf}( -10^{99}, 6, 7.25, 1.47) \quad \text{ENTER} \quad 0.1975679529
\]

Note that \(-10^{99}\) is playing the part of negative infinity.

Therefore \( P(X < 6) = 0.198 \)

**Finding Percentiles of a Normal Distribution Using TI-83**

**Example.** Find the 95th percentile of a standard normal distribution.

\[
\text{invNorm}(0.95, 0, 1) \quad \text{ENTER} \quad 1.644853626
\]

**Example.** If \( X \sim N(\mu = 7.25, \sigma = 1.47) \), find the 95th percentile.

\[
\text{invNorm}(0.95, 7.25, 1.47) \quad \text{ENTER} \quad 9.66793483
\]

**Assessing normality from the data in a random sample**

- Is the histogram more-or-less bell-shaped?
- Is the stem plot more-or-less bell-shaped?
- Is \( \text{IQR}/s \approx 1.3? \)
- Does the empirical rule hold for the data?
- Does the data distribution lack significant skewness?
  
  Use Pearson index for skewness: \( PI = \frac{3(X - \text{median})}{s} \).
  
  If \( PI \leq -1 \) or \( PI \geq 1 \), then we can conclude that data are significantly skewed.

- Is the normal probability plot (or Q-Q plot) very linear?
Distribution of Sample Means

Prior to observing the values in a random sample of size $n$ from a population with mean $\mu$ and standard deviation $\sigma$, the sample mean $\bar{X}$ is itself a random variable. Thus it makes sense to talk about the mean of $\bar{X}$ and the standard deviation of $\bar{X}$. In fact,

$$\mu_{\bar{X}} = \text{population mean} = \mu$$

$$\sigma_{\bar{X}} = \text{scaled down version of population standard deviation} = \frac{\sigma}{\sqrt{n}}$$

or

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{n-1}}$$

(if sampling without replacement from a finite population of size $N$)

$$\left(\sqrt{\frac{N-n}{n-1}}\right) \text{ is called the finite population correction factor}$$

This gives us a handle on $\bar{X}$. However, we can get an even better grip on $\bar{X}$ if the sample size $n$ is sufficiently large. The following theorem says, in essence, that $\bar{X}$ has an approximately normal distribution when $n$ is large, say $n \geq 30$.

**Central Limit Theorem.** Let $\bar{X}$ be the sample mean based on a random sample of size $n$ from a population with mean $\mu$ and standard deviation $\sigma$. Then the distribution of $Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$ approaches a standard normal distribution as $n \to \infty$.

**NOTE.** No restrictions are placed on the distribution of the population sampled, only that it has a mean and a standard deviation.

- **Approximating binomial probabilities with normal probabilities**

  Rule of thumb: if $n \geq 9$(odds for success) and $n \geq 9$(odds for failure), then $P(r \leq X \leq s) \approx P(r - .5 \leq Y \leq s + .5)$,

  where $X \sim \text{binomial}(n, p)$ and $Y \sim N(\mu = np, \sigma = \sqrt{np(1-p)})$

**Lagniappe** (a little extra stuff)

- **Chebyshev's Inequality.** For any random variable $X$ with mean $\mu$ and standard deviation $\sigma$, $P(|X - \mu| < k\sigma) \geq 1 - k^{-2}$ for any $k > 1$.

- **(Weak) Law of Large Numbers.** Let $\bar{X}$ be the sample mean of a random sample from a population with mean $\mu$ and standard deviation $\sigma$. Then, for any $\epsilon > 0$, $\lim_{n \to \infty} P(|\bar{X} - \mu| < \epsilon) = 1$.

  In other words, for large samples, it is highly likely that the sample mean will close to the population mean. Applied to identically repeated trials, the proportion of "successes" will be close to the probability of "success" with high likelihood.