Confidence Interval for the Difference of Two Means

- A confidence interval for two population means is used when you have two independent random samples and you wish to make a comparison of the difference $(\mu_1 - \mu_2)$.

- The assumptions that need to be satisfied are:
  1. Both samples must be independent SRSs from the populations of interest.
  2. Both sets of data must come from normally distributed populations. If this is not the case or if we are unsure whether the population is normally distributed, the sampling distributions of $\bar{x}_1$ and $\bar{x}_2$ must be normally distributed. $n_1 \geq 30$, $n_2 \geq 30$

- When the population standard deviations are known, we use the formula

$$\left(\bar{x}_1 - \bar{x}_2\right) \pm z^* \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

- When the population standard deviations are unknown, we will need to find the sample standard deviations, s_1 and s_2, and use the formula

$$\left(\bar{x}_1 - \bar{x}_2\right) \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

where t^* is the t-critical value based on the smaller of $n_1 - 1$ or $n_2 - 1$ degrees of freedom.

Example:

1. The height (in inches) of men at UH is assumed to have a normal distribution with a standard deviation of 3.6 inches. The height (in inches) of women at UH is also assumed to have a normal distribution with a standard deviation of 2.9 inches. A random sample of 49 men and 38 women yielded respective means of 68.3 inches and 64.6 inches. Find the 90% confidence interval for the difference in the heights of men at UH and women at UH.

men = sample 1 $\sigma_1 = 3.6$, $n_1 = 49$, $\bar{x}_1 = 68.3$

women = sample 2 $\sigma_2 = 2.9$, $n_2 = 38$, $\bar{x}_2 = 64.6$

critical value: $z^* = \text{invNorm}(.95) = 1.645$

$$(68.3 - 64.6) \pm 1.645\sqrt{\frac{3.6^2}{49} + \frac{2.9^2}{38}}$$

$$3.7 \pm 1.15$$

$$[2.55, 4.85]$$

Example:

2. A researcher wants to see if birds that build larger nests lay larger eggs. He selects two random samples of nests: one of small nests and the other of large nests. He weighs one egg from each nest. The data are summarized below:

	Small nests 1	Large nests 2
Sample size	$n_1 =$ 60	$n_2 =$ 159
Sample mean (g)	$\bar{x}_1 =$ 37.2	$\bar{x}_2 =$ 35.6
Sample variance	24.7	39.0
	$s_1 = 4.97$	$s_2 = 6.24$

Find the 95% confidence interval for the difference between the average mass of eggs in small and large nests.

$t^* = \text{invT}(.975, 59) = 2.00$

$(37.2 - 35.6) \pm 2.00\sqrt{\frac{4.97^2}{60} + \frac{6.24^2}{159}}$

1.6 ± 1.62

$[-.02, 3.22]$