10.2 Comparing Two Means

- ✓ DESCRIBE the shape, center, and spread of the sampling distribution of the difference of two sample means.
- ✓ DETERMINE whether the conditions are met for doing inference about $\mu_1 - \mu_2$.
- ✓ CONSTRUCT and INTERPRET a confidence interval to compare two means.
- ✓ PERFORM a significance test to compare two means.
- ✓ DETERMINE when it is appropriate to use two-sample *t* procedures versus paired *t* procedures.

Our parameters of interest are the population means μ_1 and μ_2. The best approach is to take separate random samples from each population and to compare the sample means.
Suppose we want to compare the average effectiveness of two treatments in a completely randomized experiment. We use the mean response in the two groups to make the comparison.

Population or treatment	Parameter	Statistic	Sample size
1	μ_1	$\bar{x}_1$	n_1
2	μ_2	$\bar{x}_2$	n_2

μ_1
μ_2

To explore the sampling distribution of the difference between two means, let's start with two Normally distributed populations having known means and standard deviations.
Based on information from the U.S. National Health and Nutrition Examination Survey (NHANES), the heights (in inches) of ten-year-old girls follow a Normal distribution N(56.4, 2.7) The heights (in inches) of ten-year-old boys follow a Normal distribution N(55.7, 3.8).
Suppose we take independent SRSs of 12 girls and 8 boys of this age and measure their heights.

What can we say about the difference $\bar{x}_f - \bar{x}_m$ in the average heights of the sample of girls and the sample of boys?

$N(55.7, 3.8)$
Normal μ σ

Choose an SRS of size n_1 from Population 1 with mean μ_1 and standard deviation σ_1 and an independent SRS of size n_2 from Population 2 with mean μ_2 and standard deviation σ_2.

Shape When the population distributions are Normal, the sampling distribution of $\bar{x}_1 - \bar{x}_2$ is approximately Normal. In other cases, the sampling distribution will be approximately Normal if the sample sizes are large enough ($n_1 \geq 30, n_2 \geq 30$).

Center The mean of the sampling distribution is $\mu_1 - \mu_2$.

$\mu_1 - \mu_2$

Spread The standard deviation of the sampling distribution of $\bar{x}_1 - \bar{x}_2$ is

$$SE = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

as long as each sample is no more than 10% of its population (10% condition).

Confidence Intervals for $\mu_1 - \mu_2$

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

2-Sample t interval

Significance Tests for $\mu_1 - \mu_2$

$$H_o: \mu_1 - \mu_2 = 0 \quad or \quad H_o: \mu_1 = \mu_2$$

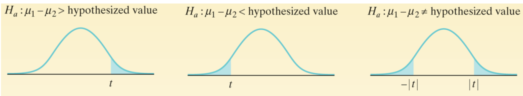
$$H_a: \mu_1 - \mu_2 > 0, \quad H_a: \mu_1 - \mu_2 < 0 \quad or \quad H_a: \mu_1 \neq \mu_2$$

$\mu_1 > \mu_2$ $\mu_1 < \mu_2$

To do a test, standardize $\bar{x}_1 - \bar{x}_2$ to get a two-sample t statistic:

$$\text{test statistic} = \frac{\text{statistic} - \text{parameter}}{\text{standard deviation of statistic}}$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$



(32)

$\mu_m = 69.3 \quad \sigma = 2.8$

$\mu_w = 64.5 \quad \sigma = 2.5$

$n_1 = 16$ (men) $\quad n_2 = 9$ (women)

a) Normal - stated the Pop. Normal

b) $\mu_m - \mu_w = 69.3 - 64.5 = 4.8$

c) $\sqrt{\frac{2.8^2}{16} + \frac{2.5^2}{9}} = 1.09$

(41)

$H_0: \mu_1 - \mu_2 = 0$

$H_a: \mu_1 - \mu_2 < 0$

μ_1 - Natural supply
μ_2 - Suppl. Supply

Conditions:

1. SRS - Randomly assigned
2. 10% Rule - Not need - Because Randomly assigned
3. Normal - Graph them Approx Normal - No Strong Skewness or outliers

2-sample t test

$t = -3.74$ p-val $= .0016$

$.0016 < .05$

Reject the H_0

We have strong evidence that the birds adjust their breeding habits based on food Supply.

.0016 is the probability of getting breeding lines as great or smaller than our sample if breeding really don't change (Same).

* 2 sample t interval 95%

$(-11.6, -3)$

Same Results!