CHAPTER 6

Study Questions

1. What is the advantage of a paired experimental design?

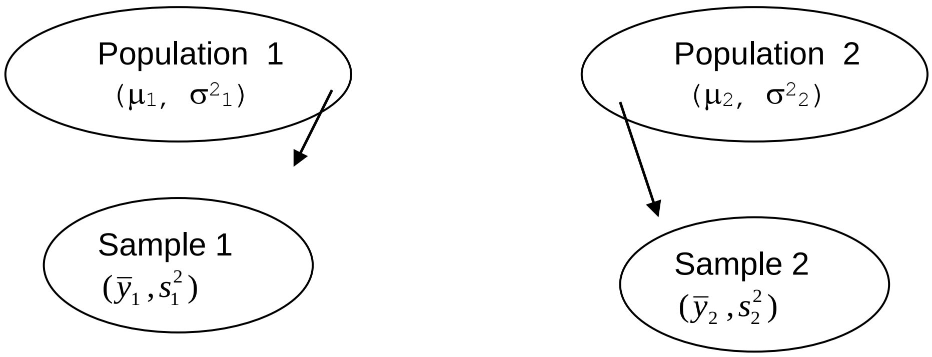
2. When the variances for two populations are assumed unequal, what adjustments do we make to the hypothesis test?

3. What is a pooled variance?

4. What assumptions are made when t_{obs} is used as the test statistic?

Inferences About μ_1 - μ_2

Frequently, we are interested in comparing the means of two populations. For example, a company may want to compare the mean life of its battery with the mean life of a competitor's battery. Another example would be a research lab that may be interested in comparing the effectiveness of two drug treatments for potential use in heart transplants.



Some Questions Concerning the Comparison of Two Population Means

1. Is the mean of the first population larger than the mean of the second population?
   $\mu_1 > \mu_2$ or $\mu_1 - \mu_2 > 0$
2. Is the mean of the second population larger than the mean of the first population?
   $\mu_1 < \mu_2$ or $\mu_1 - \mu_2 < 0$
3. Is there a difference in the population means?
   $\mu_1 \neq \mu_2$ or $\mu_1 - \mu_2 \neq 0$
4. Is the mean of the first population over 5 units larger than the mean of the second population?
   $\mu_1 - \mu_2 > 5$

Properties of the Sampling Distribution of the Difference of Sample Means

1. The sampling distribution of $(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)$ is approximately normal for large samples.
2. The mean of the sampling distribution is equal to the difference between the population means, $(\mu_1 - \mu_2)$.
3. The standard error of the sampling distribution is

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

Situations to be Considered

I. Independent Samples

Either samples are drawn from normal distributions or large sample sizes are used.

1. Variances for the two populations are equal. (Example 1)
2. Variances for the two populations are unequal. (Example 2)

II. Dependent or Paired Samples

Either samples are drawn from normal distributions or large sample sizes are used. (Example 3)

Example

Company D claims that its batteries last, on average, over 5 hours longer than its competitor Company E. To support its claim, the company tests ten of each type of battery. Below are the results of the study.

Company D Product Comparison Study	
Company D Battery Lifetime (hrs.)	**Company E Battery Lifetime (hrs.)**
20.21	14.75
18.94	14.47
20.92	12.50
21.36	10.53
19.83	13.09
19.07	16.53
18.17	13.47
20.34	15.71
21.30	14.26
19.05	14.43

Summary Statistics

$\bar{y}_D = 19.919$ $\bar{y}_E = 13.975$

$s_D^2 = 1.194$ $s_E^2 = 2.862$

Hypothesis Tests for Comparing Two Population Means
(Independent Samples and Equal Variances)

Hypotheses

H_0: $\mu_1 - \mu_2 = D_0$	H_0: $\mu_1 - \mu_2 = D_0$	H_0: $\mu_1 - \mu_2 = D_0$
H_a: $\mu_1 - \mu_2 < D_0$	H_a: $\mu_1 - \mu_2 > D_0$	H_a: $\mu_1 - \mu_2 \neq D_0$

Test Statistic

$$t_{obs} = \frac{(\bar{y}_1 - \bar{y}_2) - D_0}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$\text{where } S_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

Distribution for the Rejection Region

The rejection region is found using a t-distribution with $df = n_1 + n_2 - 2$.

Example 1

Hypothesis Test for the Company D Comparison Study

(Assuming Equal Variances and Normal Populations)

What are the hypotheses? H_o: $\mu_1 - \mu_2 = 5$ D=1, E=2

H_a: $\mu_1 - \mu_2 > 5$

What value of α should be used? Why? $\alpha=0.05$, both errors same consequence

Draw the rejection region for this problem.

$t_{crit} = t_{(18,0.05)} = 1.734$

Calculation of the test statistic

Pooled Variance

$$S_p^2 = \frac{(10-1)1.194 + (10-1)2.862}{10+10-2} = 2.028$$

Test Statistic

$$t_{obs} = \frac{19.919 - 13.975 - 5}{1.424\sqrt{\frac{1}{10}+\frac{1}{10}}} = 1.482$$

What is your conclusion?

FTR H_0; insufficient evidence to conclude that battery D lasts on average 5 hours more than battery E

Hypothesis Test for Comparing Two Population Means
(Independent Samples and Unequal Variances)

Hypotheses

H_0: $\mu_1 - \mu_2 = D_0$	H_0: $\mu_1 - \mu_2 = D_0$	H_0: $\mu_1 - \mu_2 = D_0$
H_a: $\mu_1 - \mu_2 < D_0$	H_a: $\mu_1 - \mu_2 > D_0$	H_a: $\mu_1 - \mu_2 \neq D_0$

Test Statistic

$$t'_{obs} = \frac{(\bar{y}_1 - \bar{y}_2) - D_o}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Distribution for the Rejection Region

The rejection region is found using a t-distribution with df given by

$$df = \frac{(n_1 - 1)(n_2 - 1)}{(n_2 - 1)c^2 + (1 - c)^2 (n_1 - 1)}$$

$$\textbf{where} \quad c = \frac{\frac{s_1^2}{n_1}}{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Example 2

Hypothesis Test for the Company D Comparison Study

(Assuming Unequal Variances and Normal Populations)

What are the hypotheses? H_o: μ_1-μ_2 = 5 D=1, E=2

H_a: μ_1-μ_2 > 5

What value of α should be used? Why? α=0.05, both errors same consequence

Draw the rejection region for this problem.

Degrees of Freedom

$$c = \frac{\frac{1.194}{10}}{\frac{1.194}{10} + \frac{2.862}{10}} = .294$$

$$df = \frac{(10 - 1)(10 - 1)}{(10 - 1)(.294)^2 + (1 - .294)^2 (10 - 1)} = 15.38$$

$t_{crit} = t_{(15,0.05)} = 1.753$

Calculation of the test statistic

Test Statistic

$$t'_{obs} = \frac{19.919 - 13.975 - 5}{\sqrt{\frac{1.194}{10} + \frac{2.862}{10}}} = 1.482$$

What is your conclusion? FTR H_0; insufficient evidence to conclude that battery D lasts on average 5 hours more than battery E