MTH 1210
Lab 6 Two-Mean and Paired-Samples Tests for Two Population Means μ_1 and μ_2

Part A: Two-Mean *t* Test for μ_1 and μ_2.

The ***two-mean t test*** is appropriate for comparing two (unknown) population means μ_1 and μ_2 when we have two *independent* samples from the populations.

We want to decide if there is convincing evidence against the **null** hypothesis

$$H_0: \mu_1 = \mu_2$$

versus one of the **alternatives**

1. $H_0: \mu_1 > \mu_2$
2. $H_0: \mu_1 < \mu_2$
3. $H_0: \mu_1 \neq \mu_2$

If both populations are normal *or* the sample sizes n_1 and n_2 are large (both ≥ 30), then when H_0 is true the ***two-mean t test statistic***

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

follows a *t*-distribution with degrees of freedom

$$df = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}}$$

Above, $\bar{x}_1$ and $\bar{x}_2$ are the sample means and s_1 and s_2 are the standard deviations. P-values are tail areas under this *t* distribution.

A random sample of **fruits** and another of **vegetables** were taken and the **moisture content** (by percent) measured in each. The data are shown below and in the file **fruits_veggies.mpx**.

Fruits		Vegetables	
Apricot	86	Artichoke	85
Banana	75	Bamboo Shoots	91
Avocado	72	Beets	88
Blackberry	88	Broccoli	89
Clementine	87	Cucumber	95
Fig	79	Iceberg Lettuce	96
Pink Grapefruit	92	Mushroom	92
Mango	84	Radish	95
		Tomato	94

We want to perform a test to decide which food type, if any, has the higher mean moisture content.

1. Open the data file **fruits_veggies.mpx**.
2. Make side-by-side boxplots of the two samples (fruits and vegetables) (GRAPH > BOXPLOT > MULTIPLE Y'S > SIMPLE).
3. **Copy and paste the side-by-side boxplots into the Answer Sheet**.
4. **Please answer Questions 1, 2 and 3 on the Answer Sheet.**
5. Assuming that the two samples came from normal distributions, carry out the appropriate test of the hypotheses:

STAT > BASIC STATISTICS > TWO-SAMPLE T …

In the dialog box: From the drop-down list: "Each sample is in its own column."
SAMPLE 1: Select the variable (double click).
SAMPLE 2: Select the variable (double click).
< OPTIONS >
HYPOTHESIZED DIFFERENCE: Make sure that the null-hypothesized difference is 0.0.
ALTERNATIVE HYPOTHESIS: Choose the form of the alternative hypothesis from the drop-down list using the following guideline:

	Hypothesis	**Minitab Notation**
Null	H_0: $\mu_1 = \mu_2$	
Alternative	H_a: $\mu_1 < \mu_2$	"Difference < hypothesized difference"
	H_a: $\mu_1 \neq \mu_2$	"Difference ≠ hypothesized difference"
	H_a: $\mu_1 > \mu_2$	"Difference > hypothesized difference"

< OK >

< OK >

The desired calculation will appear in the Session window.

6. **Please answer Questions 4, 5, and 6 on the Answer Sheet.**

Part B: Paired *t* Test for μ_1 and μ_2.

The ***paired-samples t test*** is appropriate for comparing two (unknown) population means μ_1 and μ_2 when we have ***paired samples*** (e.g. two measurements made on each of *n* individuals).

We'll want to test

$$H_0: \mu_1 = \mu_2,$$

Taking the ***difference*** for each pair converts the two samples into a *single* sample of differences. When H_0 is true, the ***paired t test statistic***

$$t = \frac{\bar{d} - 0}{s_d / \sqrt{n}},$$

follows a *t*-distribution with *n*-1 degrees of freedom. Above, $\bar{d}$ and s_d are the sample mean and standard deviation in the sample of *n* ***differences***. P-values are tail areas under this *t* distribution.

A study was conducted to decide if college students retain information better by reading the material in a **book** or by watching it on a **DVD**. A sample of $n = 7$ students from a college was selected and each was assigned one 19th-century novel to read and a different one to watch. Afterward, they were given a 20-point written **quiz** on each novel.

The test results are given below and in the file **books_dvds.mpx**.

Student	Book Test Result	DVD Test Result	Difference
1	90	85	5
2	80	72	8
3	90	80	10
4	75	80	-5
5	80	70	10
6	90	75	15
7	84	80	4

1. Open the data file **books_dvds.mpx**.
2. Make side-by-side boxplots of the two samples (book and DVD test results) as in Step 2 of Part A.
3. **Copy and paste the side-by-side boxplots into the Answer Sheet**.
4. **Please answer Question 1 on the Answer Sheet.**
5. Assuming that the two samples came from normal distributions, carry out the appropriate test of the hypotheses:

STAT > BASIC STATISTICS > PAIRED T …

In the dialog box: From the drop-down menu: "Each sample is in a column."
SAMPLE 1: Select the variable (double click).
SAMPLE 2: Select the variable (double click).
< OPTIONS >
HYPOTHESIZED DIFFERENCE: Make sure that the null-hypothesized value is 0.0.
ALTERNATIVE HYPOTHESIS: Choose the form of the alternative hypothesis from the drop-down list using the following guideline:

	Hypothesis	**Minitab Notation**
Null	H_0: $\mu_1 = \mu_2$	
Alternative	H_a: $\mu_1 < \mu_2$	"Difference < hypothesized difference"
	H_a: $\mu_1 \neq \mu_2$	"Difference ≠ hypothesized difference"
	H_a: $\mu_1 > \mu_2$	"Difference > hypothesized difference"

< OK >

< OK >

The desired calculation will appear in the Session window.

6. **Please answer Questions 2, 3, and 4 on the Answer Sheet.**