MATH 4720 / MSSC 5720

Instructor: Mehdi Maadooliat

Chapter 6 (Part A)

Department of Mathematical and Statistical Sciences

CHAPTER 6 (PART A)

- **Comparing Two Population Means**
  - **Independent Samples**
  - **Dependent Samples**
- **Two sample t-test (Independent Samples)**
  - **Pooled t-test**
  - **Unequal variance t-test**
- **Paired t-test (Dependent Samples)**
- **Power Analysis**
  - **Independent Samples**
  - **Dependent Samples**
- **Non-parametric Tests**
  - **Sign test (from Chapter 5, test for median M)**
  - **Wilcoxon Rank-Sum (or Mann–Whitney) Test (two independent samples)**
  - **Wilcoxon Signed-Rank Test (dependent samples)**

COMPARING TWO POPULATION MEANS

- **We discussed an example earlier to test if a diet used for losing weight is effective.**

- **Clinical Trial**

	Placebo	**New Diet**
Sample	$y_{11}, y_{12}, \dots, y_{1n_1}$	$y_{21}, y_{22}, \dots, y_{2n_2}$

- **If these two samples are drawn from populations with means μ_1 and μ_2 respectively, then the problem is testing**

- $H_0: \mu_1 = \mu_2$ **(Diet is not effective)**
- $H_a: \mu_1 > \mu_2$ **(Diet is effective)**

COMPARING TWO POPULATION MEANS CONT'D

- The two samples collected can be
  - Independent Samples
  - Dependent samples.

- The statistical methods are different for these two situations.

- **Independent Samples**

- If the samples are drawn from two different groups of population independently, samples are Independent.

	Group 1	Group 2
• Sample	$y_{11}, y_{12}, \ldots, y_{1n_1}$	$y_{21}, y_{22}, \ldots, y_{2n_2}$

- Group 1 could, for example, group of males, and Group 2 of females.

COMPARING TWO POPULATION MEANS CONT'D

- **Dependent Samples**

- **Here, there is only one group of subjects, but two different measurements are taken from this group.**

Subject	Before	After
1	y_{11}	y_{21}
2	y_{12}	y_{22}
.	.	.
.	.	.
n	y_{1n}	y_{2n}

- **Since the subjects are same for before and after measurements, the two samples are dependent.**

HYPOTHESIS TESTING FOR INDEPENDENT SAMPLES (EQUAL VARIANCES)

- **Independent Samples**

- $H_0: \mu_1 = \mu_2$
- $H_a: \mu_1 > \mu_2$

  or $\mu_1 < \mu_2$

  or $\mu_1 \neq \mu_2$

- **TS** $t = \frac{\bar{y}_1 - \bar{y}_2}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$ **(This is called pooled t-test)**

- **Here** $s_p = \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}$

HYPOTHESIS TESTING FOR INDEPENDENT SAMPLES CONT'D

- **Decision Rule:** $(df = n_1 + n_2 - 2)$
  - $H_a: \mu_1 > \mu_2$: **Reject in favor of** H_a **if** $t > t_\alpha$
  - $H_a: \mu_1 < \mu_2$: **Reject in favor of** H_a **if** $t < -t_\alpha$
  - $H_a: \mu_1 \neq \mu_2$: **Reject in favor of** H_a **if** $|t| > t_{\alpha/2}$
- **Note that the decision rule is same as we discussed for Chapter 5. Only difference is the degrees of freedom.**
- **P-values can be calculated is a similar ways.**

ASSUMPTIONS:

- **Note that, whenever we use t-statistics, there are assumptions. Also note that two samples are drawn from two populations. μ_1 is the mean of population 1, and σ_1 is its standard deviation. μ_2 is the mean of population 2, and σ_2 its standard deviation**

- **Assumption 1:** $\sigma_1 = \sigma_2$

- **Assumption 2:** $n_1 \geq 30, n_2 \geq 30$. **If not, we assume that both samples are drawn from normal populations.**

EXAMPLE

- **A study was conducted to see he effectiveness of a weight loss program. Two different groups of 10 subjects each were selected. The control group did not participate in the program. The data on weight loss was collected at the end of six months.**

Control	**Experimental**
$n_1 = 10$	$n_2 = 10$
$\bar{y}_1 = 2.1\ lb$	$\bar{y}_2 = 4.2\ lb$
$s_1 = 0.5\ lb$	$s_2 = 0.7\ lb$

- **Is there a sufficient evidence at $\alpha = 0.05$ to conclude that the program is effective? State all assumptions.**

EXAMPLE CONT'D

- **Assumptions:** **1.** $\sigma_1 = \sigma_2$
- **2. The weight loss for both groups are normally distributed.**

- $H_0: \mu_1 = \mu_2 \quad vs. \quad H_a: \mu_1 < \mu_2$

- **T.S.** $t = \frac{\bar{y}_1 - \bar{y}_2}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$

- $s_p = \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}} = \sqrt{\frac{9*0.5^2+9*0.7^2}{18}} = 0.6083$

- $t = \frac{2.1-4.2}{0.6083*\sqrt{\frac{1}{10}+\frac{1}{10}}} = -7.72$

EXAMPLE CONT'D

- **Decision Rule: Reject H_0 in favor of H_a if $t < -t_\alpha$**

- $df = n_1 + n_2 - 2 = 18,\ t_\alpha = 1.734$

- **Conclusion: If $t < -1.734$? Yes, since $t = -7.72$. Thus we reject H_0 in favor of H_a, and conclude that the program is effective.**

- **P-value Approach:** $p - value = P(t < -7.72) =$

- **Since $p - value < \alpha = 0.05$, we reject H_0 in favor of H_a, and conclude that the program is effective.**

HYPOTHESIS TESTING FOR INDEPENDENT SAMPLES (UNEQUAL VARIANCES)

- If we reject the equality of variances ($\sigma_1 \neq \sigma_2$):
  - We will talk about the test $H_0: \sigma_1 = \sigma_2$ vs $H_a: \sigma_1 \neq \sigma_2$ in Chapter 7
- We cannot use pooled t-test anymore.
- We should use the following t'- statistics.
- $t' = \dfrac{\bar{y}_1 - \bar{y}_2}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$
- Decision rules and p-value methods are same, except,
- $df = \dfrac{(n_1 - 1)(n_2 - 1)}{(1-c)^2(n_1 - 1) + c^2(n_2 - 1)}$, where $c = \dfrac{\frac{s_1^2}{n_1}}{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$.

CONFIDENCE INTERVAL FOR $(\mu_1 - \mu_2)$

- **Estimating the difference in the means $(\mu_1 - \mu_2)$**
- **Formula:**
- $\bar{y}_1 - \bar{y}_2 \pm t_{\alpha/2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$ **If** $\sigma_1 = \sigma_2$
- $\bar{y}_1 - \bar{y}_2 \pm t_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$ **if** $\sigma_1 \neq \sigma_2$
- **The df are same as we discussed earlier. Note that in both cases, we are assuming $n_1 \geq 30, n_2 \geq 30$, or the populations distributions of both populations are normal.**

BACK TO THE EXAMPLE

- **As we tested that the weight loss program is effective significantly. The next question is: how much effective?**

- **We can answer this by estimating $\mu_1 - \mu_2$ using 95% confidence interval.**

- $\bar{y}_1 - \bar{y}_2 \pm t_{\alpha/2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$ **since** $\sigma_1 = \sigma_2$
- $t_{\alpha/2}(df = 18) = 2.101$ **,** $s_p = 0.6083$

- **95% C.I. of $\mu_1 - \mu_2$:**

$$(2.1 - 4.2) \pm 2.101 * 0.6083 * \sqrt{\frac{1}{10} + \frac{1}{10}}$$

-2.1 ± 0.57, **i.e.** $-1.53\ lb \leq \mu_1 - \mu_2 \leq -2.67\ lb$

BOOK EXAMPLE 6.4

- **Does an oversized tennis racket exert less stress on the elbow?**

	Oversized	Conventional
Sample size	33	12
Sample mean	25.2	33.9
Sample St. Dev.	8.6	17.4

- $H_0: \mu_1 = \mu_2 \quad vs. \quad H_a: \mu_1 < \mu_2$

- **Note that since one of the sample size < 30, we are assuming that the samples are drawn from normal populations.**

BOOK EXAMPLE 6.4 CONT'D

- **To decide which test to use, pooled-t test or t'-test,**
- **We should test $\sigma_1 = \sigma_2$.**
  - **Next Chapter**

- **Let's say, we reject $H_0: \sigma_1 = \sigma_2$.**

- **Thus, we should use t'-test to test for the means.**

- $H_0: \mu_1 = \mu_2 \quad vs. \quad H_a: \mu_1 < \mu_2$

- **T.S.** $$t' = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{25.2 - 33.9}{\sqrt{\frac{8.6^2}{33} + \frac{17.4^2}{12}}} = -1.66$$

- **What is df?**

BOOK EXAMPLE 6.4 CONT'D

- **Degrees of Freedom:**

- $df = \frac{(n_1-1)(n_2-1)}{(1-c)^2(n_1-1)+c^2(n_2-1)}$, **where** $c = \frac{\frac{s_1^2}{n_1}}{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$.

- $c = \frac{\frac{8.6^2}{33}}{\frac{8.6^2}{33}+\frac{17.4^2}{12}} = 0.0816$

- $df = \frac{(33-1)(12-1)}{(1-0.0816)^2(33-1)+0.0816^2(12-1)} = 13.01$

- $\alpha = 0.05, \;\; t_\alpha(df = 13) = 1.771$

- **Note: If computed value of df is not an integer, always *round down* to the nearest integer.**

BOOK EXAMPLE 6.4 CONT'D

- **Decision Rule:**
- **Reject H_0 in favor of H_a if $t < -1.771$.**
- **Conclusion: Is $t < -1.771$? No, since $t = -1.66$. We fail to reject H_0, and cannot conclude that the oversized racket.**
- **Note: If we use the pooled t-test, we would reach the opposite conclusion (see the book)**

BOOK EXAMPLE 6.4 CONT'D

- **WRONG analysis based on pooled t-test:**

The standard practice in many studies is to always use the pooled t test. To illustrate that this type of practice may lead to improper conclusions, we will conduct the pooled t test on the above data. The estimate of the common standard deviation in repair costs σ is

$$\boxed{s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}} = \sqrt{\frac{(33 - 1)(8.6)^2 + (12 - 1)(17.4)^2}{33 + 12 - 2}} = \boxed{11.5104}$$

$$\boxed{\text{T.S.:} \quad t = \frac{(\bar{y}_1 - \bar{y}_2)}{s_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}} = \frac{(25.2 - 33.9)}{11.5104\sqrt{\dfrac{1}{33} + \dfrac{1}{12}}} = \boxed{-2.24}$$

The t-percentile for $\alpha = .05$ and $\boxed{\text{df} = 33 + 12 - 2 = 43}$ is given in Table 2 of the Appendix as 1.684 (for df = 40). We can now construct the rejection region.

R.R.: For $\alpha = .05$, and df = 43, reject H_0 if $t < -1.684$

Because $t = -2.24$ is less than -1.684, we would reject H_0 and conclude that there is significant evidence that the mean force of oversized rackets is smaller than the mean force of conventionally sized rackets. Using a software package, the p-value

BOOK EXAMPLE 6.4 CONT'D (CONFIDENCE INTERVAL)

- **95% Confidence Interval for $\mu_1 - \mu_2$:**

$$\bar{y}_1 - \bar{y}_2 \pm t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

- $t_{\alpha/2}(df = 13) = 2.16$

- **95% C.I. of $\mu_1 - \mu_2$:**

$$(25.2 - 33.9) \pm 2.16\sqrt{\frac{8.6^2}{33} + \frac{17.4^2}{12}}$$

- **i.e.** -8.7 ± 11.32,

$$-20.02 < \mu_1 - \mu_2 \leq 2.62$$

BOOK EXAMPLE 6.3 (PLUS R-CODE)

EXAMPLE 6.3

An experiment was conducted to evaluate the effectiveness of a treatment for tapeworm in the stomachs of sheep. A random sample of 24 worm-infected lambs of approximately the same age and health was randomly divided into two groups. Twelve of the lambs were injected with the drug, and the remaining 12 were left untreated. After a 6-month period, the lambs were slaughtered, and the worm counts recorded are listed in Table 6.3:

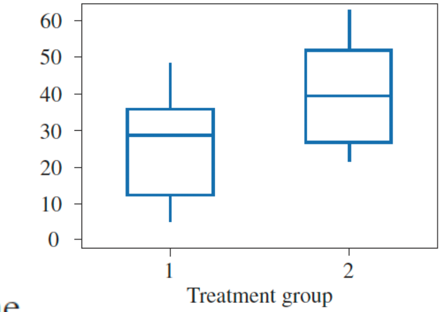
Drug-Treated Sheep	18	43	28	50	16	32	13	35	38	33	6	7
Untreated Sheep	40	54	26	63	21	37	39	23	48	58	28	39

a. Is there significant evidence that the untreated lambs have a mean tapeworm count that is more than five units greater than the mean count for the treated lambs? Use an $\alpha = .05$ test.
b. What is the level of significance for this test?
c. Place a 95% confidence interval on $\mu_1 - \mu_2$ to assess the size of the difference in the two means.

MARQUETTE UNIVERSITY
Be The Difference.

EXAMPLE 6.3 (SOLUTION)

Drug-Treated Lambs	**Untreated Lambs**
$n_2 = 12$	$n_1 = 12$
$\bar{y}_2 = 26.58$	$\bar{y}_1 = 39.67$
$s_2 = 14.36$	$s_1 = 13.86$



The sample standard deviations are of a similar size, so from this and from our observation from the boxplot, the pooled estimate of the common population standard deviation σ is now computed:

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} = \sqrt{\frac{11(13.86)^2 + 11(14.36)^2}{22}} = 14.11$$

The test procedure for evaluation of the research hypothesis that the untreated lambs have a mean tapeworm count (μ_1) that is more than five units greater than the mean count (μ_2) of the treated lambs is as follows:

H_0: $\mu_1 - \mu_2 \le 5$ (drug does not reduce the mean tapeworm count by more than 5 units)

H_a: $\mu_1 - \mu_2 > 5$ (drug does reduce the mean tapeworm count by more than 5 units)

$$\text{T.S.: } t = \frac{(\bar{y}_1 - \bar{y}_2) - D_0}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{(39.67 - 26.58) - 5}{14.11\sqrt{\frac{1}{12} + \frac{1}{12}}} = 1.404$$

R.R.: Reject H_0 if $t \ge 1.717$, where 1.717 is the value from Table 2 in the Appendix for a critical t-value with $\alpha = .05$ and df $= n_1 + n_2 - 2 = 22$.

EXAMPLE 6.3 (SOLUTION CONT'D)

b. Using Table 2 in the Appendix with $t = 1.404$ and df $= 22$, we can bound the level of significance (p-value) in the range $.05 < p\text{-value} < .10$.

Using the R function $\boldsymbol{pt}(\mathbf{t_c}, \mathbf{df})$, which calculates $P(t \le t_c)$, we can obtain the p-value for the calculated value of the T.S., $t_c = 1.404$.

$$p\text{-value} = P(t \ge 1.404) = 1 - P(t \le 1.404) = 1 - \boldsymbol{pt}(1.404, 22) = 1 - .913 = .087$$

```
tst1 <- t.test(exmp6.3$untreated, exmp6.3$treated,
               var.equal = T, mu = 5, alternative = "greater")
```

c. A 95% confidence interval on $\mu_1 - \mu_2$ provides the experimenter with an estimate of the size of the reduction in mean tapeworm count obtained by using the drug. This interval can be computed as follows:

$$(\bar{y}_1 - \bar{y}_2) \pm t_{.025} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$(39.67 - 26.58) \pm (2.074)(14.11)\sqrt{\frac{1}{12} + \frac{1}{12}} = 13.09 \pm 11.95 = (1.14, 25.4)$$

```
tst2 <- t.test(exmp6.3$untreated, exmp6.3$treated, var.equal = T)
tst2$conf.int
```

SUMMARY

Two Numerical Variables – Population Standard Deviations unknown

	Independent Samples	
	$\sigma_1 = \sigma_2$	$\sigma_1 \neq \sigma_2$
Parameters of Interest:	Mean Difference, $\mu_1 - \mu_2$	Mean Difference, $\mu_1 - \mu_2$
Confidence Interval Formula:	$\bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$ $df = n_1 + n_2 - 2$, where $s_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}$	$\bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$ $df = \frac{(n_1-1)(n_2-1)}{(1-c)^2(n_1-1) + c^2(n_2-1)}$, where $c = \frac{\frac{s_1^2}{n_1}}{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$
Name of Hypothesis Test, H_0	Two-sample T Test, $H_0: \mu_1 = \mu_2$ or $H_0: \mu_1 - \mu_2 = 0$	Two-sample T Test, $H_0: \mu_1 = \mu_2$ or $H_0: \mu_1 - \mu_2 = 0$
Test Statistic Formula:	$t = \frac{\bar{x}_1 - \bar{x}_2}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$	$t' = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$
P-value:	$H_a: \mu_1 \neq \mu_2$, p-value = $2P(T \geq \lvert t \rvert)$ $H_a: \mu_1 > \mu_2$, p-value = $P(T \geq t)$ $H_a: \mu_1 < \mu_2$, p-value = $P(T \leq t)$	