STAT 400 Stepanov
UIUC Dalpiaz

Examples for 7.2, 8.2

Two **independent** samples:

$X_1, X_2, \ldots, X_{n_1}$ from population 1, mean μ_1, std. dev. σ_1

$Y_1, Y_2, \ldots, Y_{n_2}$ from population 2, mean μ_2, std. dev. σ_2

If n_1 and n_2 are large, or population 1 and population 2 are approximately normal, then $\left(\overline{X} - \overline{Y}\right)$ is (approximately) normal with mean $\mu_1 - \mu_2$ and standard deviation $\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$.

A confidence interval for $\mu_1 - \mu_2$ is
$$\left(\overline{X} - \overline{Y}\right) \pm z_{\alpha/2} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}.$$

Test statistic for testing $H_0: \mu_1 - \mu_2 = \delta_0$ is
$$Z = \frac{\left(\overline{X} - \overline{Y}\right) - \delta_0}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}.$$

If σ_1 and σ_2 are unknown, use
$$\left(\overline{X} - \overline{Y}\right) \pm t_{\alpha/2} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

"Conservative" approach:

the number of degrees of freedom = the smaller of $n_1 - 1$ and $n_2 - 1$.

Welch's *T*:

$$\text{the number of degrees of freedom} = \left\lfloor \frac{\left(\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}\right)^2}{\dfrac{1}{n_1 - 1}\left(\dfrac{s_1^2}{n_1}\right)^2 + \dfrac{1}{n_2 - 1}\left(\dfrac{s_2^2}{n_2}\right)^2} \right\rfloor$$

If n_1 and n_2 are large, $t_{\alpha/2}$ can be approximated by $z_{\alpha/2}$.

1. Dr. Statman claims that his new revolutionary study method "Study While You Sleep" is more effective than the traditional study methods. In an experiment, 250 students enrolled in the same section of STAT 100 at UIUC were divided into two groups. One hundred students volunteered to study using SWYS method, and the other 150 students did whatever students usually do. At the end of the semester, the averages of the total number of points (out of 500) were compared for the two groups.

Note: This is **NOT** a good experiment design!

	SWYS	Traditional
(sample) average total points	450	410
(sample) standard deviation	20	45

a) Construct a 95% confidence interval for the difference in the average total points for SWYS and traditional study methods.

b) Perform the appropriate test at a 1% level of significance.

c) Test $H_0: \mu_S - \mu_T \le 30$ vs. $H_1: \mu_S - \mu_T > 30$ at $\alpha = 0.05$.

2. Two work designs are being considered for possible adoption in an assembly plant. A time study is conducted with 10 workers using design A and 12 workers using design B. The sample means and sample standard deviations of their assembly times (in minutes) are

	Design A	Design B
Sample Mean	78.3	85.6
Sample Standard deviation	4.8	6.5

Construct a 90% confidence interval for the difference in the mean assembly times between design A and Design B. Use Welch's *T*.

If we can assume that population 1 and population 2 standard deviations are equal (i.e., $\sigma_1 = \sigma_2 = \sigma$), then we can use

$$\left(\overline{X} - \overline{Y}\right) \pm t_{\alpha/2} \cdot s_{\text{pooled}} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

where

$$s_{\text{pooled}}^2 = \frac{(n_1 - 1)\cdot s_1^2 + (n_2 - 1)\cdot s_2^2}{n_1 + n_2 - 2}.$$

Then the number of degrees of freedom = $n_1 + n_2 - 2$.

3. A national equal employment opportunities committee is conducting an investigation to determine if women employees are as well paid as their male counterparts in comparable jobs. Random samples of 14 males and 11 females in junior academic positions are selected, and the following calculations are obtained from their salary data.

	Male	Female
Sample Mean	$48,530	$47,620
Sample Standard deviation	780	750

Assume that the populations are normally distributed with equal variances.

a) Construct a 95% confidence interval for the difference between the mean salaries of males and females in junior academic positions.

b) What is the p-value of the test $H_0: \mu_{\text{Male}} = \mu_{\text{Female}}$ vs. $H_1: \mu_{\text{Male}} \neq \mu_{\text{Female}}$?

The t Distribution

r	$t_{0.40}$	$t_{0.25}$	$t_{0.10}$	$t_{0.05}$	$t_{0.025}$	$t_{0.01}$	$t_{0.005}$
7	0.263	0.711	1.415	1.895	2.365	2.998	3.499
19	0.257	0.688	1.328	1.729	2.093	2.539	2.861
23	0.256	0.685	1.319	1.714	2.069	2.500	2.807
∞	0.253	0.674	1.282	1.645	1.960	2.326	2.576

Matched Pair Comparison:

Pair			Difference
1	X_1	Y_1	$D_1 = X_1 - Y_1$
2	X_2	Y_2	$D_2 = X_2 - Y_2$
.	.	.	.
.	.	.	.
.	.	.	.
n	X_n	Y_n	$D_n = X_n - Y_n$

Assume that the differences $D_i = X_i - Y_i$ are a random sample from normal distribution with mean δ and standard deviation σ_D.

A confidence interval for δ is $\overline{D} \pm t_{\alpha/2} \cdot \frac{s_D}{\sqrt{n}}$. The number of degrees of freedom = $n - 1$.

Test statistic for testing $H_0: \delta = \delta_0$ is $T = \frac{\overline{D} - \delta_0}{s_D / \sqrt{n}}$.

4. A new revolutionary diet-and-exercise plan is introduced. Eight participants were weighed in the beginning of the program, and then again a week later. The results were as follows:

Participant	1	2	3	4	5	6	7	8
Weight Before	213	222	232	201	230	188	218	182
Weight After	207	220	224	198	219	183	220	175
Pounds Lost	6	2	8	3	11	5	–2	7

a) Construct a 90% confidence interval for the average number of pounds lost during one week on that plan.

b) Is there enough evidence to conclude that the average weight loss is less than 7 pounds per week? (Use $\alpha = 0.05$.) What is the p-value of this test?