(9 points : 10 minutes)

1. To compare the quality of cakes cooked in a microwave oven to cakes cooked in a conventional oven, developers prepared seven cake recipes. Half of each recipe was cooked in a microwave oven and the other half of each recipe was cooked in a conventional oven. The quality of the cakes was judged by an expert panel and scored from 0 to 10. Treat the data below as quantitative and use them to make a 90% confidence interval for the difference in quality between microwave and conventional ovens.

Quality of Cakes			
	Oven		Diff.
Recip	Microwave	Conventional	C - M
1	5	9	4
2	6	9	3
3	6	9	3
4	8	6	-2
5	8	10	2
6	8	10	2
7	5	7	2

Claim: ____________________

H_0: ____________________

H_1: ____________________

(9 points : 10 minutes)

2. Use the data from the cake study in problem #1 to test the claim that conventional ovens are better than microwave ovens for cooking cakes, so the quality of cakes from conventional ovens is more than 1/2 point better on average than the quality of cakes from microwave ovens. Use a 5% significance level for your test.)

Claim: ____________________

H_0: ____________________

H_1: ____________________

(8 points; 7 minutes)

3. Use the summary statistics for a random selection of Saturdays in September at Folsom Lake to test whether the mean number of boats on the lake decreases by more than 100 boats when the temperature is less than 70°F compared to the number of boats when the temperature is greater than 90°F. Use a 0.025 significance level for your test. Variability in the number of boats is greater during cold weather than during hot heather.

Sample Statistic	Temp. < 70	Temp. > 90
days =	12	18
$\overline{x}$ =	217	336
s =	68	45

Solutions for problems on 1st page.

① This is a matched-pairs problem, because the values in each row are attached to the recipe listed in the first column.

~~"Test"~~ ~~claim that~~ Make a 90% CI $(\mu_c - \mu_m)$ where "c" is for conventional oven and "m" is for microwave oven.

Let $\mu_d = \mu_c - \mu_m$. Then 90% $CI(\mu_c - \mu_m) = 90\% CI(\mu_d)$

$$= \bar{d} \pm t_{\alpha/2}\left(\frac{s_d}{\sqrt{n}}\right) = 2 \pm 1.943\left(\frac{1.915}{\sqrt{7}}\right)$$

Put the differences in the calculator to get

$\bar{d} = 2$

$s_d = 1.915$

$n = 7$ d.f. $= 6$

$\alpha = 1 -$ confidence

$= 1 - 0.90$

$= 0.10$ in 2 tails

or 0.05 in 1 tail

$t = 1.943$

$$= 2 \pm 1.41$$

$$= [0.59 < \mu_d < 3.41]$$

(equivalently)

$$= [0.59 < (\mu_c - \mu_m) < 3.41]$$

Extra question: Is it reasonable to claim that conventional oven results are at least one quality point better than ~~the~~ microwave results?

Answer: No. If the statement is true, then $(\mu_c - \mu_m) \geq 1$. But, there are values less than 1 in the CI, which is the reasonable range for the true $(\mu_c - \mu_m)$.

#2. Same data as used for problem #1.
This is a hypothesis test.

Claim: quality from conventional ovens is more than ½ point better than quality from microwaves.

Symbolically: $\underbrace{\mu_c}_{\text{conventional quality}} \underbrace{>}_{\text{more than}} \underbrace{\mu_M + \tfrac{1}{2}}_{\text{1/2 point better than microwave}}$

$d = C - M$
$\bar{d} = 2$
$s_d = 1.915$
$n = 7$
$df = 6$

$t = 1.943$

so: $(\mu_c - \mu_m) > 0.5$

$H_0: (\mu_c - \mu_m) \leq 0.5$

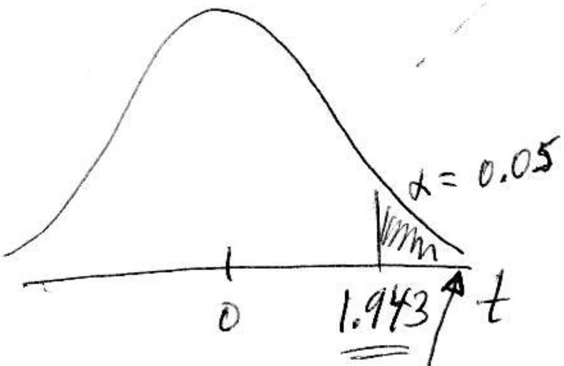
$H_1: (\mu_c - \mu_m) > 0.5$

$\alpha = 0.05$ in right tail

Test Statistic:

$$\frac{\bar{d} - \overbrace{\mu_d}^{= (\mu_c - \mu_m)}}{s_d/\sqrt{n}} = \frac{2 - 0.5}{(1.915/\sqrt{7})}$$

$$= \frac{1.5}{0.7238} = 2.072$$



Reject H_0!

#3. Test of difference between two means. For convenience, I will ~~use~~ use C = cold = (<70°) and H = hot = (>90°).

<u>Test whether</u>

$\mu_C < \mu_H - 100$

$\mu_H - 100 > \mu_C$

$(\mu_H - \mu_C) > 100$

$H_0: (\mu_H - \mu_C) \le 100$

$H_1: (\mu_H - \mu_C) > 100$

$\alpha = 0.025$ right tail

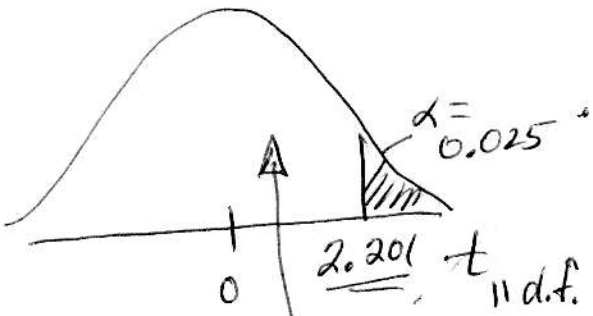
Note: $\sigma_C^2 > \sigma_H^2$ so $\neq$. Do Not pool variances. Use smaller of the two d.f.

Data

	Cold	Hot
$n =$	12	18
$df =$	11	17
$\bar{X} =$	217	336
$S =$	68	45

<u>Test Statistic</u>

$$\frac{(\bar{x}_H - \bar{x}_C) - (\mu_H - \mu_C)_0}{\sqrt{\dfrac{S_H^2}{n_H} + \dfrac{S_C^2}{n_C}}}$$



$$= \frac{(336-217) - 100}{\sqrt{\dfrac{(45)^2}{18} + \dfrac{(68)^2}{12}}} = \frac{19}{22.3} = 0.852$$

Do Not reject H_0!

(8 points : 10 minutes)

5. **Use the data on people's experience using two medications to help reduce the duration of a cold. Test the claim that the proportion of people who still have a cold after 4 days is smaller with Medication A than with medication B.** (Use a Type 1 error rate of 5% for this test.)

4-day Condition	Medication A	Medication B
No cold	27	22
Still Sick	63	88
Total	90	110

Claim: ______________________

H_0: ______________________

H_1: ______________________

(7 points : 8 minutes)

6. **Use the data from problem #5 to make a 95% confidence interval for (p_B - p_A), the difference between the proportions of people still having a cold after 4 days of taking medication B and medication A. Then, use your confidence interval to answer the question at the end of this problem.**

Based on your confidence interval, is it reasonable to say that $p_B < p_A$? Circle answer and explain.

YES **Why?** ______________________

NO ______________________

____

(8 points; 9 minutes)

7. **Based on the statistics shown below, test the claim that the percentage of 15 year old girls that have a personal cell phone is greater than the percentage of 15 year old boys that have a personal cell phone.** (For the test, use a Type 1 error rate of 0.05.)

Sample Statistics

Personal Cell Phone	15 Year Old Girls	15 Year Old Boys
Yes	90	57
No	45	35

H_0: ______________________

H_1: ______________________

Solutions for problems on 2^{nd} page.

(5) Hypothesis test concerning the proportion (p) of people who still have a cold after 4 days.

Claim: $p_A < p_B$ — A = used medication A, B = used medication B

$H_0: (p_A - p_B) \geq 0$

$H_1: (p_A - p_B) < 0$

$\alpha = 0.05$ left tail

$$\boxed{\begin{aligned} \hat{p}_A &= \frac{63}{90} = 0.7 \\ \hat{p}_B &= \frac{88}{110} = 0.8 \end{aligned}}$$

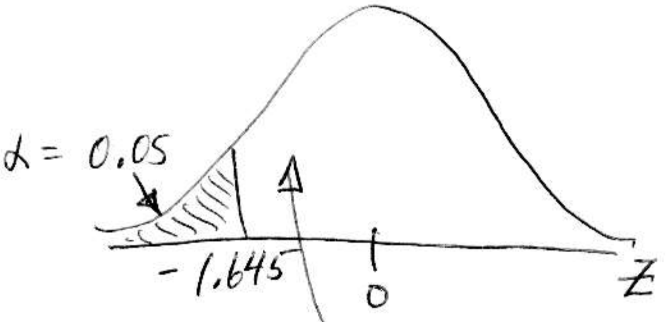
Because "0" is the hypothesized difference, we use $\bar{p}$ and $\bar{q}$:

$$\bar{p} = \frac{\text{all "successes"}}{\text{all trials}} = \frac{63+88}{\cancel{90}+110} = \frac{151}{200} = \boxed{0.755}$$

$$\bar{q} = \boxed{0.245}$$

Test Statistic:

$$\frac{(\hat{p}_A - \hat{p}_B) - \underbrace{(p_A - p_B)_0}_{\boxed{=0}}}{\sqrt{\dfrac{\bar{p}\bar{q}}{N_A} + \dfrac{\bar{p}\bar{q}}{N_B}}}$$



$$= \frac{(0.7 - 0.8) - 0}{\sqrt{\dfrac{(.755)(.245)}{90} + \dfrac{(.755)(.245)}{110}}} = \frac{-0.1}{0.0611} = \boxed{-1.64}$$

$\boxed{\text{Do Not reject } H_0\text{:}}$

#6 Using same data as in #5, make

$$95\%\ CI(p_B - p_A) = (\hat{p}_B - \hat{p}_A) \pm Z_{\alpha/2}\sqrt{\frac{\hat{p}_B\hat{q}_B}{N_B} + \frac{\hat{p}_A\hat{q}_A}{N_A}}$$

$\hat{p}_A = 0.7$	$\hat{p}_B = 0.8$
$\hat{q}_A = 0.3$	$\hat{q}_B = 0.2$
$N_A = 90$	$N_B = 110$

$$= (0.8 - 0.7) \pm 1.96\sqrt{\frac{(0.8)(0.2)}{110} + \frac{(0.7)(0.3)}{90}}$$

confidence $= 0.95$

$\alpha = 0.05$

$\alpha/2 = 0.025$

$Z_{\alpha/2} = 1.96$

$$= 0.1 \pm (1.96)(0.0615)$$

$$= 0.1 \pm \cancel{0.0121}\ 0.121$$

$$= \left[-0.021 < (p_B - p_A) < 0.221\right]$$

Is it reasonable to say that p_B is less than p_A?

$\boxed{\text{Yes}}$ Because, if $(p_B < p_A)$ then $(p_B - p_A) < 0$, and there are values in the CI $(p_B - p_A)$ that are < 0. The CI is the reasonable range for the true difference $(p_B - p_A)$.

#7 Test claim: % of 15-year old girls that have personal cell phone is greater than % of 15-year-old boys with such phones.

Personal Cell Phone	15 year old Girls	Boys
Yes	90	57
No	45	35
Total	135	92

$\hat{p} = 0.667 \quad 0.620$

$\hat{q} = 0.333 \quad 0.380$

$$\boxed{\bar{p} = \frac{90+57}{135+92} = 0.648 \qquad \bar{q} = 0.352}$$

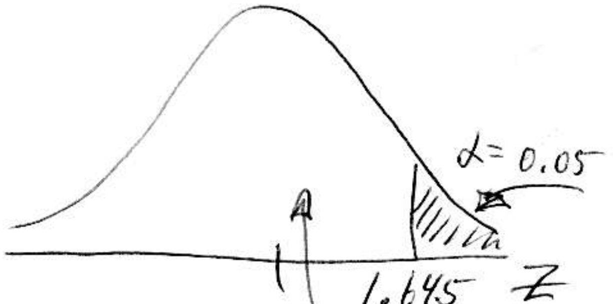
claim: $p_g > p_b \Rightarrow (p_g - p_b) > 0$

$H_0: (p_g - p_b) \leq 0$

$H_1: (p_g - p_b) > 0$

$\alpha = 0.05$ right tail

Because the hypothesized difference = 0, use $\bar{p}$ and $\bar{q}$ in denominator of the test statistic



Test statistic

$$\frac{(\hat{p}_g - \hat{p}_b) - \overbrace{(p_g - p_b)_0}^{\boxed{=0}}}{\sqrt{\frac{\bar{p}\bar{q}}{N_g} + \frac{\bar{p}\bar{q}}{N_b}}} = \frac{(0.667 - 0.620) - 0}{\sqrt{\frac{(.648)(.352)}{135} + \frac{(.648)(.352)}{92}}}$$

$$= \frac{0.047}{0.0646} = \boxed{0.73}$$

Do Not Reject Ho!

(10 points; 10 minutes)

1. **Compare the amount of energy per kilogram of two "biomass fuels" by testing the possibility that the average energy for Fuel A is more than 5 more than the average energy for Fuel B. The fuels are from very different sources, so it is likely that the variability of energy content from sample to sample is not the same. (Use a = 0.05 for this hypothesis test.)**

Sample	Fuel A	Fuel B
1	90	88
2	79	72
3	87	66
4	86	74
5	73	
$\overline{X}$ =		
s =		
n =		

H_0: ______________________

H_1: ______________________

(9 points; 10 minutes)

5. **Use the data below to prepare a 98% confidence interval for the difference between the proportion of college graduates that own a home and the proportion of those who have not graduated from college that own a home. The data come from random samples of 500 from each group.**

College Degree	Own a Home: Yes	Own a Home: No
Yes	313	187
No	218	282

Based on your confidence interval, is it reasonable for someone to claim that the proportion of college graduates that own a home is 60% but the proportion of non-graduates is 45% ?

YES **NO** **Why?** ______________________

____

____

Solutions for problems on 3rd page.

(#1) Test claim: average energy for Fuel A is more than 5 more than the average energy for Fuel B

Symbolic form: $\mu_A > \mu_B + 5$

(more than) (5 more than Fuel B)

so: $(\mu_A - \mu_B) > 5$

$H_0: (\mu_A - \mu_B) \le 5$

$H_1: (\mu_A - \mu_B) > 5$

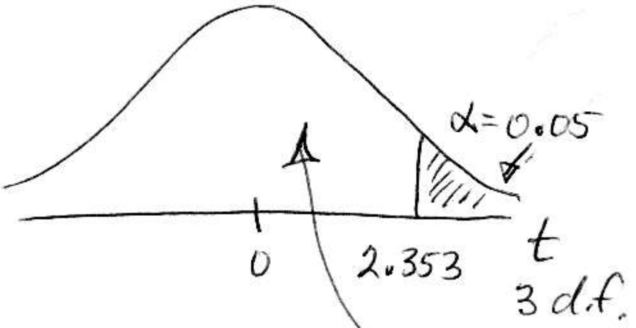
$\alpha = 0.05$ right tail

Problem says that variability is likely to be different for the two fuels, so: $\boxed{\sigma_A^2 \neq \sigma_B^2}$

Do not pool variances.

Use smaller of d.f. A and d.f. B

	A	B
$\bar{X} =$	83	75
$S =$	6.89	9.31
$n =$	5	4
$df =$	4	3



Test Statistic

$$\frac{(\bar{X}_A - \bar{X}_B) - \overbrace{(\mu_A - \mu_B)_0}^{\boxed{=5}}}{\sqrt{\frac{S_A^2}{n_A} + \frac{S_B^2}{n_B}}} = \frac{(83-75)-5}{\sqrt{\frac{(6.89)^2}{5} + \frac{(9.31)^2}{4}}} = \frac{3}{5.58} = 0.54$$

$\boxed{\text{Do not reject } H_0!}$

#5 Use the data to make: 98% $CI(p_Y - p_N)$ proportions that own home
where Y = yes for college degree
N = No for college degree

Degree	Own Home Y	Own Home N	total	$\hat{p}$	$\hat{q}$
Y	313	187	500	0.626	0.374
N	218	282	500	0.436	0.564

$$98\%(p_Y - p_N) = (\hat{p}_Y - \hat{p}_N) \pm Z_{\alpha/2}\sqrt{\frac{\hat{p}_y\hat{q}_y}{N_y} + \frac{\hat{p}_N\hat{q}_N}{N_N}}$$

$$\boxed{\begin{array}{l}\alpha = 0.02 \\ \alpha/2 = 0.01 \\ Z_{\alpha/2} = 2.33\end{array}}$$

$$= (0.626 - 0.436) \pm 2.33\sqrt{\frac{(.626)(.374)}{500} + \frac{(.436)(.564)}{500}}$$

$$= 0.19 \pm 2.33(0.031)$$

$$= 0.19 \pm 0.072$$

$$= \boxed{0.118 < (p_Y - p_N) < 0.262}$$

Is it reasonable for someone to claim that $p_y = 0.60$ but $p_N = 0.40$? YES Because then $(p_y - p_N) = 0.60 - 0.40 = \underline{0.20}$ and 0.20 is in the CI.

(8 points : 9 minutes)

9. **When the economy declines, more people have to ride public transit to work (I suppose). Thirty-eight people were selected in a study of the transit time from meadowview to downtown. Use the data given here to test the claim that the time needed to ride light rail downtown from the Meadowview area is less than 10 minutes more than the time needed to drive downtown from the Meadowview area. Variability in travel minutes is not the same for driving and for taking light rail.** (Use a 5% significance level for this test.)

Sample Results for Travel Minutes needed from Meadowview to Downtown

	Driving	Light Rail
$\overline{X}$ =	28.2	40.3
S =	4.6	6.3
N =	16	22

Claim: ______________________

H_0: ______________________

H_1: ______________________

(8 points : 8 minutes)

10. **When the economy declines, more people have to ride public transit to work (I suppose). Thirty-eight people were selected in a study of the transit time from meadowview to downtown. Use the data given here to construct a 95% confidence interval for the difference in travel time needed to go from the Meadowview area to Downtown. Assume for this analysis that variation in travel times is the same for both modes of transportation.**

Sample Results for Travel Minutes needed from Meadowview to Downtown

	Driving	Light Rail
$\overline{X}$ =	28.2	40.3
S =	4.6	6.3
N =	16	22

(7 points; 7 minutes)

9. **In a recent poll of Americans, 723 had private health insurance and 288 did not. Use the results below to make a 98% confidence interval for the difference between the proportions of these two groups with respect to their approval of President Obama's Health Care policies.**

Have Private Health Insurance	Approve of President's Health Care Policies		
	Yes	No	Total
Yes	209	514	723
No	165	123	288
Total	374	637	1011

Solutions for problems on 4th page

(#9) Test the claim that time need to travel by transit from Meadowview to Downtown (μ_T) is less than 10 minutes more than the time needed to drive from meadowview to Downtown (μ_D).

symbolic form: $\mu_T < \mu_D + 10$

$$\underbrace{\mu_T <}_{\text{less than}} \underbrace{\mu_D + 10}_{\text{10 more than driving}}$$

So: $(\mu_T - \mu_D) < 10$

$H_0: (\mu_T - \mu_D) \geq 10$

$H_1: (\mu_T - \mu_D) < 10$

$\alpha = 0.05$ left tail

Assume (says the problem) that variability is not the same for Transit times & driving times.

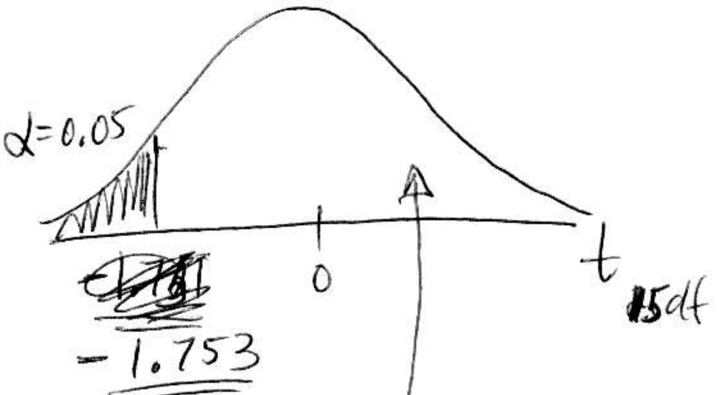
$\sigma_T^2 \neq \sigma_D^2$

So Do not pool variances.
Use smaller of $d.f._T$ and $d.f._D$.

$d.f = 15$

Test statistic

$$\frac{(\bar{x}_T - \bar{x}_D) - \overbrace{(\mu_T - \mu_D)_0}^{\boxed{=10}}}{\sqrt{\dfrac{S_T^2}{N_T} + \dfrac{S_D^2}{N_D}}}$$



$$= \frac{(40.3 - 28.2) - 10}{\sqrt{\dfrac{(6.3)^2}{22} + \dfrac{(4.6)^2}{16}}} = \frac{2.1}{1.768} = \boxed{1.187}$$

Do not reject H_0:

#10 Using the same data as for #9, make a 95% CI($\mu_T - \mu_D$) <u>BUT</u> assume variances are the same ($\sigma_T^2 = \sigma_D^2$), so pool variances and add the degrees of freedom.

$$S_{pool}^2 = \frac{(n_T-1)S_T^2 + (n_D-1)S_D^2}{(n_T-1)+(n_D-1)} = \frac{(22-1)(6.3)^2 + (16-1)(4.6)^2}{(22-1)+(16-1)}$$

$$= 31.97 \qquad df = 15+21 = 36$$

$$95\%\,CI(\mu_T - \mu_D) = (\bar{X}_T - \bar{X}_D) \pm t_{\alpha/2}\sqrt{\frac{S_{pool}^2}{N_T} + \frac{S_{pool}^2}{N_D}}$$

confidence = .95

$\alpha = 0.05$ in 2 tails

$\alpha/2 = 0.025$ in one tail

$t = 2.028$ for 36 df.

$$= (40.3 - 28.2) \pm 2.028\sqrt{\frac{31.97}{22} + \frac{31.97}{16}}$$

$$= 12.1 \pm (2.028)(1.858)$$

$$= 12.1 \pm 3.77$$

$$= [8.33 < (\mu_T - \mu_D) < 15.87]$$

#9 $98\% \ CI(p_Y - p_N) = (\hat{p}_Y - \hat{p}_N) \pm Z_{\alpha/2}\sqrt{\dfrac{\hat{p}_Y \hat{q}_Y}{n_Y} + \dfrac{\hat{p}_N \hat{q}_N}{N_N}}$

Y = yes, have private health insurance

N = NO, do not have private health insurance

Have private Health ins.	Approve Yes	No	(N) total
Yes	209	514	723
No	165	123	288

these totals at the bottom are not needed.

$\hat{p}$	$\hat{q}$
0.289	0.711
0.573	0.427

confidence = 0.98
$\alpha = 0.02$
$\alpha/2 = 0.01$
$Z_{\alpha/2} = 2.33$

$$CI = (0.289 - 0.573) \pm 2.33\sqrt{\frac{(.289)(.711)}{723} + \frac{(.573)(.427)}{288}}$$

$$= (-0.284) \pm (2.33)(0.0337)$$

$$= (-0.284) \pm 0.0785$$

$$= [-0.3625 < (p_Y - p_N) < -0.2055]$$

If you did $CI(p_N - p_Y)$, then your CI would be

$$[0.2055 < (p_N - p_Y) < 0.3625]$$