STAT 206: Chapter 10 (Fundamentals of Hypothesis Testing: Two-Sample Tests)

Ideas in Chapter 10:

- How to use hypothesis testing for comparing the difference between
  - The means of two related populations (dependent)
  - The means of two independent populations
  - The proportions of two independent populations
  - Relative Risk

- What we know about one (1) sample confidence intervals for population mean…
  - CI = point estimate ± margin of error
  - To use this method, you need:
    - Data obtained by randomization
    - Approximately normal data population distribution

- Confidence Interval for Mean Difference in ***Two*** Populations (***DEPENDENT*** Samples)
  - Want to compare two groups that are ***related to one another***, for example:
    - Ages of husbands and wives (i.e., couples ➔ ages probably NOT independent)
    - Before and after treatment of measurements of some medical test (i.e., same subjects/patients with before/after measurements ➔ "pair" is the same subject ➔ not independent)
    - Effectiveness of sunscreen in a left-arm / right-arm experiment (i.e., same subjects/individuals ➔ "pair" is the same subject ➔ not independent)
    - Braking distance for cars in wet / dry conditions (i.e., same cars, but conditions change ➔ not independent)

- Confidence Interval for Population Mean Difference (***DEPENDENT***)
  - Confidence Interval for dependent pairs is derived in same manner as for one sample mean – except that you're using the difference in your pairs ($value_1 – value_2$), ALWAYS in the same order, to find the DIFFERENCES to calculate the sample statistic, $\bar{X}_d$

- _____is the point estimate (statistic) for_____(parameter)

- Confidence Interval for _____is given by:

- ______is the number of pairs and degrees of freedom =_______

- **Interpretation of confidence intervals for two samples mean differences (i.e., DEPENDENT)**
  - Let **LL** = lower limit and **UL** = upper limit of a confidence interval for **(group A – group B).**
    That is, CI (μ_d) = CI($\mu_A - \mu_B$) = (LL , UL)
    - If LL and UL are both greater than 0, this suggests that group A has the greater mean

    - **Interpretation**: We are x%* confident that the population mean for group A is at least **LL** and at most **UL** units greater than the population mean for group B.

  - Let **LL** = lower limit and **UL** = upper limit of a confidence interval for **(group A – group B).**
    That is, CI (μ_d) = CI($\mu_A - \mu_B$) = (LL , UL)
    - If LL and UL are both less than 0, this suggests that group B has the greater mean

    - **Interpretation**: We are x%* confident that the population mean for group B is at least |**LL**| and at most |**UL**| units greater than the population mean for group A.

  - Let **LL** = lower limit and **UL** = upper limit of a confidence interval for **(group A – group B).**
    That is, CI (μ_d) = CI($\mu_A - \mu_B$) = (LL , UL)
    - If LL is less than 0 and UL is greater than 0, then neither group clearly has a greater mean

    - **Interpretation**: With x% confidence, it is unclear whether group A or group B has the greater population mean. If group A has the greater population mean, it is by at most UL units and if group B has the greater population mean, it is by at most |LL| units.

- **EXAMPLE**: Assume you send your salespeople to a "customer service" training workshop. Has the training made a difference in the number of complaints? You collect the following data:

Salesperson	# complaints BEFORE	# complaints AFTER	AFTER - BEFORE Difference
C.B	6	4	-2
T.F	20	6	-14
M.H.	3	2	-1
R.K.	0	0	0
M.O.	4	0	-4
			(sum) -21

Difference	sample diff.	Std. De.	Std. Err.	DF	Critical Value
after-before	-4.2	5.6745	2.5377	4	4.604

- **Example**: We are interested in whether there is a difference in the mean age at which men marry and the age at which women marry. The following data was collected from a random sample of 24 couples. Compute and interpret a 90% confidence interval on the mean difference between husbands' and wives' age at marriage. Assume that ages at marriage follow a normal distribution.

Difference	Sample Diff.	Std. De.	Std. Err.	DF	Critical Value
husband-wif	1.875	4.8122	0.9823	23	1.71387

couple	husband	wife	husband-wife
1	25	22	3
2	25	32	-7
3	51	50	1
4	25	25	0
5	38	33	5
6	30	27	3
7	60	45	15
8	54	47	7
9	31	30	1
10	54	44	10
11	23	23	0
12	34	39	-5
13	25	24	1
14	23	22	1
15	19	16	3
16	71	73	-2
17	26	27	-1
18	31	36	-5
19	26	24	2
20	62	60	2
21	29	26	3
22	31	23	8
23	29	28	1
24	35	36	-1

Hypothesis Tests for Population Mean Difference (***DEPENDENT***)

- Steps for a Hypothesis Test:
  - Check assumptions

  - Set up hypotheses:

  - Calculate test statistic. (Use software, and/or it will be given in output.)

  - Calculate p-value (Use software. It will be given in output)

  - Draw conclusion and **interpret the results**
    - If p-value ≤ α (or if p-value is less than .01 when no α is given), ➔ reject H_0 (With p-value = _______, we have sufficient evidence that_____________________
    - If p-value > α (or if p-value is greater than .10 when no α is given), ➔ do not reject H_0 (With p-value = _______, we do not have sufficient evidence that_____________________

  - **EXAMPLE**: Has the training made a difference in the number of complaints (at the 0.01 level)?

Difference	Sample Diff.	Std. De.	Std. Err.	DF	T. Statistic	P value
husband-wif	-4.2	5.6745	2.5377	4	-1.655	0.0866

- **EXAMPLE:** We are interested in whether there is a difference in the mean age at which men marry and the age at which women marry. The following data was collected from a random sample of 24 couples. Assume that ages at marriage follow a normal distribution. Test whether there is a difference in ages at which men and women marry using α = .10.

Difference	Sample Diff.	Std. De.	Std. Err.	DF	T. Statistic	P value
husband-wif	1.875	4.8122	0.9823	23	1.9088	0.0688

Example: (Salt Free Diet) Salt-free diets are often prescribed for people with high blood pressure. The following data are from an experiment designed to **estimate the reduction** in diastolic blood pressure (in units called millimeters of mercury (mm Hg)) as a result of following such a diet for 2 weeks. Assume diastolic readings follow a normal distribution.

a. Find and interpret the 99% confidence interval for the true mean reduction on the blood pressure.

Difference	Sample Diff	Std. Err.	DF	L. limit	U. limit
After-Before	-1	0.845	7	-3.9576	1.957

b. Test whether there is a reduction in diastolic blood pressure as a result of following a salt-free diet for 2 weeks.

Difference	Sample Diff	Std. Err.	DF	T. Statistic	P-value
After-Before	-1	0.845	7	-1.183	0.1377

- **But what if your populations are Independent?**
  Question: A recent study found that 51 children who watched a commercial for Walker Crisps (potato chips) featuring a well-known celebrity endorser ate a mean of 36 grams of Walker Crisps, but 41 children who watched a commercial for an alternative food snack ate a mean of 25 grams of Walker Crisps.
  - Is the assumption that the populations are dependent or independent?
    A. Dependent
    B. Independent

- **Two populations ➔ Two variances**
  - Methods to combine the variances:
    - If those variances are **UNEQUAL ➔ UNPOOLED**
    - If variances are **EQUAL ➔ POOLED**
  - UNPOOLED Std err =

    - Degrees of Freedom estimated by Welch-Satterthwaite equation (AWFUL! That's why the d.f. in the output looks so "strange")
  - POOLED Std err =

    - Degrees of Freedom =

  - Some sources point to the following **Rule of Thumb:** If the larger sample standard deviation is MORE THAN twice the smaller sample standard deviation then perform the t-test using the UNPOOLED method.
- **Comparing Two Means: INDEPENDENT**
  - Same steps as previous hypothesis tests:
    1. Check assumptions
    2. Set up hypotheses
    3. Calculate test statistic
    4. Calculate p-value
    5. Draw conclusion and interpret result

    - We summarize the test by reporting and interpreting the P-value
      - Smaller p-values give stronger evidence against the null hypothesis
      - If p-value $\leq \alpha$ ➔ REJECT H_0
      - If p-value $> \alpha$ ➔ FAIL to reject H_0

- With p-value = _______, we <*have / do not have*> sufficient evidence that <*state Ha in the context of the problem*>

- **Confidence Intervals for comparing two means using independent samples**
  - Formula for 95% confidence interval:

- **Interpretation of Confidence Intervals for comparing two means using independent samples**
  - Let LL = lower limit and UL = upper limit of a confidence interval for (group A – group B). That is, $\mu_A - \mu_B$= (LL , UL)

  - If LL and UL are both greater than 0, this suggests that group A has the greater mean.
  - **Interpretation:** We can be **x** % confident that the population mean for group A is at least LL and at most UL units greater than the population mean for group B.

  - If LL and UL are both less than 0, this suggests that group B has the greater mean.
  - **Interpretation:** We can be **x** % confident that the population mean for group B is at least |UL| and at most |LL| units greater than the population mean for group A.

  - If LL is less than 0, and UL is greater than 0, neither group clearly has the greater mean.
  - **Interpretation:** With **x** % confidence, it is unclear whether group A or group B has the greater population mean. If group A has the greater population mean, it is by at most UL units and if group B has the greater population mean, it is by at most |LL| units.

***Use correct Level of Confidence**

- **Example:** (variances assumed equal)You and some friends have decided to test the validity of an advertisement by a local pizza restaurant, which says it delivers to the dormitories faster than a local brand of a national chain. Both the local pizza restaurant and national chain are located across the street from your college campus. You define the variable of interest as the delivery time, in minutes, from the time the pizza is ordered to when it is delivered. You collect the data by ordering 10 pizzas from the local pizza restaurant and 10 pizzas from the national chain at different times. You organize and store the data in the excel spreadsheet shown. At the α=0.05 level, is there evidence that the mean delivery time for the local pizza restaurant is less than the mean delivery time for the national pizza chain?

Local	Chain		
16.8	22.0		
11.7	15.2		
15.6	18.7		
16.7	15.6		
17.5	20.8		
18.1	19.5		
14.1	17.0		
21.8	19.5		
13.9	16.5		
20.8	24.0		
n1=	n2=		
10	10		
means:			
16.7	18.88		=AVERAGE(data_string)
variances:			
9.58222	8.21511		=VAR.S(data_string)
std deviations:			
3.09552	2.8662		=SQRT(variance)
degrees of freedom:			
9	9	18	=SUM((n1-1)+(n2-1))
pooled variance:			
8.89867			=(variance1*df1+variance2*df2))/(df1+df2)
pooled standard error:			
1.33407			=SQRT(pooled_variance*(1/n1 + 1/n2))
mean diff:			
-2.18			=mean1 - mean2
t-stat:			
-1.6341			=(mean_diff - hypothesis_diff)/(pooled_std_err)
P-value:			
0.0598			=T.DIST(t-stat,df_pooled,TRUE)

Question:

Are the populations of delivery times for local and national pizzerias independent or dependent?

A. Independent
B. Dependent

- **95% CI =** $(\overline{X_1} - \overline{X_2}) \pm$ **t(std_error$_{\text{pool}}$)**

- **Interpretation:**

Example – *Ebay Sales*: Example 7 from Chapter 7 which compared the Ebay selling prices of the Palm M515 PDA. Some were sold using the Buy it Now option and some were sold using through the bidding option. The table shows data for both options. (The data was obtained from May 2003.)

a. Is there evidence, at the .05 level of significance, that there is a difference in the mean selling price of the two methods?
b. Find and interpret a 95% confidence interval for the difference in the mean selling price of the two methods

Difference	Sample Mean	Std. Err.	DF	T-Stat	P-value
$\mu_1 - \mu_2$	1.9603175	7.57266	16.589735	0.25886774	0.7989

Column	n	Mean	Variance	Std. Dev.	Std. Err.	Median	Range	Min	Max	Q1	Q3
Buy-It-Now	7	233.57143	214.28572	14.638501	5.5328336	235	40	210	250	225	250
Bidding	18	231.61111	481.1928	21.936108	5.17039	240	77	178	255	225	246

Buy-It-Now	Bidding
235	250
225	249
225	255
240	200
250	199
250	240
210	228
	255
	232
	246
	210
	178
	246
	240
	245
	225
	246
	225

Comparing two proportions…

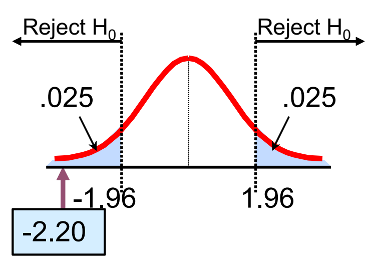
- Same steps as previous hypothesis tests:
  - Check assumptions
  - Set up hypotheses
  - Calculate test statistic
  - Calculate p-value
  - Draw conclusion and interpret results

- **If we wanted to do hypothesis testing…**
  - Hypothesis Test:
    H_0: $p_1 = p_2$ (that is, $(p_1 - p_2) = 0$)
    H_A: $p_1 \neq p_2$ (2-tailed test, > or < would be 1-tailed test)

  - Test Statistic: $z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{se_0}$

  - Where $\hat{p} = \frac{X_1 + X_2}{n_1 + n_2}$, with $se_0 = \sqrt{\hat{p}(1-\hat{p})(\frac{1}{n_1} + \frac{1}{n_2})}$

- **Example**: 2 Population Proportions – Is there a significant difference between the proportion of men and the proportion of women who will vote Yes on Proposition A? In a random sample, 36 of 72 men and 35 of 50 women indicated they would vote "Yes." Test at the .05 level of significance. Let p_1 be the proportion of men and p_2 be the proportion of women.
- Hypotheses: H_0: $p_1 - p_2 = 0$
  H_A: $p_1 - p_2 \neq 0$



- $z^* = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}_{pool}(1-\hat{p}_{pool})(\frac{1}{n_1} + \frac{1}{n_2})}}$

- Critical values =

- P-value =

- Decision:

- Conclusion:

- **CI for Two Population Proportions $p_1 - p_2$**
  - $CI = (\hat{p}_1 - \hat{p}_2) \pm z\sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2}}$
  - **EXAMPLE**: 95% CI for Men/Women voters on Proposition A (previous)
    - $CI = (\hat{p}_1 - \hat{p}_2) \pm z\sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2}}$
    - **Interpretation**:

- **Extra Topic: Comparing two proportions…**
  - Aspirin the wonder drug…
    - Pain reliever
    - Increase chance of survival for heart attack
    - Reduce the risk of colorectal cancer?
    - Reduce risk of other cancers, as well?
    - Meta-analysis combines results of several studies
      - Evidence that daily aspirin reduced deaths due to several common cancers during and after the trials (randomized and double-blind)

- **Aspirin the wonder drug…**
  - Contingency Table:

Death from Cancer			
group	**Yes**	**No**	**Total**
Placebo	347	11,188	11,535
Aspirin	327	13,708	14,035

  - Placebo group: proportion of cancer deaths

    $\hat{p}_1$ =

  - Aspirin group: proportion of cancer deaths

    $\hat{p}_2$ =

  - ($\hat{p}_1$ - $\hat{p}_2$)=

  - Standard error of ($\hat{p}_1$ - $\hat{p}_2$): se =

  - $99\% CI$ =($\hat{p}_1$ - $\hat{p}_2$) ± z (*se*)

- **Ratio of Proportions: Relative Risk**
  - Because the proportions in a clinical setting are often very small, sometimes it is hard to interpret what the tiny numbers may mean…
  - Ratio of proportions describes the sizes of the proportions ***relative*** to each other…

  - *Sample relative risk* $= \frac{\hat{p}1}{\hat{p}2} = \frac{0.030}{0.023} = 1.30$

  - Means that the proportion of the placebo group who had a cancer death was 1.30 times the proportion of the aspirin group who had a cancer death
  - Medical journals often report relative risk, especially when both proportions are close to 0 (e.g., cancer is a fairly “rare event” so the proportion of the population is small)