EDUC/PSY 6600 Name: **Unit 6 Assignment**

Research Design & Analysis I **Categorical Data Analysis**

---

*Please complete the following exercises. Feel free to work with classmates, but each student must turn in **UNIQUE** work, not photocopies or identical replicates. When applicable, use **APA format** in communicating your results in text. **Show your work!** If any question involves any math at all, show your work. When it doubt, write it out. Always show more than you think you need.*

1) WRITE-UP - Textbook Problems

Cohen Chap		Exercises	Pts	Off
19	A	3,*6	4	
	B	2,*5	4	
20	A	3,*7, 8	7	
	B	4, 8	5	
	C	1, 3	5	

2) SUMMARY – Supplementary Reading

		Pts	Off
Half Page	Read the Unit 6 Journal Article on Canvas. Summarize any mention or use/abuse of the concepts in the above chapters.	10	

4) SPSS SYNTAX – Section C: Ihno's data set – add to the skeleton SPSS syntax file (.sav)

Cohen Chap		Exercises	Pts	Off
20	A	3,*7, 8	5	
	B	4, 8	5	
	C	1, 3	5	

Grading		Earned	Possible
CORRECTNESS	*a subset of spot-checked items: must show work, especially items from back of book or done in class*		50
COMPLETENESS	*more than one item is missing or skipped: 25/50 roughly half the assignment is completed: 10/50*		50
			100

19 A	3. Use Table A.13 "Prob for the Binomial Distribution for p=.5"

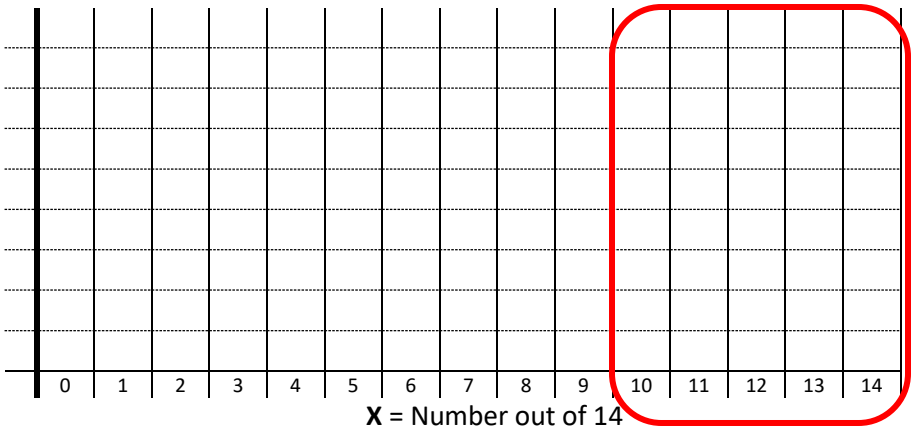
Fourteen infants are simultaneously shown a picture of a human face and a colorful ball of the same size (relative positions of the pictures are varied).

a) If 10 of the infants spend more time looking at the face than the ball, can the null hypothesis (no preference between the face and the ball) be **rejected** at the **.05 level (2-tailed test)?**

N = _____
X = _____
H_0: P = _____

Table A.13
Probabilities of the Binomial Distribution for $P = .5$

n	X	p
14	0	.0001
	1	.0009
	2	.0056
	3	.0222
	4	.0611
	5	.1222
	6	.1833
	7	.2095
	8	.1833
	9	.1222
	10	.0611
	11	.0222
	12	.0056
	13	.0009
	14	.0001



P(X >= 10) = _____
1-sided p-value = _____
2-sided p-value = _____

☐ Infants PREFER faces over balls
☐ No evidence of a preference

b) In a **two-tailed** test, **how many infants** must spend more time looking at the face than the ball to allow the researcher to reject the null hypothesis at the **.05 level**? At the **.01 level**?

X	"Probability" P(X)	1-tail p-value	2-tail p-value
14			
13			
12			
11			
10			
9			
…			

.05 level: _____
.01 level: _____

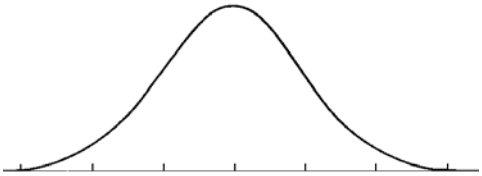
19 A	*6. Normal approximation to the Binomial Distribution

Suppose that 85% of the population is right handed (Q=.85) and 15% are left-handed (P=.15).

a) If out of **120** randomly selected civil engineers, **27** are found to be left-handed, what is the z score for testing the null hypothesis?

N = _____
X = _____
H_0: P = _____

Formula 19.1
$$z = \frac{X - NP}{\sqrt{NPQ}}$$



z= _____
p-value= _____

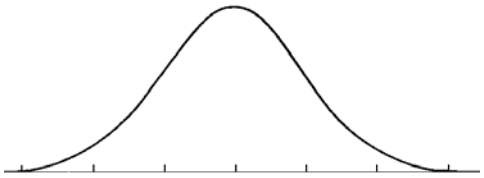
Can we reject the null hypothesis that **P = .15** for this profession?

- ☐ Evidence of more left handedness among civil engineers
- ☐ Evidence of less left handedness among civil engineers
- ☐ This sample provides no such evidence

b) If **480** civil engineers are sampled and **108** are found to be left-handed, what is the z score for testing the null hypothesis?

N = _____
X = _____
H_0: P = _____

Formula 19.1
$$z = \frac{X - NP}{\sqrt{NPQ}}$$



z= _____
p-value= _____

c) How does the z score in part (a) **compare** to the z score in part (b)?

Can you determine a **general rule** that is being illustrated?

19 B | 2. Sign Test (page 700) vs. Matched Pairs test

Perform the sign test on the data from exercise 11B #6 using the same alpha level and number of tails.

Boy	110	100	120	90	108	115	122	110	127	118
Girl	105	108	110	95	105	125	118	116	118	126
sign										

Hint: X = # of pairs with boy > girl, then use the binomial distribution

N = _____
X = _____
H_0: P = _____

☐ Evidence boys score higher than girls
☐ Evidence boys score lower than girls
☐ This sample provides no such evidence

Table A.13
Probabilities of the Binomial Distribution for $P = .5$

n	X	p
10	0	.0010
	1	.0098
	2	.0439
	3	.1172
	4	.2051
	5	.2461
	6	.2051
	7	.1172
	8	.0439
	9	.0098
	10	.0010

Did you reach the **same** conclusion with the sign test as with the matched t test?

☐ **Same**
☐ **Not the same**

If not, **explain** the discrepancy.

```
        Paired t-test

data:  score by sex
t = -0.25048, df = 9, p-value = 0.8078
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -6.018694  4.818694
sample estimates:
mean of the differences
                  -0.6
```

19 B | *5. Sign Test (ignore ties)

Does the mental condition of a chronic schizophrenic tend to **deteriorate over time** spent in a mental institution? TO answer this question, nine patients were assessed clinically after 2 years on a ward and again 1 year later. These clinical ratings appear below:

Patient #	1	2	3	4	5	6	7	8	9
Time 1	5	7	4	2	5	3	5	6	4
Time 2	3	6	5	2	4	4	6	5	3
sign									

Assume that these clinical ratings are so **crude** that it would be misleading to calculate the difference score for each subject and perform a matched t test. However, the **direction** (i.e. sign) of each difference is considered meaningful, so the sign test can be performed. Test the null hypothesis (alpha = .05, two-tailed) that there is **no difference over time** for such patients.

Table A.13
Probabilities of the Binomial Distribution for $P = .5$

n	X	p
8	0	.0039
	1	.0312
	2	.1094
	3	.2188
	4	.2734
	5	.2188
	6	.1094
	7	.0312
	8	.0039

N = _____
X = _____
H_0: P = _____

☐ Evidence of deterioration in mental condition over time
☐ This sample provides no such evidence

20 A	3. One-way Chi-Squared Test: Goodness-of-Fit	Code: R notebook

A soft drink manufacturer is conducting a blind taste test to compare its best-selling product (X) with two leading competitors (Y and Z). Each subject tastes all three and selects the one that tastes best to him or her.

a) DESCRIBE the appropriate **null hypothesis** for this study?

b) If 27 subjects prefer product X, 15 prefer product Y, and 24 prefer product Z, **can you reject the null** hypothesis at the .05 level?

N = _____
k = _____
df = _____

	X	Y	Z	total
OBSERVED				
EXPECTED				

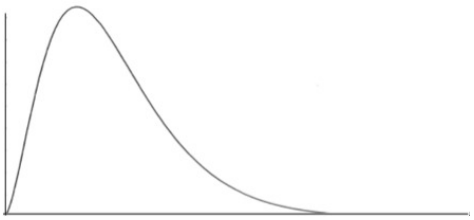
1-way Chi-Square Test

$k = number\ of\ groups$

$df = k - 1$

Formula 20.1

$$\chi^2 = \sum \frac{(OBS - EXP)^2}{EXP}$$



α=.05 → χ^2_{cv}(___) = _____

χ^2(___) = _____, p = _____

Write your conclusion regarding the evidence this sample provides/doesn't provide pertaining to **soft drink preference.**

20 A	*7. One-way Chi-Squared Test: Goodness-of-Fit	Code: R notebook

It has been suggested that admissions to psychiatric hospitals may vary by season. One hypothetical hospital admitted 100 patients last year: 30 in the spring, 40 in the summer, 20 in the fall, and 10 in the winter. Use the chi-squared test to evaluate the hypothesis that mental illness emergencies are evenly distributed throughout the year.

DESCRIBE how to find the **expected values** in this situation.

N = _____
k = _____
df = _____

	Spring	Summer	Fall	Winter	total
OBSERVED					
EXPECTED					

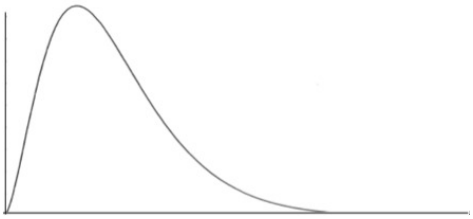
1-way Chi-Square Test

$k = number\ of\ groups$

$df = k - 1$

Formula 20.1

$$\chi^2 = \sum \frac{(OBS - EXP)^2}{EXP}$$



α=.05 → χ^2_{cv}(___) = _____

χ^2(___) = _____, p = _____

Write your conclusion regarding the evidence this sample provides/doesn't provide pertaining to **seasonal mental illness emergencies.**

20 A	8. One-way Chi-Squared Test: Goodness-of-Fit	Code: R notebook

Of the 100 psychiatric patients referred to in the previous exercise, 60 were diagnosed as schizophrenic, 30 were severely depressed, and 10 had bipolar disorder.

Assuming the **national percentages** for psychiatric admissions are 55% schizophrenic, 39% depressive, and 6% bipolar, use the chi-squared test to evaluate the null hypothesis that this particular hospital is receiving a random selection of psychiatric patients from the national population.

DESCRIBE the appropriate **null hypothesis** now?

Use the chi-squared test to evaluate the hypothesis.

N = _____
k = _____
df = _____

	Schizophrenic	Depressed	Bipolar	total
OBSERVED (sample)				
EXPECTED (nation)				

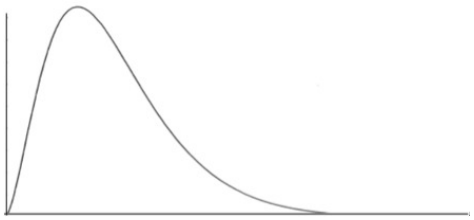
1-way Chi-Square Test

$k = number\ of\ groups$

$df = k - 1$

Formula 20.1

$$\chi^2 = \sum \frac{(OBS - EXP)^2}{EXP}$$



α=.05 → χ^2_{cv}(___) = _____

χ^2(___) = _____, p = _____

Write your conclusion regarding the evidence this sample provides/doesn't provide pertaining to **the patients being representative of the national admissions.**

20 B	4. Two-way Chi-Squared Test: Independence	Code: R notebook

A social psychologist is studying whether people are more likely to help a poor person or a rich person who they find lying on the floor. The three conditions all involve an elderly woman who falls down in a shopping mall (when only one person at a time is nearby). The independent variable concerns the apparent wealth of the woman; she is dressed to appear either poor, wealthy, or middle class. The reaction of each bystander is classified in one of three ways: ignoring her, asking if she is all right, and helping her to her feet. The data appear in the **contingency table** below: *(hint –write the expected count in each cell of the table below)*

a) Test the null hypothesis at the **.01 level**.

	Poor	Middle Class	Wealthy	
Ignores	**16**	**10**	**7**	33
Talks to her	**8**	**6**	**5**	19
Helps her up	**6**	**14**	**18**	38
	30	30	30	90

Formula 20.3

$r = row\ groups$

$c = column\ groups$

$df = (r-1)(c-1)$

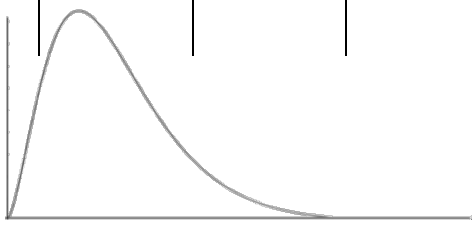
Formula 20.4

$$EXP = \frac{n_r \times n_c}{n_t}$$

Formula 20.1

$$\chi^2 = \sum \frac{(OBS - EXP)^2}{EXP}$$

N = _____
r = _____
c = _____
df = _____



α=.01 → χ^2_{cv}(___) = _____
χ^2(___) = _____, p = _____

Is there **evidence for an association** between the apparent wealth of the victim and the amount of help provided by a bystander?

b) Calculate **Cramer's phi** for these data.

Formula 20.8

$$L = \min(r,c) \rightarrow \phi_C = V = \sqrt{\frac{\chi^2}{N(L-1)}}$$

ϕ_C = _____

What can you say about the **strength** of the relationship between the two variables?

20 B	8. Two-way Chi-Squared Test: Independence	Code: R notebook

In exercise 12B #4, the dependent variable was the amount of time a subject listened to tape recorded cries for help from the next room before getting up to do something. If some subjects never responded within the time allotted for the experiment, the validity of using parametric statistical techniques could be questioned. As an alternative, subjects could be classified as fast or slow responders (and possibly, non-responders). The data were classified as fast (less than 12 seconds) and slow (12 or more seconds). The resulting contingency is shown below: *(hint – write the expected count in each cell of the table below)*

a) Test the null hypothesis (**.05 level**) that speed of response is **independent** of type of voice heard.

	Child's Voice	Adult Female	Adult Male	
Fast	**5**	**3**	**1**	9
Slow	**2**	**4**	**6**	12
	7	7	7	21

Formula 20.3

$r = row\ groups$

$c = column\ groups$

$df = (r-1)(c-1)$

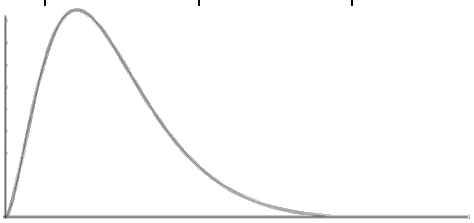
Formula 20.4

$$EXP = \frac{n_r \times n_c}{n_t}$$

Formula 20.1

$$\chi^2 = \sum \frac{(OBS - EXP)^2}{EXP}$$

N = _____
r = _____
c = _____
df = _____



α=.05 → χ^2_{cv}(___) = _____

χ^2(___) = _____, p = _____

Write your conclusion regarding the evidence of an association between the **speed of response** and the **reaction time.**

b) How does your conclusion in part a **compare** with the conclusion you drew in 12B #4?

```
Anova Table (Type 3 tests)

Response: seconds
      num Df den Df    MSE      F     ges  Pr(>F)
voice      2     18 26.476 3.7464 0.29392 0.04362 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Categorizing the dependent variable throws away information; how do you think that loss of information **affects power**?

20 C | 1. One-way Chi-Squared Test: Goodness-of-Fit — Code: R notebook

a) Perform a one-way chi-squared test to determine whether you can reject the null hypothesis that, an Ihno's university, there are the **same number of students majoring in each of the five areas** represented in Ihno's class, if you assume that Ihno's students represent a random sample with respect to major area.

	Psych	PreMed	Bio	Soc	Econ	total
OBSERVED						
EXPECTED						

Write your conclusion regarding the evidence this sample provides/doesn't provide pertaining to **equal representation of all five majors.**

χ^2(___) = _____, p = _____

b) Perform the test in part (a) separately for both the **MALES**…

MALES: χ^2(___) = _____, p = _____

☐ Evidence of at least one major having more or fewer students

☐ This sample provides no evidence of inequality in major

…and the **FEMALES** in Ihno's class.

FEMALES: χ^2(___) = _____, p = _____

☐ Evidence of at least one major having more or fewer students

☐ This sample provides no evidence of inequality in major

20 C | 3. Two-way Chi-Squared Test: Independence — Code: R notebook

Conduct a two-way chi-squared analysis of Ihno's data to test the null hypothesis that the **proportion of females is the <u>same</u> for each of the five represented majors** in the entire university population. Request a statistic to describe the strength of the relationship between gender and major.

	Psych	PreMed	Bio	Soc	Econ	total
Female						
Male						

Formula 20.8

$$L = \min(r, c)$$

$$\phi_C = V = \sqrt{\frac{\chi^2}{N(L-1)}}$$

χ^2(___) = _____ , p = _____ , ϕ_C = _____

☐ Evidence females enroll in different majors than males

☐ This sample provides no evidence of inequality in major