Homework 10: Chi-Square, Data Transformations, and Non-Parametric Tests

Due at the final exam

Homework
Factorial ANOVA Questions 1-2
Chi-Square Questions 3-6
Omit Data Transformations Questions 7-8
Non-Parametric Questions 9-10

Homework hints:

- Draw the tables for each of the Chi-Square problems.
- Clearly label graphs and transformation steps.

SPSS Assignment
SPSS *Chi-Square*: **There is no SPSS component for Chi-Square.**

Note: It is good practice to learn how to conduct the χ^2 tests because several of you will use it to analyze data for your Research Methods projects. This is often the case when you collect naturalistic data that you have no way of experimentally manipulating the behavior (e.g., candy preference or political persuasion).

Factorial ANOVA

1. A sports psychologist studied the effects of a motivational program compared to no program on injuries among players of three different sports (baseball, football, and basketball). The dependent variable in the study is the number of injuries per person over a period of 10 weeks.
   a. Identify the design and the levels of the independent variables.
   b. Set up a matrix of the design with levels of independent variables.
   c. Fill in the matrix with hypothetical values that show a main effect for type of sport.

2. Describe a between-subjects factorial ANOVA, a repeated-measures factorial ANOVA, and a mixed-factorial ANOVA.

Chi-Square

3. Carry out a chi-square test for *goodness-of-fit* for the following study (use $\alpha = 0.05$ level).

Category	Number in the past	Observed
1	100	38
2	300	124
3	50	22
4	50	16

4. A director of a small psychotherapy clinic is wondering whether there is any difference in the use of the clinic during different seasons of the year. Last year, there were 28 new clients in the winter, 33 in the spring, 16 in the summer, and 51 in the fall. On the basis of last year's data, should the director conclude that season makes a difference? (use α=0.05). Carry out the steps for testing the hypothesis to answer this question.

5. A political psychologist is interested in determining whether the community a person lives in is related to that person's opinion on a water-conservation ballot initiative. The psychologist surveys 90 people by phone with the results in the table below. Is opinion related to community at the α= 0.05 level? Do only two tasks for this problem:

    a. Explain why you would analyze these data with the χ^2 test for independence
    b. Compute the test and determine the outcome.

	Community A	**Community B**	**Community C**
For	12	6	3
Against	18	3	15
No Opinion	12	9	12

6. A researcher wants to be sure that the sample in her study is representative of the distribution of ethnic groups in her community. Her sample includes 300 whites, 80 African Americans, 100 Latinos, 40 Asians, and 80 others. According to census records, there are 48% whites, 12% African Americans, 18% Latinos, 9% Asian, and 13% others in her community. Is her sample representative of the population in the community? Do only two tasks for this problem:

    a. Explain why you would analyze these data with the χ^2 test for goodness-of-fit
    b. Compute the test so that you can determine the outcome.

~~Data Transformations (When parametric assumptions might not be met)~~

7. ~~For the distribution of the following 30 scores,~~

    ~~9, 28, 4, 16, 0, 7, 25, 1, 4, 10, 4, 2, 1, 9, 16, 11, 12, 1, 18, 2, 5, 10, 3, 17, 6, 4, 2, 23, 21, 20~~

    a. ~~Make a grouped frequencies histogram of the scores as they are (intervals 0-4.9,5-9.9,10-14.9, etc.)~~
    b. ~~Carry out a square-root transformation~~
    c. ~~Make a grouped histogram of the transformed scores (0-.9, 1-1.9, etc.)~~

8. ~~Miller (1997) conducted a study of commitment to a romantic relationship and how much attention a person pays to attractive alternatives. In this study, participants were shown a set of slides of attractive individuals. At the start of the results section, Miller notes, "The self-reports on the Attractiveness to Alternative Index and the time spent actually inspecting the attractive opposite-sex slides…were positively skewed, so logarithmic transformations of the data were performed" (p.760). Explain what is being~~

~~described here (and why it is being done) to a person who understands ordinary parametric statistics but has never heard of data transformations.~~

Nonparametric Tests (Understanding data in research reports)

9. Ford and colleagues (1997) were interested in the relation of certain personality factors to treatment for post-traumatic stress disorder (a psychological condition resulting from a traumatic event such as might be experienced during war or as a result of a violent attack). The personality factor of interest to the researchers was based on the modern version of Freudian psychoanalytic theory called "object relations." This refers to the psychological impact of our earliest relationships, mainly our parents (the "objects" of these early relationships). The researchers based their measure of object relations on a clinical interview focusing on such things as the ability to invest in a close relationship and the ability to see others in a complex way (e.g., not seeing a person as all good or all bad). In reporting their results, they abbreviated the object relations clinical interview measure as "OR-C." The distribution of scores on the OR-C was not normal (it was bimodal).

One of their analyses focused on the relation of object relations to whether a person stays in treatment to completion or terminates prematurely. They reported their results as follows:

> Six of the 74 participants prematurely terminated…. The six premature terminators did not differ from the rest of the sample on any demographic or pretest variable…. They did differ significantly from completers on OR-C ratings, scoring lower as tested by the nonparametric Mann-Whitney *U* Test ($Z = -3.34$, $p < .001$). (p. 554)

Explain, to a person who is familiar with the *t* test, but not familiar with rank-order tests, the general idea of what these researchers are doing (and why they didn't use an ordinary *t* test).

10. June and colleagues (1990) surveyed black students at a Midwestern university about problems in their use of college services. Surveys were conducted of about 250 students each time, at the end of the spring quarter over five different years. The researchers ranked the nine main problem areas for each of the years. One of their analyses then proceeded as follows: "A major question of interest was whether the ranking of most serious problems and use of services varied by years. Thus, a Kruskal-Wallis test instead of an ordinary analysis of variance (ANOVA) was performed on the rankings but was not significant…" (p. 180).

    Explain why the researchers used the Kruskal-Wallis test instead of an ordinary ANOVA. What conclusions can be drawn from this result?

Homework 10
Answers

Factorial ANOVA

1. **Identify the factorial ANOVA design**
   a. Factor 1: Program Status (Program or No Program)
   Factor 2: Sport (Baseball, Football, Basketball)
   DV: Number of injuries per player over 10 weeks
   The design is a 2 x 3 between-subjects factorial ANOVA .
   b. Create the matrix
   c. Fill in the matrix to show a main effect for type of sport.

		Sport		
		Baseball	Football	Basketball
Program Status	With Program	1	5	10
	Without Program	1	5	10

2. **Factorial Designs**
   a. Between-subjects factorial ANOVA
   Research design that manipulates factors between subjects; each subject is randomly assigned to only one condition. An example of a 2 x 2 between-subjects ANOVA could be a 2 (Test Difficulty: Hard, Easy) x 2 (Room Temperature: Hot, Cold) design, where each person is tested in only one condition.

		Test Difficulty	
		Hard	Easy
Room Temperature	Hot	Person 1	Person 2
	Cold	Person 3	Person 4

   b. Within-subjects factorial ANOVA
   Research design that manipulates only within-subjects factors (repeated measures).

   c. Mixed-Factorial ANOVA
   Research design that has at least one between-subjects factor and at least one within-subjects factor. An example of 2x2 mixed-factorial ANOVA is a 2 (Therapy Status: Therapy, Control) x 2 (Test Score: Time 1, Time 2) design, for which each person assigned randomly to type of therapy and is tested on two separate occasions.

		Personality Test Score	
		Time 1	Time 2
Therapy Status	Therapy	Person 1	Person 1
	Control	Person 2	Person 2

Chi-Square

3. **Chi-Square Goodness of Fit**

Categ	Observed	# in past (known data)	Expect. Prob. (based on *N*=500)	Expected (based on *N*=200)	$(O-E)$	$(O-E)^2$	$\frac{(O-E)^2}{E}$
1	38	100	100/500 = .20	(200*.20) = 40	38-40 = -2	4	4/40=.1
2	124	300	300/500 = .60	120	4	16	16/120=.133
3	22	50	50/500 = .10	20	2	4	4/120=.2
4	16	50	50/500 = .10	20	-4	16	16/20=.80
	Tot=200	Tot=500	Tot=500				

$$\chi^2 = \sum\frac{(O-E)^2}{E} = .1+.133+.2+.80=1.23$$

df = *Number of* categ - 1 ; *df* = 4 -1 = 3
At α = .05, with *df* = 3, χ^2 cutoff is 7.815
Null hypothesis: No significant difference between the observed and the expected values.
Alternative hypothesis: There is a significant difference between observed and expected values.

Because we **know** the frequencies based on past data, we should not compare all categories with equal expected frequencies. All **expected** frequencies per cell are **greater than 5**, so we have met that **assumption** of the test. However, we fail to reject the null hypotheses. Our frequency data (obtained) do not differ significantly from those data gathered in the past (expected); thus, the populations are the same.

Note: If expected probabilities were equivalent (e.g., all categories .25*200 = 50), however, you should find Category 2 frequencies occur more often than expected.

4. **Psychotherapy in different seasons (Chi-Square goodness-of-fit)**

Winter	Spring	Summer	Fall
28	33	16	51

Step 1.
Population I: Clients in psychotherapy clinic
Population II: Clients with no concern for season

H_0: The number of clients at any given time is independent of season
H_1: The number of clients at the clinic *depends* on (or is related to) the season (in other words, there is a relationship between the number of clients and season).

Step 2.

The comparison distribution is distributed according to χ^2 with 3 degrees of freedom (χ^2_{df} = *N*categ - 1 = 4-1= 3).

Step 3.

The critical cut-off score for χ^2 $\alpha = .05$, $df = 3$ is 7.815

Step 4.

Total # of patients = 28+33+16+51=128

Expected # of patients each season if H_0 is true = $\frac{N}{K} = \frac{128}{4} = 32$

$$\chi^2 = \frac{(28-32)^2}{32} + \frac{(33-32)^2}{32} + \frac{(16-32)^2}{32} + \frac{(51-32)^2}{32} = .5 + .03125 + 8 + 11.28 = 19.81$$

Step 5.

χ^2 of 19.81 is greater than the critical χ^2 of 7.815, so we *reject the null hypothesis* that the number of patients at any given time is independent of season. In other words, client visits are related to the season.

5. **Political Opinion Question**

   a.

   We analyze these data using the χ^2 test of independence because our data are nominal in scale and because we want to know if response frequencies are independent of the test categories. In the context of our problem, we want to test if a person's opinions are independent of the community where he or she lives. In other words, does individual opinion depend on community beliefs? Under the null hypothesis, we assume that opinion is unrelated to community.

b.

Observed Frequencies

	Community A	Community B	Community C	*Total:*
For	12	6	3	21
Against	18	3	15	36
No Opinion	12	9	12	33
Total:	42(47%)	18(20%)	30(33%)	N= 90

Expected Frequencies

	Community A	Community B	Community C
For	(.47)(12)= **9.8**	(.2)(6)=**4.2**	(.33)(3)=**6.83**
Against	(.47)(18)=**16.92**	(.2)(3)=**7.2**	(.33)(15)=**11.88**
No Opinion	(.47)(12)= **15.51**	(.2)(9)=**6.6**	(.33)(12)=**10.89**

$$\chi^2 = \frac{(12-9.8)^2}{9.8} + \frac{(18-16.92)^2}{16.92} + \frac{(12-15.51)^2}{15.51} + \frac{(6-4.2)^2}{4.2} + \frac{(3-7.2)^2}{7.2} + \frac{(9-6.6)^2}{6.6} + \frac{(3-6.93)^2}{6.93} + \frac{(15-11.88)^2}{11.88} + \frac{(12-10.98)^2}{10.98} = 8.58$$

$df = (N_{Colums}\text{-}1)(N_{Rows}\text{-}1) = (2)(2) = 4$; χ^2 cutoff(α=.05) = 9.488

χ^2 of 8.58 is less than the critical χ^2 of 9.488. We are unable to reject the Null Hypothesis that response and community are independent of each other. Thus, there is no evidence that there is a relationship between individual opinion and community type.

6. **Representative Community Question**

a.

We analyze these data using the χ^2 test for goodness-of-fit because our data are **nominal**, responses represent **frequencies** corresponding to categories, and we want to know whether our data are a good fit to the known population distribution.

b.

Category	Community	Observed	Expected
White	48%	300	(.48)(600)=288
African Am.	12%	80	(.12)(600)=72
Latinos	18%	100	(.18)(600)=108
Asians	9%	40	(.09)(600)=54
Others	13%	80	(.13)(600)=78
	Total:	600	

$$\chi^2 = \frac{(300-288)^2}{288} + \frac{(80-72)^2}{72} + \frac{(100-108)^2}{108} + \frac{(40-54)^2}{54} + \frac{(80-78)^2}{78} = 5.64$$

df = Ncateg-1 = 5-1 = 4; χ^2 cutoff(α=.05) = 9.488

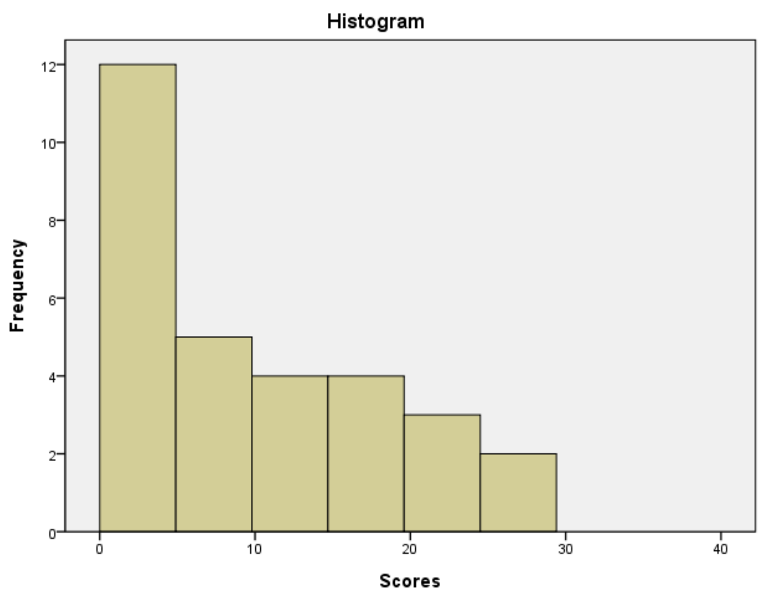
χ^2 of 5.64 is less than the critical χ^2 of 9.488. We are unable to reject the Null Hypothesis that the sample distribution and community distribution of ethnicity represent the same population. Hence, we conclude there is no evidence that the researcher's sample is not representative of the population from which she sampled.

Note: You can use the chi-square goodness of fit test to verify that your sample is representative of the population. This analysis is often helpful when you are collecting data that depend community distributions of race, gender, age, political beliefs, ethnicity, or other categorical variables that are important to the question that you are asking.

Data Transformations (When parametric assumptions might not be met)

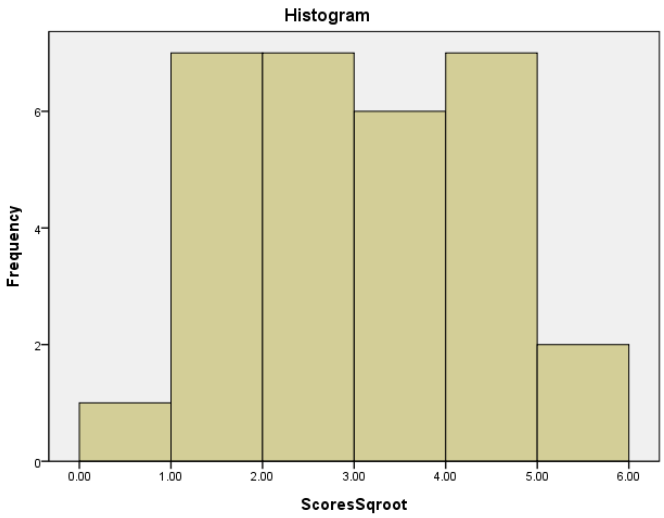
7. **Transform a Distribution**

Original distribution:



Graphing Tips:
Adjust X and Y labels to have the right scale of the graph.
Double click each axis to modify it → Editor Window → Scale → Enter scale as seen above.
Adjust bars to have correct binning.
Double click bars to open the editor window → Binning → X axis → Interval Width → Custom → Enter 4.9 (the recommended bin size in the problem).

Transformed distribution:



Transformation Tips:

You create a new variable in this transformation step (here it is labeled ScoresSqroot). To do the transformation click Transform → Compute Variable →type in a Target Variable Name (the variable you want to create) → select scores variable and move into the Numeric Experession field → indicate you want the new variable to be SQRT(scores) → OK.

8. Miller Data Transformation

Miller wanted to examine the relationships among the variables he was studying, probably including various parametric hypothesis-testing techniques such as the *t* test, analysis of variance, or testing the significance of bivariate or multiple correlation or regression results. Such procedures are based on the assumption that the distributions of the variables in the population follow a normal curve. However, Miller checked the distributions of the variables he was studying before he did parametric tests, and found that the scores on two key measures were skewed. This suggested that the population distributions for these variables probably violated the normal distribution assumption. Logarithmic transformations of the data made the sample distributions more normal. Thus, Miller could make the assumption that the population distributions of *logarithmically transformed* data were normally distributed, and he could then use parametric tests to analyze the data.

Nonparametric Tests (Understanding data in research reports)

9. Ford et al. and the Mann-Whitney U Test

The assumption of a normal population distribution was not met because the sample distribution was not normally distributed. Because the sample distribution was bimodal, a transformation would not help much. Hence, the researchers turned to a non-parametric rank-order test; "Mann-Whitney U-test." The Mann-Whitney U-test is based on the order (from smallest to largest value) of the observations and not on scale values. The number of ways we can *order* the observed data is calculated using permutations. One divided by the number of possible orders gives us the likelihood of observing any particular order of the data. Using this probability one can estimate the likelihood of achieving the observed order of the data or *more* *extreme* orders. If the probability (2-tailed or 1-tailed) is less than .05 we reject the null hypothesis that the average *rank* of scores in each condition is the same for each population.

10. June et al. and the Kruskal-Wallis Test

The rank-order non-parametric Kruskal-Wallis test was used because the data to be analyzed was ordinal and did not meet the assumptions of parametric tests. The statistic the researchers used is an extension of the Mann-Whitney U-test that can be applied to more than 2 groups. The researchers found that there was no significant difference between the medians of the group; the average rank of scores in each condition was too similar to reject the null hypothesis that they were the same in the populations.