

8.E: Chi Square (Exercises)

Q8.1

For each of the following questions, determine the appropriate test that should be used. Pick from the following three tests.

- A. Goodness of Fit
- B. Test for Independence
- C. Test for Homogeneity

- a. The tutor center maintains a list of student who use their services. These students are classified as drop-in students or appointment students. At the end of the term, the director of the tutor center randomly select students from each of the groups then looks up the grade they received in the class for which they were being tutored. The objective is to determine if there is a difference in the distribution of grades for the two groups. Grades are classified as A,B,C,F.
- b. Historically, a teacher found that 33% of the students in a class earned an A, 47% a B, 15% a C, and 5% a D or F. After modifying the way she teaches, she wants to know if her most recent class of students was consistent with past students.
- c. Students are given a math assessment and a musical assessment with the objective of determining if there is a correlation between mathematical ability and musical ability.
- d. Quantitative data are grouped by the number of standard deviations they are from the mean (e.g. z intervals of [-3, -2), [-2, -1),... [2, 3)). The objective is to determine if the distribution is normal and is based on the probability that a value would fall within each of those ranges.
- e. A researcher with the Department of Social and Health Services reviewed records of families who were receiving government assistance two years early. The researcher recorded if it was a one-parent household or a two-parent household. The researcher also recorded if the family was currently receiving government assistance. The objective was to determine if there is a correlation between then number of the parents in the household and whether the household was still receiving government assistance.
- f. A researcher with the Department of Social and Health Services wants to know if the number of parents in a household affects the length of time a family receives government assistance. The researcher identifies one-parent families and two-parent families then randomly selects from each of these two groups to determine the number of years in which they receive government assistance. A comparison will be made between the distribution of one-parent families and two-parent families.

Q8.2

For each of the following problems, identify the test that should be done, then write the hypotheses, conduct the test to find chi square and the p-value, and then write a concluding sentence.

1. Bunko is a dice game that serves as the motivation for a group of people to get together for an evening of socializing and eating. One regular Bunko player called it a mindless dice game, because it doesn't require much thinking and players can talk (or eat!) while playing. A normal game of Bunko involves 12 players, but other multiples of 4 can work nicely. If there are three tables, with the head table being number 1, then after each round of play, winning players move up one table with the goal of being at the head table (the one closest to the food!). The losers from the head table go to table 3. Three dice are used at each table. On each turn a player will roll all three dice. The first round the objective is to get ones, the second round the objective is to get twos, etc. If one of the desired numbers is obtained, the player gets a point. If two of the desired numbers are obtained on the same roll, the player gets 2 points. If all three of the dice are the desired number, the player yells bunko and gets 21 points. If none of the dice show the desired number, the dice are passed to the next person. When the head table reaches 21 points, the round is over for everyone.

As with any game of chance, there is an expected probability distribution. The expected distribution for the probability of having 0,1,2 or 3 successes can be found using the binomial distribution. Complete the probability distribution table. Refer to Chapter 4 if you can't remember the process.

$X = x$	0	1	2	3
$P(X = x)$				

Three dice were rolled 158 times and the number of ones was recorded for each turn. If the dice are fair, the sample distribution should be a good fit with the expected distribution. The sample data is shown in the table below.

$X = x$	0	1	2	3
Sample Results	100	44	13	1

Which test is appropriate for this problem?

Conduct the test then write a concluding sentence.

2. In Major League Soccer, is there a correlation between the number of shots a forward attempts and the number of goals he scores? A systematic random sample was taken from the 2013 MLS season results for all players classified as forward. The number of shots the player took was categorized as high (20 or more) and low (less than 20). The number of goals he scored was categorized as high (5 or more) and low (less than 5). The contingency table shows the results of the sample.

	High Shots	Low Shots	
High Goals	13	0	
Low Goals	15	21	

Which test is appropriate?

Complete the test using tables. Test the hypothesis at the 0.1 level of significance. Write a complete concluding sentence.

3. Lower back pain can be treated with a variety of approaches including using drugs and non-drug therapies. Data from a clinic that specializes in pain management was used to determine if there was a difference in the change in pain level for the patients being treated with a combination of drugs (local anesthetic, anti-inflammatory and a muscle relaxer) and those receiving physical therapy (lumbar traction, heat and ultrasound therapy and transcutaneous electrical nerve stimulation). Pain levels, on a scale of 1 – 5, were determined during the initial visit. The change in pain level was assessed at the 4- week period. If pain improved by 4 or 5 levels it was classified as substantial improvement. If pain improved by 1,2 or 3 levels it was classified as moderate improvement. If pain was unchanged or got worse, it was classified as no improvement. The table below shows the changes. Use this data to determine if there is a difference in pain reduction using drugs vs non-drug therapy. (data from unpublished student statistics class project)

Observed	Drug	Non-Drug	
Substantial improvement	9	6	
moderately improvement	22	23	
no improvement	9	11	

Which test is appropriate?

Complete the test using tables. Test the hypothesis at the 0.1 level of significance. Write a complete concluding sentence.

Expected	Drug	Non-Drug	
Substantial improvement			
moderately improvement			
no improvement			

Observed	Expected	$O - E$	$(O - E)^2$	$\frac{(O - E)^2}{E}$

				$\chi^2 =$
--	--	--	--	------------

4. Nationwide, for Native American tribal members with college degrees, 37% are associate degrees, 48% are bachelor degrees and 15% are Masters or PhDs. The distribution of degrees in one of the Puget Sound area tribes is 36 associate degrees, 22 bachelor degrees and 7 masters or PhDs. Is the distribution of degrees in the Puget Sound area tribe different than the national distribution?(data from unpublished student statistics class project)

Which test is appropriate?

Complete the test using tables or calculator. Test the hypothesis at the 0.1 level of significance. Write a complete concluding sentence. Show work (either tables or calculator inputs).

5. Why Statistical Reasoning Is Important for a Criminal Justice Student and Professional Developed in collaboration with Teresa Carlo, Professor of Criminal Justice

This topic is discussed in CJ 200 and others (Conflict view of Injustice).

The table below shows the racial distribution for Washington State. The data is from the WA State Government, Office of Financial Management. These percentages include those of Hispanic origin. (<http://www.ofm.wa.gov/pop/census2010/data.asp>)

White	Asian	Black	Native	Other
77.3%	7.2%	3.6%	1.5%	10.4%

In theory, the racial distribution of prisoners in WA state prisons should be consistent with this distribution. To determine if this is the case, a sample of prisoners can be taken. The random variable that will be measured is race. The hypotheses to be tested are:

H_0 : The racial distribution in WA prisons is the same as the racial distribution of the WA population

H_1 : The racial distribution in WA prisons is not the same as the racial distribution of the WA population.

Use a 5% level of significance. If the data are not significant then we will consider that society and justice are blind to race. If the data are significant, then we will seek a solution to this injustice.

There are 12 prison facilities in WA of which eight are major prisons and four are minimum- security. There is the possibility that the racial distribution varies based on location and security level and because of this random samples will be taken from each prison.

a. What type of sampling method is being used? _____

b. One prison has 2156 prisoners. If thirty prisoners will be selected from this prison, what are the first three random numbers that would be selected if the calculator were seeded with the number 12?

_____, _____, _____

Suppose the entire sample included prisoners from all the prisons. In total, 300 prisoners were selected. The number of prisoners of each race in this sample is shown in the table below. (This distribution is based on the actual distribution in WA prisons.) (www.doc.wa.gov/facilities/prison/)

White	Asian	Black	Native	Other
216	11	56	12	5

c. Which test is appropriate for this problem?

d. Make a double bar graph that shows a comparison between the observed and expected number of prisoners for each race.

e. Make a table to find the χ^2 value. Use the χ^2 table to find the p-value.

f. Write a concluding sentence.

g. Explain this conclusion in English. What do you think is the reason for this result?

h. If a solution is needed, what solution would you suggest?

This page titled [8.E: Chi Square \(Exercises\)](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Peter Kaslik](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.