Project Homework 5Solutions

STAT011 with Prof Suzy

Due by Tuesday, 12/5/23 11:59pm

Problem 1

Define your population of interest and define a relevant proportion related to this population. Conduct a hypothesis test about this unknown proportion using the data. (Don't repeat Problem 5 from PHW #4.) Make sure your answer includes:

1) the definition of your parameter and your significance level;

2) your null and alternative hypotheses;

3) an assessment of the required assumptions and conditions;

4) the calculated test statistic;

5) the p-value and conclusion of your test in context.

Solution to Problem 1

Answers will vary.

Problem 2

Define your population of interest and define a relevant population mean that you will estimate with a confidence interval. (Don't repeat Problem 1 from PHW #4.) Make sure your answer includes:

1) the definition of your parameter and your confidence level;

2) the sample mean;

3) an assessment of the required assumptions and conditions;

4) the critical value corresponding to your confidence level;

5) the standard error of your sample mean;

6) the lower and upper bounds of your interval interpreted in context.

Solution to Problem 2

Answers will vary.

Problem 3

Assess whether or not the problems listed below deal with count data for a categorical variable. (Hint: There are four problems that do not deal with count data.)

Scenario 1. A brokerage firm wants to see whether the type of account a customer has (Silver, Gold, or Platinum) affects the type of trades that customer makes (in person, by phone, or on the Internet). It collects a random sample of trades made for its customers over the past year and performs a test.

Scenario 2. The brokerage firm from (1) also wants to know if the type of account affects the size of the account (in dollars). It performs a test to see if the mean size of the account is the same for the three account types.

Scenario 3. The academic research office at a large community college wants to see whether the distribution of courses chosen (Humanities, Social Science, or Science) is different for its residential and nonresidential students. It assembles last semester's data and performs a test.

Scenario 4. A medical researcher wants to know if blood cholesterol level is related to heart disease. She examines a database of patients, testing whether the cholesterol level (in milligrams) is related to whether or not a person has heart disease.

Scenario 5. Is the quality of a car affected by what day it was built? A car manufacturer examines a random sample of the warranty claims filed over the past two years to test whether defects are randomly distributed across days of the workweek.

Scenario 6. A student wants to find out whether political leaning (liberal, moderate, or conservative) is related to choice of major. He surveys randomly chosen students and performs a test.

Scenario 7. A sales representative who is on the road visiting clients thinks that, on average, he drives the same distance each day of the week. He keeps track of his mileage for several weeks and discovers that he averages miles on Mondays, miles on Tuesdays, miles on Wednesdays, miles on Thursdays, and miles on Fridays. He wonders if this evidence contradicts his belief in a uniform distribution of miles across the days of the week.

Scenario 8. A study was performed examining epidurals as one factor that might inhibit successful breastfeeding of newborn babies. Suppose a broader study included several additional issues, including whether the mother drank alcohol, whether this was a first child, and whether the parents occasionally supplemented breastfeeding with bottled formula.

Scenario 9. Two different professors teach an introductory statistics course. The table shows the distribution of final grades they reported. We wonder whether one of these professors is an "easier" grader.

	Prof A	Prof B
A	3	9
B	11	12
C	14	8
D	9	2
F	3	1

Scenario 10. A student studying the music preferences of her classmates wants to test her theory that men and women have different music preferences. For each respondent, she records their gender and what their favorite music genre is (hip hop, trap, rock, rap, folk, indie, other).

Solution to Problem 3

The following scenarios all deal with count data: 1, 3, 5, 6, 8, 9, 10

Note for graders: As long as a student states at least six of the scenarios in the solution, please give them full credit and grade Problems 4 and 5 based on their answer. (There was a typo in the instructions that stated only six scenarios were count data.)

Problem 4

For the six scenarios in Problem 3 that do deal with count data, determine which **(if any)** chi-square test is appropriate for each.

Solution to Problem 4

Scenario 1 - test of independence

Scenario 3 - test of homogeneity

Scenario 5 - goodness of fit

Scenario 6 - test of independence

Scenario 8 - none of the three chi-square procedures are appropriate

Scenario 9 - test of homogeneity OR none of the three chi-square procedures are appropriate

Scenario 10 - test of homogeneity OR test of independence

Problem 5

State the null and alternative hypotheses for each of the scenarios in problem 4.

Solution to Problem 5

Scenario 1 - H_0 : Account type is independent of trade type; H_A : Account type is not independent of trade type

Scenario 3 - H_0 : The distribution of course type is the same for the residential population as it is for the nonresidential population; H_A : The distribution of course type for the residential population is different from the distribution of course type for the nonresidential population

(The null could be stated symbolically, e.g.: H_0 : $p_{humanities,res} = p_{humanities,nonres}, p_{social,res} = p_{social,nonres}, p_{sci,res} = p_{sci,nonres}$)

Scenario 5 - Let p be the probability that a car is defective. Then $H_0 : p_{Monday} = p_{Tuesday} = p_{Wednesday} = p_{Thursday} = p_{Friday} = 1/5$ and H_A : The probability a car is defective is different depending on which day of the week the car was built.

Scenario 6 - H_0 : Political alignment is independent of major; H_A : Political alignment is not independent of major

Scenario 9 - H_0 : The distribution of grades for Prof A matches the distribution of grades for Prof B; H_A : The distribution of grades for Prof A is different from the distribution of grades for Prof B

Scenario 10 - H_0 : Music preference is independent of gender; H_A : Music preference is not independent of gender OR H_0 : The distribution of preferred music genre is the same for men as it is for women; H_A : The distribution of preferred music of men is different from the distribution of preferred music of women.