

STAT 477/577 - Homework Assignment 5

Due Thursday, March 24, by 11:59 PM

General homework guidelines: All homework assignments should be submitted using Canvas. Please submit your answers separately, for each of the problems, as set-up in the Canvas submission portal. You are allowed to either type in your answers, as well as submit graphs directly, within each submission item. You are also allowed to submit a scanned copy of your answer, as long as the answers are submitted separately for each question, as instructed. You can either scan your answers (ask if you don't have access to a scanner or do not know how to use your phone to do so), or just submit a picture of your answer. Please note that if we can't read your answer, we won't be able to award any (partial) credit.

You have one attempt to submit your answers. If technical issues appear and your submission portal has closed for some reason, please email Prof. Caragea explaining the situation and requesting permission to resubmit. Please note that such requests must be made before the deadline.

Please note that for this homework only I have changed the due date from Tuesday to Thursday, and there is no grace period for late submissions with penalty. The due date of Thursday, March 24 is firm and final. No submissions will be accepted past this date.

Homework problems.

1. [6 pts.] The following table of counts was obtained from a random sample of 1397 respondents from the population of adults (more than 18 years old) in the United States in 1982. Each respondent was cross-classified with respect to opinions regarding gun registration as a part of comprehensive gun control legislation and imposing the death penalty on adults convicted of certain violent acts.

Gun Registration	Death Penalty		Total
	Favor	Oppose	
Favor	784	236	1020
Oppose	311	66	377
Total	1095	302	1397

- (a) Conduct a test of independence for the two variables. You may use R to answer this question, as long as you clearly explain/show how you obtained your results.
 - (b) Calculate the correlation (ϕ) coefficient for the two variables.
2. [9 pts.] The operations manager of a company that manufactures tires wants to determine whether there are any differences in the quality of workmanship among the three daily shifts. She randomly selects 496 tires and carefully inspects them. Each tire is either classified as perfect or non-perfect, and the shift that produced it is also recorded. The two categorical variables of interest are: shift and condition of the tire produced. The data can be found in the file **tires.csv**.
 - (a) Explain why it is reasonable to conduct a test of independence for these two variables even though the shift variable is a grouping variable.
 - (b) Describe a different data collection that would have required the use of a test for the equality of proportions.
 - (c) Use R to create a contingency table and mosaic plot for the two variables and write a description of their relationship.
 - (d) Conduct a test of independence for the two variables. Clearly write the hypotheses, the value of the test statistics and the p -value, along with your conclusions.
 - (e) Calculate the ϕ coefficient and Cramer's V for this contingency table. You may do this by hand or using R.

3. [10 pts.] In the 2002 General Social Survey of adults in the United States conducted by the National Opinion Research Center, respondents were asked about their perceived happiness (Not Too Happy, Pretty Happy, Very Happy) and their family income level (Below Average, Average, Above Average). The data can be found in the file **norc.csv**.
- (a) Use R to calculate a contingency table for the two variables.
 - (b) Calculate a test of independence for the two variables.
 - (c) Which two cells are concordant cells with the cell Very Happy - Average Income?
 - (d) Which two cells are discordant cells with the cell Pretty Happy - Below Average Income?
 - (e) Calculate the Goodman-Kruskal Gamma statistic for these data to determine the strength of the directional relationship between the two variables and report its confidence interval. What does this tell you about the directional relationship between perceived happiness and income of adults in the United States?