

Homework 3
Due : October 14

Problem 1 (20pts)

Download dataset data1D2D2017.xlsx from my webpage.

1. Test the hypothesis that mean spending is the same for men and for women.
2. Consider the following spending categories: < 100 , $[100, 200)$, $[200, 300)$, $[300 - 400)$, ≥ 400 . Create a table that shows the number of observations in each of those spending categories for men and for women.
3. Use that table to perform a Chi2 test of the hypothesis that spending is independent of gender.

Problem 2 (10pts)

Consider the following data for spending for a representative sample of customers in a particular market during a recent period.

<i>Spending</i>	<i>Female</i>	<i>Male</i>
$< 50K$	700	601
$[50K - 100K)$	513	557
$[100K - 200K)$	410	518
$\geq 200K$	227	309

Use these data and a chi-squared test to test the hypothesis that spending is independent of gender in this market.

With how much confidence can you reject the hypothesis that spending is independent of gender based on the table above?

Problem 3 (10pts)

Consider the following probability space and random variable.

S	s_1	s_2	s_3
p	0.2	0.3	0.5
X	100	110	90

You are given 200 draws supposedly from the X shown in problem 1 above where the number of 100s is 26, the number of 110s is 68 and the remaining 106 draws are 90s. Based on a Chisquare goodness of fit test, can you reject the hypothesis that these data came from X with 95% confidence?

Problem 4 (20 pts)

Download dataset data4D2D.xlsx from my webpage.

1. Regress Admit on GPA and GRE using a traditional regression. Produce probabilities of admission for each observation using the resulting linear probability model.
2. A probit regression of Admit on the variables produced the following coefficients:

<i>Variable</i>	<i>Coefficient</i>
<i>Constant</i>	-3.003536
<i>GPA</i>	0.454575
<i>GRE</i>	0.0016425

use this model to forecast probabilities of admission for each observation.

3. Plot the two forecast probabilities against each other, fit a line through the resulting cloud, show the line equation and the R^2 .
4. Produce a gain chart that compares the performance of the probit model above to a naive model.

Problem 5 (20 pts)

Using the exact same data and probit model as in problem 1:

1. Use the probit output to estimate the likelihood that a student whose GPA is 3.60 and whose GRE is 660 will get admitted.
2. Given the same data and model, what is the likelihood of observing exactly one admission among two students both of whose GPA is 3.60 and GRE is 660? Two admissions? Zero?
3. The same student retakes her GRE and improves her score to a 700. By how much have her chances of admission increased according to the Probit model?