

Exam 3 for MGMT 230: Introductory Statistics

Name (20 points)

For every problem, clearly state if any further assumptions need to be made to solve the problem, and if so, what assumption. To show your work, whenever you use a formula to answer a question, write down the formula first, then write down the formula with the correct numbers plugged in.

For every hypothesis test, conduct every step of a hypothesis test, and also report and interpret the p-value for every test. Use any significance level you think is appropriate.

Rules: You must not work on this exam nor discuss this exam with any other person.

1. Our automobile insurer is again interested in determining whether there is a relationship between the distance someone lives from where they work and the amount the insurance company must pay in claims are related. To answer this question, they gather data on 564 individual policy holders that have lived and worked at the same place for five years. Over these five years they find out how much has been paid out in insurance claims to each person (i.e. how much money the insurance company has paid out as a result of automobile accidents), and the distance, in miles, each person lives from their workplace. They find the following information:
 - The average amount paid out in insurance claims is \$342.56.
 - The sample standard deviation of insurance claims payouts is \$32.12.
 - The average distance people live from their workplace is 3.4 miles.
 - The sample standard deviation of the distance people live from their workplace is 4.4 miles.
 - The covariance of insurance claims paid out and distance from workspace is 352.89.
- (a) (10 points) Test the hypothesis that insurance claims payouts is positively (linearly) correlated with distances from the workplace.
- (b) Compute the linear regression (using ordinary least squares method - the only method you know) that explains the relationship between distance and insurance claims, and answer the following questions:
 - i. (5 points) Which variable is the dependent variable and which is independent? Explain.
 - ii. (10 points) Compute and report the regression equation.
 - iii. (5 points) For every one mile farther away someone lives from their workplace, how much does the predicted amount of insurance claims increase or decrease by?
 - iv. (5 points) Suppose Joe Schmo lives 12.3 miles away from his workplace. What is the expected amount in benefits that he is likely to claim in a five year period?

2. Download the Excel sheet from the class website named `exam3.xls`. The data given is a sample of data from February 15th through April 30th 1993 from the Albuquerque Board of Realtors. The variables are the selling price of houses (in hundreds of dollars), the size of the house (in square feet), the age of the house (in years), the number of features out of a possible 11 the house includes, and a 1 or 0 indicating whether the house is on a corner lot.
- (a) (10 points) Run a regression (attach the regression output) that attempts to explain the price of houses using the size of the house, the age of the house, the number of features, and whether or not the house is on a corner lot as the explanatory variables. Write down your regression equation line.
 - (b) (5 points) What does your regression predict is the effect on the price of a house from living on a corner lot?
 - (c) (5 points) Using your regression results, test the hypothesis that increasing the number of features increases the price of a house.
 - (d) (5 points) Construct and interpret a 95% confidence interval for the effect an additional year of age has on the price of a house.
 - (e) (5 points) What percentage of the variability in selling price is explained by the given factors (the size of the house, the age of the house, the number of features, and whether the house is on a corner lot)?
3. A sample of 78 gasoline purchases yields an estimate for the price elasticity of demand for gasoline equal to 0.075. The sampling distribution of this estimate is normally distributed, unbiased, and the standard error is estimated to equal 0.07 (which has 43 degrees of freedom). Let the sample elasticity be denoted by $e = 0.075$, the estimated standard error be denoted by s_e , and the population elasticity be denoted by ϵ .
- (a) (10 points) Test the hypothesis that gasoline has an inelastic demand (inelastic demand means the price elasticity of demand is less than 1).
 - (b) (5 points) Construct and interpret a 95% confidence interval for the price elasticity of demand.