DECS 430-A: Business Analytics I

Sample Waiver Exam

1) This exam is not to be copied or circulated by anyone except the exam administrator.

2) Materials allowed: Excel, and reference materials such as books, etc. You will need Excel (or an alternative product that does normal distribution calculations) though there is no need to manipulate data files.

3) The exam consists of 10 questions.

4) The time allowed for the exam is 3 hours.

5) The Kellogg Honor Code applies. In particular, the exam must be solely your own work.

GOOD LUCK!

Question 1 (5 points)

Let A and B be two events. Suppose P(A) = 0.2, P(B) = 0.3 and $P(A \cap B) = 0.1$.

a) Compute $P(A \cup B)$.

b) Are A and B independent events? Explain.

c) Compute P(A | B).

Question 2 (5 points)

The finance department of a particular company has 15 employees, 6 of whom have MBA degrees. Suppose we select three different employees sequentially at random from this department. Determine the probability of the following events.
a) The first employee has an MBA, given that there is a total of one MBA among all three employees.
b) There are exactly two employees with an MBA, given that the first employee has an MBA.
c) The first employee has an MBA, given that there is at least one MBA among all three employees.

Question 3 (5 points)

A stock has a current share price of $100/share. Let X represent the dollar amount by which its price will change over the next one week. Assume X is normally distributed, with E(X)=0.10, and standard deviation of 0.50.
a) What is the probability that after one week, the stock's value will be between $99.50 and $101.00?
b) Now examine the price of this stock one year from now (52 weeks). Suppose that each weekly price change of the stock is described by this distribution of X. In addition, suppose weekly
price changes are independent. Find the probability that after one year, the stock's value will be between $99 and $111.

Question 4 (10 points)

A fair coin is tossed 5 times. Each toss is independent of the other.
a) What is the probability of getting exactly 3 heads?
b) What is the probability of getting at least one heads?
c) What is the expected number of heads?
d) What is the variance of the number of heads?

Question 5 (5 points)

A major credit rating agency rates a bond of a company as CCC if the probability of default is 0.10 in any given year and defaults across years are independent.

a) What is the probability that a CCC bond defaults in the 4th year (that is, the company goes for 3 years without default, then defaults in the fourth)?
b) Suppose you own a pool of 20 CCC bonds from 20 different companies. Assume the default risks for different companies are independent. What is the probability that between 1 and 3 of these CCC bonds (that is either one or two or three CCC bonds) default in the first year?

Question 6 (10 points)

Unoccupied seats on flights cause airlines to lose revenues. A large airline wants to estimate its average number of unoccupied seats per flight over the past year. To accomplish this, the records of 225 flights are randomly selected, and the number of unoccupied seats is noted for each of the flights in the sample. The sample mean is 14.5 seats and the sample standard deviation is $s = 8.2$ seats.

a) Provide an 80% confidence interval for the mean number of unoccupied seats per flight during the past year.

b) Can you demonstrate, at a 2% level of significance, that the average number of unoccupied seats per flight during the last year was smaller than 15.5.

Question 7 (10 points)

With the plan to develop a chain of boutique pet supply stores, you have obtained data on pet ownership among residents of Chicago's Gold Coast. You are interested in the percentage of households in that area that own at least one pet. Within a random sample of 300 households, 30 of the households own exactly one pet, and 12 households own two or more pets. The rest own none.

a) What is your estimate of the percentage of Gold Coast households that own at least one pet? (2 points)
b) What is the 90% confidence interval for this estimate? (Note: either provide the actual interval or (if you are not using Excel) rigorously and completely explain how you would obtain all of the numerical values needed to compute the interval. (8 points)

Question 8 (5 points)

A certain type of cold vaccine is known to be only 25% effective after a period of 2 years. That is there is a 25% probability that someone who takes the vaccine will be free of the cold virus after two years.

To determine if a new, more expensive vaccine is superior in providing protection against the cold virus for a longer period of time, 200 people are chosen at random and inoculated with the new vaccine. If fewer than 120 of those receiving the new vaccine contract the virus within a two year period, the new vaccine will be considered superior to the present one.

a) Suppose the new vaccine has exactly the same effectiveness as the present one. What is the chance that the outcome of the test with the 200 subjects will conclude that the new vaccine is superior to the present one?

b) Suppose the new vaccine is in fact twice as effective as the present one. That is 50% of those who receive the vaccine are protected for 2 years or more. What is the probability that our test will declare the new vaccine NOT superior to the present one?

Question 9 (10 points)

A company sells an insurance policy against losses incurred during travel for $15. 40,000 customers buy the insurance policy. The policy pays the full amount claimed up till $5,000. Most policy holders claim nothing, the rest either $500 or $5,000.

Each customer has the following probabilities of claiming a certain amount for damages, independent of the other customers.

Claim amount	$0	$500	$5,000
Probability	99%	0.8%	0.2%

Use the central limit theorem to estimate the probability that the company makes a positive net profit on these contracts.

Question 10 (10 points)

A survey is carried out by the Chicago Public School District to determine the distribution of household size. They draw a simple random sample of 1,000 households and send interviewers to each household to collect information. After several visits, the interviewers find people at home in only 691 of the sample households. Rather than face such a high non-response rate, the District draws a second batch of households, and uses the first 309 interviews in the second batch to bring the sample up to its planned size of 1,000 households. The District counts 3,121 people in these 1,000 households, and estimates the average household size in the District to be about 3.1 persons.

Is this estimate likely to be too low, too high, or about right? Why?

Department of Managerial Economics and Decision Sciences

DECS 430-A: Business Analytics I
Sample Waiver solutions

(Note: some solutions simply provide an Excel formula that yields the correct answer.)

1a) 0.2 + 0.3 – 0.1 = 0.4

1b) No, because $P(A \cap B)$ is not the same as P(A)*P(B)

1c) 1/3

2a) 1/3

2b) 45/91

2c) 78/159

3a) 0.85

3b) 0.9

4a) BINOM.DIST(3,5,0.5, FALSE)

4b) 1- (0.5)^5

4c) 2.5

4d) 5/4

5a) 0.9*0.9*0.9*0.1

5b) BINOM.DIST(3,20,0.1, TRUE) – BINOM.DIST(0,20,0.1, TRUE)

6a) In Excel: 14.5 +/- T.INV.2T(0.2 , 224)* 8.2/sqrt(225) , or roughly [13.8 , 15.2]

6b) The test statistic is t = (15.5-14.5) / (8.2/SQRT(225)) = 1.83

The p-value is T.DIST.RT(1.82 , 224) = .035 > .02 so the answer is NO, we cannot reject the null hypothesis that the average is more than 15.5.

7a) 42/300 = 14%

7b) 0.14 +/- 0.03295

8a) Either BINOM.DIST(120, 200, 0.75, TRUE) or BINOM.DIST(119, 200, 0.75, TRUE) is acceptable.
8b) Either 1 – BINOM.DIST(119, 200, 0.5, TRUE) or 1 – BINOM.DIST(120, 200, 0.5, TRUE) is acceptable.

9) If T is total profit then the expected value of T is $40,000. The standard deviation of T is 45,600. By the Central Limit Theorem, T will be approximately normally distributed (since it is the sum of independent random variables). Then, Pr(T > 0) will be
1- NORM.DIST(0, 40,000, 45,600, TRUE) = 0.81.

10) (A full credit answer should be of similar quality to the following.)
There is a strong potential for sample bias in this case. It would depend on whether there is a correlation between the size of a household and the likelihood that someone is home to answer the door when an interviewer arrives. It may be reasonable to expect that these two things are positively correlated: larger households are more likely to have at least one household member at home, and households with children could also be more likely to be at home. By dismissing the observations for which no one is home, the School District may be underrepresenting households with just one or two members. In this case, the District's household size estimate likely would be too high.