Petter Mostad
Mathematical Statistics
Chalmers

MSA830 Statistical analysis and experimental design

Reexam 9 June 2011, 8:30 - 13:30
Examiner: Petter Mostad, phone 0707163235,
visits the exam at 9.30 and at 11.30.
Allowed to use during the exam: Pocket calculator, books, copies, and notes.
Number of points on the exam: 30. To pass the exam, at least 12 points are needed

1. Esteban needs to complete a large number of similar experiments as part of his research. Each experiment can be of type X, Y, or Z, and each experiment can be either a success or a failure. The probabilities of success are 0.7, 0.4, and 0.4 for experiments of types X, Y, and Z, respectively. We assume that all successes or failures of experiments happen independently.

   (a) One evening, Esteban decides to do one experiment of type X, one of type Y, and one of type Z. What is the probability that at least one of his experiments is a success? (1 point)

   (b) The next evening, Esteban has time for only one experiment, so he writes X, Y, and Z on three pieces of paper, randomly draws one of them, and then performes an experiment of this type. The experiment is a success. Given this information, what is the probability that the experiment was of type X? (2 points)

2. Eliza is studying a particular type of tornadoes. In her study area, there has been an average of 5.6 of these tornadoes each year for the 40 years up to and including 2009. In parts (a) and (b) below, we assume that the rate of tornadoes is stable from year to year, and that each tornado appears independently from other tornadoes.

   (a) What was the probability, at the beginning of 2010, of observing zero or one tornadoes during 2010? (2 points)

   (b) Assume that, in fact 12 tornadoes were observed during 2010. Compute the approximate probability, at the beginning of 2010, of observing 12 or more tornadoes during 2010. (2 points)

   (c) Eliza main interest is studying to what extent tornadoes happen independently from each other and independently from other trends in the weather or climate. Formulate a conclusion Eliza could draw based on the information above. (1 point)

3. Emmy compares two different types of grass seeds: She wants to find out which type is best suited for the soil and growing conditions where she lives. She finds 5 different locations of equal size, and at each location, she plants half the area with grass seed A and half with grass seed B. After 7 weeks, she measures the amount of grass cut at each location and from each grass type, getting the results in the following table:

	Location 1	Location 2	Location 3	Location 4	Location 5
Seed A	13	15	13	10	20
Seed B	13	11	12	10	13

(a) Making some assumptions, find the expected difference in amount of grass cut after 7 weeks when using seed A compared to using seed B at a new location. Find also a 95% credibility interval for this expected difference. (3 points)

(b) What assumptions did you make in part (a) to derive at your conclusions? (1 point)

(c) Emmy is still not sure which grass type is best, so she decides to extend her experiment with 15 new locations, so that she gets a total of 20 locations. Roughly what length of the 95% credibility interval should she expect after analyzing the new data? Why? (1 point)

4. Emmett is trying to optimize his recipe for muffins, by varying the recipe and asking groups of people to score the result. In particular, he is studying how the score depends on the amount of sugar. His results are :

Sugar	0	0.3	0.6	1	1.5	2
Score	2.1	3.4	3.5	3.2	2.7	1.1

Emmett would like to analyze the results using a linear model by fitting a second degree polynomial to the data. Write down the design matrix Emmett should use. (2 points)

5. Evelyn is comparing a property of two different materials, X and Y. For X, she has 7 measurements, and she has computed an average 11.34 and a sample variance 0.0235. For Y, she has 11 measurements, and she has computed an average 11.52 and a sample variance 0.0921.

(a) Find a 95% credibility interval for the precision of the observations for the property for material X. (1 point)

(b) Make a hypothesis test of whether the two population variances are equal, for the measurements for materials X and Y. (1 point)

6. Pernilla is planning an experiment over 8 days where she will vary a number of factors to see the effect on the daily income from sales of ice cream from her ice cream stand. The factors she would like to vary are:

- Displaying a billboard with a picture of an ice cream or not.
- Cutting the prices below those of a nearby shop, or having the same prices.
- Playing some music.
- Calling out to passing people or not.

(a) Propose a fractional factorial design for Pernillas experiment. (2 points)

(b) Give advice to Pernilla about how she should perform her experiment in practice for the conclusions to be as valid as possible. (2 points)

7. Christer is growing a cell culture as part of his research, and he wants to optimize the growth rate. He is investigating how type of nutrient, temperature, and Ph value is influencing the growth rate in an experiment where he uses three types of nutrients (A, B, or C), two temperatures (high or low) and two Ph values (high or low). For each possible

combination of these factors he does two experiments, resulting in a total of 24 measurements of growth rates. The results are given in the table below. For your convenience, the averages for each type of nutrient and for each temperature are given in the table, together with the grand average. The average for observations with high Ph values is 6.725 and the average for observations with low Ph values is 7.3333. The variance of all the growth rates is 2.557808.

	Temp high		Temp low		Average
	Ph high	Ph low	Ph high	Ph low	
Nutrient A	4.0 7.2	6.1 8.7	8.1 7.4	6.9 7.9	7.0375
Nutrient B	7.1 7.0	6.1 7.1	8.4 8.4	10.9 8.4	7.925
Nutrient C	3.5 5.0	5.3 7.0	6.6 8.0	7.7 5.9	6.125
Average	6.175		7.8833		7.0292

(a) Set up an ANOVA table for these observations; do not include interactions. Compute intervals for p-values. (4 points)

(b) What assumptions do you need to make before you can make interpretations of the p-values? (1 point)

(c) Making these assumptions, formulate the conclusions you can make, and the recommendations for Christer. (1 point)

(d) Assume the Temperature was disregarded as a predictive variable, so that only Ph and Nutrient was used as predictors. Make an ANOVA table for this analysis. Include intervals for p-values. (2 points)

(e) Go back to the original ANOVA table (from part (a)) and assume Christer would like to change it to include the interaction effect between Temperature and Nutrient. Make an outline for this table: You do not have to compute the sum of squares for the interaction, but you should write down the table with all the headings and line names, and include all the numbers that can be computed directly from the numbers in your ANOVA table for part (a). (1 point)