MATHEMATICS 150

(2001-2002)

PROBLEM SET 8

Solutions may be submitted in class on **Monday, March 25** or may be submitted to the Instructor's office (CC F30) **by 3:00 pm Tuesday, March 26**. Problem 7 is for practice; solutions for problem 7 are not to be submitted for marking. ***Print your name on the upper right-hand corner of the front page.***

1. Is there sufficient evidence at the 5% level of significance that over 25% of consumers will stop using a product because of one bad experience if 63 of 200 consumers in random survey said that they would stop using a product because of one bad experience? What is the empirical level of significance?

2. The data below are data for weeks 1 and 2 for the glucosamine capsule data of Problem Set 7 problem 5. If you wish to use computer software to work with these data, you may obtain these data electronically in the usual way from

http://www.trentu.ca/academic/math/courses/stat/files

The data sets for the individual weeks are in data files capsule1.dat, and capsule2.dat.

Sample capsule glucosamine amounts (mg)

week 1	week 2
499.4	498.9
500.1	500.6
498.9	499.6
499.9	499.9
500.4	500.3
502.2	499.9
502.4	499.6
499.0	499.0
498.9	499.2
501.6	499.7
501.7	501.4
500.8	499.7
	500.0
	500.9
	499.7
	499.8
	499.4
	500.6

a) Test the null hypothesis of equal process (population) variances for weeks 1 and 2 using a 5% level of significance.
b) Use the conclusion from a) to determine the appropriate procedure for comparing means and determine 95% confidence limits for the difference in the process (population) means for weeks 1 and 2.
c) Does the interval determined in b) provide sufficient evidence at the 5% level of significance that the process (population) means differed for weeks 1 and 2?

3. Consider a study to investigate the effects of different structures for on-line courses to learn specific computer software applications with the following setup and results. Each of 15 test subjects followed two courses. The courses were for two different applications judged to be of equal difficulty. Approximately half the subjects followed the first course then, after a delay to reduce learning carry-over, followed the second; the remaining subjects followed the second course then, after a delay to reduce learning carry-over, the first. Within each of these arrangements approximately half the presentations of the first course were with a current on-line course structure and the other half were with a proposed new structure. The presentations were arranged so that each subject used each structure once. Each subject followed the course with on-line feedback until each step of the course had been completed successfully. An internal log kept a record of the times required to complete the course successfully. With the balancing of first and second courses and the use of each structure for approximately half of each arrangement, it was assumed that any differences in times could be attributed to differences in the course structures. The resulting times are listed overleaf. Determine whether the data provide evidence at the 5% level of significance that the new structure reduces the learning time by over 15 minutes "on the average". ***If*** there is sufficient evidence, indicate the *P*-value.

OVER

3. c) (Continued)

Sample On-Line Course Times

subject	current	new
1	49.3	26.3
2	56.5	31.6
3	42.8	26.3
4	67.1	34.2
5	35.7	19.7
6	44.0	24.2
7	38.6	27.4
8	30.7	22.0
9	59.8	33.1
10	45.6	31.8
11	33.6	23.1
12	78.3	38.3
13	46.1	24.7
14	55.9	36.9
15	59.8	25.9

4. In a study on a high risk population, a control group of 200 individuals included 87 who had experienced a disease being investigated. A corresponding experimental group of 100 individuals who had been given a proposed preventive treatment included 29 who subsequently experienced the disease. Determine a 95% confidence limit to indicate at least how much of an improvement the proposed treatment produces in the percentage of the high risk population experiencing the disease.

5. In a survey on attitudes towards group homes, 250 people in a large population were asked whether they would support or oppose the establishment of a group home in their neighbourhood. (A non-preference was not accepted as a survey response.) The results were: 90 supported a group home and 160 opposed a group home. After a series of debates on the issue, these same 250 people were surveyed again. Of the original 160 opponents, 19 changed to supporters. Of the original 90 supporters, 11 became opponents. Is there evidence that the proportion of supporters increased in the population as a whole? If so, indicate the *P*-value.

6. (NOTE: This is an extension of Problem Set 3 problem 1.) It is difficult and expensive to classify accurately subjects who should be excluded from a particular experimental project. For any subject, a classification procedure could produce a clear 'suitable' classification (correctly or not), an unclear result, or a clear 'unsuitable' classification (correctly or not). Four relatively inexpensive preliminary classification procedures are to be compared for accuracy on the basis of sample testing of 400 subjects *known to be unsuitable*. Procedure A was used with 150 of the subjects; procedures B, C and D were used with 100, 75 and 75 of the subjects, respectively. The procedures were assigned randomly to subjects until the numbers above were obtained. The resulting classifications were stored in a data file **classify.dat** which is available in the usual way. Each line of the file contains the procedure used — A, B, C or D — in column 2 and the resulting classification suitable, unclear or unsuitable, left justified in columns 4 through 13. You must use an appropriate FORMAT subcommand with the READ command in MINITAB to enter the data.

a) Display these data with a cross-tabulation display with appropriate detail. Arrange the display so that the classifications are in rows and the procedures are in columns.

b) Do the procedures differ at the 5% level of significance with regard to classification accuracy.

7. In an analysis of field equipment battery lifetimes under one set of environmental operating conditions, six batteries of each of three types (one type from group A and three types from group B) were monitored for useful life. The resulting lifetimes were as listed below

Battery Type

A	B1	B2	B3
443	565	528	562
459	556	542	580
477	579	557	597
482	600	566	617
515	680	615	632
542	718	623	695

a) Do these results indicate that the mean battery lifetimes are not all the same for the four types of batteries?

b) Which battery types differ with regard to mean battery life?