

STATISTICS 4C03 Assignment # 1

Due 2006-01-09 13:30

The following questions are from the final exam in an introductory course. Calculate the answer to each question three ways: (1) using a calculator and tables; (2) using Splus or R; (3) using SAS.

- [3] 1. (a) Name the statisticians who (i) pioneered medical statistics during the Crimean War; (ii) produced statistical tables during WW I; (iii) developed modern exploratory data analysis.
- [1] (b) What Latin dance is etymologically related to the contingency table?
- [10] (c) The mean drying time of a certain paint is known to be 12 min. From past experience, you know that the time varies with a standard deviation of 2 min. An additive is supposed to reduce the mean drying time to 11.5 min. If you want the Type I and Type II error rates both to be 5%, how many paint samples would you have to try to test this claim? If you could only afford to test 50 paint samples, what would be your probability of Type II error? Would the test still be worth doing?
- [11] (d) A company produces 20% of the windshields for a certain model of car at Plant A and 80% at Plant B. The mean number of flaws (small bubbles) per windshield is 2.1 at Plant A and 4.3 at Plant B. If a given windshield has 3 flaws, what is the probability that it was produced at Plant A? State any assumptions you make.
- [25] 2. Carry out appropriate analyses for the following two data sets. Give graphs. State any assumptions you make. As far as possible, test each assumption. State your conclusions.
- (a) Two microprocessors were compared on a sample of six benchmark codes to determine if there is a difference in speed. Here are the times (in sec) used by each processor on each code.

Code:	1	2	3	4	5	6
Processor A:	27.2	18.1	27.2	19.7	24.5	22.1
Processor B:	24.1	19.3	26.8	20.1	27.6	29.8

- (b) The data below give the resilient modulus at 40° C (in 10⁶ kPa) in 7 sections of rutted pavement and 12 sections of non-rutted pavement.

Rutted:	1.48	1.88	1.90	1.29	3.53	2.43	1.00					
Nonrutted:	3.06	2.58	1.70	5.76	2.44	2.03	1.76	4.63	2.86	2.82	1.04	5.92

- [25] 3. Carry out appropriate analyses for the following two data sets. State any assumptions you make. State your conclusions.

(a) A machine shop wants to compare the quality of work of its 4 machinists. A sample of 300 machined parts was cross-classified according to whether a part was defective or not, and according to which machinist did the work.

Machinist:	A	B	C	D
Number of Defective Parts:	18	7	12	6
Number of Nondefective Parts:	62	81	58	56

(b) An experiment was conducted to determine the effect of sintering time on the compressive strength of two metals. The results are shown in the following table..

		Sintering time	
		100 min	200 min
Metal	1	17.1, 16.5, 14.9	19.4, 18.9, 20.1
	2	12.3, 13.8, 10.8	15.6, 17.2, 16.7

- [25] 4. The following data give the load (in lb/ft) at which the first crack in a concrete pipe specimen was observed, and the age (in days) of the specimen. Does load at first crack decrease linearly with age? Present your analysis in an ANOVA table with F -tests for non-linearity and for the slope of the regression line. Graph the data and the fitted line. Give a 99% confidence interval for the residual variance. State your assumptions and your conclusions.

Age:	20	20	20	25	25	25	31	31	31
Load:	11450	10420	11142	10840	11170	10540	9470	9190	9540

THE END