Original

Priv.-Doz. Dr. B. Heiligers

February 16, 1999

Statistical Analysis I

Examination

Please note the following

- The examination consists of 8 (differently weighted) problems for solution; for passing the exams you need
  - a total of **30 points** from all problems, **among those**
  - (at least) **12 points from problems 1 – 4, AND**
  - (at least) **12 points from problems 5 – 8.**

  For each problem, the number in parenthesis gives the maximum number of achievable points.
- You do not have to solve the problems completely, partial solutions are also possible. It is not enough, however, to display simply the result, but you should clearly display your approach and way to solution.
- You are allowed to use
  - pocket calculators,
  - text books, mathematical and/or statistical tables,
  - manuscripts and notes from the lectures and/or exercises.

Good luck !

Problem 1 (8 pts)

The annual incomes (in thousands of dollars) for 30 secretaries are given in the following stem-leaf diagram (stems represent thousands, leaves hundreds of dollars):

8	1 9 9
9	4 7 7 8
10	1 2 3 5 5 5 6 8 9
11	0 2 8
12	0 1 2 4 6
13	2 2 6
14	5
15	4
16	
17	6

(a) Draw a corresponding relative frequency histogram with five classes.

(b) Calculate the mode, the median, the lower and upper quartiles, and the interquartile range of the data.

Problem 2 (10 pts)

Last year a company placed 8 orders of particular units. The batch sizes (i.e., total numbers of ordered units) of these orders are: 55, 25, 30, 100, 75, 60, 120, 45.

(a) Calculate the mean, variance, and standard deviation of the variable "batch size".

(b) Assume that each unit costs \$3, and that each order causes fixed costs of \$50. Calculate mean and standard deviation of the variable "total costs of an order".

(c) A second sample of 12 batch sizes has mean 71.0 and variance 930.0. Find the overall mean and overall variance of "batch size" for the combined sample of 20 orders.

Problem 3 (12 pts)

91 stores were categorized according to the number of employees. For each store it was also reported whether or not there was a formal marketing plan. The following contingency table gives the results.

		formal plan	no formal plan
Employees	< 10	13	10
	$10 - 50$	18	12
	> 50	32	6

(a) Calculate all corresponding marginal and all conditional frequencies. What is the percentage of stores with formal marketing plan having more than 50 employees? What is the percentage of stores with less than 10 employees having no formal marketing plan ?

(b) Compute Cramer's V from the contingency table and interpret your result.

(c) For an arbitrary data set, can Cramer's V exceed the value 1 ?

Problem 4 (12 pts)

A company sets different prizes for a particular stereo system in 8 different cities. The accompanying table shows the numbers of sold units (y_i) and the corresponding prizes (x_i, in hundreds of dollars); the additional columns are given for convenience.

city	x_i	y_i	x_i^2	y_i^2	$x_i \cdot y_i$
1	5.5	420	30.25	176400	2310
2	6.0	380	36.00	144400	2280
3	6.5	350	42.25	122500	2275
4	6.0	400	36.00	160000	2400
5	5.0	440	25.00	193600	2200
6	6.5	380	42.25	144400	2470
7	4.5	450	20.25	202500	2025
8	5.0	420	25.00	176400	2100
Total	45.0	3240	257.00	1320200	18060

(a) Compute the correlation coefficient between "prize" and "number of sold units".

(b) Fit a regression line to the data points, taking "prize" as the regression, and "number of sold units" as the response variable. Calculate the coefficient of determination.

(c) Predict the number of sold units when the prize is \$590.

Problem 5 (10 pts)

From experience, the probabilities for the number X of breakdowns (per week) of a particular machine are known. These are given by the following probability function:

x	0	1	2	3	4
$f(x) = P(X = x)$	.10	.26	.42	.16	.06

(a) Find the cumulative probability function of X.

(b) Find the respective probabilities of the events

A: "There will be at least one breakdown per week", and of

B: "There will be at most one breakdown per week".

What are the probabilities of the the complements of A and B, respectively ?

(c) Find the probabilities of the union of A and B, and of intersection of A and B.

(d) Are A and B mutually exclusive ? Are they exhaustive ?

Problem 6 (12 pts)

An inspector examines items coming from an assembly line. A review of his records reveals that he accepts 90% of all items from this line. Furthermore, 99% of those items which he has not accepted are actually defective, and 98% of the accepted items are intact.

(a) What is the probability that an item is defective given it has been accepted ?

(b) Find the probability that an item from the assembly line is defective.

(c) What is the probability that an item is accepted given it is defective?

(d) Are the events "item accepted" and "item defective" independent ?

Problem 7 (12 pts)

The number of shirts bought by a consumer at a men's clothing store is described by the random variable X with probability function

x	0	1	2	3	4
$f(x) = P(X = x)$	.50	.20	.15	.10	.05

(a) Compute the expected value μ and the standard deviation σ of X.

(b) Find the probability that X falls in the interval $[\mu - 2\sigma, \mu + 2\sigma]$. What is the lower bound for this probability from Chebychev's theorem?

(c) Assume that at a particular day 100 consumers enter the shop. Their purchases are independent. What are the expected value and the standard deviation of the total number of shirts sold that day?

Problem 8 (8 pts)

10% of the items produced by a machine are defective. In a sample, 5 independently produced items are inspected for defects.

(a) What is the probability that exactly three items in the sample are defective?

(b) Find the mean and standard deviation of the total number of defective items in the sample.