

Original

Examination: **Statistical Analysis I**  
Examiner: Professor Dr. B. Heiligers

Summersemester 2000

**Please note the following:**

- The exams consists of 8 problems for solution; for each problem you can get at most 10 points. You do not have to solve the individual problems completely, partial solutions are also possible. It is not enough, however, to state the result only, but you should clearly display your approach and way to solution.
  - For passing the exams you have to fulfill **all** of the following three requirements. You should achieve
    - a total of (at least) **25 points** from all problems, **among those**
    - (at least) **10 points from problems 1 – 4**, and
    - (at least) **10 points from problems 5 – 8**.
  - You are allowed to use: Pocket calculators, text books, mathematical and/or statistical tables, manuscripts and notes from lectures and/or exercises.
  - **Good luck !**
- 

**Problem 1 (10 pts)**

The stem-leaf diagram gives the ages of 23 men who married in the last week.

6		0	
5		6	
4		0 4 6	
3		0 1 1 3 4 7	
2		0 2 3 3 3 5 6 6 7 8 9	
1		9	(stem units are 10 years, leaves units are years)

- Find the mode, the median, and the range of the data.  
Up to what age married (at least) 75% of these men?
- Draw a corresponding relative frequency histogram with four classes.

**Problem 2 (10 pts)**

Due to a large number of complaints, the manager of a fast-food restaurant was concerned about the length of time customers had to wait from arrival until being served. The waiting times (in minutes) of 40 customers are reported in the table.

waiting time		$m_k$	4	5	6	7	8	9	10
number of customers with that waiting time		$f_k$	3	7	11	10	5	3	1

- Find the sample mean  $\bar{x}$  and the sample standard deviation  $s$  of “waiting time”.
- Applying Chebyshev’s Theorem, what percentage of waiting times can be expected to fall in the interval  $[\bar{x} - 1.5s, \bar{x} + 1.5s]$ ? Is the actual percentage smaller?
- A second (independent) sample of 20 waiting times had mean 6.25 and variance 1.70. What are the overall mean and overall variance of the combined sample?

**Problem 3** (10 pts)

A broker firm wants to check whether the service it provides to customers depends on the customers' incomes. Each of 500 customers was asked to classify his or her annual income, and to rate his or her broker; the results are shown in the crosstab.

		income (in 1000 USD)		
		< 20	20 – 50	> 50
broker rating	outstanding	45	66	43
	average	89	124	51
	poor	34	37	11

- Find all marginal and all conditional frequencies from the crosstab. Among the customers with an income less than 20000 USD, what is the percentage of those rating their broker as "outstanding"?
- Calculate Cramer's  $V$  from the crosstab and interpret your result.

**Problem 4** (10 pts)

For investigating the effect of fertilizer on the yield of potatoes, the R & D department of a chemical company applied different amounts,  $x_i$ , of the fertilizer, and recorded the obtained yields,  $y_i$ ; (both in pounds). The data are shown in the table; the additional columns are given for convenience.

$i$	$x_i$	$y_i$	$x_i^2$	$y_i^2$	$x_i \cdot y_i$
1	1.0	24	1.00	576	24.0
2	1.5	32	2.25	1024	48.0
3	2.0	27	4.00	729	54.0
4	2.5	29	6.25	841	72.5
5	3.0	35	9.00	1225	105.0
6	3.5	36	12.25	1296	126.0
Total	13.5	183	34.75	5691	429.5*

- Fit a regression line to the data, taking "amount of fertilizer" as regression, and "yield of potatoes" as response variable. Find the coefficient of determination, and comment upon the goodness of fit.
- Predict the yield of potatoes when the amount of fertilizer is 4.0.

**Problem 5** (10 pts)

The amounts of 500 loans (in 1000 USD) made by a bank last year are described as follows:

amount of loan	1	2	3	4	5	6
number of loans of that amount	50	75	150	125	75	25

An auditor will select randomly one of these loans for inspection, (i.e., each loan has the same probability of being selected). Let  $X$  denote the amount of the selected loan.

- What are the possible values of  $X$ ? Find the probability function of  $X$ .
- Calculate the respective probabilities of the two events  
 $A$ : "the selected loan will be for 3000 USD or more",  
 $B$ : "the selected loan will be for 5000 USD or less".
- Find the probabilities of the union of  $A$  and  $B$ , and of intersection of  $A$  and  $B$ .
- Are  $A$  and  $B$  mutually exclusive? Are they exhaustive?

**Problem 6** (10 pts)

A soccer star scores (at least) one goal in 70% of all national, but only in 20% of all international soccer games. Half of his games are national, and half are international.

- (a) What is the soccer star's probability of scoring a goal in a game?
- (b) The star did not score a goal in his last game. What is the probability that the game was an international soccer game?
- (c) Are the two events "score a goal" and "play a national game" independent?

**Problem 7** (10 pts)

From his experience, a salesman knows the following probabilities of the number of business contracts  $X$  he will make next month.

$x$	5	6	7	8	9	10
$f(x) = P(X = x)$	.10	.20	.25	.30	.10	0.05

- (a) Find the expectation  $\mu$  and the standard deviation  $\sigma$  of  $X$ .
- (b) The salesman receives a monthly salary of \$2000 (fixed), plus \$100 for each contract he made in that month. What are the mean and the standard deviation of the salesman's monthly salary?

**Problem 8** (10 pts)

A small commuter airline flies planes that can seat up to eight passengers. Since the airline knows from history that 20% of ticketed passengers will cancel their reservations at the last minute, they accept "overbooking" on the flights.

- (a) Suppose the airline sold  $n = 10$  tickets on a particular flight. What is the probability to "run out of seats", i.e., what is the probability that more than eight passengers will show up for this flight?
- (b) What are expectation and variance of the number of passengers which will show up for the flight?