Otto–von–Guericke–Universität Magdeburg **July 8, 2003**

Examination: **Statistics II (Bachelor in Economics & Management)**

Examiner: PD Dr. Th. Müller-Gronbach

Please note the following:

- The exam consists of 5 problems for solution; the points you can get for each problem are given in parentheses next to the number of the problem. You need not solve the individual problems completely, partial solutions are also accepted. For problems 2–5 it is not enough, however, to state the result only, but you should clearly display your approach and your way to the solution. Also, *conclusions* are to be drawn where applicable.
- For passing the exam you have to achieve a total of (at least) **20 points** out of all problems.
- 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. (7 points)

In the following $X_1, \ldots, X_n \sim \mathrm{B}(1, p)$ is an i.i.d. sample of Bernoulli variables with probability of success p, i.e. $P(X_i = 1) = p = 1 - P(X_i = 0)$. Are the following statements correct or wrong? (Just answer *Yes* or *No.*)

a) The larger the sample size n, the smaller the sample mean.

b) The smaller the sample size n, the larger the mean of the sample variance.

c) For n large, $\sqrt{n} \cdot \dfrac{\overline{X} - p}{\sqrt{p \cdot (1-p)}}$ is approximately standard-normally distributed.

d) $\dfrac{1}{n} \sum_{i=1}^{n} (X_i - \overline{X})^2$ is an unbiased estimator of the variance $p \cdot (1-p)$.

e) Let $\widehat{p}_1$ be an unbiased estimator of p, and let $\widehat{p}_2$ be a biased estimator of p. Then $\mathrm{Var}(\widehat{p}_1) \leq \mathrm{Var}(\widehat{p}_2)$, i.e. the unbiased estimator has a smaller variance.

f) Let Y be a continuous random variable, and denote by q_α an α-quantile from the distribution of Y. Then $P(q_{0.1} \leq Y \leq q_{0.7}) = 0.6$.

g) If φ is a level-α-test (for a testing problem H_0 versus H_1), then the probability of a type-II-error is at most α.

Problem 2. (9 points)

Let $X_1, \ldots, X_{n_x}$ and $Y_1, \ldots, Y_{n_y}$ be two independent i.i.d. samples of sizes n_x and n_y, resp., from a normally distributed population with (unknown) mean μ and variance $\sigma^2 = 6$. Consider the two estimators of μ,

$$\widehat{\mu}_1 := \frac{1}{3}\overline{X} + \frac{2}{3}\overline{Y} \qquad \text{and} \qquad \widehat{\mu}_2 := \frac{1}{n_x + n_y}\left(n_y \cdot \overline{X} + n_x \cdot \overline{Y}\right) .$$

We are given the following relevant data from the samples:

	sample X	sample Y
sample size	$n_x = 10$	$n_y = 20$
sample mean	$\overline{x} = -3.0816$	$\overline{y} = -2.8117$

a) Show that $\widehat{\mu}_1$ and $\widehat{\mu}_2$ are unbiased for μ. Calculate the according estimates.

b) Calculate the variances of $\widehat{\mu}_1$ and $\widehat{\mu}_2$. Which of the estimators is more efficient?

Problem 3. (15 points)

A company producing refrigerators is interested in comparing the electric power consumption of their own product with that of a similar refrigerator (having the same capacity) produced by a competitor. The power consumptions of the own and the competitor's refrigerator, measured per month, are assumed to be independent and normally distributed with unknown means μ_x, μ_y and variances σ_x^2, σ_y^2, respectively.

The following data have been observed from two samples of sizes 12, each:

i	x_i	x_i^2	y_i	y_i^2
1	56.50	3192.25	59.75	3570.06
2	58.78	3455.09	53.30	2840.89
3	69.05	4767.90	63.09	3980.35
4	68.97	4756.86	56.52	3194.51
5	52.17	2721.71	65.65	4309.92
6	60.29	3634.88	54.71	2993.18
7	59.38	3525.98	45.94	2110.48
8	67.59	4568.41	53.13	2822.80
9	55.18	3044.83	69.38	4813.58
10	53.64	2877.25	57.11	3261.55
11	56.46	3187.73	71.34	5089.40
12	60.57	3668.73	52.43	2748.91
$\sum$	718.58	43401.63	702.35	41735.64

a) Test, at a level of 5%, the hypothesis that the two population variances are equal.

In the following, assume that the population variances are equal, $\sigma_x^2 = \sigma_y^2$.

b) Calculate a 90%-confidence interval estimate of the difference of the two population means, $\mu_x - \mu_y$. Interpret this confidence interval.

c) At level $\alpha = 0.1$, is there enough evidence that the mean power consumption of the company's own refrigerator is smaller than that of the competitor's refrigerator?

Problem 4. (12 points)

Does the participation in tutorials have a significant influence on passing or not passing the exam? A sample of $n = 56$ students yielded the following data:

		# of attended tutorials			
frequency		0–4	5–9	10–18	19–20
passing the exam	no	8	10	8	2
	yes	2	6	12	8

At a level of 5%, test whether

a) *'passing the exam'* is independent of *'attendance in tutorials'*,

b) the probability of *'passing the exam'* is 0.6.

Problem 5. (8 points)

A company producing light bulbs is interested in comparing a new type of light bulbs with the established one. The quality assurance department observed the following lifetimes of $n_x = 8$ light bulbs of the established type:

$$38.44\,,\quad 77.42\,,\quad 56.59\,,\quad 23.04\,,\quad 282.71\,,\quad 50.61\,,\quad 299.04\,,\quad 87.86\,,$$

and the following lifetimes of $n_y = 7$ light bulbs of the new type:

$$138.74\,,\quad 33.43\,,\quad 439.91\,,\quad 46.40\,,\quad 22.96\,,\quad 21.00\,,\quad 43.72\,,$$

all measured in hours.

State suitable assumptions on the distributions of the lifetimes of the two types of light bulbs. Test, at a level of $\alpha = 0.05$, whether the two population means coincide.