Statistics II in English
Exercises from *Übungsaufgaben und Lösungen zu Statistik I und II*
Part 2

Translation made by Barbara Choroś and Steffen Dähne

Institut für Statistik and Ökonometrie
CASE - Center for Applied Statistics and Economics
Humboldt-Universität zu Berlin

Exercise 7-1:
Let the population have the mean μ and variance σ^2. Let (X_1, X_2, X_3) be a random sample from this population. The following three estimators are given:

$$\begin{aligned}\hat{\theta}_1 &= \frac{1}{3}(X_1 + X_2 + X_3)\\ \hat{\theta}_2 &= \frac{1}{4}(2X_1 + 2X_3)\\ \hat{\theta}_3 &= \frac{1}{3}(2X_1 + X_3)\end{aligned}$$

1. Which of these functions are unbiased?
2. Which of these functions would you prefer according to the criterion of effectiveness?

Exercise 7-2:
A random sample of size n=10 drawn from a population yields the following values: 2; 4; 3; 5; 1; 4; 5; 1; 2; 3. Estimate from these data as good as possible mean and variance of the population.

Exercise 7-4:
Students of a big university are asked for their motivation of studying. This random variable (rv) has three values: 0 (not motivated), 1 (motivated), 2 (very motivated). The unknown proportions of these degrees of motivation in the population are π_0, π_1, π_2. For estimating them a simple random sample of size $n = 10$ is drawn. The rvs X_i represent the degree of motivation of the i-th student in the sample. The following estimators for π_1 and π_2

are proposed:

$$\hat{\pi}_1 = \frac{1}{10}\sum_{i=1}^{10}(2X_i - X_i^2) \quad (1)$$

$$\hat{\pi}_2 = \frac{1}{20}\sum_{i=1}^{10}(X_i^2 - X_i) \quad (2)$$

$$(3)$$

1. Are these two estimators unbiased?

2. Give an unbiased estimation function for the proportion of unmotivated students.

3. Estimate π_0, π_1, π_2 if the following answers are in the sample:
0 2 1 0 0 2 2 1 0 2.

Exercise 7-5:
A one-armed bandit (coin machine) has the following probability distribution for the gain X per game (in EUR). The producer of these gaming

x	-1	0	+1
P(X=x)	p	p	1-2p

machines asks a statistician to estimate p in order to check, if the value of p changed since the starting of the machines.

1. The statistician draws a random sample of size $n = 6$, i.e. he operates the gaming machine exactly 6 times and writes down the gain.
The sample $(X_1, X_2, X_3, X_4, X_5, X_6)$ has the following realisation: $(-1, 1, -1, 0, 1, 1)$. Verbalise this sampling result.

2. Determine the following probabilities:
$P(X = 0)$, $P(X = -1)$, $P(X = 1)$.

3. How would you determine the probability for the gain X per game based on the sample mentioned above if you had no information about the probability distribution of X at all?

4. What is $P\{(X_1, X_2, X_3, X_4, X_5, X_6) = (-1, 1, -1, 0, 1, 1)\}$ if the above probability distribution is supposed.

5. Determine the maximum likelihoood (ML) estimator for p for this problem.

6. Estimate p based on the sample at hand with the ML method.

7. Estimate p based on the sample at hand with the least squares method.

Exercise 7-9:
You are interested in the frequency of accident at an important junction. Insurance companies as well as the police already studied this in the past, so that you can access their results.

x	0	1	2	$x \neq 0, 1, 2$
P(X=x)	0,7	0,1	0,2	0

Table 1: Insurance company

x	0	1	2	$x \neq 0, 1, 2$
P(X=x)	0,1	0,4	0,5	0

Table 2: Police

Unfortunately, both probability distributions differ severely. In order to decide between the two distributions, you by yourself observe the junction on 5 randomly drawn days and note the number of accidents. The following sample results: $(0, 2, 0, 2, 1)$.

1. What is your decision when using the ML principle?

2. What is your decision when using the Least Squares principle?

Exercise 7-12:

A For a test 49 randomly drawn cars of the same type were fuelled with the same amount of fuel. With this amount of fuel the cars went on average 50km. Suppose the std. dev. of the population to be known with 7km.

   (a) Give an explicit confidence interval $[V_L, V_U]$ for the average mileage (in km) μ for this type of car and the confidence level $1 - \alpha$.
   (b) Determine the estimation interval for μ when $1 - \alpha = 0.95$.
   (c) What sample size n is needed, if for the same confidence level the estimation interval for shall have a width of 2km?

B Some visitors of this test event are randomly chosen by journalists and asked for their membership in the ADAC (Allgemeiner Deutscher Automobil Club - German automobile club). Among the 200 persons asked 40 were ADAC members. Determine the estimation interval for π when $1 - \alpha = 0.99$.

C A vending machine company has installed coffee vending machine for this test event. This machine fills the $0,2l$ goblet with coffee. Assume the filling quantity to be approximately normally distributed. A random sample of size $n = 5$ yields the following values in l: 0.18, 0.25, 0.12, 0.20, 0.25.

(a) Determine explicitly the confidence interval $[V_L, V_U]$ for the average filling quantity μ of this vending machine at the confidence level $1 - \alpha$.

(b) Determine the estimation interval for μ when $1 - \alpha = 0.95$.