Statistical Techniques for CS/BIT 2021-1B

Practice test #1

Time: 2hrs 15min

Instructions. This test consists of 7 exercises. The formula sheet and the probability tables are provided. An ordinary calculator is allowed, not a programmable one (GR).

If a question asks "is there a statistical test that ...", this should be interpreted as "a statistical test *covered in this course* that ...".

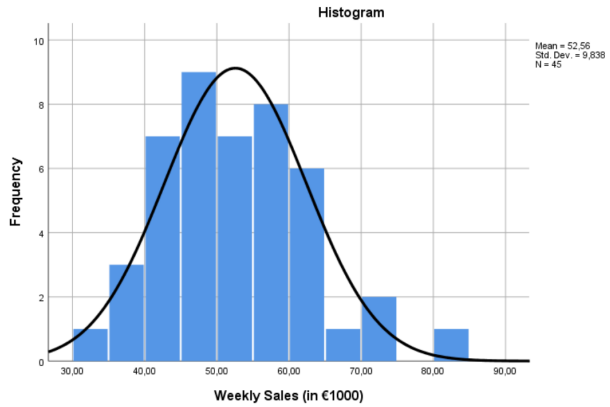
1. Consider the following data from a bookstore for the weekly sales (in € 1000) from 45 consecutive weeks. The sales are arranged in order:

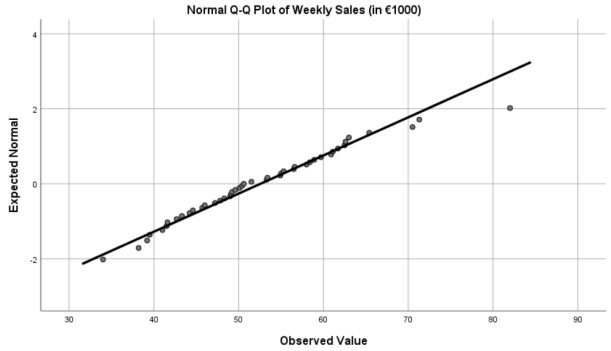
34.0	38.2	39.2	39.5	41.0	41.5	41.6	42.7	43.3
44.2	44.6	45.7	46.0	47.2	47.8	48.3	49.0	49.1
49.2	49.6	50.1	50.4	50.6	51.5	53.3	53.4	54.9
55.0	55.3	56.5	56.6	58.0	58.4	58.9	59.7	60.9
61.1	61.7	62.6	62.5	63.0	65.4	70.5	71.3	82.0

The classical numerical summary for this sample is determined together with the histogram and the normal Q-Q plot:

Sample Size	Mean	Std. Dev.	Variance	Skewness	SE Skewness	Kurtosis	SE Kurtosis
45	52.56	9.84	96.77	0.601	0.354	0.573*	0.695

*Kurtosis is adjusted as in SPSS (kurtosis − 3).





a. Determine the Z-score of the highest weekly sale and argue whether this value is "extreme" or not, using the "Empirical Rule".

b. Provide

(1) 20^{th} and 80^{th} percentile,

(2) the 5-number summary,

(3) outliers according to the 1.5 x IQR-rule.

c. Do you consider a normal model reasonable for these dataset of the weekly book sales? Motivate your answer with all of the following:

(1) numerical measures,

(2) the histogram,

(3) the normal Q-Q plot.

d. The value of Shapiro-Wilk's test statistic was reported as $w = 0.974$. Give

(1) the hypotheses,

(2) the Rejection Region,

(3) the conclusion of this test, for $\alpha = 0.10$.

2. Let T_1 and T_2 be estimators for the some parameter θ. Which of the following statements is true (check all that apply)?

(a) $\text{MSE}(T_1) = \text{Bias}(T_1) + \text{Var}(T_1)$

(b) $\text{MSE}(T_1) = \text{E}(T_1 - \theta)^2$

(c) The standard error of T_1 is equal to $\text{MSE}(T_1)$

(d) T_1 is better than T_2 if $\text{Var}(T_1) < \text{Var}(T_2)$

3. Consider a one-sided test regarding some population parameter μ. Suppose the hypotheses are $H_0 \colon \mu \leq \mu_0$ vs $H_1 \colon \mu > \mu_0$, the test statistic is called T, and the null hypothesis is rejected whenever $T \geq c$. Suppose the true value of the parameter is $\mu = \mu_1$. Which expression correctly expresses the power of the test (check all that apply)?

(a) $1 - P(T \leq c \mid \mu = \mu_0)$

(b) $P(T \leq c \mid \mu = \mu_0)$

(c) $1 - P(T \leq c \mid \mu = \mu_1)$

(d) $P(T \leq c \mid \mu = \mu_1)$

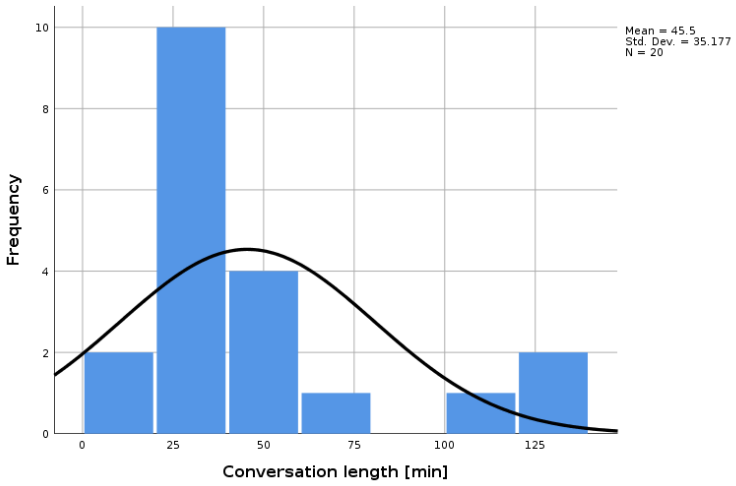
4. A mortgage provider wants to determine the prices for mortgage consultations. One of the things he examines for this purpose, is the length of the conversations. The results of a random sample of $n = 20$ consultations are as follows:

21	26	106	32	33	122	42	16	20	23
78	39	40	32	21	130	43	13	50	23

The classical numerical summary for this sample is determined together with a histogram:

Sample Size	Mean	Std. Dev.	Variance	Skewness	SE Skewness	Kurtosis	SE Kurtosis
20	45.50	35.18	1237.42	1.443	0.419	0.745*	0.391

*Kurtosis is adjusted as in SPSS (kurtosis − 3).



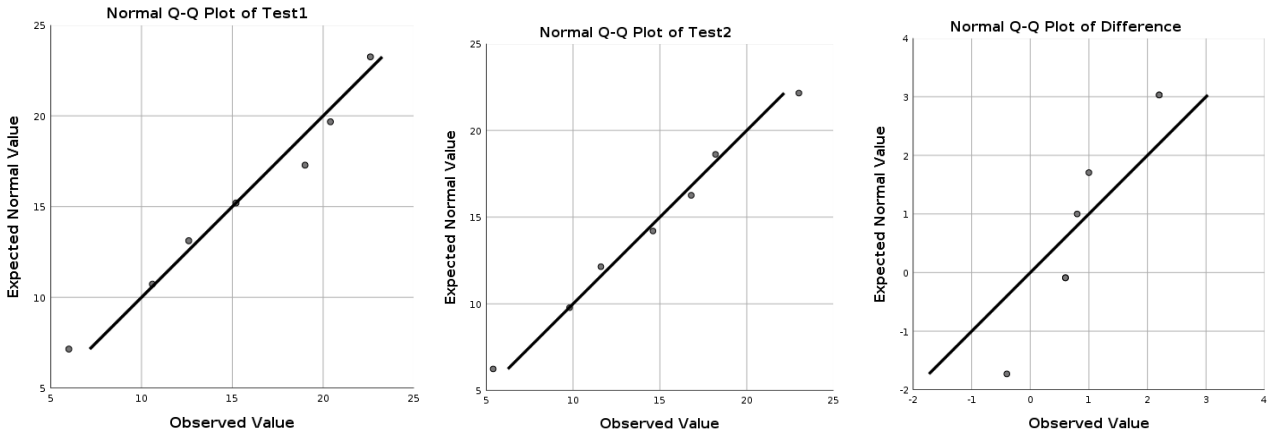
a. Can you apply a suitable method to construct a 95% confidence interval for the expected conversation length? If so, construct the interval. If not, explain why not.

b. Someone interprets the confidence interval as follows: "If we repeat the experiment often (observing the conversation times of 20 random customers each time), then in about 95 of the 100 repetitions the sample mean is included in the interval." Is this a correct interpretation? Why (not)?

5. A teacher develops two exams for the same course with the expectation that the Test 1 is easier than Test 2. The teacher asks seven random students to complete both exams in a random order. The data of the grades per student is shown below, together with some descriptive statistics and graphs.

Student	1	2	3	4	5	6	7
Test 1	15.2	20.4	19.0	22.6	6.0	12.6	10.6
Test 2	14.6	18.2	16.8	23.0	5.4	11.6	9.8
Difference	0.6	2.2	2.2	-0.4	0.6	1.0	0.8

	Sample Size	Mean	Std. Dev.	Variance
Test 1	7	15.20	5.89	34.80
Test 2	7	14.20	5.82	33.98
Difference	7	1.00	0.93	0.86



a. Is there a statistical test that can be correctly applied here to decide whether Test 1 is indeed easier than Test 2? Justify your answer. If your answer is yes, compute the test statistic and the p-value for the test you chose.

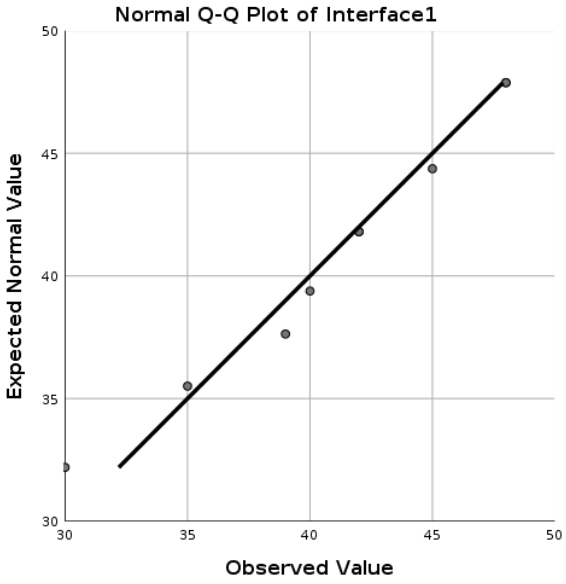
b. Some people interpret the p-value of a test as: "p is the probability that the null hypothesis is true". Is this statement correct? Why (not)?

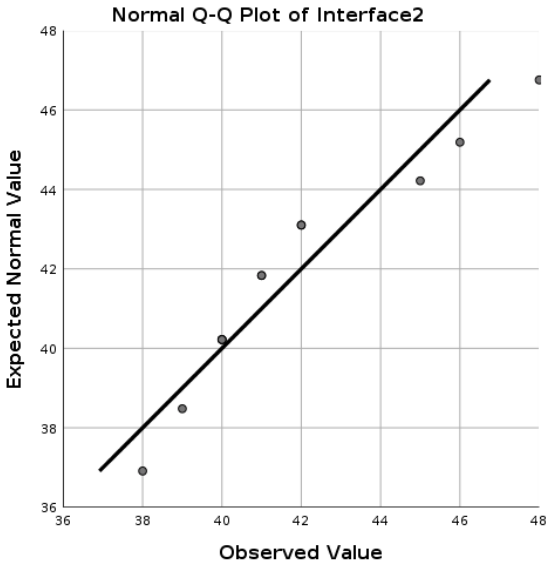
6. A student runs an experiment to study the effect of the interfaces of two different technologies developed in Human Computer Interaction (HCI). In order to compare the interfaces, the student asked 21 randomly selected interface users to provide responses on the interface they tested. The responses are the total grades of the 50 questions in the questionnaire with 1 point for each correct answer and 0 point for each incorrect answer. There are 9 cases for the first and 12 cases for the second interface. For each of the two interfaces, the number of correct responses (out of 50) are observed as follows:

Interface 1	30	35	40	39	42	45	48	42	48			
Interface 2	42	40	48	46	41	40	45	38	39	40	41	42

	Sample Size	Mean	Std. Dev.	Variance	Skewness	SE Skewness	Kurtosis	SE Kurtosis
Int.1	9	41.00	5.895	34.750	-0.631	0.717	0.090	1.400
Int.2	12	41.83	3.010	9.061	0.960	0.637	.119	1.232

*Kurtosis is adjusted as in SPSS (kurtosis − 3).





The student is interested in knowing if the variances of the responses are different for the two interfaces.

a. Explain in words what should be the null and alternative hypotheses for a statistical test in order to answer the student's question.

b. Suppose that a suitable test is applied and the test statistic is not in the rejection region. Does this prove that the variances of the responses are equal? Why (not)?

c. Is there a statistical test that can be correctly applied to answer the student's question? If so, which test would you apply? Justify your answer.

7. During the corona crisis in 2020 the managers of the universities decided to switch to online education and online testing as to make sure that students were not delayed in their study programs. But many of the lecturers objected against online testing. *Scienceguide* organized a survey among lecturers of technical studies and among lecturers of social sciences to verify whether the opinions on digital testing are different:

- Among 95 social science lecturers 40 consider online testing acceptable and 35 were against. The remaining 20 persons did not have a (clear) opinion.
- Among 120 technical lecturers 35 think that online testing is acceptable and 60 were against (the remaining 25 had no opinion).

Can we state that there is a difference in opinion on digital testing among lecturers in technical and social sciences?

a. Which is a suitable test to answer the question at hand? Justify your answer.

b. For the test you chose above, conduct the test applying the testing procedure with $\alpha = 5\%$.

$$\text{Grade} = 1 + \frac{\#\text{ points}}{36} \times 9$$

Rounded to 1 decimal

question	1	2	3	4	5	6	7	Total
points possible	11	2	2	5	5	5	6	36