Test "Statistics" for IEM in module 3 (201300108) – April 16, 2020, 13.45-16.00 V2

Lecturer: Dick Meijer – module coordinator: Adina Aldea.

This online test consists of 6 exercises. You can use **the reader**, including the formula sheet and the probability tables. **Your own notes** are allowed as well. Other resources are not allowed. Use a common (non-programmable) calculator, not a graphical one (GR). The test should be hand written. Number the pages before copying and uploading them.

Above the resources for this test are mentioned. Read the following carefully:
By testing you remotely in this fashion, we express our trust that you will adhere to the ethical standard of behaviour expected of you. This means that we trust you to answer the questions and perform the assignments in this test to the best of your own ability, without seeking or accepting the help of any source that is not explicitly allowed by the conditions of this test.
Copy the following statement to your first answering sheet and **sign it with your name and student number**: "*I will make this test to the best of my own ability, without seeking or accepting the help of any source not explicitly allowed by the conditions of the test.*"

1. We determined the sodium content of twenty-five boxes organic cornflakes.
The ordered measurements (in milligram) are as follows:

128.24 128.33 128.71 128.77 128.77 129.00 129.29 129.39 129.53 129.54 129.64 129.65 129.73
129.74 129.78 130.12 130.14 130.42 130.69 130.72 130.80 130.91 130.92 131.15 133.15

Using certain software, we obtained the (classical) numerical summary of the data.
The results are presented in the following table:

Sample size	Mean	Standard deviation	Variance	Skewness	Kurtosis
25	129.885	1.076	1.157	0.994	5.151

a. Determine the 5-numbers-summary of the data and the 90th percentile.
b. Before confidence intervals can be determined and tests can be carried out, we need to examine whether the normal distribution is a reasonable assumption for these data.
Comment on the normality on the basis of the numerical summaries (in a. and in the table above).
c. Using the same software, we also determined the value of Shapiro Wilk's test statistic: $W = 0.932$
Provide for this test the rejection region for α = 5%. Draw your conclusion concerning the normality.

In the following parts you may, if necessary, assume normality:
d. Determine a 95%-confidence interval for the expected sodium content.
Give an **interpretation** for this numerical interval.
e. Determine a 95%-confidence interval for the variance of the sodium content.

2. A population has an unknown mean μ and an unknown variance σ^2.
Consider a (very) small random sample X_1 and X_2 (sample size 2), drawn from this population and consider the following two estimators of μ: $T_1 = \frac{X_1+X_2}{2}$ and $T_2 = \frac{X_1+3X_2}{4}$.

a. Show whether T_1 and T_2 are unbiased.
b. Which of these two estimators is the best, in terms of the Mean Squared Error? Why?

Consider for the remaining exercises 3, 4, 5 and 6 the following problems (indicated with A, B and C):

A. Many people suffer from FNE (´Fear of Negative Evaluation´). To find out whether eating habits have any influence, a psychologist carried out an experiment with two groups of 11 students each. Students of the first group suffer from the eating disorder bulimia. The other students have normal eating habits. Each student fills out a survey. Based on the results of this survey, a FNE-score is calculated. The higher the score, the higher the FNE. The results are as follows:

With bulimia x_1	21	13	10	20	25	19	16	21	24	13	14	$\bar{x}_1 = 17.82, s_1 = 4.92$
Normal eating habits x_2	13	6	16	13	8	19	23	18	11	15	7	$\bar{x}_2 = 13.55, s_2 = 5.34$
$x_3 = x_1 - x_2$	+8	+7	-6	+7	+17	0	-7	+3	+13	-2	+7	$\bar{x}_3 = 4.27, s_3 = 7.52$

B. 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 programmes. But many of the lecturers objected against online testing since the educational goals are different (open book) and online tests are very fraud sensitive. *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 sciences 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 (25 had no opinion).

C. Is a training effective in increasing the speed of reaction? A sample of 10 arbitrarily chosen car drivers was subjected to a training. Before and after the training the speed of reaction (in milliseconds) was measured. The table below shows the results (averages of repeated tests).

Car driver	1	2	3	4	5	6	7	8	9	10
Before the training	175	252	252	280	256	190	276	371	270	178
After the training	168	253	247	286	235	185	261	375	261	166

3. First we will apply **parametric tests** to the problems A, B and C: these tests include binomial and t-tests for one and two samples, the F-test on the equality of variances and Chi-square tests on the variance and for one or two categorical variables (amongst which the test on independence and the test on homogeneity). Determine for each of the questions in parts a., b. and c. below (only!):

1. The type of test. Motivate your choice in one sentence.
2. The hypotheses
3. The (formula of the) test statistic and its distribution under H_0.

a. Problem A: Does this survey prove that the expected FNE score for students with bulimia is higher?

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

c. Problem C: Does the training improve the speed of reaction of car drivers?

4. Use in this exercise 5% as a level of significance:

a. Determine for problem A the observed value of the test statistic, the p-value of the test as stated in 3a. and decide whether H_0 is rejected.

b. Determine for problem B the observed value of the test statistic, the rejection region for the test in 3b. and draw your conclusion in words.

5. Review problem C: which test is the **non-parametric** alternative for the test in question 3c.? Conduct this test in 8 steps and with $\alpha = 5\%$. Start in step 1 by defining the observed variable(s).

6. Conduct for problem A the test on the equality of variances with the testing procedure and with $\alpha = 5\%$.

-- END--

Grade = $1 + \frac{\#\text{ points}}{50} \times 9$,

rounded on 1 decimal.

1					2		3			4		5	6	Tot
a	b	c	d	e	a	b	a	b	c	a	b			
3	2	3	4	2	2	2	4	4	4	4	4	6	6	50

	1					2		3			4		5	6		
	a	b	c	d	e	a	b	a	b	c	a	b			Tot	Test
	3	2	3	4	2	2	2	4	4	4	4	4	6	6	50	Result
Your score	**3**	**2**	**1**	**4**	**2**	**2**	**1**	**3**	**3**	**3**	**3**	**4**	**6**	**5**	**42**	**8.6**

Solutions

Exercise 1

a. Median = $X_{(13)} = 129.73$.
25% of 25 is 6.25, so $Q_1 = x_{(7)} = 129.29$, likewise: $Q_3 = x_{(19)} = 130.69$
5-numbers-summary: min = 128.24, $Q_1 = 129.29$, $M = 129.73$, $Q_3 = 130.69$, max = 133.15.
90% of 25 is 22.5, the 90th percentile is $x_{(23)} = 130.92$.

b. Both the skewness 0.99 and the kurtosis 5.2 deviate from the reference values 0 and 3 of the normal distribution: these values do not supply a decisive answer about normality.
Furthermore the difference of the median 129.73 and the mean 129.885 is not large, which does not point towards a strong skewness (to the right in this case according to the skewness coefficient), but rather to symmetry, like in the case of a normal distribution.

c. For $\alpha = 5\%$ and $n = 25$ we obtain the rejection region $W \leq c = 0.918$.
So $W = 0.932$ does not lie in the rejection region: it cannot be proven that the distribution is not normal, at a 5% significance level.

d. 95%-CI(μ) = $\left(\overline{X} - c\frac{S}{\sqrt{n}}, \overline{X} + c\frac{S}{\sqrt{n}}\right)$, where $n = 25$, $\overline{x} = 129.885$, $s \approx 1.076$ and $c = 2.064$ from the t_{25-1}-table, such that $P(T_{24} > c) = 0.025$
95%-CI(μ) = (129.885– 2.064× $\frac{1.076}{\sqrt{25}}$, 129.885– 2.064× $\frac{1.076}{\sqrt{25}}$) ≈ (129.44, 130.33).
Interpr.: "We are 95% confident that the expected sodium content is between 129. 44 and 130.33 mg"

e. $95\% - CI(\sigma^2) = \left(\frac{(n-1)S^2}{c_2}, \frac{(n-1)S^2}{c_1}\right)$, where $P(\chi^2_{25-1} \leq c_1) = P(\chi^2_{25-1} \geq c_2) = \frac{\alpha}{2} = 0.025$
So $c_1 = 12.40$ and $c_2 = 39.36$ and $s^2 = 1.157$.
$95\% - CI(\sigma^2) = \left(\frac{24 \cdot 1.157}{39.36}, \frac{24 \cdot 1.157}{12.40}\right) \approx (0.71, 2.24)$

Exercise 2

a. Both estimators are unbiased, because:
$E(T_1) = E(\overline{X}) = \mu$ and
$E(T_2) = E\left(\frac{X_1 + 3X_2}{4}\right) = \frac{1}{4}E(X_1 + 3X_2) = \frac{1}{4}(EX_1 + 3 \cdot EX_2) = \frac{1}{4} \cdot 4\mu = \mu$

b. For unbiased estimators, the mean squared error equals the variance. Then T_1 is better, because:
$var(T_1) = var(\overline{X}) = \frac{\sigma^2}{2} = \frac{1}{2}\sigma^2$ and
$var(T_2) = var\left(\frac{X_1 + 3X_2}{4}\right) = \frac{1}{4^2}var(X_1 + 3X_2) = \frac{1}{16}[var(X_1) + 3^2 var(X_2)] = \frac{5}{8} \cdot \sigma^2$
So $var(T_1) < var(T_2)$.

Exercise 3

a. 1. In this case, we are dealing with two independent samples of 11 observations each, drawn from two different populations (!). (*note that the pairs in the table are not related to each other, so do not examine "accidental" differences!*): we will apply the **two independent samples *t*-test** on the equality of the population means (assuming equal, but unknown variances).
2. We test $H_0: \mu_1 = \mu_2$ against $H_1: \mu_1 > \mu_2$ with α = 5% (*index 1 is "with bulimia"*)
3. Test statistic: $T = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{S^2\left(\frac{1}{11} + \frac{1}{11}\right)}}$ with $S^2 = \frac{10S_1^2 + 10S_2^2}{11 + 11 - 2}$
Under H_0 T has a t-distribution with 11 + 11 – 2 = 20 degrees of freedom (*short*: $T \overset{H_0}{\sim} t_{20}$)

b. 1. In this situation we observe the variable "opinion on online testing" with 3 categories for two populations of lecturers: Social sciences and Technical sciences. The sample sizes are chosen separately and the samples are independent. Hence we can use the **χ^2-test on homogeneity**.

2. Test $H_0: p_{11} = p_{21}$ and $p_{12} = p_{22}$ and $p_{13} = p_{23}$ against

$H_1: p_{1i} \neq p_{2i}$ for some at least one value of i

3. $\chi^2 = \sum_{j=1}^{c} \sum_{i=1}^{r} \frac{(N_{ij} - \hat{E}_0 N_{ij})^2}{\hat{E}_0 N_{ij}}$, with $\hat{E}_0 N_{ij} = \frac{\text{row total} \times \text{column total}}{n}$,

where χ^2 has under H_0 a χ^2-distribution with $df = (r-1)(c-1) = 2$

c. 1. Paired samples, since we have two dependent observations per person. Hence we will apply the **one sample t-test on the differences**, the reductions before – after of the time of reaction.

2. Test $H_0: \mu = 0$ against $H_1: \mu > 0$ (*where μ is the expected reduction.*)

3. Test statistic $T = \frac{\overline{X}}{S/\sqrt{n}}$ has under H_0 a t_{10-1}-distribution.

Exercise 4

a. Since $S^2 = \frac{10}{20} \cdot 4.92^2 + \cdot \frac{10}{20} 5.34^2 = 26.361$, the observed value of the test statistic is:

$$t = \frac{17.82 - 13{,}55}{\sqrt{26.361\left(\frac{1}{11} + \frac{1}{11}\right)}} \approx 1.95$$

The test is right-sided ("*Is the expected FNE score with bulimia is higher*"): reject H_0 if the p-value $\leq \alpha$

The p-value = $P(T_{20} \geq 1.95)$ is between 5% (1.725) and 2.5% (2.086).

The p-value < 5% = $\alpha \Rightarrow$ reject H_0.

b. The table of observed and expected values is as follows, e.g. $E_{11} = \frac{95 \times 75}{215} \approx 33.1$

	Pro	Contra	No opinion	Total
Social Sciences	40 (33.1)	35 (42.0)	20 (19.9)	95
Technical Sciences	35 (41.9)	60 (53.0)	25 (25.1)	120
Total	75	95	45	215

Observed value: $\chi^2 = \frac{(40-33.1)^2}{33.1} + \cdots + \frac{(25-25.1)^2}{25.1} \approx 4.67$

Reject H_0 if $\chi^2 \geq c = 5.99$, since $P(\chi^2_2 \geq c) = 0.05 = \alpha$.

$\chi^2 = 4.67 < 5.99 = c$, so we cannot reject H_0, in words:

"At a 5% level of significance there is not sufficient proof to claim that lecturers in Social Sciences and in Technical Studies have different opinions on online testing."

Exercise 5

1. $X =$ "number of positive reductions in the time of 10 persons in the sample" $\sim B(10, p)$.

2. Test $H_0: p = \frac{1}{2}$ (or median reduction = 0) against $H_1: p > \frac{1}{2}$ with $\alpha = 5\%$.

3. X

4. Under H_0 X is $B\left(10, \frac{1}{2}\right)$-distributed.

5. Observed $X = 7$ positive reductions (before – after): +7, -1, +5, -6, +21, +5, +15, -4, +9, +12.

6. Right-sided test: reject H_0 if the p-value = $P(X \geq 7 | H_0) \leq \alpha = 5\%$.

$P(X \geq 7 | H_0) = 1 - P(X \leq 6 | H_0) = 1 - 0.828 = 17.2\%$

7. The p-value = 17.2% > α: we fail to reject H_0.

8. At a 5% level of significance the sample did not sufficiently show that the time of reaction is reduced after the training.

An alternative is determining the right-sided rejection region: $X \geq c = 9$, since for $p = \frac{1}{2}$ we have $P(X = 10) = \left(\frac{1}{2}\right)^{10} \approx 0.001$, $P(X = 9) = 10\left(\frac{1}{2}\right)^{10} \approx 0.010$ and $P(X = 8) = \binom{10}{8}\left(\frac{1}{2}\right)^{10} \approx 0.044$.
Then $P(X \geq 8) \approx 5.5\% > \alpha$ and $P(X \geq 9) \approx 1.1\%$, so $c = 9$.

Exercise 6

1. Model: all FNE-scores $X_1, \dots, X_{11}$, of the students with bulimia and $Y_1, \dots, Y_{11}$ of students with normal eating habits are independent and $X_i \sim N(\mu_1, \sigma_1^2)$ and $Y_j \sim N(\mu_2, \sigma_2^2)$.
   (*note that "$X_1, \dots, X_{11}$ and $Y_1, \dots, Y_{11}$ are independent" implies that the two samples are independent and each sample is random.*)
2. Test $H_0: \sigma_1^2 = \sigma_2^2$ against $H_1: \sigma_1^2 \neq \sigma_2^2$ with $\alpha = 5\%$.
3. Test statistic $F = \frac{S_1^2}{S_2^2}$.
4. Distribution under H_0: $F \sim F_{11-1}^{11-1}$
5. Observed value: $F = \frac{s_1^2}{s_2^2} = \frac{4.92^2}{5.34^2} \approx 0.849$
6. We have a two-sided test: reject H_0 if $F \leq c_1$ or $F \geq c_2$.
   $P(F_{10}^{10} \geq c_2) = \frac{\alpha}{2} = 0.025$, so according to the F_{10}^{10}: $c_2 = 3.72$
   $P(F_{10}^{10} \leq c_1) = P\left(F_{10}^{10} \geq \frac{1}{c_1}\right) = \frac{\alpha}{2} = 0.025$, so $\frac{1}{c_1} = 3.72$, or $c_1 \approx 0.269$
7. Since $F = 0.849$ does not lie in the Rejection Region, we cannot reject H_0.
8. At a significance level of 5% we cannot prove that the variances of the FNE-scores of the two groups are different.