Test Statistical Techniques for BIT-TCS (Module 6 -201800421) , Thursday the 20th of December 2018, 8.45-11.00 h. Lecturer Dick Meijer, module-coordinator Dennis Reidsma

This test consists of 4 exercises. A formula sheet and the probability tables are provided. An ordinary scientific calculator is allowed, not a programmable one (GR).

1. The effectiveness of a helpdesk is assessed by researchers, by measuring the service times of customers who asked for support. The researchers intend to estimate the expected service time, using a confidence interval, and want to compare the mean to the bench mark for helpdesks.
In the table below you will find the observed service times (the order statistics) of 30 customers.

0.20	0.62	1.02	1.08	1.23	1.24	1.45	1.80	1.85	1.91
1.93	2.10	2.11	2.21	2.24	2.26	2.37	2.41	2.49	2.57
2.94	3.10	3.34	3.69	3.81	4.52	4.67	5.22	5.76	6.44

The numerical summary of the observations is given below (SPSS-output): note that the reported value of the kurtosis is, as usual in SPSS, "Kurtosis – 3".

Descriptive Statistics

	N	Mean	Std. Deviation	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Service Time (sec.)	30	2,6193	1,50601	2,268	0,912	0,427	0,486	0,833
Valid N (listwise)	30							

a. Compare the mean and the median: what does the difference tell you about the distribution of the service times?

b. Determine the 5-numbers-summary of the data set and use the quartiles to check whether there are outliers.

c. Assess the assumption of normality of the service times, using the numerical summary.

d. In the SPSS-output the value of Shapiro-Wilk's test statistic was reported: $W = 0.930$.
1. Give the null and the alternative hypothesis for this test.
2. Determine the coefficients $a_1$and a_{29} in the formula of Shapiro Wilk's W.
3. Determine the rejection region of the test.
4. What is your conclusion (in words) at a 5% significance level?

e. Determine a 95%-confidence interval for the expected service time (assuming normality) and give the proper interpretation of the numerical interval.

2. Do students at the UT in majority support the opinion that English should be the sole official language at their university. In a random sample of 200 UT-students 111 supported the opinion.

a. Conduct an appropriate test to verify whether the sample proves that a majority of all UT-students supports the opinion. Apply the testing procedure with $\alpha = 5\%$, using the **p-value** of the test.

b. Determine, additionally, the rejection region of the test in a.

c. Determine the power of the test in a. (*If you could not find the rejection region use* $\{112, 113, \dots\}$)

3. A computer scientist is investigating the usefulness of two different design languages in improving programming tasks. Sixteen expert programmers, familiar with both languages, are asked to code a standard

Expert	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	$\bar{x}$	s
Language 1	17	16	21	14	27	19	18	15	18	24	16	14	21	24	13	20	18.56	4.05
Language 2	18	14	19	11	18	18	13	12	23	21	10	14	19	23	12	17	16.38	4.18
Difference	**−1**	+2	+2	+3	+9	+1	+5	+3	**−5**	+3	+6	0	+2	+1	+1	+3	2.19	3.08

function in both languages, and the time (in minutes) is recorded. The data follow:

a. Assume for this part that normal distribution can be assumed for the observed times.
Conduct an appropriate (parametric) test in 8 steps at a 1% significance level, to check whether there is a structural difference in time needed to complete the task in language 1 and in language 2.

b. If the computer scientist considers it unreasonable to assume the normal distribution for the observed values, which test would you advise him as an alternative for the test in a.
Additionally, (only) define the test statistic of this test and give its observed value, the hypotheses and the distribution of the test statistic if H_0 is true.

4. Suppose we have two independent random samples, both with sample size 12.
We want to test whether there is a difference in the population means, but to apply the two independent samples t-test we need to check the assumption of normality for both samples and the assumption of equal variances.

a. Which test should we use to check the equality of variances?
Give for this test the rejection region if $\alpha = 5\%$ (use the given sample sizes 12).

b. For the test in **a.** a p-value of 1.2% is reported: what conclusion can you draw from that result?

c. Considering the result of the test on normality, it is obvious that the normal distribution does not apply for one of the samples. Which alternative non-parametric test can we apply here? Give the formula for the appropriate test statistic and determine the distribution under H_0 (if both sample sizes are 12).

5. The yearly return (in %) of state bonds is modelled as a normally distributed variable with unknown expected return μ and standard deviation $\sigma = \mu$. Based on a random sample of four returns X_1, X_2, X_3 and X_4 we consider two estimators of μ: $T_1 = \bar{X} = \frac{1}{4}\sum_{i=1}^{4} X_i$ and $T_2 = \frac{1}{5}\sum_{i=1}^{4} X_i$.

a. Are T_1 and T_2 unbiased estimators of μ? Motivate both answers.

b. Determine the mean squared errors of T_1 and T_2

c. For which values of n is T_2 a better estimator of μ than T_1?

Grade = $1 + \frac{\#\text{ points}}{45} \times 9$,

rounded at 1 decimal

1					2			3		4			5			Tot
a	b	c	d	e	a	b	c	a	b	a	b	c	a	b	c	
2	3	1	4	4	6	2	2	6	4	2	2	2	2	2	1	45

Formula sheet "Statistics for Engineers"

Rules Probability Theory: $var(X) = E(X^2) - (EX)^2$

$E(aX + b) = aE(X) + b$ and $var(aX + b) = a^2 var(X)$

$E(X \pm Y) = E(X) \pm E(Y)$

If X and Y are independent: $var(X \pm Y) = var(X) + var(Y)$

Bounds for Confidence Intervals:

* $\hat{p} \pm c\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$, with $\Phi(c) = 1 - \frac{1}{2}\alpha$
* $\overline{X} \pm c\dfrac{S}{\sqrt{n}}$, with $P(T_{n-1} \geq c) = \frac{1}{2}\alpha$
* $\left(\dfrac{(n-1)S^2}{c_2}, \dfrac{(n-1)S^2}{c_1}\right)$, with $P(\chi^2_{n-1} \leq c_1) = P(\chi^2_{n-1} \geq c_2) = \dfrac{\alpha}{2}$
* $\overline{X} - \overline{Y} \pm c\sqrt{S^2\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}$, with $S^2 = \dfrac{n_1 - 1}{n_1 + n_2 - 2}S_X^2 + \dfrac{n_2 - 1}{n_1 + n_2 - 2}S_Y^2$

  and $P\left(T_{n_1+n_2-2} \geq c\right) = \frac{1}{2}\alpha$

  or: $\overline{X} - \overline{Y} \pm c\sqrt{\dfrac{S_X^2}{n_1} + \dfrac{S_Y^2}{n_2}}$, with $\Phi(c) = 1 - \frac{1}{2}\alpha$
* $\hat{p}_1 - \hat{p}_2 \pm c\sqrt{\dfrac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \dfrac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$, with $\Phi(c) = 1 - \frac{1}{2}\alpha$

Testing procedure in 8 steps

1. Give a probability model of the observed values (the statistical assumptions).
2. State the null hypothesis and the alternative hypothesis, using parameters in the model.
3. Give the proper test statistic.
4. State the distribution of the test statistic if H_0 is true.
5. Compute (give) the observed value of the test statistic.
6. State the test and **a.** Determine the rejection region or
   b. Compute the p-value.
7. State your statistical conclusion: reject or fail to reject H_0 at the given significance level.
8. Draw the conclusion in words.

Test statistics and distributions under H_0:

* Binomial test: $X \sim B(n, p_0)$: $P(X = x) = \binom{n}{x} p_0^x (1 - p_0)^{n-x}$ or use the binomial table,

  or for large n approximately $N\left(np_0, np_0(1 - p_0)\right)$
* $T = \dfrac{\overline{X} - \mu_0}{S/\sqrt{n}} \sim t_{n-1}$

* S^2 , where $\dfrac{(n-1)S^2}{\sigma_0^2} \sim \chi^2_{n-1}$
* $T = \dfrac{(\overline{X}-\overline{Y})-\Delta_0}{\sqrt{S^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \sim t_{n_1+n_2-2}$ (and S^2 as given above) or $Z = \dfrac{(\overline{X}-\overline{Y})-\Delta_0}{\sqrt{\frac{S_X^2}{n_1}+\frac{S_Y^2}{n_2}}} \sim N(0,1)$
* $Z = \dfrac{\hat{p}_1-\hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \sim N(0,1)$, with $\hat{p} = \dfrac{X_1+X_2}{n_1+n_2}$
* $F = \dfrac{S_X^2}{S_Y^2} \sim F^{n_1-1}_{n_2-1}$

Analysis of categorical variables

* 1 row and k columns: $\chi^2 = \sum_{i=1}^{k} \dfrac{(N_i - E_0N_i)^2}{E_0N_i}$ $(df = k-1)$
* $r \times c$ – cross table: $\chi^2 = \sum_{j=1}^{c}\sum_{i=1}^{r} \dfrac{\left(N_{ij}-\hat{E}_0N_{ij}\right)^2}{\hat{E}_0N_{ij}}$, with $\hat{E}_0N_{ij} = \dfrac{\text{row total} \times \text{column total}}{n}$

  and $df = (r-1)(c-1)$.

Non-parametric tests

* Sign test: $X \sim B\left(n, \frac{1}{2}\right)$ under H_0
* Wilcoxon`s Rank sum test: $W = \sum_{i=1}^{n_1} R(X_i)$,

  under H_0 with: $E(W) = \dfrac{1}{2}n_1(N+1)$ and $var(W) = \dfrac{1}{12}n_1n_2(N+1)$

Test on the normal distribution

* Shapiro – Wilk`s test statistic: $W = \dfrac{\left(\sum_{i=1}^{n} a_iX_{(i)}\right)^2}{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}$