The University of The Gambia
School of Arts and Sciences

Division of Physical and Natural Sciences

Exam "Applied Statistics" Friday 1st June, 2018, 11.15-13.15 hrs in Faraba C3/4

This exam consists of 4 exercises. The formula sheet and tables are added separately.
A simple scientific calculator is allowed (and advised), but a mobile phone is not.

1. An insurance company can get damaged cars repaired at two different garages.
The company suspects that **garage 1 is more expensive than garage 2**.
To find out whether this is indeed the case, the damage of seven randomly chosen cars is estimated by both garage 1 and garage 2. The damage assessments (in hundreds of Euro´s) are as follows:

	Damage assessments (in 100 Euro)							$\bar{x}$	s^2
Garage 1	15.2	20.4	19.0	22.6	6.0	12.6	10.6	15.2	34.80
Garage 2	14.6	18.2	16.8	23.0	5.4	11.6	9.8	14.2	33.99
Difference	0.6	2.2	2.2	-0.4	0.6	1.0	0.8	1.0	0.87

Columns $\bar{x}$ and s^2 show the means and sample variances of the measurements in each row.

a. Carry out a **parametric** test to examine if garage 1 is, on average, more expensive than garage 2. Use the testing procedure with $\alpha = 0.05$.

b. Suppose that there are reasons to choose for a **non-parametric** test: which test should be carried out in this case?
Give (only): 1. the test statistic and its observed value,
2. the distribution under H_0,
3. the p-value and
4. the values of α (1%, 5% or 10%) for which the null hypothesis should be rejected.

2. Many people suffer from FNE (´Fear of Negative Evaluation´). To find out whether eating habits have any influence, a psychologist carries 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 questionnaire. 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	$\bar{x}_3 = 4.27, s_3 = 7.52$

Is there a difference between the expected FNE-scores of students with bulimia and students with normal eating habits? To answer this question, carry out a suitable statistical test. Use level of significance $\alpha = 5\%$ and give in step 1 all relevant statistical assumptions that are necessary for the chosen test.

3. A certain type of vitamin allegedly prevents colds. In an experiment, designed to verify the claim, 100 subjects are given the vitamin and 100 subjects are given a placebo. All 200 participants are told that they were given the vitamin. The results are as follows:

	Less colds	More colds	No difference
Control group	39	21	40
Treated group	51	20	29

a. Determine a 95%-confidence interval for the difference in proportions “Less colds” in the populations of treated and non-treated persons and give the proper **interpretation** of the interval.

b. Suppose we want to test the influence of the vitamin on susceptibility to colds. Should we apply, for these observations, a test on independence or a test on homogeneity?

c. Apply the test selected in b., but only give:

1. The test statistic and its observed value.
2. The rejection region for $\alpha = 10\%$.
3. Your conclusion with respect to the effectiveness of the vitamin in preventing colds,

4. An event occurs at a rate of p $(0 < p < 1)$: we define $X = 1$ if the event occurs and $X = 0$ otherwise: $P(X = 1) = p$ and $P(X = 0) = 1 - p$.
$x_1, \dots, x_{100}$ is a realization of a random sample $X_1, \dots, X_{100}$ of X.
For simplicity of notation we can use $x = \sum_{i=1}^{100} x_i$.

a. Show that $\hat{p} = \frac{1}{100}\sum_{i=1}^{100} X_i$ is the maximum likelihood estimator of p.

b. Show that $\hat{p}$ is a consistent estimator.

c. Show that the test that rejects $H_0: p = 0.2$ in favour of $H_1: p = 0.3$ for large values of $X = \sum_{i=1}^{100} X_i$ is most powerful, using Neyman-Pierson’s lemma (at given α).

d. Determine for the test in c. the significance level and the power of the test when we will reject H_0 if $X \geq 28$.

e. Find the uniformly most powerful test on $H_0: p = 0.2$ versus $H_1: p > 0.2$

-------------------------------- END---------------------------------------

Grade = $1 + \frac{\#\ of\ points}{41} \times 9$, rounded at one decimal

1		2	3			4					
a	b		a	b	c	a	b	c	d	e	Total
6	4	7	4	1	4	3	2	4	3	3	41

Solutions - Exercise 1

a. In this case, we are dealing with paired samples (the damage assessments pairwise concern the same levels of damage!) Let $X_1, X_2,..., X_7$ denote the seven differences (assessment garage 1 minus assessment garage 2). In this case, the eight steps are as follows:

(1) The differences $X_1, ... X_7$ are independent and all $N(\mu, \sigma^2)$-distributed, with unknown μ and σ^2.

(2) We test $H_0: \mu = 0$ against $H_1: \mu > 0$ with $\alpha = 5\%$

(3) Test statistic $T = \frac{\overline{X}}{S/\sqrt{7}}$.

(4) Under H_0, we have: $T \sim t_6$

(5) Observed value of the test statistic: $t = \frac{1}{\sqrt{0.87}/\sqrt{7}} \approx 2.85$.

(6) We reject H_0 if $T \geq c$, with $c = 1.943$ from the t_6-table such that $P(T_6 > c) = 0.05$

(7) Reject H_0, since 2.85 > 1.943.

(8) At a 5% level of significance we showed that garage 1 is structurally more expensive than garage 2.

Using the p-value: The p-value is $P(T_6 \geq 2.85)$. According to the t_6-table, it lies between 1% and 2.5% (interpolation yields 1.5%). So for $\alpha \geq 0.025$, we should reject H_0. (or for $\alpha \geq 0.015$ if you used linear interpolation.)

b. We apply the sign test:

- Test statistic: $X =$ "the number of positive differences in favour of garage 1". Observed value $X = 6$.
- X is $B\left(7, \frac{1}{2}\right)$ distributed under H_0.
- the p-value = $P(X \geq 6|H_0) = P\left(X = 6 \middle| p = \frac{1}{2}\right) + P\left(X = 7 \middle| p = \frac{1}{2}\right) = \binom{7}{6}\left(\frac{1}{2}\right)^7 + \left(\frac{1}{2}\right)^7 = 0.0625$
- So H_0 should only be rejected (garage 1 is structurally more expensive) if $\alpha = 0.10$

Exercise 2

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: "accidental" differences!*):

1. The 11 + 11 FNE-scores of the students who suffer from bulimia (sample 1) and the students who don't (sample 2) are independent. Sample 1 is taken from the $N(\mu_1, \sigma^2)$- distribution and sample 2 from the $N(\mu_2, \sigma^2)$ -distribution. (equal variances assumed)
2. We test $H_0: \mu_1 = \mu_2$ against $H_1: \mu_1 \neq \mu_2$ with $\alpha = 5\%$
3. Test statistic: $T = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{S^2\left(\frac{1}{11} + \frac{1}{11}\right)}}$ with $S^2 = \frac{10S_1^2 + 10S_2^2}{11+11-2}$
4. Under H_0, test statistic T has a t-distribution with 11 + 11 – 2 = 20 degrees of freedom
5. Observed value of the test statistic: $S^2 = \frac{1}{2}(4.92^2 + 5.34^2) \approx 26.361$, so $t = \frac{17.82 - 13.55}{\sqrt{26.361\left(\frac{1}{11} + \frac{1}{11}\right)}} \approx 1.95$
6. Two-sided test ("*show there is a difference*"): we reject H_0 if $T \leq -c$ or $T \geq c$

   Level of significance 5%, so use tail probability 2.5% in the t-table for $df = 20$, so $c = 2.086$
7. Result 1.95 does not lie in the rejection region $\Rightarrow$ H_0 is not rejected.
8. At a 5% significance level, we cannot prove that the FNE-scores of the two groups differ systematically.

Exercise 3

a. We are interested in the difference of two proportions, based on two independent samples:

We use the formula $\hat{p}_1 - \hat{p}_2 \pm c\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$ with $\Phi(c) = 1 - \frac{1}{2}\alpha$,

where $n_1 = n_2 = 100$, $\hat{p}_1 = \frac{39}{100}$ and $\hat{p}_2 = \frac{51}{100}$ and $c = 1.96$, such that $\Phi(c) = 0.975$

95%-CI($p_1 - p_2$) = (0.11 – 0.136, 0.11 + 0.136) = (-0.026, 0.246)
"We are 95% confident that the difference (treatment – placebo) in proportions is between -2.6% and +24.6%".

b. We want to compare two samples with respect to the distribution of a variable "colds", which attains the values "less colds" (1), "more colds" (2) and "no difference" (3) respectively. So choose for a test on homogeneity of the two distributions of colds, (the sample size were determined beforehand).

c. We define N_{ij} = "number of persons suffering from cold j" for the control group ($i = 1$) and the treated group ($i = 2$) respectively.

1. The test statistic is: $\chi^2 = \sum_{j=1}^{3}\sum_{i=1}^{2}\frac{(N_{ij}-\hat{E}_0N_{ij})^2}{\hat{E}_0N_{ij}}$ where $\hat{E}_0N_{ij} = \frac{row\ total \times column\ total}{n}$

   the values of $E_{ij} = \hat{E}_0N_{ij}$ are computed in the table:

	Less colds (1)	More colds (2)	No difference (3)	Total
Control group (1)	$\boldsymbol{N_{11}}$= 39, $\boldsymbol{E_{11}}$ = 45	$\boldsymbol{N_{12}}$= 21 , $\boldsymbol{E_{12}}$= 20.5	$\boldsymbol{N_{13}}$= 40, $\boldsymbol{E_{13}}$ = 34.5	100
Treated group (2)	$\boldsymbol{N_{21}}$= 51, $\boldsymbol{E_{21}}$ = 45	$\boldsymbol{N_{22}}$= 20, $\boldsymbol{E_{22}}$ = 20.5	$\boldsymbol{N_{23}}$= 29, $\boldsymbol{E_{23}}$= 34.5	100
Total	90	41	69	200

   Observed value: $\chi^2 = \frac{(39-45)^2}{45} + \cdots + \frac{(29-34.5)^2}{34.5} \approx 3.38$

2. We reject H_0 if $\chi^2 \geq c$. $\alpha = 0.10$, so from the χ^2_2-table, we obtain c = 4.61, such that $P(\chi^2_2 \geq c) = \alpha$
3. The observed value $\chi^2 \approx 3.38 < 5.99$, so we do not reject H_0: at a 5% significance level, we cannot prove that the vitamin has any effect on the susceptibility to colds.

Exercise 4

a. The likelihood function is $L(p) = \prod_{i=1}^{100} P(X_i = x_i|p) = p^{\sum x_i}(1-p)^{\sum(1-x_i)} = p^x(1-p)^{100-x}$
Log-likelihood: $\ln L(p) = x \ln p + (100 - x)\ln(1-p)$, where $0 < p < 1$

$$\frac{d}{dp}\ln L(p) = \frac{x}{p} - \frac{100-x}{1-p} = 0 \quad \text{if} \quad (1-p)x = (100-x)p \quad \text{or} \quad p = \frac{x}{100} = \frac{1}{100}\sum_{i=1}^{100} x_i$$

L attains is maximum at this value of p in $(0,1)$, since $\frac{d}{dp}\ln L(p) = -\frac{x}{p^2} - \frac{100-x}{(1-p)^2} < 0$ on $(0,1)$.

Hence the *mle* of p is $\hat{p} = \frac{1}{100}\sum_{i=1}^{100} X_i$.

b. We know that $E(\hat{p}) = p$ and $var(\hat{p}) = \frac{p(1-p)}{n}$, so that $MSE(\hat{p}) = (E(\hat{p}) - p)^2 + var(\hat{p}) = 0 + \frac{p(1-p)}{n}$
Since $\lim_{n\to\infty} MSE(\hat{p}) = 0$, $\hat{p}$ is a consistent estimator of p

c. Neymann-Pierson's ratio:

$$r(x_1,\dots,x_{100}) = \frac{L(0.2)}{L(0.3)} = \frac{0.2^x 0.8^{100-x}}{0.3^x 0.7^{100-x}} = \left(\frac{8}{7}\right)^{100}\left(\frac{2}{3}\right)^x\left(\frac{8}{7}\right)^{-x} = \left(\frac{8}{7}\right)^{100}\left(\frac{2}{3}\times\frac{7}{8}\right)^x = \left(\frac{8}{7}\right)^{100}\left(\frac{14}{24}\right)^x$$

r is a decreasing function in $x = \sum_{i=1}^{100} x_i$. Hence rejecting H_0 for small values of $r(X_1,\dots,X_{100})$ is equivalent to rejecting H_0 for large values of $X = \sum_{i=1}^{100} X_i$: then the test that rejects H_0 for large values of X is according to Neyman Pierson's lemma (for given α) most powerful.

d. X is under H_0 approximately $N(np_0, np(1-p_0)) = N(20, 16)$ and under H_1: $p = 0.3$ $N(30, 21)$

$$\alpha = P(X \geq 28|H_0) = P(X \geq 27.5|p = 0.2) \approx P\left(Z \geq \frac{27.5-20}{4}\right) \approx 1 - \Phi(1.88) = 3.01\%$$

$$\beta(0.3) = P(X \geq 28|p = 0.3) = P(X \geq 27.5|p = 0.3) \approx P\left(Z \geq \frac{27.5-30}{\sqrt{100 \cdot 0.3 \cdot 0.7}}\right) \approx \Phi(0.59) \approx 0.722$$

e. The most powerful test on $H_0: p = 0.2$ versus $H_1: p = p_1$, where $p_1 > 0.2$ rejects H_0 for small

$$r(x_1,\dots,x_{100}) = \frac{L(0.2)}{L(p_1)} = \frac{0.2^x 0.8^{100-x}}{p_1^x(1-p_1)^{100-x}} = \left(\frac{8}{1-p_1}\right)^{100}\left(\frac{2}{p_1}\times\frac{(1-p_1)}{8}\right)^x = \left(\frac{8}{1-p_1}\right)^{100}\left(\frac{1-p_1}{4p_1}\right)^x$$

Since $\frac{1-p_1}{4p_1} < 1 \Leftrightarrow 1 - p_1 < 4p_1 \Leftrightarrow \frac{1}{5} < p_1$ (is what we assumed), $r(X_1,\dots,X_{100})$ attains small values for large values of $X = \sum X_i$: choosing the rejection region $X \geq c$ such that $P(X \geq c|H_0) = \alpha$ for all $p_1 > 0.2$, we found the uniformly most powerful test.