



# 1













Ens. : Michael Herzog  
 Exam: Understanding Statistics Experimental Design - MAN  
 Date: 15th Jan. 2020  
 Time : 12:15 - 14:30 (135 minutes)

## Student One

SCIPER: 111111

Do not turn the page before the start of the exam. This document is double-sided, has 10 pages, the last ones possibly blank. Do not unstaple.

- Place your student card on your table.
- **No other paper materials** are allowed to be used during the exam.
- Using a **calculator** or any other electronic device is not permitted during the exam.
- Make sure your mobile phone is switched off and not on the table.
- In **Part I** (one choice questions), we give :
  - +2 points if your answer is correct,
  - 0 points if your answer is incorrect or if more than 1 answer is ticked.
- In **Part II** (multiple choices questions), we give :
  - +1 point each for your correct tick (Hit) and correct non-tick (correct rejection).
- Use a **black or dark blue ballpen** and clearly erase with **correction fluid** if necessary.
- If a question is wrong, the teacher may decide to nullify it.

Respectez les consignes suivantes   Observe this guidelines   Beachten Sie bitte die unten stehenden Richtlinien		
choisir une réponse   select an answer Antwort auswählen	ne PAS choisir une réponse   NOT select an answer NICHT Antwort auswählen	Corriger une réponse   Correct an answer Antwort korrigieren
  		 
ce qu'il ne faut <b>PAS</b> faire   what should <b>NOT</b> be done   was man <b>NICHT</b> tun sollte		
     		



## Part I: one choice questions

In part I, there is only **ONE** correct response for each question, which you should tick. A correct answer earns you 2 points.

**Question 1** A recent study found that 25 out of 100 students who joined a statistics class via zoom suffered from eye itches, determined on a ratio scale. This was true for another 5 out of 25 students who joined the lecture in the class room. Students were free to choose whether or not to join via zoom. What is the Odds Ratio for this example?

- 0.083
- None of them
- 0.750
- 1.250
- 0.166
- 0.450
- 1.333

**Question 2** You want to find out whether there is a significant difference between the two groups in the example in Question 1. Tick the most appropriate test, taking into account that the variances between the groups are different.

- 2-way ANOVA with 1 level
- Two sample dependent t-test
- None of them
- Welch's t-test
- Two sample independent t-test
- One sample t-test

**Question 3** Related to Question 1, which one is the true statement?

- We need to take the incidence rate into account to make a proper conclusion
- None of them
- There is a multiple testing problem
- The sample may not be i.i.d

**Question 4** The researchers asked the students in Question 1 three more questions to get a better overview of the sample, namely about their gender (male, female, other), their academic status (BA, MA, PhD), their body height (*cm*), and, as mentioned, whether or not they were in the class room. They want to test how all these factors determine eye itch strength. What type of statistical test(s) is/are best to use?

- 4 independent sample t-tests with Bonferroni correction
- 4 dependent sample t-tests with Bonferroni correction
- 4 - way repeated measures ANOVA
- None of them
- 4 - way independent measures ANOVA
- Pearson correlation



**Question 5** Which one is the correct equation for the empirical  $d'$  as used in SDT?

- $(\mu_1 - \mu_2)/\sigma$
- $(z(\mu_1) - z(\mu_2))/\sigma$
- $z(Hit) - z(FA)$
- $-[z(Hit) + z(FA)]/2$
- None of them
- $z(Hit) - z(CR)$
- $(Hit + CR)/2$

**Question 6** You are comparing two treatments against eye itches. You take two samples of 18 people each and measure the strength of eye itch pain. Calculate the t-value for an independent samples t-test with an estimated difference of means of 0.2 and an estimated standard deviation of 0.10.

- 6
- 8.49
- 2.4
- 2
- 4.47
- None of them

**Question 7** Which of the equations below is the equation for the standard error of the mean with standard deviation  $\sigma$  and sample size  $n$ ?

- $\frac{\sigma}{\sqrt{n-1}}$
- $\sigma\sqrt{\frac{n-1}{n}}$
- $\sigma\sqrt{\frac{2}{n}}$
- None of them
- $\frac{\sigma}{\sqrt{n}}$
- $\sigma\sqrt{\frac{1}{n(n-1)}}$

**Question 8** An experiment delivered a significant result with a p-value of 0.049. What can you tell about the effect size in terms of *Cohen's d* given Cohen's classification of low, medium, and high effects sizes?

- It is high because the result is significant
- It is low
- It is medium high because the result is barely significant
- None of them



**Question 9** An experiment delivered a significant result with a p-value of 0.049. Which of the following statements is true?

- The effect size is lower than in an experiment with the same number of participants and a p-value of 0.051
- The effect size is higher than in an experiment with the same number of participants and a p-value of 0.0049
- None of them
- The effect size is higher than in an experiment with the same number of participants and a p-value of 0.051

**Question 10** The table below shows the outcome of an experiment, which had something to do with eye itches. How many participants joined the experiment?

Source	Sum of Squares (SS)	Degrees of freedom (df)	Mean of the Squares (MS)	F-value	p-value
Between	546.50	5	109.30	5.43	< 0.05
Within	483.12	24	20.13		

- 29
- 28
- 24
- 21
- 30
- 4

**Question 11** The following table lists the details of 5 experiments. Each experiment used a two-tailed t-test with  $\alpha = 0.05$  to determine if there was a significant difference between 2 population means. N is the sample size for each group. What is the correct ranking of power of the experiment in descending order?

Experiment	N	Mean <sub>Group1</sub>	Mean <sub>Group2</sub>	$\hat{\sigma}$
A	10	25	30	6
B	10	25	30	8
C	50	25	30	8
D	100	25	30	8
E	100	25	55	8

- None of them
- E>D>C>A>B
- E>D>C>B>A
- D>E>C>B>A
- E>C>D>B>A
- C>D>E>A>B
- E>D>A>C>B



**Question 12** A vendor offers a test to determine whether or not you are a statistics genius. The test has a sensitivity of 50% and a specificity of 50%. You take the test and it is positive. How likely (in percent) is it that you are a stats genius taking into account that geniuses are rare, namely, 1%?

- 50%
- 0.1%
- This is partial information
- 5%
- 1%

**Question 13** You run an experiment where the data is analysed with a 2x2 ANOVA. Which of the following analyses has the *highest* power?

- Determining a main effect of each factor and a post-hoc comparison of three means
- Determining the effects of each factor and a significant interaction
- Determining a main effect of factor 1
- Determining a main effect of factor 1 and a significant main effect of factor 2

**Question 14** A large-scale meta-analysis shows that 85 out of 100 investigations reported a significant impact of Visine eye drops on eye itchiness. If the studies were done properly (e.g., no optional stopping, no publication bias), what would you estimate as the typical (or average) power of these studies?

- Around 85%
- Around 70%
- Close to 100%, because everyone knows that Visine eye drops actually do work
- Around 50%
- Around 10%

**Question 15** What is the strongest argument for creating a preregistration document for a study?

- It is the only way to justify your analysis and sample size
- It allows you to interpret a null finding
- It prevents critics from attacking you
- It lends credence to your predicted outcome regardless of the source of the prediction
- It motivates you to think careful about your analysis and sample size



## Part II: multiple choice questions

In part II, each response may or may not be correct. However, at least one response is correct. **Tick all** correct response alternatives and leave incorrect ones blank. You will get 1 point for each correct tick (Hit) and each correct non-tick (Correct rejection).

**Question 16** What type(s) of scale(s) can be used for the independent variable when *Pearson's r* is computed?

- Nominal
- Interval
- Ordinal
- Ratio

**Question 17** Which of Cohen's d value(s) is (are) counted as large effect(s)?

- 0.5
- 1.0
- 0.7
- 1.5

**Question 18** Which of the following is (are) estimate (s) of the effect size?

- Hedge's g
- $\eta^2$  (*Eta*<sup>2</sup>)
- $R^2$
- Cohen's d

**Question 19** In an exploratory study, researchers determined eye itches of 100 students before and after taking a statistics class. A subgroup of 47 students had eye itches at the beginning whereas the rest had not. All 100 students joined the same statistics class for 14 weeks. The itches became significantly worse in the first group during the 14 weeks as shown by a dependent two sample t-test. No significant change was found for the other group of 53 students as shown by a one sample t-test based on the difference of pre- vs. post-itch ratings. The authors concluded that students with eye itches should not take statistics classes. Which of the following statement(s) is (are) true?

- One should have used an independent t-test between the two groups using the pre/post-differences of each participant
- It is plain wrong to use two different t-tests for the different groups
- There is HARKing
- There is a multiple comparison problem
- The conclusion is not supported by the data



**Question 20** A study wants to investigate whether eye itches change in a sinusoidal fashion during a statistics lecture. As a marker for the intensity of the eye itches, pupil size was measured in 100 students 6 times during the class of 90min and 6 times during 90min of leisure time. How to best compute statistics?

- For each participant, fit a sine function to the class and leisure time conditions, extract the phases, take the difference of phases, and use a one sample t-test to test whether this value is significantly different from zero
- For each participant, fit a sine function to the class and leisure time conditions, extract the amplitudes, and compute a dependent t-test to test whether there is significant difference
- None of them
- For each participant, fit a cosine function to the class and leisure time conditions, extract the amplitudes, take the difference of the amplitudes, and use a one sample t-test to test whether this value is significantly different from zero

**Question 21** In an exploratory study, EPFL wants to improve the online lecture system, and thus wants to learn how well students coped with the current system. Researchers randomly sampled 100 students each from Bachelor (who had on-campus lectures) and Master (who had only online lectures) classes, and measured 1) eye itchiness, 2) blood cortisol level (a stress indicator), 3) alcohol consumption per week, and 4) grades. Researchers found that 4) the grades between the two groups were significant, and 2) the cortisol level was significantly different between groups when using a 1-tailed t-test, after a 2-tailed t-test did not reach significance. 1) eye itchiness was almost significant with a 1-tailed t-test ( $p=0.053$ ), but there was no significance in 3) alcohol consumption levels ( $p=0.1$ ). To improve effect size, sample size was increased by 240 students (120 each). Researchers concluded that 'For 240 randomly sampled students, where half had on-campus courses and the other half only had online courses, students who only took online courses had suffered physically (1-tailed t-test on eye itches,  $t(119)=1.65$   $p<0.05$ ) and mentally based cortisol level (1-tailed t-test,  $t(119)=1.68$ ,  $p<0.05$ ) and also their grades deteriorated drastically (2-tailed t-test,  $t(119)=2.85$ ,  $p<0.01$ ). The results show the clear need to improve the online lecture system, hence, we request 100,000 CHF as a budget for changes.

What types of bad statistical practice are present in the above example?

- HARKing
- Optional stopping
- File Drawer
- Multiple testing
- Confusion of significance and effect size

**Question 22** Which of the following statements about the criterion in SDT is (are) correct?

- It depends on the incidence rate
- The criterion may reflect the response strategy of participants
- None of them
- The criterion changes when lowering the standard deviation
- If sensitivity and specificity are equal, the criterion is optimal



**Question 23** Power in a two independent samples t-test can be increased by:

- Decreasing the significance level
- Decreasing the standard deviation
- Increasing sample size
- Increasing the significance level
- None of them

**Question 24** Which of the following statements is (are) correct about the p-value?

- The p-value is partial information
- It gives you the probability that the result observed is due to chance
- $1-p$  is power
- None of them
- A non-significant p-value tells you that there is evidence for  $H_0$

**Question 25** Which of the following assumptions must be met to perform an independent ANOVA?

- Homogeneity of variance
- Normally distributed populations
- Independent samples
- At least interval scale for the independent variables

**Question 26** Which statements about PCA is (are) correct?

- PCA reduces the overall variance in the data
- A scree plot can be used to decide whether factors may be neglected
- None of them
- If the PCA takes  $n$  variables, then there are  $n-1$  different eigenvalues
- If the PCA takes  $n$  variables, then there are  $n$  different eigenvalues

**Question 27** Which statements about the Bayes approach is (are) true?

- Bayes is optimal in terms of SDT
- Choosing the priors is a problem in the Bayes Approach
- The Bayes approach can provide evidence for the  $H_0$  as well as the  $H_1$
- The incidence rates of the  $H_0$  and  $H_1$  are important



PROJET



PROJET