university of groningen

faculty of behavioural and social sciences

Statistics II

Module code: PSBE2-07

2019-2020

Reader

Casper Albers & Jorge Tendeiro

Contents

Course information

The course Statistics II covers the entire first semester (blocks 1a and 1b) and consists of lectures (two hours per week) and practicals (two hours per week). The course is concluded by means of two partial exams (or one final resit exam), conditional on meeting the mandatory requirements concerning the practicals (see below). The exam will consist exclusively of multiple choice questions.

Aim of the course

After successful completion of the course the student is expected to be able to apply a variety of statistical methods.

The course focuses first and foremost on the theory behind these techniques: How does a technique work, what are its properties and background, when is a technique suitable or not, and what are the limitations of the technique.

In addition, students will learn how to apply these techniques to data from research on the social sciences: How to carry out analyses both manually and using statistical software (JASP or, if desired, R), how to generate relevant output (tables, figures), how to interpret such output, and how to draw valid conclusions.

Statistics II builds and elaborates on the concepts studied in Statistics Ia (PSBE1-05) and Statistics Ib (PSBE1-08). Confidence intervals and tests of significance will in particular play a key role in this course. Active knowledge on these topics is essential for success in Statistics II. Therefore, an overview lecture covering these topics will take place during the introduction week.

Statistics II forms an essential basis for Statistics III. Also, courses such as Research Methods and the Bachelor Thesis will be better approached upon a successful completion of Statistics II.

Prerequisites

Participation in Statistics II is only possible for students who passed both Statistic Ia and Ib, or for students who passed one of those courses and received a grade of at least 5 on the other course.
Statistics II is itself an entry requirement for Statistics III (PSBE2-12; requirement: Passing the practical requirement of Statistics II), Research Practicum (PSBE2-09; requirement: Passing the practical requirement of Statistics II), and the Bachelor Thesis (PSBE3-BT; requirement: Passing Statistics II completely). Please consult the Education and Examination Rules for further and official details.

The courses Statistics III and Research Practicum are both planned in the second semester of this year. Being eligible for participation in these courses depends on passing the practical requirement of Statistics II. In general, students know around Christmas whether they have passed or failed the practical and thus know a month in advance whether they can participate in either of these courses.
For the third-year Bachelor Thesis the requirement is that you must have passed Statistics II completely. *Only* the first opportunity of the exam takes place this semester, since the resit opportunity is after 1 February. **If you are a third year student aiming to write your Bachelor Thesis the second semester it is imperative that you pass this year's Statistics II on the first opportunity**. If you pass Statistics II in the resit, the earliest opportunity to participate in the Bachelor Thesis is Semester 1 of next year.

Course requirements

The requirements to pass Statistics II are based on the practicals and the exam. Participation in the exam is only allowed to students who meet the practical requirement. **Failing the practical requirement will lead to failing the course in the current academic year**. Students who fulfilled the practical requirement during the last academic year are now exempt, see section "Information for students retaking the course." The practical requirement has one component: **Compulsory attendance** at the practicals, outlined below. You can follow your own progress for this requirement on Nestor and you are advised to do so regularly.

Literature

- A. Agresti (2018). *Statistical Methods for the Social Sciences.* Fifth Edition (Global Edition). ISBN 978-1-292-22031-7.
- Kruschke, J. K. & Liddell, T. M. (2018). Bayesian data analysis for newcomers. *Psychonomic Bulletin & Review*, *25*, 155-177. doi:10.3758/s13423-017-1272-1.
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2001). False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. *Psychological Science*, *22*, 1359-1366. doi:10.1177/0956797611417632.
- John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the Prevalence of Questionable Research Practices With Incentives for Truth Telling. *Psychological Science*, *23*, 524-532. doi:10.1177/0956797611430953.
- All material that has been or will be placed on Nestor. Most notably, the reader (including the extra text on correlation), the practical exercises and the lecture slides.

From the book by Agresti, mainly chapters 9 – 12 are the material to be covered in this course. Please consult the table at the end of these introductory pages to see how these chapters are divided over the semester. All literature is exam material.

Software

The practical exercises can be made using either JASP or R. In either case, a practical reader will be made available on Nestor. Both software options are available through the UWP[1].

JASP is freely available from its website[2], in versions for Windows, Mac OS, and Linux. JASP is free, and will stay free, as it is published under an Open Source licence. The JASP developers have a number of short video tutorials on their website, https://jasp-stats.org/how-to-use-jasp/, explaining, e.g., how to do a *t*-test using JASP. For most of this course's purposes JASP is ready to use.

Students who are considering a career in academia are recommended to use R[3] during the practical sessions. Like JASP, R is free and runs in all three major platforms. R might be more difficult to learn than JASP, but it has many more features for advanced statistical analysis (that are out of the scope of the course Statistics 2; but might be inside your future research scope). There will be some practical groups dedicated to R: The assistant will provide assistance for this software and students are expected to use R. When enrolling for practical groups, it will be noted which groups are dedicated to R. **Students in this group are encouraged to take an R-tutorial (e.g., http://www.r-tutorial.nl) *before* the first practical session**. There will be no assistance on R in other practical groups.

Course coordination

Casper Albers (PSBA2-07)	*Jorge Tendeiro* (PSBE2-07)	*Karin Siebenga*
Tel. 050 36 38239	Tel. 050 36 36953	Tel. 050 36 39358
c.j.albers@rug.nl	j.n.tendeiro@rug.nl	k.siebenga@rug.nl
Room H.231	Room H.206	Room H.214
Lectures/coordination PSBA2-07	Lectures/coordination PSBE2-07	Practical coordination

Please contact the **practical coordinator (Karin Siebenga)** by email for any questions concerning practicals.

Questions concerning the course material may be posed on Nestor, to the assistants during practicals, or to the lecturer before/during/after the lecture. Additionally, students can make an appointment with the lecturer. **Students who email the lecturer with questions on the course contents will not be served and such emails will typically not be replied to.** Use Nestor for these queries. Also, please check before asking your question whether it has not already been asked at either the discussion board or is listed in the Frequently Asked Questions-list a few pages further on in this reader.

[1] UWP: http://uwp.rug.nl. You can log in using your student number and UG password.
[2] JASP: https://jasp-stats.org/
[3] R: http://www.r-project.org/

Lectures

There will be 14 lectures, seven per block. Attending the lectures is most beneficial when you come prepared and have read through the corresponding sections in the advised reading material. During lectures, focus lies on explaining the theory and the interpretation of results. There is also the opportunity to ask questions. Attending the lectures is not compulsory but **anything discussed in the lectures is exam material**. A refresher course lecture will take place on the introduction week of the academic year.

Practical

The lectures are accompanied by practicals in order to learn how to apply statistical theory to practical situations. There are 12 practicals in total. There are no practicals for the introductory lecture nor the overview lecture.

The aim of the practicals is:

1. To increase the insight in the theory of statistical methods.

2. To be able to apply statistical methods to empirical data (existing studies).

3. To be able to decide what statistical method to use in order to answer a particular question.

4. To be able to work with statistical software independently.

The practicals take place in groups of around 20 students. Students need to enroll for a practical group. Please observe that **signing up for the course is not sufficient to be able to participate in the practicals**. Enrollment in practical groups is going to take place on Nestor from **August 30** through **September 11**.

Students should prepare for the practicals in advance, otherwise they will not be able to complete them within two hours. Students should study the advised material before attending the practical. Also, **being present and actively participating in the practical activities is required**.

During the practicals, students shall work on the assignments listed in the syllabus. Computers will be available to everyone but students may choose to work with their own laptop, if preferred. Students are encouraged to work in teams of two on the assignments. Each practical is accompanied by an assistant who is available to answer questions in case of need.

Some practical groups will be dedicated to working with R. Enrollment for these groups is at a first come, first serve basis, but we will strive to accommodate all interested students. All other groups will be dedicated to working with JASP.

Attendance requirement

There is an attendance requirement. Students are expected **to be actively present in all practicals**. In order to accommodate for unforeseen absences (due to force majeure, such as illness), the attendance requirement stipulates that students need to attend **at least ten of the practicals** that are scheduled to take place in the class room. **Attending less than ten practicals implies failing this requirement and, thus, failing the course.** Students do **not** have to email or call in order to report a one-time absence since that does not conflict with meeting the attendance requirement. However, students who find it impossible to meet the attendance requirement must warn the practicals coordinator **in advance** (Karin Siebenga: k.siebenga@rug.nl). Only exceptional force majeure reasons may be invoked. The course coordinator, and possibly the Examinations Committee, will then decide whether and how exceptions to the attendance requirement will be granted.

It is the student's responsibility to weekly monitor his/her practical attendances. The practical attendance lists are updated weekly on Nestor, under 'My Grades'. Do not wait until the end of the semester but check weekly whether your attendance has been recorded correctly.

Exam

The exam determines the final grade. Participation in the exam is only possible after meeting the practicals requirement. There are two partial exams:

1. Partial 1, after block 1a (up to Lecture 7). This partial covers the following topics: Simple linear regression (estimation, inference); model validity, causality and association; multiple regression (interaction effects, partial correlation, standardized regression), and assumptions. See the course schedule for a detailed overview of the literature for this partial exam.

2. Partial 2, after block 1b. This partial covers all exam material, with emphasis on the topics not yet covered in partial 1: Regression with categorical predictors; multiple comparisons and contrasts; one-way ANOVA; two-way ANOVA; introduction to Bayesian statistics; questionable research practices and the replication crisis.

Each partial exam has the duration of 1 hour and consists of **18 multiple-choice** questions. **No grades will be released after the first partial exam.** The reason is that at that moment the exam is half-way; grades can only be computed after the second partial exam. We will, however, release the total number of correct answers for the first partial exam, so you will have a fair idea of how well you did.
The resit exam will focus on all material, will last 2 hours, and consist of **36 multiple-choice** questions. The exams are closed book. This syllabus shows the formula sheet that will be provided during the exam. All other formulas are to be known unless stated otherwise during the lectures. Registration for the exam is done as usual via Progress. The same rules apply to the resit.

Information for students retaking the course

Literature

This year the course underwent a major update. There is new literature, lecture slides, practical labs, and course contents (in particular the inclusion of Bayesian statistics). Students are therefore **strongly adviced** to keep close track of the course's progression in order to be updated.

Practicals

The practical requirement continues to be valid for one year. That is, students who have met the practical requirement during the last academic year are now exempt from the practicals in the current year.
Information about students who are exempt from practicals in this year's course will be made available on Nestor in late August/early September. Students retaking the course are **strongly advised** to check on time whether they are exempt from the practical in order to prevent misunderstandings.

Students who passed the practical requirement during last academic year have two options for the current academic year:

1. Not enrolling for the practicals on Nestor; therefore not attending practical sessions. This is considered the default option. Students are therefore expected to work on all the practical exercises on their own.

2. Enrolling for the practicals in Nestor (before the deadline). By doing this students automatically **wave away the 'Pass' from last year**. This means that such students must again fulfill the practicals requirement in order to pass the course. Students thus risk losing their access to the exam. These students must email the practical coordinator (Karin Siebenga: k.siebenga@rug.nl; deadline: Tuesday, September 10, 23:59) to inform her about their decision.

Information during the course

Additional information about the course and additional material is to be distributed via Nestor. Students are therefore advised to check Nestor's announcements regularly. Students are also encouraged to use the Discussion Board to ask (and answer!) questions. Previous experience shows that the Discussion Board is very useful to exchange opinions among the whole course community. The lecturer will moderate the discussions, check whether the answers provided by fellow-students are correct, and provide his own view if it is needed. Please avoid using the Discussion Board to address the lecturer directly. The lecturer will keep a bird's eye view on the Discussion Board. So, ideally, the Discussion Board is an area where students can ask and answer each other's questions, with the advantage that the lecturer will stop by regularly to make sure that no wrong explanations are being conveyed.

Credits

The material for this course has been put together over the years by various people from the Psychometrics and Statistics group at the Heymans Institute. Most notable contributors are Casper Albers, Jorge Tendeiro, and Mark Huisman.

Course schedule

Lecture	Date	Reading material	Topic
0	06 Sep.	4 - 7	Refresher Statistics I
1	09 Sep.	9.1 - 9.4	Simple linear estimation: Estimation
2	16 Sep.	9.5, A1	Simple linear estimation: Inference
3	23 Sep.	9.6, 10	Model validity. Causality & Association
4	30 Sep.	11.1 - 11.3	Multiple regression
5	07 Oct.	11.4 - 11.5	Multiple regression: Interaction effects
6	14 Oct.	11.6 - 11.7	M.R.: Partial correlation, standardized regression
7	21 Oct.	—	Assumptions
First partial exam: 08 Nov., 18:45-19:45			
8	11 Nov.	12.1	Regression with Categorical Predictors
9	18 Nov.	12.2	Multiple Comparisons and Contrasts
10	25 Nov.	12.3	ANOVA, one-way
11	02 Dec.	12.4	ANOVA, two-way
12	09 Dec.	A2	Introduction to Bayesian statistics
13	16 Dec.	A3	Good statistics, bad statistics
14	06 Jan.	—	Overview
Second partial exam: 20 Jan., 12:15-13:15			
Resit exam: 06 Apr., 12:15-14:15			

All sections refer to the book by Agresti. A1: Additional text 'Inference for Correlations', in this reader. A2: Kruschke & Liddell (2018). A3: Simmons et al. (2011); John et al. (2012).

Formula sheet

At the exam, you will receive this formula sheet.

Pooled variance for i groups

$$s_p^2 = \frac{\sum_i (n_i - 1)s_i^2}{\sum_i (n_i - 1)}$$

Confidence interval for μ

$$\overline{y} \pm t^* \frac{s}{\sqrt{n}}.$$

t-test for H: $\mu_1 = \mu_2$

Test statistic:

$$t = \frac{\bar{y}_1 - \bar{y}_2}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}.$$

Test for H: $\rho = 0$

Test statistic:

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}.$$

Contrasts

Sample estimation:

$$c = \sum_i a_i \bar{x}_i$$

Standard error:

$$\mathrm{SE}_c = s_p \sqrt{\sum_i \frac{a_i^2}{n_i}}.$$

Fisher Z-transformation

Transformation:

$$r_Z = \frac{1}{2} \ln\left(\frac{1+r}{1-r}\right).$$

Inverse transformation:

$$r = \frac{e^{2r_Z} - 1}{e^{2r_Z} + 1}.$$

(Semi-)partial correlations

Formula's valid when working with DV y and two predictors.

$$pr_1 = \frac{r_{y1} - r_{y2}r_{12}}{\sqrt{(1-r_{y2}^2)(1-r_{12}^2)}} = \sqrt{\frac{R^2 - r_{y2}^2}{1 - r_{y2}^2}}$$

$$sr_1 = \frac{r_{y1} - r_{y2}r_{12}}{\sqrt{1-r_{12}^2}} = \sqrt{R^2 - r_{y2}^2}$$

Adjusted R^2

$$R^2_{\text{adj}} = R^2 - \frac{p}{n-p-1}\left(1 - R^2\right).$$

Effect sizes

$$\eta_p^2 = \frac{SS_{\text{effect}}}{SS_{\text{effect}} + SS_{\text{error}}}, \quad \omega^2 = \frac{SS_{\text{effect}} - df_{\text{effect}} \times MSE}{MSE + SS_{\text{total}}}$$

Binomial model

$$p(X = x | N, \theta) = \binom{N}{x} \theta^x (1-\theta)^{N-x}, \quad \binom{N}{x} = \frac{N!}{(N-x)!x!}$$

FAQ

Course requirements

Q: *What do I need to do in order to pass this course?*
A: You must (i) be enrolled for the course and the exams via ProgressWWW, (ii) enroll for a practical group on Nestor (unless you passed the practical requirement last year), (iii) attend at least **ten** practicals in your own practical group, (iv) pass the regular exam (or the resit if you take it).

Q: *I am retaking the course. Do I have exemption from the practical requirement?*
A: Did you fulfill the practicals requirement last year? If so, then you are exempt. Please do confirm that this is the case by checking the information on Nestor.

Q: *I have practical exemption. What do I need to do in order to pass this course?*
A: You must (i) be enrolled for the course and the exams via ProgressWWW and (ii) pass the regular exam (or the resit if you take it).

Q: *I missed the first partial exam, can I still participate in the second partial exam?*
A: Yes, in theory you can. However, if you score all questions in the second partial correctly, you will have a total score of 50%, which is insufficient to pass the course.

Prerequisites

Q: *What do I need to do in order to be able to enroll in Statistics III (PSBE2-12) and/or Research Practicum (PSBE2-09)?*
A: You must have passed the practical requirement for Statistics II.

Q: *What do I need to do in order to be able to enroll in the Bachelor thesis in the next semester?*
A: You must pass Statistics II **during the exam opportunity in January**.

Q: *If I pass Statistics II in the resit, may I still enroll in the Bachelor thesis next semester?*
A: No.

Other

Q: *Will lecture recordings be uploaded to Nestor?*
A: This year the aim is to make the recordings available about one week before each partial exam. However, we cannot guarantee that all lectures will be recorded properly. Lectures that take place outside university buildings, e.g. in the Stadskerk or at the Pathe Cinema, are not always recorded. Furthermore, sometimes recordings have unforeseen technical problems (e.g. no sound). Moreover, recordings are not meant to replace attending lectures; instead, they are meant to complement the lectures. You are thus strongly advised to attend the lectures.

Q: *Can I attend the lectures of the other course?*
A: Apart from the language, PSBA2-07 and PSBE2-07 are 100% equivalent. If you prefer, because you wish to hear the material twice, or because it fits better with your personal schedule, you are free to attend the lectures of the course in the other language (see Ocasys for schedule).

Q: *Can I make audio/video recordings of lectures/practicals for private use?*
A: No, you are not allowed to make recordings of lectures and/or practicals, even if it is only for personal use. Unauthorised recording will be reported to the Examinations Committee and/or Legal Affairs, who might take appropriate action.

Additional reading material

In addition to the textbook, the following is reading (and exam) material as well:

- All lecture slides. Will be uploaded to Nestor as separate document(s).
- All practical labs. Will be uploaded to Nestor as separate document(s).
- Kruschke, J. K. & Liddell, T. M. (2018). Bayesian data analysis for newcomers. *Psychonomic Bulletin & Review*, *25*, 155-177. doi:10.3758/s13423-017-1272-1.
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2001). False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. *Psychological Science*, *22*, 1359-1366. doi:10.1177/0956797611417632.
- John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the Prevalence of Questionable Research Practices With Incentives for Truth Telling. *Psychological Science*, *23*, 524-532. doi:10.1177/0956797611430953.

All the articles above can be downloaded for free from the university network.

- Albers, C. J. (2015). Inference for correlations. Can be found in the following few pages.

Inference for correlations

Casper Albers[*]

Last update: October 17, 2016

Introduction. In the first year, it was taught how to interpret the correlation coefficient ρ and how to compute the sample correlation coefficient r. In short, ρ measures the strength and direction of the linear relation ('co-relation') between two variables (x and y, say). No relation corresponds to $\rho = 0$, a perfect positive relation to $\rho = 1$, and a perfect negative relation to $\rho = -1$. For a given sample of paired observations x_i and y_i ($i = 1, \ldots, n$), the sample correlation coefficient r is computed via

$$r = \frac{1}{n-1} \sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{s_x} \cdot \frac{y_i - \bar{y}}{s_y} \right).$$

In case you need to refresh your memory on the computation and interpretation of correlations, please reread your first year material on these topics (e.g. Moore, McCabe & Craig, Section 2.3). You will need a good understanding of these basics in order to understand the following.

In this text, the basics of inference for correlation(s) is discussed. A few references to technical notes are placed (in the format $^{\mathrm{TN}x}$). At the end of the document, these technical notes are listed. These notes serve as additional explanation and are not part of the exam material. Depending on your mathematical proficiency, you might find these notes either helpful or confusing. Feel free to skip these notes if they confuse you.

Each section closes with a few basic exercises. At the end of the document, the solutions to the exercises are presented.

Hypothesis testing for ρ. The computation of the correlation coefficient r tells something about the strength and direction of the linear relation between x and y in the sample. The next step is relating what this can tell us about the population correlation. Of special interest is the null hypothesis H_0: $\rho = 0$, versus either the two sided alternative H_A: $\rho \neq 0$, or a one sided alternative.

As you (should) know, there is a strong relation between the correlation coefficient ρ and the slope β_1 of the simple linear regression equation $y = \beta_0 + \beta_1 x + \varepsilon$; expression via

$$\beta_1 = \rho \frac{\sigma_y}{\sigma_x} \quad \Leftrightarrow \quad \rho = \beta_1 \frac{\sigma_x}{\sigma_y}.$$

From these equations, you can directly see that $\rho = 0$ if and only if $\beta_1 = 0$. Thus, testing H_0: $\rho = 0$ is equivalent to testing H_0: $\beta_1 = 0$ (against either a one- or two-sided alternative), performed through the test statistic

$$t = \frac{b_1}{SE_{b_1}}. \tag{1}$$

Under the null hypothesis, t follows the t-distribution with $n-2$ degrees of freedom. An alternative formula$^{\mathrm{TN}1}$ to compute the same value is

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}. \tag{2}$$

Exercises Consider Table 1 which contains data on 8 patients suffering from agoraphobic problems. Here, x denotes the number of therapy sessions a patient has received, and y is the grade the patient gave himself for dealing with the problems after the last session (a higher grade indicates better capability of dealing with the problems). Use this data set throughout in the exercises.

1. Confirm (either by hand or by using software) that the regression equation $y = b_0 + b_1 x$ has

[*]University of Groningen; c.j.albers@rug.nl

x	1	2	3	4	5	6	7	8
y	4	2	6	5	6	8	9	6

Table 1: Data used in the exercises

$b_0 = 2.750$ and $b_1 = 0.667$; and that $SE_{b_1} = 0.243$. Interpret the value for b_1.

2. Compute (by hand) the correlation coefficient. Use $\bar{y} = 5.75$, $\bar{x} = 4.5$, $SD(y) = 2.188$, and $SD(x) = 2.450$.

3. Test, at $\alpha = 5\%$, H_0: $\rho = 0$ versus H_a: $\rho > 0$ using approaches (1) and (2). Explain why it is chosen to use the one-sided alternative. Interpret the outcome of the test.

Confidence Intervals for ρ. For many parameters, such as the slope of a regression equation and the mean of a sample, confidence intervals (CIs) are constructed via

$$(\text{estimator}) \pm (\text{critical value}) \times (\text{standard error}).$$

This construction of CIs relies, amongst others, on (approximate) normality of the sampling distribution of the estimator. In general, however, the distribution of r is not symmetric around ρ. This can best be made clear on basis of an example. From a given population with known correlation $\rho = 0.90$, 1000 random samples of size $n = 100$ have been drawn. For each of the 1000 samples, the correlation coefficient r has been computed. Figure 1 displays the histogram of the 1000 estimates for ρ. It is clear that this histogram is not symmetric: the 'tail' on the left hand side is heavier than that on the right.

This is understandable: By definition, a correlation coefficient can never be larger than 1. Thus, no estimate can deviate more than 0.10 to the right from $\rho = 0.90$. However, they can deviate more than 0.10 to the left from ρ and this indeed happens (in 36 of the 1000 cases).

To solve this problem and to be able to create a confidence interval after all, we use the so-called *Fisher z-transformation*, named after the 20th century statistician Sir R.A. Fisher. In short, this technique transforms all values to a 'parallel world' where the assumption of (approximate) normality of the sampling distribution *does* hold. In that parallel world, the standard technique to create a CI is employed, after which the transformation is applied 'the other way round' to obtain a CI in the 'real world'.

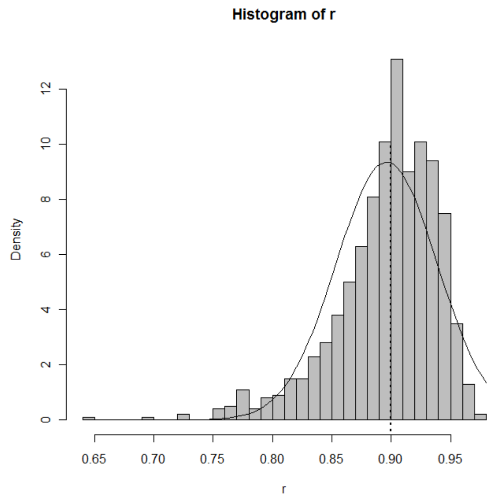

Figure 1: Histogram of the 1000 r-scores.

The Fisher z-transformation transforms an r-value into a new value, denoted by r' (sometimes it is also denoted by z or r_z) through[1]

$$r' = \frac{1}{2} \log\left(\frac{1+r}{1-r}\right).$$

A visualisation of this transformation is given in Figure 2. A few things can be observed: (i) whenever r is positive, r' will be positive (and similarly, negative r yield negative r'); (ii) there is a monotonic relation between r and r', i.e. every r-score corresponds to a unique r'-score, and vice versa; (iii) the r'-scores are not restricted to be inside the interval $[-1, 1]$.

Values of r closer to ± 1 will be affected more by the transformation than values close to 0. This ensures that, in our example, the right half of the plot, which was restricted to $[0.9, 1.0]$ now has a

[1] Recall that in statistics the convention is to always use the natural logarithm (ln) and denoting this as log, unless explicitly stated otherwise.

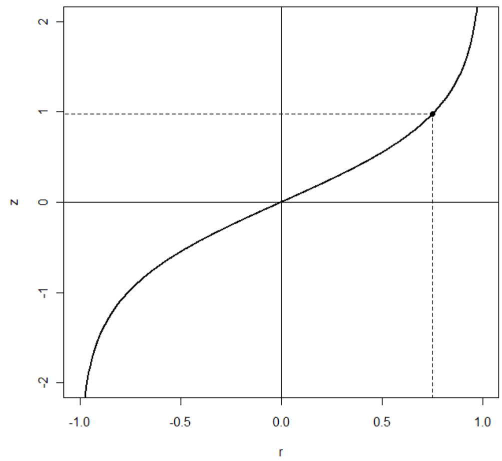

Figure 2: Fisher z-transformation.

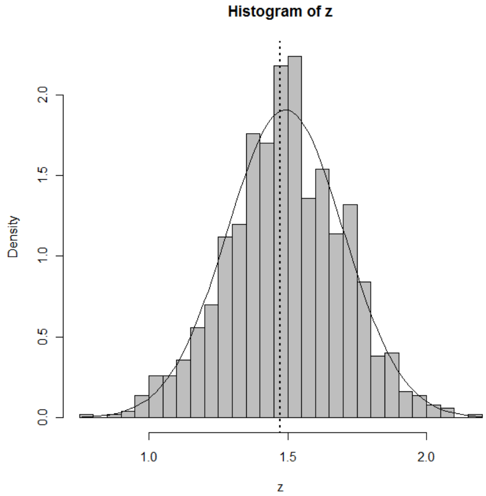

Figure 3: Histogram of the 1000 r'-scores.

wider range of values, whereas the left half of the plot isn't changed too much: It already had a wide range of values. In Figure 3, the same 1000 correlations from Figure 1 are shown again, but this time transformed into r'-scores. As you can see, now the histogram does look nicely symmetric. Indeed, Fisher showed that, after the z-transformation, the sampling distribution of the correlation coefficient is (approximately) normal[TN2], with mean ρ' (the Fisher z-transform of ρ) and variance $\frac{1}{n-3}$. This can be used to construct the confidence intervals.

The $(1-\alpha)\%$ CI for ρ_z is given by

$$\left(r' - z^*_{\alpha(2)}\frac{1}{\sqrt{n-3}}, r' + z^*_{\alpha(2)}\frac{1}{\sqrt{n-3}}\right)$$

where $z^*_{\alpha(2)}$ is the two-sided level-α critical value for the standard normal distribution, e.g. 1.96 when $\alpha = 0.05$. Now we have the CI for ρ', we are not finished yet: We are not interested in ρ', we only use that as a technical tool to obtain an interval for ρ. Thus, we need to transform the CI for ρ' to one for ρ. The inverse transformation of the Fisher z-distribution is

$$r = \frac{\mathrm{e}^{2r'} - 1}{\mathrm{e}^{2r'} + 1}.$$

By applying this inverse transformation to both the lower and the upper bound of the CI for ρ', one obtains the interval for ρ.

Exercises

4. Construct the 95% CI for ρ for the data of the previous exercises.

A test to compare two ρ's. Often in statistics, you want to compare different experiments and decide whether there are population differences or not. Suppose for instance, that you conduced an experiment to measure the correlation between the amount of time spent studying some material and the grade obtained for an exam based on the material. Suppose as well that there are two distinct groups (e.g. men/women, or those who visited training sessions/those who studied by themselves). Group A has sample size n_1 and correlation ρ_1, and group B has n_2 and r_2. It is of interest to test the hypothesis H_0: $\rho_1 = \rho_2$ versus H_a: $\rho_1 \neq \rho_2$ (or, perhaps, versus a one-sided alternative).

Also here, the Fisher z-transformation helps us out. Obviously, when $\rho_1 = \rho_2$ then also $\rho'_1 = \rho'_2$. Thus, testing H_0: $\rho_1 = \rho_2$ is essentially equivalent to testing H_0: $\rho'_1 = \rho'_2$. For this hypothesis the test statistic

$$Z = \frac{r'_1 - r'_2}{\sqrt{\frac{1}{n_1-3} + \frac{1}{n_2-3}}}$$

can be used, and, using a table with critical values (or a computer), the z score can be converted into a p-value.

Exercises

5. Consider again the data in Table 1. For an adapted version of the therapy, data has been collected. In this study, $n = 12$ and $r = 0.500$. Test the null hypothesis of equality of correlations versus the two-sided alternative.

Conclusion. After working through this document, you should: (i) be able to construct hypothesis tests and confidence intervals for the correlation coefficient; (ii) be able to construct hypothesis tests for comparing correlation coefficients; (iii) understand why the Fisher z-transformation is needed.

Technical notes* Feel free to skip this section. It will not be examined. It can help, however, in gaining a better understanding of the material that is part of the exam material.

1. That both formulas give the same t-value can be seen from $b_1 = r s_y / s_x$ and

$$SE_{b_1} = \frac{1}{\sqrt{n-2}} \sqrt{\frac{s_y^2}{s_x^2} - b_1^2}.$$

Then,

$$\begin{aligned} t &= \frac{b_1}{SE_{b_1}} \\ &= \frac{r \frac{s_y}{s_x}}{\frac{1}{\sqrt{n-2}} \sqrt{\frac{s_y^2}{s_x^2} - b_1^2}} \\ &= \frac{\frac{s_y}{s_x}}{\frac{s_y}{s_x}} \frac{r\sqrt{n-2}}{\sqrt{1 - \frac{s_x^2}{s_y^2} b_1^2}} \\ &= \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}. \end{aligned}$$

2. The text provides a heuristic explanation as to why Fisher's transformation works. A full, detailed, and very technical explanation as to why exactly this transformation is the best one (and not any other transformation that affects high correlations more than near-zero ones), is not presented here. It can be found in Fisher, R.A. (1915), Frequency distribution of the values of the correlation coefficient in samples of an indefinitely large population, *Biometrika*, 10,
3. I haven't described the Bayesian approach to inference on correlations. The main reason is that I've restricted attention to approaches that can be computed manually which the Bayesian approach generally can't. In short, the Bayesian approach entails the following steps: (i) assume a bivariate normal distribution for x and y and impose (non-informative) priors on μ_x, μ_y, σ_x and σ_y; (ii) impose a prior on ρ. Common choices are the uniform $U(-1, 1)$ one and the symmetrized reference prior; (iii) through MCMC (and this is the step you don't want to do manually) estimate the posterior density and (iv) use this to obtain a credible interval for ρ or as a step towards computing a Bayes Factor.

Answers to exercises

1. These values are indeed correct. The slope, 0.667, can be interpreted as: 'for each additional session, the self-reported grade increases by two-thirds of a unit'.
2. $r = 0.746$.
3. The test is one sided, because one is only interested in therapies with a positive effect (mathematically it would be equivalent to test H$_0$: $\rho \leq 0$ vs H$_a$: $\rho > 0$). The approach based on the regression context provides

$$t = \frac{b}{SE_{b_1}} = \frac{0.667}{0.243} = 2.748,$$

the approach based on correlation provides

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.746\sqrt{6}}{\sqrt{1-0.746^2}} = 2.748.$$

They are indeed both the same. To obtain the p-value, look at a table with critical values (and df $= n-2 = 6$), to find $0.01 < p < 0.02$. When using software, the p-value can be computed more precisely as $p = 0.0167$. This is smaller than $\alpha = 0.05$ thus the null hypothesis is rejected. We conclude that there is a significant effect of therapy.

4. From $r = 0.746$ we have

$$r' = \tfrac{1}{2} \log \tfrac{1.746}{0.254} = 0.964.$$

Since we need the 95% interval, $z^*_{\alpha(2)} = 1.96$, thus the interval is

$0.964 \pm 1.96/\sqrt{5} = 0.964 \pm 0.877 = (0.087, 1.841)$.

The inverse transformation of the lower bound is

$$r = \frac{e^{2 \cdot 0.087} - 1}{e^{2 \cdot 0.087} + 1} = \frac{1.190 - 1}{1.190 + 1} = 0.087$$

(note: for values r close to zero, it holds that $r \approx r'$, this explains why both r and r' coincide up to three decimal places). Similarly, the upper bound is computed via

$$r = \frac{e^{2 \cdot 1.841} - 1}{e^{2 \cdot 1.841} + 1} = \frac{39.726 - 1}{39.726 + 1} = 0.951.$$

Thus, the 95% CI is $(0.087, 0.951)$. This shows that, even though the correlation differs significantly from zero, the interval covers almost fully all positive values: We can not accurately tell what the value of the correlation is (which is obviously due to the small sample size).

5. For both studies, r' needs to be computed:

Study	n	r	r'
A	8	0.746	0.964
B	12	0.500	0.549

Next, compute

$$Z = \frac{0.964 - 0.549}{\sqrt{\frac{1}{5} + \frac{1}{9}}} = \frac{0.415}{0.558} = 0.746.$$

Next, use a table to find the corresponding p-value: $p = 0.46$. Thus, H_0 cannot be rejected.