

Exercises day 5

Basic Statistics for health researchers 2026

March 2nd, 2026

Warming up

Before starting the exercise below, learn from the R-demo of Lecture 5 (available from the course webpage):

1. Read and run the code.
2. Check that the output matches the results presented on the slides.
3. Do not hesitate to add your own comments into the script.

Exercise A (Sample size and power calculation)

Assume that we aim to study a new fertility treatment. We plan to compare the chances of pregnancy within 6 months. Assume that, according to previous studies, we can reasonably think that without fertility treatment the chance of pregnancy within 6 months is approximately 40%, for the (specific) population of women that we study.

1. How many women could we enroll in our study, if we want a statistical power of 85% and expect that the treatment results in 60% chance of pregnancy. As usual, we plan to control the risk of false positive finding at 5%. **Note:** we “could” does not mean that we “should”, as it might become apparent from the next questions.
2. Using the sample size suggested by the result to the previous question:
 - (a) What does the power become if the treatment is actually less effective than we thought and results in only 50% or 55% chance of pregnancy?
 - (b) What is the smallest chance difference that leads to a “decent” power of at least 75%?

Exercise B (Inference in 2×2 tables, subgroup analyses)

For this exercise we will work with the “Smoking” data (available from the course webpage).

Part I

In this first part we assume that we would like to (naively) study the (overall) association between smoking and 20-years risk of death.

1. Read the description of the data and load the data into R/SAS. How many women were still alive 20 years after the initial survey? How many women were smoking at the time of the initial survey?
2. Compute a “2 by 2” table and investigate the association between smoking (initial survey) and 20-years risk of death. **Hint:** use the `table()` function.
 - (a) How many smokers and non smokers were alive and dead 20 years after the initial survey?
 - (b) Estimate the 20-year risk of death for smokers and non smokers. Are the results surprising?
 - (c) Compute confidence intervals for these two 20-year risk of death. **Hint:** use the `binom.test()` function.
 - (d) Do these first results “suggest” that the direction of the results can likely be explained by small sample random variation (i.e. “chance”)?
3. Use a statistical test to investigate whether smoking is significantly associated with 20-year risk of death. Conclude.
4. Use the `table2x2()` function to estimate each of the following association measure, together with confidence intervals, to complement the p-value you have just computed: risk difference, risk ratio, odds ratio.
Note: to obtain the risk of death (and not survival) it is necessary to relevel *death*, with ‘yes’ as reference (the function is calculating the probability for the event in the first column of the table). You can use the code

```
smoking$death<- relevel(smoking$death, ref="yes")
```
5. Write a conclusion sentence quoting the risk ratio to describe the association between smoking and death.
6. Produce a barplot to compare the number of women included in each age group, between smokers and non smokers. What do you observe? **Hint:** use the `barplot()` function with the option `beside=TRUE`.
7. Produce a barplot to compare the 20-years probability of death in each age group. What do you observe?

8. Visually compare the 20-year probabilities between smokers and non smokers, in each age group. If you have substantial programming experience, you can try to write some code to produce an appropriate plot. If you have less experience (which is completely fine), just look at Figure 1 below (the solutions will present the code to produce the plot). What do you observe?
9. Can you now provide a plausible explanation for the results obtained at question 2?

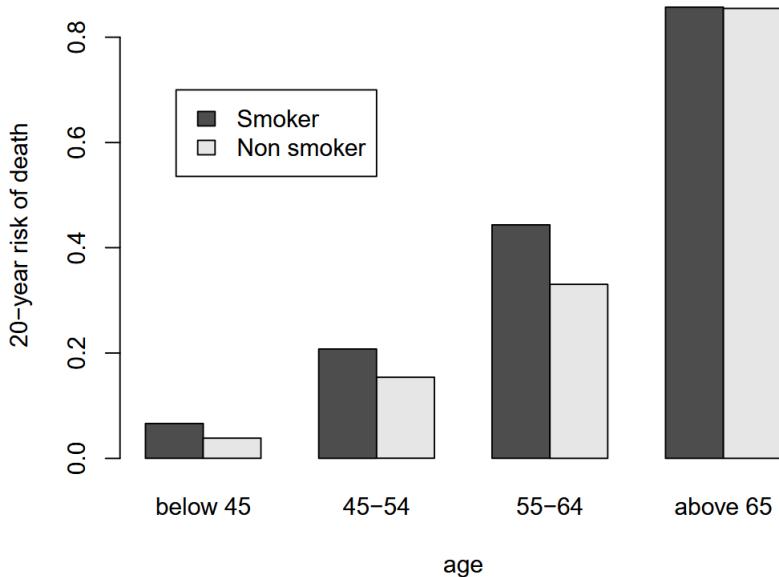


Figure 1: 20-year risk of death in each age group, for smokers and non smokers.

Part II (if time allows)

In this second, part we assume that we would like to study the association between smoking and 20-year risk of death separately within each age group.

10. Compute the risk ratio and the corresponding confidence intervals to describe the association in each age group. Conclude.
11. Why do the results of Part II seem more interesting than those of Part I?

Exercise C (Crossover trial, if time allows)

Background, data and aim:

“There is growing concern regarding the side effects of sedative hypnotics especially in regard to arousability and fall risk, which are critical for safety when responding to nighttime stimuli, such as a fire alarm or a crying baby.” (Drake et al (2017)).

Drake et al (2017) wanted to examine and compare the arousability threshold of two drugs for sleep disorders: doxepin (6 mg) and zolpidem (10 mg). To do so, they included

$n = 52$ healthy males in their study. They conducted an experimental procedure during which each subjects was monitored while sleeping during four nights. In this exercise, we work with data from two of these four nights, during which each participant received either zolpidem or doxepin. Each night, at a specific time, the investigators started to present acoustic stimuli to the subjects through audiometric earphones. Tones began at 30 dB and increased by 5 dB until the participant woke up or the maximum dB-level (110 dB) was reached. In this exercise, our aim is to compare the risk of the participant not waking until receiving the loudest tone (110 dB), under each treatment.

Questions:

1. Read the description of the “sedative” data (available from the course webpage) and load the data into R/SAS.
2. What are the proportions of subjects who did not wake before receiving the loudest tone (110 dB) after receiving each treatment?
3. Compute 95% confidence intervals for the risk of not waking before receiving the loudest tone (110 dB) after receiving each treatment.
4. Name a (statistical) hypothesis test that can be used with these “sedative” data, to compare the risk of not waking after receiving each treatment. Justify your choice.
5. Compute an appropriate p-value and write a conclusion sentence to summarize the main results.
6. Find the paper by Drake et al (2017) (<https://doi.org/10.1093/sleep/zsx086>) and read the abstract. Do your results and conclusion sentence match those in the abstract? ¹

¹Note: it seems that there is a minor typo in the abstract, 17.6% instead of 17.3% as reported in the result section ‘Primary Outcomes’, subsection ’Auditory Awakening Threshold’, on page 4.