

# Home Exercises 5

Your Name

25.10.2021

Write your name at the beginning of the file as “author:”.

1. Return to Moodle by **9.00am, Mon 25.10.** (to section “BEFORE”).
2. Watch the exercise session video available in Moodle by **10.00am, Mon 25.10.**
3. If you observe during the exercise session that your answers need some correction, return a corrected version to Moodle (to section “AFTER”) by **9.00 am, Mon 1.11.**

## Problem 1 & 2. (Definition of statistical power.)

- (a) An existing treatment helps patients in 60% of the cases. We want to study whether a new treatment performs better (or worse) than the existing treatment. We are planning collecting  $n$  patients and apply the new treatment to each of them. To understand the statistics of the planned experiment, plot the null distribution (using probability mass function) ( $H_0$ : new treatment is like the old one and has success rate 60%) of the possible outcomes when  $n = 70$ . Hint: The possible outcomes are the numbers of successes in the binomial experiment with  $n = 70$  and  $p = 0.6$ . To plot the distribution you can use simply a plot command where  $x$  is a vector of possible outcome values and  $y$  has the corresponding probabilities from `dbinom( )` with appropriate parameters.
- (b) For the sake of power calculation, let's hypothesize that the new treatment would help 80% of the patients ( $p_{new} = 0.80$ ). Plot the distribution of possible outcomes under this hypothesized distribution in the same Figure where the null distribution is. Based on the Figure, do you expect to have a good power to detect benefits of the treatment that has success rate of 80% over the old treatment with this sample size. Hint: You can use `lines( )` command and a different color from previous distribution to show the alternative distribution. If you instead call `plot( )`, it will make a completely new plot.
- (c) If we set a significance level  $\alpha = 0.05$ , what would be the possible outcome values that would give a significant result (under the null hypothesis)? Hint: Apply `qbinom( )` to both tails of the null distributions to find the two cutpoints that separate a tail probability of  $\alpha/2$  from the lower and upper tails.
- (d) Under the hypothesized alternative distribution ( $p_{new} = 0.80$ ), what is the probability of getting a significant result given the region computed in (c)? In other words, what is the power of the planned experiment. Hint: Use `pbinom( )` to the alternative distribution at the cutpoints and sum the two tails up.
- (e) Compute the power at significance level  $\alpha = 0.05$  with the parameters used in (a-d) also using `pwr.p.test( )`. Does it approximately agree with the value you computed in part (d)?

NOTE: Before you can apply `pwr.p.test( )`, you must first install `pwr` package using `install.packages("pwr")`, and then load it in your current Rmd file using command `library(pwr)`.

### **Problem 3 (Power of t-test)**

Consider study of hypertension treatment by Chow et al. [https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(21\)01922-X/fulltext](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(21)01922-X/fulltext) about a phase 3 trial among Australian adults with hypertension, who were untreated or receiving monotherapy.

Participants were randomly assigned to either treatment, that started with the quadpill (containing irbesartan at 37.5 mg, amlodipine at 1.25 mg, indapamide at 0.625 mg, and bisoprolol at 2.5 mg) or an indistinguishable monotherapy control (irbesartan 150 mg).

The primary outcome was difference in unattended office systolic blood pressure at 12 weeks.

They report: "Before the study, we estimated a sample size of 650 patients would provide 90% power at an  $\alpha$  of 0.05 to detect a difference of 4 mm Hg in the primary outcome, assuming an SD of 15 mm Hg. The calculations allowed for a 10% data-loss rate."

Repeat the power calculation for t-test assuming equal sample sizes in treatment and control groups. Do you approximately agree with their estimated sample size? Note that 10% data-loss rate means that the estimated sample size is larger by 10% compared to the theoretical sample size because some patients may drop out from the study.

### **Problem 4. (Proportions in a clinical trial.)**

Read the Summary of this study from Lancet 397: 2487-2496 June 2021

[https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(21\)01063-1/fulltext](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(21)01063-1/fulltext)

They did a randomised study in South Korea on patients aged at least 20 years who maintained dual antiplatelet therapy without clinical events for 6–18 months after percutaneous coronary intervention with drug-eluting stents. Patients were randomly assigned (1:1) to receive a monotherapy agent of clopidogrel 75 mg once daily or aspirin 100 mg once daily for 24 months. The primary endpoint was a composite of all-cause death, non-fatal myocardial infarction, stroke, readmission due to acute coronary syndrome, and Bleeding Academic Research Consortium (BARC) bleeding type 3 or greater at 24 months.

They report: "The sample size calculation was based on the assumption that the event rates of the primary endpoint at 24 months would be 9.6% for the clopidogrel group and 12.0% for the aspirin monotherapy group. With a sampling ratio of 1:1 and an estimated rate of follow-up loss as 5% in each group for 24 months, 5530 patients were needed to ensure a power of at least 80% with a two-sided  $\alpha$  of 5%.

Repeat the power calculation assuming that statistical method is test of difference in proportions in two samples of same size. Do you get similar sample size? (OK to have difference of some tens of patients from the reported.)