06. Hypothesis testing

To be completed Friday 9th of December, by 12 noon sharp. You are encouraged to submit in pairs, but are also allowed to submit alone. **One person** must submit the solution as **a single pdf file**, to the folder on Ilias, by naming *exercise_03_name1_matriculationnumber1_name2_matriculationnumber2.pdf*, with the obvious replacements in the strings. **Not adhering to the formatting requirements may result in the submission not being graded.**
Exercises are only graded in a binary fashion as sufficient or insufficient. To be graded as sufficient, you do not necessarily have to have correct solutions to every sub-question, but you must have made a clear and earnest effort to solve the entire exercise. Ultimately, what constitutes sufficient is at the discretion of the tutors. To be admitted to the exam, you must have submitted sufficient answers to at least 5 of the 6 exercise sheets.

1. **EXAMple Question — Interpreting null-hypothesis significance testing**

   Two separate groups of students, from courses A and B, take the same mathematics test. We want to determine whether or not there is a difference in the two groups' performance on the test.

   For the following statements, explain if they are True or False. State your reasoning.

   a) A reasonable null hypothesis is that the expected value of the performance-score of students in group A and B is the same.

   b) If we obtain a p-value of .09 when comparing the performance of the two groups, then we can conclude there is no difference in performance.

   c) If we obtain a p-value of .00001 when comparing the performance of the two groups, then we have proven that there is a difference in their performance.

   d) If we take a randomly chosen student from group A, and a randomly chosen student from group B, then the p-value will tell us the probability that the student from group A has a lower performance score.

2. **Theory Question — Constructing an exact hypothesis test**

   Consider i.i.d. samples from a normal distribution $x_1, \dots, x_n \sim \mathcal{N}(\mu, \sigma^2)$. We want to test the hypothesis

   $$H_0 : \mu = 0 \qquad\qquad H_1 : \mu \neq 0$$

   We choose the test statistic $T(x_1, \cdots, x_n) = \frac{1}{n}\sum_{i=1}^{n} x_i$. Clearly, the more the test statistic deviates from zero, the more likely the test should reject the null hypothesis. Hence, we define a *rejection region* $R = [-\infty, -c] \cup [c, \infty]$ for some $c > 0$, and if the test statistic falls into R, we reject H_0, otherwise, we retain it.

   a) Compute the distribution of the test statistic.
   Hint: You can use that if $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$ and $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$, and X and Y are independent, then $X + Y \sim \mathcal{N}(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2)$. In addition, for any $a \in \mathbb{R}$, we have $aX = \mathcal{N}(a\mu_X, a^2\sigma_X^2)$

   b) A useful quantity to construct a hypothesis test is the *power function* $\beta(\mu)$. Show that the power function here is given by

   $$\beta(\mu) = P_\mu\left(T(x_1, \dots, x_n) \in R\right) = 2\left(1 - \Phi\left(\frac{\sqrt{n}}{\sigma}(c - \mu)\right)\right)$$

   were Φ denotes the CDF (cumulative distribution function) of a standard Normal $\mathcal{N}(0, 1)$.
   Hint: The Normal distribution is symmetric, thus $\Phi(-c) = 1 - \Phi(c)$.

c) Let $\alpha \in (0, 1)$. Show that the value $c = c_\alpha$ with

$$c_\alpha = \frac{\sigma}{\sqrt{n}} \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$$

leads to a test with a significance level of α. Here Φ^{-1} denotes the inverse CDF of a standard Normal.

d) Show that the p-value for any dataset $X = (x_1, \ldots, x_n)$ can be computed by

$$p_X = 2\left(1 - \Phi\left(\frac{\sqrt{n}}{\sigma} T(X)\right)\right)$$

3. **Practical Question — Permutation testing**

In this exercise, we will implement permutation testing and interpret the respective results. Open the notebook to continue.