

EDDA: Final Assignment

- *Submit in Canvas a well readable report before the specified deadline, as one pdf-file. Submitting an R-script file is also allowed, it must contain relevant plot commands.*
- *The submitted report must follow the Guidelines for assignments posted on Canvas.*
- *Unless otherwise specified, tests should be performed using a level of 0.05.*

Exercise Diet

To investigate the effect of 3 types of diet, 78 persons were divided randomly in 3 groups, the first group following diet 1, second group diet 2 and the third group diet 3. Next to some other characteristics, the weight was measured before diet and after 6 weeks of diet for each person in the study. The collected data is summarized in the data frame `diet.txt` with the following columns: `person` – participant number, `gender` – gender (1 = male, 0 = female), `age` – age (years), `height` – height (cm), `preweight` – weight before the diet (kg), `diet` – the type of diet followed, `weight6weeks` – weight after 6 weeks of diet (kg).

- (1.0 point) By using only the columns `preweight` and `weight6weeks`, test the claim that the diet affects the weight loss. Give and check the assumptions of the test(s) applied.
- (1.0) Compute and add to the data frame the variable `weight.lost` expressing the lost weight. Test the claim that the median lost weight is bigger than 3 kg.
- (1.5) Using now `weight.lost` as response variable, apply one-way ANOVA to test whether type of diet has an effect on the lost weight. Do all three types diets lead to weight loss? Which diet was the best for losing weight?
- (1.2) Use two-way ANOVA to investigate effect of the `diet` and `gender` (and possible interaction) on the lost weight. Is the Friedman test relevant here? Can it be at all applied?
- (1.3) Apply appropriate model to investigate effect(s) of `diet` and `height` (and possibly their interaction) on the lost weight. Is the effect of `height` the same for all 3 types of diet?
- (1.5) Which of the two approaches, the one from c) or from e), do you prefer? Why? For the preferred model, predict the lost weight for all three types of diet for an average person.
- (2.5) Create the categorical (binary) variable `lost.4kg` expressing whether the lost weight is bigger than 4kg or not. (Hint: if `a=c(1,-1,2)`, then `a>0 = TRUE FALSE TRUE`.)

Suppose now that we observed the variable `lost.4kg` instead of `weight6weeks`. Could we still test the hypothesis in b), c), d), e)? If yes, propose relevant models and perform relevant tests.