02424 Assignment 1: Clothing insulation level

This is the first of three mandatory assignments for the course 02424. It must handed in using the lear.dtu (time and date is given there). The submissions most contain one collected attached file in Portable Document Format (PDF), other document formats will not be accepted.

The report of each assignment must be prepared in groups of 3, and the final grading wil be based on the reports and the (individual) oral exam.

When writing the report please explain carefully what you did in each step, back up your statements with quantitative measures when possible, explicitly write down all models used in mathematical notation, and last keep it short and concise.

The problem and data

The level of clothing a person is wearing at office is one of the key factors influencing their level of comfort [1]. In addition, the level of comfort influences the need for cooling and/or heating.

The data (in the file `clothingSum.csv`) include the variable given in Table 1 below.

Table 1: List of included variables in the data-set

Variable	Type	Description
`clo`	Continuous	Level of clothing
`tOut`	Continuous	Outdoor temperature
`tInOp`	Continuous	Indoor operating temperature
`sex`	Factor	Sex of the subject
`subjId`	Factor	Identifier for subject
`day`	Factor	Day (within the subject)

Problem A: General linear model

The first part of the assignment concerns modeling of clothing insulation level based on Indoor operating temperature and outdoor temperature (i.e. you should ignore `subjId`, and `day`). You should consider the following steps

1. Make an exploratory analysis of the data.

2. Fit general linear models to the data (i.e. find the best model).
3. Give an interpretation of your model, including some graphical presentation.
4. Argue that a weighted analysis should be used and estimate the weight parameters (hint: consider sex differences)
5. Present the final model by some plots, including confidence interval and prediction intervals.
6. Use the residuals to check if the subject Id can be ignored in your final model.

Problem B: Including subject Id

In this part you should include subject Id in the modelling.

1. Develop a general linear model that include subject Id as an explanatory variable.
2. Give a visual presentation of the parameters, do you see?
3. Give an interpretation of the parameters, and the model. Can the model e.g. be used to predict the behavior of a new subject?

Problem C: The full data-set

The data you have used so far is an aggregated version of a larger data-set the include multiple observations pr. day, in this part you should use the full data-set (`clothingFull.csv`), which include a number of oberservations (around 6) for each day that a subject visit the lab. I addition to the variables in the first part of the assignment, the dataset also include an indentifier of observation number (within subject and day).

1. Fit a model with the same structure as in Part A/B, and discuss the assumption on the residuals (in particular the assumption of independence).

Problem D: Conclussion

Write a small conclusion of your finding.

References

[1] Fanger, P.O. (1970). *Thermal Comfort Analysis and Applications in Environmental Engineering*. McGraw-Hill, New York.

[2] Schweiker, M. and Wagner, A. (2015). *A framework for an adaptive thermal heat balance model (ATHB)*. Building and Environment (94), Elsevier Ltd.