

---

Toenail Infection [toenail]

---

## Assignment

### Dataset

The dataset considers information from a longitudinal clinical trial in dermatology which was set up to compare the efficacy of two oral treatments (*testing* and *standard*) for toenail infection (De Backer et al., 1998). One of the end points of the study was the degree of onycholysis which expresses the degree of separation of the nail plate from the nail-bed (0, *absent*; 1, *mild*; 2, *moderate*; 3, *severe*) and was evaluated at seven visits (approximately on weeks 0, 4, 8, 12, 24, 36 and 48). In total, 1908 measurements on 294 patients are available. In this dataset, only a dichotomized onycholysis (0, *absent or mild*; 1, *moderate or severe*) is given.

The data have kindly been made available for statistical research by Novartis, Belgium. The source of the data must be acknowledged in any publication which uses them, see Lesaffre and Spiessens (2001) for more details.

### Problem

Compare efficacy in treatment of onycholysis of the *testing* treatment in comparison to the *control* one.

### Requirements

- (i) Use standard logistic model (in which independence of observations is assumed) and develop reasonable model capturing evolution of probabilities of infection in the two groups over time. **Write down the linear predictor of the model and provide one table with estimated model parameters, standard errors, Wald p-values and the model deviance.**
- (ii) **Plot** estimated evolution of probabilities of infection, related 95% pointwise confidence bands and empirical probabilities of infection (all in one plot).
- (iii) Based on the model **specify the null hypothesis** that will express a hypothesis of no difference in treatment efficacy between the two groups. Perform the test (using standard methods for MLE estimated GLM with independence assumption), **report the P-value and your conclusion.**

The report in the pdf format (file named as Surname\_Firstname\_4.pdf) and the related R script (file named as Surname\_Firstname\_4.R) have to be submitted in Moodle by **Monday May 8, 2023 [23:59 CEST]**.

### Dataset

The dataset (in ASCII format, space separated values) can be downloaded from [https://www2.karlin.mff.cuni.cz/~komarek/vyuka/2022\\_23/nmst412/Problem\\_4/toenail.txt](https://www2.karlin.mff.cuni.cz/~komarek/vyuka/2022_23/nmst412/Problem_4/toenail.txt)

The dataset contains 1908 rows (visits) conducted on 294 patients and 5 variables.

*Variable list:* See Table 1.

Table 1: Variable coding table

Variable Name	Variable Label	Variable Coding
<code>idnr</code>	identification number of the patient	integer
<code>infect</code>	dichotomized onycholysis	0: <i>absent or mild</i> ; 1: <i>moderate or severe</i>
<code>trt</code>	treatment group	0: <i>control</i> ; 1: <i>testing</i>
<code>time</code>	time of measurement (in months)	numeric $\in [0, 18.5]$
<code>visit</code>	visit number	integer $\in \{1, \dots, 7\}$