02424 Assignment 3

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

Problem 1: Weight of rats

Part A

The following experiment was carried out to investigate the effect of *thyroxin* and *thiouracil* on the weight of rats. In a period of 5 weeks 7 rats had *thyroxin* added to their drinking water, 10 rats had *thiouracil* added to their drinking water, and another 10 rats got plain drinking water. The data originates from Box(1950).

	Time	1	2	3	4	5
Rat	Treatment	Log-weight (g)				
1	Plain	4.0431	4.4543	4.7362	4.9345	5.1475
2	Plain	4.0943	4.5326	4.8122	4.9836	5.1761
3	Plain	3.9512	4.3438	4.7095	4.9698	5.2204
4	Plain	3.8918	4.2047	4.6052	4.8598	5.0999
5	Plain	4.0254	4.3944	4.6444	4.7958	5.0173
6	Plain	3.8286	4.2485	4.6250	4.8752	5.0304
7	Plain	3.9318	4.2627	4.5433	4.7005	4.9488
8	Plain	4.1431	4.5109	4.7185	4.8675	5.0370
9	Plain	3.8918	4.2047	4.4998	4.7185	4.9416
10	Plain	4.0431	4.4067	4.7005	4.9345	5.1299
11	Thyroxin	4.0775	4.4427	4.7958	4.9836	5.1985
12	Thyroxin	3.9890	4.2627	4.4998	4.7005	4.9273
13	Thyroxin	4.0254	4.3175	4.6821	5.0173	5.2417
14	Thyroxin	4.0775	4.4427	4.7536	4.9972	5.1761
15	Thyroxin	4.0431	4.2767	4.5747	4.7875	4.9698
16	Thyroxin	3.9512	4.2905	4.5747	4.7536	4.9416
17	Thyroxin	3.9512	4.2485	4.6540	4.9273	5.1417
18	Thiouracil	4.1109	4.4543	4.6913	4.7875	4.8598
19	Thiouracil	4.0775	4.3820	4.6151	4.7095	4.8040
20	Thiouracil	3.9703	4.3694	4.6052	4.6634	4.8903
21	Thiouracil	4.0775	4.4773	4.6052	4.7095	4.8040
22	Thiouracil	3.9318	4.3175	4.6151	4.8122	4.9416
23	Thiouracil	3.9318	4.3175	4.5218	4.6052	4.7791
24	Thiouracil	4.0254	4.3567	4.5539	4.6347	4.6821
25	Thiouracil	4.0604	4.2341	4.5326	4.7536	4.9416
26	Thiouracil	3.8286	4.1109	4.3567	4.4998	4.6728
27	Thiouracil	3.9703	4.2767	4.4886	4.6444	4.8040

The goal is to find a suitable model to describe the data and draw the correct conclusion about the three treatments, so briefly conclude on the effects of the treatments after each of the following questions.

1. Make a plot of the average log-weights for the three treatments plotted against time, and possibly other relevant exploratory plots.

2. Analyze the data via a summary variable for each rat. The main difficulty with this data set is that observations taken from the same rat will likely be correlated. To overcome this we can reduce the data set to one observation per rat, which will then be independent. For instance the increase in log-weight $\Delta \log W_i$ over the 5 weeks for each rat for each rat $i = 1 \ldots 27$.

3. Analyze the data via a simple linear mixed model. One way to account for correlated observations on the same rat is to set up a model where the base level for each rat is considered a random variable. Investigate also if it is possible to describe the log-weight development over time as a second degree polynomial over time (possibly with treatment dependent coefficients).

4. Analyze the data via a repeated measurement model. Set up a model where the correlation the measurements within each rat is explicitly specified.

5. Which of the previous models is the most suitable for this data set?

Problem 2: Poisson mixed effect models

This problem deals with Poisson mixed effect models, the data used for the exercise is simulated data (see file `simdat3.csv`). The data contain an outcome `y`, two regressors `x1`, and `x2` and a grouping variable `group`.

1. Make an explorative analysis of the data.
2. Fit a generalized linear model and argue that the grouping structure needs to be taken into account
3. Fit a generalized linear mixed effect model, with groups as the random effect. You may use the R-package `glmmTMB` to do this.
    - Write down the model and the estimates, and the interpretation of the parameters
    - Plot the estimated random effects in some appropriate way
    - Check the accuracy of the Laplace approximation by importance sampling
4. Using the Laplace approximation, fit a hierarchical model (similar to the model estimated above) with the second stage model being a Gamma model, with mean value equal 1 (in line with the formulation in Theorem 6.3).
    - Write down the model
    - Estimate parameters and compare with the result in Question 3
    - Check the accuracy of the Laplace approximation by importance sampling
5. (Difficult) The model in Question 4 allows an explicit calculation of the integral in the hierarchical likelihood.
    - Find the explicit formulation of the likelihood (Hint: consult the proof of Theorem 6.1 and the formulation in Theorem 6.3 and Example 6.6, note that the there is a misprint here ($U = \log(V)$ should be $V = \log(U)$)).
    - Implement the above formulation and estimate the parameters.
    - Find the conditional mean and variance of the random effects (Hint: see (the proof of) Theorem 6.2)
6. Compare and discuss the models and estimaton methods you have used in this problem, e.g. which method/model would you prefer, how difficult is it to generalize the results to more complicated structures, etc.