Non-Life Insurance: Mathematics and Statistics

Exercise sheet 10

Exercise 10.1 Log-Linear Gaussian Regression Model (R Exercise)
Suppose that a car insurance portfolio of an insurance company has been divided according to two tariff criteria

- vehicle type: {passenger car, delivery van, truck} = {1,2,3},
- driver age: {21-30 years, 31-40 years, 41-50 years, 51-60 years} = {1,2,3,4}.

For simplicity, we set the number of policies $v_{i,j} = 1$ for all risk classes (i, j), $1 \leq i \leq 3$, $1 \leq j \leq 4$. Moreover, we assume that we work with a multiplicative tariff structure and that we observed the following claim amounts:

	21-30y	31-40y	41-50y	51-60y
passenger car	2'000	1'800	1'500	1'600
delivery van	2'200	1'600	1'400	1'400
truck	2'500	2'000	1'700	1'600

Table 1: Observed claim amounts in the $3 \cdot 4 = 12$ risk classes.

Calculate the tariffs using the log-linear Gaussian regression model.

(a) Determine the design matrix Z of the log-linear Gaussian regression model.

(b) Calculate the tariffs using the MLE method within the log-linear Gaussian regression model framework under the assumption that the variance correction term can be avoided.

(c) Is there statistical evidence that the classification into different types of vehicles could be omitted?

Exercise 10.2 Method of Bailey & Simon
Consider the same setup as in Exercise 10.1. Calculate the tariffs using the method of Bailey & Simon. Comment on the results.

Exercise 10.3 Method of Bailey & Jung
Consider the same setup as in Exercise 10.1. Calculate the tariffs using the method of Bailey & Jung (i.e. the method of total marginal sums). Compare the results.

Exercise 10.4 Tweedie's Compound Poisson Model
Let $S \sim \text{CompPoi}(\lambda v, G)$, where $\lambda > 0$ is the unknown claim frequency parameter, $v > 0$ the known volume and G the distribution function of a gamma distribution with known shape parameter $\gamma > 0$ and unknown scale parameter $c > 0$. Then, S has a mixture distribution with a point mass of $\mathbb{P}[S = 0]$ in 0 and a density f_S on $(0, \infty)$.

(a) Calculate $\mathbb{P}[S = 0]$ and the density f_S of S on $(0, \infty)$.

(b) Show that S belongs to the exponential dispersion family with

$$
\begin{aligned}
w &= v, \\
\phi &= \frac{\gamma+1}{\lambda\gamma}\left(\frac{\lambda v \gamma}{c}\right)^{\frac{\gamma}{\gamma+1}}, \\
\theta &= -(\gamma+1)\left(\frac{\lambda v \gamma}{c}\right)^{-\frac{1}{\gamma+1}}, \\
\boldsymbol{\Theta} &= (-\infty, 0), \\
b(\theta) &= \frac{\gamma+1}{\gamma}\left(\frac{-\theta}{\gamma+1}\right)^{-\gamma}, \\
c(0, \phi, w) &= 0 \quad \text{and} \\
c(x, \phi, w) &= \log\left(\sum_{n=1}^{\infty}\left[\frac{(\gamma+1)^{\gamma+1}}{\gamma}\left(\frac{\phi}{w}\right)^{-\gamma-1}\right]^{n}\frac{1}{\Gamma(n\gamma)n!}\,x^{n\gamma-1}\right), \quad \text{if } x > 0.
\end{aligned}
$$

Exercise 10.5 Log-Linear Gaussian Regression Model (R Exercise)
Interpret the following R output of Exercise 10.1.

Listing 1: R output of Exercise 10.1.

```
Call:
lm(formula = observation ~ van + truck + X31_40y + X41_50y + X51_60y, data = data)

Residuals:
      Min        1Q    Median        3Q       Max
-0.087095 -0.019871  0.006206  0.022773  0.064464

Coefficients:
            Estimate Std. Error t value Pr(>!t!)
(Intercept)  7.68800    0.04233 181.610 1.88e-12 ***
van         -0.05625    0.04233  -1.329 0.232227
truck        0.11342    0.04233   2.679 0.036575 *
X31_40y     -0.21565    0.04888  -4.412 0.004511 **
X41_50y     -0.37511    0.04888  -7.674 0.000256 ***
X51_60y     -0.37381    0.04888  -7.647 0.000261 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 0.05987 on 6 degrees of freedom
Multiple R-squared:  0.941,     Adjusted R-squared:  0.8918
F-statistic: 19.13 on 5 and 6 DF,  p-value: 0.001261
```