Technische Universität München
Zentrum Mathematik
Lehrstuhl für Mathematische Statistik
Prof. Claudia Czado, Ph.D. (cczado@ma.tum.de)
Ariane Hanebeck (ariane.hanebeck@tum.de)

WS 2022/2023

Exercise Sheet 5
Generalized Linear Models

Discussion of the tutorial exercises on November 21 and 24, 2021

Problem 1 (*)

a) Assume $Y_i \sim Binom(n_i, p_i), i = 1, ..., n$ independent and $Y_i^S = Y_i/n_i$.

Show Eq. (3.28) from the book, i.e. show that the (unscaled) deviance for the scaled binomial regression model is given by

$$D(\widehat{\boldsymbol{\mu}}, \boldsymbol{y}) = 2 \sum_{i=1}^{n} \left\{ y_i \log\left(\frac{y_i}{\hat{\mu}_i}\right) + (n_i - y_i) \log\left(\frac{n_i - y_i}{n_i - \hat{\mu}_i}\right) \right\}.$$

b) Determine the (unscaled) deviance and the canonical link function for the following distribution. Assume that n realizations $y_1, \ldots, y_n$ are given.

$Y \sim \text{Gamma}(r, \lambda)$ with density function

$$f_Y(y) \quad = \quad \frac{\lambda^r}{\Gamma(r)} y^{r-1} e^{-\lambda y} \quad y \geq 0,\, r > 0,\, \lambda > 0\,.$$

We assume that r and λ are unknown.

Problem 2 (*)

a) Plot the three functions F_0 from Definition 4.3 in the book as a function of the linear component η in **one** graph. Describe the differences.

b) Now, assume we have two covariates X_1 and X_2 and given values $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$. How does $\hat{p} = \hat{\mu}^S$ depend on x_1 and x_2 using F_0?

Given $(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2) = (1, 1, 1)$, plot a surface plot of $\hat{\mu}^S(x_1, x_2)$ over all the values $x_1 \in (-5, 5)$ and $x_2 \in (-5, 5)$ for the three models from Definition 4.3 respectively.

Describe what you see. Also play around with the values for $(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2)$ to see what changes.

Hint: A surface plot can be plotted like this:

```
>library(plotly)

>fig <- plot_ly(type = 'surface', x = x, y = y, z = z)

>fig
```

Problem 3 (Additional) Let $\boldsymbol{H}$ denote the Hessian matrix (i.e. the matrix of the second derivatives) of the log-likelihood function. Show that the Hessian matrix $\boldsymbol{H}$ of a GLM with a canonical link function and known ϕ is independent of the data $\boldsymbol{y}$.