A5 Consider the gamma GLM.

(a) Calculate the expected Fisher information $I(\beta)$ in case the design matrix is of the form

$$X = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}$$

and an arbitrary link is used. How does $I(\beta)$ simplify when the canonical link is used?

(b) Under the setup of part (a), calculate the asymptotic variance of $\hat{\beta}_1$ and $\hat{\beta}_2$.

(c) Suppose that the gamma GLM contains the intercept and one one-factor level predictor, and that the canonical link is used. Write down $\hat{\beta}_1$ and $\hat{\beta}_2$, as well as their asymptotic variances. *You can use any result proved in the previous assignment.*

(d) How do the results in part (c) change when the log link is used?

(e) Derive the deviance for an arbitrary gamma GLM.

(f) For a gamma GLM, calculate the Pearson, Anscombe and deviance residuals.

A6 For a GLM assuming $Y_i \sim \mathcal{N}(\mu_i, \sigma^2)$, find the form of the difference between the deviances for nested models M_0 and M_1. How does your finding relate to comparing nested models in classical linear regression?

A7 `R` **excercise.** Load the data set on Horseshoe Crab Mating (*Ethology* 102, 1–21, 1996).

```
library(glmbb)
data(crabs)
```

The response is `satell`, the number of satellites (male crabs) attached to the nest of the female crab.

(a) Use `weight` (weight of the female crab) and `color` (her color) as predictors. Fit an appropriate GLM with `weight`, `color` and the interaction between `weight` and `color` as predictors. Print the summary of the fit.

(b) Using the Wald test, test whether the interaction between `weight` and `color` is significant.

(c) Using deviance tests, decide whether the model from part (a) can be simplified.

(d) By looking at the summary of the model in part (a), think of simplifying it by modifying the `color` predictor. Fit this simpler model, say `mod2`.

(e) Plot the Pearson and deviance residuals for model `mod2` from part (d), and comment on the plot. Check the deviance of `mod2`. Based on the residual plots and the deviance, do you think that the model fits the data well?