Exercise 1

Let Y be a count variable which is modeled by a Poisson regression model

$$Y|x \sim Poisson\Big(\lambda(x)\Big)$$

where $\mathbb{E}(Y|x) = \lambda(x) = \exp(\beta_0 + x\beta_1)$ and where x is the covariate of interest.
We assume a random sample with independent observation pairs (y_i, x_i) for $i = 1, ..., n$.
Define $\beta = (\beta_0, \beta_1)^\top$ and $Z_i = (1, x_i)^\top$ so that $Z_i^\top \beta = \beta_0 + x_i\beta_1$.

(a) Write down the log-likelihood for β.

(b) Write down the score equation (= first derivative of the log-likelihood).

(c) Write down the Fisher information matrix.

Exercise 2

Consider the simple linear regression model

$$Y = X\beta + \varepsilon,$$

where $Y = (Y_1, ..., Y_n)^\top$, $\varepsilon = (\varepsilon_1, ..., \varepsilon_n)^\top$ and

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

We assume that x_i are the values of the univariate covariate of interest and that $\varepsilon \sim N(0, \sigma^2 I_n)$ where I_n denotes the n-dimensional unit matrix.

(a) Derive the ordinary least squares (OLS) estimate for β.

(b) Show that the OLS is unbiased and write down the variance of the estimate.

(c) Show that the OLS estimate equals the maximum likelihood estimate.

(d) Assume now that the residuals ε are correlated: $\varepsilon \sim N(0, \Sigma)$ for some invertible covariance matrix $\Sigma \in \mathbb{R}^{n \times n}$. Derive the resulting maximum likelihood estimate and calculate its variance by deriving the Fisher information matrix.

(e) Consider the OLS estimate from a). Give the conditions under which this estimate is still consistent. Justify your statements.

Exercise 3

In a study on the undernutrition of children in developing countries, a random sample of 100 children under 5 years is drawn. It is recorded whether a child is undernourished or not, leading to the binary response variable

$$Y_i = \begin{cases} 1 & \text{child was undernourished,} \\ 0 & \text{otherwise.} \end{cases}$$

As explanatory variables the study recorded

- x_1 family income per head per year (in \$), and
- x_2 indicator: $x_2 = 1$ for families that live in urban or city areas; $x_2 = 0$ for families that live in a rural areas.

The parameter estimates of the resulting Logit model are

Variable	Estimate	p-Value
Intercept	3.3084	0.0000048
x_1	-0.0379	0.0000004
x_2	0.9298	0.2180000

(a) Write down the model equation of the Logit model.

(b) Interpret the effect of x_1 ($\hat{\beta}_1 = -0.0379$). Calculate the odds ratio for undernutrition comparing children living in the city/urban area compared to children living in a rural area.

(c) Interpret the p-values. Which effects are significant on a $\alpha = 0.05$ significance level?

(d) Explain the problem of multiple testing and suggest a solution.

Exercise 4

(a) Consider the following matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \\ 3 & 2 & 1 \end{pmatrix}.$$

Calculate the determinant of A manually, i.e. without software or a calculator.

(b) Assume that $Y \sim N(0, \Sigma)$ whereby the covariance matrix $\Sigma \in \mathbb{R}^{p \times p}$ is of full rank. Moreover, the eigenvalue decomposition of Σ is given by

$$\Sigma = U \Lambda U^\top,$$

where $U \in \mathbb{R}^{p \times p}$ is the matrix of eigenvectors and $\Lambda = diag(\lambda_1, ..., \lambda_p)$ contains the eigenvalues of Σ. Standardize Y. I.e., construct a matrix A such that $\tilde{Y} = AY$ is standard normal:

$$\tilde{Y} \sim N(0, I_p).$$

(c) Construct a univariate random variable $X = a^\top Y$ with $a^\top = (a_1, ..., a_p)$ such that

$$Var(X) = \max_{\|\tilde{a}\|=1} Var(\tilde{a}^\top Y)$$

that is, the variance is maximal for all linear combinations of Y.

Exercise 5

(a) Assume that λ is gamma distributed, i.e.

$$\lambda \sim \frac{\lambda^{k-1}\exp(-\lambda/\theta)}{\theta^k \Gamma(k)},$$

where $k > 0$, $\theta > 0$ are the parameters of the gamma distribution. In short: $\lambda \sim \Gamma(k, \theta)$. For a given λ (i.e. conditional on λ)

$$Y|\lambda \sim Poisson(\lambda).$$

Derive the distribution of λ given Y.

(b) In a Bayesian model, the prior distribution of the parameter θ is denoted by $g(\theta)$. Given θ, Y is observed; i.e., we assume

$$Y|\theta \sim f(y|\theta).$$

Under which condition is the posterior mode (i.e. the maximum of the posterior distribution of θ conditional on Y) equal to the maximum likelihood estimate?