Exercises 9

1. Refer to the material presented in Sections 3.1 and 3.2, and define $z_i = \eta_i + g'(\mu_i)(y_i - \mu_i)$ and $w_i = \{V(\mu_i)g'(\mu_i)^2\}^{-1}$, where $g'(\mu_i) = d\eta_i/d\mu_i$. Show that:

   - $\mathrm{E}(z_i) = \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}$.
   - $var(\mathbf{z}) = \mathbf{W}^{-1}\phi$.
   - $var(\hat{\boldsymbol{\beta}}) = (\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X})^{-1}\phi$.

2. Consider the saturated form of the log-linear model for a 2×2 contingency table given in the lecture notes. With the form of constraints on the parameters used in the course, show that the interaction term
$$\phi_{22} = \log\left(\frac{\pi_{11}\pi_{22}}{\pi_{12}\pi_{21}}\right),$$
where the $\pi_{ij}s$ are cell probabilities.

   Interpret the right hand side of the above equation in terms of an odds ratio and hence give the value of this odds ratio if there is no interaction. Give another explanation of this value of the odds ratio directly from the independence hypothesis.

3. $Y_1, \ldots, Y_N$ are independent random variables such that Y_i has a Poisson distribution with mean μ_i. Show that the joint distribution of $Y_1, \ldots, Y_N$ conditional on their sum being n has a multinomial distribution whose parameters you should obtain.

4. With reference to Section 3.4.2, consider the log-linear model for a 3-way contingency table in which B and C are independent conditional on the category of A.

   (i) Verify the algebraic formula for the fitted values given in Table 3 (Section 3.4.2) of the notes.

   (ii) Write down algebraically the log-linear model which in the `R` notation has predictor A + B + C + A:B + A:C and verify algebraically that it satisfies the formula in part (i).

5. Refer to the material presented in Section 3.5, show that:

   - $\hat{\boldsymbol{\beta}} = (\mathbf{X}^{\mathrm{T}}\mathbf{X} + \mathbf{S})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{y}$, when the response variable is assumed to follow a Gaussian distribution.
   - for a function $f(x_1, x_2) = \sum_k \beta_k b_k(x_1, x_2)$,
$$\int\int \left(\frac{\partial^2 f}{\partial x_1^2}\right)^2 + 2\left(\frac{\partial^2 f}{\partial x_1 \partial x_2}\right)^2 + \left(\frac{\partial^2 f}{\partial x_2^2}\right)^2 dx_1 dx_2 = \boldsymbol{\beta}^{\mathrm{T}}\mathbf{S}\boldsymbol{\beta},$$
where $\mathbf{S}$ is a matrix of known coefficients, similar to that defined in the lecture notes.