Exercises 9 solutions

1.
   - $\mathrm{E}(z_i) = \eta_i + g'(\mu_i)(\mathrm{E}(y_i) - \mu_i) = \eta_i + 0 = \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}$.
   - For each observation $var(z_i) = var(g'(\mu_i)y_i) = g'(\mu_i)^2 var(y_i) = g'(\mu_i)^2 V(\mu_i)\phi = w_i^{-1}\phi$. Because the y_i are assumed to be independent, the z_i are independent too. It follows that $var(\mathbf{z}) = \mathbf{W}^{-1}\phi$.
   - We know that $\hat{\boldsymbol{\beta}} = (\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{z}$. Using standard results on transformation of covariance matrices like those seen in lectures and in Rice, we have that
   $$var(\hat{\boldsymbol{\beta}}) = (\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{W}^{-1}\mathbf{W}\mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X})^{-1}\phi = (\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X})^{-1}\phi.$$

2. The model is
$$\log \mu_{ij} = \lambda + \alpha_i + \beta_j + \phi_{ij}$$
for $i, j = 1, 2$ where $\alpha_1 = \beta_1 = \phi_{1j} = \phi_{i1} = 0$. Hence,
$$\log\left(\frac{\pi_{11}\pi_{22}}{\pi_{12}\pi_{21}}\right) = \log\mu_{11} + \log\mu_{22} - \log\mu_{12} - \log\mu_{21} = \phi_{22}.$$
Let A and B denote two variables. The left hand side in the above is the log of the odds ratio
$$\frac{odds\ on\ category\ 2\ of\ B\ for\ category\ 2\ of\ A}{odds\ on\ category\ 2\ of\ B\ for\ category\ 1\ of\ A}.$$
(Similarly with A and B swapped in this description)
When $\phi_{22} = 0$, there is no interaction and the odds ratio is 1.
Under the independence hypothesis, $\pi_{ij} = \pi_{i+}\pi_{+j}$ and the odds ratio is 1.

3. The joint probability mass function of $Y_1, \ldots, Y_N$ is
$$\prod_{i=1}^{N} \frac{\mu_i^{y_i}\exp(-\mu_i)}{y_i!}.$$
The distribution of $Y_+ = \sum_{i=1}^N Y_i$ is a Poisson with mean $\mu_+ = \sum_{i=1}^N \mu_i$ and so
$$P(Y_+ = n) = \frac{\mu_+^n \exp(-\mu_+)}{n!}.$$
Dividing these two pmf's gives the required conditional pmf, i.e.
$$n!\prod_{i=1}^{N}\left(\frac{\mu_i}{\mu_+}\right)^{y_i}/y_i!,$$
which is multinomial with index n and cell probabilities μ_i/μ_+.

4. (i) From the lecture notes, under this model the cell probabilities are given by
$$\pi_{ijk} = \pi_{ij+}\pi_{i+k}/\pi_{i++}.$$
For the expected values, $\mu_{ijk} = n\pi_{ijk}$, $\mu_{ij+} = n\pi_{ij+}$, etc. Substitute for the π in the equation above to give $\mu_{ijk} = \mu_{ij+}\mu_{i+k}/\mu_{i++}$ as in Table 3 of the notes.

   (ii) Algebraically, the log-linear model is
$$\log\mu_{ijk} = \lambda + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \gamma_k + (\alpha\gamma)_{ik}$$
for $i = 1, \ldots, I; j = 1, \ldots, J; k = 1, \ldots, K$. Hence, $\mu_{ij+}\mu_{i+k}/\mu_{i++}$ is
$$\frac{e^{\lambda+\alpha_i+\beta_j+(\alpha\beta)_{ij}}\left(\sum_k e^{\gamma_k+(\alpha\gamma)_{ik}}\right)e^{\lambda+\alpha_i+\gamma_k+(\alpha\gamma)_{ik}}\left(\sum_j e^{\beta_j+(\alpha\beta)_{ij}}\right)}{e^{\lambda+\alpha_i}\left(\sum_{jk} e^{\beta_j+(\alpha\beta)_{ij}+\gamma_k+(\alpha\gamma)_{ik}}\right)}$$
which is equal to $e^{\lambda+\alpha_i+\beta_j+(\alpha\beta)_{ij}+\gamma_k+(\alpha\gamma)_{ik}} = \mu_{ijk}$ as required.

5.

- Maximisation of (3.27) is equivalent to minimisation of $S(\boldsymbol{\beta}) = ||\mathbf{y} - \mathbf{X}\boldsymbol{\beta}||^2 + \boldsymbol{\beta}^{\mathrm{T}}\mathbf{S}\boldsymbol{\beta}$. $S(\boldsymbol{\beta})$ can be written as $\mathbf{y}^{\mathrm{T}}\mathbf{y} - 2\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{y} + \boldsymbol{\beta}^{\mathrm{T}}(\mathbf{X}^{\mathrm{T}}\mathbf{X} + \mathbf{S})\boldsymbol{\beta}$. Now, differentiating this w.r.t. $\boldsymbol{\beta}$ and setting to zero gives $(\mathbf{X}^{\mathrm{T}}\mathbf{X} + \mathbf{S})\hat{\boldsymbol{\beta}} = \mathbf{X}^{\mathrm{T}}\mathbf{y}$ which produces the required result.

- $$\int\int \left(\frac{\partial^2 f}{\partial x_1^2}\right)^2 + 2\left(\frac{\partial^2 f}{\partial x_1 \partial x_2}\right)^2 + \left(\frac{\partial^2 f}{\partial x_2^2}\right)^2 dx_1 dx_2$$

  is equivalent to

  $$\int\int \left(\frac{\partial^2 f}{\partial x_1^2}\right)^2 dx_1 dx_2 + 2\int\int \left(\frac{\partial^2 f}{\partial x_1 \partial x_2}\right)^2 dx_1 dx_2 + \int\int \left(\frac{\partial f^2}{\partial x_2^2}\right)^2 dx_1 dx_2.$$

  Similarly as in the lecture notes, this can be written as $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{S}\boldsymbol{\beta}$, with $\mathbf{S}$ given as

  $$\int\int \mathbf{b}''_{x_1,x_1}(x_1,x_2)\mathbf{b}''_{x_1,x_1}(x_1,x_2)^{\mathrm{T}} + 2\mathbf{b}''_{x_1,x_2}(x_1,x_2)\mathbf{b}''_{x_1,x_2}(x_1,x_2)^{\mathrm{T}} + \mathbf{b}''_{x_2,x_2}(x_1,x_2)\mathbf{b}''_{x_2,x_2}(x_1,x_2)^{\mathrm{T}} dx_1 dx_2,$$

  where the k^{th} component of $\mathbf{b}''_{x_j,x_h}(x_1,x_2)$ is $\partial^2 b_k / \partial x_j \partial x_h$, for $j, h = 1, 2$.