i) $Y_{ij} = \alpha + a_i + (\beta + b_i)x_{ij} + \varepsilon_{ij}$

$\begin{pmatrix} a_i \\ b_i \end{pmatrix} \sim N(0, \Sigma)$ where $\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$

$\varepsilon_{ij} \sim N(0, \sigma^2)$

[Not necessary ~~at~~ to label each element in Σ; enough to say it is 2x2 nondiagonal matrix]

ii) Fixed effects are used to define the average intercept (which is the alcohol use at age 15) & gradient, which are the primary quantities of interest. Random effects are used to ~~indicate~~ represent how individuals deviate from the ~~expe~~ average case

iii) $\hat{\alpha} = 0.922$ $\quad\hat{\sigma}_1^2 = 0.648$

$\hat{\beta} = 0.271$ $\quad\hat{\sigma}_2^2 = 0.155$

$\rho = 0.26$

$\sigma^2 = 0.337$

(iv) $Cov(Y_{i1}, Y_{i3}) = Cov(\alpha + a_i + (\beta + b_i)(-1) + \varepsilon_{i1},$

$\alpha + a_i + (\beta + b_i)(1) + \varepsilon_{i3})$

$= Cov(a_i - b_i, a_i + b_i)$ ~~Var~~

$= Var(a_i) + Cov(a_i, b_i) - Cov(a_i, b_i)$

$- Var(b_i)$

$= Var(a_i) - Var(b_i)$

$= 0.648 - 0.155$

$= 0.493$

(v) $\hat{\beta} = \frac{1}{IJ} \sum_{i=1}^{I} \sum_{j=1}^{3} Y_{ij}$ $J = 3$

$Var(\hat{\beta}) = \frac{1}{I^2 J^2} Var\left(\sum_{i=1}^{I} \sum_{j=1}^{3} Y_{ij}\right)$

but $\sum_{j=1}^{3} Y_{ij} \perp\!\!\!\perp \sum_{j=1}^{3} Y_{i'j}$

for $i \neq i'$

$= \frac{1}{IJ^2} Var\left(\sum_{j=1}^{3} Y_{ij}\right)$

(v ctd)

$$= \frac{1}{IJ^2} Var\left((a_1 - b_1) + a_1 + (a_1 + b_1) + \sum_{j=1}^{3} \varepsilon_{ij}\right)$$

$$= \frac{1}{IJ^2}\left(Var(3a_1) + Var\left(\sum_{j=1}^{3} \varepsilon_{ij}\right)\right)$$

$$= \frac{1}{IJ^2}\left(9\sigma_1^2 + 3\sigma^2\right)$$

$$= \frac{1}{IJ}\left(J\sigma_1^2 + \sigma^2\right)$$

(vi) We are testing whether we need a correlation term between the two random effects

i.e. ~~whether~~ $H_0 : \rho = 0$

vs $H_1 : \rho \neq 0$

This is an example of a bootstrap generalized likelihood ratio test.

We have found a p-value of 0.53 which suggests there is no evidence to reject H_0.

MAS473 Extended linear models
2015-16 Exam Solutions

(2)(i) Could use the R command glm($T \sim \log D + A +$ B, family=poisson) **[2 marks, 1 for log 1 for the rest]** . The Poisson distribution for the counts is specified by the family argument **[1 mark]**. It fits a generalised linear model with a log link (the default)**[1 mark]**.

(2)(ii) $L = \prod_{ij} \frac{\mu_{ij}{}^{t_{ij}} e^{-\mu_{ij}}}{t_{ij}!}$ **[2 marks, 1 for Poisson, 1 for product over i, j]** so $l = \sum_{ij} t_{ij} log\mu_{ij} - \mu_{ij} + \text{constant}$ **[1 mark]**

(2)(iii)

$$\begin{aligned} l &= \sum_{ij} \left\{ t_{ij} \left(\gamma log d_{ij} + \alpha_i + \beta_j \right) - d_{ij}{}^{\gamma} e^{\alpha_i + \beta_j} \right\} + \text{constant} \textbf{[1 mark]} \\ \frac{\partial l}{\partial \alpha_i} &= \sum_{j} \left\{ t_{ij} - d_{ij}{}^{\gamma} e^{\alpha_i + \beta_j} \right\} \textbf{[2 marks, 1 for sum over j, one for correct diff]} \end{aligned}$$

At the mle $\frac{\partial l}{\partial \alpha_i} = 0$ **[1 mark]** so $\sum_j t_{ij} = \sum_j \hat{t}_{ij}$ **[1 mark]**.

(2)(iv) In model 2 the gradient is forced to be 1 whilst in model 1 it would be estimated via the parameter γ **[2 marks]**. In model 2 the intercept just depends on the origin of travel, whilst in model 1 it depends on both the origin and the destination of travel. **[2 marks]**.

(2)(vi) Find the difference in the residual deviances for models 1 and 3 $(w_1 - w_3)$ **[1 mark]**. Calculate the difference in the degrees of freedom between models 1 and 3 $(d_1 - d_3)$ **[1 mark]**. Calculate the p value as $Pr\left(\chi^2_{d_1 - d_3} \geq w_1 - w_3\right)$ **[1 mark]**. Small p values indicate the null that null is unlikely to be true.**[1 mark]**

(3)(i) $\beta_0 + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik}$ **[1 mark]** where i is 'diet', j is 'drink' and k is 'low'**[1 mark]**.

$$\begin{aligned} \alpha_1 &= \beta_1 = \gamma_1 = 0 \\ (\alpha\beta)_{ij} &= 0 \quad \text{if } i = 1 \text{ or } j = 1 \\ (\alpha\gamma)_{ik} &= 0 \quad \text{if } i = 1 \text{ or } k = 1 \end{aligned}$$

[1 mark for each line]

(3)(ii) Parameters are $\beta_0, \alpha_2, \alpha_3, \beta_2, \gamma_2, (\alpha\beta)_{22}, (\alpha\beta)_{32}, (\alpha\gamma)_{22}, (\alpha\gamma)_{32}$ **[1 mark]**
So $n - p = 12 - 9 = 3$.**[1 mark]**

(3)(iii) M1 is nested within M2. M2 is nested within both M3 and M4.**[1 mark]**

$M1 \to M2$	$\Delta D = 5.72$	$\Delta df = 1$	$p < 0.05$
$M2 \to M3$	$\Delta D = 5.72$	$\Delta df = 1$	$p < 0.05$
$M2 \to M4$	$\Delta D = 0.53$	$\Delta df = 2$	$p > 0.05$

so Model 3 looks best **[3 marks]**.

A residual deviance of 2.88 on 4 df shows little evidence of a poor fit **[1 mark]** so conclude that the distribution of low birth weight babies varies with drinking habit but not diet during pregnancy. **[1 mark]**

(3)(iv) Require $E(Y_{122})$ **[1 mark]**. Collapse over diet $\hat{\mu}_{122} = n_{12}\bar{\pi}_{122} = 50 \times 40/93 = 21.5$ **[3 marks; 1 for n, 1 for pi, 1 for answer]**

(3)(v) This is just a confidence interval for $E(Y_{111})$ (i.e. a confidence interval for β_0) **[1 mark]**. So calculate a 95% CI for $\log\mu_{111}$ using $\beta_0 \pm 2 \times$ standard error from `summary` output **[1 mark]** and then take e to the power of the interval ends **[1 mark]**.