Theory Exam Section II 2012

1). n responses iid

$Y_{i1} \to$ primary infection (binary)

$Y_{i2} \to$ secondary infection (binary)

$\alpha = P(Y_{i1}=1)$; $\beta = P(Y_{i2}=1 \mid Y_{i1}=1)$

$\alpha, \beta \in (0,1)$

X_1 patients w/ $(Y_{i1}=1, Y_{i2}=1)$

X_2 patients w/ $(Y_{i1}=1, Y_{i2}=0)$

X_3 patients w/ $(Y_{i1}=0, Y_{i2}=0)$

$X_1 + X_2 + X_3 = n$

<u>Multivariate Dist</u>

(a) Does the dist of the data have the form of an exponential family? Give details.

$$f(\underset{\sim}{x}) = \binom{n}{x_1, x_2, x_3} \pi_1^{x_1} \pi_2^{x_2} \pi_3^{x_3}$$

$$\pi_1 = P(Y_{i1}=1, Y_{i2}=1) = P(Y_{i2}=1 \mid Y_{i1}=1)P(Y_{i1}=1)$$
$$= \alpha\beta$$
$$\pi_2 = P(Y_{i1}=1, Y_{i2}=0) = (1-P(Y_{i2}=1 \mid Y_{i1}=1))P(Y_{i1}=1)$$
$$= \alpha(1-\beta)$$
$$\pi_3 = P(Y_{i1}=0, Y_{i2}=0) = 1-\alpha$$

Check: $\pi_1 + \pi_2 + \pi_3 = \cancel{\alpha\beta} + \cancel{\alpha} - \cancel{\alpha\beta} + 1 - \cancel{\alpha} = 1$ ✓

$$f(\underset{\sim}{x}) = \binom{n}{x_1\, x_2\, x_3} (\alpha\beta)^{x_1} (\alpha(1-\beta))^{x_2} (1-\alpha)^{x_3}$$

$$= \binom{n}{x_1\, x_2\, x_3} \exp\left(x_1 \log(\alpha\beta) + x_2 \log(\alpha(1-\beta)) + \overset{n-x_1-x_2}{x_3} \log(1-\alpha) \right)$$

$$\cancel{= \binom{n}{x_1\, x_2\, x_3} \exp\left(\log\alpha\,(x_1 + x_2) + x_1 \log\beta + x_2 \log(1-\beta) + x_3 \log(1-\alpha) \right)}$$

$$= \binom{n}{x_1\, x_2\, x_3} \exp\left(x_1 (\log(\alpha\beta) - \log(1-\alpha)) + x_2 (\log(\alpha(1-\beta)) - \log(1-\alpha)) + n \log(1-\alpha) \right)$$

$$= \exp\left(x_1 \log\left(\frac{\alpha\beta}{1-\alpha}\right) + x_2 \log\left(\frac{\alpha(1-\beta)}{1-\alpha}\right) + n\log(1-\alpha) - c(y) \right)$$

form of $\exp\left(\sum_{i=1}^{k} x_i \theta_i - b(\theta) - c(x) \right)$, $k = 2$

$$c(x) = -\log \binom{n}{x_1\, x_2\, x_3}$$

$$b(\theta) = -n \log(1-\alpha)$$

Note: $$\begin{cases} \theta_1 = \log(\overset{\log\alpha + \log\beta}{\alpha\beta}) - \log(1-\alpha) \\ \theta_2 = \log\alpha + \log(1-\beta) - \log(1-\alpha) \end{cases}$$

$$(\alpha\beta) + \alpha(1-\beta) + (1-\alpha) = 1$$

$$\Rightarrow \left(\frac{\alpha\beta}{1-\alpha}\right) + \left(\frac{\alpha(1-\beta)}{1-\alpha}\right) + 1 = \frac{1}{1-\alpha}$$

$$\Rightarrow e^{\theta_1} + e^{\theta_2} + 1 = \frac{1}{1-\alpha}$$

$$\Rightarrow (1-\alpha) = \frac{1}{e^{\theta_1} + e^{\theta_2} + 1}$$

$$b(\theta) = +n \log(e^{\theta_1} + e^{\theta_2} + 1)$$

(b) Derive the MLE's of α & B

$$\log f(\underset{\sim}{x}) = \log \binom{n}{x_1\, x_2\, x_3} + x_1 \log \alpha + x_1 \log B +$$

$$+ x_2 \log \alpha + x_2 \log (1-B) + x_3 \log (1-\alpha)$$

$$\frac{\partial}{\partial \alpha} \log f(\underset{\sim}{x}) = \frac{x_1}{\alpha} + \frac{x_2}{\alpha} - \frac{x_3}{1-\alpha} \overset{\text{set}}{=} 0$$

$$\Rightarrow \frac{(x_1 + x_2)(1-\alpha) - x_3(\alpha)}{\alpha\cancel{(1-\alpha)}} = 0$$

$$\Rightarrow (x_1 + x_2) - \alpha \underbrace{(x_1 + x_2 + x_3)}_{n} = 0$$

$$\Rightarrow \boxed{\hat{\alpha} = \frac{x_1 + x_2}{n}}$$

$$\frac{\partial}{\partial B} \log f(\underset{\sim}{x}) = \frac{x_1}{B} - \frac{x_2}{1-B} \overset{\text{set}}{=} 0$$

$$\Rightarrow \frac{x_1(1-B) - x_2 B}{\cancel{B(1-B)}} = 0$$

$$\Rightarrow x_1 - B(x_1 + x_2) = 0$$

$$\Rightarrow \boxed{\hat{B} = \frac{x_1}{x_1 + x_2} = \frac{x_1}{n - x_3}}$$

(c) Derive the asymptotic covariance matrix of the estimators derived in (b)

By MLE theory,

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N(0, I(\theta)^{-1})$$

$$I(\theta) = \lim_{n\to\infty} \frac{1}{n} I_n(\theta)$$

$$I_n(\theta) = E[-\partial^2/\partial\theta^2 \, l(\theta)]$$

$$\frac{\partial^2}{\partial\alpha^2} l(\alpha, \beta) = -\frac{x_1}{\alpha^2} - \frac{x_2}{\alpha^2} - \frac{x_3}{(1-\alpha)^2}(-1)(-1)$$

$$= \frac{-1}{\alpha^2}(x_1 + x_2) - \frac{x_3}{(1-\alpha)^2}$$

$$E[-\partial^2/\partial\alpha^2 \, l(\alpha,\beta)] =$$

$$= \frac{E[x_1 + x_2]}{\alpha^2} - \frac{E[x_3]}{(1-\alpha)^2}$$

$$E[x_1] = n(\alpha\beta)$$

$$E[x_2] = n(\alpha)(1-\beta) = n\alpha - n\alpha\beta$$

$$E[x_1 + x_2] = \cancel{n\alpha\beta} + n\alpha - \cancel{n\alpha\beta} = n\alpha$$

$$E[x_3] = n(1-\alpha)$$

$$= \frac{n\alpha}{\alpha^2} - \frac{n(1-\alpha)}{(1-\alpha)^2}$$

$$= \frac{n}{\alpha} - \frac{n}{(1-\alpha)}$$

$$\lim_{n\to\infty} \frac{1}{n}\left(\frac{n}{\alpha} - \frac{n}{1-\alpha}\right) = \frac{1}{\alpha} - \frac{1}{1-\alpha} = \frac{(1-\alpha) - \alpha}{\alpha(1-\alpha)} = \frac{1}{\alpha(1-\alpha)}$$

$$\frac{\partial^2}{\partial\alpha\partial\beta} \ell(\alpha,\beta) = 0$$

$$\frac{\partial^2}{\partial\beta^2} \ell(\alpha,\beta) = -\frac{x_1}{\beta^2} - \frac{x_2}{(1-\beta)^2}(-1)(-1)$$

$$= -\frac{x_1}{\beta^2} - \frac{x_2}{(1-\beta)^2}$$

$$E[-\partial^2/\partial\beta\, \ell(\alpha,\beta)]$$

$$= \frac{E(x_1)}{\beta^2} + \frac{E(x_2)}{(1-\beta)^2}$$

$$= n\left(\frac{\alpha\beta}{\beta^2} + \frac{\alpha(1-\beta)}{(1-\beta)^2}\right)$$

$$= \frac{n\alpha}{\beta} + \frac{n\alpha}{(1-\beta)}$$

$$= n\alpha\left(\frac{1-\beta+\beta}{\beta(1-\beta)}\right)$$

$$= n\alpha\left(\frac{1}{\beta(1-\beta)}\right)$$

$$\lim_{n\to\infty} \frac{1}{n}\left(\frac{n\alpha}{\beta(1-\beta)}\right) = \frac{\alpha}{\beta(1-\beta)}$$

$$I(\alpha,\beta) = \begin{bmatrix} 1/\alpha(1-\alpha) & 0 \\ 0 & \alpha/\beta(1-\beta) \end{bmatrix}$$

$$[I(\alpha,\beta)]^{-1} = \begin{bmatrix} \alpha(1-\alpha) & 0 \\ 0 & \beta(1-\beta)/\alpha \end{bmatrix}$$

Therefore,

$$\sqrt{n}\left(\begin{bmatrix}\hat{\alpha} \\ \hat{\beta}\end{bmatrix} - \begin{bmatrix}\alpha \\ \beta\end{bmatrix}\right) \xrightarrow{d} N\left(\underset{\sim}{0}, \begin{bmatrix} \alpha(1-\alpha) & 0 \\ 0 & \beta(1-\beta)/\alpha \end{bmatrix}\right)$$

(e) Derive the LRT test statistic for testing

$H_0: \alpha - \beta = 0$ vs. $H_A: \alpha - \beta \neq 0$

$\Leftrightarrow H_0: \alpha = \beta$ vs. $H_A: \alpha \neq \beta$.

$$\Lambda = \frac{\sup_{\theta \in \Theta_0} L(\theta)}{\sup_{\theta \in \Theta} L(\theta)} \qquad \theta = (\alpha, \beta)^T$$

Θ_0: $\alpha = \beta$

Θ : general case

Case 1: Θ_0 (H_0)

$$\ell(\alpha,\beta) = \log\binom{n}{x_1 x_2 x_3} + x_1 \log\alpha + x_1 \log\beta +$$

$$+ x_2 \log\alpha + x_2 \log(1-\beta) + x_3 \log(1-\alpha)$$

– plug in $\alpha = \beta$ (as in H_0 case)

$$= \log\binom{n}{x_1 x_2 x_3} + x_1 \log\alpha + x_1 \log\alpha +$$

$$+ x_2 \log\alpha + x_2 \log(1-\alpha) + x_3 \log(1-\alpha)$$

$$= \log c(\underset{\sim}{x}) + (2x_1 + x_2)\log\alpha + (x_2 + x_3)\log(1-\alpha)$$

$$\frac{\partial}{\partial\alpha}\ell(\alpha) = \frac{2x_1 + x_2}{\alpha} - \frac{x_2 + x_3}{1-\alpha} \overset{set}{=} 0$$

$$\Rightarrow \frac{(1-\alpha)(2x_1 + x_2) - \alpha(x_2 + x_3)}{\alpha\cancel{(1-\alpha)}} = 0$$

$$\Rightarrow 2x_1 + x_2 - \alpha(2x_1 + 2x_2 + x_3) = 0$$

$$\Rightarrow \tilde{\alpha} = \frac{2x_1 + x_2}{2x_1 + 2x_2 + x_3}$$

note: $x_1 + x_2 + x_3 = n$

$$\boxed{\tilde{\alpha} = \frac{2x_1 + x_2}{2n - x_3}}$$

$$\Lambda = \frac{L(\alpha = \tilde{\alpha}, \beta = \tilde{\alpha})}{L(\alpha = \hat{\alpha}, \beta = \hat{\beta})}$$

↑ from general MLE case

$$= \frac{(\tilde{\alpha})^{2x_1} \cdot (\tilde{\alpha}(1-\tilde{\alpha}))^{x_2} \cdot (1-\tilde{\alpha})^{x_3}}{(\hat{\alpha}\hat{\beta})^{x_1} (\hat{\alpha}(1-\hat{\beta}))^{x_2} (1-\hat{\alpha})^{x_3}} < K$$

$$= \frac{\left(\frac{2x_1 + x_2}{n - x_3}\right)^{2x_1} \cdots}{\left(\frac{x_1 + x_2}{n}\right)^{x_1}}$$

Likelihood ratio test:

$\Lambda < K \Leftrightarrow -2\log\Lambda > K^*$

$$\log\Lambda = 2[\ell(\hat{\alpha}, \hat{\beta}) - \ell(\tilde{\alpha}, \tilde{\alpha})]$$

- plug in $\alpha = \hat{\alpha}$, $\beta = \hat{\beta}$ into $\ell(\alpha, \beta)$
- plug $\alpha = \tilde{\alpha}$, $\beta = \tilde{\alpha}$ into $\ell(\alpha, \beta)$
- find difference & then multiply by 2.

(f) Derive the score test for the hypothesis in part (e)

Score test:

$$SC_n = (\dot{\ell}(\tilde{\theta}))^T I_n(\tilde{\theta})^{-1} (\dot{\ell}(\tilde{\theta}))$$

→ evaluated at $\tilde{\theta}$ = MLE under the hypothesis

$$I_n(\theta)^{-1} = [n I(\theta)]^{-1} = \frac{1}{n} I(\theta)^{-1}$$

$$\ell(\tilde{\theta}) = \ell(\tilde{\alpha}, \tilde{\alpha}) =$$

$$= \log\binom{n}{x_1\, x_2\, x_3} + x_1 \log\left(\frac{\tilde{\alpha}^2}{1-\tilde{\alpha}}\right) + x_2 \log\left(\frac{\tilde{\alpha}\cancel{(1-\tilde{\alpha})}}{\cancel{1-\tilde{\alpha}}}\right) + n\log(1-\tilde{\alpha})$$

$$= \log\binom{n}{x_1\, x_2\, x_3} + x_1 \log\left(\frac{\tilde{\alpha}^2}{1-\tilde{\alpha}}\right) + x_2 \log\tilde{\alpha} + n\log(1-\tilde{\alpha})$$

$$\tilde{\alpha} = \frac{2x_1 + x_2}{2n - x_3} \rightarrow \text{plug in}$$

$$I_n(\theta)^{-1} = \begin{bmatrix} \frac{\tilde{\alpha}(1-\tilde{\alpha})}{n} & 0 \\ 0 & \frac{(1-\tilde{\alpha})}{n} \end{bmatrix}$$

$$\frac{\tilde{\alpha}(1-\tilde{\alpha})}{n\tilde{\alpha}} = \frac{1-\tilde{\alpha}}{n}$$

$$\text{-plug in } \tilde{\alpha} = \frac{2x_1 + x_2}{2n - x_3}$$

→

$$\dot{\ell}(\tilde{\theta}) = \begin{bmatrix} \partial/\partial\alpha \, \ell(\alpha,\beta) \big|_{\alpha,\beta=\tilde{\alpha}} \\ \partial/\partial\beta \, \ell(\alpha,\beta) \big|_{\alpha,\beta=\tilde{\alpha}} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{x_1}{\tilde{\alpha}} + \frac{x_2}{\tilde{\alpha}} - \frac{x_3}{1-\tilde{\alpha}} \\ \frac{x_1}{\tilde{\alpha}} - \frac{x_2}{1-\tilde{\alpha}} \end{bmatrix}$$ plug in $\tilde{\alpha} = \frac{2x_1 + x_2}{2n - x_3}$

$$SC_n = (\dot{\ell}(\tilde{\alpha}))^T (I_n(\tilde{\alpha}))^{-1} \dot{\ell}(\tilde{\alpha})$$

Reject H_0 if $SC_n > \chi^2(df=1, 1-\alpha)$

(g) Derive the Wald test statistic for the hypotheses in part (e)

Wald test:

$$(\hat{\theta}-\theta_0)^T I_n(\hat{\theta})(\hat{\theta}-\theta_0)$$

$$\Rightarrow (R\hat{\theta}-R\theta_0)^T[R\, I_n(\hat{\theta})^{-1}R^T]^{-1}(R\hat{\theta}-R\theta_0)$$

$$R\hat{\theta} = [1 \ -1]\begin{bmatrix}\hat{\alpha}\\ \hat{\beta}\end{bmatrix} = \hat{\alpha}-\hat{\beta}$$

$$R\theta_0 = [1 \ -1]\begin{bmatrix}\alpha\\ \beta\end{bmatrix} = \alpha-\beta=0$$

$$R\, I_n(\hat{\theta})^{-1}R^T = [1 \ -1]\begin{bmatrix}\hat{\alpha}(1-\hat{\alpha})/n & 0\\ 0 & \hat{\beta}(1-\hat{\beta})/n\hat{\alpha}\end{bmatrix}\begin{bmatrix}1\\ -1\end{bmatrix}$$

$$= \begin{bmatrix}\hat{\alpha}(1-\hat{\alpha})/n & -\hat{\beta}(1-\hat{\beta})/n\hat{\alpha}\end{bmatrix}\begin{bmatrix}1\\ -1\end{bmatrix}$$

$$= \frac{\hat{\alpha}(1-\hat{\alpha})}{n} + \frac{\hat{\beta}(1-\hat{\beta})}{n\hat{\alpha}}$$

$$(R\, I_n(\hat{\theta})R^T)^{-1} = \left[\frac{\hat{\alpha}(1-\hat{\alpha})}{n} + \frac{\hat{\beta}(1-\hat{\beta})}{n\hat{\alpha}}\right]^{-1}$$

Wald test statistic:

$$\frac{(\hat{\alpha}-\hat{\beta})^2}{\left[\frac{\hat{\alpha}(1-\hat{\alpha})}{n} + \frac{\hat{\beta}(1-\hat{\beta})}{n\hat{\alpha}}\right]}$$

plug in $\hat{\alpha}$ & $\hat{\beta}$ results from part (b) (MLE's under general case)

$$\hat{\alpha} = \frac{X_1+X_2}{n} \qquad \hat{\beta} = \frac{X_1}{n-X_3}$$

(h) Conditional likelihood

nuisance $= \alpha$ (interest $= \beta$)

$$l_c(\beta) = l(\text{joint dist}) - l(\text{suff. stat of } \alpha)$$

$$f(\underset{\sim}{x}) = \exp\Big(x_1 \log\left(\frac{\alpha}{1-\alpha}\right) + x_1 \log \beta + x_2 \log\left(\frac{\alpha}{1-\alpha}\right) + x_2 \log(1-\beta) + - b(\theta) - c(\underset{\sim}{x}) \Big)$$

$$= \exp\left((x_1+x_2) \log\left(\frac{\alpha}{1-\alpha}\right) + x_1 \log \beta + x_2 \log(1-\beta) - b(\theta) - c(\underset{\sim}{x}) \right)$$

Sufficient statistic for α is $(x_1 + x_2)$

$$(X_1, X_2, X_3) \sim \text{Mult}(\alpha\beta, \alpha(1-\beta), 1-\alpha, n)$$

Marginally,

$$(X_1 + X_2, X_3) \sim \text{Mult}(n, (\alpha\beta + \alpha(1-\beta)), 1-\alpha)$$

$$\Rightarrow X_1 + X_2 \sim \text{Bin}(n, \alpha)$$

$$l(\text{joint dist}) - l(\text{suff stat of } \alpha)$$

$$= l(\alpha, \beta) - \log\left[\left(\frac{n!}{(x_1+x_2)!\, x_3!} \right) \alpha^{(x_1+x_2)} (1-\alpha)^{x_3} \right]$$

note: $n - x_1 - x_2 = x_3$

$$= \log\left(\frac{\cancel{n!}}{x_1!\, x_2!\, \cancel{x_3!}} \right) + x_1 \cancel{\log \alpha} + x_1 \log \beta + x_2 \cancel{\log \alpha} + x_2 \log(1-\beta) + + x_3 \cancel{\log(1-\alpha)}$$

$$- \log\left(\frac{\cancel{n!}}{(x_1+x_2)!\, \cancel{x_3!}} \right) - \cancel{(x_1+x_2) \log \alpha} - \cancel{x_3 \log(1-\alpha)}$$

$$\boxed{\begin{aligned} &= -\log x_1! - \log x_2! + \log(x_1+x_2)! + x_1 \log \beta + x_2 \log(1-\beta) \\ &= l_c(\beta) \end{aligned}}$$

$$\frac{\partial}{\partial B} \ell_c(B) = \frac{x_1}{B} - \frac{x_2}{1-B} \overset{\text{set}}{=} 0$$

$$\Rightarrow \frac{x_1(1-B) - Bx_2}{B\cancel{(1-B)}} = 0$$

$$\Rightarrow x_1 - B(x_1 + x_2) = 0$$

$$\Rightarrow \boxed{B_c = \frac{x_1}{x_1+x_2} = \frac{x_1}{n-x_3} = \hat{B}}$$

- same as before.

Result is intuitive because:

- same result as before
- ?

Theory Exam Section 2 2012

3. Consider independent observations $y_1, \ldots, y_n$, where $y_i = (y_{i1}, y_{i2})'$ is a bivariate binary random vector such that y_{ij} takes values 0 and 1 for $j = 1, 2$. Suppose that $y_i \sim QE(\theta, \lambda)$, where $QE(\theta, \lambda)$ is a bivariate binary distribution of quadratic exponential form

(*) See Exam #2 from 762 Fall 2017

$$p(y_i|\theta, \lambda) = \Delta(\theta, \lambda)^{-1} \exp\{y_{i1}\theta_1 + y_{i2}\theta_2 + y_{i1}y_{i2}\lambda - C(y_{i1}, y_{i2})\},$$

where $\Delta(\theta, \lambda)$ is a normalizing constant and $C(y_{i1}, y_{i2})$ is a 'shape' function independent of $\theta = (\theta_1, \theta_2)'$ and λ.

✓ (a) Derive both the marginal distribution of y_{i1} and the conditional distribution of y_{i2} given y_{i1}. Specify a sufficient and necessary condition such that y_{i1} and y_{i2} are independent.

✓ (b) Calculate the marginal mean of y_i, denoted by $\mu = (\mu_1, \mu_2)' = E(y_i)$, the marginal product moment of $y_{i1}y_{i2}$, denoted by $\eta_{12} = E(y_{i1}y_{i2})$, and the marginal product centered moment of $\underbrace{(y_{i1} - \mu_1)(y_{i2} - \mu_2)}_{\mathrm{cov}(y_{i1}, y_{i2})}$, denoted by $\sigma_{12} = E\{(y_{i1} - \mu_1)(y_{i2} - \mu_2)\}$.

✓ (c) Calculate the Jacobian of the transformation from the canonical parameters θ and λ to the marginal parameters μ and η_{12}, denoted by $V = \partial(\theta, \lambda)/\partial(\mu, \eta_{12})$. Use V^{-1} to characterize the covariance matrix of $(y_i', y_{i1}y_{i2})'$ and specify a sufficient and necessary condition such that this transformation is one-to-one.

✓ (d) Suppose that we also observe a $p \times 1$ column vector x_i for each i and that conditionally on x_i, $y_i \sim QE(\theta_i, \lambda_i)$, where $\theta_i = (\theta_{i1}, \theta_{i2})'$ and λ_i may depend on x_i, for $i = 1, \ldots, n$. Consider the model

$$E[y_i|x_i] = \mu_i = (\mu_{i1}, \mu_{i2})' = \mu(x_i, \beta), E[(y_{i1} - \mu_{i1})(y_{i2} - \mu_{i2})|x_i] = \sigma_{i12} = \sigma_{12}(x_i, \beta, \alpha),$$

where β is an unknown $p \times 1$ regression parameter and α is an unknown scalar parameter. Derive the likelihood score equations for $(\alpha, \beta')'$ and simplify them using the result obtained in part (c). Please clarify whether such estimating equations explicitly involve $C(y_{i1}, y_{i2})$.

✗ (e) Consider generalized estimation equations for α and β given by

$$\sum_{i=1}^{n} \frac{\partial(\mu_i, \sigma_{i12})}{\partial(\alpha, \beta')} \frac{\partial \ell(y_i|\theta_i, \lambda_i)}{\partial(\theta_i, \lambda_i)} = 0$$

Compare the estimate of $(\alpha, \beta')'$ in part (d) with that in part (e) in terms of the statistical efficiency. To do so, provide an explicit comparison of the asymptotic variances of these estimators.

✓ (f) Will the results in parts (a)-(e) be changed if y_{i1} and y_{i2} are continuous variables instead of binary variables? Please explain. If so, then derive the corresponding results and compare with those obtained above.

762 Take-Home Exam II 2017

3. Suppose that $y_1, \ldots, y_n$ are independent bivariate observations, where $y_i = (y_{i1}, y_{i2})'$ is a bivariate binary random vector such that y_{ij} takes values 0 and 1 for $j = 1, 2$. Suppose that $y_i \sim \mathrm{QE}(\theta, \lambda)$, where $\mathrm{QE}(\theta, \lambda)$ is a bivariate distribution of quadratic exponential form

$$p(y_i|\theta, \lambda) = (\Delta(\theta, \lambda))^{-1} \exp\{y_{i1}\theta_1 + y_{i2}\theta_2 + y_{i1}y_{i2}\lambda - C(y_{i1}, y_{i2})\}, \tag{3}$$

where $\Delta(\theta, \lambda)$ is the normalizing constant and $C(y_{i1}, y_{i2})$ is a function that does not depend on $\theta = (\theta_1, \theta_2)'$ and λ.

✓(a) (2 points) Find an explicit expression for $\Delta(\theta, \lambda)$.

✓ (b) (5 points) Derive the marginal distribution of y_{i1} and the conditional distribution of y_{i2} given y_{i1}. Specify a sufficient and necessary condition for y_{i1} and y_{i2} to be independent.

✓ (c) (7 points) Let $\mu = E(y_i)$, $\eta_{12} = E(y_{i1}y_{i2})$, and $\sigma_{12} = \mathrm{Cov}(y_{i1}, y_{i2})$. Derive explicit expressions for μ, η_{12}, and σ_{12} based on the distribution in (3).

(d) (7 points) Suppose we consider a reparameterization from (θ, λ) to (μ, η_{12}). Calculate the Jacobian of the transformation from the parameters (θ, λ) to the parameters (μ, η_{12}), in which the Jacobian is denoted by $V = \partial(\theta, \lambda)/\partial(\mu, \eta_{12})$. Use V^{-1} to characterize the covariance matrix of $(y_i', y_{i1}y_{i2})'$ and specify a sufficient and necessary condition such that this transformation is one-to-one. Find determinant of V^{-1}

✓(e) (11 points) Suppose that we also observe a $p \times 1$ column vector x_i for each i and that conditionally on x_i, $y_i \sim QE(\theta_i, \lambda_i)$, where $\theta_i = (\theta_{i1}, \theta_{i2})'$ and λ_i may depend on x_i, for $i = 1, \ldots, n$. Consider the model

$$E[y_i|x_i] = \mu_i = (\mu_{i1}, \mu_{i2})' = \mu(x_i, \beta), E[(y_{i1} - \mu_{i1})(y_{i2} - \mu_{i2})|x_i] = \sigma_{i12} = \sigma_{12}(x_i, \beta, \alpha),$$

where β is an unknown $p \times 1$ regression parameter and α is an unknown scalar parameter. Derive the likelihood score equations for (α, β) and the Fisher information matrix for (α, β), and simplify them using the result obtained in part (d). Please clarify whether the score equations explicitly involve $C(y_{i1}, y_{i2})$.

✓(f) (3 points) Will the results in parts (b)-(e) be changed if y_{i1} and y_{i2} are continuous variables instead of binary variables? Please explain. If so, then derive the corresponding results and compare with those obtained above.

Hillary Heiling

Midterm # 2

BIOS 762

Take Home 2017

3). $y_1, \ldots, y_n$ indep. bivariate obs.

$y_i = (y_{i1}, y_{i2})'$

$y_{ij} = \{0, 1\}$ for $j = 1, 2$

$y_i \sim QE(\theta, \lambda)$ = bivariate dist of quadratic exponential form

$p(y_i | \theta, \lambda) =$

$$(\Delta(\theta,\lambda))^{-1} \exp\left(y_{i1}\theta_1 + y_{i2}\theta_2 + y_{i1}y_{i2}\lambda - c(y_{i1}, y_{i2})\right)$$

$\Delta(\theta, \lambda)$ = normalizing constant

$c(y_{i1}, y_{i2})$ = function that does not depend on $\theta = (\theta_1, \theta_2)^T$ & λ.

(a) Find an explicit expression for $\Delta(\theta, \lambda)$

$$\sum_{y_{i1}} \sum_{y_{i2}} p(y_i | \theta, \lambda) = 1$$

Both y_{i1} & y_{i2} can be 0 or 1

4 combinations:

$$\begin{bmatrix} y_{i1} \\ y_{i2} \end{bmatrix} \in \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}$$

Continued. →

$$\sum_{y_{i1}} \sum_{y_{ij}} p(y_i \mid \theta, \lambda)$$

$$= (\Delta(\theta,\lambda))^{-1} \cdot \Big[\exp(\theta_1 + \theta_2 + \lambda - C(1,1)) + \exp(\theta_1 - C(1,0)) + \exp(\theta_2 - C(0,1)) + \exp(-C(0,0)) \Big]$$

$$= (\Delta(\theta,\lambda))^{-1} \left[\frac{\exp(\theta_1 + \theta_2 + \lambda)}{\exp(C(1,1))} + \frac{\exp(\theta_1)}{\exp(C(1,0))} + \frac{\exp(\theta_2)}{\exp(C(0,1))} + \frac{1}{\exp(C(0,0))} \right]$$

$$\Rightarrow \Delta(\theta,\lambda) = \frac{\exp(\theta_1 + \theta_2 + \lambda)}{\exp(C(1,1))} + \frac{\exp(\theta_1)}{\exp(C(1,0))} + \frac{\exp(\theta_2)}{\exp(C(0,1))} + \frac{1}{\exp(C(0,0))}$$

✓

$\boxed{2/2}$

(b) Derive the marginal dist of y_{i1} (i) and the conditional dist of $y_{i2} | y_{i1}$ (ii)

Specify sufficient + necessary condition for y_{i1} + y_{i2} to be indep.

(i) $P(y_{i1}) = \sum_{y_{i2}} P(y_{i1}, y_{i2}) \qquad (y_{i2} = \{0,1\})$

$$= [\Delta(\theta,\lambda)]^{-1} \left[\exp(y_{i1}\theta_1 + \theta_2 + y_{i1}\lambda - c(y_{i1}, 1)) + \exp(y_{i1}\theta_1 - c(y_{i1}, 0)) \right]$$

(ii) $P(y_{i2}|y_{i1}) = \dfrac{P(y_{i1}, y_{i2})}{P(y_{i1})}$

$$= \frac{[\exp(\cancel{y_{i1}\theta_1} + y_{i2}\theta_2 + y_{i1}y_{i2}\lambda - c(y_{i1}, y_{i2}))]}{[\exp(\cancel{y_{i1}\theta_1} + \theta_2 + y_{i1}\lambda - c(y_{i1}, 1)) + \exp(\cancel{y_{i1}\theta_1} - c(y_{i1}, 0))]}$$

$$= \frac{\exp(\cancel{y_{i2}\theta_2} + y_{i2}\theta_2 + y_{i1}y_{i2}\lambda - c(y_{i1}, y_{i2}))}{\exp(\theta_2 + y_{i1}\lambda - c(y_{i1}, 1)) + \exp(-c(y_{i1}, 0))}$$

(iii) When is y_{i1} + y_{i2} indep?

Options: $P(y_{i1}, y_{i2}) = P(y_{i1}) P(y_{i2})$

OR

$P(y_{i1}|y_{i2}) = P(y_{i1})$

OR $P(y_{i2}|y_{i1}) = P(y_{i2})$.

Continued. →

When is $P(y_{i1}, y_{i2}) = P(y_{i1})\,P(y_{i2})$?

Let $\lambda = \infty$ and $C(y_{i1}, y_{i2}) =$ constant K for all combinations of y_{i1}, y_{i2}.

$\Delta(\theta, \lambda) =$
$$e^{-K}\left[\exp(\theta_1 + \theta_2) + \exp(\theta_1) + \exp(\theta_2) + 1\right]$$

$p(y_i \mid \theta, \lambda) =$
$$[\Delta(\theta, \lambda)]^{-1} \exp(y_{i1}\theta_1 + y_{i2}\theta_2 - K)$$
$$= \frac{\cancel{e^{-K}} \exp(y_{i1}\theta_1 + y_{i2}\theta_2)}{\cancel{e^{-K}} \underbrace{\left[\exp(\theta_1 + \theta_2) + \exp(\theta_1) + \exp(\theta_2) + 1\right]}_{\gamma}}$$
$$= \frac{\exp(y_{i1}\theta_1)\exp(y_{i2}\theta_2)}{\gamma}$$

Check: Does this $= P(y_{i1})\,P(y_{i2})$?

$P(y_{i1}) =$
$$(\Delta(\theta, \lambda))^{-1}\left[\exp(y_{i1}\theta_1 + \theta_2 - K) + \exp(y_{i1}\theta_1 - K)\right]$$
$$= \frac{\cancel{e^{-K}}\left[\exp(y_{i1}\theta_1)(\exp(\theta_2) + 1)\right]}{\cancel{e^{-K}}\left[\exp(\theta_1 + \theta_2) + \exp(\theta_1) + \exp(\theta_2) + 1\right]}$$

$$P(y_{i2}) = \sum_{y_{i1}} P(y_{i1}, y_{i2}) \qquad (y_{i1} = \{0, 1\})$$
$$= (\Delta(\theta, \lambda))^{-1}\left[\exp(\theta_1 + y_{i2}\theta_2 - K) + \exp(y_{i2}\theta_2 - K)\right]$$
$$= \frac{\cancel{e^{-K}}\left[\exp(y_{i2}\theta_2)(\exp(\theta_1) + 1)\right]}{\cancel{e^{-K}}\left[\exp(\theta_1 + \theta_2) + \exp(\theta_2) + \exp(\theta_2) + 1\right]}$$

Continued $\longrightarrow$

$P(y_{i1})\, P(y_{i2})$

$$= \frac{\exp(y_{i1}\theta_1 + y_{i2}\theta_2)\left[(\exp(\theta_1)+1)(\exp(\theta_2)+1)\right]}{\left[\exp(\theta_1+\theta_2) + \exp(\theta_2) + \exp(\theta_1) + 1\right]^2}$$

Note: $[\exp(\theta_1)+1][\exp(\theta_2)+1]$

$= \exp(\theta_1+\theta_2) + \exp(\theta_1) + \exp(\theta_2) + 1$

$$= \frac{\exp(y_{i1}\theta_1 + y_{i2}\theta_2)}{\exp(\theta_1+\theta_2) + \exp(\theta_1) + \exp(\theta_2) + 1}$$

$= P(y_{i1}, y_{i2})$ given $\lambda = 0$, $c(y_{i1}, y_{i2}) = K$ ✓

Therefore, under these conditions, y_{i1} & y_{i2} are independent. ✓

Dan

$\boxed{4.0/5}$

(−0.5) Replace $c(y_{i1}, y_{i2}) = K$ for $c(y_{i1}, y_{i2}) = c_1(y_{i1}) + c_2(y_{i2})$

(−0.5) Error in conditional distn

(c) $\mu = E(y_{i\cdot})$
$\eta_{12} = E(y_{i1} y_{i2})$
$\sigma_{12} = Cov(y_{i1}, y_{i2})$
$= E(y_{i1} y_{i2}) - E(y_{i1}) E(y_{i2})$

Find explicit expressions.

(i) $\mu = E(y_{i\cdot}) = \begin{bmatrix} E(y_{i1}) \\ E(y_{i2}) \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$

$$\mu_1 = \sum_{y_{i1}} \sum_{y_{i2}} y_{i1} P(y_{i1}, y_{i2}) \qquad (y_{i1}, y_{i2} \in \{0,1\})$$

$$= [(1)[P(1,1) + P(1,0)] + \underbrace{0[P(0,1) + P(0,0)]}_{0}$$

$$= [\Delta(\theta,\lambda)]^{-1} [\exp(\theta_1 + \theta_2 + \lambda - c(1,1)) + \exp(\theta_1 - c(1,0))]$$

$$= \frac{\left[\dfrac{\exp(\theta_1+\theta_2+\lambda)}{\exp(c(1,1))} + \dfrac{\exp(\theta_1)}{\exp(c(1,0))}\right]}{\underbrace{\left[\dfrac{\exp(\theta_1+\theta_2+\lambda)}{\exp(c(1,1))} + \dfrac{\exp(\theta_1)}{\exp(c(1,0))} + \dfrac{\exp(\theta_2)}{\exp(c(0,1))} + \dfrac{1}{\exp(c(0,0))}\right]}_{\Delta(\theta,\lambda)}}$$

$$\mu_2 = (1)[P(1,1) + P(0,1)] + \underbrace{0[P(1,0) + P(0,0)]}_{0}$$

$$= [\Delta(\theta,\lambda)]^{-1} [\exp(\theta_1 + \theta_2 + \lambda - c(1,1)) + \exp(\theta_2 - c(0,1))]$$

$$= \frac{\left[\dfrac{\exp(\theta_1+\theta_2+\lambda)}{e^{c(1,1)}} + \dfrac{\exp(\theta_2)}{e^{c(0,1)}}\right]}{\underbrace{\left[\dfrac{\exp(\theta_1+\theta_2+\lambda)}{e^{c(1,1)}} + \dfrac{\exp(\theta_1)}{e^{c(1,0)}} + \dfrac{\exp(\theta_2)}{e^{c(0,1)}} + \dfrac{1}{e^{c(0,0)}}\right]}_{\Delta(\theta,\lambda)}}$$

$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$ specified above.

$$\eta_{12} = E[y_{i1}\, y_{i2}]$$

$$= \sum_{y_{i1}} \sum_{y_{i2}} y_{i1} y_{i2}\, P(y_{i1}, y_{i2})$$

$$= (1\cdot 1)P(1,1) + (1\cdot 0)P(1,0) + (0\cdot 1)P(0,1) + (0\cdot 0)P(0,0)$$

$$= P(1,1)$$

$$= \frac{\left[\dfrac{\exp(\theta_1+\theta_2+\lambda)}{e^{c(1,1)}}\right]}{\left[\dfrac{\exp(\theta_1+\theta_2+\lambda)}{e^{c(1,1)}} + \dfrac{\exp(\theta_1)}{e^{c(1,0)}} + \dfrac{\exp(\theta_2)}{e^{c(0,1)}} + \dfrac{1}{e^{c(0,0)}}\right]}$$

$$\sigma_{12} = E[y_{i1}\, y_{i2}] - E[y_{i1}]E[y_{i2}]$$

$$= \eta_{12} - \mu_1 \mu_2$$

$$= \frac{\left[\dfrac{\exp(\theta_1+\theta_2+\lambda)}{e^{c(1,1)}}\right]}{\Delta(\theta,\lambda)} - \frac{\left[\dfrac{e^{\theta_1+\theta_2+\lambda}}{e^{c(1,1)}} + \dfrac{e^{\theta_1}}{e^{c(1,0)}}\right]\left[\dfrac{e^{\theta_1+\theta_2+\lambda}}{e^{c(1,1)}} + \dfrac{e^{\theta_2}}{e^{c(0,1)}}\right]}{[\Delta(\theta,\lambda)]^2}$$

$\boxed{7/7}$

ⓓ Suppose we consider a reparameterization from (θ, λ) to (μ, η_{12})

Jacobian $V = \dfrac{\partial(\theta, \lambda)}{\partial(\mu, \eta_{12})}$

In general for scalar variables,

$$\frac{\partial y}{\partial x} = \frac{1}{\partial x / \partial y}$$

great!

Therefore,

$$V^{-1} = \frac{\partial(\mu, \eta_{12})}{\partial(\theta, \lambda)} \quad \text{and} \quad V = (V^{-1})^{-1}$$

assuming inverse transformation exists.

Use V^{-1} to characterize the cov. matrix of

$$\begin{bmatrix} y_{i1} \\ y_{i2} \\ y_{i1}y_{i2} \end{bmatrix}$$

Specify a sufficient & necessary condition for this transformation from (θ, λ) to (μ, η_{12}) is one-to-one

cov matrix of $[y_{i1}, y_{i2}, y_{i1}y_{i2}]^T$

$$= \begin{bmatrix} Var(y_{i1}) & Cov(y_{i1}, y_{i2}) & Cov(y_{i1}, y_{i1}y_{i2}) \\ Cov(y_{i1}, y_{i2}) & Var(y_{i2}) & Cov(y_{i2}, y_{i1}y_{i2}) \\ Cov(y_{i1}, y_{i1}y_{i2}) & \uparrow & Var(y_{i1}y_{i2}) \end{bmatrix}$$

$Cov(y_{i2}, y_{i1}y_{i2})$

Continued. $\longrightarrow$

$$V^{-1} = \begin{bmatrix} \partial\mu_1/\partial\theta_1 & \partial\mu_1/\partial\theta_2 & \partial\mu_1/\partial\lambda \\ \partial\mu_2/\partial\theta_1 & \partial\mu_2/\partial\theta_2 & \partial\mu_2/\partial\lambda \\ \partial\pi_{12}/\partial\theta_1 & \partial\pi_{12}/\partial\theta_2 & \partial\pi_{12}/\partial\lambda \end{bmatrix}$$

Let μ_1 be of form $\frac{a}{a+b}$. $\quad a+b = \Delta(\theta,\lambda)$

$$\frac{\partial\mu_1}{\partial\theta_1} = \frac{(a+b)\,\partial/\partial\theta_1\, a - a\,\partial/\partial\theta_1 (a+b)}{(a+b)^2}$$

$$\frac{\partial}{\partial\theta_1} a = \frac{\partial}{\partial\theta_1}\left[\frac{\exp(\theta_1+\theta_2+\lambda)}{e^{C(1,1)}} + \frac{\exp(\theta_1)}{e^{C(1,0)}}\right]$$

$$= \left[\frac{\exp(\theta_1+\theta_2+\lambda)}{e^{C(1,1)}} + \frac{\exp(\theta_1)}{e^{C(1,0)}}\right]$$

$$= a$$

$$\frac{\partial}{\partial\theta_1}(a+b) = \frac{\partial}{\partial\theta_1}\Delta(\theta,\lambda)$$

$$= \left[\frac{\exp(\theta_1+\theta_2+\lambda)}{e^{C(1,1)}} + \frac{\exp(\theta_1)}{e^{C(1,0)}} + 0 + 0\right]$$

$$= a$$

$$\Rightarrow \frac{\partial\mu_1}{\partial\theta_1} = \frac{\Delta(\theta,\lambda)\cdot a - a(a)}{(\Delta(\theta,\lambda))^2}$$

$$= \frac{a(\Delta(\theta,\lambda) - a)}{(\Delta(\theta,\lambda))^2}$$

$$= \frac{\left[\frac{\exp(\theta_1+\theta_2+\lambda)}{e^{C(1,1)}} + \frac{\exp(\theta_1)}{e^{C(1,0)}}\right]\left[\frac{\exp(\theta_2)}{e^{C(0,1)}} + \frac{1}{e^{C(0,0)}}\right]}{(\Delta(\theta,\lambda))^2} \checkmark$$

Likewise, if μ_2 is re-written as $\dfrac{l}{l+k'}$, then

$$\frac{\partial \mu_2}{\partial \theta_2} = \frac{(l+k)\,\partial/\partial\theta_2\, l - l\,\partial/\partial\theta_2\,(l+k)}{(l+k)^2}$$

$$= \frac{\Delta(\theta,\lambda)(l) - l(l)}{(\Delta(\theta,\lambda))^2} \quad \text{since } l+k = \Delta(\theta,\lambda)$$

$$= \frac{a[\Delta(\theta,\lambda) - l]}{(\Delta(\theta,\lambda))^2}$$

$$= \frac{\left[\overbrace{\frac{\exp(\theta_1+\theta_2+\lambda)}{e^{C(1,1)}} + \frac{\exp(\theta_2)}{e^{C(0,1)}}}^{l}\right]\left[\frac{\exp(\theta_1)}{e^{C(1,0)}} + \frac{1}{e^{C(0,0)}}\right]}{(\Delta(\theta,\lambda))^2}$$

$$\frac{\partial \mu_1}{\partial \theta_2} = \frac{\Delta(\theta,\lambda)\cdot\partial/\partial\theta_2(a) - a\,\partial/\partial\theta_2(\Delta(\theta,\lambda))}{(\Delta(\theta,\lambda))^2}$$

$$\frac{\partial}{\partial\theta_2}(a) = \frac{\exp(\theta_1+\theta_2+\lambda)}{e^{C(1,1)}} + 0$$

$$\frac{\partial}{\partial\theta_2}\Delta(\theta,\lambda) = \frac{\exp(\theta_1+\theta_2+\lambda)}{e^{C(1,1)}} + \frac{\exp(\theta_2)}{e^{C(0,1)}}$$

$$\frac{\partial \mu_1}{\partial \theta_2} = \frac{\left(\Delta(\theta,\lambda)\left[\frac{\exp(\theta_1+\theta_2+\lambda)}{e^{C(1,1)}}\right] - \left[\frac{\exp(\theta_1+\theta_2+\lambda)}{e^{C(1,1)}} + \frac{\exp(\theta_1)}{e^{C(1,0)}}\right] * \left[\frac{\exp(\theta_1+\theta_2+\lambda)}{e^{C(1,1)}} + \frac{\exp(\theta_2)}{e^{C(0,1)}}\right]\right)}{[\Delta(\theta,\lambda)]^2}$$

✓

$$\frac{\partial \mu_1}{\partial \lambda} = \frac{\Delta(\theta,\lambda)\left(\partial a/\partial \lambda\right) - a\, \partial \Delta(\theta,\lambda)/\partial \lambda}{\left(\Delta(\theta,\lambda)\right)^2}$$

$$\frac{\partial a}{\partial \lambda} = \frac{\partial \Delta(\theta,\lambda)}{\partial \lambda} = \frac{\exp(\theta_1+\theta_2+\lambda)}{e^{C(1,1)}}$$

$$\frac{\partial \mu_1}{\partial \lambda} = \frac{\left[\frac{\exp(\theta_1+\theta_2+\lambda)}{e^{C(1,1)}}\right]\left[\Delta(\theta,\lambda) - a\right]}{\left(\Delta(\theta,\lambda)\right)^2}$$

$$= \frac{\left[\frac{\exp(\theta_1+\theta_2+\lambda)}{e^{C(1,1)}}\right]\left[\frac{\exp(\theta_2)}{e^{C(0,1)}} + \frac{1}{e^{C(0,0)}}\right]}{\left(\Delta(\theta,\lambda)\right)^2}$$

Likewise,

$$\frac{\partial \mu_2}{\partial \lambda} = \frac{\left[\frac{\exp(\theta_1+\theta_2+\lambda)}{e^{C(1,1)}}\right]\left[\Delta(\theta,\lambda) - l\right]}{\left(\Delta(\theta,\lambda)\right)^2}$$

$$= \frac{\left[\frac{\exp(\theta_1+\theta_2+\lambda)}{e^{C(1,1)}}\right]\left[\frac{\exp\theta_1}{e^{C(1,0)}} + \frac{1}{e^{C(0,0)}}\right]}{\left[\Delta(\theta,\lambda)\right]^2}$$

(Summary: Have now found $\partial\mu_1/\partial\theta_1$, $\partial\mu_1/\partial\theta_2$, $\partial\mu_1/\partial\lambda$, $\partial\mu_2/\partial\theta_2$, $\partial\mu_2/\partial\lambda$)

$$\frac{\partial \mu_2}{\partial \theta_1} = \frac{\Delta(\theta,\lambda)\left(\partial l/\partial \theta_1\right) - l\, \partial \Delta(\theta,\lambda)/\partial \theta_1}{\left(\Delta(\theta,\lambda)\right)^2}$$

$$\frac{\partial l}{\partial \theta_1} = \frac{\exp(\theta_1+\theta_2+\lambda)}{e^{C(1,1)}}$$

$$\frac{\partial \Delta(\theta,\lambda)}{\partial \theta_1} = \frac{\exp(\theta_1+\theta_2+\lambda)}{e^{C(1,1)}} + \frac{\exp(\theta_1)}{e^{C(1,0)}}$$

Continued →

$$\frac{\partial \mu_2}{\partial \theta_1}$$

$$= \frac{\left(\Delta(\theta,\lambda)\left[\frac{\exp(\theta_1+\theta_2+\lambda)}{e^{C(1,1)}}\right] - \left[\frac{\exp(\theta_1+\theta_2+\lambda)}{e^{C(1,1)}} + \frac{\exp(\theta_2)}{e^{C(0,1)}}\right] * \left[\frac{\exp(\theta_1+\theta_2+\lambda)}{e^{C(1,1)}} + \frac{\exp(\theta_1)}{e^{C(1,0)}}\right] \right)}{[\Delta(\theta,\lambda)]^2}$$

let $\eta_{12} = \frac{m}{\Delta(\theta,\lambda)}$ $\qquad m = \frac{\exp(\theta_1+\theta_2+\lambda)}{e^{C(1,1)}}$

$$\frac{\partial \eta_{12}}{\partial \theta_1} = \frac{\Delta(\theta,\lambda)(\partial m/\partial\theta_1) - m(\partial \Delta(\theta,\lambda)/\partial\theta_1)}{[\Delta(\theta,\lambda)]^2}$$

$$\frac{\partial m}{\partial \theta_1} = \frac{\partial m}{\partial \theta_2} = \frac{\partial m}{\partial \lambda} = \frac{\exp(\theta_1+\theta_2+\lambda)}{e^{C(1,1)}} = m$$

Have found $\frac{\partial \Delta(\theta,\lambda)}{\partial\theta_1}$, $\frac{\partial \Delta(\theta,\lambda)}{\partial\theta_2}$, + $\frac{\partial \Delta(\theta,\lambda)}{\partial\lambda}$

in previous pgs.

$$\frac{\partial \eta_{12}}{\partial \gamma} = \frac{\Delta(\theta,\lambda)\overbrace{(\partial m/\partial\gamma)}^{=m} - m(\partial\Delta(\theta,\lambda)/\partial\gamma)}{(\Delta(\theta,\lambda))^2}$$

generic parameter

$$= \frac{m[\Delta(\theta,\lambda) - \partial\Delta(\theta,\lambda)/\partial\gamma]}{(\Delta(\theta,\lambda))^2}$$

$$\Delta(\theta,\lambda) - \frac{\partial\Delta(\theta,\lambda)}{\partial\theta_1} = \frac{\exp(\theta_2)}{e^{C(0,1)}} + \frac{1}{e^{C(0,0)}}$$

$$\Delta(\theta,\lambda) - \frac{\partial\Delta(\theta,\lambda)}{\partial\lambda} = \left[\frac{\exp(\theta_1)}{e^{C(1,0)}} + \frac{\exp(\theta_2)}{e^{C(0,1)}} + \frac{1}{e^{C(0,0)}}\right]$$

$$\Delta(\theta,\lambda) - \partial\Delta(\theta,\lambda)/\partial\theta_2 = \frac{\exp(\theta_1)}{e^{C(1,0)}} + \frac{1}{e^{C(0,0)}}$$

$$\frac{\partial n_{12}}{\partial \theta_1} = \frac{\left[\dfrac{\exp(\theta_1+\theta_2+\lambda)}{e^{C(1,1)}}\right]\left[\dfrac{e^{\theta_2}}{e^{C(0,1)}}+\dfrac{1}{e^{C(0,0)}}\right]}{(\Delta(\theta,\lambda))^2}$$

$$\frac{\partial n_{12}}{\partial \theta_2} = \frac{\left[\dfrac{\exp(\theta_1+\theta_2+\lambda)}{e^{C(1,1)}}\right]\left[\dfrac{e^{\theta_1}}{e^{C(1,0)}}+\dfrac{1}{e^{C(0,1)}}\right]}{[\Delta(\theta,\lambda)]^2}$$

$$\frac{\partial n_{12}}{\partial \lambda} = \frac{\left[\dfrac{\exp(\theta_1+\theta_2+\lambda)}{e^{C(1,1)}}\right]\left[\dfrac{e^{\theta_1}}{e^{C(1,0)}}+\dfrac{e^{\theta_2}}{e^{C(0,1)}}+\dfrac{1}{e^{C(0,0)}}\right]}{(\Delta(\theta,\lambda))^2}$$

How do we use V^{-1} to characterize the covariance matrix of $\begin{bmatrix} y_{i1} \\ y_{i2} \\ y_{i1}y_{i2} \end{bmatrix}$?

we know from results of MLE's,

$$\sqrt{n}\left(\begin{bmatrix} \hat{\theta}_1 - \theta_1 \\ \hat{\theta}_2 - \theta_2 \\ \hat{\theta}_3 - \theta_3 \end{bmatrix}\right) \xrightarrow{d} N(0, I(\theta_1, \theta_2, \lambda)^{-1})$$

$$\updownarrow$$

$$\lim_{n \to \infty} \left[\frac{1}{n} I_n(\theta_1, \theta_2, \theta_3)\right]^{-1}$$

Fisher info matrix

Let $\gamma = [\theta_1 \ \theta_2 \ \lambda]^T$

let $\varepsilon = [\mu_1 \ \mu_2 \ \eta_{12}]^T$

Assuming one-to-one transformations apply,

$g(\gamma) = \varepsilon$

Therefore, by delta method,

In order for this proof to work, you would need to demonstrate that $\frac{1}{n}\sum[y_i]$ is the MLE of the data or that $g(\hat{\gamma})$ = their means

$$\sqrt{n}\left(\begin{bmatrix} \frac{1}{n}\sum_{i=1}^n y_{i1} \\ \frac{1}{n}\sum_{i=1}^n y_{i2} \\ \frac{1}{n}\sum_{i=1}^n y_{i1}y_{i2} \end{bmatrix} - \begin{bmatrix} \mu_{i1} \\ \mu_{i2} \\ \eta_{12} \end{bmatrix}\right) =$$

$$= \sqrt{n}(g(\hat{\gamma}) - g(\gamma)) \xrightarrow{d} N(0, \nabla g(\gamma) I(\theta_1, \theta_2, \lambda)^{-1} \nabla g(\gamma)^T)$$

where $\nabla g(\gamma) = \dfrac{\partial(\mu_1, \mu_2, \eta_{12})}{\partial(\theta_1, \theta_2, \lambda)} = V^{-1}$ as defined earlier.

Therefore,

$$\text{Cov}(\mu_1, \mu_2, \eta_{12}) = (V^{-1}) I_n(\theta_1, \theta_2, \lambda)^{-1} (V^{-1})^T$$

X

The transformation from (θ, λ) to (μ, η_{12}) is one to one if the inverse transformation exists. In other words, if V & V^{-1} are non-singular (determinant $\neq 0$).

When is determinant of $|V^{-1}| \neq 0$?

I don't have the time or inclination to find this determinant, but if I did, I would take the determinant of V^{-1} and find when/if it ever $= 0$.

$$|V^{-1}| = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}$$

$$= a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix}$$

$$= a(ei - fh) - b(di - gf) + c(dh - eg).$$

set $= 0$.

6/7

(e) $p \times 1$ column vector x_i $\forall i$

$y_i | x_i \sim QE(\theta_i, \lambda_i)$

$\theta_i = (\theta_{i1}, \theta_{i2})'$, λ_i may depend on x_i

$i = 1, \dots, n$

$$E[y_i | x_i] = \mu_i = \begin{bmatrix} \mu_{i1} \\ \mu_{i2} \end{bmatrix} = \mu(x_i, \beta)$$

$$\text{Cov}(y_{i1}, y_{i2} | x_i) = \sigma_{i12} = \sigma_{12}(x_i, \beta, \alpha)$$

β = unknown $p \times 1$ regression parameter,

α = unknown scalar parameter.

(i) Derive likelihood score eqns. for (α, β).

$$P(y_{i1}, y_{i2} | x_i) = (\Delta(\theta_i, \lambda_i))^{-1} \exp(y_{i1}\theta_{i1} + y_{i2}\theta_{i2} + y_{i1}y_{i2}\lambda_i(x_i) - c(y_{i1}, y_{i2}))$$

$\Delta(\theta_i, \lambda_i)$ is same as before, but

$\theta_1 \rightarrow \theta_{i1}$

$\theta_2 \rightarrow \theta_{i2}$

$\lambda \rightarrow \lambda_i(x_i) \rightarrow$ possible function of x_i

$$l_n(\theta_i, \lambda_i) = \sum_{i=1}^{n} \log P(y_{i1}, y_{i2} | x_i)$$

$$= \sum_{i=1}^{n} \left[-\log \Delta(\theta_i, \lambda_i(x_i)) + y_{i1}\theta_{i1} + y_{i2}\theta_{i2} + y_{i1}y_{i2}\lambda_i(x_i) + - c(y_{i1}, y_{i2}) \right]$$

Continued. $\longrightarrow$

As noted in part 'd',

$$\begin{bmatrix}\theta_1\\ \theta_2\\ \theta_3\end{bmatrix} = \gamma, \quad \begin{bmatrix}\mu_1\\ \mu_2\\ \eta_{12}\end{bmatrix} = \varepsilon = g(\gamma)$$

let $\mu_1 = g_1(\gamma)$,
$\mu_2 = g_2(\gamma)$,
$\eta_{12} = g_3(\gamma)$.

$$\frac{\partial\, l_n(\alpha, B)}{\partial B} = \Bigg[\overbrace{\underbrace{\frac{\partial\, l_n(\theta_1, \theta_2, \lambda)}{\partial \gamma}}_{1\times 3} \cdot \underbrace{\frac{\partial \gamma}{\partial \varepsilon}}_{3\times 3}}^{\gamma} \cdot \underbrace{\frac{\partial \varepsilon}{\partial B}}_{3\times p}\Bigg]^T \rightarrow p\times 1$$

where $\varepsilon = [\mu_1, \mu_2, \eta_{12}]$

I can find $\frac{\partial\, l_n(\theta_1, \theta_2, \lambda)}{\partial \theta_1}$, $\frac{\partial\, l_n(\theta_1, \theta_2, \lambda)}{\partial \theta_2}$, + $\frac{\partial\, l_n(\theta_1, \theta_2, \lambda)}{\partial \lambda}$.

To find $\frac{\partial \gamma}{\partial \varepsilon}$, assuming one-to-one case,

I can use $V = [V^{-1}]^{-1}$, with V^{-1} specified in part d.

Then, I can also find

$$\frac{\partial \mu_1}{B}, \frac{\partial \mu_2}{B}, + \frac{\partial \eta_{12}}{B} := \frac{\partial \varepsilon}{\partial B}.$$

$$\frac{\partial\, l_n(\theta, \lambda)}{\partial \theta_{i1}} = \Sigma_{i=1}^{n}\left[-\frac{1}{\Delta(\theta_i, \lambda_i(x_i))} \cdot \left(\frac{e^{\theta_1+\theta_2+\lambda_i(x_i)}}{e^{c(1,1)}} + \frac{e^{\theta_{i1}}}{e^{c(1,0)}}\right) + y_{i1}\right]$$

$$\frac{\partial\, l_n(\theta, \lambda)}{\partial \theta_{i2}} = \Sigma_{i=1}^{n}\left[-\frac{1}{\Delta(\theta_i, \lambda_i(x_i))} \cdot \left(\frac{e^{\theta_1+\theta_2+\lambda_i(x_i)}}{e^{c(1,1)}} + \frac{e^{\theta_{i2}}}{e^{c(0,1)}}\right) + y_{i2}\right]$$

Continued. $\longrightarrow$

$$\frac{\partial \ln(\theta,\lambda)}{\partial \lambda_i(x_i)} = \Sigma_{i=1}^{n}\left[\frac{-1}{\Delta(\theta_i, \lambda_i(x_i))} \cdot \left(\frac{e^{\theta_{i1}+\theta_{i2}+\lambda_i(x_i)}}{e^{C(1,1)}} \right) \cdot \lambda_i'(x_i) + \right.$$
$$\left. + \; y_{i1} y_{i2} \cdot \lambda_i'(x_i) \right]$$

$$\frac{\partial \ln(\gamma)}{\partial \gamma} = \left[\frac{\partial \ln(\gamma)}{\partial \theta_{i1}} \quad \frac{\partial \ln(\gamma)}{\partial \theta_{i2}} \quad \frac{\partial \ln(\gamma)}{\partial \lambda_i(x_i)} \right]_{1\times 3}$$

$\frac{\partial \gamma}{\partial \varepsilon} = V = [V^{-1}]^{-1}$ from part (d).

$$\frac{\partial \varepsilon}{\partial B} = \begin{bmatrix} (\partial \mu_{i1}/\partial B)^T \\ (\partial \mu_{i2}/\partial B)^T \\ (\partial \eta_{i12}/\partial B)^T \end{bmatrix} = 3\times p \text{ dimensions.}$$

~~let $\mu_i(x_i, B) = \mu_i(h(x_i, B))$~~

$$\frac{\partial \mu_{i1}}{\partial B} = \mu_{i1}'(x_i, B)$$

$$\frac{\partial \mu_{i2}}{\partial B} = \mu_{i2}'(x_i, B)$$

> depending on function of x_i, B inside, will need chain rule, which here remains unspecified.

Note: $\sigma_{i12} = \eta_{i12} - \mu_{i1}\mu_{i2}$

$\Rightarrow \eta_{i12} = \sigma_{i12} + \mu_{i1}\mu_{i2}$

$$\frac{\partial \eta_{i12}}{\partial B} = \sigma_{i12}'(x_i, B, \alpha) + (\mu_{i1}'(x_i, B))(\mu_{i2}'(x_i, B))$$

↑ derivative in relation to B.

Score eqns:

$$\begin{bmatrix} \frac{\partial \ln(\alpha, \beta)}{\partial \beta} \\ \frac{\partial \ln(\alpha, \beta)}{\partial \alpha} \end{bmatrix}$$

Have alreed found $\frac{\partial \ln(\alpha, \beta)}{\partial \beta}$.

$$\frac{\partial \ln(\alpha, \beta)}{\partial \alpha} = \left[\frac{\partial \ln(\gamma)}{\partial \gamma} \cdot \frac{\partial \gamma}{\partial \varepsilon} \cdot \frac{\partial \varepsilon}{\partial \alpha} \right]^T$$

$\frac{\partial \ln(\gamma)}{\partial \gamma}$ & $\frac{\partial \gamma}{\partial \varepsilon}$ remain the same as those used in the making of $\frac{\partial \ln(\alpha, \beta)}{\partial \beta}$.

$$\frac{\partial \varepsilon}{\partial \alpha} = \begin{bmatrix} \underset{\sim}{0}_{1 \times p} \\ \underset{\sim}{0}_{1 \times p} \\ (\partial \eta_{12} / \partial \alpha)^T \end{bmatrix} = 3 \times p$$

$$\frac{\partial \eta_{12}}{\partial \alpha} = \sigma_{12}'(x_i, \beta, \alpha) + 0$$

$$= \sigma_{i12}'(x_i, \beta, \alpha)$$

↑ derivative in relation to α

Fisher info matrix for α, β.

Can find $I_n(\theta_1, \theta_2, \lambda) = I_n(\gamma)$

Then $I_n(\mu_1, \mu_2, \eta_{12}) = I_n(\varepsilon)$

$$= \underset{2\times 3}{\left[\frac{\partial \gamma}{\partial \alpha, \beta}\right]^T} \; \underset{3\times 3}{I_n(\gamma)} \; \underset{3\times 2}{\left[\frac{\partial \gamma}{\partial \alpha, \beta}\right]}$$

(would have to be done numerically, plus censoring) guide apply here

$$\frac{\partial \gamma}{\partial \alpha, \beta} = \frac{\partial \gamma}{\partial \varepsilon} \cdot \frac{\partial \varepsilon}{\partial \alpha, \beta} \rightarrow (3\times 3)(3\times 2) = 3\times 2$$

$$\frac{\partial \gamma}{\partial \varepsilon} = V = (V^{-1})^{-1} \text{ as described in (d).}$$

$$\frac{\partial \varepsilon}{\partial \alpha, \beta} = \begin{bmatrix} \partial \mu_{i1}/\partial \alpha & \partial \mu_{i1}/\partial \beta \\ \partial \mu_{i2}/\partial \alpha & \partial \mu_{i2}/\partial \beta \\ \partial \eta_{i12}/\partial \alpha & \partial \eta_{i12}/\partial \beta \end{bmatrix}$$

$$= \begin{bmatrix} 0 & \partial \mu_{i1}/\partial \beta \\ 0 & \partial \mu_{i2}/\partial \beta \\ \partial \eta_{i12}/\partial \alpha & \partial \eta_{i12}/\partial \beta \end{bmatrix}$$

↑ parts described earlier during work to find score function.

Need to find $I_n(\gamma)$:

$$\frac{\partial \ln(\gamma)}{\partial \theta_1} = \Sigma_{i=1}^{n} \left[\overbrace{\frac{-1}{\Delta(\theta_{i}, \lambda_i(x_i))}}^{a} \overbrace{\left(\frac{e^{\theta_{i1} + \theta_{i2} + \lambda_i(x_i)}}{e^{C(1,1)}} + \frac{e^{\theta_{i1}}}{e^{C(1,0)}} \right)}^{b} + y_{i1} \right]$$

$$\frac{\partial^2 \ln(\gamma)}{\partial \theta_1^2} = \Sigma_{i=1}^{n} \left[a(\partial b/\partial \theta_1) + b(\partial a/\partial \theta_1) \right]$$

Continued →

$$= \sum_{i=1}^{n} \left[\frac{(-1)}{\Delta(\theta_i, \lambda_i(x_i))} \cdot b + b \left(\frac{(-1)(-1)}{\Delta(\theta_i, \lambda_i(x_i))^2} (b) \right) \right]$$

$$\rightarrow \frac{\partial b}{\partial \theta_1} = b.$$

$$= \sum_{i=1}^{n} \left[\left(\frac{e^{\theta_{i1} + \theta_{i2} + \lambda_i(x_i)}}{e^{C(1,1)}} + \frac{e^{\theta_{i1}}}{e^{C(1,0)}} \right) \left[\frac{-1}{\Delta(\theta_i, \lambda_i(x_i))} + \frac{1}{(\Delta(\theta_i, \lambda_i(x_i)))^2} \left(\frac{e^{\theta_{i1} + \theta_{i2} + \lambda_i(x_i)}}{e^{C(1,1)}} + \frac{e^{\theta_{i1}}}{e^{C(1,0)}} \right) \right] \right]$$

Likewise,

$$\frac{\partial^2}{\partial \theta_2^2} \ln(\gamma) = \sum_{i=1}^{n} \left[\left(\frac{e^{\theta_{i1} + \theta_{i2} + \lambda_i(x_i)}}{e^{C(1,1)}} + \frac{e^{\theta_{i2}}}{e^{C(0,1)}} \right) \left[\frac{-1}{\Delta(\theta_i, \lambda_i(x_i))} + \frac{1}{(\Delta(\theta_i, \lambda_i(x_i)))^2} \left(\frac{e^{\theta_{i1} + \theta_{i2} + \lambda_i(x_i)}}{e^{C(1,1)}} + \frac{e^{\theta_{i2}}}{e^{C(0,1)}} \right) \right] \right]$$

$$\frac{\partial^2 \ln(\gamma)}{\partial \theta_1 \partial \theta_2} = \sum_{i=1}^{n} \left(a(\partial b / \partial \theta_2) + b(\partial a / \partial \theta_2) \right)$$

$$= \sum_{i=1}^{2} \left[\frac{-1}{\Delta(\theta_i, \lambda_i(x_i))} \left(\frac{e^{\theta_{i1} + \theta_{i2} + \lambda_i(x_i)}}{e^{C(1,1)}} \right) + \frac{1}{(\Delta(\theta_i, \lambda_i(x_i)))^2} \underbrace{\left(\frac{e^{\theta_{i1} + \theta_{i2} + \lambda_i(x_i)}}{e^{C(1,1)}} + \frac{e^{\theta_{i2}}}{e^{C(0,1)}} \right)}_{\partial a / \partial \theta_2} \underbrace{\left(\frac{e^{\theta_{i1} + \theta_{i2} + \lambda_i(x_i)}}{e^{C(1,1)}} + \frac{e^{\theta_{i1}}}{e^{C(1,0)}} \right)}_{b} \right]$$

$$\frac{\partial \ln(\gamma)}{\partial \theta_1 \partial \lambda_i(x_i)} = \sum_{i=1}^{n} \left(a \, \partial b / \partial \lambda_i(x_i) + b(\partial a / \partial \lambda_i(x_i)) \right)$$

$$= \sum_{i=1}^{n} \left[\frac{-1}{\Delta(\theta_i, \lambda_i(x_i))} \left(\frac{e^{\theta_{i1} + \theta_{i2} + \lambda_i(x_i)}}{e^{C(1,1)}} \cdot \lambda_i'(x_i) \right) + \frac{1}{(\Delta(\theta_i, \lambda_i(x_i)))^2} \left(\frac{e^{\theta_{i1} + \theta_{i2} + \lambda_i(x_i)}}{e^{C(1,1)}} \lambda_i'(x_i) \right) \left(\frac{e^{\theta_{i1} + \theta_{i2} + \lambda_i(x_i)}}{e^{C(1,1)}} + \frac{e^{\theta_{i1}}}{e^{C(1,0)}} \right) \right]$$

Continued →

Likewise,

$$\frac{\partial \ln(\gamma)}{\partial\theta_2 \partial\lambda_i(x_i)} = \left[\frac{-1}{\Delta(\theta_i, \lambda_i(x_i))} \left(\frac{e^{\theta_{i1}+\theta_{i2}+\lambda_i(x_i)}}{e^{C(1,1)}} \cdot \lambda_i'(x_0) \right) + \frac{1}{\Delta(\theta_i, \lambda_i(x_0))^2} \left(\frac{e^{\theta_{i1}+\theta_{i2}+\lambda_0(x_0)}}{e^{C(1,1)}} \cdot \lambda_i'(x_0) \right) \left(\frac{e^{\theta_{i1}+\theta_{i2}+\lambda_i(x_0)}}{e^{C(1,1)}} + \frac{e^{\theta_{i2}}}{e^{C(0,1)}} \right) \right]$$

$$\frac{\partial \ln(\lambda_i(x_i))}{\partial \lambda_i(x_i)} =$$

$$= \sum_{i=1}^{n} \left[\underbrace{\frac{-1}{\Delta(\theta_i, \lambda_i(x_i)}}_{a} \cdot \underbrace{\frac{e^{\theta_{i1}+\theta_{i2}+\lambda_i(x_i)}}{e^{C(1,1)}}}_{b} \cdot \underbrace{\lambda_i'(x_i)}_{c} + \underbrace{y_{i1} y_{i2} \lambda_i'(x_i)}_{d} \right]$$

$$\frac{\partial^2 \ln(\lambda_0(x_i))}{\partial \lambda_i(x_i)^2} = \sum_{i=1}^{n} \left[a \left[\partial b \cdot c / \partial\lambda_i(x_0) \right] + (b \cdot c) \left[\partial a / \partial\lambda_i(x_0) \right] + \partial d / \partial\lambda_i(x_i) \right]$$

$$\sum_{i=1}^{n} \left[a \left[b \frac{\partial c}{\partial\lambda_i(x_0)} + c \frac{\partial b}{\partial\lambda_i(x_0)} \right] + bc \left[\frac{\partial a}{\partial\lambda_i(x_i)} \right] + y_{i1} y_{i2} \lambda_i''(x_0) \right]$$

$$= \sum_{i=1}^{n} \left[\frac{-1}{\Delta(\theta_i, \lambda_i(x_i))} \left[\frac{e^{\theta_{i1}+\theta_{i2}+\lambda_i(x_0)}}{e^{C(1,1)}} \left(\lambda_i''(x_i) \right) + (\lambda_i'(x_0))^2 \left(\frac{e^{\theta_{i1}+\theta_{02}+\lambda_i(x_i)}}{C(1,1)} \right) \right] + \left(\frac{e^{\theta_{i1}+\theta_{i2}+\lambda_i(x_i)}}{e^{C(1,1)}} \right)^2 \left(\lambda_i'(x_0) \right)^2 \left(\frac{1}{(\Delta(\theta_i, \lambda_i(x_0)))^2} \right) + y_{i1} y_{i2} \lambda_i''(x_0) \right]$$

Since all but $\frac{\partial^2 \ln(\gamma)}{\partial\lambda_i(x_i)^2}$ do not contain y_{i1} or y_{i2},

$E\left[-\partial^2/\partial\gamma^2 \ln(\gamma)\right]$ = just the negative of the specified results for all other derivatives.

$$E[y_{i1} y_{i2}] = [E(y_{i1} y_{i2}) - E(y_{i1}) E(y_{i2})] + E(y_{i1}) E(y_{i2})$$
$$= \sigma_{i12} + \mu_{i1} \mu_{i2}$$

$\Rightarrow E[-\partial^2/\partial\lambda_i(x_i)^2 \, l_n(\gamma)]$
= negative of result with $(\sigma_{i12} + \mu_{i1}\mu_{i2})$
replacing the term $(y_{i1} y_{i2})$.

$$I_n(\gamma) = -E \begin{bmatrix} \dfrac{\partial^2 l_n(\gamma)}{\partial\theta_1^2} & \dfrac{\partial^2 l_n(\gamma)}{\partial\theta_1 \partial\theta_2} & \dfrac{\partial^2 l_n(\gamma)}{\partial\theta_1 \partial\lambda_i(x_i)} \\ \dfrac{\partial^2 l_n(\gamma)}{\partial\theta_1 \partial\theta_2} & \dfrac{\partial^2 l_n(\gamma)}{\partial\theta_2^2} & \dfrac{\partial^2 l_n(\gamma)}{\partial\theta_2 \partial\lambda_i(x_i)} \\ \dfrac{\partial^2 l_n(\gamma)}{\partial\theta_1 \partial\lambda_i(x_i)} & \dfrac{\partial^2 l_n(\gamma)}{\partial\theta_2 \partial\lambda_i(x_i)} & \dfrac{\partial^2 l_n(\gamma)}{\partial(\lambda_i(x_i))^2} \end{bmatrix}$$

Plug in results specified in earlier pgs.

Do score eqns explicity involve $C(y_{i1}, y_{i2})$?

The $\Delta(\theta_0, \lambda_i)$ term and some of the chain rule parts of the derivatives contain the C function evaluated at specific y_{i1}, y_{i2} combos of 1 & 0. However, these instances evaluate the C function as a constant, not a statistic. Nowhere in the score function is the $C(y_{i1}, y_{i2})$ evaluated with the given observed y_{i1}, y_{i2} data.

$\boxed{6/11}$

Cant use your desired Formula for Fishers info in this case. We could if we were reparametrizing β. However, here we change variables unique to each subject (θ_i) so this formula wont work (f). Will results in parts b-e change if y_{i1} & y_{i2} are continuous?

Part (b) will change because

$$P(y_{i1}) \text{ will } = \int_{\Omega_{y_{i2}}} P(y_{i1}, y_{i2}) \cancel{dy_{i1}}\, dy_{i2}$$

$$\sum_{\Omega_{y_{i2}}} P(y_{i1}, y_{i2}) \text{ ?}$$

Likewise,

$P(y_{i2} | y_{i1})$ will necessarily change because $\frac{P(y_{i1}, y_{i2})}{P(y_{i1})}$ will change as a result.

Part (c) will change slightly because again,

$$E(y_{ij}) = \int_{\Omega_{y_{i1}}} \int_{\Omega_{y_{i2}}} y_{ij}\, P(y_{i1}, y_{i2})\, dy_{i1}\, dy_{i2}$$

which will be different than

$$\sum_{y_{i1}} \sum_{y_{i2}} P(y_{i1}, P_{y_{i2}})$$ in the discrete case

Likewise, $E(y_{i1} y_{i2})$ will change.

Part (d) will change slightly because if $\underset{\sim}{\mu}$ & n_{i2} change, then the derivatives in V^{-1} will change slightly. Otherwise, the rest of the problem will remain the same.

Part (e) will only change when V or V^{-1} needs to be used. Otherwise, the other functions should not change in the search for score & Fisher info functions.

$\boxed{\frac{2.5}{3}}$