2004 QUESTION 4

$\underbrace{(X_1,\dots,X_n)'}_{k\ 1s,\ n-k\ 0s}$, $\underbrace{(Y_1,\dots,Y_n)'}_{k\ 1s,\ n-k\ 0s}$ iid random vectors.

their joint pmfs are $P_r\{X_1=i_1,\dots,X_n=i_n\} = \dfrac{1}{\binom{n}{k}}$, $\sum_{j=1}^{n} i_j = k$

$\rightarrow$ each of the $\binom{n}{k}$ ways are equally likely given that $(X_1,\dots,X_n)'$ has k 1s and $n-k$ 0s

Let $N = \sum_{i=1}^{n} |X_i - Y_i|$ denote the number of coordinates at which the two vectors differ. Let M = number of coordinates i for which $X_i = 1$ and $Y_i = 0$. Let L = the number of coordinates i for which $X_i = 0$ and $Y_i = 1$. The $N = L + M$.

(a) X_1 can be 0 or 1. Know that in $X_1, \dots, X_n$ exactly k are 1 and $n-k$ are 0.

$E(X_1) = 0 \cdot P(X_1 = 0) + 1 \cdot P(X_1 = 1)$
$\quad = 1 \cdot \frac{k}{n} = \frac{k}{n}$

Similarly, $E(X_1^2) = 0^2 P(X_1 = 0) + 1^2 \cdot P(X_1 = 1) = \frac{k}{n}$

$\Rightarrow Var(X_1) = E(X_1^2) - E(X_1)^2 = \frac{k}{n} - \left(\frac{k}{n}\right)^2 = \frac{k}{n}\left(1 - \frac{k}{n}\right)$

$E(X_1 X_2) = 0 \cdot P(X_1 \cdot X_2 = 0) + 1 \cdot P(X_1 \cdot X_2 = 1)$
$\quad = 1 \cdot P(X_1 \cdot X_2 = 1) = 1 \cdot P(X_1 = 1, X_2 = 1) = \frac{k}{n} \cdot \frac{(k-1)}{n-1}$

$\Rightarrow Cov(X_1, X_2) = E(X_1 X_2) - E(X_1) \cdot E(X_2)$
$\quad = \frac{k(k-1)}{n(n-1)} - \left(\frac{k}{n}\right)^2 = \frac{k}{n}\left[\frac{k-1}{n-1} - \frac{k}{n}\right]$
$\quad = \frac{k}{n}\left[\frac{n(k-1) - (n-1)k}{n(n-1)}\right] = \frac{k}{n}\left[\frac{\cancel{nk} - n - (\cancel{nk} - k)}{n(n-1)}\right]$
$\quad = \frac{k(k-n)}{n^2(n-1)}$

b) $E[X_1 Y_1] = E_{Y_1}\{E[X_1 Y_1 | Y_1]\} = E_{Y_1}\{Y_1 \cdot \frac{k}{n}\} = \left(\frac{k}{n}\right)^2$

$Var(X_1 Y_1) = E_{Y_1}\{Var(X_1 Y_1 | Y_1)\} + Var_{Y_1}\{E(X_1 Y_1 | Y_1)\}$
$\quad E_{Y_1}\{Y_1^2 [\frac{k}{n}(1 - \frac{k}{n})]\} + Var_{Y_1}\{Y_1 \cdot (\frac{k}{n})\}$
$\quad = \left(\frac{k}{n}\right)^2 \left(1 - \frac{k}{n}\right) + \left(\frac{k}{n}\right)^2 \cdot \frac{k}{n}\left(1 - \frac{k}{n}\right)$
$\quad \left(\frac{k}{n}\right)^2 \left(1 - \frac{k}{n}\right)\left[1 + \frac{k}{n}\right] = \left(\frac{k}{n}\right)^2 \left[1 - \left(\frac{k}{n}\right)^2\right]$

$Cov(X_1 Y_1, X_2 Y_2) = E(X_1 Y_1 X_2 Y_2) - E(X_1 Y_1) \cdot E(X_2 Y_2)$
$\quad E(X_1 X_2) \cdot E(Y_1 Y_2) - E(X_1 Y_1) \cdot E(X_2 Y_2)$
$\quad \left[\frac{k(k-1)}{n(n-1)}\right]^2 - \left(\frac{k}{n}\right)^4 = \left(\frac{k}{n}\right)^2 \left[\left(\frac{k-1}{n-1}\right)^2 - \left(\frac{k}{n}\right)^2\right]$

$\Rightarrow Corr(X_1 Y_1, X_2 Y_2) = \dfrac{Cov(X_1 Y_1, X_2 Y_2)}{\sqrt{Var(X_1 Y_1) Var(X_2 Y_2)}} = \dfrac{\cancel{\left(\frac{k}{n}\right)^2}\left[\left(\frac{k-1}{n-1}\right)^2 - \left(\frac{k}{n}\right)^2\right]}{\cancel{\left(\frac{k}{n}\right)^2}\left[1 - \left(\frac{k}{n}\right)^2\right]} = \dfrac{\left(\frac{k-1}{n-1}\right)^2 - \left(\frac{k}{n}\right)^2}{1 - \left(\frac{k}{n}\right)^2}$

$$\underset{X}{\begin{bmatrix}0\\1\\0\\0\\1\end{bmatrix}} \quad \underset{Y}{\begin{bmatrix}1\\0\\1\\0\\0\end{bmatrix}} \longrightarrow L = 2,\ M = 2,\ N = 4 \qquad \begin{bmatrix}0\\1\\1\\1\end{bmatrix} \quad \begin{bmatrix}1\\1\\1\\0\end{bmatrix}$$

(c) Find the exact distribution of $N-2M$. $N = \sum |X_i - Y_i| = L + M$, where L = number of i where $X_i = 0$, $Y_i = 1$, M = number of i where $X_i = 1$, $Y_i = 0$.

Think hypergeometric. Have $2n$ total observations. Know each group of n contains exactly k 1s.

$$\begin{array}{cc|cc|c} & & Y & & \\ & & 0 & 1 & \\ \hline X & 0 & S_1 & S_2 & \\ & 1 & S_3 & S_4 & k \\ \hline & & & k & \end{array}$$

S_1: # positions where both $= 0$.

$\Rightarrow S_2 = S_3 = M$

$N = S_2 + S_3 = 2M$

$\Rightarrow N - 2M = 0.$

(d) $N = \sum |X_i - Y_i|$ where $\sum X_i = \sum Y_i = k$. If n total and k 1s, $N \in \{0, 1, 2, \dots, n\}$. $\min(2k, 2(n-k))$

$E(N) = 2k\left(1 - \frac{k}{n}\right)$

$$E\left\{\sum_j I_j [X_j \neq Y_j]\right\} = \sum_j E\{I_j[X_j \neq Y_j]\}$$
$$= \sum_j P[X_j \neq Y_j]$$
$$= n \cdot 2\left[\frac{k}{n}\left(1 - \frac{k}{n}\right)\right] = 2k\left(1 - \frac{k}{n}\right)$$

$$\mathrm{Var}(N) = \mathrm{Var}\Big\{\sum_j \underbrace{I_j[X_j \neq Y_j]}_{Z_j}\Big\} = \sum \mathrm{Var}\, Z_j + \sum_{j \neq j'} \mathrm{Cov}(Z_j, Z_{j'})$$

$$= n\, E(Z_j)[1 - E(Z_j)] + \sum_{j \neq j'} E(Z_j Z_{j'}) - \underbrace{E(Z_j)E(Z_{j'})}_{[E(Z_j)]^2}$$

$$E[Z_j Z_{j'}] = 2\left[\frac{\binom{n-1}{k-1}^2}{\binom{n}{k}^2} + \frac{\binom{n-2}{k-2}\binom{n-2}{k}}{\binom{n}{k}^2}\right]$$

(e) dist. of M

M = number of i where $X_i = 1$, $Y_i = 0$ given $\sum X_i = \sum Y_i = k$.

Hypergeometric

Instead of $X \mid X + Y$

can use

$Z = \frac{X}{X+Y}$ if $Z \perp (X+Y)$

[2×2 table with fixed margins]

→ hypergeometric

2005 QUESTION 1

$X_1, \dots, X_n$ iid $f(x|\theta,\nu) = \theta \nu^{\theta} x^{-(\theta+1)} I(x \geq \nu)$, $\theta > 0, \nu > 0$

(a) $L(\theta, \nu | \underset{\sim}{X}) = \theta^n \nu^{n\theta} \left(\frac{1}{\prod X_i}\right)^{\theta+1} I(\nu \leq X_{(1)})$

$\Rightarrow \hat{\nu} = X_{(1)}$

$\ell(\theta, \nu | \underset{\sim}{X}) = n \log \theta + n\theta \log \nu - (\theta+1) \log(\prod X_i)$

$\frac{\partial \ell}{\partial \theta} = \frac{n}{\theta} + n \log \nu - \log(\prod X_i) = 0$

$\Rightarrow \frac{n}{\hat{\theta}} = \log\left[\frac{\prod X_i}{X_{(1)}^n}\right] \Rightarrow \frac{1}{\hat{\theta}} = T(\underset{\sim}{x})$ where $T(\underset{\sim}{x}) = \frac{1}{n} \log\left[\frac{\prod X_i}{X_{(1)}^n}\right]$

$\Rightarrow \hat{\theta} = \frac{1}{T(x)}$

(b) θ, ν unknown. wish to test $H_0: \theta = 1$ against $H_1: \theta \neq 1$.

$$LRT = \frac{\sup_{H_0} L(\theta, \nu|\underset{\sim}{x})}{\sup_{H_0 \cup H_1} L(\theta, \nu | \underset{\sim}{x})} = \frac{X_{(1)}^n \left(\frac{1}{\prod X_i}\right)^2}{\hat{\theta}^n X_{(1)}^{n\hat{\theta}} \left(\frac{1}{\prod X_i}\right)^{\hat{\theta}+1}} < c$$

$$= X_{(1)}^{n - \frac{n}{T(x)}} T(x)^n \left(\frac{1}{\prod X_i}\right)^{1 - \frac{1}{T(x)}} < c$$

$$= T(x)^n \left[\frac{X_{(1)}^n}{\prod X_i}\right] \left[\frac{X_{(1)}^n}{\prod X_i}\right]^{-\frac{1}{T(x)}} < c \quad \text{note } e^{nT(x)} = \frac{\prod X_i}{X_{(1)}^n}$$

$$= T(x)^n e^{-nT(x)} \cdot e^n < c$$

$$= T(x)^n e^{-nT(x)} < ce^{-n} = c_*$$

Consider the function $f(x) = x^n e^{-nx}$ $\quad f(0) = 0, \quad f(\infty) = 0$

and f is unimodal

$$f'(x) = (nx^{n-1}) e^{-nx} + x^n (e^{-nx})(-n) = 0 = nx^{n-1} e^{-nx} (1 - x)$$

$$\Rightarrow nx^{n-1} e^{-nx} = nx^n e^{-nx} \Rightarrow x = 1$$

$$1 = x$$

$$f''(x) = n(n-1)x^{n-2} \left[e^{-nx}(1-x)\right] + nx^{n-1} \left[-ne^{-nx}(1-x) - e^{-nx}\right]$$

$$f''(1) = -ne^{-n} \quad (< 0 \Rightarrow \text{conc. down})$$

so that LRT is a unimodal function of $T(x)$.

$\Rightarrow$ reject if $LRT < c$

$\Rightarrow$ reject if $T(x) \leq c_1$ or $T(x) \geq c_2$

(c) θ, ν unknown and wish to test $H_0: \nu = \nu_0, \theta = \theta_0$ vs $H_1: \nu < \nu_0, \theta < \theta_0$.

To find UMP test, look for statistic with MLR. let $\theta_2 > \theta_1$, $\nu_2 > \nu_1$

$$\frac{f(\underset{\sim}{x} | \theta_2, \nu_2)}{f(\underset{\sim}{x} | \theta_1, \nu_1)} = \frac{\theta_2 \nu_2^{n\theta_2} \left(\frac{1}{\prod X_i}\right)^{\theta_2+1} I(X_{(1)} \geq \nu_2)}{\theta_1 \nu_1^{n\theta_1} \left(\frac{1}{\prod X_i}\right)^{\theta_1+1} I(X_{(1)} \geq \nu_1)} = c^* \left(\frac{1}{\prod X_i}\right)^{\theta_2 - \theta_1} \frac{I(X_{(1)} \geq \nu_2)}{I(X_{(1)} \geq \nu_1)}$$

So that $\dfrac{f(\underset{\sim}{x}|\theta_2, \nu_2)}{f(\underset{\sim}{x}|\theta_1, \nu_1)} = \begin{cases} \text{undef} & x_{(1)} \le \nu_1 \\ 0 & x_{(1)} \in (\nu_1, \nu_2) \\ c^*\left(\frac{1}{\Pi x_i}\right)^{\theta_2 - \theta_1} & x_{(1)} \ge \nu_2 \end{cases}$

So that this ratio is monotone in $T(X) = \frac{1}{\Pi X_i}$

A UMP test is given by

$\phi[T(X)] = \begin{cases} 1 & \text{if } T(X) < c \\ 0 & \text{if } T(X) > c \end{cases}$

to make the test size α, note that

$Pr_{H_0}\left[\frac{1}{\Pi X_i} < c\right] = Pr_{H_0}\left[-\log(\Pi X_i) < c_1\right] = \alpha \qquad \log \Pi X_i = \Sigma \log X_i$

$= Pr_{H_0}\left[\log \Pi X_i > c_1\right] = \alpha$

Find dist of $\Sigma \log X_i$.

Note that $\int_0^t f(x|\theta,\nu)\,dx = \nu^\theta\left[-x^{-\theta}\right]_\nu^t = \nu^\theta\left[\nu^{-\theta} - t^{-\theta}\right] = 1 - \left(\frac{\nu}{t}\right)^\theta, \quad t > \nu$

$\Rightarrow P(\log X < t)$

$= P(X < e^t) = 1 - \left(\frac{\nu}{e^t}\right)^\theta = 1 - \nu^\theta e^{-t\theta}$

$= 1 - e^{\theta \log \nu} e^{-t\theta}$

$= 1 - e^{-\theta(t - \log \nu)}$, a shifted exponential

$\Rightarrow \log X - \log \nu \sim$ exponential (mean $= \frac{1}{\theta}$) $\Rightarrow \Sigma \log\left(\frac{x_i}{\nu}\right) \sim \text{Gamma}\left(n, \frac{1}{\theta}\right)$

$\log\left(\frac{X}{\nu}\right)$

$\Rightarrow Pr_{H_0}\left[\log \Pi X_i > c_\alpha\right] = Pr_{H_0}\left[\Sigma \log\left(\frac{x_i}{\nu}\right) + n \log \nu > c_\alpha\right]$

$= Pr_{H_0}\Big[\underbrace{\Sigma \log\left(\tfrac{x_i}{\nu}\right)}_{J \sim \text{Gamma}(n, \frac{1}{\theta})} > \underbrace{c_\alpha - n \log \nu_0}_{c^*}\Big]$

$\Rightarrow c^*$ chosen so that $\underbrace{\int_{c^*}^\infty \frac{\theta^n}{\Gamma(n)} j^{n-1} e^{-j\theta}\,dj}_{\text{call this } 1 - G(c^*)} = \alpha$

$\Rightarrow G_{n,\frac{1}{\theta}}(c^*) = 1 - \alpha \Rightarrow c^* = G^{-1}_{n,\frac{1}{\theta}}(1-\alpha)$

$c^* = G^{-1}_{n,\frac{1}{\theta}}(1-\alpha) \Rightarrow$ the UMP size α test is

$\phi\left[\Sigma \log\left(\frac{x_i}{\nu}\right)\right] = \begin{cases} 1 & \text{if } \Sigma \log\left(\frac{x_i}{\nu}\right) > c^* \\ 0 & \text{otherwise} \end{cases}$

(d) θ known, ν unknown. test $H_0: \nu = \nu_0$ vs $H_1: \nu \ne \nu_0$. Consider $Y = X^{-\theta}$

$P(Y < t) = P(X^{-\theta} < t)$

$= P(X \ge t^{-\frac{1}{\theta}}) = 1 - \nu^\theta (t^{-\frac{1}{\theta}})^{-\theta} = 1 - \nu^\theta t \Rightarrow Y \sim U(0, \nu^{-\theta})$

[illegible] ν_1, $\frac{f(\ldots)}{[illegible]} = \frac{\nu_2}{\nu_1} \frac{I\{y_{(n)} < \nu_2\}}{I\{y_{(n)} < \nu_1\}}$ [illegible] MLR in $y_{(n)}$

Know UMP test will have form

$$\phi(Y) = \begin{cases} 1 & \text{if } Y_{(n)} < c_1 \text{ or } Y_{(n)} > c_2 \\ 0 & \text{otherwise} \end{cases} \qquad (*)$$

Must optimally choose c_1, c_2. Consider $H_0: \nu = \nu_0$ vs. $H_1: \nu < \nu_0$.
The UMP test has form

$$\phi_1(Y) = \begin{cases} 1 & \text{if } Y_{(n)} < c_\alpha \\ 0 & \text{otherwise} \end{cases}$$

where c_α chosen so that $Pr_{H_0}[Y_{(n)} < c_\alpha] = \left(\frac{c_\alpha}{\nu_0}\right)^n = \alpha \Rightarrow c_\alpha = \nu_0 \alpha^{\frac{1}{n}}$

$\Rightarrow$ Power function is $\beta_{\phi_1}(\nu) = Pr_{H_1}[Y_{(n)} < c_\alpha] = \left(\frac{c_\alpha}{\nu}\right)^n, \quad \nu < \nu_0$

$$= \alpha\left(\frac{\nu_0}{\nu}\right)^n$$

Consider $H_0: \nu = \nu_0$ vs. $H_1: \nu > \nu_0$. The UMP test has form

$$\phi_2(Y) = \begin{cases} 1 & \text{if } Y_{(n)} > c_\alpha \\ 0 & \text{otherwise} \end{cases}$$

where c_α chosen so that $Pr_{H_0}[Y_{(n)} > c_\alpha] = 1 - Pr_{H_0}[Y_{(n)} < c_\alpha] = \alpha$

$$\Rightarrow Pr_{H_0}[Y_{(n)} < c_\alpha] = 1-\alpha$$

$$\Rightarrow \left(\frac{c_\alpha}{\nu_0}\right)^n = (1-\alpha) \Rightarrow c_\alpha = \nu_0(1-\alpha)^{\frac{1}{n}}$$

$\Rightarrow$ power function is $\beta_{\phi_2}(\nu) = Pr_{H_1}[Y_{(n)} > c_\alpha]$

$$= 1 - Pr_{H_1}[Y_{(n)} < c_\alpha] = 1 - \left(\frac{c_\alpha}{\nu}\right)^n, \quad \nu > \nu_0$$

$$= 1 - \left(\frac{\nu_0}{\nu}\right)^n(1-\alpha)$$

If we can find a decision rule ϕ_3 whose power function is

$$\beta_{\phi_3}(\nu) = \alpha\left(\frac{\nu_0}{\nu}\right)^n I(\nu < \nu_0) + \left\{1 - \left(\frac{\nu_0}{\nu}\right)^n(1-\alpha)\right\} I(\nu > \nu_0)\},$$

and whose form is given above in $(*)$, then ϕ_3 is the UMP test of $H_0: \nu = \nu_0$ vs. $H_1: \nu \neq \nu_0$.

Choose c_1, c_2 so that

(1) $Pr_{H_0}[Y_{(n)} < c_1] + Pr_{H_0}[Y_{(n)} > c_2] = \alpha \Rightarrow \left(\frac{c_1}{\nu_0}\right)^n + 1 - \left(\frac{c_2}{\nu_0}\right)^n = \alpha \Rightarrow \underline{\left(\frac{c_1}{\nu_0}\right)^n - \left(\frac{c_2}{\nu_0}\right)^n = \alpha - 1}$

(2) for $\nu < \nu_0$, $\beta_{\phi_3}(\nu) = \underline{\alpha\left(\frac{\nu_0}{\nu}\right)^n = \left(\frac{c_1}{\nu}\right)^n} \Rightarrow c_1 = \nu_0 \alpha^{\frac{1}{n}}$

(3) for $\nu > \nu_0$, $\beta_{\phi_3}(\nu) = \underline{1 - \left(\frac{\nu_0}{\nu}\right)^n(1-\alpha) = 1 - \left(\frac{c_2}{\nu}\right)^n} \Rightarrow c_2 = \nu_0(1-\alpha)^{\frac{1}{n}}$

$$\Rightarrow \underline{\left(\frac{\nu_0}{\nu}\right)^n(1-\alpha) = \left(\frac{c_2}{\nu}\right)^n, \quad \nu > \nu_0}$$

c_1, c_2 satisfying (1), (2), (3) will lead to a UMP test

2005 QUESTION 4

$X_1, \dots, X_n$ iid $N(\mu, \sigma^2)$ $\perp$ $Y_1, \dots, Y_n$ iid $N(\mu, \gamma\sigma^2)$, $\gamma > 0$

(a) $L(\mu, \sigma^2 | \gamma, \underset{\sim}{X}, \underset{\sim}{Y}) \propto (2\pi\sigma^2)^{-n} \exp\left\{ \frac{-1}{2\sigma^2}\left[\sum X_i^2 - 2\mu \sum X_i + \frac{1}{\gamma}\left(\sum Y_i^2 - 2\mu \sum Y_i \right) \right] \right\}$

$\propto \exp\left\{ \frac{-1}{\sigma^2}\left[\frac{1}{2}\sum X_i^2 + \frac{1}{2\gamma}\sum Y_i^2 - \mu \underbrace{\left(\sum X_i + \frac{1}{\gamma}\sum Y_i \right)}_{\text{CS for } \mu} \right] \right\}$

$E\left[\sum X_i + \frac{1}{\gamma}\sum Y_i \right] = n\mu\left(1 + \frac{1}{\gamma}\right) = n\mu\left(\frac{\gamma+1}{\gamma}\right) = n\left(\frac{\gamma+1}{\gamma}\right)\mu$

$\Rightarrow$ The UMVUE of μ is $\hat{\mu}_n = \frac{\gamma}{n(1+\gamma)}\left[\sum X_i + \frac{1}{\gamma}\sum Y_i \right] = \frac{\gamma}{1+\gamma}\bar{X}_n + \frac{1}{1+\gamma}\bar{Y}_n$

(b) γ, μ, σ^2 unknown. Classic results: $S_X^2 \xrightarrow{a.s.} \sigma^2$, $S_Y^2 \xrightarrow{a.s.} \gamma\sigma^2$, $S_X^2 \perp S_Y^2$,
$\bar{X} \perp \bar{Y}$, $S_X^2 \perp \bar{X}$, $S_Y^2 \perp \bar{Y}$. Let $\hat{\gamma}_n = \frac{S_X^2}{S_Y^2}$, $\hat{\mu}_n = \frac{\hat{\gamma}_n}{1+\hat{\gamma}_n}\bar{X}_n + \frac{1}{1+\hat{\gamma}_n}\bar{Y}_n$.

(i) $E(\hat{\mu}_n) = E\left[\frac{\hat{\gamma}_n}{1+\hat{\gamma}_n}\bar{X}_n + \frac{1}{1+\hat{\gamma}_n}\bar{Y}_n \right]$

$= E[\bar{X}_n]\, E\left[\frac{\hat{\gamma}_n}{1+\hat{\gamma}_n}\right] + E[\bar{Y}_n]\, E\left[\frac{1}{1+\hat{\gamma}_n}\right]$ as $\hat{\gamma}_n \perp \bar{X}$, $\hat{\gamma}_n \perp \bar{Y}$

$= \mu\left\{ E\left[\frac{\hat{\gamma}_n}{1+\hat{\gamma}_n}\right] + E\left[\frac{1}{1+\hat{\gamma}_n}\right] \right\}$

$= \mu\left\{ \underbrace{E\left[\frac{\hat{\gamma}_n}{1+\hat{\gamma}_n} + \frac{1}{1+\hat{\gamma}_n} \right]}_{=1} \right\} = \mu.$

(ii) $\sqrt{n}(\hat{\mu}_n - \mu)$

$= \sqrt{n}\left[\frac{\hat{\gamma}_n}{1+\hat{\gamma}_n}\bar{X} + \frac{1}{1+\hat{\gamma}_n}\bar{Y} - \mu\left(\frac{\hat{\gamma}_n}{1+\hat{\gamma}_n} + \frac{1}{1+\hat{\gamma}_n} \right) \right]$

$= \frac{\hat{\gamma}_n}{1+\hat{\gamma}_n}\underbrace{\sqrt{n}(\bar{X}-\mu)}_{N(0,\sigma^2)} + \frac{1}{1+\hat{\gamma}_n}\underbrace{\sqrt{n}(\bar{Y}-\mu)}_{N(0,\gamma\sigma^2)}$

Note $\hat{\gamma}_n \xrightarrow{a.s.} \gamma \Rightarrow \frac{\hat{\gamma}_n}{1+\hat{\gamma}_n} \xrightarrow{a.s.} \frac{\gamma}{1+\gamma}$

$\Rightarrow \frac{1}{1+\hat{\gamma}_n} \xrightarrow{a.s.} \frac{1}{1+\gamma}$

(Slutsky's) $\xrightarrow{d} N\left[0, \underbrace{\left(\frac{\gamma}{1+\gamma}\right)^2\sigma^2 + \left(\frac{1}{1+\gamma}\right)^2\gamma\sigma^2}\right]$

$= \sigma^2\left[\frac{\gamma^2+\gamma}{(1+\gamma)^2}\right] = \sigma^2\left(\frac{\gamma}{1+\gamma}\right)$

Then, a 95% CI for μ is $\hat{\mu}_n \pm 1.96\sqrt{\frac{\hat{\sigma}^2\left(\frac{\hat{\gamma}}{1+\hat{\gamma}}\right)}{n}}$

(iii) Find information bound for μ.

$\ell_n \propto -\frac{1}{2\sigma^2}\left\{ \sum(X_i-\mu)^2 + \frac{1}{\gamma}\sum(Y_i-\mu)^2 \right\} - n\log(\sigma^2)$

$\frac{\partial \ell}{\partial \mu} = \frac{1}{\sigma^2}\sum(X_i-\mu) + \frac{1}{\gamma\sigma^2}\sum(Y_i-\mu)$ $\quad \frac{1}{\sigma^2}\left[\sum X_i + \frac{1}{\gamma}\sum Y_i\right] - n\mu\left(\frac{1}{\sigma^2} + \frac{1}{\gamma\sigma^2}\right)$ $\quad \frac{n}{\sigma^2}\left(1+\frac{1}{\gamma}\right) = \frac{n}{\sigma^2}\left(\frac{\gamma+1}{\gamma}\right)$

$\frac{\partial \ell}{\partial \sigma^2} = \frac{1}{2\sigma^4}\left\{ \sum(X_i-\mu)^2 + \frac{1}{\gamma}\sum(Y_i-\mu)^2 \right\} - \frac{n}{\sigma^2}$

$\frac{\partial \ell}{\partial \gamma} = \frac{1}{\gamma^2}\frac{1}{2\sigma^2}\sum(Y_i-\mu)^2$

$\Rightarrow I_1^{-1}(\mu) = \frac{\sigma^2\gamma}{\gamma+1}$

✓ ($\hat{\mu}_n$ asy efficient).

$E\left[\dots \right]$ [illegible]

When $\delta = 1$, $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2) \perp Y_1, \ldots, Y_n \overset{iid}{\sim} N(\mu, \sigma^2)$

$$\Rightarrow \frac{S_X^2}{\sigma^2} \sim \chi^2_{n-1} \quad \perp \quad \frac{S_Y^2}{\sigma^2} \sim \chi^2_{n-1}$$

(iv) $\hat{\delta}_n = \dfrac{S_Y^2}{S_X^2} \Rightarrow \dfrac{1}{1+\hat{\delta}_n} = \left[\dfrac{\frac{1}{\sigma^2} S_X^2}{\frac{1}{\sigma^2}(S_X^2 + S_Y^2)}\right] = \dfrac{\text{Gamma}\left(\frac{n-1}{2}, 2\right)}{\text{Gamma}\left(\frac{n-1}{2}, 2\right) + \text{Gamma}\left(\frac{n-1}{2}, 2\right)} \sim \text{Beta}\left(\frac{n-1}{2}, \frac{n-1}{2}\right)$

$$\Rightarrow \frac{\hat{\delta}_n}{1+\hat{\delta}_n} = \frac{S_Y^2}{S_X^2 + S_Y^2} \sim \text{Beta}\left(\frac{n-1}{2}, \frac{n-1}{2}\right)$$

$$\left.\begin{aligned} &\Rightarrow E\left[\frac{\hat{\delta}_n}{1+\hat{\delta}_n}\right] = E\left[\frac{1}{1+\hat{\delta}_n}\right] = \frac{1}{2} \\ &\Rightarrow E\left[\left(\frac{\hat{\delta}_n}{1+\hat{\delta}_n}\right)^2\right] = E\left[\left(\frac{1}{1+\hat{\delta}_n}\right)^2\right] = \frac{n+1}{4n} \end{aligned}\right\} \Rightarrow \text{Var}\left(\frac{\hat{\delta}_n}{1+\hat{\delta}_n}\right) = \frac{n+1}{4n} - \frac{1}{4} = \frac{1}{4n}$$

(v) $\text{Var}(\hat{\mu}_n) = \text{Var}\left[\overbrace{\dfrac{\hat{\delta}_n}{1+\hat{\delta}_n}}^{A} \bar{X}_n + \overbrace{\dfrac{1}{1+\hat{\delta}_n}}^{B} \bar{Y}_n\right]$

$$= E_{A,B}\left\{\text{Var}\left(\frac{\hat{\delta}_n}{1+\hat{\delta}_n}\bar{X}_n + \frac{1}{1+\hat{\delta}_n}\bar{Y}_n \,\middle|\, A, B\right)\right\} + \text{Var}_{A,B}\left\{E\left(\frac{\hat{\delta}_n}{1+\hat{\delta}_n}\bar{X}_n + \frac{1}{1+\hat{\delta}_n}\bar{Y}_n \,\middle|\, A, B\right)\right\}$$

$$= E_{A,B}\left\{(A^2 + B^2)\left(\frac{\sigma^2}{n}\right)\right\} + \text{Var}_{A,B}\left\{(A+B)\mu\right\}$$

$\cancel{A \perp B}$ Var(A+B) · Var(A) · Var(B) − Var(A)

$$= \frac{n+1}{2n^2}\sigma^2 + \mu^2 \cdot \frac{1}{2n}$$

$$= \left(\frac{1}{2n} + \frac{1}{2n^2}\right)\sigma^2 + \mu^2\left(\frac{1}{2n}\right) = \frac{1}{2n}\left(\frac{\sigma^2}{n} + \mu^2\right) + \frac{1}{2}\frac{\sigma^2}{n}$$

$$\lim_{n \to \infty} \text{Var}(\hat{\mu}_n) = \sigma^2 \underbrace{\lim_{n \to \infty}\left(\frac{1}{2n} + \frac{1}{2n^2}\right)}_{0} + \mu^2 \underbrace{\lim_{n \to \infty}\left(\frac{1}{2n}\right)}_{0} = 0.$$

$X|\theta$
$\downarrow$

$X \sim Bin(n, \theta)$, $n \geq 2$. Test $H_0: \theta \leq \theta_0$ vs. $H_1: \theta > \theta_0$. Let $\mathfrak{I} = \{T_j : j = 0, 1, \ldots, n-1\}$ be a class of nonrandomized dec. rules, s.t. $T_j(X) = 1$ iff $X \geq j+1$. Consider the 0-1 loss function.

(i) with $\theta \sim U(0,1)$ as a prior, show that the Bayes rule within $\mathfrak{I}$ is $T_{j^*}(X)$, where j^* is the largest $j \in \{0, 1, \ldots, n-1\}$ such that $B_{j+1, n-j+1}(\theta_0) \geq \frac{1}{2}$ i.e. $P(\theta \leq \theta_0) \geq \frac{1}{2}$.

(ii) Derive a minimax rule over the class $\mathfrak{I}$.

(i) Find Bayes risk under 0-1 loss. Recall

$$\phi(T_j(x)) = \begin{cases} 1 & X \geq j+1 \quad (\text{reject } H_0) \\ 0 & \text{otherwise} \quad (\text{accept } H_0) \end{cases}$$

An error is made with the the following probability (frequentist risk)

$$= Pr_{H_0}[X \geq j+1] + Pr_{H_1}[X \leq j]$$
$$= Pr[X \geq j+1 \mid \theta \leq \theta_0] + Pr[X \leq j \mid \theta > \theta_0] = \sum_{x=j+1}^{n} \binom{n}{x} \theta_0^x (1-\theta_0)^{n-x} \times I_{(0,\theta_0]}(\theta) + \sum_{x=0}^{j} \binom{n}{x} \theta_1^x (1-\theta_1)^{n-x} \times I_{(\theta_0,1)}(\theta)$$
$$= R(\phi, \theta)$$

The Bayes risk is given as

$$r(\phi) = \int_\Theta R(\phi, \theta)\, p(\theta)\, d\theta$$

To find the Bayes rule, can minimize posterior expected loss

$X|\theta \sim Bin(n, \theta)$, $\theta \sim U(0,1)$ [a Beta(1,1) dist.] $\Rightarrow \theta|X \sim$ Beta (conjugate prior)

$$p(\theta|x) = \frac{p(x|\theta) \cdot p(\theta)}{\int_\Theta p(x|\theta) \cdot p(\theta)} = \frac{\binom{n}{x}\theta^x(1-\theta)^{n-x} \cdot 1}{\int \cancel{\binom{n}{x}}\theta^x(1-\theta)^{n-x}\, d\theta} = \frac{\frac{\Gamma(n+2)}{\Gamma(x+1)\Gamma(n-x+1)} \theta^{x+1-1}(1-\theta)^{n-x+1-1}}{\int \frac{\Gamma(n+2)}{\Gamma(x+1)\Gamma(n-x+1)} \theta^x (1-\theta)^{n-x}}$$

$\theta|x \sim Beta(x+1, n-x+1)$

A loss is incurred if $X \geq j+1$ while $\theta \leq \theta_0$ or if $X \leq j$ while $\theta > \theta_0$

$$r(\phi) = \int_\Theta R(\phi, \theta)\, p(\theta)\, d\theta$$
$$= \sum_{x=j+1}^{n} \binom{n}{x} \int_0^{\theta_0} \theta^x (1-\theta)^{n-x}\, d\theta + \sum_{x=0}^{j} \binom{n}{x} \int_{\theta_0}^{1} \theta^x (1-\theta)^{n-x}\, d\theta$$
$$= \sum_{x=j+1}^{n} Pr[\theta \leq \theta_0 | x] + \sum_{x=0}^{j} Pr[\theta > \theta_0 | x]$$

$\theta|x \sim Beta(x+1, n-x+1)$ $\qquad$ $\theta|x \sim Beta(x+1, n-x+1)$

The family $\{B_{\beta+1, n-\beta+1}(y) : \beta > 0\}$ is exp. fam. with MLR in $\log y - \log(1-y)$

For [illegible] $(T_{j-1}$ [illegible] $T_j)$ [illegible] $B_{j+1, n-j+1}(\theta)$ [illegible]

$P(Y \leq t) = P(\log Y - \log(1-Y) \leq \log t - \log(1-t))$ decreasing in β for $t \in (0,1)$

[illegible] (θ_0) decreasing [illegible]

If j^* is largest integer s.t. $B_{j+1, n+j+1}(\theta_0) \geq \frac{1}{2}$, then

$r_{T_{j-1}} - r_{T_j} \geq 0$, $j = 1, \dots, j^*$

$r_{T_{j-1}} - r_{T_j} < 0$, $j = j^*+1, \dots, n-1$

$\Rightarrow r_{T_{j^*}} = \min_{j=0,1,\dots,n-1} r_{T_j}$ (Bayes rule).

(ii) $P_\theta(X \leq j)$ decreasing in θ, and $P_\theta(X > j)$ increasing in θ.

$\Rightarrow \sup_{\theta \in (0,1)} R_{T_j}(\theta) = P_{\theta_0}(X > j) = \sum_{k=j+1}^{n} \binom{n}{k} \theta_0^k (1-\theta_0)^{n-k}$

Then, the minimax rule over $\mathcal{J}$ is T_{n-1}

2006 Question 1

(1) Consider the linear model $Y = X\beta + \varepsilon$, where

$$\underbrace{\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix}}_{Y} = \underbrace{\begin{bmatrix} 2 & 6 \\ 4 & 12 \\ 6 & 18 \end{bmatrix}}_{X} \underbrace{\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}}_{\beta} + \underbrace{\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix}}_{\varepsilon}$$

(a) Check whether the following functions are estimable. Justify all answers.

Note that a function of β, $\lambda\beta$, is estimable if and only if λ can be written as ρX for some vector or matrix ρ.

(i) $\beta_1 = \underbrace{[1 \quad 0]}_{\lambda} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$. Must find $\rho = [a \ b \ c]$ such that

$$[a \quad b \quad c] \begin{bmatrix} 2 & 6 \\ 4 & 12 \\ 6 & 18 \end{bmatrix} = [1 \quad 0]$$

so that $\boxed{1}$ $2a + 4b + 6c = 1$ and $\underbrace{6a + 12b + 18c = 0}$

$\boxed{2}$ $\Rightarrow 2a + 4b + 6c = 0$

Cannot find a, b, c to satisfy both equations, so β_1 is not estimable.

(ii) $3\beta_1 + 9\beta_2 = \underbrace{[3 \quad 9]}_{\lambda} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$. Must find $\rho = [a \quad b \quad c]$ such that

$$[a \quad b \quad c] \begin{bmatrix} 2 & 6 \\ 4 & 12 \\ 6 & 18 \end{bmatrix} = [3 \quad 9]$$

so that $\boxed{1}$ $2a + 4b + 6c = 3$ and $\boxed{2}$ $6a + 12b + 18c = 9$

$\Rightarrow 2a + 4b + 6c = 3$ (same equation as $\boxed{1}$)

$\Rightarrow$ any $\rho = [a \quad b \quad c]$ such that $2a + 4b + 6c = 3$, so $3\beta_1 + 9\beta_2$ is estimable.

(b) To find BLUE of $\beta_1 - 3\beta_2 = [1 \quad -3] \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$, first rewrite $\lambda\beta$ as

$= [a \quad b \quad c]$

$\swarrow$

ρ such that $\rho X = \lambda = [-1 \quad -3]$

$$[a \quad b \quad c]\begin{bmatrix} 2 & 6 \\ 4 & 12 \\ 6 & 18 \end{bmatrix} = [-1 \quad -3]$$

$\Rightarrow 2a + 4b + 6c = -1$ in both equations

Choose $a = \frac{1}{2}$, $b = \frac{1}{4}$, $c = \frac{-1}{2}$ as one possible solution, $\Rightarrow \rho = [\frac{1}{2} \quad \frac{1}{4} \quad \frac{-1}{2}]$.

The BLUE of $\rho X \beta = \lambda \beta = [-\beta_1, -3\beta_2]$ is $\rho M Y$

$$= [\tfrac{1}{2} \quad \tfrac{1}{4} \quad \tfrac{-1}{2}]\, M \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix},$$

as X is not full rank ~~use generalized inverse~~

where $M = X(X'X)^{-}X'$, the orthogonal projection operator onto $C(X)$,

where $C(X) = \text{span}\left\{\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}\right\}$

The variance of the BLUE is $\text{Var}(\rho M Y)$

$$= (\rho M)\, \text{Var}\, Y\, (\rho M)'$$

$$= \sigma^2 \rho M \rho', \quad \rho = [\tfrac{1}{2} \quad \tfrac{1}{4} \quad \tfrac{1}{2}]$$

(c) Wish to test $H_0: \beta_1 + 3\beta_2 = 0$ vs. $H_A: \beta_1 + 3\beta_2 \neq 0$.

Under H_0, rewrite model as $Y = X_0 \gamma_0 + \varepsilon$

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} = \begin{bmatrix} 2 & 6 \\ 4 & 12 \\ 6 & 18 \end{bmatrix}\begin{bmatrix} -3\beta_2 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}}_{X_0} \underbrace{\beta_2}_{\gamma_0}$$

$\Rightarrow M_0 = X_0(X_0'X_0)^{-}X_0' = 0$ [orthogonal projection operator onto $C(X_0)$]

(i) To test these hypotheses, use

$$F = \frac{\left[\frac{Y'(M-M_0)Y}{r(M-M_0)}\right]}{\left[\frac{Y'(I-M)Y}{r(I-M)}\right]} = \frac{\left[\frac{Y'MY}{1}\right]}{\left[\frac{Y'(I-M)Y}{2}\right]} \sim F_{1,2} \text{ under } H_0$$

reject if $F > F_{1-\alpha, 1, 2}$.

(ii) Numerator $= \dfrac{\left(\dfrac{Y'(M-M_0)Y}{\sigma^2}\right)}{r(M-M_0)}$ is a $\dfrac{\chi^2_{r(M-M_0)}}{r(M-M_0)}$ random variable,

as $Y \sim N(X\beta, \sigma^2 I) \Rightarrow \frac{Y}{\sigma} \sim N\left(\frac{X\beta}{\sigma}, I\right)$

$$\Rightarrow \left(\frac{Y}{\sigma}\right)' \underbrace{(M-M_0)}_{\text{orthogonal projection operator}} \left(\frac{Y}{\sigma}\right) \sim \chi^2_{r(M-M_0)},\ \delta = \frac{(X\beta)'(M-M_0)(X\beta)}{2\sigma^2}$$

Denominator $= \dfrac{\left(\dfrac{Y'(I-M)Y}{\sigma^2}\right)}{r(I-M)}$ is a $\dfrac{\chi^2_{r(I-M)}}{r(I-M)}$ random variable,

as $Y \sim N(X\beta, \sigma^2 I) \Rightarrow \frac{Y}{\sigma} \sim N\left(\frac{X\beta}{\sigma}, I\right)$

$$\Rightarrow \left(\frac{Y}{\sigma}\right)' \underbrace{(I-M)}_{\text{orthogonal projection operator}} \left(\frac{Y}{\sigma}\right) \sim \chi^2_{r(I-M)},\ \delta = \frac{\overbrace{(X\beta)'(I-M)(X\beta)}^{=0}}{2\sigma^2}$$

also, numerator and denominator are independent, as

$$(M-M_0)(I-M) = \cancel{M-M} - M_0 + M_0 M$$
$$= M_0 M - M_0 = \underbrace{-M_0(I-M)}_{\text{orthogonal to one another}} = 0$$

Then $F = \dfrac{\left[\dfrac{Y'(M-M_0)Y}{r(M-M_0)}\right]}{\left[\dfrac{Y'(I-M)Y}{r(I-M)}\right]} \sim F_{\underbrace{r(M-M_0)}_{=1},\ \underbrace{r(I-M)}_{=2}}$ under H_0

(iii) under H_A, $F \sim F_{1,2}$, $\delta = \dfrac{(X\beta)'(M-M_0)(X\beta)}{2\sigma^2}$

(d) If $Y \sim N(X\beta, \sigma^2 V)$, and V is positive definite, can write $V = QQ'$

Then it follows that $Q^{-1}VQ'^{-1} = I$. Analyze equivalent model

$Q^{-1}Y = Q^{-1}X\beta + Q^{-1}e$, so that $\underbrace{\dfrac{Q^{-1}Y}{\sigma^2}}_{=Y^*} \sim N\Big(\underbrace{\dfrac{Q^{-1}X\beta}{\sigma^2}}_{E(Y^*)}, I\Big)$

so that an F-test can be constructed to test the hypotheses of interest. let $X^* = Q^{-1}X$, $M^* = X^*(X^{*\prime}X^*)^- X^{*\prime}$. then, with $Y^* = Q^{-1}Y$, and noting that $M_0^* = 0$,

$$F = \frac{\left[\frac{Y^{*\prime}(M)Y^*}{r(M)}\right]}{\left[\frac{Y^{*\prime}(I-M)Y^*}{r(I-M)}\right]} \sim F_{r(M)=1,\ r(I-M)=2} \text{ under } H_0,$$

the ratio of two independent scaled χ^2 random variables

as $\frac{Q^{-1}Y}{\sigma} \sim N\left(\frac{Q^{-1}X\beta}{\sigma}, I\right) \Rightarrow \left(\frac{Q^{-1}Y}{\sigma}\right)' M \left(\frac{Q^{-1}Y}{\sigma}\right) \sim \chi^2_{r(M)},\ \delta = \frac{\overbrace{(Q^{-1}X\beta)'M(Q^{-1}X\beta)}^{=0 \text{ under } H_0}}{2\sigma^2}$

$\Rightarrow \left(\frac{Q^{-1}Y}{\sigma}\right)' (I-M) \left(\frac{Q^{-1}Y}{\sigma}\right) \sim \chi^2_{r(I-M)},\ \delta = \frac{(Q^{-1}X\beta)'\overbrace{(I-M)(Q^{-1}X\beta)}^{=0}}{2\sigma^2}$

and numerator $\perp$ denominator, as

$M(I-M) = M-M = 0$

Note that under H_1, the non-centrality parameter in the numerator is not equal to 0, and

$F \sim F_{1,2},\ \delta = \frac{(Q^{-1}X\beta)'M(Q^{-1}X\beta)}{2\sigma^2}$

2006 QUESTION 2

(a) Let $\delta(X,j)$ be the probability of taking action j under the rule δ. Let E_j be the expectation under f_j. Then

$$R_\delta(j) = E_j\left[\sum_{k=1}^{J} L(j,k)\,\delta(X,k)\right] = \sum_{k\neq j} E_j[\delta(X,k)] = 1 - E_j[\delta(X,j)]$$

$$\left(\text{as } \sum_{k=1}^{J} \delta(X,k) = 1\right)$$

(b) The Bayes risk of a decision rule δ is

$$R_\delta = \sum_{j=1}^{J} \pi_j R_\delta(j) = 1 - \sum_{j=1}^{J} \pi_j E_j[\delta(X,j)]$$

(c) Let δ^* be a rule satisfying $\delta^*(X,j) = 1$ if and only if $\pi_j f_j(x) = g(x)$, where $g(x) = \max_{k\in\{1,\dots,J\}} \pi_k f_k(x)$. Then, δ^* is a Bayes rule, since for any rule δ,

$$\begin{aligned} R_\delta &= 1 - \sum_{j=1}^{J} \int \pi_j\,\delta(x,j) f_j(x)\,d\nu \\ &\geq 1 - \sum_{j=1}^{J} \int \delta(x,j)\, g(x)\, d\nu \\ &= 1 - \int g(x)\, d\nu \\ &= 1 - \sum_{j=1}^{J} \int_{g(x) = \pi_j f_j(x)} \pi_j f_j(x)\, d\nu = R_{\delta^*} \end{aligned}$$

so that δ^* minimizes the Bayes risk wrt prior π.

(d) From (c), the Bayes rule $\delta^*(X,j) = 1$ if and only if $\phi(x-\mu_j) > \phi(x-\mu_k)$, $k \neq j$. As $\phi(x-\mu_j) = (2\pi)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(x-\mu_j)^2\right\}$, a nonrandomized Bayes rule takes action 1 if and only if $|X-\mu_1| < |X-\mu_2|$.

(e) Let c be a positive constant and consider a rule δ_c such that $\delta_c(X,1) = 1$ if $f_1(x) > c f_2(x)$, $\delta_c(X,2) = 1$ if $f_1(X) < c f_2(X)$, and $\delta_c(X,1) = \gamma$ if $f_1(x) = c f_2(x)$.

Since $\delta_c(X,j) = 1$ if and only if $\pi_j f_j(x) = \max_k \pi_k f_k(x)$, where $\pi_1 = \frac{1}{c+1}$ and $\pi_2 = \frac{c}{c+1}$, it follows from (c) that δ_c is a Bayes rule.

Let P_j be the probability corresponding to f_j. The risk of δ_c is

$$R_{\delta_c}(1) = P_1[f_1(X) < c f_2(X)] + \gamma P_1[f_1(X) = c f_2(X)]$$

$$R_{\delta_c}(2) = 1 - P_2[f_1(X) < c f_2(X)] + \gamma P_2[f_1(X) = c f_2(X)]$$

constant freq. risk + Bayes $\to$ minimax

Let $\Psi(c) = P_1[f_1(X) \le c f_2(X)] + P_2[f_1(X) \le c f_2(X)] - 1$. Then, Ψ is nondecreasing in c, $\Psi(0) = -1$, $\lim_{c\to\infty} \Psi(c) = 1$, and $\Psi(c) - \Psi(c^-) = P_1[f_1(X) = c f_2(X)] + P_2[f_1(X) = c f_2(X)]$. Let $c_* = \inf\{c : \Psi(c) \ge 0\}$. If $\Psi(c_*) = \Psi(c_*^-)$, we set $\gamma = 0$; otherwise, we set $\gamma = \frac{\Psi(c_*)}{\Psi(c_*) - \Psi(c_*^-)}$. Then, $R_{\delta_{c_*}}(j)$ is constant. For any rule δ, $\sup_j R_\delta(j) \ge R_\delta \ge R_{\delta_{c^*}} = \underbrace{R_{\delta_{c_*}}(j) = \sup_j R_{\delta_{c_*}}(j)}_{\text{due to constant freq. risk}}$.

thus, δ_{c_*} is a minimax rule

2006 QUESTION 3

3. (a) $P(T>t) = P(T>t, \text{no mutation}) + P(T>t, \text{mutation})$

$= (1-\pi) + \pi\exp(-\lambda t)$

(b) dominating measure: Lebesgue on $\mathbb{R}$ + mass on τ

$Y = \begin{cases} \tau, & T \geq \tau \\ t, & t \in [0,\tau) \end{cases}$. must find density

$f_Y(y) = \begin{cases} (1-\pi) + \pi\exp(-\lambda\tau), & y = \tau \\ \frac{\partial}{\partial y}[1 - P(T>y)], & y \in [0,\tau) \end{cases}$

$= \frac{\partial}{\partial y}[\pi - \pi\exp(-\lambda y)]$

$= \pi\lambda\exp(-\lambda y)$

So that the density of Y can be written as

$f_Y(y) = [\pi\lambda\exp(-\lambda y)]^{I[y\in[0,\tau)]}\,[(1-\pi) + \pi\exp(-\lambda\tau)]^{I(y=\tau)}$

(c) Note τ fixed. Show if (π,λ) and $(\tilde{\pi},\tilde{\lambda})$ are equal, then

$\dfrac{f_Y(y|\pi,\lambda)}{f_Y(y|\tilde{\pi},\tilde{\lambda})} = \dfrac{\pi\lambda\exp[y(\tilde{\lambda}-\lambda)]}{\tilde{\pi}\tilde{\lambda}}, \quad y \in [0,\tau) \qquad (*)$

$= \dfrac{(1-\pi) + \pi\exp(-\lambda\tau)}{(1-\tilde{\pi}) + \tilde{\pi}\exp(-\tilde{\lambda}\tau)}, \quad y = \tau$

Only have to examine density part, not point mass. In $(*)$ if $\lambda \neq \tilde{\lambda}$, then $\dfrac{\pi\lambda}{\tilde{\pi}\tilde{\lambda}}\exp[y(\tilde{\lambda}-\lambda)] \neq 1$ over whole range $y \in [0,\tau)$

no matter what value is chosen for π.

(d) $\log f_Y(y_i) = I[y_i \in [0,\tau)]\cdot\log[\pi\lambda\exp(-\lambda y_i)] + I[y_i=\tau]\log[(1-\pi) + \pi\exp(-\lambda\tau)]$

$= I[y_i \in [0,\tau)]\{\log\pi + \log\lambda - \lambda y_i\} + I[y_i=\tau]\log[(1-\pi) + \pi\exp(-\lambda\tau)] = \ell_i$

To find likelihood equations,

$\dfrac{\partial \ell_i}{\partial\lambda} = I[y_i \in [0,\tau)]\{\frac{1}{\lambda} - y_i\} - I[y_i=\tau]\cdot\dfrac{\pi\tau\exp(-\lambda\tau)}{(1-\pi) + \pi\exp(-\lambda\tau)}$

$\dfrac{\partial \ell_i}{\partial\pi} = \dfrac{1}{\pi} I[y_i \in [0,\tau)] - I[y_i=\tau]\cdot\dfrac{[1-\exp(-\lambda\tau)]}{(1-\pi) + \pi\exp(-\lambda\tau)}$

no closed form for π, λ. Algebra to solve for π, the closed form of λ

Now, note the following:

$P(y_i<\tau) = 1 - P(y_i=\tau) = \pi[1-\exp(-\lambda\tau)]$, $P(y_i=\tau) = [(1-\pi) + \pi\exp(-\lambda\tau)]$ and

$P(Y_i = \tau) = [(1-\pi) + \pi \exp(-\lambda\tau)]$

$P(Y_i < \tau) = \pi\,[1 - \exp(-\lambda\tau)]$

$E\{y_i\, I(y_i < \tau)\} = \pi \int_0^{\tau} y_i \cdot \lambda \exp(-\lambda y_i)\, dy_i = \pi\left[\frac{1}{\lambda}(1-\exp(-\lambda\tau)) - \tau\exp(-\lambda\tau)\right]$

The likelihood equations are

$$\frac{\partial \underset{\sim}{\ell}}{\partial \lambda} = \sum_{i=1}^{n} \frac{\partial \ell_i}{\partial \lambda} \quad \text{and} \quad \frac{\partial \underset{\sim}{\ell}}{\partial \pi} = \sum_{i=1}^{n} \frac{\partial \ell_i}{\partial \pi}$$

(e) From MLE theory,

$$\sqrt{n}\left(\begin{bmatrix}\hat\lambda\\ \hat\pi\end{bmatrix} - \begin{bmatrix}\lambda\\ \pi\end{bmatrix}\right) \xrightarrow{d} N_2\left(\begin{bmatrix}0\\0\end{bmatrix}, \underbrace{I^{-1}(\lambda,\pi)}_{\text{inverse Fisher information.}}\right)$$

To find $I^{-1}(\lambda,\pi)$, note that

$$\frac{\partial^2 \ell}{\partial \lambda^2} = I(y_i<\tau)\left(\frac{-1}{\lambda^2}\right) - I(y_i=\tau)\left\{\frac{[(1-\pi)+\pi e^{-\lambda\tau}][-\tau^2\pi e^{-\lambda\tau}] - (\pi\tau e^{-\lambda\tau})^2}{[(1-\pi)+\pi e^{-\lambda\tau}]^2}\right\}$$

$$\frac{\partial^2 \ell}{\partial\lambda\partial\pi} = -I(y_i=\tau)\left\{\frac{[(1-\pi)+\pi e^{-\lambda\tau}][\tau e^{-\lambda\tau}] + [\pi\tau e^{-\lambda\tau}][1-e^{-\lambda\tau}]}{[(1-\pi)+\pi e^{-\lambda\tau}]^2}\right\}$$

$$\frac{\partial^2 \ell}{\partial\pi^2} = \frac{-1}{\pi^2} I(y_i<\tau) - I(y_i=\tau)\left\{\frac{[(1-\pi)+\pi e^{-\lambda\tau}]0 + (1-e^{-\lambda\tau})^2}{[(1-\pi)+\pi e^{-\lambda\tau}]^2}\right\}$$

$$\frac{\partial^2 \ell}{\partial\pi\partial\lambda} = -I(y_i=\tau)\left\{\frac{[(1-\pi)+\pi e^{-\lambda\tau}][\tau e^{-\lambda\tau}] + [1-e^{-\lambda\tau}][\pi\tau e^{-\lambda\tau}]}{[(1-\pi)+\pi e^{-\lambda\tau}]^2}\right\}$$

Then, $I(\lambda,\pi)$ is equal to

$$E\begin{bmatrix} -\frac{\partial^2\ell}{\partial\lambda^2} & -\frac{\partial^2\ell}{\partial\lambda\partial\pi} \\ & -\frac{\partial^2\ell}{\partial\pi^2}\end{bmatrix}, \text{ where}$$

$$E\left[-\frac{\partial^2\ell}{\partial\lambda^2}\right] = \frac{1}{\lambda^2}P(y_i<\tau) - \tau^2\pi e^{-\lambda\tau} - \frac{(\pi\tau e^{-\lambda\tau})^2}{P(y_i=\tau)}$$

$$E\left[\frac{-\partial^2\ell}{\partial\lambda\partial\pi}\right] = \tau e^{-\lambda\tau}\left[1 + \frac{P(Y_i<\tau)}{P(Y_i=\tau)}\right]$$

$$E\left[-\frac{\partial^2\ell}{\partial\pi^2}\right] = \frac{1}{\pi^2}\left[P(y_i<\tau) + \frac{P(Y_i<\tau)^2}{P(Y_i=\tau)}\right]$$

→ Invert to get $I^{-1}(\lambda,\pi)$.

(f) $B \sim \text{Bernoulli}(\pi) \rightarrow P[B=1] = \pi$ and $P(B=0] = (1-\pi)$

$t \mid B=0$ ~ , $t \mid B=1$ ~ exponential [mean $= \frac{1}{\lambda}$]

→ $f(B,t) = [\pi\lambda\exp(-\lambda t)]^{I(B=1)}\,[1-\pi]^{I(B=0)}$

observed data $= \{y_1, \ldots, y_n\} \in [0, \tau]$

complete data

log-likelihood, ℓ_i, is equal to:

$$\ell_i = I[B_i = 1]\left\{\log \pi + \log \lambda - \lambda t_i\right\} + I[B_i = 0]\log(1-\pi)$$

(i) Complete data score vector (likelihood equations)

$$\frac{\partial \ell_i}{\partial \lambda} = I[B_i = 1]\left\{\frac{1}{\lambda} - t_i\right\} \longrightarrow \frac{\partial \ell}{\partial \lambda} = \sum_{i=1}^{n} \frac{\partial \ell_i}{\partial \lambda}$$

$$\frac{\partial \ell_i}{\partial \pi} = I[B_i = 1]\left\{\frac{1}{\pi}\right\} - I[B_i = 0]\left[\frac{1}{1-\pi}\right] \longrightarrow \frac{\partial \ell}{\partial \pi} = \sum_{i=1}^{n} \frac{\partial \ell_i}{\partial \pi}$$

(ii) In E-step, must compute

(1) $E\left\{I[B_i = 1] \cdot t_i \mid \text{observed}, \lambda^{(t)}, \pi^{(t)}\right\}$

(2) $E\left\{I[B_i = 1] \mid \text{observed}, \lambda^{(t)}, \pi^{(t)}\right\}$

(3) $E\left\{I[B_i = 0] \mid \text{observed}, \lambda^{(t)}, \pi^{(t)}\right\}$

To maximize, set likelihood equations equal to 0 so that

$$\frac{1}{\lambda}\sum_{i=1}^{n} I[B_i = 1] = \sum_{i=1}^{n} I[B_i = 1] t_i$$

$$\Rightarrow \hat{\lambda}^{(t+1)} = \frac{\sum_{i=1}^{n} E\left\{I[B_i = 1] \mid \text{observed}, \lambda^{(t)}, \pi^{(t)}\right\}}{\sum_{i=1}^{n} E\left\{I[B_i = 1] \cdot t_i \mid \text{observed}, \lambda^{(t)}, \pi^{(t)}\right\}}$$

and $(1-\pi)\sum_{i=1}^{n} I[B_i = 1] - \pi \sum_{i=1}^{n} I[B_i = 0] = 0$

$$\Rightarrow \sum_{i=1}^{n} I[B_i = 1] - \pi\Big\{\underbrace{\sum_{i=1}^{n} I[B_i = 1] + \sum_{i=1}^{n} I[B_i = 0]}_{= n}\Big\} = 0$$

$$\Rightarrow \hat{\pi}^{(t+1)} = \frac{1}{n}\sum_{i=1}^{n} E\left\{I[B_i = 1] \mid \text{observed}, \lambda^{(t)}, \pi^{(t)}\right\}$$

$$E[\sum B_i \mid \quad] = \sum_{i=1}^{n}\left[I(y_i < \tau) + \frac{\pi e^{-\lambda\tau}}{1-\pi+\pi e^{-\lambda\tau}} I(y_i = \tau)\right]$$

$$E[\sum t_i B_i \mid \quad] = \sum_{i=1}^{n} I(y_i < \tau) y_i + \sum_{i=1}^{n} \frac{\pi[\tau + \frac{1}{\lambda}] e^{-\lambda\tau}}{1-\pi+\pi e^{-\lambda\tau}} I(y_i = \tau)$$

(c). In order to get $\pi_{ij} = \pi_{ji}$ we need $e^{\alpha_i + \beta_j + \gamma_{ij}} = e^{\alpha_j + \beta_i + \gamma_{ji}}$

From marginal homogeneity $\pi_{i+} = \pi_{+i}$ for all i

$$\sum_{j=1}^{k} e^{\alpha_i + \beta_j + \gamma_{ij}} = \sum_{j=1}^{k} e^{\alpha_j + \beta_i + \gamma_{ji}}$$

$$\Rightarrow e^{\alpha_i} \sum_{j=1}^{k} e^{\beta_j + \gamma_{ij}} = e^{\beta_i} \sum_{j=1}^{k} e^{\alpha_j + \gamma_{ji}}$$

by quasi-symmetry $\gamma_{ij} = \gamma_{ji}$

$$\Rightarrow \frac{e^{\alpha_i}}{e^{\beta_i}} = \frac{\sum_{j=1}^{k} e^{\alpha_j + \gamma_{ji}}}{\sum_{j=1}^{k} e^{\beta_j + \gamma_{ji}}} = \text{const} = C.$$

i.e. $e^{\alpha_i} = C e^{\beta_i}$ for $i = 1, \ldots, k$.

So $e^{\alpha_i + \beta_j + \gamma_{ij}} = C \cdot e^{\beta_i} \cdot \frac{e^{\alpha_j}}{C} \cdot e^{\gamma_{ji}} = e^{\alpha_j + \beta_i + \gamma_{ji}}$

i.e. $\pi_{ij} = \pi_{ji}$

Hence marginal homogeneity and quasi-symmetry imply symmetry

(b) (i) $\ln(\pi) = \sum_i^K \sum_j^K z_{ij1} \log \pi_{ij1} + \sum_i^K \sum_j^K z_{ij2} \log \pi_{ij2} + \text{const}$

$$= \sum_i^K \sum_j^K z_{ij1} \log \pi_{ij1} + \sum_i^K \sum_j^K z_{ji2} \log \pi_{ji2} + \text{const}$$

$$= \sum_i \sum_j z_{ij1} \log \pi_{ij1} + \sum_i \sum_j Y_{ji} \log \pi_{ji2} + \text{const}$$

$$= \sum_i \sum_j Y_{ij} \log \pi_{ij1} + \sum_i \sum_j Y_{ij} \log \pi_{ji2} + \text{const}$$

$$= \sum_i \sum_j Y_{ij} \left[(\delta_{ij} + \delta_{ji}) + (\varepsilon_{i1} + \varepsilon_{j2}) + (\phi_{j1} + \phi_{i2}) \right] + \text{const}$$

Set $\alpha_i = \varepsilon_{i1} + \phi_{i2}$ $\quad \beta_j = \phi_{j1} + \varepsilon_{j2}$ $\quad \gamma_{ij} = \delta_{ij} + \delta_{ji}$

$$= \sum_i \sum_j Y_{ij} (\alpha_i + \beta_j + \gamma_{ij}) + \text{const}$$

with $\gamma_{ij} = \gamma_{ji}$

Hence the model setting of (4) is equivalent to those for the quasi-symmetry model of part (a)

So does the score equations.

ii) We should divide the deviance computed from the software by 2

4 (a) Use the reference cell coding

i) $\log(\pi_{ij}) = \log\left(\frac{\mu_{ij}}{n}\right) = \alpha_i + \beta_j + \gamma_{ij}$

$$\Longrightarrow \log\mu_{ij} = \log n + \sum_{s=1}^{k-1}\alpha_s I(s=i) + \sum_{t=1}^{k-1}\beta_t I(t=j) + \sum_{s=1}^{k-1}\sum_{t=1}^{k-1}\gamma_{st} I(s=i,\ t=j)$$

$$\beta = (\lambda_0, \alpha_1, \alpha_2, \cdots, \alpha_{k-1}, \beta_1, \beta_2, \cdots, \beta_{k-1}, \gamma_{11}, \gamma_{12}, \cdots, \gamma_{1\,k-1}, \gamma_{21}, \gamma_{22}, \cdots, \gamma_{2\,k-1}, \cdots, \gamma_{k-1\,1}, \gamma_{k-1\,2}, \cdots, \gamma_{k-1\,k-1})^T$$

$$X = \begin{array}{ll} \mu_{11} & \\ \mu_{12} & s=1 \\ \mu_{1\,k-1} & \\ \mu_{1k} & \\ & \\ & s=2 \\ & \\ & \\ & \\ & s=k-1 \\ & \\ & \\ \mu_{k1} & \\ \mu_{k2} & s=k \\ & \\ \mu_{kk} & \end{array}
\begin{pmatrix}
1 & 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\
1 & 1 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\
1 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\
1 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\
1 & 0 & 1 & \cdots & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\
1 & 0 & 1 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\
1 & 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & \cdots & 0 & 0 & \cdots & 0 \\
1 & 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\
1 & 0 & 0 & \cdots & 1 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \cdots & 1 & 0 & \cdots & 0 \\
1 & 0 & 0 & \cdots & 1 & 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 1 & \cdots & 0 \\
1 & 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 1 \\
1 & 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\
1 & 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\
1 & 0 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\
1 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\
1 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0
\end{pmatrix}$$

$g(\cdot) = \log(\cdot)$

Thus $g(\mu_{ij}) = X\beta$ for $i = 1, \ldots, k$, $j = 1, \ldots, k$

with $\gamma_{ij} = \gamma_{ji}$

ii) $l_f(\pi) = \sum_i^K \sum_j^K y_{ij} \log \pi_{ij} + \text{const}$ under $\sum_i^K \sum_j^K \pi_{ij} = 1$

So we can consider the function $l(\pi) = \sum_i^K \sum_j^K y_{ij} \log \pi_{ij} + \lambda\left(\sum\sum \pi_{ij} - 1\right)$

$$= \sum_i \sum_j y_{ij}(\alpha_i + \beta_j + \gamma_{ij}) - \lambda\left(\sum\sum e^{\alpha_i + \beta_j + \gamma_{ij}} - 1\right)$$

$$\frac{\partial l}{\partial \pi_{ij}} = \frac{y_{ij}}{\pi_{ij}} + \lambda = 0 \implies \lambda = -n$$

So $y_{ij} = n \cdot \pi_{ij}$ where $\pi_{ij} = e^{\alpha_i + \beta_j + \gamma_{ij}}$

2006 QUESTION 5

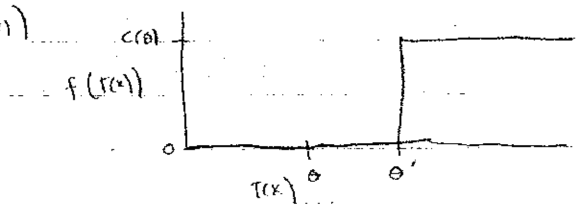
$X_1, \dots, X_n$ iid $f(x|\lambda,\theta) = \begin{cases} \frac{1}{\lambda} e^{-\frac{1}{\lambda}(x-\theta)} & x \geq \theta, \ \lambda > 0, \ -\infty < \theta < \infty \\ 0 & \text{otherwise} \end{cases}$

(a) λ known, θ unknown. Show this family of distributions has the MLR property in some statistic $T(X)$.

$f(\underset{\sim}{x}) = \lambda^{-n} \exp\left\{ -\frac{1}{\lambda} \sum_{i=1}^{n} (x_i - \theta) \right\} \cdot I[x_{(1)} \geq \theta]$

Consider $T(X) = X_{(1)}$. Show for $\theta' > \theta$, $f(\underset{\sim}{x}|\theta') / f(x|\theta)$ is increasing in $T(X)$.

$$\underbrace{\frac{f(x|\theta')}{f(x|\theta)}}_{\text{call this } f(T(x))} = \underbrace{\exp\left\{ \frac{n}{\lambda}(\theta' - \theta) \right\}}_{c(\theta)} \cdot \frac{I[X_{(1)} \geq \theta']}{I[X_{(1)} \geq \theta]}$$



$\left.\right\}$ $\frac{f(x|\theta')}{f(x|\theta)}$ increasing in $X_{(1)}$

$\Rightarrow$ MLR property holds

(b) θ, λ unknown. Let ℓ denote log-likelihood, L likelihood

$\Rightarrow L(\theta,\lambda) = \lambda^{-n} \exp\left\{ -\frac{1}{\lambda} \sum_{i=1}^{n} (x_i - \theta) \right\} \cdot I[\theta \leq x_{(1)}]$

$\hat{\theta} = X_{(1)}$ maximizes $L(\theta,\lambda)$ because it is the largest value θ can take and thus the value that minimizes $\sum_{i=1}^{n} (x_i - \hat{\theta})$ [or maximizes $-\frac{1}{\lambda} \sum_{i=1}^{n} (x_i - \hat{\theta})$]

Then, $\ell(\hat{\theta},\lambda) = -n\log\lambda - \frac{1}{\lambda} \sum_{i=1}^{n} (x_i - \hat{\theta})$

$\Rightarrow \frac{\partial \ell}{\partial \lambda} = \frac{-n}{\lambda} + \frac{1}{\lambda^2} \sum_{i=1}^{n} (x_i - \hat{\theta}) = 0$

$\Rightarrow \hat{\lambda} = \frac{1}{n} \sum_{i=1}^{n} (x_i - X_{(1)})$

(c) The joint distribution of the order statistics of $(X_1, \dots, X_n)$ is:

$f(\underset{\sim}{X}_{(\cdot)}) = \frac{n!}{\lambda^n} \exp\left\{ -\frac{1}{\lambda} \sum_{i=1}^{n} (x_i - \theta) \right\} I[X_1 \geq \theta], \quad X_1 < X_2 < \dots < X_n$

let $Y_1 = X_{(1)}$ $\quad \Rightarrow X_{(1)} = Y_1$

$Y_2 = (n-1)(X_{(2)} - X_{(1)}) \quad \Rightarrow X_{(2)} = \frac{Y_2}{n-1} + Y_1$

$Y_i = (n-i+1)(X_{(i)} - X_{(i-1)}) \quad \Rightarrow X_{(3)} = \frac{Y_3}{n-2} + X_{(2)} = \frac{Y_3}{n-2} + \frac{Y_2}{n-1} + Y_1 \qquad \boxed{\sum_{j=2}^{i} \frac{Y_j}{n-j+1}}$

$Y_n = X_{(n)} - X_{(n-1)} \quad \Rightarrow X_{(n)} = Y_n + X_{(n-1)} = Y_n + \frac{Y_{n-1}}{2} + \frac{Y_{n-2}}{3} + \dots + \frac{Y_3}{n-2} + \frac{Y_2}{n-1} + Y_1$

So that $X_{(1)} = Y_1$, $X_{(i)} = Y_1 + \sum_{j=2}^{i} \frac{Y_j}{n-j+1}$

$$\Rightarrow J = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & \frac{1}{n-1} & 0 & & 0 \\ \vdots & \frac{1}{n-1} & \frac{1}{n-2} & 0 & 0 \\ 1 & \vdots & \vdots & \vdots & 1 \end{bmatrix} \Rightarrow |\det J| = \frac{1}{(n-1)!}$$

$$\sum_{i=2}\sum_{j=2}\frac{Y_j}{n-j+1} = \frac{Y_2}{n-1} + \frac{Y_2}{n-1} + \frac{Y_3}{n-2} + \frac{Y_2}{n-1} + \frac{Y_3}{n-2} + \frac{Y_4}{n-3} \qquad (n-1)\frac{Y_2}{(n-1)} + \cdots + 1\frac{Y_n}{Y_n} = \sum_{i=2}^{n} Y_i$$

Then $f(Y_1, \dots, Y_n)$ is equal to:

$$\frac{n!}{(n-1)!}\cdot\frac{1}{\lambda^n} I[Y_1 \geq \theta]\cdot \exp\left\{-\frac{1}{\lambda}\sum_{i=2}^{n}\left(Y_1 + \sum_{j=2}^{i}\frac{Y_j}{n-j+1}\right) + \frac{n\theta}{\lambda}\right\}$$

$$= n\cdot\lambda^{-n} I[Y_1 \geq \theta]\cdot \exp\left\{\frac{-n}{\lambda}Y_1 + \frac{n}{\lambda}\theta - \frac{1}{\lambda}\sum_{i=2}^{n}\sum_{j=2}^{i}\frac{Y_j}{n-j+1}\right\}$$

$$= \underbrace{\frac{n}{\lambda}\cdot \exp\left\{\frac{-n}{\lambda}(Y_1-\theta)\right\} I[Y_1 \geq \theta]}_{Y_1 \sim E[\theta, \frac{\lambda}{n}]}\cdot \frac{1}{\lambda^{n-1}}\exp\left\{\frac{-1}{\lambda}\sum_{i=2}^{n}\sum_{j=2}^{i}\frac{Y_j}{n-j+1}\right\}$$

$$\underbrace{\frac{1}{\lambda^{n-1}}\exp\left\{\frac{-1}{\lambda}\left(Y_2+Y_3+\cdots+Y_n\right)\right\}}_{\Rightarrow Y_i,\ i=2,\dots,n \ \sim \text{Exponential (mean} = \lambda)}$$

Note that $\sum_{i=2}^{n} Y_i = \sum_{i=2}^{n}(n-i+1)(X_{(i)} - X_{(i-1)})$

$= (n-1)(X_{(2)} - X_{(1)}) + (n-2)(X_{(3)} - X_{(2)}) + (n-3)(X_{(4)} - X_{(3)})$

$= -(n-1)X_{(1)} + X_{(2)} + X_{(3)} + \cdots + X_{(n)}$

$= \sum_{i=1}^{n}(X_i - X_{(1)})$

And that $Y_1 \sim E[\theta, \frac{\lambda}{n}] \perp \sum_{i=2}^{n} Y_i = \sum_{i=1}^{n}(X_i - X_{(1)}) \sim \text{Gamma}(n-1, \lambda)$

as $Y_2, \dots, Y_n \overset{iid}{\sim}$ Exponential (mean $=\lambda$)

(d) θ, λ unknown. Derive a 95% joint confidence region for (θ, λ).

KEY POINT: FIND A PIVOTAL QUANTITY INVOLVING (θ, λ).

$X_{(1)} \sim E[\theta, \frac{\lambda}{n}] \Rightarrow f(X_{(1)}) = \frac{n}{\lambda}\exp\{\frac{-n}{\lambda}(X_{(1)} - \theta)\} I[X_{(1)} \geq \theta]$

$\Rightarrow P(X_{(1)} \leq t) = \int_{\theta}^{t}\frac{n}{\lambda}\exp\{\frac{-n}{\lambda}(x-\theta)\}\,dx$

$= \left[-\exp\{-\frac{n}{\lambda}(x-\theta)\}\right]_{\theta}^{t} = 1 - \exp\left[-\frac{n}{\lambda}(t-\theta)\right]$

Consider $e^{\frac{1}{\lambda}(X_{(1)}-\theta)}$

$P[e^{\frac{1}{\lambda}(X_{(1)}-\theta)} < t] = P[\frac{1}{\lambda}(X_{(1)}-\theta) < \log t]$

$= P[X_{(1)} < \theta + \lambda \log t] = 1 - \exp\left[-\frac{n}{\lambda}\left(\cancel{\theta - \theta} + \lambda\log t\right)\right]$

$= 1 - \exp[-n\log t]$

$= 1 - \exp[\log(t^{-n})] = 1 - t^{-n}$

Thus, $e^{\frac{1}{\lambda}(X_{(1)}-\theta)}$ is a pivotal quantity.

A 95% joint confidence region for (θ, λ) is:

$\{(\theta, \lambda) : a^* < e^{\frac{1}{\lambda}(X_{(1)}-\theta)} < b^*\}$, where a^*, b^* such that

$1 - (a^*)^{-n} = 0.025 \Rightarrow 0.975 = (a^*)^{-n} \Rightarrow a^* = (0.975)^{-\frac{1}{n}} \qquad \frac{1}{0.975^{\frac{1}{n}}}$

$1 - (b^*)^{-n} = 0.975 \Rightarrow b^* = (0.025)^{-\frac{1}{n}} \qquad \frac{1}{0.025^{\frac{1}{n}}}$

(e) λ known and θ unknown. Show that power function of UMP

test of $H_0: \theta \geq \theta_0$ against $H_1: \theta < \theta_0$ is given by:

$\pi(\theta) = 1 - (1-\alpha)\exp\left\{-\frac{n}{\lambda}(\theta_0 - \theta)\right\}$

Know that $X_{(1)}$ is sufficient for θ. thus, the UMP test, of $H_0: \theta \geq \theta_0$ against $H_1: \theta < \theta_0$ has the form

$$\phi(X_{(1)}) : \begin{cases} 1 & \text{if } X_{(1)} < c_\alpha \\ 0 & \text{otherwise} \end{cases}$$

where c_α chosen so that $Pr_{H_0}[X_{(1)} < c_\alpha] = \alpha$

$$\Rightarrow 1 - \exp\left[-\frac{n}{\lambda}(c_\alpha - \theta_0)\right] = \alpha$$

$$\Rightarrow \log(1-\alpha) = -\frac{n}{\lambda}(c_\alpha - \theta_0)$$

$$\Rightarrow c_\alpha = \theta_0 - \frac{\lambda}{n}\log(1-\alpha)$$

then, the power function is

$$\begin{aligned} Pr_\theta[X_{(1)} < c_\alpha] &= 1 - \exp\left[-\frac{n}{\lambda}(c_\alpha - \theta)\right] \\ &= 1 - \exp\left[-\frac{n}{\lambda}c_\alpha\right]\exp\left[\frac{n}{\lambda}\theta\right] \\ &= 1 - \exp\left[-\frac{n}{\lambda}\left(\theta_0 - \frac{\lambda}{n}\log(1-\alpha)\right)\right]\exp\left[\frac{n}{\lambda}\theta\right] \\ &= 1 - \exp\left[-\frac{n}{\lambda}\theta_0 + \log(1-\alpha) + \frac{n}{\lambda}\theta\right] \\ &= 1 - (1-\alpha)\exp\left[-\frac{n}{\lambda}(\theta_0 - \theta)\right] \end{aligned}$$

(f) Derive the joint asymptotic distribution of $\left[X_{(1)}, \sum_{i=1}^{n}(X_i - X_{(1)})\right]$, properly normalized

From (c), $X_{(1)} \sim E\left[\theta, \frac{\lambda}{n}\right] \perp \underbrace{\sum_{i=1}^{n}(X_i - X_{(1)})}_{\text{sum of } (n-1) \text{ iid } Exp(\lambda) \text{ random variables}} \sim Gamma(n-1, \lambda)$

$\rightarrow$ by CLT, $\frac{1}{\sqrt{n}}\sum_{i=2}^{n}(Y_i - \lambda) \xrightarrow{d} N(0, \lambda^2)$

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(X_i - X_{(1)}) = \frac{1}{\sqrt{n}}\sum_{i=2}^{n} Y_i - \lambda + \lambda$$

$$= \frac{1}{\sqrt{n}}\sum_{i=2}^{n}(Y_i - \lambda) + \underbrace{\frac{\lambda}{\sqrt{n}}}_{\xrightarrow{p} 0} \xrightarrow{d} N(0, \lambda^2)$$

by slutsky

$$\begin{aligned} Pr[n(X_{(1)} - \theta) < t] &= Pr\left[X_{(1)} < \frac{t}{n} + \theta\right] \\ &= 1 - \exp\left\{-\frac{n}{\lambda}\left(\frac{t}{n} + \theta - \theta\right)\right\} \\ &= 1 - \exp\left\{-\frac{1}{\lambda}t\right\} \end{aligned}$$

so that $n(X_{(1)} - \theta) \sim \text{Exponential}(\text{mean} = \lambda)$

$$\sqrt{n}\begin{pmatrix} \sqrt{n}X_{(1)} - \frac{1}{\sqrt{n}}\theta \\ \frac{1}{n}\sum(X_i - X_{(1)}) - \lambda \end{pmatrix} \xrightarrow{d} \begin{pmatrix} N(0, \lambda) \\ Exp(\text{mean} = \lambda) \end{pmatrix}$$

2007 QUESTION 1

(1) $z_1, \dots, z_n$ constants with at least one $z_i > 0$, at least one $z_i < 0$. X_i, $i = 1, \dots, n$ $\sim$ Poisson $(\lambda_0 e^{\beta z_i})$, independent.

(a) $$f(X_i = x_i) = \frac{(\lambda_0 e^{\beta z_i})^{x_i}}{x_i!} \exp\{-\lambda_0 e^{\beta z_i}\} = \exp\{\beta z_i x_i - \lambda_0 e^{\beta z_i} + \underbrace{x_i \log\lambda_0 - \log x_i!}_{h(x_i)}\}$$

$\to f(\underset{\sim}{X})$, the joint distribution, is equal to

$$\prod_{i=1}^{n} \exp\{\beta z_i x_i - \lambda_0 e^{\beta z_i} + x_i \log\lambda_0 - \log x_i!\}$$

$$= \exp\{\underset{\Theta}{\beta} \underset{T}{\sum_{i=1}^{n} z_i x_i} - \underset{B(\Theta)}{\lambda_0 \sum_{i=1}^{n} e^{\beta z_i}} + \sum_{i=1}^{n} h(x_i)\}$$

A complete sufficient statistic for the canonical parameter β is $\sum_{i=1}^{n} X_i z_i$

(b) $\ell(\beta) \propto \beta \sum_{i=1}^{n} z_i x_i - \lambda_0 \sum_{i=1}^{n} e^{\beta z_i}$

$\dot\ell(\beta) = \sum_{i=1}^{n} z_i x_i - \lambda_0 \sum_{i=1}^{n} z_i e^{\beta z_i} = 0 \Rightarrow \sum_{i=1}^{n} z_i e^{\hat\beta z_i} = \underbrace{\frac{1}{\lambda_0} \sum_{i=1}^{n} z_i x_i}_{\text{call this } C^*,\ \text{a horizontal line in } \mathbb{R}^2}$

Must show that $f(\beta) = \sum_{i=1}^{n} z_i e^{\beta z_i} = C^*$ for some $\beta = \hat\beta$.

conditions z_i → When $\beta = \infty$, $f(\beta) = \infty$. When $\beta = -\infty$, $f(\beta) = -\infty$ also. Thus $f(\beta)$ is monotonically increasing as a function of β, and must equal C^* when $\beta = \hat\beta$ (existence of MLE)

$\ddot\ell(\beta) = -\lambda_0 \sum_{i=1}^{n} z_i^2 e^{\beta z_i} < 0$ for all β (concave down) implying that the MLE $\hat\beta$ is unique.

(c) Can solve for $\hat\beta$ using Newton-Raphson.

$$\hat\beta^{(t+1)} = \hat\beta^{(t)} + [-\ddot\ell(\beta)]^{-1} \dot\ell(\beta)\,|_{\beta = \hat\beta^{(t)}}$$

$$= \hat\beta^{(t)} + \frac{\left[\sum_{i=1}^{n} z_i x_i - \lambda_0 \sum_{i=1}^{n} z_i e^{\hat\beta^{(t)} z_i}\right]}{\sum_{i=1}^{n} z_i^2 e^{\hat\beta^{(t)} z_i}}$$

(d) Assume $m = pn$ of the n z_i's are 1, and $(n-m) = (1-p)n$ of the n z_i's are -1. Show that the MLE of β is strongly consistent as $n \to \infty$.

Know that $\sum_{i=1}^{n} z_i e^{\hat\beta z_i} = \frac{1}{\lambda_0}\left[\sum_{i=1}^{n} z_i x_i\right]$ with condition on z_i

$\Rightarrow \sum_{i=1}^{m} e^{\hat\beta} - \sum_{i=m+1}^{n} e^{-\hat\beta} = \frac{1}{\lambda_0}\left[\sum_{i=1}^{m} A_i - \sum_{i=m+1}^{n} B_i\right]$ where $A_i \sim$ Poisson$(\lambda_0 e^{\beta})$

$$\Rightarrow np e^{\hat\beta} - n(1-p) e^{-\hat\beta} = \frac{1}{\lambda_0}\left[\sum_{i=1}^{np} A_i - \sum_{i=1}^{n(1-p)} B_i\right]$$

$$\Rightarrow p e^{\hat\beta} - (1-p) e^{-\hat\beta} = \frac{1}{\lambda_0}\underbrace{\left[p \cdot \frac{1}{np}\sum_{i=1}^{np} A_i - (1-p)\cdot \frac{1}{n(1-p)}\sum_{i=1}^{n(1-p)} B_i\right]}$$

$$\xrightarrow{a.s.} \frac{1}{\lambda_0}\left[p \cdot E(A_i) - (1-p) E(B_i)\right] = p e^{\beta} - (1-p) e^{-\beta} \quad \text{By SLLN + Slutsky}$$

As $pe^{\hat\beta} - (1-p)e^{-\hat\beta} \xrightarrow{a.s.} pe^{\beta} - (1-p)e^{-\beta}$, can use continuous mapping theorem to conclude $\hat\beta \xrightarrow{a.s.} \beta$.

(e) $\sqrt{n}(\hat\beta - \beta) \rightarrow_d N(0, I^{-1}(\beta))$

Recall $-\ddot{\ell}_1(\beta) = \lambda_0 z_i^2 e^{\beta z_i}$. Need to find $I(\beta) = E[-\ddot{\ell}_1(\beta)]$

$E[-\ddot{\ell}_1(\beta)]$ can be approximated by $\lim_{n\to\infty} \frac{1}{n}[-\ddot{\ell}_n(\beta)] = \lim_{n\to\infty} \frac{\lambda_0}{n}\left[\sum_{i=1}^{n} z_i^2 e^{\beta z_i}\right]$

$$= \lim_{n\to\infty} \frac{\lambda_0}{n}\left[np e^{\beta} + n(1-p) e^{-\beta}\right]$$

$$= \lambda_0 \left[p e^{\beta} + (1-p) e^{-\beta}\right]$$

$$\Rightarrow I(\beta) = \lambda_0\left[pe^{\beta} + (1-p)e^{-\beta}\right] \Rightarrow I^{-1}(\beta) = \frac{1}{\lambda_0\left[pe^{\beta} + (1-p)e^{-\beta}\right]}$$

(f) Test $H_0: \beta = 0$ versus $H_1: \beta > 0$. From exponential family results, $\sum_{i=1}^{n} z_i x_i = T$ is a complete sufficient statistic for β, and the UMP test is

$$\phi(T) = \begin{cases} 1 & \text{if } T > k_\alpha \\ \gamma & \text{if } T = k_\alpha \\ 0 & \text{otherwise} \end{cases}$$

where 1 indicates reject H_0, 0 accept, and γ reject w.p. γ.

γ, k_α chosen so that $E_{H_0}[\phi(T)] = \alpha$.

To find k_α satisfying asymptotic type I error α, find asymptotic distribution of $T = \sum_{i=1}^{n} z_i x_i$ under conditions in (d)

$$\sum_{i=1}^{np} A_i - \sum_{i=1}^{n(1-p)} B_i, \quad A_i \sim \text{Poisson}(\lambda_0 e^{\beta}), \; B_i \sim \text{Poisson}(\lambda_0 e^{-\beta})$$

Note that $\frac{1}{\sqrt{np}}\sum_{i=1}^{np}[A_i - E(A_i)] \xrightarrow{d} N(0, Var(A_i)) \Rightarrow \frac{1}{\sqrt{n}}\sum_{i=1}^{np}[A_i - E(A_i)] \xrightarrow{d} N(0, p \cdot Var(A_i))$

$$\frac{1}{\sqrt{n(1-p)}}\sum_{i=1}^{n(1-p)}[B_i - E(B_i)] \xrightarrow{d} N(0, Var(B_i)) \Rightarrow \frac{1}{\sqrt{n}}\sum_{i=1}^{n(1-p)}[B_i - E(B_i)] \xrightarrow{d} N(0, (1-p)Var(B_i))$$

Then, $\frac{1}{\sqrt{n}}\left(\sum_{i=1}^{np}[A_i - E(A_i)] - \sum_{i=1}^{n(1-p)}[B_i - E(B_i)]\right) \xrightarrow{d} N\Big(0, \underbrace{p\,Var(A_i) + (1-p)\,Var(B_i)}_{\sigma^2}\Big)$

$\Rightarrow \frac{1}{\sqrt{n}}\Big[\underbrace{\sum_{i=1}^{np} A_i - \sum_{i=1}^{n(1-p)} B_i}_{T} - \underbrace{\left(np\,E(A_i) - n(1-p)\,E(B_i)\right)}_{\mu}\Big] \xrightarrow{d} N(0, \sigma^2)$

$\Rightarrow \frac{1}{\sqrt{n}}(T-\mu) \xrightarrow{d} N(0, \sigma^2)$

Want to choose k_α so that under $H_0: \beta = 0 \Rightarrow E(A_i) = E(B_i) = Var(A_i) = Var(B_i) = \lambda_0$

$\Rightarrow \sigma^2 = \lambda_0, \quad \Rightarrow \mu = n(2p-1)\lambda_0$

$E[\phi(T)] = P(T > k_\alpha) = \alpha$

$= P\left[\frac{1}{\sqrt{n}}\frac{(T-\mu)}{\sigma} > \frac{1}{\sqrt{n}}\frac{(k_\alpha - \mu)}{\sigma}\right] = \alpha$

$\Rightarrow 1 - \Phi\left[\frac{1}{\sqrt{n}}\frac{(k_\alpha - \mu)}{\sigma}\right] = \alpha$

$\Rightarrow \Phi^{-1}(1-\alpha) = \frac{1}{\sqrt{n}}\frac{(k_\alpha - \mu)}{\sigma} \quad \Rightarrow k_\alpha = \mu + \sigma\sqrt{n}\,\Phi^{-1}(1-\alpha)$

$= n(2p-1)\lambda_0 + \sqrt{n\lambda_0}\,\Phi^{-1}(1-\alpha)$

2007 QUESTION 3

(3.) $Y = X_1\beta_1 + X_2\beta_2 + \varepsilon$, $\varepsilon \sim N_n(0, \sigma^2 I)$, $X = (X_1, X_2)$ of rank $p = p_1 + p_2$

σ^2 unknown

$H_0: \beta_2 = 0 \Rightarrow Y = X_1\beta_1 + \varepsilon$ (REDUCED)

$H_1: \beta_2 \neq 0 \Rightarrow Y = X\beta + \varepsilon$ $X = (X_1, X_2)$, $\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$ (FULL)

(a) let $M = X(X'X)^- X'$, $M_0 = X_1(X_1'X_1)^{-1}X_1'$

$$F = \frac{\left[\dfrac{Y'(M-M_0)Y}{r(M-M_0)}\right]}{\left[\dfrac{Y'(I-M)Y}{r(I-M)}\right]} \begin{array}{l} \sim F_{r(M-M_0),\, r(I-M)} \text{ under } H_0 \\ \sim F_{r(M-M_0),\, r(I-M),\ \delta = \frac{(X\beta)'(M-M_0)(X\beta)}{2\sigma^2}} \text{ under } H_1 \end{array}$$

In the numerator, note that $W = \dfrac{Y'(M-M_0)Y}{\sigma^2} \sim \chi^2_{r(M-M_0),\ \delta = \frac{(X\beta)'(M-M_0)(X\beta)}{2\sigma^2}}$

$$\Rightarrow \boxed{E(W) = r(M-M_0) + 2\delta}$$

then $E\left[\dfrac{\sigma^2}{r(M-M_0)} W\right] = \dfrac{\sigma^2}{r(M-M_0)} E(W) = \sigma^2 + \dfrac{\sigma^2 \cdot \delta \cdot 2}{r(M-M_0)}$

$$= \sigma^2 + \frac{(X\beta)'(M-M_0)(X\beta)}{r(M-M_0)}$$

(b) X_2 fixed so $r(M-M_0) = p - p_1$ fixed, the expected mean square from (a) is proportional to

$$\beta'X'(M-M_0)X\beta = \beta'X'X\beta - \beta'X'M_0X\beta$$
$$= \beta'[X'X - X'M_0X]\beta$$

$X'X = \begin{bmatrix} X_1' \\ X_2' \end{bmatrix}[X_1 \ X_2] = \begin{bmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{bmatrix}$, $X'M_0X = \begin{bmatrix} X_1'M_0X_1 & X_1'M_0X_2 \\ X_2'M_0X_1 & X_2'M_0X_2 \end{bmatrix}$

$$= \begin{bmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'M_0X_2 \end{bmatrix}$$

$$\Rightarrow \beta'[X'X - X'M_0X]\beta = \beta'\left\{\begin{bmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{bmatrix} - \begin{bmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'M_0X_2 \end{bmatrix}\right\}\beta$$

$$\underbrace{\beta_2'X_2'X_2\beta_2}_{\text{constant}} - \underbrace{\beta_2'X_2'M_0X_2\beta_2}_{\text{have freedom to manipulate, as it involves } M_0 = X_1(X_1'X_1)^- X_1'}$$

second term is always ≥ 0, when $X_1 \perp X_2$, $X_1'X_2 = 0$ and $\beta_2'X_2'M_0X_2\beta_2 = 0$.

thus $X_1 \perp X_2$ maximizes the expected mean square

(c) X_2 is not null but $\text{rank}(X) = \text{rank}(X_1) = p_1 \Rightarrow X_2 = X_1 C$

Then, $X'X = \begin{bmatrix} X_1' \\ C'X_1' \end{bmatrix} [X_1 \;\; X_1 C] = \begin{bmatrix} X_1'X_1 & X_1'X_1 C \\ C'X_1'X_1 & C'X_1'X_1 C \end{bmatrix}$, $X'Y = \begin{bmatrix} X_1' \\ C'X_1' \end{bmatrix}$

and with $\hat{\beta} = \begin{bmatrix} (X_1'X_1)^{-1} X_1' \\ 0 \end{bmatrix} Y$,

$$(X'X)\hat{\beta} = \begin{bmatrix} X_1'X_1 & X_1'X_1 C \\ C'X_1'X_1 & C'X_1'X_1 C \end{bmatrix} \begin{bmatrix} (X_1'X_1)^{-1}X_1' \\ 0 \end{bmatrix} Y = \begin{bmatrix} X_1' \\ C'X_1' \end{bmatrix} Y = X'Y$$

(d) X_2 is a null matrix, so that $Y \sim N_n(X_1\beta_1, \sigma^2 I_n)$

(i) $\Rightarrow \ell(\beta_1, \sigma^2) \propto -\frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}(Y - X_1\beta_1)'(Y - X_1\beta_1)$ let $\psi = \frac{\beta_1}{\sigma^2} \Rightarrow \beta_1 = \sigma^2\psi$

$$\Rightarrow \ell(\sigma^2, \psi) \propto -\frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}(Y - X_1\psi\sigma^2)'(Y - X_1\psi\sigma^2)$$
$$= -\frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\left[Y'Y - 2Y'X_1\psi\sigma^2 + \sigma^4\psi'X_1'X_1\psi\right]$$
$$= -\frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}Y'Y + Y'X_1\psi - \frac{\sigma^2}{2}\psi'X_1'X_1\psi$$

$$\Rightarrow \frac{\partial \ell}{\partial \sigma^2} = \frac{-n}{2\sigma^2} + \frac{1}{2\sigma^4}Y'Y - \frac{1}{2}\psi'X_1'X_1\psi = 0$$
$$\Rightarrow \frac{-n}{\sigma^2} + \frac{1}{\sigma^4}Y'Y - \psi'X_1'X_1\psi = 0$$
$$\Rightarrow \frac{-n}{\sigma^2} + \frac{1}{\sigma^4}Y'Y - \frac{1}{\sigma^4}\underbrace{Y'X_1(X_1'X_1)^{-1}X_1'}_{M}Y = 0 \Rightarrow n\sigma^2 = Y'(I-M)Y \Rightarrow \hat{\sigma}^2 = \frac{Y'(I-M)Y}{n}$$

$$\Rightarrow \frac{\partial \ell}{\partial \psi} = X_1'Y - \frac{\sigma^2}{2} 2(X_1'X)\psi = 0 \Rightarrow X_1'Y = \sigma^2(X_1'X)\psi$$
$$\Rightarrow \hat{\psi} = \frac{(X_1'X_1)^{-1}X_1'Y}{\hat{\sigma}^2} = \frac{(X_1'X_1)^{-1}X_1'Y}{\left(\frac{Y'(I-M)Y}{n}\right)}$$

(ii) $\hat{\psi}$ is a ratio of two independent random variables, $\hat{\beta}_1 = (X_1'X_1)^{-1}X_1'Y$ and $\hat{\sigma}^2 = \frac{1}{n}Y'(I-M)Y$. They are independent as $(I-M)[X_1(X_1'X_1)^{-1}] = 0$.

Thus, $E(\hat{\psi}) = E(\hat{\beta}_1) \cdot E\left(\frac{1}{\hat{\sigma}^2}\right) = \beta_1 \cdot E\left(\frac{1}{\hat{\sigma}^2}\right)$

$\hat{\sigma}^2 = \frac{1}{n}Y'(I-M)Y = \frac{\sigma^2}{n}W$, where $W \sim \chi^2_{r(I-M)=n-p}$

$\Rightarrow E\left(\frac{1}{\hat{\sigma}^2}\right) = \frac{n}{\sigma^2}E\left(\frac{1}{W}\right)$, where $W \sim \text{Gamma}\left(\frac{n-p}{2}, 2\right)$

$$E(W^{-1}) = \int_0^\infty w^{-1}\frac{w^{\left(\frac{n-p}{2}-1\right)}e^{(-w/2)}}{\Gamma\left(\frac{n-p}{2}\right)2^{\left(\frac{n-p}{2}\right)}} = \int_0^\infty \frac{w^{\left(\frac{n-p}{2}-1-1\right)}e^{(-w/2)}}{\Gamma\left(\frac{n-p}{2}\right)2^{\left(\frac{n-p}{2}\right)}}\,dw = \frac{\Gamma\left(\frac{n-p}{2}-1\right)2^{\left(\frac{n-p}{2}-1\right)}}{\Gamma\left(\frac{n-p}{2}\right)2^{\left(\frac{n-p}{2}\right)}}$$
$$= \frac{1}{\frac{n-p}{2}-1}\cdot\frac{1}{2} = \frac{1}{n-p-2}$$

Then $E(\hat{\psi}) = \beta_1 \cdot \frac{n}{\sigma^2}\left(\frac{1}{n-p-2}\right) = \frac{n}{n-p-2}\cdot\frac{\beta_1}{\sigma^2} = \frac{n}{n-p-2}\psi$

(iii) $\lim_{n\to\infty} E(\hat{\psi}) = \lim_{n\to\infty}\frac{n}{n-p-2}\psi = \psi$

2007 QUESTION 4

$Y = (Y_0, Y_1, Y_2)' \sim$ multinomial $(m, (\gamma_0, \gamma_1, \gamma_2)')$ where

$$\gamma_0 = (1-\pi)^2 \,/\, [(1-\pi)^2 + 2\pi(1-\pi)\theta^{-1} + \pi^2] = \frac{a}{a+b+c}$$

$$\gamma_1 = 2\pi(1-\pi)\theta^{-1} \,/\, [(1-\pi)^2 + 2\pi(1-\pi)\theta^{-1} + \pi^2] = \frac{b}{a+b+c}$$

$$\gamma_2 = \pi^2 \,/\, [(1-\pi)^2 + 2\pi(1-\pi)\theta^{-1} + \pi^2] = \frac{c}{a+b+c}$$

$$\Rightarrow f(\underset{\sim}{Y} | \underset{\sim}{\gamma}) = \underbrace{\frac{m!}{y_0!\,y_1!\,y_2!}}_{h(y)} \left(\frac{a}{a+b+c}\right)^{y_0} \left(\frac{b}{a+b+c}\right)^{y_1} \left(\frac{c}{a+b+c}\right)^{n-y_0-y_1}$$

$$= h(y) \left(\frac{a}{c}\right)^{y_0} \left(\frac{b}{c}\right)^{y_1} \left(\frac{c}{a+b+c}\right)^{n}$$

Now, note that $\frac{a}{c} = \left(\frac{1-\pi}{\pi}\right)^2 = \exp\left\{2\log\left(\frac{1-\pi}{\pi}\right)\right\} = \exp\{-2\lambda\}$

$\frac{b}{c} = \frac{2\pi(1-\pi)}{\theta \cdot \pi^2} = 2\left(\frac{1-\pi}{\pi}\right)\theta^{-1} = 2\exp\{-\lambda-\psi\}$

and that $\left(\frac{c}{a+b+c}\right)^n = \left(1 + \frac{a}{c} + \frac{b}{c}\right)^{-n} = \left(1 + \exp\{-2\lambda\} + 2\exp\{-\lambda-\psi\}\right)^{-m}$

$= -\log h(y)$ $\Rightarrow f(\underset{\sim}{Y} | \underset{\sim}{\gamma}) = \exp\left\{ y_0(-2\lambda) + y_1(-\lambda-\psi) - m\log[1+\exp(-2\lambda)+2\exp(-\lambda-\psi)] \underbrace{- c(y) - y_1\log 2}_{c(y)} \right\}$

$$= \exp\left\{(-2y_0 - y_1)\lambda + (-y_1)\psi - m\log[1+\exp(-2\lambda)+2\exp(-\lambda-\psi)] - c(y)\right\}$$

Note that $y_1 = m - y_0 - y_2$

(i) $= \exp\left\{ \underbrace{(y_2-y_0)\lambda + (-y_1)\psi} - \overbrace{m[\log[1+\exp(-2\lambda)+2\exp(-\lambda-\psi)] + \lambda]}^{b(\eta)} - c(y) \right\}$

(ii) $Q(y)'\eta$, where $Q(y) = [y_2 - y_0, -y_1]'$, $\eta = [\lambda \;\; \psi]'$

(iii) The log-likelihood is equal to

$$\ell(\lambda, \psi) = (y_2-y_0)\lambda + (-y_1)\psi - m\{\log[1+\exp(-2\lambda)+2\exp(-\lambda-\psi)] + \lambda\} - c(y)$$

$$= (y_2-y_0)\lambda + (-y_1)\psi - m\{\log[(1+\exp(-2\lambda)+2\exp(-\lambda-\psi))\cdot\exp(\lambda)]\}$$

$$(y_2-y_0)\lambda + (-y_1)\psi - m\{\log[e^{\lambda} + e^{-\lambda} + 2e^{-\psi}]\} - c(y)$$

(b) $\frac{\partial \ell}{\partial \lambda} = (y_2 - y_0) - m\left[\frac{e^{\lambda} - e^{-\lambda}}{e^{\lambda} + e^{-\lambda} + 2e^{-\psi}}\right] = 0$ $\Rightarrow \frac{e^{\lambda} - e^{-\lambda}}{e^{\lambda} + e^{-\lambda} + 2e^{-\psi}} = \frac{y_2 - y_0}{m}$

$\frac{\partial \ell}{\partial \psi} = (-y_1) - m\left[\frac{-2e^{-\psi}}{e^{\lambda} + e^{-\lambda} + 2e^{-\psi}}\right] = 0$ $\Rightarrow \frac{-2e^{-\psi}}{e^{\lambda} + e^{-\lambda} + 2e^{-\psi}} = \frac{-y_1}{m} = \frac{y_2 + y_0}{m} - 1$

$\Rightarrow \frac{y_2 + y_0}{m} = \frac{e^{\lambda} + e^{-\lambda}}{e^{\lambda} + e^{-\lambda} + 2e^{-\psi}}$ $\Rightarrow \frac{y_2 + y_0}{e^{\lambda} + e^{-\lambda}} = \frac{y_2 - y_0}{e^{\lambda} - e^{-\lambda}}$

$\Rightarrow \frac{y_2 + y_0}{y_2 - y_0} = \frac{e^{2\lambda} + 1}{e^{2\lambda} - 1}$ $\Rightarrow \frac{\frac{y_2}{y_0} + 1}{\frac{y_2}{y_0} - 1} = \frac{e^{2\lambda} + 1}{e^{2\lambda} - 1}$

$\Rightarrow e^{2\lambda} = \frac{y_2}{y_0}$, $\hat{\lambda} = \frac{1}{2}\log\left(\frac{y_2}{y_0}\right)$

$$\hat{\lambda} = \frac{1}{2}\log\left(\frac{y_2}{y_0}\right) = \log\left[\left(\frac{y_2}{y_0}\right)^{\frac{1}{2}}\right]$$

$$-\hat{\lambda} = -\frac{1}{2}\log\left(\frac{y_2}{y_0}\right) = \log\left[\left(\frac{y_2}{y_0}\right)^{-\frac{1}{2}}\right] = \log$$

Also, $\dfrac{y_2 - y_0}{e^{\lambda} - e^{-\lambda}} = \dfrac{y_1}{2e^{-\psi}}$

$$\Rightarrow 2e^{-\psi} = \frac{y_1}{y_2 - y_0}\, e^{\lambda} - e^{-\lambda} = \frac{y_1}{y_2 - y_0}\left\{\left(\frac{y_2}{y_0}\right)^{\frac{1}{2}} - \left(\frac{y_0}{y_2}\right)^{\frac{1}{2}}\right\}$$

$$\Rightarrow 2e^{-\psi} = \frac{y_1}{y_2 - y_0}\left\{\left(\frac{y_0}{y_2}\right)^{\frac{1}{2}}\left(\frac{y_2}{y_0} - 1\right)\right\}$$

$$\Rightarrow 2e^{-\psi} = \frac{y_1}{y_0}\left(\frac{y_0}{y_2}\right)^{\frac{1}{2}}$$

$$\Rightarrow e^{-\psi} = \frac{y_1}{2y_0}\left(\frac{y_0}{y_2}\right)^{\frac{1}{2}}$$

$$\Rightarrow -\psi = \log\left[\frac{y_1}{2y_0}\left(\frac{y_0}{y_2}\right)^{\frac{1}{2}}\right] \Rightarrow \hat{\psi} = \log\left[\left(\frac{y_1}{2y_0}\right)^{-1}\left(\frac{y_2}{y_0}\right)^{\frac{1}{2}}\right]$$

$$\hat{\psi} = \frac{1}{2}\log\left(\frac{y_2}{y_0}\right) - \log\left(\frac{y_1}{2y_0}\right)$$

$$= \hat{\lambda} - \psi_c$$

(c) The asymptotic covariance matrix of $(\hat{\lambda}, \hat{\psi})$ is $I^{-1}(\lambda, \psi)$, where $I(\lambda,\psi) = E\left[-\frac{\partial^2 \ell}{\partial \eta}\right]$, $\eta = (\lambda, \psi)$

$$e^{\lambda} \cdot 1 + 1 + \cancel{e^{-2\lambda}} + 2e^{-\psi-\lambda} + 2e^{-\psi+\lambda} \quad (\cancel{e^{2\lambda}} - 2 + \cancel{e^{-2\lambda}})$$

$$\frac{\partial^2 \ell}{\partial \lambda^2} = \frac{-1}{(e^{\lambda}+e^{-\lambda}+2e^{-\psi})^2}\left\{(e^{\lambda}+e^{-\lambda}+2e^{-\psi})(e^{\lambda}+e^{-\lambda}) - (e^{\lambda}-e^{-\lambda})^2\right\}$$

$$\left.\begin{aligned}\frac{\partial^2 \ell}{\partial\lambda\,\partial\psi} &= \frac{-(e^{\lambda}-e^{-\lambda})}{(e^{\lambda}+e^{-\lambda}+2e^{-\psi})^2}\left\{-2e^{-\psi}\right\}\\ \frac{\partial^2 \ell}{\partial\psi\,\partial\lambda} &= (2e^{-\psi})\cdot\frac{e^{\lambda}-e^{-\lambda}}{-(e^{\lambda}+e^{-\lambda}+2e^{-\psi})^2}\end{aligned}\right\} = \frac{-1}{(e^{\lambda}+e^{-\lambda}+2e^{-\psi})^2}\left\{2e^{-\psi}(e^{\lambda}-e^{-\lambda})\right\}$$

$$\frac{\partial^2 \ell}{\partial \psi^2} = \frac{-1}{(e^{\lambda}+e^{-\lambda}+2e^{-\psi})^2}\left\{(e^{\lambda}+e^{-\lambda}+2e^{-\psi})(2e^{-\psi}) - (2e^{-\psi})^2\right\}$$

$$\Rightarrow I(\lambda,\psi) = \frac{1}{(e^{\lambda}+e^{-\lambda}+2e^{-\psi})^2}\begin{bmatrix} 4 + 2e^{\lambda-\psi} + 2e^{-\lambda-\psi} & 2e^{\lambda-\psi} - 2e^{-\lambda-\psi} \\ / & 2e^{\lambda-\psi} + 2e^{-\lambda-\psi}\end{bmatrix}$$

$$\Rightarrow I^{-1}(\lambda,\psi) = \frac{(e^{\lambda}+e^{-\lambda}+2e^{-\psi})^2}{(4+2e^{\lambda-\psi}+2e^{-\lambda-\psi})(2e^{\lambda-\psi}+2e^{-\lambda-\psi}) - (2e^{\lambda-\psi}-2e^{-\lambda-\psi})^2}\begin{bmatrix} 2e^{\lambda-\psi}+2e^{-\lambda-\psi} & -2e^{\lambda-\psi}+2e^{-\lambda-\psi} \\ / & 4+2e^{\lambda-\psi}+2e^{-\lambda-\psi}\end{bmatrix}$$

Can plug in MLEs as an estimator. Note $e^{\hat{\lambda}} = \left(\frac{y_2}{y_0}\right)^{\frac{1}{2}}$, $e^{-\hat{\lambda}} = \left(\frac{y_0}{y_2}\right)^{\frac{1}{2}}$,

$e^{-\hat{\psi}} = \left(\frac{y_1}{2y_0}\right)\left(\frac{y_0}{y_2}\right)^{\frac{1}{2}}$, $e^{\hat{\lambda}-\hat{\psi}} = \frac{y_1}{2y_0}$, $e^{-\hat{\lambda}-\hat{\psi}} = \exp\left\{-2\hat{\lambda} + \log\left(\frac{y_1}{2y_0}\right)\right\} = \exp\left\{\log\left(\frac{y_0}{y_2}\cdot\frac{y_1}{2y_0}\right)\right\} = \frac{y_1}{2y_2}$

$$I^{-1}(\hat{\lambda},\hat{\psi}) = \frac{\left[\left(\frac{y_2}{y_0}\right)^{\frac{1}{2}} + \left(\frac{y_0}{y_2}\right)^{\frac{1}{2}} + \left(\frac{y_1}{y_0}\right)\left(\frac{y_0}{y_2}\right)^{\frac{1}{2}}\right]^2}{\left(4+\frac{y_1}{y_0}+\frac{y_1}{y_2}\right)\left(\frac{y_1}{y_0}+\frac{y_1}{y_2}\right) - \left(\frac{y_1}{y_0}-\frac{y_1}{y_2}\right)^2}\begin{bmatrix}\frac{y_1}{y_2}+\frac{y_1}{y_0} & \frac{y_1}{y_2}-\frac{y_1}{y_0} \\ / & 4+\frac{y_1}{y_0}+\frac{y_1}{y_2}\end{bmatrix}$$

$$\cdots \begin{bmatrix}\frac{1}{4y_2}\left(1+\frac{y_2}{y_0}\right) & \frac{1}{4y_2}\left(1-\frac{y_2}{y_0}\right) \\ / & \frac{1}{4y_2}+\frac{1}{y_1}+\frac{1}{4y_0}\end{bmatrix}$$

(d) $y_2 - y_0$ is sufficient for λ assuming $\psi = \psi_0$ is known, and y_1 is sufficient for $-\psi$. A conditional likelihood for ψ is found as follows. Recall,

$$f(y_0, y_1, y_2 \mid \lambda, \psi) = \frac{m!}{y_0! y_1! y_2!} 2^{y_1} e^{\lambda(y_2 - y_0)} e^{\psi(-y_1)} \left(\frac{1}{e^{\lambda} + e^{-\lambda} + 2e^{-\psi}}\right)^m$$

then

$$P(y_1 \mid y_2 - y_0 = t) = \frac{P(y_1, y_2 - y_0 = t)}{P(y_2 - y_0 = t)} = \frac{\frac{\cancel{m!}}{y_0! y_1! y_2!} 2^{y_1} \cancel{e^{\lambda(y_2 - y_0)}} e^{\psi(-y_1)} \cancel{\left(\frac{1}{e^{\lambda} + e^{-\lambda} + 2e^{-\psi}}\right)^m}}{\sum_{\substack{y_1 \text{ such that} \\ y_2' - y_0' = t}} \frac{\cancel{m!}}{y_0! y_1! y_2!} 2^{y_1} \cancel{e^{\lambda(y_2 - y_0)}} e^{-y_1 \psi} \cancel{\left(\frac{1}{e^{\lambda} + e^{-\lambda} + 2e^{-\psi}}\right)^m}}$$

$$= \frac{\frac{1}{y_0! y_1! y_2!} 2^{y_1} e^{-y_1 \psi}}{\sum_{\substack{y_1 \text{ such} \\ \text{that } y_2 - y_0 = t}} \frac{1}{y_0! y_1! y_2!} 2^{y_1} e^{-y_1 \psi}}$$

(e) $y_0 = 3$, $y_1 = 0$, $y_2 = 2$ were observed (i.e. $m = 5$, and $y_2 - y_0 = -1$, and $y_1 = 0$). Compute the exact one-sided p-value for testing $H_0: \theta = 1$ vs. $H_1: \theta > 1$ with λ unspecified.

$H_0: \theta = 1 \Rightarrow \psi = 0 \Rightarrow -\psi = 0$
$H_1: \theta > 1 \Rightarrow \psi > 0 \Rightarrow -\psi < 0$

$y_1 \mid y_2 - y_0 = t$ is the appropriate conditional likelihood, y_1 suff. for $-\psi$

One-sided exact p-value $= P_{H_0}(y_1 \le 0 \mid y_2 - y_0 = -1)$. Note $P_{H_0}(y_1 \mid y_2 - y_0 = -1) = \dfrac{2^{y_1} \frac{1}{y_0! y_1! y_2!}}{\sum_{y_1 \text{ s.t. } y_2 - y_0 = -1} 2^{y_1} \frac{1}{y_0! y_1! y_2!}}$

y_1	y_0	y_2	Possible?	$2^{y_1} \frac{1}{y_0! y_1! y_2!}$	cond. prob.
0	3	2	✓	$\frac{1}{6 \cdot 2} = \frac{1}{12}$	$\frac{1/12}{21/12} = 1/21$
1			x		
2	2	1	✓	$4 \cdot \frac{1}{2 \cdot 2} = 1$	$\frac{12/12}{21/12} = \frac{12}{21}$
3			x		
4	1	0	✓	$16 \cdot \frac{1}{24} = \frac{8}{12}$	$\frac{8/12}{21/12} = \frac{8}{21}$
5			x		

$\Rightarrow P(y_1 \le 0 \mid y_2 - y_0 = -1) = \frac{1}{21} \approx 0.0476$

2007

Q5 solution

~~Q7~~ Solution

(a) Let $Y = \frac{\theta X_1}{X_2}$. Also, let $Z = X_2$.

Thus

$$X_1 = \frac{yz}{\theta}$$

$$X_2 = z$$

$$f_{Y,Z}(y,z) = f_{X_1,X_2}\left(\frac{yz}{\theta}, z\right) = f_{X_1}\left(\frac{yz}{\theta}\right) f_{X_2}(z)\,|J|$$

$$= \lambda_1 \exp\left\{-\frac{\lambda_1 yz}{\frac{\lambda_1}{\lambda_2}}\right\} \lambda_2 \exp\left\{-\lambda_2 z\right\} |J|$$

Where

$$J = \det\begin{bmatrix} \frac{\partial x_1}{\partial y} & \frac{\partial x_1}{\partial z} \\ \frac{\partial x_2}{\partial y} & \frac{\partial x_2}{\partial z} \end{bmatrix} = \det\begin{bmatrix} z/\theta & y/\theta \\ 0 & 1 \end{bmatrix}$$

$$= z/\theta$$

Thus

$$f_{Y,Z}(y,z) = \lambda_1 \exp\{-\lambda_2 yz\}\ \lambda_2 \exp\{-\lambda_2 z\}\left(\frac{z\lambda_2}{\lambda_1}\right)$$

$$= z\,\lambda_2^2 \exp\{-\lambda_2 z(1+y)\}$$

Thus

$$f_Y(y) = \int_0^\infty z\lambda_2^2 e^{-\lambda_2(1+y)z}\,dz$$

$$= \lambda_2^2 \int_0^\infty z^{2-1} e^{-\lambda_2(y+1)z}\,dz$$

$$= \frac{\lambda_2^2\,\Gamma(2)}{((y+1)\lambda_2)^2} = \frac{1}{(1+y)^2}, \quad 0<y<\infty$$

Thus $F(y) = \int_0^y \frac{1}{(1+u)^2}\,du = \frac{y}{1+y}, \quad 0<y<\infty$

Since the distribution of $Y = \frac{\theta X_1}{X_2}$ does not depend on θ, it is a pivotal. A $(1-\alpha)\times 100\%$ confidence interval for θ is given by

$$a < \frac{\theta X_1}{X_2} < b$$

Where $F_Y(b) = 1-\frac{\alpha}{2}$

$F_Y(a) = \frac{\alpha}{2}$

Thus

$$F_Y(b) = \frac{b}{1+b} = 1 - \frac{\alpha}{2}$$

$$\Rightarrow b = \frac{1-\alpha/2}{1-(1-\alpha/2)} = \frac{2-\alpha}{\alpha}$$

$$F_Y(a) = \alpha/2 \Rightarrow \frac{a}{1+a} = \alpha/2$$

$$\Rightarrow a = \frac{\alpha/2}{1-\alpha/2} = \frac{\alpha}{2-\alpha}$$

Thus, the interval is

$$\frac{\alpha}{2-\alpha} < \frac{\theta X_1}{X_2} < \frac{2-\alpha}{\alpha}$$

$$\boxed{\left(\frac{\alpha}{2-\alpha}\right)\left(\frac{X_2}{X_1}\right) < \theta < \left(\frac{X_2}{X_1}\right)\left(\frac{2-\alpha}{\alpha}\right)}$$

(b)

$$Y = \lambda_1 X_1 + \lambda_2 X_2$$

$$Z = X_2$$

$$f_{Y,Z}(y,z) = f_{X_1,X_2}\left(\frac{y-\lambda_2 z}{\lambda_1}, z\right)|J|$$

$$= f_{X_1}\left(\frac{y-\lambda_2 z}{\lambda_1}\right) f_{X_2}(z)\,|J|$$

$$J = \det\begin{bmatrix} \frac{\partial x_1}{\partial y} & \frac{\partial x_1}{\partial z} \\ \frac{\partial x_2}{\partial y} & \frac{\partial x_2}{\partial z} \end{bmatrix} = \det\begin{bmatrix} \frac{1}{\lambda_1} & -\frac{\lambda_2}{\lambda_1} \\ 0 & 1 \end{bmatrix} = \frac{1}{\lambda_1}$$

$$f_{Y,Z}(y,z) = \frac{1}{\lambda_1}\left(\lambda_1 e^{-\lambda_1\left(\frac{y-\lambda_2 z}{\lambda_1}\right)}\right)\left(\lambda_2 e^{-\lambda_2 z}\right)$$

$$= \lambda_2 e^{-(y-\lambda_2 z)} e^{-\lambda_2 z}$$

$$= \begin{cases} \lambda_2 e^{-y} & \lambda_2 z < y < \infty \\ & 0 < z < \infty \\ 0 & \text{otherwise} \end{cases}$$

$$f_Y(y) = \int_0^{y/\lambda_2} \lambda_2 e^{-y}\, dz$$

$$= z\lambda_2 e^{-y}\Big|_0^{y/\lambda_2} = y e^{-y}, \quad y > 0$$

Thus $Y \sim \text{Gamma}(2,1)$, which does not depend on $\theta = (\lambda_1, \lambda_2)$, so $Y = \lambda_1 X_1 + \lambda_2 X_2$ is a pivotal quantity.

$F_Y(y) =$ cdf of Gamma$(2,1)$

Thus the joint confidence region for $\theta = (\lambda_1, \lambda_2)$ is given by

$$C(\lambda_1, \lambda_2) = \{(\lambda_1, \lambda_2) : a < \lambda_1 X_1 + \lambda_2 X_2 < b\}$$

where $b = F_Y^{-1}\left(1 - \frac{\alpha}{2}\right)$

$a = F_Y^{-1}\left(\frac{\alpha}{2}\right)$

$F_Y(y) =$ cdf of Gamma$(2,1)$.

Note:

We can actually obtain $F_Y(y)$ in closed form in this case

$$F_Y(y) = \int_0^y u e^{-u}\, du = \int u \, d(-e^{-u})$$

$$= -ue^{-u}\Big|_0^y + \int_0^y e^{-u}\, du$$

$$= 1 - e^{-y} - ye^{-y} = 1 - e^{-y}(1+y)$$

Problem 1 - c).

The joint density of X_1, X_2 is

$$p(x_1, x_2 | \lambda_1, \lambda_2) = \lambda_1 \lambda_2 \exp(-\lambda_1 x_1 - \lambda_2 x_2)\, I(x_1 > 0)\, I(x_2 > 0)$$
$$= \lambda_1 \lambda_2 \exp[(\lambda_2 - \lambda_1) x_1 - \lambda_2 (x_1 + x_2)]\, I(x_1 > 0)\, I(x_2 > 0)$$

By theorem 2.7, there exist UMPU tests of testing $H_0: \lambda_1 - \lambda_2 = 0$ vs $H_1: \lambda_1 - \lambda_2 \neq 0$

i.e. H_0 $\lambda_1 = \lambda_2$ vs $H_1: \lambda_1 \neq \lambda_2$

Here we have $U = X_1$, $T = X_1 + X_2$. $\theta = \lambda_1 - \lambda_2$, $\xi = \lambda_2$

Then the UMPU test of size α for testing $H_0: \lambda_1 = \lambda_2$ against $\lambda_1 \neq \lambda_2$ is of this form

$$\phi(u) = \begin{cases} 1 & \text{if } u \leq C_1(t) \text{ or } u \geq C_2(t) \\ 0 & C_1(t) < u < C_2(t) \end{cases}$$

with $E_0[\phi | T = t] = \alpha$

$E_0(U\phi | T = t) = \alpha E_0(U | T = t)$

$$f(x_1 | x_1 + x_2) = \frac{f(x_1, x_1 + x_2)}{f(x_1 + x_2)} \quad \text{under } H_0 \cdot \lambda_1 = \lambda_2$$

let $t = x_1 + x_2$, $u = x_1$ then $\begin{cases} x_1 = u \\ x_2 = t - u \end{cases}$ $\left| \frac{\partial(x_1, x_2)}{\partial(t, u)} \right| = 1$

$$f(u, t) = \lambda_1 e^{-\lambda_1 u} \cdot \lambda_2 e^{-\lambda_2 (t - u)} \overset{\lambda_1 = \lambda_2}{=\!=} \lambda_1^2 e^{-\lambda_1 t}$$

$$-(t) = \int_0^t f(u,t)\,du = \lambda_1^2 t e^{-\lambda_1 t}$$

$$\Rightarrow f(x_1 \mid x_1 + x_2) = f(u \mid t) = \frac{\lambda_1^2 e^{-\lambda_1 t}}{\lambda_1^2 t e^{-\lambda_1 t}} = \frac{1}{t} \quad (\text{under } H_0)$$

$$E_0[\phi \mid T = t] = \alpha \Leftrightarrow \int_0^{C_1(t)} \frac{1}{t}\,du + \int_{C_2(t)}^{t} \frac{1}{t}\,du = \alpha$$

$$\Leftrightarrow C_1(t) + t - C_2(t) = \alpha t \qquad \text{i.e.} \quad C_1(t) - C_2(t) = (\alpha - 1)t \qquad ①$$

$$E_0[U\phi \mid T = t] = \alpha E_0[U \mid T = t]$$

$$\Leftrightarrow \int_0^{C_1(t)} \frac{u}{t}\,du + \int_{C_2(t)}^{t} \frac{u}{t}\,du = \alpha \cdot \int_0^t \frac{u}{t}\,du$$

$$\Leftrightarrow \frac{1}{t}\left(\frac{C_1(t)^2}{2} + \frac{t^2}{2} - \frac{C_2(t)^2}{2}\right) = \frac{\alpha t}{2}$$

$$\Leftrightarrow C_1(t)^2 + t^2 - C_2(t)^2 = \alpha t^2 \qquad \text{i.e.} \quad C_1(t)^2 - C_2(t)^2 = (\alpha - 1)t^2 \qquad ②$$

from ①, ② we can get $\begin{cases} C_1(t) = \frac{\alpha t}{2} \\ C_2(t) = (1 - \frac{\alpha}{2})t \end{cases}$

(d) The cdf of $X_{(1)}$ is $1-(1-F(x))^n$

$$F(x) = \int_a^x \lambda e^{-\lambda(u-a)}\, du$$

$$= -e^{-\lambda(u-a)}\Big|_a^x = 1 - e^{-\lambda(x-a)}$$

$$F_{X_{(1)}}(x) = 1 - [e^{-\lambda(x-a)n}]$$. Define

$$H(t) = P[\, e^{\lambda(X_{(1)}-a)} \le t \,]$$

$$= P\left(X_{(1)} \le \frac{\log t}{\lambda} + a\right)$$

$$= F_{X_{(1)}}\left[\frac{\log t}{\lambda} + a\right]$$

$$= 1 - \left[e^{-\lambda\left[\frac{\log t}{\lambda} + a - a\right]n}\right]$$

$$= 1 - e^{-n\log t} = 1 - \frac{1}{t^n}, \quad t > 0$$

Thus $e^{\lambda(X_{(1)}-a)}$ is a pivotal quantity

P. 11

Thus a $(1-\alpha)\times 100\%$ joint confidence region for (λ, a) is

$$C(\lambda, a) = \left\{(\lambda, a) : a^* < e^{\lambda(X_{(1)}-a)} < b^*\right\}$$

Where $H(b^*) = 1 - \frac{\alpha}{2}$

$H(a^*) = \alpha/2$

Thus $1 - \frac{1}{b^{*n}} = 1 - \alpha/2 \Rightarrow b^* = \left(\frac{2}{\alpha}\right)^{1/n}$

$1 - \frac{1}{a^{*n}} = \alpha/2 \Rightarrow a^* = \left(\frac{2}{2-\alpha}\right)^{1/n}$

Thus

$$C(\lambda, a) = \left\{(\lambda, a) : \left(\frac{2}{2-\alpha}\right)^{1/n} < e^{\lambda(X_{(1)}-a)} < \left(\frac{2}{\alpha}\right)^{\frac{1}{n}}\right\}$$

yinzhou@email.unc.edu

(e) The joint density of the observations is

$$P(X \mid a, \lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda(X_i - a)} I(X_i \geq a)$$

$$= \lambda^n e^{-\lambda \Sigma(X_i - a)} I(a \leq X_{(1)})$$

$$= \lambda^n e^{-n\lambda[\bar{X} - a]} I(a \leq X_{(1)}),$$

$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

The numerator of the Bayes factor is

$$\text{numerator} = \int \lambda_0^n e^{-n\lambda_0[\bar{X} - a]} I(a \leq X_{(1)}) \pi(a)\, da$$

$$= \int \frac{\lambda_0^n}{a_0} e^{-n\lambda_0[\bar{X} - a]} I(a \leq X_{(1)}) I(0 < a < a_0)\, da$$

$$= \int_0^{m_1} \frac{\lambda_0^n}{a_0} e^{-n\lambda_0[\bar{X} - a]}\, da, \quad m_1 = \min(a_0, X_{(1)})$$

$$= \frac{\lambda_0^n}{a_0} e^{-n\lambda_0 \bar{X}} \left[\frac{1}{n\lambda_0} e^{n\lambda_0 a} \Big|_0^{m_1} \right]$$

$$= \frac{\lambda_0^{n-1}}{n a_0} e^{-n\lambda_0 \bar{X}} \left[e^{n\lambda_0 m_1} - 1 \right] = B_1$$

The denominator of the Bayes factor is

$$\text{denominator} = \iint \lambda^n e^{-n\lambda[\bar{x}-a]} I(a \le X_{(1)})\, \pi(a,\lambda)\, da\, d\lambda$$

$$= \iint_0^\infty \frac{r_0^{b_0}}{\Gamma(r_0)} \lambda^{n+r_0-1} e^{-\lambda[n\bar{x}-na+b_0]} I(a \le X_{(1)}) \frac{I(a_0 < a < 1)}{(1-a_0)}\, d\lambda\, da$$

$$= \int_{a_0}^{m_2} \int_0^\infty \frac{r_0^{b_0}}{\Gamma(r_0)} \frac{\lambda^{n+r_0-1}}{(1-a_0)} e^{-\lambda[n\bar{x}-na+b_0]}\, d\lambda\, da$$

where $m_2 = \min(1, X_{(1)})$

$$\int_{a_0}^{m_2} \left[\frac{r_0^{b_0}}{\Gamma(r_0)(1-a_0)}\right] \frac{\Gamma(n+r_0)}{[n\bar{X}+b_0-na]^{n+r_0}}\, da$$

$$= \left[\frac{\Gamma(n+r_0)\, r_0^{b_0}}{\Gamma(r_0)(1-a_0)}\right]\left[\frac{(n\bar{X}+b_0-na)^{1-n-r_0}}{1-n-r_0}\left(-\frac{1}{n}\right)\Bigg|_{a_0}^{m_2}\right]$$

$$= \frac{\Gamma(n+r_0)\, r_0^{b_0}}{\Gamma(r_0)(1-a_0)}\left[\frac{(n\bar{X}+b_0-na_0)^{1-n-r_0}}{n(1-n-r_0)} - \frac{(n\bar{x}+b_0-nm_2)}{n(1-n-r_0)}\right]$$

$$= B_2$$

2007 QUESTION 5

$X_i \sim p(x_i|\lambda_i) = \lambda_i e^{-\lambda_i x_i}$, $i = 1, 2$ and $X_1 \perp X_2$ [exponential (mean $= \frac{1}{\lambda_i}$)].

Suppose λ_1, λ_2 unknown.

(a) $\theta = \frac{\lambda_1}{\lambda_2}$. Show that $\theta \frac{X_1}{X_2} = \frac{\lambda_1 X_1}{\lambda_2 X_2}$ is a pivotal quantity.

Show that $P\left(\theta \frac{X_1}{X_2} < t\right)$ does not involve θ.

Let $j = \lambda_1 X_1 \Rightarrow x_1 = \frac{j}{\lambda_1} \Rightarrow f(j) = e^{-j}$

Let $k = \lambda_2 X_2 \Rightarrow f(k) = e^{-k}$

$\Big\}$ j, k are independent so $f(j,k) = e^{-(j+k)}$, $j, k > 0$

$\theta \frac{X_1}{X_2} = \frac{j}{k}$. The distribution of $\frac{j}{k}$ can be found using a bivariate transformation. Let $U = \frac{j}{k}$, $V = k$ $\Rightarrow j = uv$, $k = v$ $\Rightarrow$ jacobian $= \left|\det\begin{bmatrix} v & u \\ 0 & 1 \end{bmatrix}\right| = v$

$\Rightarrow f_{U,V} = v \cdot e^{-(uv+v)} = v \cdot e^{-v(u+1)}$, $v > 0$, $u > 0$

$$\Rightarrow f_U = \int_0^\infty v \cdot e^{-v(u+1)}\,dv = \left[\frac{-v}{u+1} e^{-v(u+1)}\right]_0^\infty + \frac{1}{u+1}\int_0^\infty e^{-v(u+1)}\,dv$$

$$= 0 + \frac{1}{(u+1)^2}\underbrace{\left[-e^{-v(u+1)}\right]_0^\infty}_{1}$$

$$= (u+1)^{-2}$$

$$F_U(t) = \int_0^t (u+1)^{-2}\,du = \left[-(u+1)^{-1}\right]_0^t = 1 - \frac{1}{t+1}, \quad t \in 0, \infty$$

Then $P\left(\theta \frac{X_1}{X_2} < t\right) = P(U < t) = 1 - \frac{1}{t+1} = \frac{t}{t+1}$ so that $\theta \frac{X_1}{X_2}$ is a pivotal quantity.

A $(1-\alpha)$ confidence interval is constructed as follows

$P\left(a < \theta \frac{X_1}{X_2} < b\right) = 1-\alpha$, where $\int_0^a f(u)\,du = \frac{\alpha}{2}$, $\int_0^b f(u)\,du = 1 - \frac{\alpha}{2}$ $\Big\}$ $\Rightarrow \frac{a}{a+1} = \frac{\alpha}{2}$, $\frac{b}{b+1} = 1 - \frac{\alpha}{2}$; $a = \frac{\alpha}{2-\alpha}$, $b = \frac{2-\alpha}{\alpha}$

$\Rightarrow$ A $(1-\alpha)$ confidence interval for θ is $\left[\frac{\alpha}{2-\alpha}\left(\frac{X_2}{X_1}\right), \frac{2-\alpha}{\alpha}\left(\frac{X_2}{X_1}\right)\right]$

or $\left\{\theta : a < \theta \frac{X_1}{X_2} < b\right\}$

(b) Let $\theta = (\lambda_1, \lambda_2)$. Show that $\lambda_1 X_1 + \lambda_2 X_2$ is a pivotal quantity.

From (a), with $J = \lambda_1 X_1$, $K = \lambda_2 X_2$, J, K are iid exponential (mean $= 1$)

Then, $\underbrace{J+K}_{=U} \sim \text{Gamma}(2, 1)$. $\Rightarrow f_U(u) = \frac{1}{\Gamma(2) 1^2} u^{2-1} e^{-u} = u e^{-u}$

Then, $P(a < U < b) = 1-\alpha$, where $\int_0^a f(u)\,du = \frac{\alpha}{2}$, $\int_0^b f(u)\,du = 1 - \frac{\alpha}{2}$

and a $(1-\alpha)$ confidence region for λ_1, λ_2 is

$\{\lambda_1, \lambda_2 : a < \lambda_1 X_1 + \lambda_2 X_2 < b\}$

(c) $H_0: \lambda_1 = \lambda_2$ vs. $H_1: \lambda_1 \neq \lambda_2$ $\Leftrightarrow$ $H_0: \lambda_2 - \lambda_1 = 0$ vs. $H_1: \lambda_2 - \lambda_1 \neq 0$

$$f_{X_1,X_2} = \lambda_1\lambda_2 \exp\{-\lambda_1 x_1 - \lambda_2 x_2\} = \exp\Big\{\underbrace{x_1(\lambda_2 - \lambda_1)}_{\text{parameter of interest}} - \underbrace{\lambda_2}_{\text{nuisance}}(x_2 + x_1) + \log\lambda_1 + \log\lambda_2\Big\}$$

A UMPU test for the above hypotheses can be derived using the statistic $X_1 | X_1 + X_2 = t$ (from multiparameter exponential family theory)

$f(x_1 | x_1 + x_2 = t) = P(X_1 = c \,|\, X_1 + X_2 = t)$

$$= \frac{P(X_1 = c, X_2 = t-c)}{\int_0^t P(X_1 = c, X_2 = t-c)} = \frac{P(X_1 = c)P(X_2 = t-c)}{\int_0^t P(X_1 = c)P(X_2 = t-c)\,dc} = \frac{\lambda_1 e^{-\lambda_1 c}\lambda_2 e^{-\lambda_2(t-c)}}{\int_0^t \lambda_1 e^{-\lambda_1 c}\lambda_2 e^{-\lambda_2(t-c)}\,dc}$$

$\int_0^1 e^{-\lambda c}\, dc = \left[-\frac{1}{\lambda} e^{-\lambda c}\right]_0 \quad 1 - \frac{1}{\lambda} \ldots$

$$= \frac{e^{-\lambda_1 c + \lambda_2 c}}{\int_0^1 e^{-\lambda_1 c + \lambda_2 c}} \quad \frac{e^{-(\lambda_1 - \lambda_2) c}}{\int_0^t e^{-(\lambda_1 - \lambda_2) c}\, dc} \quad \frac{e^{-(\lambda_1 - \lambda_2) c}}{1 - \frac{1}{\lambda_1 - \lambda_2} e^{-(\lambda_1 - \lambda_2) t}}, \quad c \in (0, t)$$

A more precise statement is the UMPU test is

$$\phi(X_1) = \begin{cases} 1 & \text{if } X_1 < c(X_1 + X_2) \text{ or if } X_1 > c(X_1 + X_2) \\ 0 & \text{otherwise} \end{cases}$$

The joint of $f_{X_1, X_2} = \lambda_1 \lambda_2 e^{-\lambda_1 x_1 - \lambda_2 x_2}, \quad x_1, x_2 > 0$

let $U = X_1 + X_2 \Rightarrow X_2 = U - V \Rightarrow |\det[J]| = \left|\det\begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}\right| = 1$

$V = X_1 \qquad X_1 = V$

$$\Rightarrow f_{U,V} = \lambda_1 \lambda_2 e^{-\lambda_1 v - \lambda_2 (u - v)}$$

$$= \lambda_1 \lambda_2 e^{-\lambda_2 u - v(\lambda_1 - \lambda_2)}, \quad 0 < v < u < \infty$$

let $\lambda_1 = \lambda_2 = \lambda_*$ at this point

$$f_U(u) = \lambda_1 \lambda_2 e^{-\lambda_2 u} \int_0^u e^{-v(\lambda_1 - \lambda_2)}\, dv$$

$$= \lambda_1 \lambda_2 e^{-\lambda_2 u} \left[1 - \frac{1}{\lambda_1 - \lambda_2} e^{-u(\lambda_1 - \lambda_2)}\right]$$

$$\Rightarrow f_{V|U}(v) = \frac{f_{U,V}}{f_U} = \frac{e^{-\lambda_1 v - \cancel{\lambda_2 u} + \lambda_2 v}}{e^{-\cancel{\lambda_2 u}}\left[1 - \frac{1}{\lambda_1 - \lambda_2} e^{-u(\lambda_1 - \lambda_2)}\right]} = \frac{e^{-v(\lambda_1 - \lambda_2)}}{1 - \frac{1}{\lambda_1 - \lambda_2} e^{-u(\lambda_1 - \lambda_2)}} \quad 0 < v < u$$

with $\lambda_1 = \lambda_2 = \lambda$, $f_{U,V} = \lambda_*^2 e^{-\lambda_* u}, \quad 0 < v < u < \infty$

$$\Rightarrow f_U = \int_0^u \lambda_*^2 e^{-\lambda_* u}\, dv = u \cdot \lambda_*^2 e^{-\lambda_* u} \qquad \text{under } H_0$$

$$\Rightarrow f_{V|U} = f(X_1 \mid X_1 + X_2 = t) = \frac{\lambda_*^2 e^{-\lambda_* u}}{u \cdot \lambda_*^2 e^{-\lambda_* u}} = \frac{1}{t} \quad [\text{i.e. } \sim U(0, t)]$$

We have $\phi(X_1) = \begin{cases} 1 & \text{if } X_1 < c_1 \text{ or if } X_1 > c_2 \\ 0 & \text{otherwise} \end{cases}$

where (1) $E_{H_0}[\phi] = \alpha \Rightarrow P_{H_0}(X_1 < c_1 \mid X_1 + X_2 = t) + P_{H_0}(X_1 > c_2 \mid X_1 + X_2 = t) = \alpha$

$$\Rightarrow \frac{c_1}{t} + \frac{(t - c_2)}{t} = \alpha \Rightarrow c_1 + t - c_2 = t\alpha \Rightarrow c_1 - c_2 = t(\alpha - 1)$$

$$\Rightarrow c_1 = c_2 + t(\alpha - 1)$$

(2) $E_{H_0}[X_1 \cdot \phi(X_1)] = \alpha \cdot E_{H_0}[X_1]$

$$= \alpha \cdot \frac{t}{2}$$

$$\int_0^{c_1} x_1 \cdot \frac{1}{t}\, dx_1 + \int_{c_2}^{t} x_1 \frac{1}{t}\, dx_1 = \alpha \frac{t}{2}$$

$$\Rightarrow \int_0^{c_1} x_1\, dx_1 + \int_{c_2}^t x_1\, dx_1 = \alpha \frac{t^2}{2}$$

$$\frac{1}{2} c_1^2 + \frac{1}{2}(t^2 - c_2^2) = \alpha \frac{t^2}{2} \Rightarrow c_1^2 + t^2 - c_2^2 = \alpha t^2$$

$$\Rightarrow c_1^2 - c_2^2 = t^2(\alpha - 1)$$

$\therefore$ (1) + (2) $\Rightarrow [c_2 + t(\alpha - 1)]^2 - c_2^2 = t^2(\alpha - 1)$

$$\Rightarrow \cancel{c_2^2} + 2c_2 t(\alpha - 1) + t^2(\alpha - 1)^2 - \cancel{c_2^2} = t^2(\alpha - 1)$$

$$2c_2 + t(\alpha - 1) = t$$

$$\ldots \Rightarrow c_2 = \frac{2t - t\alpha}{2} = t - t\frac{\alpha}{2} = t\left(1 - \frac{\alpha}{2}\right)$$

(d) $X_1, \ldots, X_n$ iid $p(x|a,\lambda) = \lambda e^{-\lambda(x-a)}$, $x \ge a$, λ, a unknown.

The MLE of a is $\hat{a} = X_{(1)}$. Its CDF is

$(X_{(1)} < t)$ $\quad 1 - [1 - F(t)]^n$, where $F(x) = \int_a^x \lambda e^{-\lambda(x-a)} = [-e^{-\lambda(x-a)}]_a^x$

$= 1 - e^{-\lambda(x-a)}$

$= 1 - [1 - (1 - e^{-\lambda(t-a)})]^n$

$= 1 - [e^{-\lambda(t-a)}]^n$

$= 1 - e^{-\lambda n(t-a)}$

Now, need to find a function of $X_{(1)}$, say $g(X_{(1)})$, so that

$P[g(X_{(1)}) < t]$ is free of λ, a

consider $g(X_{(1)}) = e^{\lambda(X_{(1)}-a)}$

$P[e^{\lambda(X_{(1)}-a)} < t] = P[\lambda(X_{(1)}-a) < \log t]$

$= P[X_{(1)} < a + \frac{1}{\lambda}\log t] = 1 - e^{-\lambda n[\frac{1}{\lambda}\log t + \cancel{a} - a]}$

$= 1 - e^{-n\log t} = 1 - t^{-n}$

so that $e^{\lambda(X_{(1)}-a)}$ is a pivotal quantity. A $(1-\alpha)$ confidence region for (λ, a) is

$\{(\lambda, a) : a^* < \overbrace{e^{\lambda(X_{(1)}-a)}}^{G} < b^*\}$

where $P[G < a^*] = 1 - (a^*)^{-n} = \frac{\alpha}{2} \Rightarrow (1 - \frac{\alpha}{2})^{-\frac{1}{n}} = a^*$

$P[G < b^*] = 1 - (b^*)^{-n} = 1 - \frac{\alpha}{2} \Rightarrow b^* = (\frac{\alpha}{2})^{-\frac{1}{n}}$

(e) test $H_0: \lambda = \lambda_0$, $0 < a < a_0$ vs. $H_1: \lambda \ne \lambda_0$, $a_0 < a < 1$, where $a_0 \in (0,1)$, $\lambda_0 > 0$, and (λ_0, a_0) are specified values in the hypotheses. Under H_0, $a \sim U(0, a_0)$. Under H_1, $\lambda \perp a$ a priori, and $a \sim U(a_0, 1)$, $\lambda \sim \pi(\lambda) \propto \underbrace{\lambda^{r_0-1} \exp(-\lambda b_0)}$, $r_0 > 0$, $b_0 > 0$. Compute the Bayes factor in favor of H_0.

Gamma kernel must fill in blanks here.

$\frac{(H_0)}{(H_1)}: \quad B = \frac{\int_\Theta p(x|\theta)\, p(\theta|H_0)\, \partial\theta}{\int_\Theta p(x|\theta)\, p(\theta|H_1)\, \partial\theta} = \frac{B_0}{B_1}$

$B_0 = \int_a \lambda_0^n \exp[-\lambda_0 \sum(X_i - a)] \cdot \frac{1}{a_0}\, \partial a = \lambda_0^n \exp[-\lambda_0 \sum X_i] \frac{1}{a_0} \int_0^{a_0} \exp(n\lambda_0 a)\, da$

$\frac{1}{n\lambda_0}[e^{n\lambda_0 a}]_0^{a_0}$

$\frac{\lambda_0^n}{n\lambda_0} \frac{1}{a_0} \exp(-\lambda_0 \sum X_i)[\exp(n\lambda_0 a_0) - 1]$

$B_1 \quad \int_a \int_\lambda \lambda^n \exp[\lambda \sum(X_i - a)] \left(\frac{1}{1-a_0}\right) \lambda^{r_0-1} \exp[-\lambda b_0]\, \partial\lambda\, \partial a$

$\int_a \int_\lambda \lambda^{n+r_0-1} \exp\{-\lambda[\sum(X_i - a) + b_0]\}\, \partial\lambda\, \partial a \quad \left(\frac{1}{1-a_0}\right)$

$\int \Gamma(n+r_0) \left[\frac{1}{[illegible]}\right]^{n+r_0} da \quad \left(\frac{1}{[illegible]}\right)$

2008 QUESTION 1

Y_i independent $\sim$ Gamma (μ_i, ν) OR Gamma $\left(\nu, \frac{\mu_i}{\nu}\right)$

$$f(y_i|\mu_i,\nu) = \frac{y^{\nu-1}\exp\left[-y/\left(\frac{\mu_i}{\nu}\right)\right]}{\Gamma(\nu)\left(\frac{\mu_i}{\nu}\right)^{\nu}}, \quad y_i \geq 0, \quad i = 1, \dots, 12$$

(a) Show that $G(\mu_i, \nu)$ belongs to exponential family and derive MGF

$$f(y_i|\mu_i,\nu) = \exp\left\{-y_i\left(\frac{\nu}{\mu_i}\right) + (\nu-1)\log y_i - \log\Gamma(\nu) - \nu\log\left(\frac{\mu_i}{\nu}\right)\right\}$$

$$= \exp\Big\{\nu\Big[y_i\Big(\underbrace{\frac{-1}{\mu_i}}_{\theta_i = \frac{-1}{\mu_i}}\Big) - \log(\mu_i) + \log y_i\Big] \underbrace{- \log y_i - \log\Gamma(\nu) + \nu\log\nu}_{s(y_i,\nu)}\Big\}$$

$$= \exp\Big\{\nu\Big[y_i\,\theta_i \underbrace{- \log\left(\frac{-1}{\theta_i}\right)}_{b(\theta_i)} + \log y_i\Big] - s(y_i,\nu)\Big\}$$

The MGF is equal to $\exp\left\{\phi\left[b\left(\theta + \frac{t}{\phi}\right) - b(\theta)\right]\right\}$, an exponential family result, with $\phi = \nu$, $\theta_i = \frac{-1}{\mu_i}$, $b(\theta_i) = \log\left(\frac{-1}{\theta_i}\right)$, so that:

$$M_{Y_i}(t) = \exp\left\{\nu\left[\log\left(\frac{-1}{\theta_i + \frac{t}{\nu}}\right) - \log\left(\frac{-1}{\theta_i}\right)\right]\right\}$$

$$= \left[\frac{\frac{1}{\theta_i + \frac{t}{\nu}}}{\frac{1}{\theta_i}}\right]^{\nu} = \left[1 + \frac{t}{\nu\theta_i}\right]^{-\nu} = \left[1 - \frac{\mu_i}{\nu}t\right]^{-\nu}$$

Assume $\mu_i = e^{\beta_1 + \beta_2 x_i}$, $x_i = 0$ if Hospital A, 1 if Hospital B and that ν is a known integer

(b) The log-likelihood of (β_1, β_2) for one individual is proportional to:

$$\ell_i(\beta_1, \beta_2) \propto \nu\left[y_i\left(-e^{-\beta_1 - \beta_2 x_i}\right) - (\beta_1 + \beta_2 x_i)\right]$$

so that

$$\ell \propto \nu\left\{-e^{-\beta_1}\sum_{i=1}^{6} y_i - 6\beta_1 - e^{-\beta_1-\beta_2}\sum_{i=7}^{12} y_i - 6(\beta_1+\beta_2)\right\}$$

The score vector is equal to:

$$\frac{\partial\ell}{\partial\beta_1} = \nu\left\{e^{-\beta_1}\sum_{i=1}^{6} y_i - 6 + e^{-\beta_1-\beta_2}\sum_{i=7}^{12} y_i - 6\right\} = 0$$

$$\frac{\partial\ell}{\partial\beta_2} = \nu\left\{e^{-\beta_1-\beta_2}\sum_{i=7}^{12} y_i - 6\right\} = 0$$

$$\Rightarrow e^{-\beta_1-\beta_2} = \frac{6}{\sum_{i=7}^{12} y_i} = \frac{1}{\bar{y}_B}$$

$$\Rightarrow e^{-\beta_1}\sum_{i=1}^{6} y_i + \frac{1}{\bar{y}_B}\sum_{i=7}^{12} y_i = 12 \Rightarrow e^{-\beta_1} = \frac{1}{\bar{y}_A} \Rightarrow \hat{\beta}_1 = \log(\bar{y}_A)$$

$$\Rightarrow \frac{1}{\bar{y}_A}e^{-\beta_2} = \frac{1}{\bar{y}_B} \Rightarrow e^{-\beta_2} = \frac{\bar{y}_A}{\bar{y}_B} \Rightarrow \hat{\beta}_2 = \log\left(\frac{\bar{y}_B}{\bar{y}_A}\right)$$

The observed information is calculated as:

$$\frac{\partial^2\ell}{\partial\beta_1^2} = \nu\left\{-e^{-\beta_1}6\bar{y}_A - e^{-\beta_1-\beta_2}6\bar{y}_B\right\} \Rightarrow E\left[-\frac{\partial^2\ell}{\partial\beta_1^2}\right] = 12$$

$$\frac{\partial^2\ell}{\partial\beta_2\partial\beta_1} = \nu\left\{-e^{-\beta_1-\beta_2}6\bar{y}_B\right\} \Rightarrow E\left[-\frac{\partial^2\ell}{\partial\beta_2\partial\beta_1}\right] = 6$$

$$\frac{\partial^2\ell}{\partial\beta_2^2} = \nu\left\{-e^{-\beta_1-\beta_2}6\bar{y}_B\right\} \Rightarrow E\left[-\frac{\partial^2\ell}{\partial\beta_2^2}\right] = 6$$

$$I_n^*(\beta) = \lim_{n\to\infty} \frac{1}{n} I_n(\beta) = v \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

so that $I_n(\beta) = v \begin{bmatrix} 12 & 6 \\ 6 & 6 \end{bmatrix} \Rightarrow \left[I_n^*(\beta)\right]^{-1} = \frac{1}{v} \cdot \frac{1}{\frac{1}{4}} \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix} = \frac{1}{v} \begin{bmatrix} 2 & -2 \\ -2 & 4 \end{bmatrix}$

By MLE theory, $\sqrt{n} \begin{pmatrix} \hat{\beta}_1 - \beta_1 \\ \hat{\beta}_2 - \beta_2 \end{pmatrix} \xrightarrow{d} N_2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \underbrace{\left[\lim_{n\to\infty} \frac{1}{n} I_n(\beta)\right]^{-1}}_{= \frac{1}{v} \begin{bmatrix} 2 & -2 \\ -2 & 4 \end{bmatrix}} \right)$

Then, consistent estimates of $SD(\beta_1)$ and $SD(\beta_2)$ are $\sqrt{\frac{2}{v}}$ and $\sqrt{\frac{4}{v}}$, respectively.

(c) $H_0: \beta_2 = 0$ vs. $H_1: \beta_2 \neq 0$.

Let $\mu_A = e^{\beta_1}$ and $\mu_B = e^{\beta_1 + \beta_2}$. Under H_0, $\mu_A = \mu_B$, so that $\frac{1}{\mu_A} = \frac{1}{\mu_B}$, and $\frac{-1}{\mu_A} = \frac{-1}{\mu_B}$, and $\frac{-1}{\mu_A} + \frac{1}{\mu_B} = 0$. Recall that the density

$f(y_i|\mu_i, v) \propto \exp\left\{v\left[y_i\left(\frac{-1}{\mu_i}\right) - \log(\mu_i)\right]\right\}$, so that the joint density is

$$f(y_i|\mu_i, v) \propto \exp\left\{v\left[\left(\frac{-1}{\mu_A}\right)\sum_{i=1}^{6} y_i + \left(\frac{-1}{\mu_B}\right)\sum_{i=7}^{12} y_i - \log(\mu_A) - \log(\mu_B)\right]\right\}$$

$$= \exp\left\{v\left[\underbrace{\left(\frac{-1}{\mu_A} + \frac{1}{\mu_B}\right)}_{\text{interest}} \underbrace{\sum_{i=1}^{6} y_i}_{U} + \underbrace{\left(\frac{-1}{\mu_B}\right)}_{\text{nuisance}} \underbrace{\sum_{i=1}^{12} y_i}_{T} - \log(\mu_A) - \log(\mu_B)\right]\right\}$$

As $H_0: \beta_2 = 0 \Rightarrow H_0: \frac{-1}{\mu_A} + \frac{1}{\mu_B} = 0$, and U is sufficient complete statistic, from multiparameter exponential family theory, the UMPU size α test of $H_0: \frac{-1}{\mu_A} + \frac{1}{\mu_B} = 0$ vs. $H_1: \frac{-1}{\mu_A} + \frac{1}{\mu_B} \neq 0$ is

$$\phi(U) = \begin{cases} 1 & \text{if } U < c_1(T) \text{ or } U > c_2(T) \\ \gamma & \text{if } U = c_1(T) \text{ or } c_2(T) \\ 0 & \text{otherwise} \end{cases}$$

Now, $U = \sum_{i=1}^{6} y_i \sim \text{Gamma}\left(6v, \frac{\mu_A}{v}\right)$, and $T = U + V$, $V \sim \text{Gamma}\left(6v, \frac{\mu_B}{v}\right)$, $U \perp V$, and $J = \frac{U}{T} = \frac{U}{U+V} \sim \text{Beta}(6v, 6v)$ under H_0 $\mu_A = \mu_B = \mu$.

$$\phi(J) = \begin{cases} 1 & \text{if } J < c_1 \text{ or } J > c_2 \\ 0 & \text{otherwise} \end{cases}$$

c_1, c_2 chosen to satisfy (1) $P_{H_0}[J < c_1] + P_{H_0}[J > c_2] = \alpha$

(2) $E[J \cdot \phi(J)] = \alpha E(J) = \frac{\alpha}{2}$

(d) $\{\beta_2 = $ null is not rejected$\}$, $J = \frac{\sum y_i}{\sum y_i} \sim$ Beta when $\beta_2 = 0$

(e) Use cumulant generating function to find $E(Y_i^*)$, $Var(Y_i^*)$,

where $Y_i^* = \log Y_i$

$$K_{Y_i^*}(t) = \log M_{Y_i^*}(t) \qquad\qquad E[Y_i^t] = \frac{\Gamma(v+t)\left(\frac{\mu_i}{v}\right)^t}{\Gamma(v)} \underbrace{\int_0^\infty \frac{y_i^{v+t-1}\, e^{-y_i/t}}{\Gamma(v+t)\left(\frac{\mu_i}{v}\right)^{v+t}}\, dt}_{=1}$$

$$= \log\left\{E[e^{tY_i^*}]\right\}$$

$$= \log\left\{E[Y_i^t]\right\}$$

$$= \log\left\{\frac{\Gamma(v+t)}{\Gamma(v)}\left(\frac{\mu_i}{v}\right)^t\right\} = t\log\left(\frac{\mu_i}{v}\right) + \log\Gamma(v+t) - \log\Gamma(v)$$

$$E(Y_i^*) = \frac{\partial}{\partial t}K_{Y_i^*}(t)\Big|_{t=0} = \log\left(\frac{\mu_i}{v}\right) + \frac{\Gamma'(v+t)}{\Gamma(v)}\Big|_{t=0} = \frac{\Gamma'(v)}{\Gamma(v)} + \log\mu_i - \log v$$

$$= \psi(v) + \beta_1 + \beta_2 x_i - \log v$$

$$= \beta_1^* + \beta_2 x_i$$

$$Var(Y_i^*) = \frac{\partial^2}{\partial t^2}K_{Y_i^*}(t)\Big|_{t=0} = \frac{\Gamma''(v+t)}{\Gamma(v)}\Big|_{t=0} = \frac{\Gamma''(v)}{\Gamma(v)} = \psi'(v)$$

(f) Treat Y_i^*, $i=1,\dots,12$, as independent random variables with mean $\beta_1^* + \beta_2 x_i$ and variance $\psi'(v)$.

(a) Let $\underset{12\times 2}{D} = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_{12}\end{bmatrix}$, $\underset{12\times 12}{V} = \begin{bmatrix} \psi'(v) & & 0 \\ & \ddots & \\ 0 & & \psi'(v)\end{bmatrix}$, $e_{12\times 1} = \begin{bmatrix} y_1^* - (\beta_1^* + \beta_2 x_1) \\ \vdots \\ y_{12}^* - (\beta_1^* + \beta_2 x_{12})\end{bmatrix}$

With $x_i = 0$ $i=1,\dots,6$ and $x_i = 1$ $i=7,\dots,12$

$D = \begin{bmatrix} 1 & 0 \\ \vdots & 0 \\ \vdots & \vdots \\ 1 & 1\end{bmatrix}$, V is same, $e = \begin{bmatrix} y_1^* - \beta_1^* \\ y_6^* - \beta_1^* \\ y_7^* - \beta_1^* - \beta_2 \\ y_{12}^* - \beta_1^* - \beta_2\end{bmatrix}$

(i) solve $D'V^{-1}e = 0$

$$D'V^{-1}e = \begin{bmatrix} 1 & \cdots & & 1 \\ 0 \cdots & 0 & 1 \cdots & 1\end{bmatrix}\begin{bmatrix} \frac{1}{\psi'(v)} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\psi'(v)}\end{bmatrix}\begin{bmatrix} y_1^* - \beta_1^* \\ \vdots \\ y_6^* - \beta_1^* \\ y_7^* - \beta_1^* - \beta_2 \\ \vdots \\ y_{12}^* - \beta_1^* - \beta_2\end{bmatrix}$$

$$\begin{bmatrix} \frac{1}{\psi'(v)} & \cdots & \frac{1}{\psi'(v)} \\ 0 & 0 \;\frac{1}{\psi'(v)} & \frac{1}{\psi'(v)}\end{bmatrix}\begin{bmatrix} y_1^* - \beta_1^* \\ y_6^* - \beta_1^* \\ y_7^* - \beta_1^* - \beta_2 \\ \vdots \\ y_{12}^* - \beta_1^* - \beta_2\end{bmatrix}$$

$$\frac{1}{\psi'(v)}\begin{bmatrix} \sum_{i=1}^{12} y_i^* - 12\beta_1^* - 6\beta_2 \\ \sum_{i=7}^{12} y_i^* - 6\beta_1^* - 6\beta_2\end{bmatrix} = \begin{bmatrix} 0 \\ 0\end{bmatrix}$$

(ii) Find MLEs, $\hat{\beta}_1^*$, $\hat{\beta}_2$

From (1), $\beta_1^* + \beta_2 = \bar{y}_B^*$ $\Rightarrow \sum_{i=1}^{12} y_i^* = 6(\beta_1^* + \beta_2) + 6\beta_1^*$

$\Rightarrow \sum_{i=1}^{12} y_i^* = \sum_{i=7}^{12} y_i + 6\beta_1^* \Rightarrow \hat{\beta}_1^* = \bar{y}_A^*$

$\hat{\beta}_2 = \bar{y}_B^* - \hat{\beta}_1^*$

$= \bar{y}_B^* - \bar{y}_A^*$

The variances of $\hat{\beta}_1^*$ and $\hat{\beta}_2$ are found as:

$Var(\hat{\beta}_1^*) = Var(\bar{y}_A^*) = \frac{1}{6} Var(y_1^*) = \frac{1}{6}\psi'(r)$

$Var(\hat{\beta}_2) = Var(\bar{y}_B^* - \bar{y}_A^*) = Var(\bar{y}_B^*) + Var(\bar{y}_A^*)$ as $\bar{y}_B^* \perp \bar{y}_A^*$

$= \frac{1}{3}\psi'(r)$

2008 Question 2

② $X \sim Bin(n_1, p_1) \perp Y \sim Bin(n_2, p_2)$ so that

$$p(x,y \mid p_1,p_2) = \binom{n_1}{x}\binom{n_2}{y} p_1^x p_2^y (1-p_1)^{n_1-x}(1-p_2)^{n_2-y} \quad x \in \{0,\dots,n_1\},\ y \in \{0,\dots,n_2\}$$

(a) want to estimate $p_1 - p_2$. squared error loss $L(p_1,p_2,a) = (p_1-p_2-a)^2$,
where $|a| \le 1$. Derive Bayes rule wrt prior of $U(0,1)$ on p_1, p_2

$$p(p_1,p_2 \mid x,y) = \frac{p(x,y \mid p_1,p_2)\cdot p(p_1,p_2)}{\int_{p_1,p_2} p(x,y \mid p_1,p_2)\, p(p_1,p_2)} = \frac{\cancel{\binom{n_1}{x}}\cancel{\binom{n_2}{y}} p_1^x(1-p_1)^{n_1-x} p_2^y (1-p_2)^{n_2-y}}{\int_{p_1} \cancel{\binom{n_1}{x}} p_1^x (1-p_1)^{n_1-x} dp_1 \int_{p_2} \cancel{\binom{n_2}{y}} p_2^y (1-p_2)^{n_2-y} dp_2}$$

= product of two Beta distributions ($U(0,1)$ applied to binomial is a Beta conjugate prior)

$$= \frac{1}{B(x+1, n_1-x+1)} p_1^x (1-p_1)^{n_1-x} \times \frac{1}{B(y+1, n_2-y+1)} p_2^y (1-p_2)^{n_2-y}$$

under squared error loss, the Bayes estimator of p_1-p_2 is known to be the posterior mean of p_1 minus the posterior mean of p_2

$$\widehat{(p_1-p_2)}_B = E[p_1 \mid x,y] - E[p_2 \mid x,y] = \frac{x+1}{n_1+2} - \frac{y+1}{n_2+2}$$

(b) the frequentist risk of the Bayes rule in (a) is its variance plus its bias squared, under squared error loss.

$$R[\widehat{(p_1-p_2)}_B] = Var\left(\frac{X+1}{n_1+2} - \frac{Y+1}{n_2+2}\right) + \left\{E\left(\frac{X+1}{n_1+2} - \frac{Y+1}{n_2+2}\right) - (p_1-p_2)\right\}^2$$

$$= \frac{n_1p_1(1-p_1)}{(n_1+2)^2} + \frac{n_2p_2(1-p_2)}{(n_2+2)^2} + \left\{\frac{n_1p_1+1}{n_1+2} - \frac{n_2p_2+1}{n_2+2} - (p_1-p_2)\right\}^2$$

$$= \frac{n_1p_1(1-p_1)}{(n_1+2)^2} + \frac{n_2p_2(1-p_2)}{(n_2+2)^2} + \left\{\frac{n_1p_1+1-(n_1+2)p_1}{(n_1+2)} - \frac{n_2p_2+1-(n_2+2)p_2}{(n_2+2)}\right\}^2$$

$+1-4p_1+4p_1^2$

$4)\ p_1^2($

$$= \boxed{\frac{n_1p_1(1-p_1)}{(n_1+2)^2} + \frac{n_2p_2(1-p_2)}{(n_2+2)^2} + \left\{\frac{1-2p_1}{(n_1+2)} - \frac{1-2p_2}{(n_2+2)}\right\}^2}$$

(*) this expression ~~referred to~~ later

$$= \frac{n_1p_1(1-p_1)+(1-2p_1)^2}{(n_1+2)^2} + \frac{n_2p_2(1-p_2)+(1-2p_2)^2}{(n_2+2)^2} - \frac{2(1-2p_1)(1-2p_2)}{(n_1+2)(n_2+2)}$$

$$\frac{1+p_1(n_1-4)+p_1^2(4-n_1)}{(n_1+2)^2} + \frac{1+p_2(n_2-4)+p_2^2(4-n_2)}{(n_2+2)^2} - \frac{2(1-2p_2-2p_1+4p_1p_2)}{(n_1+2)(n_2+2)}$$

$$= \frac{1}{(n_1+2)^2} + \frac{1}{(n_2+2)^2} - \frac{2}{(n_1+2)(n_2+2)} + p_1\left[\frac{(n_1-4)}{(n_1+2)^2} + \frac{4}{(n_1+2)(n_2+2)}\right] + p_2\left[\frac{(n_2-4)}{(n_2+2)^2} + \frac{4}{(n_1+2)(n_2+2)}\right] - \frac{8p_1p_2}{(n_1+2)(n_2+2)} + \dots$$

1

$$+ \; p_1^2 \left[\frac{(4-n_1)}{(n_1+2)^2} \right] + p_2^2 \left[\frac{(4-n_2)}{(n_2+2)^2} \right]$$

The Bayes risk can be found by filling in the functions of p_i, $i=1,2$ with their prior expectations. Note that $p_i \sim U(0,1)$ (prior) so that

$$E[p_i] = \frac{1}{2}$$

$$E[p_i^2] = \int_0^1 p_i^2 \, dp_i = \frac{1}{3}\left[p_i^3\right]_0^1 = \frac{1}{3}$$

$$\Rightarrow R[(\widehat{p_1 - p_2})_B] = \frac{1}{(n_1+2)^2} + \frac{1}{(n_2+2)^2} - \frac{2}{(n_1+2)(n_2+2)} + \frac{1}{2}\left[\frac{(n_1-4)}{(n_1+2)^2} + \cancel{\frac{8}{(n_1+2)(n_2+2)}} + \frac{(n_2-4)}{(n_1+2)^2}\right] - \cancel{\frac{2}{(n_1+2)(n_2+2)}}$$

$$+ \frac{1}{3}\left[\frac{(4-n_1)}{(n_1+2)^2} + \frac{(4-n_2)}{(n_2+2)^2}\right]$$

$$\frac{1}{(n_1+2)^2}\left[1 + \frac{1}{2}(n_1-4) - \frac{1}{3}(n_1-4)\right] + \frac{1}{(n_2+2)^2}\left[1 + \frac{1}{2}(n_2-4) - \frac{1}{3}(n_2-4)\right]$$

$$= \frac{1}{(n_1+2)^2}\Big[\underbrace{1 + \frac{1}{6}(n_1-4)}_{\frac{1}{6}n_1 + \frac{1}{3}}\Big] + \frac{1}{(n_2+2)^2}\Big[\underbrace{1 + \frac{1}{6}(n_2-4)}_{\frac{1}{6}n_2 + \frac{1}{3}}\Big] = \frac{1}{6}\left[\frac{1}{n_1+2} + \frac{1}{n_2+2}\right]$$

Appeal to theorems about admissibility → Consider the unbiased estimator $\frac{x}{n_1} - \frac{y}{n_2} = (\widehat{p_1 - p_2})_C$ as a competitor. Then

$$R[(\widehat{p_1 - p_2})_C] = Var\left(\frac{x}{n_1} - \frac{y}{n_2}\right) = \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2} + 0$$

$$= p_1(1-p_1)\left[\frac{1}{n_1}\right] + p_2(1-p_2)\left[\frac{1}{n_2}\right] + 0$$

$(\widehat{p_1 - p_2})_B$ is inadmissible if there exists another estimator, d^*, whose frequentist risk, $R(d^*, p_1-p_2)$, is such that $R(d^*, p_1-p_2) \le R((\widehat{p_1-p_2})_B, p_1-p_2)$ for all p_1, p_2 with strict inequality for at least one (p_1-p_2) value in the parameter space.

From (*)

$$R[(\widehat{p_1 - p_2})_B] = p_1(1-p_1)\left[\frac{n_1}{(n_1+2)^2}\right] + p_2(1-p_2)\left[\frac{n_2}{(n_2+2)^2}\right] + \left\{\frac{1-2p_1}{(n_1+2)^2} - \frac{1-2p_2}{(n_2+2)^2}\right\}^2$$

For all p_1, p_2, $R[(\widehat{p_1-p_2})_B] < R[(\widehat{p_1-p_2})_C]$

$(\widehat{p_1-p_2})_C$ is the UMVUE for p_1-p_2, so it is the best unbiased estimator in terms of variance. As $(\widehat{p_1-p_2})_B$ beats this estimator, it beats all unbiased estimators. $\hat{p}$ is admissible.

F. F. B

$p_1(1-p_2) - p_2(1-p_1) \le 0$

$p_1 - p_1p_2 - p_2 + p_1p_2 \le 0$

$(p_1 - p_2) \le 0$ $\qquad\qquad$ $(p_1 - p_2) > 0$.

(c) Derive a UMPU size α test for testing $H_0: p_1 \le p_2$ versus $H_1: p_1 > p_2$ and derive an explicit expression for the p-value of the test.

First, reformulate the hypotheses of interest in terms of a different parameter

$$H_0: p_1 \le p_2 \iff p_1 - p_2 \le 0 \iff \frac{p_1 - p_2 - p_1p_2 + p_1p_2}{(1-p_1)(1-p_2)} \le 0$$

$$\iff \frac{p_1(1-p_2) - p_2(1-p_1)}{(1-p_1)(1-p_2)} \le 0$$

$$\iff \frac{p_1}{(1-p_1)} - \frac{p_2}{(1-p_2)} \le 0$$

$$\Rightarrow \log\left(\frac{p_1}{1-p_1}\right) \le \log\left(\frac{p_2}{1-p_2}\right)$$

$$\Rightarrow \log\left\{\frac{\frac{p_1}{1-p_1}}{\frac{p_2}{1-p_2}}\right\} = \psi \le 0$$

Now, write the density $p(x, y | p_1, p_2)$ as the following

$$p(x,y|p_1,p_2) = \underbrace{\binom{n_1}{x}\binom{n_2}{y}}_{h(x,y)} p_1^x (1-p_1)^{n_1-x} p_2^y (1-p_2)^{n_2-y}$$

$$= h(x,y)\exp\left\{x\log\left[\frac{p_1}{1-p_1}\right] + y\log\left[\frac{p_2}{1-p_2}\right] + n_1\log(1-p_1) + n_2\log(1-p_2)\right\}$$

$$= h(x,y)\exp\left\{x\psi + (x+y)\log\left[\frac{p_2}{1-p_2}\right] + n\log(1-p_1) + n_2\log(1-p_2)\right\}$$

An exponential family with sufficient statistic X for ψ, and sufficient statistic $(x+y)$ for the nuisance parameter the UMPU test of $H_0: \psi \le 0$ versus $H_1: \psi > 0$ has the following form, where 1 indicates rejection of H_0, 0 acceptance, and γ rejection w.p. γ.

$$\phi(x) = \begin{cases} 1 & \text{if } X > c_\alpha(x+y) \\ \gamma & \text{if } X = c_\alpha(x+y) \\ 0 & \text{if } X < c_\alpha(x+y) \end{cases}$$

(Type I error $= \alpha$)

an appropriately chosen value that is a function of $x+y$

Now, note that the conditional distribution of $X | X+y$ is found as

$$P_r[X = c | X+Y = t] = \frac{P_r[X = c, Y = t-c]}{P_r[X+Y = t]}$$

$$\frac{\Pr[X=c,\ Y=t-c]}{\Pr[X+Y=t]} = \frac{\binom{n_1}{c} p_1^{c}(1-p_1)^{n_1-c}\binom{n_2}{t-c} p_2^{t-c}(1-p_2)^{n_2-(t-c)}}{\sum_{\text{all } c}\binom{n_1}{c} p_1^{c}(1-p_1)^{n_1-c}\binom{n_2}{t-c} p_2^{t-c}(1-p_2)^{n_2-(t-c)}}$$

$$= \frac{\binom{n_1}{c}\left(\frac{p_1}{1-p_1}\right)^c\left(\frac{1-p_2}{p_2}\right)^c\binom{n_2}{t-c}}{\sum_{\text{all } c}\binom{n_1}{c}\binom{n_2}{t-c}\left(\frac{p_1}{1-p_1}\right)^c\left(\frac{1-p_2}{p_2}\right)^c}, \text{ recall } \psi = \log\left[\frac{\frac{p_1}{1-p_1}}{\frac{p_2}{1-p_2}}\right]$$

$$= \frac{\binom{n_1}{c}\binom{n_2}{t-c}\left[e^{\psi}\right]^c}{\sum_{\text{all } c}\binom{n_1}{c}\binom{n_2}{t-c}\left[e^{\psi}\right]^c}, \text{ a non-central hypergeometric distribution}$$

$$= \frac{\binom{n_1}{c}\binom{n_2}{t-c}}{\binom{n_1+n_2}{t}}, \text{ a central hypergeometric distribution when } \psi = 0 \text{ (under } H_0)$$

so that
$$\phi(x) = \begin{cases} 1 & x > c_\alpha \\ \gamma & x = c_\alpha \\ 0 & x < c_\alpha \end{cases}$$

where γ, c_α chosen so that

$$\alpha = \sum_{c=c_\alpha+1}^{\text{upper limit of } x} P(X=c \text{ under } H_0) + \gamma P(X=c_\alpha \text{ under } H_0), \text{ where}$$

$P(X=c$ under $H_0) =$

The p-value is the probability of seeing something as great or greater than the observed x under H_0, where $X+Y=t$ also observed.

$$P_{H_0}(X \geq x) = \sum_{X=x}^{\text{upper limit}} \frac{\binom{n_1}{x}\binom{n_2}{t-x}}{\binom{n_1+n_2}{t}}$$

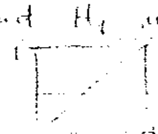
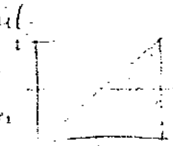
(d) the power function of the test, as a function of ψ is given as:

$$P_{\psi}(X > c_\alpha) + \gamma P_{\psi}(X = c_\alpha) = \sum_{x=c_\alpha+1}^{\text{upper limit}}\left\{\frac{\binom{n_1}{x}\binom{n_2}{t-x}\left[e^{\psi}\right]^x}{\sum_{\text{all } c}\binom{n_1}{c}\binom{n_2}{t-c}\left[e^{\psi}\right]^c}\right\} + \gamma\,\frac{\binom{n_1}{c_\alpha}\binom{n_2}{t-c_\alpha}\left[e^{\psi}\right]^{c_\alpha}}{\sum_{\text{all } c}\binom{n_1}{c}\binom{n_2}{t-c}\left[e^{\psi}\right]^c}$$

(e) Suppose $(n_1=1,\ n_2=2)$ and $(X=0,\ Y=2)$ are observed. Derive the Bayes factor in favor of H_0 in (c), where the joint prior for (p_1, p_2) is a uniform prior with appropriate support under each hypothesis.

$$B = \frac{P(\text{observed} \mid H_0)}{P(\text{observed} \mid H_1)} = \frac{\int_{p_1,p_2} P(\text{observed} \mid p_1,p_2)\, P(p_1,p_2 \mid H_0)\, dp_{1,2}}{\int_{p_1,p_2} P(\text{observed} \mid p_1,p_2)\, \underbrace{P(p_1,p_2 \mid H_1)}_{\text{priors}}\, dp_{1,2}}$$

The uniform priors under H_0 and H_1 must be found in further detail.

Under H_0, $p_1 \leq p_2$ $\Rightarrow$ density:  and under H_1, $p_1 > p_2$ $\Rightarrow$ density: 

a) Under H_0, $f(p_1, p_2) = 2$, $0 \le p_1 \le p_2 \le 1$

Under H_1, $f(p_1, p_2) = 2$, $0 \le p_2 < p_1 \le 1$

The probability of the observed data is given as:

$$p(x = 0, y = 2 \mid p_1, p_2, n_1 = 1, n_2 = 2) = \binom{1}{0}\binom{2}{2} p_1^0 (1-p_1)^1 p_2^2 (1-p_2)^0$$

so that the Bayes factor is equal to

$$\frac{\int_0^1 \int_0^{p_2} 2 \cdot p_2^2 (1-p_1)\, \partial p_1\, \partial p_2 \quad \boxed{1}}{\int_0^1 \int_0^{p_1} 2 p_2^2 (1-p_1)\, \partial p_2\, \partial p_1 \quad \boxed{2}}$$

$\boxed{1}$ $\int_0^1 2p_2^2 \int_0^{p_2} (1-p_1)\, \partial p_1\, \partial p_2 = \int_0^1 2p_2^2 \left[p_1 - \frac{p_1^2}{2} \right]_0^{p_2} \partial p_2$

$= \int_0^1 \left[2p_2^3 - p_2^4 \right] \partial p_2 = \left[\frac{1}{2} p_2^4 - \frac{1}{5} p_2^5 \right]_0^1 = \frac{1}{2} - \frac{1}{5} = \left(\frac{3}{10} \right)$

$\boxed{2}$ $\int_0^1 2(1-p_1) \int_0^{p_1} p_2^2\, \partial p_2\, \partial p_1 = \int_0^1 2(1-p_1) \left[\frac{1}{3} p_2^3 \right]_0^{p_1} \partial p_1$

$= \int_0^1 \frac{2}{3} p_1^3 - \frac{2}{3} p_1^4\, \partial p_1 = \left[\frac{1}{6} p_1^4 - \frac{2}{15} p_1^5 \right]_0^1 = \frac{1}{6} - \frac{2}{15} = \left(\frac{1}{30} \right)$

So that the Bayes factor in favor of H_0 is computed as:

$B = \frac{\left(\frac{3}{10}\right)}{\left(\frac{1}{30}\right)} = \frac{\left(\frac{9}{30}\right)}{\left(\frac{1}{30}\right)} = 9$, somewhat strong evidence in favor of H_0 (according to Jeffreys' criteria)

2008 QUESTION 3

TRUE $\quad Y_{n\times 1} = \underset{\text{rank}(X)=p}{X_{n\times p}}\beta_{p\times 1} + Z_{n\times q}\gamma_{q\times 1} + \varepsilon_{n\times 1}, \quad \varepsilon \sim N_n(0, \sigma^2 I_n) \qquad (1)$

NAIVE $\quad Y_{n\times 1} = \underset{\text{rank}(X)=p}{X_{n\times p}}\beta_{p\times 1} + \varepsilon_{n\times 1}, \quad \varepsilon \sim N_n(0, \sigma^2 I_n) \qquad (2)$

Assume σ^2 unknown. Investigators obtain $\hat\sigma^2$ and $\hat\beta$, the least squares estimates from (2).

(a) $\hat\beta = (X'X)^{-1}X'Y$

(i) $E(\hat\beta) = (X'X)^{-1}X'[X\beta + Z\gamma] = \beta + \underbrace{(X'X)^{-1}X'Z\gamma}_{\text{Bias}}$

(ii) As $(X'X)^{-1}$, γ are non-zero, $X'Z = 0$ for the bias of $\hat\beta$ to be zero. $X'Z \Rightarrow X \perp Z$ or X and Z are orthogonal

(b)(i) $Var(\hat\beta) = \sigma^2(X'X)^{-1}$

(ii) $\hat\beta \sim N_p\left(\beta + (X'X)^{-1}X'Z\gamma,\ \sigma^2(X'X)^{-1}\right)$

(c) $\hat\sigma^2 = \frac{1}{n-p}Y'(I-M)Y$. As $\frac{1}{\sigma^2}Y'(I-M)Y \sim \chi^2_{n-p,\ \delta}$, $\delta = \frac{1}{2\sigma^2}(Z\gamma)'(I-M)Z\gamma$

$$E(\hat\sigma^2) = \frac{\sigma^2}{n-p}E\left[\frac{1}{\sigma^2}Y'(I-M)Y\right] = \frac{\sigma^2}{n-p}\left\{n-p + \frac{(Z\gamma)'(I-M)Z\gamma}{\sigma^2}\right\}$$

$$= \sigma^2 + \frac{(Z\gamma)'(I-M)(Z\gamma)}{n-p}$$

(d) $\hat\beta$ has the same form in models (1) and (2). Conditions exist with zero bias:
$\hat\beta \sim N_p(\beta, \sigma^2(X'X)^{-1})$.
If $\hat\beta$ is fit using (2) when (1) is true, the LSE $\hat\sigma^2$ is biased. $\Rightarrow$ inflated variance estimate
If $\hat\beta$ is fit using (1), the LSE $\hat\sigma_1^2$ is unbiased, and the estimated variance of $\hat\beta$ is not inflated.

(e) observe (Y^*, X^*, Z^*) from the following model

$Y^*_{m\times 1} = \underset{\text{rank}(X^*)=p}{X^*_{m\times p}}\beta^*_{p\times 1} + \underset{\text{rank}(Z^*)=q}{Z^*_{m\times q}}\gamma^*_{q\times 1} + \varepsilon^*_{m\times 1}, \quad \varepsilon^* \sim N_m(0, \sigma^2_* I_m) \qquad (3)$

rank(X) $\qquad$ rank(Z)

Test $H_0: \beta = \beta^*$ and $\gamma = \gamma^*$ [ie additional m come from same model as (1)]

Under the full unrestricted model, the m+n observations can be represented as

$$\begin{bmatrix} Y_{n\times 1} \\ Y^*_{m\times 1} \end{bmatrix} = \underbrace{\begin{bmatrix} X_{n\times p} & 0 \\ 0 & X^*_{m\times p} \end{bmatrix}}_{X} \begin{bmatrix} \beta_{p\times 1} \\ \beta^*_{p\times 1} \end{bmatrix} + \underbrace{\begin{bmatrix} Z_{n\times q} & 0 \\ 0 & Z^*_{m\times q} \end{bmatrix}}_{Z} \begin{bmatrix} \gamma_{q\times 1} \\ \gamma^*_{q\times 1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{n\times 1} \\ \varepsilon^*_{m\times 1} \end{bmatrix} \quad \text{(FULL)}$$

Under the null, a nested model exists as follows:

$$\begin{bmatrix} Y_{n\times 1} \\ Y^*_{m\times 1} \end{bmatrix} = \underbrace{\begin{bmatrix} X_{n\times p} \\ X^*_{m\times p} \end{bmatrix}}_{X_0} \begin{bmatrix} \beta_{p\times 1} \end{bmatrix} + \underbrace{\begin{bmatrix} Z_{n\times q} \\ Z^*_{m\times q} \end{bmatrix}}_{Z_0} \begin{bmatrix} \gamma_{q\times 1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{n\times 1} \\ \varepsilon^*_{m\times 1} \end{bmatrix} \quad \text{(NESTED)}$$

let M be the orthogonal projection operator onto $C(X, Z)$, and let M_0 be the orthogonal projection operator onto $C(X_0, Z_0)$. to test the hypotheses of interest, consider

$$F = \frac{\left[\dfrac{Y'(M-M_0)Y}{r(M-M_0)}\right]}{\left[\dfrac{Y'(I-M)Y}{r(I-M)}\right]} \begin{array}{l} \sim F_{r(M-M_0),\, r(I-M)} \quad \text{under } H_0 \\ \sim F_{r(M-M_0),\, r(I-M),\, \delta = \frac{E(Y)'(M-M_0)E(Y)}{2\sigma^2}} \quad \text{under } H_A \end{array}$$

If $F > F_{r(M-M_0),\, r(I-M),\, 1-\alpha}$, reject H_0.

2008 QUESTION 4

$(X_1, Y_1), \ldots, (X_n, Y_n)$ iid. X_i is continuous, Y_i a count.
$a_i \sim N(0, \sigma^2)$. $Y_i|a_i \perp X_i|a_i$ and $Y_i|a_i \sim \text{Poisson}(\lambda e^{a_i})$ and
$X_i|a_i \sim N(a_i, 1)$. σ^2 unknown. Estimation of λ of interest.

$$f(X_i, Y_i) = \int_{\text{over } a_i} f(X, Y, a_i)\,da_i = \int_{a_i} f(X, Y|a_i) f(a_i)\,da_i$$
$$= \int_{a_i} f(X|a_i) f(Y|a_i) f(a_i)\,da_i \quad (\text{as } X|a_i \perp Y|a_i)$$

$$= \int_{-\infty}^{\infty} \overset{x}{N(a_i, 1)} \cdot \overset{y}{\text{Poisson}(\lambda e^{a_i})} \overset{a}{N(0, \sigma^2)}\,da_i$$

$$\int_{-\infty}^{\infty} (2\pi)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(x-a_i)^2\right\} \frac{(\lambda e^{a_i})^y}{y!} e^{-\lambda e^{a_i}} (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left\{\frac{-1}{2\sigma^2} a_i^2\right\} da_i$$
(continued later in (c))

(a) Method of Moments

(i) Find UMVUE of σ^2 using $(X_1, \ldots, X_n)$

$$f(X) = \int_{a_i} f(X|a_i) f(a_i) = \int_{-\infty}^{\infty} (2\pi)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(X-a_i)^2\right\} (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left\{\frac{-1}{2\sigma^2} a_i^2\right\} da_i$$

$$= (2\pi\sigma^2)^{-\frac{1}{2}} \int_{-\infty}^{\infty} (2\pi)^{-\frac{1}{2}} \exp\left\{\frac{-1}{2}\left[X^2 - 2a_i X + a_i^2 + \frac{a_i^2}{\sigma^2}\right]\right\} da_i$$

$$\exp\left\{-\frac{1}{2}\left[a_i^2 \underbrace{\left(1 + \frac{1}{\sigma^2}\right)}_{\frac{\sigma^2+1}{\sigma^2}} - 2a_i X + X^2\right]\right.$$

$$= \exp\left\{\frac{-1}{2\left[\frac{\sigma^2}{\sigma^2+1}\right]}\left[a_i^2 - 2a_i X\left(\frac{\sigma^2}{\sigma^2+1}\right) + X^2\left(\frac{\sigma^2}{\sigma^2+1}\right)^2\right]\right.$$
$$\left.\frac{-1}{2}\left[X^2\left(1 - \frac{\sigma^2}{\sigma^2+1}\right)\right]\right\}$$

$$= [2\pi(\sigma^2+1)]^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\left[X^2\left(\frac{1}{\sigma^2+1}\right)\right]\right\} \underbrace{\int_{-\infty}^{\infty} \left[2\pi\left(\frac{\sigma^2}{\sigma^2+1}\right)\right]^{-\frac{1}{2}} \exp\left\{\frac{-1}{2\left[\frac{\sigma^2}{\sigma^2+1}\right]}\left(a_i - X\left(\frac{\sigma^2}{\sigma^2+1}\right)\right)^2\right\} da_i}_{=1}$$

$$= [2\pi(\sigma^2+1)]^{-\frac{1}{2}} \exp\left\{\frac{-1}{2(\sigma^2+1)} X_i^2\right\}$$

$$\Rightarrow X_i \sim N(0, \sigma^2+1)$$

$\Rightarrow X_1, \ldots, X_n$ iid $N(0, \sigma^2+1) \longrightarrow \Rightarrow \frac{X_i}{\sqrt{\sigma^2+1}} \sim N(0, 1)$

$$l(\sigma^2 | \underset{\sim}{X}) = [2\pi(\sigma^2+1)]^{-\frac{n}{2}} \exp\left\{\frac{-1}{2(\sigma^2+1)} \underbrace{\sum_{i=1}^{n} X_i^2}_{\sum^n X_i^2}\right\}$$

$\sum^n X_i^2$ complete sufficient for $\frac{1}{2(\sigma^2+1)}$

$E\left[\sum^n X_i^2\right] = n E(X_i^2) = n(\sigma^2+1)$ $n\sigma^2 + n$

$\therefore \frac{1}{n}\sum X_i^2 - 1$ is the UMVUE of σ^2 denoted $\hat{\sigma}^2$

$a_i \sim N(0, \sigma^2)$. Find dist of e^{a_i}: $j_i \Rightarrow a_i = \log j_i$

$\uparrow$

$MGF = E(e^{ta_i}) = \exp\left(\frac{\sigma^2 t^2}{2}\right)$

$= E(j_i^t) = \exp\left(\frac{\sigma^2 t^2}{2}\right) \Rightarrow E(j_i^1) = e^{\frac{\sigma^2}{2}}, \quad E(j_i^2) = e^{2\sigma^2}$

$\Rightarrow Var(j_i) = e^{2\sigma^2} - e^{\sigma^2}$

(ii) $E(Y_i) = E_{a_i} E\{Y_i | a_i\}$

$= E_{a_i}(\lambda e^{a_i}) = \lambda E(e^{a_i}) = \lambda e^{\frac{\sigma^2}{2}}$

$a_i \sim N(0, \sigma^2)$

~~$\Rightarrow E(e^{a_i}) = \int_{-\infty}^{\infty} (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left\{\underbrace{-\frac{1}{2\sigma^2} a_i^2 + a_i}_{= -\frac{1}{2\sigma^2}[a_i^2 - a_i\sigma^2]}\right\}$ $(a_i^2 - a_i\sigma^2 + \frac{\sigma^4}{4}) \cdot (a_i - \frac{\sigma^2}{2})^2$~~

~~$= \int_{-\infty}^{\infty} (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left\{\frac{-1}{2\sigma^2}\left[a_i - \frac{\sigma^2}{2}\right]^2 + \frac{\sigma^2}{8}\right\}$~~

~~$= \left[e^{\frac{\sigma^2}{8}}\right]$~~

$Var(Y_i) = E_{a_i} Var(Y_i | a_i) + Var_{a_i} E(Y_i | a_i)$

$= E_{a_i}(\lambda e^{a_i}) + Var_{a_i}(\lambda e^{a_i})$

$= \lambda e^{\frac{\sigma^2}{2}} + \lambda^2 Var_{a_i}(e^{a_i})$

$= \lambda e^{\frac{\sigma^2}{2}} + \lambda^2 [e^{2\sigma^2} - e^{\sigma^2}]$

(iii) let $g(\lambda, \sigma^2)$ denote $E(Y_i)$.

Solve the following to find moment estimator of λ, $\hat{\lambda}$.

$\frac{1}{n}\sum_{i=1}^{n} Y_i = g(\lambda, \hat{\sigma}^2) = \lambda e^{\frac{\hat{\sigma}^2}{2}}$

$\Rightarrow \hat{\lambda}_1 = e^{-\frac{\hat{\sigma}^2}{2}} \cdot \bar{y}$, where $\hat{\sigma}^2 = \overline{x^2} - 1$

$= e^{-\frac{1}{2}(\overline{x^2} - 1)} \cdot \bar{y}$

(iv) Can use delta method to derive asymptotic dist. of $\hat{\lambda}_1$.

$E(y) = \lambda e^{\frac{\sigma^2}{2}}$, $Var(y) = \lambda e^{\frac{\sigma^2}{2}} + \lambda^2 [e^{2\sigma^2} - e^{\sigma^2}]$

$E(x^2) = \sigma^2 + 1$, $Var(x^2) = E(x^4) - [E(x^2)]^2$

$= 3(\sigma^2+1)^2 - (\sigma^2+1)^2 = 2(\sigma^2+1)^2$

$Cov(x^2, y) = E(x^2 y) - E(x^2)E(y)$

$E(x^2 y) = E_{a_i} E(x^2 y | a)$

$= E_{a_i} E(x^2|a) \cdot E(y|a) = E_{a_i}\left(\lambda e^{a_i}(1 + a_i^2)\right)$

$= \lambda[E(e^{a_i}) + E(a_i^2 e^{a_i})]$

$= \lambda[e^{\frac{\sigma^2}{2}} + e^{\frac{\sigma^2}{2}}(\sigma^2 + \sigma^4)]$

$\lambda e^{\frac{\sigma^2}{2}}(\sigma^4 + \sigma^2 + 1)$

$E(a_i^2 e^{a_i})$, where $a_i \sim N(0, \sigma^2)$

$\int (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left\{-\frac{a_i^2}{2\sigma^2} + a_i\right\} a_i^2 \, da_i$

$= \int a_i^2 (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma^2}[a_i^2 - 2a_i\sigma^2 + \sigma^4] + \frac{\sigma^2}{2}\right\}$

$= e^{\frac{\sigma^2}{2}} \cdot \int a_i^2 \cdot \underbrace{N(\sigma^2, \sigma^2)}_{= \sigma^2 + \sigma^4} \, da_i \qquad e^{\frac{\sigma^2}{2}}(\sigma^2 + \sigma^4)$

By multivariate CLT,

$\sqrt{n}\begin{bmatrix} \overline{x^2} - E(x^2) \\ \bar{y} - E(y) \end{bmatrix} \xrightarrow{d} N_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma = \begin{bmatrix} Var(x^2) & Cov(x^2, y) \\ & Var(y) \end{bmatrix}\right)$

$$[a \quad b]\begin{bmatrix} i & j \\ j & k \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = [ai+bj \quad aj+bk]\begin{bmatrix} a \\ b \end{bmatrix}$$

$$= a^2 i + abj + abj + b^2 k$$

$$= a^2 i + 2abj + b^2 k$$

Write $\hat{\lambda} = e^{-\frac{1}{2}(\overline{x^2}-1)} \cdot \bar{y} = g(x,y)$, where $g(x,y) = e^{-\frac{1}{2}(x-1)} \cdot y$

Then, $g'(x,y) = \left[-\frac{1}{2} e^{-\frac{1}{2}(x-1)} \cdot y \quad e^{-\frac{1}{2}(x-1)}\right]$

By delta method, $\sqrt{n}\left[\hat{\lambda}_1 - g\underbrace{\left(E(x^2), E(Y)\right)}_{=\lambda}\right] \xrightarrow{d} N_1(0, \Sigma^*)$

① $g\left[E(x^2), E(Y)\right] = e^{-\frac{1}{2}(E(x^2)-1)} E(y)$
$= e^{-\frac{1}{2}(\sigma^2+1-1)} \lambda e^{\frac{\sigma^2}{2}} = \lambda$

② $\Sigma^* = g'\left[E(x^2), E(y)\right] \cdot \begin{bmatrix} Var(x^2) & cov(x^2, y) \\ cov(x^2, y) & Var(y) \end{bmatrix} \left\{ g'\left[E(x^2), E(y)\right] \right\}^T$

$$= \left[-\frac{\lambda}{2} \quad e^{-\frac{\sigma^2}{2}}\right] \begin{bmatrix} 2(\sigma^2+1)^2 & \lambda e^{\frac{\sigma^2}{2}}(\sigma^4+\sigma^2+1) \\ \lambda e^{\frac{\sigma^2}{2}}(\sigma^4+\sigma^2+1) & \lambda e^{\frac{\sigma^2}{2}} + \lambda^2\left[e^{2\sigma^2} - e^{\sigma^2}\right] \end{bmatrix} \begin{bmatrix} -\frac{\lambda}{2} \\ e^{-\frac{\sigma^2}{2}} \end{bmatrix}$$

$$= \frac{\lambda^2}{4}\left[2(\sigma^2+1)^2\right] - \lambda e^{-\frac{\sigma^2}{2}}\left[\lambda e^{\frac{\sigma^2}{2}}(\sigma^4+\sigma^2+1)\right] + e^{-\sigma^2}\left\{\lambda e^{\frac{\sigma^2}{2}} + \lambda^2\left[e^{2\sigma^2} - e^{\sigma^2}\right]\right\}$$

$$= \frac{\lambda^2}{2}(\sigma^2+1)^2 - \lambda^2(\sigma^4+\sigma^2+1) + \lambda e^{-\frac{\sigma^2}{2}} + \lambda^2(e^{\sigma^2}-1)$$

$$= \lambda e^{-\frac{\sigma^2}{2}} + \lambda^2\left\{\frac{1}{2}(\sigma^2+1)^2 - (\sigma^4+\sigma^2+1) + (e^{\sigma^2}-1)\right\}$$

(b) a_i missing data. Derive an EM algorithm for computing MLEs of λ and σ^2. Denote MLE of λ as $\hat{\lambda}_2$.

First, need complete data likelihood, $f(\underset{\sim}{x_i}, \underset{\sim}{Y_i}, \underset{\sim}{a_i})$

$\therefore L^c = \prod_{i=1}^n f(x_i, y_i, a_i) = \prod_{i=1}^n f(x_i, y_i | a_i) f(a_i)$
$\prod_{i=1}^n f(x_i|a_i) \cdot f(y_i|a_i) \cdot f(a_i)$ [as $X_i|a_i \perp Y_i|a_i$

$= \prod_{i=1}^n (2\pi)^{-\frac{1}{2}} exp\left\{-\frac{1}{2}(x_i-a_i)^2\right\} (2\pi\sigma^2)^{-\frac{1}{2}} exp\left\{-\frac{1}{2\sigma^2} a_i^2\right\} \cdot \frac{(\lambda e^{a_i})^{y_i}}{y_i!} exp\left\{-\lambda e^{a_i}\right\}$

$\prod_{i=1}^n (4\pi^2\sigma^2)^{-\frac{1}{2}} \frac{\lambda^{y_i}}{y_i!} exp\left\{a_i y_i - \lambda e^{a_i} - \frac{1}{2}(x_i-a_i)^2 - \frac{a_i^2}{2\sigma^2}\right\}$

so that the log likelihood of the complete data is:

$\ell^c = -\frac{n}{2}\log(4\pi^2\sigma^2) + \Sigma y_i(\log\lambda) - \Sigma\log(y_i!) + \Sigma a_i y_i - \lambda\Sigma e^{a_i} - \frac{1}{2}\Sigma(x_i - a_i)^2 - \frac{1}{2\sigma^2}\Sigma a_i^2$

And the score vector for the complete data (for λ) is

$\frac{\partial \ell^c}{\partial \lambda} = \frac{1}{\lambda}\Sigma y_i - \Sigma e^{a_i} = 0 \Rightarrow \frac{1}{\lambda}\Sigma y_i = \Sigma e^{a_i} \Rightarrow \hat{\lambda}_2 = \frac{\Sigma y_i}{\Sigma e^{a_i}}$

The problem is that Σe^{a_i} is unobserved. Can find conditional expectation of e^{a_i} given observed, or $E\left[e^{a_i} | X_i, Y_i\right]$

$$E[e^{a_i}(X_i, Y_i)] = \frac{\int_{-\infty}^{\infty} e^{a_i}(4\pi^2\sigma^2)^{-\frac{1}{2}} \frac{\lambda^{y_i}}{y_i!} \exp\{a_i y_i - \lambda e^{a_i} - \frac{1}{2}(x_i - a_i)^2 - \frac{a_i^2}{\sigma^2}\} \, da_i}{\int_{-\infty}^{\infty} (4\pi^2\sigma^2) \frac{\lambda^{y_i}}{x_i!} \exp\{a_i y_i - \lambda e^{a_i} - \frac{1}{2}(x_i - a_i)^2 - \frac{a_i^2}{\sigma^2}\} \, da_i}$$

$$= \frac{\int_{-\infty}^{\infty} \exp\{a_i(y_i+1) - \lambda e^{a_i} - \frac{1}{2}(x_i-a_i)^2 - \frac{a_i^2}{\sigma^2}\} \, da_i}{\int_{-\infty}^{\infty} \exp\{a_i y_i - \lambda e^{a_i} - \frac{1}{2}(x_i-a_i)^2 - \frac{a_i^2}{\sigma^2}\} \, da_i}$$

let $j_i = e^{a_i}$

$\Rightarrow a_i = \log j_i$

$$= \frac{\int_{-\infty}^{\infty} \exp\{a_i(y_i+1) - \lambda e^{a_i} - \frac{1}{2}(x_i^2 - 2a_i x_i + a_i^2) - \frac{a_i^2}{\sigma^2}\} \, da_i}{\int_{-\infty}^{\infty} \exp\{\frac{a_i^2}{2}\underbrace{(\frac{-1}{\sigma^2} - 1)}_{\frac{\sigma^2+1}{\sigma^2}} + a_i(y_i + x_i) - \lambda e^{a_i} - \frac{1}{2}x_i^2\} \, da_i}$$

$$= \frac{\int_{-\infty}^{\infty} \exp\{\frac{-a_i^2}{2}(+\frac{1}{\sigma^2}+1) + a_i(y_i + x_i + 1) - \lambda e^{a_i}\} \, da_i}{\int_{-\infty}^{\infty} \exp\{\frac{-a_i^2}{2}(\frac{1}{\sigma^2}+1) + a_i(y_i + x_i) - \lambda e^{a_i}\} \, da_i}$$

$$\frac{\exp\{\frac{-(y_i+x_i+1)^2}{2[\frac{\sigma^2}{\sigma^2+1}]}\}}{\exp\{\frac{-(y_i+x_i)^2}{2[\frac{\sigma^2}{\sigma^2+1}]}\}} \cdot \frac{\int_{-\infty}^{\infty} [2\pi(\frac{\sigma^2}{\sigma^2+1})]^{-\frac{1}{2}} \exp\{\frac{-1}{2[\frac{\sigma^2}{\sigma^2+1}]}(a_i - (y_i+x_i+1))^2 \cdot e^{(-\lambda e^{a_i})}\} \, da_i}{\int_{-\infty}^{\infty} [2\pi(\frac{\sigma^2}{\sigma^2+1})]^{-\frac{1}{2}} \exp\{\frac{-1}{2[\frac{\sigma^2}{\sigma^2+1}]}(a_i - (y_i+x_i))^2 \cdot e^{(-\lambda e^{a_i})}\} \, da_i}$$

$$= \exp\{\frac{-1}{2[\frac{\sigma^2}{\sigma^2+1}]}[(y_i+x_i+1)^2 - (y_i+x_i)^2]\} \cdot \frac{E[e^{-\lambda e^{a_i}}] \text{ where } a_i \sim N(y_i+x_i+1, \frac{\sigma^2}{\sigma^2+1})}{E[e^{-\lambda e^{a_i}}] \text{ where } a_i \sim N(y_i+x_i, \frac{\sigma^2}{\sigma^2+1})}$$

...

$$\hat{\lambda}_1^{(t+1)} = \frac{\sum y_i}{\sum E[e^{a_i} | X_i, Y_i, \lambda^{(t)}, \sigma^{2(t)}]}$$

(c) Conditional likelihood. Treat a_i as a parameter (also, λ and e^{a_i})

(i) $f(X_i, Y_i \mid a_i) = f(X_i|a_i) \cdot f(Y_i|a_i)$ [by independence assumption]

$$= (2\pi)^{-\frac{1}{2}} \exp\{-\frac{1}{2}(x_i - a_i)^2\} \cdot \frac{\lambda^{y_i} e^{a_i y_i}}{y_i!} \cdot \exp\{-\lambda e^{a_i}\}$$

$$= (2\pi)^{-\frac{1}{2}} \frac{\lambda^{y_i}}{y_i!} \exp\{-\frac{1}{2}(x_i^2 - 2a_i x_i + a_i^2) + a_i y_i - \lambda e^{a_i}\}$$

$$= (2\pi)^{-\frac{1}{2}} \frac{\lambda^{y_i}}{y_i!} \exp\{-\frac{1}{2}x_i^2 + a_i x_i - a_i^2 + a_i y_i - \lambda e^{a_i}\}$$

$$= (2\pi)^{-\frac{1}{2}} \frac{\lambda^{y_i}}{y_i!} \exp\{(x_i + y_i)a_i - a_i^2 - \lambda e^{a_i} - \frac{1}{2}x_i^2\}$$

sufficient statistic for a_i (treated as nuisance)

(ii) Conditional likelihood. product of $f(Y_i | X_i + Y_i)$ over $i = 1, \ldots, n$.

$$f(Y_i | X_i + Y_i) = \frac{f(Y_i, X_i + Y_i)}{f(X_i + Y_i)} = \frac{P(Y_i = c, X_i + Y_i = t)}{P(X_i + Y_i = t)}$$

$$= \frac{P(Y_i = c, X_i = t - c)}{P(X_i + Y_i = t)}$$

$$= \frac{P(Y_i = c)\cdot P(X_i = t-c)}{P(X_i + Y_i = t)} = \frac{P(Y_i = c)\cdot P(X_i = t-c)}{\sum_{c=0}^{t} P(Y_i = c)\cdot P(X_i = t-c)}$$

$$= \frac{\frac{(\lambda e^{a_i})^c}{c!}\cancel{\exp\{-\lambda e^{a_i}\}}\cdot \cancel{(2\pi)^{-\frac{1}{2}}}\exp\{-\frac{1}{2}(t-c-a_i)^2\}}{\sum_{c=0}^{t}\frac{(\lambda e^{a_i})^c}{c!}\cancel{\exp\{-\lambda e^{a_i}\}}\cdot\cancel{(2\pi)^{-\frac{1}{2}}}\exp\{-\frac{1}{2}(t-c-a_i)^2\}}$$

$$= \frac{\frac{\lambda^c}{c!}\exp\{ca_i - \cancel{\frac{1}{2}(t-a_i)^2} + (t-a_i)c - \frac{1}{2}c^2\}}{\sum_{c=0}^{t}\frac{\lambda^c}{c!}\exp\{ca_i - \cancel{\frac{1}{2}(t-a_i)^2} + (t-a_i)c - \frac{1}{2}c^2\}}$$

$$P(Y_i = c \mid X_i + Y_i = t) = \frac{\frac{\lambda^c}{c!}\exp\{-\frac{1}{2}c^2 + ct\}}{\sum_{c=0}^{t}\frac{\lambda^c}{c!}\exp\{-\frac{1}{2}c^2 + ct\}}$$

(*) Able to eliminate a_i...

Can rewrite as

$$P(Y_i = y_i \mid X_i + Y_i = t_i) = \frac{\frac{\lambda^{y_i}}{y_i!}\exp\{-\frac{1}{2}y_i^2 + y_i t_i\}}{\sum_{y_i=0}^{t_i}\frac{\lambda^{y_i}}{y_i!}\exp\{-\frac{1}{2}y_i^2 + y_i t_i\}}$$

Then the joint conditional $\overset{\log}{\vee}$ likelihood is given as

$\ell^{cond} = \sum \ell_i$, where $\ell_i = \log P(Y_i = y_i \mid X_i + Y_i = t_i)$

$$= y_i \log\lambda - \log y_i! - \frac{1}{2}y_i^2 + y_i t_i - \log\left[\sum_{y_i=0}^{t_i}\frac{\lambda^{y_i}}{y_i!}\exp\{-\frac{1}{2}y_i^2 + y_i t_i\}\right]$$

$$\Rightarrow \ell^{cond} = \log\lambda\sum y_i - \sum\log y_i! - \frac{1}{2}\sum y_i^2 + \sum y_i t_i - \sum\log\left[\sum_{y_i=0}^{t_i}\frac{\lambda^{y_i}}{y_i!}\exp\{-\frac{1}{2}y_i^2 + y_i t_i\}\right]$$

$$\Rightarrow \frac{\partial\ell}{\partial\lambda} = \frac{1}{\lambda}\sum y_i - \sum\left[\frac{1}{\sum_{y_i=0}^{t_i}\frac{\lambda^{y_i}}{y_i!}\exp\{-\frac{1}{2}y_i^2 + y_i t_i\}}\sum_{y_i=0}^{t_i}\frac{\lambda^{y_i-1}}{(y_i-1)!}\exp\{-\frac{1}{2}y_i^2 + y_i t_i\}\right]$$

$$= \frac{1}{\lambda}\sum y_i - \sum\left[1 - \frac{\frac{\lambda^{t_i}}{t_i!}\exp\{-\frac{1}{2}t_i^2 + t_i^2\}}{\sum_{y_i=0}^{t_i}\frac{\lambda^{y_i}}{y_i!}\exp\{-\frac{1}{2}y_i^2 + y_i t_i\}}\right] = 0$$

$$= \frac{1}{\lambda}\sum y_i - \sum\left[\frac{\frac{\lambda^{t_i}}{t_i!}\exp\{\frac{1}{2}t_i^2\}}{\sum_{y_i=0}^{t_i}\frac{\lambda^{y_i}}{y_i!}\exp\{-\frac{1}{2}y_i^2 + y_i t_i\}}\right] = n$$

(*) $\hat{\lambda}_3$ solves this equation

(d) advantages / disadvantages of $\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3$

$\hat{\lambda}_1$. Not as good because didn't use model assumptions (if model true, least efficient)

$\hat{\lambda}_2$ will converge to MLE. efficient

$\hat{\lambda}_3$. Good as $X_i + Y_i$ is s-ancillary maybe p-ancillary (more eff. than $\hat{\lambda}_1$, potentially less than $\hat{\lambda}_2$)

→ comparisons of efficiency and ease of computation

(b) $\frac{\partial}{\partial \sigma^2} = \frac{-n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum a_i^2 = 0$

$\Rightarrow \hat{\sigma}^2 = \frac{1}{n} \sum a_i^2$

$\Rightarrow \hat{\sigma}^{2(t+1)} = \frac{1}{n} \sum E\left(a_i^2 \mid x_i, y_i, \lambda^{(t)}, \sigma^{2(t)}\right)$

2008 QUESTION 5

$\{X(t), t \geq 0\}$ Poisson process with mean $\lambda_1 t$ independent of

$\{Y(t), t \geq 0\}$ Poisson process with mean $\lambda_2 t$. $Z(t) = X(t) - Y(t)$

(i)

(a) $P_n(t) = P[Z(t) = n]$, for $n = 0, \pm 1, \pm 2, \ldots$

Want to find $\phi_Z(s) = E[s^Z] = E[s^{X-Y}]$

$= E[s^X] \cdot E[s^{-Y}]$ (*) KEY STEP: USE INDEPENDENCE OF X, Y

$= \phi_X(s) \cdot \phi_Y(\frac{1}{s})$

where $\phi_X(s) = \sum_{x=0}^{\infty} s^x \cdot \frac{(\lambda_1 t)^x}{x!} \cdot e^{-\lambda_1 t}$

$= \sum_{x=0}^{\infty} \frac{(\lambda_1 t s)^x}{x!} e^{-\lambda_1 t} = e^{-\lambda_1 t} \cdot e^{\lambda_1 t s} = e^{\lambda_1 t(s-1)}$

and $\phi_Y(\frac{1}{s}) = e^{\lambda_2 t(\frac{1}{s}-1)}$

Then $\phi_Z(s) = \phi_X(s) \cdot \phi_Y(\frac{1}{s}) = e^{\lambda_1 t(s-1)} \cdot e^{\lambda_2 t(\frac{1}{s}-1)}$

$= \exp\{-(\lambda_1+\lambda_2)t\} \cdot \exp\{\lambda_1 t \cdot s + \lambda_2(\frac{t}{s})\}$

ALSO, IMPORTANT TO RECOGNIZE

$\phi_Z(s) = E(s^Z) = \sum_{z=-\infty}^{\infty} P[Z(t)=z] \cdot s^z$

(ii) Note that the j^{th} factorial moment is given by

$E[Z(t)^{(j)}] = \frac{\partial^{(j)}}{\partial s} \phi_Z(s)_{|s=1}$

$\Rightarrow E[Z(t)^{(1)}] = \frac{\partial}{\partial s} \phi_Z(s)_{|s=1}$

$= \exp\{-(\lambda_1+\lambda_2)t\} \cdot \exp\{\lambda_1 t \cdot s + \lambda_2(\frac{t}{s})\} \cdot [\lambda_1 t - \lambda_2 t s^{-2}]_{|s=1}$

$= t(\lambda_1 - \lambda_2)$

$E[Z(t)(Z(t)-1)] = \frac{\partial^2}{\partial s^2} \phi_Z(s)_{|s=1} = \exp\{-(\lambda_1+\lambda_2)t\} \cdot \Big[\exp\{\lambda_1 t \cdot s + \lambda_2(\frac{t}{s})\} \cdot (2\lambda_2 t s^{-3})$

$+ \exp\{\lambda_1 t \cdot s + \lambda_2(\frac{t}{s})\} \cdot (\lambda_1 t - \lambda_2 t s^{-2})^2\Big]_{|s=1}$

$= 2\lambda_2 t + [(\lambda_1 - \lambda_2)t]^2$

$\Rightarrow Var(Z(t)) = 2\lambda_2 t + (\lambda_1-\lambda_2)^2 t^2 + (\lambda_1-\lambda_2)t - \cancel{(\lambda_1-\lambda_2)^2 t^2}$

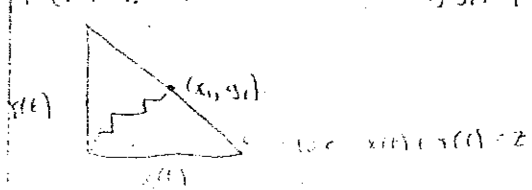
$= (\lambda_1+\lambda_2)t$

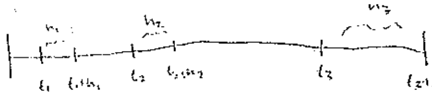
EASIER WAY: As $X(t) \perp Y(t)$, $E[Z(t)] = E[X(t)] - E[Y(t)] = (\lambda_1 - \lambda_2)t$

$Var[Z(t)] = Var[X(t)] + Var[Y(t)] = (\lambda_1+\lambda_2)t$

(b) $[X(t), Y(t)]$ stochastic process in two-dimensional space

P(intersection occurs at (x_1, y_1) | intersects $x+y=z$)





(*) length of time with no events is $t - h_1 - h_2 - h_3$

Note $X(t)+Y(t) \sim \text{Poisson}[(\lambda_1+\lambda_2)t]$

$$P[X(t)=x_1 \mid X(t)+Y(t) = x_1+y_1 = z]$$

$$= \frac{P[X(t)=x,\ X(t)+Y(t)=z]}{P[X(t)+Y(t)=z]} = \frac{P[X(t)=x]\cdot P[Y(t)=z-x]}{P[X(t)+Y(t)=z]}$$

if don't know dist, can sum numerator over $x \in [0,\dots,z]$

$$= \frac{\frac{[\lambda_1 t]^x}{x!} e^{-\lambda_1 t} \cdot \frac{[\lambda_2 t]^{z-x}}{(z-x)!} e^{-\lambda_2 t}}{\frac{[(\lambda_1+\lambda_2)t]^z}{z!} e^{-(\lambda_1+\lambda_2)t}}$$

$$= \frac{z!}{x!(z-x)!}\left[\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)t\right]^x \left[\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)t\right]^{z-x}, \quad x \in [0,\dots,z]$$

$$\Rightarrow X \sim \text{Bin}\left(z, \left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)t\right)$$

(c) $X(t)$ is a Poisson process with mean $\lambda_1 t$. let Z_n = time between $(n-1)^{st}$ event and n^{th} event, and let $S_n = \sum_{i=1}^n Z_i$. Derive

(i) the distribution of S_n start with Z_1

$$P(Z_1 > t) = P\{0 \text{ events in } [0,t)\}$$
$$= P[X(t)=0] = e^{-\lambda_1 t} \Rightarrow Z_1 \sim \text{Exponential}\left(\text{mean} = \frac{1}{\lambda_1}\right)$$

$$P(Z_2 > t \mid Z_1 = s) = P\{0 \text{ events in } (s, s+t) \mid \text{one event in } (0,s)\}$$
$$P\{0 \text{ events in } (s,s+t)\} = e^{-\lambda_1 t} \Rightarrow Z_2 \sim \text{Exponential}\left(\text{mean} = \frac{1}{\lambda_1}\right)$$

$\Rightarrow Z_1,\dots,Z_n$ iid Exponential $\left(\text{mean} = \frac{1}{\lambda_1}\right)$

$\Rightarrow S_n = \sum_{i=1}^n Z_i \sim \text{Gamma}\left(n, \frac{1}{\lambda_1}\right)$

(ii) let $0 < t_1 < t_2 < \cdots < t_{n+1} = t$, and let h_i be small enough so that $t_i + h_i < t_{i+1}$, $i = 1,\dots,n$. Now,

$$P\{t_i \le S_i \le t_i + h_i,\ i=1,\dots,n \mid X(t) = n\}$$

$$= \frac{P\{\text{one event in each } [t_i, t_i+h_i],\ i=1,\dots,n, \text{ and no events elsewhere in } [0,t]\}}{P\{X(t)=n\}}$$

$$= \frac{(\lambda_1 h_1)e^{-\lambda_1 h_1}\cdots(\lambda_1 h_n)e^{-\lambda_1 h_n} \times e^{-\lambda_1(t-h_1-\cdots-h_n)}}{e^{-\lambda_1 t}\frac{[\lambda_1 t]^n}{n!}}$$

$$\frac{\prod h_i \cdot n!}{t^n} \longrightarrow \frac{n!}{t^n} \quad \text{as } h_i \to 0$$

(: joint distribution of order statistics of n iid $U(0,t)$ random variables)

2009 THEORY 1 QUESTION 1

	A	A^c	Total
B	X_{11}	X_{12}	n_1
	X_{21}	X_{22}	n_2
Total	m_1	m_2	n

$X = (X_{11}, X_{12}, X_{21}, X_{22}) \sim \text{Multinomial}\left(n, (p_{11}, p_{12}, p_{21}, p_{22})\right)$

$$\Rightarrow f(x_{11}, x_{12}, x_{21}, x_{22}) = \underbrace{\frac{n!}{x_{11}!\,x_{12}!\,x_{21}!\,x_{22}!}}_{h(\underset{\sim}{x})} p_{11}^{x_{11}} p_{12}^{x_{12}} p_{21}^{x_{21}} p_{22}^{x_{22}}$$

(a) $f(\underset{\sim}{x}) = \exp\left\{ x_{11}\log p_{11} + x_{12}\log p_{12} + x_{21}\log p_{21} + (n - x_{11} - x_{12} - x_{21})\log p_{22} + \log h(\underset{\sim}{x}) \right\}$

$$= \exp\left\{ x_{11}\underbrace{\log\left(\frac{p_{11}}{p_{22}}\right)}_{\theta_1} + x_{12}\underbrace{\log\left(\frac{p_{12}}{p_{22}}\right)}_{\theta_2} + x_{21}\underbrace{\log\left(\frac{p_{21}}{p_{22}}\right)}_{\theta_3} + n\log(p_{22}) + \log h(\underset{\sim}{x}) \right\}$$

$$n\log(p_{22}) = -n\log\left(\frac{1}{p_{22}}\right) = \underbrace{-n\log\left[1 + e^{\theta_1} + e^{\theta_2} + e^{\theta_3}\right]}_{= b(\underset{\sim}{\theta})}$$

(b) A and B are independent $\Rightarrow P(A \cap B) = P(A)\cdot P(B) = (p_{11} + p_{21})(p_{11} + p_{12})$

$$\Rightarrow \log\left(\frac{p_{11}\cdot p_{22}}{p_{12}\cdot p_{21}}\right) = 0 \qquad (*)$$

$$\log\left(\frac{p_{12}}{p_{22}}\right) + \log\left(\frac{p_{21}}{p_{22}}\right) = \log\left(\frac{p_{11}}{p_{22}}\right) \qquad (*)(*)$$

must show that $(*) \Leftrightarrow (*)(*)$

" $\Rightarrow$ "

$$\log\left(\frac{p_{11}\cdot p_{22}}{p_{12}\cdot p_{21}}\right) = 0 \Rightarrow \log\left(\frac{p_{12}\cdot p_{21}}{p_{11}\cdot p_{22}}\right) = 0$$

$$\Rightarrow \log\left(\frac{p_{12}}{p_{22}}\right) + \log\left(\frac{p_{21}}{p_{11}}\right) = 0$$

$$\Rightarrow \log\left(\frac{p_{12}}{p_{22}}\right) + \log\left(\frac{p_{21}\cdot p_{22}}{p_{11}\cdot p_{22}}\right) = 0 \Rightarrow \log\left(\frac{p_{12}}{p_{22}}\right) + \log\left(\frac{p_{21}}{p_{22}}\right) = \log\left(\frac{p_{11}}{p_{22}}\right)$$

" $\Leftarrow$ "

follows exact same argument above in reverse

(c) $\theta = a_0\log\left(\frac{p_{11}}{p_{22}}\right) + a_1\log\left(\frac{p_{12}}{p_{22}}\right) + a_2\log\left(\frac{p_{21}}{p_{22}}\right)$ with $a_0 = 1$, $a_1 = a_2 = -1$,

$$\theta = \log\left(\frac{p_{11}}{p_{22}}\right) - \log\left(\frac{p_{12}}{p_{22}}\right) - \log\left(\frac{p_{21}}{p_{22}}\right) = \log\left(\frac{p_{11}}{\cancel{p_{22}}}\cdot\frac{\cancel{p_{22}}}{p_{12}}\cdot\frac{p_{22}}{p_{21}}\right) = \log\left(\frac{p_{11}\cdot p_{22}}{p_{12}\cdot p_{21}}\right)$$

Test $H_0: \theta = 0$ $(A \perp B)$ versus $H_1: \theta \neq 0$ using UMPU theory for multiparameter exponential families. Write $f(\underset{\sim}{x})$ as follows

involves n

$$f(\underset{\sim}{x}) = \exp\left\{ x_{11}\underbrace{\left[\log\left(\frac{p_{11}}{p_{22}}\right) - \log\left(\frac{p_{12}}{p_{22}}\right) - \log\left(\frac{p_{21}}{p_{22}}\right)\right]}_{\theta} + (x_{12} + x_{11})\underbrace{\log\left(\frac{p_{12}}{p_{22}}\right)}_{\eta_1} + (x_{21} + x_{11})\underbrace{\log\left(\frac{p_{21}}{p_{22}}\right)}_{\eta_2} - b(\underset{\sim}{\theta}) + \log h(\underset{\sim}{x})\right.$$

$\Rightarrow x_{11}$ sufficient for θ, $x_{12} + x_{11} = n_1$, $x_{21} + x_{11} = m_1$, and n are sufficient for nuisance [para]meters (η_1, η_2)

$\Rightarrow$ can test H_0 using $x_{11} \mid m_1, n_1, n$

$f(x_{11} \mid m_1, n_1, n) = \exp\{x_{11}\theta + n_1\eta_1 + \cancel{m_1\eta_2} - \cancel{b(\theta)}\}\ \frac{n!}{x_{11}!\,x_{12}!\,x_{21}!\,x_{22}!}$

[illegible]

$$\Rightarrow f_\theta(x_{11} \mid n_1, m_1, n) = \frac{[e^\theta]^{x_{11}} \frac{n!}{x_{11}!\,x_{12}!\,x_{21}!\,x_{22}!}}{\sum_{x_{11}} [e^\theta]^{x_{11}} \frac{n!}{x_{11}!\,x_{12}!\,x_{21}!\,x_{22}!}} = \frac{\binom{n_1}{x_{11}}\binom{n-n_1}{m_1-x_{11}}[e^\theta]^{x_{11}}}{\sum_{x_{11}} \binom{n_1}{x_{11}}\binom{n-n_1}{m_1-x_{11}}[e^\theta]^{x_{11}}}$$

(non-central hypergeometric)

under $H_0: \theta = 0$,

$$f_{H_0}(x_{11} \mid n_1, m_1, n) = \frac{\binom{n_1}{x_{11}}\binom{n-n_1}{m_1-x_{11}}}{\binom{n}{m_1}}, \text{ the central hypergeometric}$$

A UMPU test is given as

$$\phi(x_{11}) = \begin{cases} 1 & \text{if } x_{11} < c_1 \text{ or } x_{11} > c_2 \\ \gamma_i & \text{if } x_{11} = c_1 \text{ or } c_2 \\ 0 & \text{otherwise.} \end{cases}$$

where γ, c_1, c_2 chosen to satisfy

(1) $\displaystyle\sum_{\substack{\text{lower range}\\ \text{of } x_{11}}}^{c_1-1} f_{H_0}(x_{11} \mid n_1, m_1, n) + \gamma_1 P_{H_0}(x_{11} = c_1 \mid n_1, m_1, n) + \gamma_2 P_{H_0}(x_{11} = c_2 \mid n_1, m_1, n) + \sum_{x_{11}=c_2+1}^{\substack{\text{upper range}\\ \text{of } x_{11}}} f_{H_0}(x_{11} \mid n_1, m_1, n) = \alpha$

(2) $E_{H_0}[x_{11} \cdot \phi(x_{11})] = \alpha \cdot E_{H_0}[x_{11}]$

The conditional power function is given as:

$$\beta(\theta) = \sum_{\text{lower } x_{11}}^{c_1-1} f_\theta(x_{11} \mid n_1, m_1, n) + \gamma_1 P_\theta(x_{11} = c_1 \mid n_1, m_1, n) + \gamma_2 P_\theta(x_{11} = c_2 \mid n_1, m_1, n) + \sum_{x_{11}=c_2+1}^{\text{upper } x_{11}} f_\theta(x_{11} \mid n_1, m_1, n)$$

(d) test $H_0: P(A) \geq P(B)$ versus $H_1: P(A) < P(B)$ using UMPU theory.

$P(A) = p_{11} + p_{21}$, $P(B) = p_{11} + p_{12}$

$$P(A) \geq P(B) \Rightarrow p_{21} \geq p_{12} \Rightarrow \log\left(\frac{p_{21}}{p_{22}}\right) \geq \log\left(\frac{p_{12}}{p_{22}}\right)$$

$$\Rightarrow \log\left(\frac{p_{21}}{p_{22}}\right) - \log\left(\frac{p_{12}}{p_{22}}\right) \geq 0$$

Rewrite $f(\underset{\sim}{x})$ as follows:

$$f(\underset{\sim}{x}) = \exp\Big\{ x_{11} \underbrace{\log\left(\frac{p_{11}}{p_{22}}\right)}_{\eta_1} + (x_{12}+x_{21}) \underbrace{\log\left(\frac{p_{12}}{p_{22}}\right)}_{\eta_2} + x_{21} \underbrace{\left[\log\left(\frac{p_{21}}{p_{22}}\right) - \log\left(\frac{p_{12}}{p_{22}}\right)\right]}_{\theta} - \underbrace{b(\underset{\sim}{\theta})}_{\text{involves } n} \Big\} \frac{n!}{x_{11}!\,x_{12}!\,x_{21}!\,x_{22}!}$$

By UMPU theory for multiparameter exponential families, can test $H_0: P(A) \geq P(B)$ or $H_0: \theta \geq 0$ using $X_{21} \mid x_{11}, x_{12}+x_{21}, n$

(*) see p.4

$$f(x_{21} \mid x_{11}, x_{12}+x_{21}, n) = \frac{\exp\{x_{11}\eta_1 + (x_{12}+x_{21})\eta_2 + x_{21}\theta - b(\underset{\sim}{\theta})\} \frac{n!}{x_{11}!\,x_{12}!\,x_{21}!\,x_{22}!}}{\sum_{x_{21}} \exp\{x_{11}\eta_1 + (x_{12}+x_{21})\eta_2 + x_{21}\theta - b(\underset{\sim}{\theta})\} \frac{n!}{x_{11}!\,x_{12}!\,x_{21}!\,x_{22}!}}$$

$$= \frac{[e^\theta]^{x_{21}} \frac{n!}{x_{11}!\,(j-x_{21})!\,x_{21}!\,x_{22}!}}{\sum [e^\theta]^{x_{21}} \binom{j}{x_{21}}\binom{n-j}{x_{11}} \text{[illegible]}}$$

let $j = n - (x_{11} + x_{22})$

$$\frac{[e^\theta]^{x_{21}} \binom{j}{x_{21}}\binom{n-j}{x_{11}}}{\sum [e^\theta]^{x_{21}} \binom{j}{\text{[illegible]}}\binom{n-j}{\text{[illegible]}}}, \text{ a non-central hypergeometric}$$

Note, under $H_0: \theta \geq 0$,

$$f_{H_0}(x_{21} \mid x_{11}, j, n) = \frac{\binom{j}{x_{21}}\binom{n-j}{x_{11}}}{\binom{n}{x_{21}+x_{11}}}, \quad \text{a central hypergeometric}$$

$$\binom{x_{12}+x_{21}}{x_{12}} \left(\frac{p_{12}}{p_{12}+p_{21}}\right)^{x_{12}} \left(\frac{p_{21}}{p_{12}+p_{21}}\right)^{x_{21}}$$

thus, the UMPU test is given as

$$\phi(x_{21}) = \begin{cases} 1 & \text{if } x_{21} > c_\alpha \\ \gamma & \text{if } x_{21} = c_\alpha \\ 0 & \text{otherwise} \end{cases}$$

where c_α, γ satisfy

$$\sum_{c_\alpha+1}^{\text{upper } x_{21}} f_{H_0}(x_{21} \mid x_{11}, j, n) + \gamma P_{H_0}[x_{21} = c_\alpha \mid x_{11}, j, n] = \alpha$$

(c) Let $\hat{p}_{ij}$ denote MLE of p_{ij} under $H_0: A \perp B$

let $\hat{\pi}_{ij}$ denote MLE of p_{ij} under $H_1: A \not\perp B$

can easily find closed forms of $\hat{\pi}_{ij}$ using constrained maximization.

$$\ell(\pi_{11}, \pi_{12}, \pi_{21}, \pi_{22}) \propto x_{11}\log\pi_{11} + x_{12}\log\pi_{12} + x_{21}\log\pi_{21} + x_{22}\log\pi_{22} + \lambda(1 - \pi_{11} - \pi_{12} - \pi_{21} - \pi_{22})$$

$$= \frac{\partial \ell}{\partial \pi_{ij}} = \frac{x_{ij}}{\pi_{ij}} - \lambda \Rightarrow \frac{x_{11}}{\pi_{11}} = \frac{x_{12}}{\pi_{12}} = \frac{x_{21}}{\pi_{21}} = \frac{x_{22}}{\pi_{22}}$$

$$\Rightarrow \hat{\pi}_{ij} = \frac{x_{ij}}{n}$$

then $\Lambda_n = \dfrac{\text{maximized } L \text{ under } H_0}{\text{maximized } L \text{ under } H_1}$

$$= \left(\frac{\hat{p}_{11}}{\hat{\pi}_{11}}\right)^{x_{11}} \left(\frac{\hat{p}_{12}}{\hat{\pi}_{12}}\right)^{x_{12}} \left(\frac{\hat{p}_{21}}{\hat{\pi}_{21}}\right)^{x_{21}} \left(\frac{\hat{p}_{22}}{\hat{\pi}_{22}}\right)^{x_{22}} =$$

$$\Rightarrow -2\log\Lambda_n = 2\left[x_{11}\log\left(\frac{\hat{\pi}_{11}}{\hat{p}_{11}}\right) + x_{12}\log\left(\frac{\hat{\pi}_{12}}{\hat{p}_{12}}\right) + x_{21}\log\left(\frac{\hat{\pi}_{21}}{\hat{p}_{21}}\right) + x_{22}\log\left(\frac{\hat{\pi}_{22}}{\hat{p}_{22}}\right)\right]$$

$$= 2n \sum_{i,j} \hat{\pi}_{ij} \log\left(\frac{\hat{\pi}_{ij}}{\hat{p}_{ij}}\right)$$

Perform a second-order taylor series expansion of $f(a) = a\log\left(\frac{a}{b}\right)$ about b, where $a = \hat{\pi}_{ij}$, $b = \hat{p}_{ij}$ (for all i, j)

$$\Rightarrow f(a) \approx f(b) + f'(b)(a-b) + \frac{f''(b)}{2}(a-b)^2$$

$$\approx 0 + \left[\log\left(\frac{a}{b}\right) + \underbrace{\cancel{a\left(\frac{b}{a}\right)\left(\frac{1}{b}\right)}}_{1}\right]_{a=b}(a-b) + \frac{1}{2}\left[\frac{1}{a}\right]_{a=b}(a-b)^2$$

with $a = \hat{\pi}_{ij}$, $b = \hat{p}_{ij}$

$$\hat{\pi}_{ij}\log\left(\frac{\hat{\pi}_{ij}}{\hat{p}_{ij}}\right) \approx \underbrace{(\hat{\pi}_{ij} - \hat{p}_{ij})}_{\to 0} + \frac{1}{2\hat{p}_{ij}}(\hat{\pi}_{ij} - \hat{p}_{ij})^2$$

(i)

$$\Rightarrow -2\log\Lambda_n = 2n\sum_{i,j}\frac{1}{2\hat{p}_{ij}}(\hat{\pi}_{ij} - \hat{p}_{ij})^2 \quad \sum_{i,j}\frac{n^2(\hat{\pi}_{ij} - \hat{p}_{ij})^2}{n\hat{p}_{ij}} = \sum_{i,j}\frac{(x_{ij} - n\hat{p}_{ij})^2}{n\hat{p}_{ij}}$$

(ii) under H_0, $-2\log\Lambda_n \xrightarrow{d} \chi^2_1$

under H_1, $-2\log\Lambda_n \xrightarrow{d} \chi^2_{1,\delta}$ [illegible]

$$f(x_{21} \mid x_{11}, x_{12}+x_{21}=t, n) = \frac{\exp\{\cancel{x_{11}\eta_1} + \cancel{t\eta_2} + x_{21}\log\left(\frac{p_{21}}{p_{12}}\right) - b(\underset{\sim}{\theta})\} \frac{\cancel{n!}}{\cancel{x_{11}!}\, x_{12}!\, x_{21}!\, \cancel{x_{22}!}}}{\sum_{\substack{x_{21}\ \text{s.t.}\\ x_{12}+x_{21}=t}}^{t} \exp\{\cancel{x_{11}\eta_1} + \cancel{t\eta_2} + x_{21}\log\left(\frac{p_{21}}{p_{12}}\right) - b(\underset{\sim}{\theta})\} \frac{\cancel{n!}}{\cancel{x_{11}!}\, x_{12}!\, x_{21}!\, \cancel{x_{22}!}}}$$

$(x_{21}=0)$

$$= \frac{\left(\frac{p_{21}}{p_{12}}\right)^{x_{21}} \binom{t}{x_{21}} \left(\frac{p_{12}}{p_{21}+p_{12}}\right)^{t}}{\sum_{x_{21}=0}^{t} \binom{t}{x_{21}} \left(\frac{p_{21}}{p_{12}}\right)^{x_{21}} \left(\frac{p_{12}}{p_{21}+p_{12}}\right)^{t}} = \binom{\overset{= x_{21}+x_{12}}{t}}{x_{21}} \left(\frac{p_{21}}{p_{21}+p_{12}}\right)^{x_{21}} \left(\frac{p_{12}}{p_{21}+p_{12}}\right)^{x_{12}}$$

Under H_0, $p_{21}=p_{12} \Rightarrow$

$$\frac{p_{21}}{p_{21}+p_{12}} = \frac{p_{12}}{p_{21}+p_{12}} = \frac{1}{2}$$

2009 QUESTION 2

$\lambda \sim$ Exponential (mean $= \frac{1}{\theta}$). Conditional on λ, $X \sim$ Poisson (λ) $\perp$ of $Y \sim$ Poisson $(\beta\lambda)$. $(X_1, Y_1), \ldots, (X_n, Y_n)$ are an iid sample from the unconditional joint distribution of (X, Y).

(a) the properties of the unconditional distribution of (X, Y)

(i) $E(X) = E_\lambda\{E[X|\lambda]\} = E_\lambda[\lambda] = \frac{1}{\theta} = \theta^{-1}$

$$Var(X) = E_\lambda\{Var(X|\lambda)\} + Var_\lambda\{E(X|\lambda)\}$$
$$= E_\lambda\{\lambda\} + Var_\lambda\{\lambda\} = \theta^{-1} + \theta^{-2}$$

$$E(Y) = E_\lambda\{E[Y|\lambda]\} = E_\lambda[\beta\lambda] = \beta\theta^{-1}$$

$$Var(Y) = E_\lambda\{Var(Y|\lambda)\} + Var_\lambda\{E(Y|\lambda)\}$$
$$= E_\lambda(\beta\lambda) + Var_\lambda(\beta\lambda) = \beta\theta^{-1} + \beta^2\theta^{-2}$$

$$cov(X,Y) = E(XY) - E(X)E(Y)$$
$$= E_\lambda[\underbrace{E(XY|\lambda)}_{\text{conditionally independent}}] - (\theta^{-1})(\beta\theta^{-1})$$
$$= E_\lambda[E(X|\lambda)\cdot E(Y|\lambda)] - \beta\theta^{-2} = E_\lambda[\beta\lambda^2] - \beta\theta^{-2}$$
$$= \beta\{Var(\lambda) + [E(\lambda)]^2\}$$
$$= \beta\{\theta^{-2} + \theta^{-2}\} - \beta\theta^{-2} = \beta\theta^{-2}$$

(ii) $f(X,Y) = \int_\lambda f(X,Y,\lambda)\, d\lambda$

$$= \int_\lambda f(X,Y|\lambda)\cdot f(\lambda)\, d\lambda$$
$$= \int_\lambda f(X|\lambda)\cdot f(Y|\lambda)\cdot f(\lambda)\, d\lambda \quad , \text{ as } X|\lambda \perp Y|\lambda$$
$$= \int_0^\infty \left[\frac{\lambda^x}{x!} e^{-\lambda}\right]\left[\frac{(\beta\lambda)^y}{y!} e^{-\beta\lambda}\right]\left[\theta e^{-\theta\lambda}\right] d\lambda$$
$$= \frac{\beta^y\theta}{x!\,y!}\int_0^\infty \lambda^{x+y} e^{-\lambda/\left(\frac{1}{\theta+\beta+1}\right)}\, d\lambda$$
$$= \left[\int_0^\infty \frac{\lambda^{(x+y+1)-1} e^{-\lambda/\left(\frac{1}{\theta+\beta+1}\right)}}{\Gamma(x+y+1)\left(\frac{1}{\theta+\beta+1}\right)^{x+y+1}}\, dy\right] \cdot \frac{\beta^y\theta}{x!\,y!}\cdot\frac{\overbrace{\Gamma(x+y+1)}^{(x+y)!}}{(\theta+\beta+1)^{x+y+1}}$$

(Gamma density)

$$= \frac{(x+y)!}{x!\,y!}\left(\frac{\theta}{\theta+\beta+1}\right)\left(\frac{1}{\theta+\beta+1}\right)^x\left(\frac{\beta}{\theta+\beta+1}\right)^y, \quad x, y = 0, 1, 2, \ldots$$

(b) Show that the MLEs (based on a sample of size n from $f(X,Y)$) are
$\hat{\theta}_n = \bar{X}_n^{-1}$, $\hat{\beta}_n = \bar{Y}_n / \bar{X}_n$.

Must maximize the following likelihood for θ, β

$$L(\theta,\beta \mid \underset{\sim}{X}, \underset{\sim}{Y}) = \prod_{i=1}^{n} \frac{\theta \beta^{y_i}}{(\theta+\beta+1)^{x_i+y_i+1}} \frac{(x_i+y_i)!}{x_i!\,y_i!}$$

$$\propto \theta^n \beta^{\sum y_i} (\theta+\beta+1)^{-(\sum x_i + \sum y_i + n)}$$

Or equivalently can maximize ℓ, the log-likelihood

$$\ell(\theta,\beta \mid \underset{\sim}{X}, \underset{\sim}{Y}) \propto n \log\theta + \sum y_i \log\beta - \left(\sum x_i + \sum y_i + n\right)\log(\theta+\beta+1)$$

$\hat{\theta}_n$ and $\hat{\beta}_n$ solve the following score equations

$$\frac{\partial \ell}{\partial \theta} = \frac{n}{\theta} - \frac{(\sum x_i + \sum y_i + n)}{\theta+\beta+1} = 0 \quad \Rightarrow \quad n(\theta+\beta+1) - \theta\left(\sum x_i + \sum y_i + n\right) = 0$$
$$\Rightarrow (\beta+1)n - \theta\left(\sum x_i + \sum y_i\right) = 0$$

$$\frac{\partial \ell}{\partial \beta} = \frac{\sum y_i}{\beta} - \frac{(\sum x_i + \sum y_i + n)}{\theta+\beta+1} = 0 \quad \Rightarrow \quad \sum y_i(\theta+\beta+1) - \beta\left(\sum x_i + \sum y_i + n\right) = 0$$
$$\Rightarrow (\theta+1)\sum y_i - \beta\left(\sum x_i + n\right) = 0$$
$$\therefore \beta = \frac{(\theta+1)\sum y_i}{\sum x_i + n}$$

Then $$\beta + 1 = \frac{(\theta+1)\sum y_i + \sum x_i + n}{\sum x_i + n}$$

$$\Rightarrow \left(\frac{(\theta+1)\sum y_i + \sum x_i + n}{\sum x_i + n}\right) n - \theta\left(\sum x_i + \sum y_i\right) = 0$$

$$\Rightarrow \left[(\theta+1)\sum y_i + \sum x_i + n\right] n - \theta\left(\sum x_i + \sum y_i\right)\left(\sum x_i + n\right) = 0$$

$$\Rightarrow \cancel{n\theta \sum y_i} + n\sum y_i + n\sum x_i + n^2 - \theta\left[\left(\sum x_i\right)^2 + n\sum x_i + \left(\sum x_i\right)\left(\sum y_i\right) + \cancel{n\sum y_i}\right] = 0$$

$$\Rightarrow \hat{\theta} = \frac{n\left(\sum y_i + \sum x_i + 1\right)}{(\sum x_i)^2 + n\sum x_i + (\sum x_i)(\sum y_i)} = \frac{n\left(\sum y_i + \sum x_i + 1\right)}{\sum x_i\left(\sum x_i + 1 + \sum y_i\right)} = \bar{X}_n^{-1}$$

Then $$\hat{\beta} = \frac{(\hat{\theta}+1)\sum y_i}{\sum x_i + n} = \frac{\left(\frac{n}{\sum x_i} + 1\right)\sum y_i}{\sum x_i + n} = \frac{\left(\frac{n + \sum x_i}{\sum x_i}\right)\sum y_i}{(n + \sum x_i)} = \frac{\sum y_i}{\sum x_i} = \bar{Y}_n / \bar{X}_n$$

(c) By MLE theory, as $(X_1, Y_1), \dots, (X_n, Y_n)$ are iid,

$$\sqrt{n}\begin{pmatrix}\hat{\theta}_n - \theta \\ \hat{\beta}_n - \beta\end{pmatrix} \xrightarrow{d} N_2\left(0,\ I^{-1}(\theta,\beta)\right)$$

where $I(\theta,\beta)$ is the Fisher information in one observation from $f(X,Y)$.

$I(\theta,\beta) = -E\left(\frac{\partial \dot{l}}{\partial \eta}\right)$, where $\eta = (\theta, \beta)$, and l is the log-likelihood for one observation

$$l = l(\theta,\beta) = x_i \log\theta + y_i \log\beta - (x_i + y_i + 1)\log(\theta+\beta+1)$$

$$\Rightarrow \dot{l}(\theta) = \frac{x_i}{\theta} - \frac{(x_i+y_i+1)}{(\theta+\beta+1)} \quad \Rightarrow \quad \ddot{l}(\theta) = \frac{-x_i}{\theta^2} + \frac{(x_i+y_i+1)}{(\theta+\beta+1)^2}$$

$$\dot{l}(\beta) = \frac{y_i}{\beta} - \frac{(x_i+y_i+1)}{(\theta+\beta+1)} \qquad \ddot{l}(\theta,\beta) = \frac{(x_i+y_i+1)}{(\theta+\beta+1)^2}$$

$$\ddot{l}(\beta) = \frac{-y_i}{\beta^2} + \frac{(x_i+y_i+1)}{(\theta+\beta+1)^2}$$

$$\Rightarrow I(\theta,\beta) = E\begin{bmatrix} \frac{x_i}{\theta^2} - \frac{x_i+y_i+1}{(\theta+\beta+1)^2} & -\frac{(x_i+y_i+1)}{(\theta+\beta+1)^2} \\ & \frac{y_i}{\beta^2} - \frac{(x_i+y_i+1)}{(\theta+\beta+1)^2}\end{bmatrix}$$

$$\text{As } \quad = E\begin{bmatrix} \frac{1}{\theta^2} - \frac{\frac{1}{\theta}\cancel{(1+\beta+\theta)}}{(\theta+\beta+1)^2} & -\frac{\frac{1}{\theta}\cancel{(1+\beta+\theta)}}{(\theta+\beta+1)^2} \\ & \frac{1}{\beta\theta} - \frac{\frac{1}{\theta}\cancel{(1+\beta+\theta)}}{(\theta+\beta+1)^2}\end{bmatrix}$$

$$= \frac{1}{\theta+\beta+1}\begin{bmatrix} \frac{\theta+\beta+1}{\theta^2} - \frac{1}{\theta} & -\frac{1}{\theta} \\ -\frac{1}{\theta} & \frac{\theta+\beta+1}{\beta\theta} - \frac{1}{\theta}\end{bmatrix} = \frac{1}{\theta+\beta+1}\begin{bmatrix} \frac{\beta+1}{\theta^2} & -\frac{1}{\theta} \\ -\frac{1}{\theta} & \frac{\theta+1}{\beta\theta}\end{bmatrix}$$

$$\Rightarrow \underline{I^{-1}(\theta,\beta)} = \frac{\theta+\beta+1}{\left(\frac{\beta+1}{\theta^2}\right)\left(\frac{\theta+1}{\beta\theta}\right) - \left(\frac{1}{\theta}\right)^2}\begin{bmatrix} \frac{\theta+1}{\beta\theta} & \frac{1}{\theta} \\ \frac{1}{\theta} & \frac{\beta+1}{\theta^2}\end{bmatrix} \qquad \beta\theta^3\begin{bmatrix} \frac{\theta+1}{\beta\theta} & \frac{1}{\theta} \\ \frac{1}{\theta} & \frac{\beta+1}{\theta^2}\end{bmatrix}$$

$$\frac{\beta\theta+\beta+\theta+1}{\beta\theta^3} \qquad \frac{1}{\theta^2}$$

$\beta\theta^3$ [illegible] $\beta\theta$

$$\begin{bmatrix} \theta^2(\theta+1) & \beta\theta^2 \\ & \theta\beta(\beta+1)\end{bmatrix}$$

(d) Let $T_1 = \sqrt{n\bar{X}_n/2}\left(\bar{Y}_n/\bar{X}_n - 1\right)$, and $T_2 = \sqrt{n\bar{X}_n/2}\,\ln\left(\bar{Y}_n/\bar{X}_n\right)$. All of the following derivations are performed under $H_0: \beta = 1$.

First, note that from the multivariate CLT, and from results in (a),

$$\sqrt{n}\left(\begin{pmatrix}\bar{X}_n \\ \bar{Y}_n\end{pmatrix} - \begin{pmatrix}E(X) \\ E(Y)\end{pmatrix}\right) \xrightarrow{d} N_2\left(0, \begin{bmatrix} Var(X) & Cov(X,Y) \\ & Var(Y)\end{bmatrix}\right)$$

Under $H_0: \beta = 1$, $E(X) = E(Y) = \theta^{-1}$, $Var(X) = Var(Y) = \theta^{-1} + \theta^{-2}$, $Cov(X,Y) = \theta^{-2}$

$$\Rightarrow \sqrt{n}\left(\begin{pmatrix}\bar{X}_n \\ \bar{Y}_n\end{pmatrix} - \begin{pmatrix}\theta^{-1} \\ \theta^{-1}\end{pmatrix}\right) \xrightarrow[\text{under } H_0]{d} N_2\Bigg(0, \underbrace{\begin{bmatrix} \theta^{-1}+\theta^{-2} & \theta^{-2} \\ & \theta^{-1}+\theta^{-2}\end{bmatrix}}_{\theta^{-2}\begin{bmatrix}\theta+1 & 1 \\ 1 & \theta+1\end{bmatrix}, \text{ call this } \Sigma}\Bigg)$$

(i) $T_1 = \sqrt{n}\left[\underbrace{\sqrt{\frac{\bar{X}_n}{2}}\left(\frac{\bar{Y}_n}{\bar{X}_n} - 1\right)}_{g(x,y) = \sqrt{\frac{x}{2}}\left(\frac{y}{x} - 1\right)}\right]$

$\left(\frac{x}{2}\right)^{\frac{1}{2}}\left(\frac{y}{x} - 1\right) = \left(\frac{x}{2}\right)^{\frac{1}{2}}\left(\frac{y-x}{x}\right) = \left(\frac{1}{2}\right)^{\frac{1}{2}} x^{-\frac{1}{2}}(y-x)$

$\Rightarrow \nabla g(x,y) = \left(\left(\frac{1}{2}\right)^{\frac{1}{2}}\left[-\frac{1}{2}x^{-\frac{3}{2}}(y-x) - x^{-\frac{1}{2}}\right], \left(\frac{1}{2}\right)^{\frac{1}{2}} x^{-\frac{1}{2}}\right)$

Can use delta method to find asymptotic distribution of $g(\bar{X}, \bar{Y})$.

Note $g(E(X), E(Y)) = 0$ as $E(X) = E(Y)$

$\nabla g(E(X), E(Y)) = \left(\frac{1}{2}\right)^{\frac{1}{2}}\left(-\theta^{\frac{1}{2}}, \theta^{\frac{1}{2}}\right)$

By delta, $\underbrace{\sqrt{n}\left[g(\bar{X}, \bar{Y}) - g(E(X), E(Y))\right]}_{= T_1} \xrightarrow{d} N_1\left(0, \Sigma^*\right)$

where $\Sigma^* = \nabla g[E(X), E(Y)]\, \Sigma\, \nabla g[E(X), E(Y)]'$

$= \frac{1}{2}\theta^{-2}\begin{bmatrix}-\theta^{\frac{1}{2}} & \theta^{\frac{1}{2}}\end{bmatrix}\begin{bmatrix}\theta+1 & 1 \\ 1 & \theta+1\end{bmatrix}\begin{bmatrix}-\theta^{\frac{1}{2}} \\ \theta^{\frac{1}{2}}\end{bmatrix}$

$\begin{bmatrix}a & b\end{bmatrix}\begin{bmatrix}i & j \\ j & k\end{bmatrix}\begin{bmatrix}a \\ b\end{bmatrix}$

$(ai + bj, aj + bk)\begin{pmatrix}a \\ b\end{pmatrix}$

$a^2 i + 2abj + b^2 k$

$= \frac{1}{2}\theta^{-2}\left[\theta(\theta+1) - 2\theta + \theta(\theta+1)\right] = \frac{1}{2}\theta^{-2}\left[2\theta^2\right] = 1$

$\Rightarrow T_1 \xrightarrow{d} N_1(0, 1)$

(ii) $t_1 - t_2 = \sqrt{\frac{n\bar{X}_n}{2}}\left[\underbrace{\frac{\bar{Y}_n}{\bar{X}_n} - 1 - \log\left(\frac{\bar{Y}_n}{\bar{X}_n}\right)}_{= f(\hat{\beta}_n)}\right]$

$f(\hat{\beta}_n)$, where f is a continuous function ($f(a) = a - 1 - \log a$)

that $\hat{\beta}_n = \frac{\bar{Y}_n}{\bar{X}_n} \xrightarrow[a.s.]{} \beta = 1$ (under H_0) as $n \to \infty$. from MLE theory.

consequently, from continuous mapping theorem, $f(\hat{\beta}_n) \xrightarrow[a.s]{} f(1) = 0$.

Thus, $t_1 - t_2 \xrightarrow[p]{} 0$. What about $\sqrt{\frac{n}{2}\bar{X}_n}$? $\bar{X}_n \xrightarrow[a.s.]{} E(X) = \theta^{-1}$. From continuous map,

(*) see p.6 $\quad \sqrt{\frac{\bar{X}_n}{2}} \xrightarrow[a.s.]{} \sqrt{\frac{\theta^{-1}}{2}}$ $\quad t_1 - t_2 = \underbrace{(\sqrt{n})}_{\text{what to do?}} \underbrace{\sqrt{\frac{\bar{X}_n}{2}}}_{\xrightarrow[a.s.]{} \sqrt{\frac{\theta^{-1}}{2}}} \cdot \underbrace{f(\hat{\beta}_n)}_{\xrightarrow[a.s]{} 0}$

(iii) $t_2 = \sqrt{n}\left[\underbrace{\sqrt{\frac{\bar{X}_n}{2}} \ln\left(\frac{\bar{Y}_n}{\bar{X}_n}\right)}\right]$

$g(x,y) = \left(\frac{1}{2}\right)^{\frac{1}{2}} x^{\frac{1}{2}} \ln\left(\frac{y}{x}\right) \Rightarrow \nabla g(x,y) = \left(\frac{1}{2}\right)^{\frac{1}{2}}\left[\underbrace{\frac{1}{2}x^{-\frac{1}{2}}\ln\left(\frac{y}{x}\right) + x^{\frac{1}{2}} \cdot \left(\frac{x}{y}\right)\left(\frac{-y}{x^2}\right)}_{\frac{1}{2}x^{-\frac{1}{2}}\ln\left(\frac{y}{x}\right) - x^{-\frac{1}{2}}}, \underbrace{x^{\frac{1}{2}}\frac{1}{y}}_{\frac{x^{\frac{1}{2}}}{y}}\right]$

under H_0, $E(X) = E(Y) = \theta^{-1} \Rightarrow g(E(X), E(Y)) = 0$

$\Rightarrow \nabla g[E(X), E(Y)] = \left(\frac{1}{2}\right)^{\frac{1}{2}}\left[-\theta^{\frac{1}{2}}, \theta^{\frac{1}{2}}\right)$

By delta, $\underbrace{\sqrt{n}\left[g(\bar{X}, \bar{Y}) - g(E(X), E(Y))\right]}_{= t_2} \xrightarrow[d]{} N_1(0, \Sigma^*)$

Σ^* is exactly the same as in (i), as ∇g are equivalent in (i) and (iii), and $\Sigma^* = \nabla g \Sigma \nabla g'$.

thus $\Sigma^* = 1$, and $t_2 \xrightarrow[d]{} N_1(0,1)$.

(e) $\tau = \frac{\theta}{\theta + 2} \Rightarrow \tau(\theta + 2) = \theta \Rightarrow 2\tau = \theta - \theta\tau \Rightarrow \theta = \frac{2\tau}{1-\tau}$. with $\beta = 1$,

$f(X, Y \mid \tau) = \frac{(x+y)!}{x!\,y!}\left(\frac{2\tau}{1-\tau}\right) \cdot \frac{1}{\left(\frac{2\tau}{1-\tau} + 2\right)^{x+y+1}} = \frac{(x+y)!}{x!\,y!}\left(\frac{2\tau}{1-\tau}\right)\left(\frac{1-\tau}{2}\right)^{x+y+1}$

$\frac{2\tau + 2(1-\tau)}{1-\tau} \quad \frac{2}{1-\tau})$ $\qquad = \frac{(x+y)!}{x!\,y!}\,\tau\left(\frac{1-\tau}{2}\right)^{x+y}$, $x, y = 0, 1, \ldots$

The Bayes estimator of τ under squared error loss will be the posterior mean of τ, where there is a Beta(a_0, b_0) prior on τ,

i.e. $\pi(\tau) = \frac{\Gamma(a_0+b_0)}{\Gamma(a_0)\Gamma(b_0)}\tau^{a_0-1}(1-\tau)^{b_0-1}$

... posterior distribution $\tau \mid X, Y$ ... found as

$$f(\tau \mid \underset{\sim}{X}, \underset{\sim}{Y}) = \frac{f(\tau, \underset{\sim}{X}, \underset{\sim}{Y})}{f(\underset{\sim}{X}, \underset{\sim}{Y})}$$

$$= \frac{f(\underset{\sim}{X}, \underset{\sim}{Y} \mid \tau) f(\tau)}{\int_\tau f(\underset{\sim}{x}, \underset{\sim}{Y} \mid \tau) f(\tau) \, \partial\tau}$$

$$= \frac{\left[\prod_{i=1}^{n} \frac{(x_i+y_i)!}{x_i! \, y_i!}\right] \left(\frac{1}{2}\right)^{\cancel{\sum x_i + \sum y_i}} \tau^n (1-\tau)^{\sum x_i + \sum y_i} \cdot \frac{\Gamma(a_0+b_0)}{\Gamma(a_0)\Gamma(b_0)} \tau^{a_0-1} (1-\tau)^{b_0-1}}{\int_0^1 \left[\prod_{i=1}^{n} \frac{(x_i+y_i)!}{x_i! \, y_i!}\right] \left(\frac{1}{2}\right)^{\sum x_i + y_i} \frac{\Gamma(a_0+b_0)}{\Gamma(a_0)\Gamma(b_0)} \tau^{n+a_0-1} (1-\tau)^{\sum x_i + \sum y_i + b_0 - 1} \, \partial\tau}$$

$$\Rightarrow \tau \mid \underset{\sim}{X}, \underset{\sim}{Y} \sim \text{Beta}\left(n+a_0, \sum X_i + \sum Y_i + b_0\right)$$

And the Bayes estimator is $E[\tau \mid \underset{\sim}{X}, \underset{\sim}{Y}] = \dfrac{n+a_0}{\sum X_i + \sum Y_i + n + a_0 + b_0} = \hat{\tau}_B$.

As the loss function is squared error, the Bayes estimator is unique. If we can show that the Bayes estimator has finite frequentist risk, then it is admissible. Under squared error loss, the frequentist risk of $\hat{\tau}_B$ is $Var(\hat{\tau}_B) + Bias^2(\hat{\tau}_B)$. Or can try to compute directly.

$$R(\hat{\tau}_B, \tau) = \sum_{x,y} (\hat{\tau}_B - \tau)^2 \cdot f(\underset{\sim}{X}, \underset{\sim}{Y} \mid \tau) \, \partial\tau$$

$$= \sum_{x,y} \left[\frac{n+a_0}{\sum X_i + \sum y_i + n + a_0 + b_0} - \tau\right]^2 \left\{\prod_{i=1}^{n} \frac{(x_i+y_i)!}{x_i! \, y_i!} \left(\frac{1}{2}\right)^{\sum x_i + \sum y_i} \tau^n (1-\tau)^{\sum X_i + \sum Y_i}\right\} \overset{?}{<} \infty$$

(d)(ii) $T_1 - T_2 = \sqrt{\frac{\bar{X}}{2}} \sqrt{n} \left[\frac{\bar{Y}}{\bar{X}} - 1 - \log\left(\frac{\bar{Y}}{\bar{X}}\right)\right]$

By LLN, $\bar{X} \xrightarrow[a.s.]{} E(X) = \frac{1}{\theta}$. By continuous map, $\sqrt{\frac{\bar{X}}{2}} \xrightarrow[a.s.]{} \sqrt{\frac{1}{2\theta}}$

we know $\hat{\beta} = \frac{\bar{Y}}{\bar{X}}$, and $\hat{\beta} \xrightarrow[a.s.]{} \beta = 1$.

consider a Taylor expansion of $f(\hat{\beta}) = \hat{\beta} - 1 - \log(\hat{\beta})$ about 1, so that:

$$f(\hat{\beta}) \approx f(1) + f'(1)(\hat{\beta} - 1) + \frac{f''(1)(\hat{\beta}-1)^2}{2} + \text{remainder}$$

$$\approx 0 + \underbrace{\left(1 - \frac{1}{1}\right)}_{0}(\hat{\beta} - 1) + \frac{1}{2}(\hat{\beta} - 1)^2 + \text{remainder}$$

then $\sqrt{n}\left[\frac{\bar{Y}}{\bar{X}} - 1 - \log\left(\frac{\bar{Y}}{\bar{X}}\right)\right] \approx \frac{(\hat{\beta} - 1)}{2} \sqrt{n}(\hat{\beta} - 1)$

$\sqrt{n}(\hat{\beta} - 1) \to N(0, I^{-1}(\beta))$, By (1), (2) + Slutsky,

$\Rightarrow \sqrt{n}(\hat{\beta} - 1) \to 0 \cdot N(0, I^{-1}(\beta)) = 0$

[illegible] $\sqrt{n}\left[\frac{\bar{Y}}{\bar{X}} - 1 - \log\left(\frac{\bar{Y}}{\bar{X}}\right)\right] \to \sqrt{\frac{1}{2\theta}} \cdot 0 \cdot 0 \to 0$.

2009 QUESTION 3

$B_{n+1} = \sum_{k=1}^{B_n} A_{nk}$

$P_{ij} = P(B_{n+1} = j \mid B_n = i) = P\left(\sum_{k=1}^{i} A_{nk} = j\right)$

(a) $E(B_{n+1}) = E_{B_n}\left[E(B_{n+1} \mid B_n)\right]$

$$= E_{B_n}\left[\sum_{k=1}^{B_n} A_{nk}\right] = E_{B_n}\left[B_n E(A_{nk})\right] = E(B_n)E(A_{nk}) \qquad (*)$$

$\therefore Var(B_{n+1}) = E_{B_n}\left[Var(B_{n+1} \mid B_n)\right] + Var_{B_n}\left[E(B_{n+1} \mid B_n)\right]$

$$= E_{B_n}\left[Var\left[\sum_{k=1}^{B_n} A_{nk}\right]\right] + Var_{B_n}\left[E\left(\sum_{k=1}^{B_n} A_{nk}\right)\right]$$

$$= E_{B_n}\left[B_n Var(A_{nk})\right] + Var_{B_n}\left[B_n E(A_{nk})\right]$$

$$= E(B_n)Var(A_{nk}) + \left[E(A_{nk})\right]^2 Var(B_n) \qquad (*)(*)$$

(b) let μ $\to$ For $n \geq 1$, will use induction to show that $\underline{E(B_n) = \mu^n}$

As $B_0 = 1$, $E(B_1) = E\left[\sum_{k=1}^{B_0=1} A_{nk}\right] = E(A_{nk}) = \mu^1$

Assume $E(B_{n-1}) = \mu^{n-1}$

If from $(*)$, $E(B_n) = E(B_{n-1}) \cdot E(A_{nk}) = \mu^{n-1} \cdot \mu = \mu^n$

$\to$ Now, for $\mu \neq 1$, will use induction to show that $\underline{Var(B_n) = \sigma^2\mu^{n-1}(1-\mu^n)(1-\mu)^{-1}}$, $n \geq$

$Var(B_1) = E(B_0)Var(A_{nk}) + Var(B_0)\left[E(A_{nk})\right]^2$ from $(*)(*)$

$$= 1 \cdot \sigma^2 + 0\mu^2 = \sigma^2 = \sigma^2\mu^0(1-\mu)(1-\mu)^{-1}$$

$$= \sigma^2$$

Now, assume $Var(B_{n-1}) = \sigma^2\mu^{n-2}(1-\mu^{n-1})(1-\mu)^{-1}$. Using $(*)(*)$,

$Var(B_n) = E(B_{n-1})Var(A_{nk}) + \left[E(A_{nk})\right]^2 Var(B_{n-1})$

$$= \mu^{n-1} \cdot \sigma^2 + \mu^2\left[\sigma^2\mu^{n-2}(1-\mu^{n-1})(1-\mu)^{-1}\right]$$

$$(1-\mu)^{-1}\left[(1-\mu)\mu^{n-1}\sigma^2 + \sigma^2\mu^n(1-\mu^{n-1})\right]$$

$$= (1-\mu)^{-1}\sigma^2\mu^{n-1}\left[1-\mu + \mu(1-\mu^{n-1})\right] = (1-\mu)^{-1}\sigma^2\mu^{n-1}\left[1-\mu^n\right]$$

$\mu - \mu^n$

$\to$ Now, for $\mu = 1$, will use induction to show that $V(B_n) = n\sigma^2$, $n \geq 1$

$Var(B_1) = E(B_0)Var(A_{nk}) + Var(B_0)\left[E(A_{nk})\right]^2$ (from $(*)(*)$)

$1 \cdot \sigma^2 + 0 \cdot \mu^2 = \sigma^2 = 1\sigma^2$

Assume $Var(B_{n-1}) = (n-1)\sigma^2$ using $(*)(*)$

[illegible]

(c) From (b), with $\mu < 1$, we know

(i) $E(B_n) = \mu^n$, and thus $\lim_{n\to\infty} E(B_n) \to 0$

(ii) $Var(B_n) = \sigma^2 \mu^{n-1}(1-\mu^n)(1-\mu)^{-1}$

$= \sigma^2 (1-\mu)^{-1} \{\mu^{n-1} - \mu^{2n-1}\}$, and thus $\lim_{n\to\infty} Var(B_n)$

$= \sigma^2 (1-\mu)^{-1} \lim_{n\to\infty} \{\mu^{n-1} - \mu^{2n-1}\} \to 0$

From Chebyshev's inequality, we know that

$$P\left[|B_n - E(B_n)| > \varepsilon\right] \le \frac{Var(B_n)}{\varepsilon^2}$$

Taking limits on both sides of the above, we have (from (i), (ii)):

$$P\left[|B_n - 0| > \varepsilon\right] \le \frac{0}{\varepsilon^2} \Rightarrow P\left[B_n > \varepsilon\right] \le 0 \text{ as } n\to\infty$$

(*) Markov discreteness

$$\Rightarrow P(B_n = 0) \to 1 \text{ as } n\to\infty$$

(d) $P_{1j}^{(n)} = P(B_n = j \mid B_0 = 1)$.

$$\phi_n(z) = E\left[z^{B_n}\right] = \sum_{j=0}^{\infty} P(B_n = j)\cdot z^j$$

$$= \sum_{j=0}^{\infty} \left[\sum_{\text{all } k} P(B_n = j, B_0 = k)\right] z^j$$

$$= \sum_{j=0}^{\infty} \left[\sum_{\text{all } k} P(B_n = j \mid B_0 = k)\cdot P(B_0 = k)\right] z^j$$

(we assume that $B_0 = 1$ with probability 1)

$$= \sum_{j=0}^{\infty} \left[\sum_{k=1}^{1} P(B_n = j \mid B_0 = k)\ \underbrace{P(B_0 = k)}_{=1}\right] z^j$$

$$= \sum_{j=0}^{\infty} P(B_n = j \mid B_0 = 1)\cdot z^j = \sum_{j=0}^{\infty} P_{1j}^{(n)} z^j$$

(e) Define $\phi(z) = E\left[z^{A_{nk}}\right] = \sum_{j=0}^{\infty} P(A_{nk} = j) z^j$. Then,

$$\sum_{j=0}^{\infty} P\left(\sum_{k=1}^{B_n} A_{nk} = j \mid B_n = k\right) z^j = \sum_{j=0}^{\infty} P\left(A_{n1} + A_{n2} + \cdots + A_{nk} = j\right) z^j$$

$$= E\left[z^{\sum_{k=1}^{k} A_{nk}}\right] = E\left[z^{A_{n1} + A_{n2} + \cdots + A_{nk}}\right], \quad A_{ni},\ i = 1, \ldots, k \text{ iid } (*)$$

$$= E\left[z^{A_{n1}}\right] \cdots E\left[z^{A_{nk}}\right]$$

$$= \left[E\left(z^{A_{n1}}\right)\right]^k$$

$$= \left[\phi(z)\right]^k$$

(f) Establish that the following recursive relationship holds

$$\Phi_n(z) = \Phi_{n-1}\{\Phi(z)\}, \quad n \geq 1$$

$$\Phi_n(z) = E\left[z^{B_n}\right] = E_{B_{n-1}}\left\{E\left[z^{B_n} \mid B_{n-1} = k\right]\right\}$$

$$= E_{B_{n-1}}\left\{\sum_{j=0}^{\infty} P\left[B_n = \sum_{k=1}^{B_{n-1}} A_{nk} = j \mid B_{n-1} = k\right] z^j\right\}$$

$$= E_{B_{n-1}}\left\{\underbrace{\sum_{j=0}^{\infty} P\left[A_{n1} + A_{n2} + \cdots + A_{nk} = j\right] z^j}_{=\{\Phi(z)\}^k,\ \text{from (e)}}\right\}$$

$$= E_{B_{n-1}}\left\{\Phi(z)^{B_{n-1}}\right\}$$

$$= E\left[\Phi(z)^{B_{n-1}}\right] = \Phi_{n-1}\{\Phi(z)\}$$

argument in pgf (instead of z)

2009 THEORY 2 QUESTION 1

(a) Likelihood proportional to (where X_{ij} denotes number out of n with $y_{i1} = i, y_{i2} = j$, $i = 0,1$ $j = 0,1$)

$$\propto \pi_{11}^{X_{11}} \pi_{12}^{X_{12}} \pi_{21}^{X_{21}} \pi_{22}^{X_{22}}, \quad \sum_{j,k} \pi_{jk} = 1 \text{ (constraint)}$$

$$\Rightarrow \ell \propto X_{11} \log \pi_{11} + X_{12} \log \pi_{12} + X_{21} \log \pi_{21} + X_{22} \log \pi_{22} + \lambda (1 - \pi_{11} - \pi_{12} - \pi_{21} - \pi_{22})$$

$$\Rightarrow \frac{\partial \ell}{\partial \pi_{jk}} = \frac{X_{jk}}{\pi_{jk}} - \lambda \Rightarrow \frac{X_{11}}{\pi_{11}} = \frac{X_{12}}{\pi_{12}} = \frac{X_{21}}{\pi_{21}} = \frac{X_{22}}{\pi_{22}} \Rightarrow \hat{\pi}_{jk} = \frac{X_{jk}}{n} \quad \text{(MLEs)}$$

Write $\hat{\pi}_{jk} = \dfrac{\sum_i I(Y_{i1} = j, Y_{i2} = k)}{n}$ (written as a mean)

From assumptions,

$$\left.\begin{aligned} E[I(Y_{i1}=j, Y_{i2}=k)] &= P[I(Y_{i1}=j, Y_{i2}=k)] = \pi_{jk} \\ E[I(Y_{i1}=j, Y_{i2}=k)^2] &= P[I(Y_{i1}=j, Y_{i2}=k)] = \pi_{jk} \end{aligned}\right\} \Rightarrow \operatorname{Var}[I(Y_{i1}=j, Y_{i2}=k)] = \pi_{jk}(1-\pi_{jk})$$

Also, $E\{I(Y_{i1}=j, Y_{i2}=k) \cdot I(Y_{i1}=j', Y_{i2}=k')\} = 0$

$$\Rightarrow \operatorname{Cov}\{I(Y_{i1}=j, Y_{i2}=k), I(Y_{i1}=j', Y_{i2}=k')\} = 0 - \pi_{jk}\pi_{j'k'}$$

Then, from the multivariate CLT,

$$\sqrt{n} \begin{pmatrix} \hat{\pi}_{00} - \pi_{00} \\ \hat{\pi}_{01} - \pi_{01} \\ \hat{\pi}_{10} - \pi_{10} \\ \hat{\pi}_{11} - \pi_{11} \end{pmatrix} \xrightarrow{d} N_4 \left(0, \ \underbrace{\Sigma = \begin{bmatrix} \pi_{00}(1-\pi_{00}) & -\pi_{00}\pi_{01} & -\pi_{00}\pi_{10} & -\pi_{00}\pi_{11} \\ \cdot & \pi_{01}(1-\pi_{01}) & -\pi_{01}\pi_{10} & -\pi_{01}\pi_{11} \\ \cdot & \cdot & \pi_{10}(1-\pi_{10}) & -\pi_{10}\pi_{11} \\ \cdot & \cdot & \cdot & \pi_{11}(1-\pi_{11}) \end{bmatrix}}_{= \operatorname{diag}(\underset{\sim}{\pi}) - \underset{\sim}{\pi}\underset{\sim}{\pi}'} \right)$$

$$\underset{\sim}{\pi} = (\pi_{00}, \pi_{01}, \pi_{10}, \pi_{11})$$

(b) $\pi_{1+} = \exp(\alpha) / [1 + \exp(\alpha)] \Rightarrow \alpha = \log\left(\dfrac{\pi_{1+}}{1-\pi_{1+}}\right) \Rightarrow \hat{\alpha} = \log\left(\dfrac{\hat{\pi}_{10} + \hat{\pi}_{11}}{\hat{\pi}_{00} + \hat{\pi}_{01}}\right)$

$\pi_{+1} = \exp(\alpha+\beta) / [1 + \exp(\alpha+\beta)] \Rightarrow \alpha + \beta = \log\left(\dfrac{\pi_{+1}}{1-\pi_{+1}}\right)$

$$\Rightarrow \hat{\beta} = \log\left(\frac{\pi_{+1}}{1-\pi_{+1}}\right) - \log\left(\frac{\pi_{1+}}{1-\pi_{1+}}\right)$$

$$= \log\left(\frac{\pi_{+1}(1-\pi_{1+})}{\pi_{1+}(1-\pi_{+1})}\right) = \log\left(\frac{(\hat{\pi}_{01}+\hat{\pi}_{11})(\hat{\pi}_{00}+\hat{\pi}_{01})}{(\hat{\pi}_{10}+\hat{\pi}_{11})(\hat{\pi}_{00}+\hat{\pi}_{10})}\right)$$

Delta

$$g(a,b,c,d) = \left\{ \log\left(\frac{c+d}{a+b}\right), \log\left[\frac{(b+d)}{(a+c)} \frac{(a+b)}{(c+d)}\right] \right\}'$$

$$= \{\log(c+d) - \log(a+b), \ \log(b+d) + \log(a+b) - \log(a+c) - \log(c+d)\}'$$

[illegible]

By multivariate delta,

$$\sqrt{n}\left(\underbrace{g(\hat{\underset{\sim}{\pi}})}_{=\begin{pmatrix}\hat{\tau}_n\\ \hat{\beta}_n\end{pmatrix}} - \underbrace{g(\underset{\sim}{\pi})}_{=\begin{pmatrix}\tau\\ \beta\end{pmatrix}}\right) \xrightarrow{d} N_2\left(0,\ \Sigma^* = \nabla g\, \Sigma\, \nabla g'\right)$$

$$\Sigma^* = \left[\quad\right]\left[\quad\right]\left[\;\begin{pmatrix}\ \end{pmatrix}\;\right]$$

$\Rightarrow$ asymptotic variance of $\hat{\beta}_n$ is calculated as

$$\left[\underbrace{\frac{\pi_{10}-\pi_{01}}{\pi_{+0}\pi_{0+}}}_{\frac{1}{\pi_{0+}}-\frac{1}{\pi_{+0}}},\ \underbrace{\frac{1}{\pi_{1+}\pi_{0+}}}_{\frac{1}{\pi_{1+}}+\frac{1}{\pi_{0+}}},\ \underbrace{\frac{-1}{\pi_{+0}\pi_{+1}}}_{\frac{-1}{\pi_{+0}}\cdot\frac{1}{\pi_{+1}}},\ \underbrace{\frac{\pi_{01}-\pi_{10}}{\pi_{1+}\pi_{+1}}}_{\frac{1}{\pi_{1+}}-\frac{1}{\pi_{+1}}}\right]\begin{bmatrix}\pi_{00}(1-\pi_{00}) & -\pi_{00}\pi_{01} & -\pi_{00}\pi_{10} & -\pi_{00}\pi_{11}\\ -\pi_{00}\pi_{01} & \pi_{01}(1-\pi_{01}) & -\pi_{01}\pi_{10} & -\pi_{01}\pi_{11}\\ -\pi_{00}\pi_{10} & -\pi_{01}\pi_{10} & \pi_{10}(1-\pi_{10}) & -\pi_{10}\pi_{11}\\ -\pi_{00}\pi_{11} & -\pi_{01}\pi_{11} & -\pi_{10}\pi_{11} & \pi_{11}(1-\pi_{11})\end{bmatrix}\nabla g'$$

$$= \left[\pi_{00}(1-\pi_{00})\left[\frac{1}{\pi_{0+}}-\frac{1}{\pi_{+0}}\right] - \pi_{00}\pi_{01}\left[\frac{1}{\pi_{1+}}+\frac{1}{\pi_{0+}}\right] + \pi_{00}\pi_{10}\left[\frac{1}{\pi_{+0}}+\frac{1}{\pi_{+1}}\right] - \pi_{00}\pi_{11}\left[\frac{1}{\pi_{1+}}-\frac{1}{\pi_{+1}}\right],\right.$$

$$\frac{1}{\pi_{0+}}\left(\pi_{00}(1-\pi_{00}) - \pi_{00}\pi_{01}\right) + \frac{1}{\pi_{+0}}\left(-\pi_{00}(1-\pi_{00}) + \pi_{00}\pi_{10}\right) + \frac{1}{\pi_{1+}}\left(-\pi_{00}\pi_{01} - \pi_{00}\pi_{11}\right) + \frac{1}{\pi_{+1}}\left(\pi_{00}\pi_{10} + \pi_{00}\right.$$

$$\pi_{00}\left(\frac{1-\pi_{0+}}{\pi_{0+}} + \frac{-1+\pi_{10}}{\pi_{+0}} + \frac{\quad}{\pi_{1+}}\right.$$

$$= \left(\frac{1}{\pi_{10}\pi_{0+}\pi_{+1}\pi_{1+}}\right)^2\left[(\pi_{10}-\pi_{01})\pi_{+1}\pi_{1+},\ \pi_{+0}\pi_{+1},\ -\pi_{0+}\pi_{1+},\ (\pi_{01}-\pi_{10})\pi_{0+}\pi_{+0}\right]\begin{bmatrix}\pi_{00}(1-\pi_{00}) & \\ -\pi_{00}\pi_{01} & \\ -\pi_{00}\pi_{10} & \\ -\pi_{00}\pi_{11} & \end{bmatrix}\nabla$$

$$\left(\frac{1}{\quad}\right)^2\left[(\pi_{10}-\pi_{01})\pi_{00}(1-\pi_{00})\pi_{+1}\pi_{1+} - \pi_{+0}\pi_{+1}\pi_{00}\pi_{01} + \pi_{0+}\pi_{1+}\pi_{00}\pi_{10} - (\pi_{01}-\pi_{10})\pi_{0+}\pi_{10}\pi_{00}\pi_{11}\right]$$

$$\frac{\quad}{\quad}\quad \frac{\quad}{\quad}\quad \frac{(\pi_{11}\pi_{00}-\pi_{10}\pi_{01})}{\quad} \quad \text{asy. var}(\hat{\beta}_n)$$

(c) $\underbrace{P(Y_{i1}=1)}_{\pi_{1+,i}} = \dfrac{\exp(\alpha_i)}{1+\exp(\alpha_i)}$, $\underbrace{P(Y_{i2}=1)}_{\pi_{+1,i}} = \dfrac{\exp(\alpha_i+\beta)}{1+\exp(\alpha_i+\beta)}$

Y_{i1} \ Y_{i2}	0	1	
0	π_{00}	π_{01}	$\frac{1}{1+\exp(\alpha_i)}$
1	π_{10}	π_{11}	$\frac{\exp(\alpha_i)}{1+\exp(\alpha_i)}$
	$\frac{1}{1+\exp(\alpha_i+\beta)}$	$\frac{\exp(\alpha_i+\beta)}{1+\exp(\alpha_i+\beta)}$	1

$\pi_{00}+\pi_{01} = \dfrac{1}{1+\exp(\alpha_i)}$

$\pi_{00}+\pi_{10} = \dfrac{1}{1+\exp(\alpha_i+\beta)}$

$\pi_{01}+\pi_{11} = \dfrac{\exp(\alpha_i+\beta)}{1+\exp(\alpha_i+\beta)}$

$\pi_{10}+\pi_{11} = \dfrac{\exp(\alpha_i)}{1+\exp(\alpha_i)}$

$$L(\alpha_i,\beta \mid y_{i1},y_{i2}) \propto \left[\frac{\exp(\alpha_i)}{1+\exp(\alpha_i)}\right]^{y_{i1}} \left[\frac{1}{1+\exp(\alpha_i)}\right]^{1-y_{i1}} \left[\frac{\exp(\alpha_i+\beta)}{1+\exp(\alpha_i+\beta)}\right]^{y_{i2}} \left[\frac{1}{1+\exp(\alpha_i+\beta)}\right]^{1-y_{i2}}$$

$$= \exp\Big\{ y_{i1}\left[\alpha_i - \log(1+\exp(\alpha_i))\right] + (1-y_{i1})\left[-\log(1+\exp(\alpha_i))\right] + y_{i2}\left[\alpha_i+\beta-\log(1+\exp(\alpha_i+\beta))\right] + (1-y_{i2})\left[-\log(1+\exp(\alpha_i+\beta))\right]\Big\}$$

$$= \exp\Big\{ \alpha_i \underbrace{(y_{i1}+y_{i2})}_{=s_i \text{ is sufficient for } \alpha_i} + \beta y_{i2} - \log(1+\exp(\alpha_i)) - \log(1+\exp(\alpha_i+\beta_i)) \Big\}$$

$$\Rightarrow f(y_{i2}\mid s_i) = \frac{\exp\{\cancel{\alpha_i s_i} + \beta y_{i2} - \cancel{\log(\cdot)}\}}{\sum_{y_{i2}} \exp\{\cancel{\alpha_i s_i} + \beta y_{i2} - \cancel{\log(\cdot)}\}} = \frac{\exp\{\beta y_{i2}\}}{\sum_{y_{i2}} \exp\{\beta y_{i2}\}}$$

(so that $s_i = 1$)

$\Rightarrow$ If $s_i = 0 \Rightarrow f(y_{i2}\mid s_i=0) = 1,\ y_{i2}=0$

If $s_i = 1 \Rightarrow f(y_{i2}\mid s_i=1) = \underbrace{\left[\dfrac{e^\beta}{1+e^\beta}\right]^{y_{i2}}}_{\pi_{01}} \underbrace{\left[\dfrac{1}{1+e^\beta}\right]^{1-y_{i2}}}_{\pi_{10}}$

If $s_i = 2 \Rightarrow f(y_{i2}\mid s_i=2) = 1,\ y_{i2}=1$

Thus $f(\underset{\sim}{y_{i2}} \mid \underset{\sim}{s_i}) = \left(\dfrac{e^\beta}{1+e^\beta}\right)^{\sum y_{i2}} \left(\dfrac{1}{1+e^\beta}\right)^{n^*-\sum y_{i2}}$

$\Rightarrow \ell(\beta \mid \underset{\sim}{y_{i2}}, \underset{\sim}{s_i}) = \sum y_{i2}\left[\beta - \log(1+e^\beta)\right] + (n^* - \sum y_{i2})\left[-\log(1+e^\beta)\right]$

$= \beta \sum y_{i2} - n^* \log(1+e^\beta)$

$\dfrac{\partial \ell}{\partial \beta} = \sum y_{i2} - \dfrac{n^* \cdot e^\beta}{1+e^\beta} \Rightarrow \dfrac{e^\beta}{1+e^\beta} = \dfrac{1}{n^*}\sum y_{i2}$, where n^* corresponds to number with $s_i = 1$, or

$n_{01}+n_{10} = n^*$

$\sum y_{i2} = n^*$

... $\dfrac{e^\beta}{1+e^\beta}$, $e^\beta = \dfrac{n_{01}}{\ldots} \quad \dfrac{n_{01}}{n^* - n_{01}}$

... $\dfrac{n_{01}}{\ldots}$

then, $\frac{\partial l^2}{\partial \beta^2} = -n^* \left(\frac{(1+e^\beta)e^\beta - (e^\beta)^2}{(1+e^\beta)^2} \right)$

$$\Rightarrow -\frac{\partial l^2}{\partial \beta} = n^* \frac{e^\beta}{(1+e^\beta)^2} \Rightarrow I_1(\beta) = \frac{1}{n^*} E\left(-\frac{\partial l^2}{\partial \beta}\right) = \frac{e^\beta}{(1+e^\beta)^2} = \frac{e^\beta}{1+e^\beta} \cdot \frac{1}{1+e^\beta}$$

$= \pi_{01} \cdot \pi_{10}$

$$\Rightarrow I^{-1}(\beta) = \frac{1}{\pi_{01} \cdot \pi_{10}} = \frac{\overbrace{\pi_{01}+\pi_{10}}^{=1}}{\pi_{01} \cdot \pi_{10}} = \frac{1}{\pi_{10}} + \frac{1}{\pi_{01}}$$

By MLE theory, $\sqrt{n^*}\left(\hat{\beta}_C - \beta\right) \xrightarrow{d} N_1(0, I^{-1}(\beta))$

(d) $\hat{\beta}_{MLE}$ maximizes

$$l = \sum_{i=1}^{n} \underbrace{\{\alpha_i \cdot s_i + \beta y_{i2} - \log(1+\exp(\alpha_i)) - \log(1+\exp(\alpha_i+\beta))\}}_{l_i}$$

By MLE theory, $\sqrt{n}\left(\hat{\beta}_{MLE} - \beta\right) \xrightarrow{d} N(0, I^{-1}(\beta))$, where $I^{-1}(\beta)$ is found as follows:

$$\frac{\partial l}{\partial \alpha_i} = s_i - \frac{e^{\alpha_i}}{1+e^{\alpha_i}} - \frac{e^{\alpha_i+\beta}}{1+e^{\alpha_i+\beta}} \qquad \frac{\partial^2 l}{\partial \alpha_i^2} = \frac{-e^{\alpha_i}}{(1+e^{\alpha_i})^2} - \frac{e^{\alpha_i+\beta}}{(1+e^{\alpha_i+\beta})^2}$$

$$\frac{\partial l}{\partial \beta} = \sum\left(y_{i2} - \frac{e^{\alpha_i+\beta}}{1+e^{\alpha_i+\beta}}\right) \qquad \frac{\partial^2 l}{\partial \alpha_i \partial \beta} = \frac{-e^{\alpha_i+\beta}}{(1+e^{\alpha_i+\beta})^2}$$

$$\frac{\partial^2 l}{\partial \beta^2} = \frac{-e^{\alpha_i+\beta}}{(1+e^{\alpha_i+\beta})^2}$$

$$\Rightarrow I(\alpha_i, \beta) = \left[\begin{matrix} \frac{e^{\alpha_i}}{(1+e^{\alpha_i})^2} + \frac{e^{\alpha_i+\beta}}{(1+e^{\alpha_i+\beta})^2} & \frac{e^{\alpha_i+\beta}}{(1+e^{\alpha_i+\beta})^2} \\ & \end{matrix} \right.$$

$1 = \frac{e^{\alpha_i}}{1+e^{\alpha_i}} + \frac{e^{\alpha_i+\beta}}{1+e^{\alpha_i+\beta}}$

$\frac{1}{1+e^{\alpha_i+\beta}} = \frac{e^{\alpha_i}}{1+e^{\alpha_i}} \Rightarrow ?$

$$\frac{\partial l}{\partial \alpha_i} = 0 \Rightarrow s_i = \frac{e^{\alpha_i}}{1+e^{\alpha_i}} + \frac{e^{\alpha_i+\beta}}{1+e^{\alpha_i+\beta}}, \quad i = 1, \dots, n$$

$s_i = 0 \Rightarrow \alpha_i = -\infty$

$s_i = 2 \Rightarrow \alpha_i = \infty$

$s_i = 1 \Rightarrow \frac{1}{1+e^{\alpha_i}} = \frac{1}{1+e^{-\alpha_i-\beta}}$

$\Rightarrow \alpha_i = -\alpha_i - \beta$

$\Rightarrow \beta = -2\alpha_i$

$\alpha_i = -\frac{1}{2}\beta$

$$\frac{\partial l}{\partial \beta} = 0 \Rightarrow \underbrace{\sum y_{i2}}_{n_{01}+n_{11}} = \sum\left(\frac{e^{\alpha_i+\beta}}{1+e^{\alpha_i+\beta}}\right)$$

$$\Rightarrow \sum_{i=1}^{n} \frac{e^{\alpha_i+\beta}}{1+e^{\alpha_i+\beta}} = n_{00} \cdot 0 + (n_{01}+n_{10}) \frac{e^{\frac{\beta}{2}}}{1+e^{\frac{\beta}{2}}} + n_{11} \cdot 1 \Rightarrow \hat{\beta}_{MLE} = 2\log\left(\frac{n_{01}}{n_{10}}\right)$$

$2[\log(\hat{\pi}_{01}) - \log(\hat{\pi}_{10})]$

(e) $(\pi_{11}\pi_{0+})^{-1} + (\pi_{11}\pi_{10})^{-1} \le (\pi_{11}\pi_{+0})^{-1} + (\pi_{11}\pi_{0+})^{-1}$ Use this to show

$\dots [\sqrt{n}(\hat{\beta}_{M})] \le \text{var}[\sqrt{n}(\hat{\beta}_C)]$ when $Y_{i1} \perp Y_{i2}$ for $i = 1, \dots, n$ and $\alpha_i = \tau$ $i = 1, \dots$

[illegible]

showing $\pi_{11}\pi_{+0} + \pi_{11}\pi_{0+} \le \pi_{11}\pi_{0+} + \pi_{11}\pi_{+0}$ [illegible]

[illegible] TRUE [illegible]

More on Problem 1

1 (a) $P(Y_{i1}=j, Y_{i2}=k) = \pi_{jk}$ $j=0,1$ $k=0,1$

$$l(\pi) = \log\left(\frac{n!}{n_{00}!\,n_{01}!\,n_{10}!\,n_{11}!}\right) + n_{00}\log\pi_{00} + n_{10}\log\pi_{10} + n_{01}\log\pi_{01} + n_{11}\log\pi_{11}$$

$$= n_{00}\log\pi_{00} + n_{10}\log\pi_{10} + n_{01}\log\pi_{01} + n_{11}\log(1-\pi_{00}-\pi_{01}-\pi_{10}) + \text{const}$$

Set $\begin{cases} \frac{\partial l}{\partial \pi_{00}} = 0 \\ \frac{\partial l}{\partial \pi_{01}} = 0 \\ \frac{\partial l}{\partial \pi_{10}} = 0 \end{cases} \Longrightarrow \hat{\pi}_{jk} = \frac{n_{jk}}{n}$

Let $Z_i = \left(I(y_{i1}=0, y_{i2}=0),\ I(y_{i1}=0, y_{i2}=1)\ \ I(y_{i1}=1, y_{i2}=0),\ I(y_{i1}=1, y_{i2}=1)\right)^T$

$\sim \text{Multinomial}(1, \pi)$

$Cov(Z_i) = diag(\pi) - \pi\pi^T = \Sigma$ $\quad \pi = (\pi_{00}, \pi_{01}, \pi_{10}, \pi_{11})^T$

$E(Z_i) = \pi$

by CLT $\sqrt{n}(\hat{\pi}-\pi)^T = \sqrt{n}\left(\frac{\sum_{i=1}^{n} Z_i}{n} - \pi\right)^T \xrightarrow{d} N(0, Cov(Z_i))$

i.e $\sqrt{n}(\hat{\pi}-\pi)^T \xrightarrow{d} N(0\ \ \Sigma)$

(b) $R = \frac{\pi_{00}\pi_{11}}{\pi_{01}\pi_{10}} \Rightarrow \log R = \log\pi_{00} + \log\pi_{11} - \log\pi_{10} - \log\pi_{01}$

$\hat{R} = \frac{\hat{\pi}_{00}\hat{\pi}_{11}}{\hat{\pi}_{01}\hat{\pi}_{10}} = \frac{n_{00}n_{11}}{n_{01}n_{10}} \Longrightarrow \widehat{\log(R)} = \log\hat{R} = \log\frac{n_{00}n_{11}}{n_{01}n_{10}}$

Let $g(\pi) = \log R$ then $\frac{\partial g}{\partial \pi} = \left(\frac{1}{\pi_{00}}\ \ -\frac{1}{\pi_{01}}\ \ -\frac{1}{\pi_{10}},\ \frac{1}{\pi_{11}}\right)^T$

Apply Delta method

$\sqrt{n}[\log\hat{R} - \log R] = \sqrt{n}[g(\hat{\pi}) - g(\pi)] \xrightarrow{d} [\partial g(\pi)]^T N(0, \Sigma)$

$\xrightarrow{d} N\left(0, [\partial g(\pi)]^T \Sigma [\partial g(\pi)]\right)$

$$[\partial g(\pi)]^T \Sigma [\partial g(\pi)] = \frac{1}{\pi_{00}} + \frac{1}{\pi_{01}} + \frac{1}{\pi_{10}} + \frac{1}{\pi_{11}}$$

(c) $\operatorname{Var} \log \hat{R} = \frac{1}{n}\hat{\sigma}^2 = \frac{1}{n_{00}} + \frac{1}{n_{01}} + \frac{1}{n_{10}} + \frac{1}{n_{11}}$

So 95% CI for $\log R$ is $\log \hat{R} \pm 1.96 \cdot \sqrt{\frac{1}{n_{00}} + \frac{1}{n_{01}} + \frac{1}{n_{10}} + \frac{1}{n_{11}}}$

then 95% CI for R is

$$\left[\frac{n_{00} n_{11}}{n_{01} n_{10}} e^{-1.96 \cdot \sqrt{\frac{1}{n_{00}} + \frac{1}{n_{01}} + \frac{1}{n_{10}} + \frac{1}{n_{11}}}} , \frac{n_{00} n_{11}}{n_{01} n_{10}} e^{1.96 \sqrt{\frac{1}{n_{00}} + \frac{1}{n_{01}} + \frac{1}{n_{10}} + \frac{1}{n_{11}}}} \right]$$

(d) $\alpha = \operatorname{logit}(\pi_{1+}) = \log \frac{\pi_{10} + \pi_{11}}{\pi_{01} + \pi_{00}}$ and $\alpha + \beta = \operatorname{logit}(\pi_{+1}) = \log \frac{\pi_{01} + \pi_{11}}{\pi_{10} + \pi_{00}}$

$$\Rightarrow \beta = \log\left(\frac{\pi_{01} + \pi_{11}}{\pi_{00} + \pi_{10}} \cdot \frac{\pi_{01} + \pi_{00}}{\pi_{10} + \pi_{11}} \right)$$

by the invariance of MLE. we have

$$\hat{\alpha} = \log \frac{n_{10} + n_{11}}{n_{01} + n_{00}} \qquad \hat{\beta} = \log\left(\frac{n_{01} + n_{11}}{n_{00} + n_{10}} \quad \frac{n_{01} + n_{00}}{n_{10} + n_{11}} \right)$$

(e) $$\frac{\partial(\alpha, \beta)}{\partial \pi} = \begin{pmatrix} -\frac{1}{\pi_{0+}} , & -\frac{1}{\pi_{0+}} , & \frac{1}{\pi_{1+}} & \frac{1}{\pi_{1+}} \\ \frac{1}{\pi_{0+}} - \frac{1}{\pi_{+0}} , & \frac{1}{\pi_{0+}} + \frac{1}{\pi_{+1}} , & -\frac{1}{\pi_{+0}} - \frac{1}{\pi_{1+}} & \frac{1}{\pi_{+1}} - \frac{1}{\pi_{1+}} \end{pmatrix}$$

Apply Delta method we can get

$$\sqrt{n}\left[\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} - \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \right] \xrightarrow{d} N(0, \Sigma_1)$$

where $\Sigma_1 = \frac{\partial(\alpha,\beta)}{\partial \pi} \Sigma \left(\frac{\partial(\alpha,\beta)}{\partial \pi} \right)^T$

$$= \frac{\partial(\alpha,\beta)}{\partial \pi} \operatorname{diag}(\pi) \left(\frac{\partial(\alpha,\beta)}{\partial \pi} \right)^T$$

since $\frac{\partial(\alpha,\beta)}{\partial \pi} \pi = 0$

(f) In order to show $(\pi_{1+}\pi_{0+})^{-1} + (\pi_{+1}\pi_{+0})^{-1} \le (\pi_{1+}\pi_{+0})^{-1} + (\pi_{+1}\pi_{0+})^{-1}$

we can show $\pi_{+1}\pi_{+0} + \pi_{1+}\pi_{0+} \le \pi_{+1}\pi_{0+} + \pi_{1+}\pi_{+0}$

i.e $\pi_{+1}(\pi_{+0} - \pi_{0+}) \le \pi_{1+}(\pi_{+0} - \pi_{0+})$

i.e $(\pi_{+0} - \pi_{0+})(\pi_{+1} - \pi_{1+}) \le 0$

i.e $(\pi_{10} + \pi_{00} - \pi_{00} - \pi_{01})(\pi_{11} + \pi_{01} - \pi_{10} - \pi_{11}) \le 0$

i.e. $(\pi_{10} - \pi_{01})(\pi_{01} - \pi_{10}) \le 0$

i.e $-(\pi_{10} - \pi_{01})^2 \le 0$

Since $-(\pi_{10} - \pi_{01})^2 \le 0$ is always true

So $(\pi_{1+}\pi_{0+})^{-1} + (\pi_{+1}\pi_{+0})^{-1} \le (\pi_{1+}\pi_{+0})^{-1} + (\pi_{+1}\pi_{0+})^{-1}$

2009 THEORY 2 QUESTION 2

$Y_{ij} = \mu_i + (X_{ij} - \bar{X}_i)\gamma_i + \varepsilon_{ij}$, $i = 1,2$, $j = 1, \dots, n_i$ $\quad \varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma^2)$

$$\begin{cases} Y_{11} = \mu_1 + (X_{11} - \bar{X}_1)\gamma_1 + \varepsilon_{11} \\ Y_{12} = \mu_1 + (X_{12} - \bar{X}_1)\gamma_1 + \varepsilon_{12} \\ \vdots \\ Y_{1n_1} = \mu_1 + (X_{1n_1} - \bar{X}_1)\gamma_1 + \varepsilon_{1n_1} \\ Y_{21} = \mu_2 + (X_{21} - \bar{X}_2)\gamma_2 + \varepsilon_{21} \\ Y_{22} = \mu_2 + (X_{22} - \bar{X}_2)\gamma_2 + \varepsilon_{22} \\ \vdots \\ Y_{2n_2} = \mu_2 + (X_{2n_2} - \bar{X}_2)\gamma_2 + \varepsilon_{2n_2} \end{cases}$$

(a) Let $\beta = (\mu_1, \gamma_1, \mu_2, \gamma_2)'$. We can write the model as:

$$\underbrace{\begin{bmatrix} Y_{11} \\ \vdots \\ Y_{1n_1} \\ Y_{21} \\ \vdots \\ Y_{2n_2} \end{bmatrix}}_{Y_{(n_1+n_2)\times 1}} = \underbrace{\begin{bmatrix} 1 & X_{11} - \bar{X}_1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & X_{1n_1} - \bar{X}_1 & 0 & 0 \\ 0 & 0 & 1 & X_{21} - \bar{X}_2 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & X_{2n_2} - \bar{X}_2 \end{bmatrix}}_{X_{(n_1+n_2)\times 4}} \underbrace{\begin{bmatrix} \mu_1 \\ \gamma_1 \\ \mu_2 \\ \gamma_2 \end{bmatrix}}_{\beta_{4\times 1}} + \underbrace{\begin{bmatrix} \varepsilon_{11} \\ \vdots \\ \varepsilon_{1n_1} \\ \varepsilon_{21} \\ \vdots \\ \varepsilon_{2n_2} \end{bmatrix}}_{\varepsilon_{(n_1+n_2)\times 1}} \qquad (*)$$

(i) X is specified above. $\varepsilon_{(n_1+n_2)\times 1} \sim N_{(n_1+n_2)}\left(0, \sigma^2 I_{(n_1+n_2)\times(n_1+n_2)}\right)$

(ii) By the Gauss-Markov theorem, the BLUE (best (minimum variance) linear unbiased estimator) of β is the least squares estimator of β, or $\hat{\beta} = (X'X)^{-1}X'Y$, $(X'X)^{-1}$ exists, as X is full rank.

(b) (i) Consider $a = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix}$. Then $a'\beta = \gamma_1 - \gamma_2$. $a'\beta$ is estimable if and only if we can find a ρ such that $\rho' X = a'$, or

$$[\rho_1 \; \rho_2 \; \cdots \; \rho_{n_1+n_2}] \begin{bmatrix} 1 & X_{11} - \bar{X}_1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & X_{1n_1} - \bar{X}_1 & 0 & 0 \\ 0 & 0 & 1 & X_{21} - \bar{X}_2 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & X_{2n_2} - \bar{X}_2 \end{bmatrix} = [0 \; 1 \; 0 \; -1] \ni$$

full rank X

$\rho_1, \rho_2, \dots, \rho_{n_1+n_2}$ must satisfy:

(1) $\sum_{i=1}^{n_1} \rho_i = 0$

(2) $\sum_{i=n_1+1}^{n_2} \rho_i = 0$

(3) $\sum_{i=1}^{n_1} \rho_i (X_{1i} - \bar{X}_1) = 1$

(4) $\sum_{i=n_1+1}^{n_2} \rho_i (X_{2i} - \bar{X}_2) = -1$

We find $\rho_1, \dots, \rho_{n_1+n_2}$ to satisfy (1), (2), (3), (4), then $\gamma_1 - \gamma_2$ is estimable.

(ii) Because $\varepsilon \sim N(0, \sigma^2 I)$, $Y = X\beta + \varepsilon \sim N(X\beta, \sigma^2 I)$, and $\hat{\beta} = (X'X)^{-1}X'Y$ [illegible] mean β, variance $\sigma^2(X'X)^{-1}$) [illegible] $a'\hat{\beta} \sim N(a'\beta, \sigma^2 a'(X'X)^{-1}a)$

thus, a $(1-\alpha)$ level confidence interval for $a'\beta = \gamma_1 - \gamma_2$ is given by

$$\left\{ \gamma_1 - \gamma_2 : \left| \frac{a'\hat{\beta} - (\gamma_1 - \gamma_2)}{\sqrt{\sigma^2 \cdot a'(X'X)^{-1} a}} \right| < z_{1-\frac{\alpha}{2}} \right\} \quad \text{where } z_{1-\frac{\alpha}{2}} = \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$$

or $a'\hat{\beta} \pm z_{1-\frac{\alpha}{2}} \sqrt{\sigma^2 \cdot a'(X'X)^{-1} a}$

(iii) σ^2 unknown. Define $M = X(X'X)^{-1}X'$. As $\operatorname{rank}(X) = 4$, $\hat{\sigma}^2 = \dfrac{Y'(I-M)Y}{n_1+n_2-4}$ is an unbiased estimator of σ^2.

We wish to find a statistic to test $H_0: \gamma_1 = \gamma_2$. Can use theory of nested models. Under H_0, we can rewrite $(*)$ as (with $\gamma_1 = \gamma_2 = \gamma$).

$$\underbrace{\begin{bmatrix} Y_{11} \\ \vdots \\ Y_{1n_1} \\ Y_{21} \\ \vdots \\ Y_{2n_2} \end{bmatrix}}_{Y} = \underbrace{\begin{bmatrix} 1 & 0 & (X_{11} - \bar{X}) \\ \vdots & \vdots & \\ 1 & 0 & (X_{1n_1} - \bar{X}) \\ 0 & 1 & (X_{21} - \bar{X}) \\ \vdots & \vdots & \\ 0 & 1 & (X_{2n_2} - \bar{X}) \end{bmatrix}}_{X_0} \underbrace{\begin{bmatrix} \mu_1 \\ \mu_2 \\ \gamma \end{bmatrix}}_{\gamma_0} + \underbrace{\begin{bmatrix} \varepsilon_{11} \\ \vdots \\ \varepsilon_{1n_1} \\ \varepsilon_{21} \\ \vdots \\ \varepsilon_{2n_2} \end{bmatrix}}_{\varepsilon} \qquad (*)(*)$$

As the null hypothesis produces a nested model with orthogonal projection operator $M_0 = X_0(X_0'X_0)^{-1}X_0'$, can test H_0 using the statistic:

$$F = \frac{\left[\dfrac{Y'(M-M_0)Y}{r(M)-r(M_0)}\right]}{\hat{\sigma}^2} \sim F_{\underbrace{r(M)-r(M_0)}_{=1},\, n_1+n_2-4} \quad (\text{under } H_0)$$

(c) (i) Results change. In (b)(ii), $\hat{\beta} \sim N_2(\beta, \sigma^2(X_1'X_1)^{-1})$, where $\beta = [\gamma_1, \gamma_2]$,

$$X_1 = \begin{bmatrix} (X_{11} - \bar{X}_1) & 0 \\ \vdots & \vdots \\ (X_{1n_1} - \bar{X}_1) & 0 \\ 0 & (X_{21} - \bar{X}_2) \\ \vdots & \\ 0 & (X_{2n_2} - \bar{X}_2) \end{bmatrix}. \quad \text{thus } a'\hat{\beta} \sim N_2(\underbrace{a'\beta}_{=\gamma_1 - \gamma_2}, \sigma^2 a'(X_1'X_1)^{-1} a).$$

Is $a'\beta$ estimable here? Yes

In (b)(iii), this additional restriction produces a different nested model

$$\begin{bmatrix} Y_{11} \\ \vdots \\ Y_{1n_1} \\ Y_{21} \\ \vdots \\ Y_{2n_2} \end{bmatrix} = \underbrace{\begin{bmatrix} (X_{11} - \bar{X}) \\ \vdots \\ (X_{1n_1} - \bar{X}) \\ (X_{21} - \bar{X}) \\ \vdots \\ (X_{2n_2} - \bar{X}) \end{bmatrix}}_{X_2} \underbrace{[\gamma]}_{\gamma_2} + \underbrace{\begin{bmatrix} \varepsilon_{11} \\ \vdots \\ \varepsilon_{1n_1} \\ \varepsilon_{21} \\ \vdots \\ \varepsilon_{2n_2} \end{bmatrix}}_{\varepsilon}$$

Let $M_2 = X_2(X_2'X_2)^{-1}X_2'$

and a different test statistic

$$F = \frac{\left[\dfrac{Y'(M-M_2)Y}{r(M)-r(M_2)}\right]}{\hat{\sigma}^2} \sim F_{\underbrace{r(M)-r(M_2)}_{4-1=3},\, n_1+n_2-4} \quad (\text{under } H_0)$$

(c)(ii) Under the constraint $\delta'\beta = \gamma_1 - \gamma_2 = 0$ [or $\gamma_1 = \gamma_2 = \gamma$], the nested model is given as (*)(*) on the previous page. With $\beta = [\mu_1, \mu_2, \gamma]'$, the least squares estimate of β is $\hat\beta = (X_0'X_0)^{-1}X_0'Y$, where [with $z_i = (x_i - \bar x)$]

$$X_0'X_0 = \begin{bmatrix} 1 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 & \cdots & 1 \\ z_{11} & \cdots & z_{1n_1} & z_{21} & \cdots & z_{2n_2} \end{bmatrix}\begin{bmatrix} 1 & 0 & z_{11} \\ \vdots & \vdots & \\ 1 & 0 & z_{1n_1} \\ 0 & 1 & z_{21} \\ \vdots & \vdots & \\ 0 & 1 & z_{2n_2} \end{bmatrix} = \begin{bmatrix} n_1 & 0 & \sum_{i=1}^{n_1} z_i = 0 \\ 0 & n_2 & \sum z_i = 0 \\ \sum_{i=1}^{n_1} z_i & \sum_{n_1+1} z_i & \sum_{i=1}^{n_1+n_2} z_i^2 \end{bmatrix} = \begin{bmatrix} n_1 & 0 & a \\ 0 & n_2 & b \\ a & b & c \end{bmatrix} \quad \begin{matrix} A & B \\ B' & C \end{matrix}$$

$= 0$

$$(X_0'X_0)^{-1} = \begin{bmatrix} (A - BC^{-1}B')^{-1} & -A^{-1}B(C - B'A^{-1}B)^{-1} \\ -C^{-1}B'(A - BC^{-1}B')^{-1} & (C - B'A^{-1}B)^{-1} \end{bmatrix}$$

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

$$\begin{bmatrix} \left\{\begin{bmatrix} n_1 & 0 \\ 0 & n_2 \end{bmatrix} - \begin{bmatrix} a \\ b \end{bmatrix} C^{-1} [a\ b]\right\}^{-1} & -\begin{bmatrix} n_1 & 0 \\ 0 & n_2 \end{bmatrix}^{-1}\begin{bmatrix} a \\ b \end{bmatrix}\left\{C - [a\ b]\begin{bmatrix} n_1 & 0 \\ 0 & n_2 \end{bmatrix}^{-1}\begin{bmatrix} a \\ b \end{bmatrix}\right\}^{-1} \\ -C^{-1}(a\ b)\underbrace{\left\{\begin{bmatrix} n_1 & 0 \\ 0 & n_2 \end{bmatrix} - \begin{bmatrix} a \\ b \end{bmatrix} C^{-1} [a\ b]\right\}^{-1}}_{(1)} & \left\{C - [a\ b]'\begin{bmatrix} n_1 & 0 \\ 0 & n_2 \end{bmatrix}^{-1}\begin{bmatrix} a \\ b \end{bmatrix}\right\}^{-1} \cdot (2) \end{bmatrix}$$

$$(X_0'Y) = \begin{bmatrix} 1 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 & \cdots & 1 \\ z_{11} & \cdots & z_{1n_1} & z_{21} & \cdots & z_{2n_2} \end{bmatrix}\begin{bmatrix} Y_{11} \\ \vdots \\ Y_{1n_1} \\ Y_{21} \\ \vdots \\ Y_{2n_2} \end{bmatrix} = \begin{bmatrix} \sum^{n_1} y_i \\ \sum^{n_2} y_i \\ \sum z_i y_i \end{bmatrix} = \begin{bmatrix} j \\ k \\ l \end{bmatrix}$$

Then $\hat\beta = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}\begin{bmatrix} j \\ k \\ l \end{bmatrix} = \begin{bmatrix} aj + bk + cl \\ dj + ek + fl \\ gj + hk + il \end{bmatrix} \quad \begin{matrix} \hat\mu_1 \\ \hat\mu_2 \\ = \hat\gamma \end{matrix}$ ← need these as scalars

For $\hat\gamma$, need (1) $\left(-\sum_{1}^{n_1+n_2} z_i^2\right)^{-1}\left[\sum_1^{n_1} z_i \ \sum_{n_1+1}^{n_2} z_i\right]\left\{\begin{matrix} n_1 - (\sum z_i^2)^{-1}(\sum^{n_1} z_i)^2, & \\ -(\sum z_i)(\sum z_i)(\sum z_i^2)^{-1} & n_2 - (\sum z_i^2)^{-1}(\sum z_i)^2 \end{matrix}\right\}^{-1}$

$\hat\mu_1, \hat\mu_2$ and $\hat\gamma$

(a) (b) (d)

$$\frac{1}{ad - b^2}\begin{bmatrix} d & -b \\ -b & a \end{bmatrix}$$

$$\left(-\sum z_i^2\right)^{-1}\left[\sum z_i\,(d) - \sum z_i\,(b),\quad -b\sum z_i + \sum z_i\,(a)\right]$$

(2) $\left\{\sum z_i^2 - \left[\sum z_i \ \sum z_i\right]'\begin{bmatrix} n_1 & 0 \\ 0 & n_2 \end{bmatrix}^{-1}\begin{bmatrix} \sum z_i \\ \sum z_i \end{bmatrix}\right\}^{-1}$

[illegible]

[illegible] efficient

[illegible]

(a) Y_ℓ = ℓ^{th} row of Y, X_ℓ' = ℓ^{th} row of X. Interested in F-test for hyp. that Y_k is not an outlier. LSE without Y_k, X_k' is denoted $\hat\beta_{(k)}$.

(i) $\hat\beta_{(k)} = \left(X_{(k)}' X_{(k)}\right)^{-1} X_{(k)}' Y_{(k)}$

need expression not involving $X_{(k)}$. Use result for inverses given in BIOS 763. Forget.

(ii) $D_k = Y_k - X_k' \hat\beta_{(k)}$. Y_k assumed $\sim N_1(X_k'\beta, \sigma^2)$

$\hat\beta_{(k)} \sim N_p\left(\beta, (X_{(k)}'X_{(k)})^{-1}\sigma^2\right) \Rightarrow -X_k'\hat\beta_{(k)}$

$\sim N_1\left(-X_k'\beta, \sigma^2 X_k'(X_{(k)}'X_{(k)})^{-1}X_k\right)$

$Y_k \perp Y_{(k)}$ so $Y_k \perp \hat\beta_{(k)}$

Have D_k as a sum of two independent normals.

$$\Rightarrow D_k \sim N_1\Big(\underbrace{X_k'\beta - X_k'\beta}_{=0},\ \underbrace{\sigma^2 + \sigma^2 X_k'(X_{(k)}'X_{(k)})^{-1}X_k}_{=\sigma^2\left(1 + X_k'(X_{(k)}'X_{(k)})^{-1}X_k\right)}\Big)$$

(iii) F-test for H_0: $E(Y_k) = X_k'\beta$

Under H_0, $D_k \sim N_1\Big(0, \underbrace{\sigma^2[1 + X_k'(X_{(k)}'X_{(k)})^{-1}X_k]}_{\sigma_*^2}\Big) \Rightarrow \frac{D_k}{\sigma_*} \sim N(0,1) \Rightarrow \left(\frac{D_k}{\sigma_*}\right)^2 \sim \chi^2_1$

Under H_1, $E(Y_k) \neq X_k'\beta$

$\left(\frac{D_k}{\sigma_*}\right)^2 \sim \chi^2_1$ under H_0. Must estimate σ^2 as σ^2 unknown.

Use $\hat\sigma^2 = \dfrac{Y'(I-M)Y}{n_1+n_2-4}$. $\hat\sigma^2 \perp$ of $\dfrac{D_k}{\sigma_*}$

Then, $$\frac{\left(\frac{D_k}{\sigma_*}\right)^2}{\hat\sigma^2} \sim F_{1, n_1+n_2-4} \text{ under } H_0$$

Estimability of β

(b) $\rho' X = \lambda' = I$

choose $\rho' = (X'X)^{-1}X'$ as inverse exists. done

[illegible] λ_1' [illegible] ρ_1' [illegible] $[\lambda_1'(X'X)^{-1}X']X = \lambda_1'$. done

2009 THEORY 2 QUESTION 3

risk factor X on count variable Y. For $k=1,2$, (Y_{ik}, X_{ik}), $i=1,\dots,n$, are iid from $f(X_{ik}, Y_{ik}) = \underbrace{f(Y_{ik} \mid X_{ik})}_{\sim \text{Poisson}(\lambda_{ik})} \underbrace{f(X_{ik})}_{\sim N(0,\sigma_k^2)}$

$\lambda_{ik} = \exp(\alpha_k + \beta X_{ik})$

$(\sigma_1^2, \sigma_2^2, \alpha_1, \alpha_2, \beta)$ are unknown

(a) the likelihood of all of the data from the two centers is:

$$L(\sigma_1^2, \sigma_2^2, \alpha_1, \alpha_2, \beta) = \prod_{k=1}^{2} \prod_{i=1}^{n} f(Y_{ik} \mid X_{ik}) f(X_{ik})$$

$$= \prod_{k=1}^{2} \prod_{i=1}^{n} \frac{(\lambda_{ik})^{y_{ik}}}{(y_{ik}!)} e^{-\lambda_{ik}} (2\pi\sigma_k^2)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma_k^2} X_{ik}^2\right\}, \quad \lambda_{ik} = \exp(\alpha_k + \beta X_{ik})$$

$$= \prod_{k=1}^{2} \prod_{i=1}^{n} \frac{(2\pi\sigma_k^2)^{-\frac{1}{2}}}{y_{ik}!} \exp\left\{-\lambda_{ik} + y_{ik}(\alpha_k + \beta X_{ik}) - \frac{1}{2\sigma_k^2} X_{ik}^2\right\}$$

$$= \prod_{k=1}^{2} \left(\prod_{i=1}^{n} \frac{1}{y_{ik}!}\right) (2\pi\sigma_k^2)^{-\frac{n}{2}} \exp\left\{-\sum_{i=1}^{n} e^{\alpha_k + \beta X_{ik}} + \sum_{i=1}^{n} y_{ik}(\alpha_k + \beta X_{ik}) - \frac{1}{2\sigma_k^2} \sum_{i=1}^{n} X_{ik}^2\right\}$$

$$= \left(\prod_{i=1}^{n} \frac{1}{y_{i1}!}\right)\left(\prod_{i=1}^{n} \frac{1}{y_{i2}!}\right) (2\pi\sigma_1^2)^{-\frac{n}{2}} (2\pi\sigma_2^2)^{-\frac{n}{2}} \exp\Big\{-\sum_{i=1}^{n} e^{\alpha_1 + \beta X_{i1}} - \sum_{i=1}^{n} e^{\alpha_2 + \beta X_{i2}} + \alpha_1 \sum_{i=1}^{n} y_{i1}$$
$$+ \alpha_2 \sum_{i=1}^{n} y_{i2} + \beta \sum_{i=1}^{n} (y_{i1} X_{i1} + y_{i2} X_{i2})$$
$$- \frac{1}{2\sigma_1^2} \sum_{i=1}^{n} X_{i1}^2 - \frac{1}{2\sigma_2^2} \sum_{i=1}^{n} X_{i2}^2 \Big\}$$

so that the log-likelihood is proportional to:

$$\ell(\sigma_1^2, \sigma_2^2, \alpha_1, \alpha_2, \beta) \propto -\frac{n}{2}\log\sigma_1^2 - \frac{n}{2}\log\sigma_2^2 - \sum_{i=1}^{n} e^{\alpha_1 + \beta X_{i1}} - \sum_{i=1}^{n} e^{\alpha_2 + \beta X_{i2}} + \alpha_1 \sum_{i=1}^{n} y_{i1} + \alpha_2 \sum_{i=1}^{n} y_{i2} + \beta \sum_{i=1}^{n} (y_{i1} X_{i1} + y_{i2} X_{i2})$$
$$- \frac{1}{2\sigma_1^2} \sum_{i=1}^{n} X_{i1}^2 - \frac{1}{2\sigma_2^2} \sum_{i=1}^{n} X_{i2}^2$$

Must find Fisher information, I

$\dot{\ell}(\sigma_1^2) = -\frac{n}{2\sigma_1^2} + \frac{1}{2\sigma_1^4} \sum X_{i1}^2$

$\dot{\ell}(\sigma_2^2) = -\frac{n}{2\sigma_2^2} + \frac{1}{2\sigma_2^4} \sum X_{i2}^2$

$\dot{\ell}(\alpha_1) = -\sum e^{\alpha_1 + \beta X_{i1}} + \sum_{i=1}^{n} y_{i1}$

$\dot{\ell}(\alpha_2) = -\sum e^{\alpha_2 + \beta X_{i2}} + \sum y_{i2}$

$\dot{\ell}(\beta) = -\sum (e^{\alpha_1 + \beta X_{i1}}) X_{i1} - \sum_{i=1}^{n} (e^{\alpha_2 + \beta X_{i2}}) X_{i2} + \sum_{i=1}^{n} (y_{i1} X_{i1} + y_{i2} X_{i2})$

$$\ddot{\ell} = \begin{bmatrix} \frac{n}{2\sigma_1^4} - \frac{1}{\sigma_1^6}\sum X_{i1}^2 & 0 & 0 & 0 & 0 \\ 0 & \frac{n}{2\sigma_2^4} - \frac{1}{\sigma_2^6}\sum X_{i2}^2 & 0 & 0 & 0 \\ 0 & 0 & -\sum e^{\alpha_1 + \beta X_{i1}} & 0 & -\sum (e^{\alpha_1 + \beta X_{i1}}) X_{i1} \\ 0 & 0 & 0 & -\sum e^{\alpha_2 + \beta X_{i2}} & -\sum (e^{\alpha_2 + \beta X_{i2}}) X_{i2} \\ [illegible] & [illegible] & [illegible] & [illegible] & [illegible] \end{bmatrix}$$

Note: $E(X_{1i}^2) = \sigma_1^2$, $E(X_{2i}^2) = \sigma_2^2$

$$E\left(e^{a_1+\beta X_{1i}}\right) = e^{a_1}\int_{-\infty}^{\infty}(2\pi\sigma_1^2)^{-\frac{1}{2}} \exp\left\{\frac{-1}{2\sigma_1^2}X_{1i}^2 + \beta X_{1i}\right\}$$

$$= e^{a_1}\int_{-\infty}^{\infty}(2\pi\sigma_1^2)^{-\frac{1}{2}} \exp\left\{\frac{-1}{2\sigma_1^2}\left(X_{1i}^2 - 2\beta\sigma_1^2 X_{1i} + \beta^2\sigma_1^4\right) + \frac{\beta^2\sigma_1^4}{2\sigma_1^2}\right\}$$

$$= e^{a_1+\frac{\beta^2\sigma_1^2}{2}}\int_{-\infty}^{\infty}\text{density of a } N\left(\beta\sigma_1^2, \sigma_1^2\right) \text{ random variable}$$

$$\Rightarrow E\left[\left(e^{a_1+\beta X_{1i}}\right)\cdot X_{1i}\right] = e^{a_1+\frac{\beta^2\sigma_1^2}{2}}\cdot E(X_{1i}) = \beta\sigma_1^2 e^{a_1+\frac{\beta^2\sigma_1^2}{2}}$$

$$\Rightarrow E\left[\left(e^{a_1+\beta X_{1i}}\right)X_{1i}^2\right] = e^{a_1+\frac{\beta^2\sigma_1^2}{2}}\cdot E(X_{1i}^2) = (\sigma_1^2+\beta^2\sigma_1^4)\, e^{a_1+\frac{\beta^2\sigma_1^2}{2}}$$

$$\Rightarrow I_{2n} = \begin{array}{c} \qquad\qquad A \qquad\qquad\qquad\qquad\qquad B \\ \left[\begin{array}{cccc|c}
\frac{n}{2\sigma_1^4} & 0 & 0 & 0 & 0 \\
0 & \frac{n}{2\sigma_2^4} & 0 & 0 & 0 \\
0 & 0 & ne^{a_1+\frac{\beta^2\sigma_1^2}{2}} & 0 & n\beta\sigma_1^2 e^{a_1+\frac{\beta^2\sigma_1^2}{2}} \\
0 & 0 & 0 & ne^{a_2+\frac{\beta^2\sigma_2^2}{2}} & n\beta\sigma_2^2 e^{a_2+\frac{\beta^2\sigma_2^2}{2}} \\
0 & 0 & n\beta\sigma_1^2 e^{a_1+\frac{\beta^2\sigma_1^2}{2}} & n\beta\sigma_2^2 e^{a_2+\frac{\beta^2\sigma_2^2}{2}} & n\left[(\sigma_1^2+\beta^2\sigma_1^4)e^{a_1+\frac{\beta^2\sigma_1^2}{2}} + (\sigma_2^2+\beta^2\sigma_2^4)e^{a_2+\frac{\beta^2\sigma_2^2}{2}}\right]
\end{array}\right] \\ \qquad\qquad\qquad\qquad B' \qquad\qquad\qquad\qquad\qquad C \end{array}$$

so that

$$I_1 = \frac{1}{2n}I_{2n} = \begin{array}{c} \qquad\qquad A \quad B \\ \left[\begin{array}{cccc|c}
\frac{1}{4\sigma_1^4} & 0 & 0 & 0 & 0 \\
0 & \frac{1}{4\sigma_2^4} & 0 & 0 & 0 \\
0 & 0 & \frac{1}{2}e^{a_1+\frac{\beta^2\sigma_1^2}{2}} & 0 & \frac{1}{2}\beta\sigma_1^2 e^{a_1+\frac{\beta^2\sigma_1^2}{2}} \\
0 & 0 & 0 & \frac{1}{2}e^{a_2+\frac{\beta^2\sigma_2^2}{2}} & \frac{1}{2}\beta\sigma_2^2 e^{a_2+\frac{\beta^2\sigma_2^2}{2}} \\
\hline
0 & 0 & \frac{1}{2}\beta\sigma_1^2 e^{a_1+\frac{\beta^2\sigma_1^2}{2}} & \frac{1}{2}\beta\sigma_2^2 e^{a_2+\frac{\beta^2\sigma_2^2}{2}} & \frac{1}{2}\left[(\sigma_1^2+\beta^2\sigma_1^4)e^{a_1+\frac{\beta^2\sigma_1^2}{2}} + (\sigma_2^2+\beta^2\sigma_2^4)e^{a_2+\frac{\beta^2\sigma_2^2}{2}}\right]
\end{array}\right] \\ \qquad\qquad B' \quad C \\ \qquad b_1 \qquad b_2 \end{array}$$

$$I^{-1}(\beta) = (C - B'A^{-1}B)^{-1}$$

$$= \left\{\frac{1}{2}\left[(\sigma_1^2+\beta^2\sigma_1^4)e^{a_1+\frac{\beta^2\sigma_1^2}{2}} + (\sigma_2^2+\beta^2\sigma_2^4)e^{a_2+\frac{\beta^2\sigma_2^2}{2}}\right] - \begin{pmatrix}0 & 0 & b_1 & b_2\end{pmatrix}\begin{bmatrix}\frac{1}{a_1} & 0 & 0 & 0 \\ 0 & \frac{1}{a_2} & 0 & 0 \\ 0 & 0 & \frac{1}{a_3} & 0 \\ 0 & 0 & 0 & \frac{1}{a_4}\end{bmatrix}\begin{bmatrix}0 \\ 0 \\ b_1 \\ b_2\end{bmatrix}\right\}^{-1}$$

$$\begin{pmatrix}0 & 0 & \frac{b_1}{a_3} & \frac{b_2}{a_4}\end{pmatrix}\begin{pmatrix}0 \\ 0 \\ b_1 \\ b_2\end{pmatrix} \qquad \frac{b_1^2}{a_3} + \frac{b_2^2}{a_4}$$

$$= \left\{\quad - \frac{1}{2}\beta^2\sigma_1^4 e^{a_1+\frac{\beta^2\sigma_1^2}{2}} - \frac{1}{2}\beta^2\sigma_2^4 e^{a_2+\frac{\beta^2\sigma_2^2}{2}}\right\}^{-1}$$

$$\left\{\frac{1}{2}\left(\sigma_1^2 e^{a_1+\frac{\beta^2\sigma_1^2}{2}} + \sigma_2^2 e^{a_2+\frac{\beta^2\sigma_2^2}{2}}\right)\right\}^{-1} \qquad \frac{2}{\sigma_1^2 e^{a_1+\frac{\beta^2\sigma_1^2}{2}} + \sigma_2^2 e^{a_2+\frac{\beta^2\sigma_2^2}{2}}}$$

By MLE theory, $\sqrt{2n}(\hat{\beta}-\beta) \xrightarrow{d} N_1\left(0, \frac{2}{\sigma_1^2 e^{a_1+\frac{\beta^2\sigma_1^2}{2}} + \sigma_2^2 e^{a_2+\frac{\beta^2\sigma_2^2}{2}}}\right)$

b) individual level data unavailable, but researchers report MLEs from cohort 1 and 2

(i) For $k=1,2$

$$L(\sigma_k^2, \alpha_k, \beta) \propto \prod_{i=1}^{n}\left(\frac{1}{y_{ik}!}\right)(2\pi\sigma_k^2)^{-n/2} \exp\left\{-\sum_{i=1}^{n} e^{\alpha_k+\beta x_{ik}} + \sum_{i=1}^{n} y_{ik}(\alpha_k+\beta x_{ik}) - \frac{1}{2\sigma_k^2}\sum_{i=1}^{n} x_{ik}^2\right\}$$

$$\Rightarrow l(\sigma_k^2, \alpha_k, \beta) \propto -\frac{n}{2}\log\sigma_k^2 - \sum_{i=1}^{n} e^{\alpha_k+\beta x_{ik}} + \alpha_k\sum_{i=1}^{n} y_{ik} + \beta\sum_{i=1}^{n} y_{ik}x_{ik} - \frac{1}{2\sigma_k^2}\sum x_{ik}^2$$

$$\dot{l}(\sigma_k^2) = \frac{-n}{2\sigma_k^2} + \frac{1}{2\sigma_k^4}\sum x_{ik}^2$$

$$\dot{l}(\alpha_k) = -\sum e^{\alpha_k+\beta x_{ik}} + \sum y_{ik}$$

$$\dot{l}(\beta) = -\sum(e^{\alpha_k+\beta x_{ik}})x_{ik} + \sum y_{ik}x_{ik}$$

$$\Rightarrow I_n = -E\begin{bmatrix} \frac{n}{2\sigma_k^4} - \frac{1}{\sigma_k^6}\sum x_{ik}^2 & 0 & 0 \\ 0 & -\sum e^{\alpha_k+\beta x_{ik}} & -\sum(e^{\alpha_k+\beta x_{ik}})x_{ik} \\ 0 & -\sum(e^{\alpha_k+\beta x_{ik}})x_{ik} & -\sum(e^{\alpha_k+\beta x_{ik}})x_{ik}^2 \end{bmatrix}$$

$$\Rightarrow I_1 = \frac{1}{\cancel{n}}\begin{bmatrix} \frac{\cancel{n}}{2\sigma_k^4} & 0 & 0 \\ 0 & \cancel{n}e^{\alpha_k+\frac{\beta\sigma_k^2}{2}} & \cancel{n}\beta\sigma_k^2 e^{\alpha_k+\frac{\beta\sigma_k^2}{2}} \\ 0 & \cancel{n}\beta\sigma_k^2 e^{\alpha_k+\frac{\beta\sigma_k^2}{2}} & \cancel{n}(\sigma_k^2+\beta^2\sigma_k^4)e^{\alpha_k+\frac{\beta^2\sigma_k^2}{2}} \end{bmatrix}$$

A, B, B', C

$$\Rightarrow I^{-1}(\beta) = (C - B'A^{-1}B)^{-1}$$

$$= \left\{(\sigma_k^2+\beta^2\sigma_k^4)e^{\alpha_k+\frac{\beta^2\sigma_k^2}{2}} - (0\ \ b_1)\begin{bmatrix}\frac{1}{a_1} & 0 \\ 0 & \frac{1}{a_2}\end{bmatrix}\begin{bmatrix}0 \\ b_2\end{bmatrix}\right\}^{-1}$$

$$= \left\{ \quad - \beta^2\sigma_k^4 e^{\alpha_k+\frac{\beta\sigma_k^2}{2}}\right\}^{-1} \quad \frac{b_1^2}{a_2}$$

$$= \left\{\sigma_k^2 e^{\alpha_k+\frac{\beta^2\sigma_k^2}{2}}\right\}^{-1} = \frac{1}{\sigma_k^2 e^{\alpha_k+\frac{\beta^2\sigma_k^2}{2}}}$$

$\Rightarrow$ By MLE theory, $\sqrt{n}(\hat{\beta}_k - \beta) \xrightarrow{d} N_1\left(0, \underbrace{\frac{1}{\sigma_k^2 e^{\alpha_k+\frac{\beta^2\sigma_k^2}{2}}}}_{V_k}\right)$ for $k=1,2$

A consistent estimator of V_k is $\hat{V}_k = \frac{1}{\hat{\sigma}_k^2 e^{\hat{\alpha}_k+\frac{\hat{\beta}^2\hat{\sigma}_k^2}{2}}}$

(ii) Want to obtain single $\hat{\beta}$ using $\hat{\beta}_1, \hat{\beta}_2$. As $\hat{\beta}_1 \perp \hat{\beta}_2$,

$$\sqrt{n}\begin{pmatrix}\hat{\beta}_1 - \beta \\ \hat{\beta}_2 - \beta\end{pmatrix} \xrightarrow{d} N_2\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{bmatrix}V_1 & 0 \\ 0 & V_2\end{bmatrix}\right)$$

∇g is gradient vector

$$\Rightarrow \sqrt{n}\left(g(\hat{\beta}_1,\hat{\beta}_2) - g(\beta,\beta)\right) \xrightarrow{d} N_1\left(0, \ 2\ \nabla g(\beta,\beta)\begin{bmatrix}V_1 & 0 \\ 0 & V_2\end{bmatrix}\nabla g'(\beta,\beta)\right)$$

Want $g_{opt}(x,y)$ so that $g_{opt}(x,x) = x$, and $g_{opt}(x,y)$ minimizes

$$\nabla g(\beta,\beta)\begin{bmatrix}V_1 & 0 \\ 0 & V_2\end{bmatrix}\nabla g'(\beta,\beta)$$

$g^*_{opt}(x,y) = \dfrac{V_2 x + V_1 y}{V_1 + V_2}$ $\Rightarrow$ $\nabla g(x,y) = \dfrac{1}{V_1+V_2}(V_2 \;,\; V_1)$

$\Rightarrow g^*_{opt}(\beta,\beta) = \beta$ ✓ $\Rightarrow$ $\nabla g(\beta,\beta) = \left(\dfrac{1}{V_1+V_2}\right)^2 [V_2 \;\; V_1] \begin{bmatrix} V_1 & 0 \\ 0 & V_2 \end{bmatrix} \begin{bmatrix} V_2 \\ V_1 \end{bmatrix}$

$D = \nabla g(\beta,\beta) \begin{bmatrix} V_1 & 0 \\ 0 & V_2 \end{bmatrix} \nabla g'$ $= \left(\dfrac{1}{V_1+V_2}\right)^2 (V_2^2 V_1 + V_1^2 V_2) = \dfrac{V_1 V_2}{V_1+V_2}$

$= \left[\left(\dfrac{\partial g}{\partial \beta}\right)^2 V_1 + \left(\dfrac{\partial g}{\partial \beta}\right)^2 V_2\right]$

(iv) $\sqrt{2n}\left(\hat{g}_{opt}(\hat\beta_1,\hat\beta_2) - \beta\right) \xrightarrow{d} N_1\left(0,\; 2\cdot\dfrac{V_1V_2}{(V_1+V_2)}\right)$, $V_k = \dfrac{1}{\sigma_2^k e^{\alpha_k + \frac{\beta^2\sigma_k^2}{2}}}$

In (a), $\sqrt{2n}(\hat\beta - \beta) \xrightarrow{d} N_1\left(0,\; 2\cdot\left[\dfrac{1}{\frac{1}{V_1}+\frac{1}{V_2}}\right]\right)$

$\Rightarrow \underline{A.R.E}\left(\hat g_{opt}(\hat\beta_1,\hat\beta_2), \hat\beta\right) = \dfrac{\cancel{2}\left[\dfrac{V_1V_2}{V_1+V_2}\right]}{\cancel{2}\left[\dfrac{1}{\frac{1}{V_1}+\frac{1}{V_2}}\right] = \left[\dfrac{V_1V_2}{V_1+V_2}\right]} = 1$

(c) Assume $\alpha_1 = \alpha_2 = \alpha$

Then, $\ell(\sigma_1^2,\sigma_2^2,\alpha,\beta) \propto -\frac{n}{2}\log\sigma_1^2 - \frac{n}{2}\log\sigma_2^2 - \sum_{i=1}^n e^{\alpha+\beta x_{i1}} - \sum_{i=1}^n e^{\alpha+\beta x_{i2}} + \alpha\sum_{i=1}^n (y_{i1}+y_{i2}) + \beta\sum_{i=1}^n (y_{i1}x_{i1} + y_{i2}x_{i2}) - \frac{1}{2\sigma_1^2}\sum x_{i1}^2 - \frac{1}{2\sigma_2^2}\sum x_{i2}^2$

$\dot\ell(\sigma_1^2) = \frac{-n}{2\sigma_1^2} + \frac{1}{2\sigma_1^4}\sum x_{i1}^2$, $\dot\ell(\sigma_2^2) = \frac{-n}{2\sigma_2^2} + \frac{1}{2\sigma_2^4}\sum x_{i2}^2$

$\dot\ell(\alpha) = -\sum_{i=1}^n\left(e^{\alpha+\beta x_{i1}} + e^{\alpha+\beta x_{i2}}\right)$, $\dot\ell(\beta) = -\sum_{i=1}^n\left[(e^{\alpha+\beta x_{i1}})x_{i1} + (e^{\alpha+\beta x_{i2}})x_{i2}\right] + \cdots$

$$\Rightarrow I_{2n} = -E\begin{bmatrix} \frac{n}{2\sigma_1^4} - \frac{1}{\sigma_1^6}\sum x_{i1}^2 & 0 & 0 & 0 \\ 0 & \frac{n}{2\sigma_2^4} - \frac{1}{\sigma_2^6}\sum x_{i2}^2 & 0 & 0 \\ 0 & 0 & -\sum_{i=1}^n(e^{\alpha+\beta x_{i1}} + e^{\alpha+\beta x_{i2}}) & -\sum_{i=1}^n\left[(e^{\alpha+\beta x_{i1}})x_{i1} + (e^{\alpha+\beta x_{i2}})x_{i2}\right] \\ 0 & 0 & -\sum_{i=1}^n\left[(e^{\alpha+\beta x_{i1}})x_{i1} + (e^{\alpha+\beta x_{i2}})x_{i2}\right] & -\sum_{i=1}^n\left[(e^{\alpha+\beta x_{i1}})x_{i1}^2 + (e^{\alpha+\beta x_{i2}})x_{i2}^2\right] \end{bmatrix}$$

A B

$$\Rightarrow I_1 = \frac{1}{2n}\begin{bmatrix} \frac{n}{2\sigma_1^4} & 0 & 0 & 0 \\ 0 & \frac{n}{2\sigma_2^4} & 0 & 0 \\ 0 & 0 & n\left[e^{\alpha_1+\frac{\beta^2\sigma_1^2}{2}} + e^{\alpha_2+\frac{\beta^2\sigma_2^2}{2}}\right] & n\left[\beta\sigma_1^2 e^{\alpha_1+\frac{\beta^2\sigma_1^2}{2}} + \beta\sigma_2^2 e^{\alpha_2+\frac{\beta^2\sigma_2^2}{2}}\right] \\ 0 & 0 & n\left[\beta\sigma_1^2 e^{\alpha_1+\frac{\beta^2\sigma_1^2}{2}} + \beta\sigma_2^2 e^{\alpha_2+\frac{\beta^2\sigma_2^2}{2}}\right] & n\left[(\sigma_1^2+\beta^2\sigma_1^4)e^{\alpha_1+\frac{\beta^2\sigma_1^2}{2}} + (\sigma_2^2+\beta^2\sigma_2^4)e^{\alpha_2+\frac{\beta^2\sigma_2^2}{2}}\right] \end{bmatrix}$$

A^* B C

$$\Rightarrow I_1^{-1}(\beta) = \left\{\frac{1}{2}\left[(\sigma_1^2+\beta\sigma_1^4)e^{\alpha_1+\frac{\beta^2\sigma_1^2}{2}} + (\sigma_2^2+\beta\sigma_2^4)e^{\alpha_2+\frac{\beta^2\sigma_2^2}{2}}\right] - \frac{1}{2}\frac{\left(\beta\sigma_1^2 e^{\alpha_1+\frac{\beta^2\sigma_1^2}{2}} + \beta\sigma_2^2 e^{\alpha_2+\frac{\beta^2\sigma_2^2}{2}}\right)^2}{e^{\alpha_1+\frac{\beta^2\sigma_1^2}{2}} + e^{\alpha_2+\frac{\beta^2\sigma_2^2}{2}}}\right\}^{-1}$$

By MLE theory, $\sqrt{2n}(\hat\beta_c - \beta) \xrightarrow{d} N_1(0, I^{-1}(\beta))$

given above

... $I^{-1}(\beta)$ to equal $I^{-1}(\beta)$ in (a), $\dfrac{(\beta\sigma_1^2 e^{\alpha_1+\frac{\beta^2\sigma_1^2}{2}} + \beta\sigma_2^2 e^{\alpha_2+\frac{\beta^2\sigma_2^2}{2}})^2}{e^{\alpha_1+\frac{\beta^2\sigma_1^2}{2}} + e^{\alpha_2+\frac{\beta^2\sigma_2^2}{2}}} = \beta^2\sigma_1^4 e^{\alpha_1+\frac{\beta^2\sigma_1^2}{2}} + \beta^2\sigma_2^4 e^{\alpha_2+\frac{\beta^2\sigma_2^2}{2}}$

... asy var$(\hat\beta_c)$ = asy var$(\hat\beta)$

PROBLEM 1

(3) (a) $(X_1, Y_1), \ldots, (X_n, Y_n)$ are iid $N_2(\mu, \Sigma)$, where $\mu = (\mu_1, \mu_2)'$ and $\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}$.

We assume $\sigma_{12} = 0$.
All other parameters are unknown.

$\Sigma y_i^2 \Delta / \sigma_1^2$

The likelihood is proportional to

$$\exp\left\{ -\frac{1}{2}\left(\frac{\sum_1^n x_i^2}{\sigma_1^2} - \frac{2\mu_1}{\sigma_1^2}\sum_1^n x_i + \sum_1^n y_i^2/\sigma_2^2 - \frac{2\mu_2}{\sigma_2^2}\sum_1^n y_i \right)\right\}$$

$$= \exp\left\{ \sum_1^4 \eta_i T_i \right\}$$

where
$\eta_1 = \frac{1}{2\sigma_2^2} - \frac{1}{2\sigma_1^2}$ $\qquad T_1 = \sum_1^n x_i^2$

$\eta_2 = -\frac{1}{2\sigma_2^2}$ $\qquad T_2 = \sum_1^n x_i^2 + \sum_1^n y_i^2$

$\eta_3 = \frac{\mu_1}{\sigma_1^2}$ $\qquad T_3 = \sum_1^n x_i$

$\eta_4 = \frac{\mu_2}{\sigma_2^2}$ $\qquad T_4 = \sum_1^n y_i$

This being the natural parameterization of an exponential family, T_1, T_2, T_3 & T_4 are sufficient for $\eta_1, \eta_2, \eta_3, \eta_4$.

Now, note that WLOG we can take $\Delta_0 = 1$ since in other cases we can scale the Y values by $\sqrt{\Delta_0}$ to reduce the ratio of variances to 1.

Thus, our null is $H_0: \frac{\sigma_2^2}{\sigma_1^2} = 1$ vs $H_1: \frac{\sigma_2^2}{\sigma_1^2} = 1$

Under the natural parameterization,

$$H_0: \eta_1 = 0 \quad \text{vs} \quad H_1: \eta_1 \neq 0.$$

η_2, η_3 & η_4 are nuisance parameters.

Hence, by theorem 2.7, the UMPU test will be

$$\phi(x,y) = \begin{cases} 1 & \text{if } T_1 < c_1(T_2,T_3,T_4) \text{ or } T_1 > c_2(T_2,T_3,T_4) \\ \gamma_i & \text{if } T_1 = c_i(T_2,T_3,T_4)\ ,\ i=1,2. \\ 0 & \text{otherwise} \end{cases}$$

[Since T_1 is continuous, we can take $\gamma_i = 0$, $i=1,2$]

where γ_i & c_i are such that

$$E_{\eta_1=0}\left[\phi \mid T_2,T_3,T_4\right] = \alpha$$

$$\& \ E_{\eta_1=0}\left[T_1\phi \mid T_2,T_3,T_4\right] = \alpha E\left(T_1 \mid T_2,T_3,T_4\right).$$

Thus, taking $\gamma_i = 0$ for $i=1,2$, we reject when

$$T_1 < c_1(T_2,T_3,T_4)$$
$$\text{or } \ T_1 > c_2(T_2,T_3,T_4)$$

Now, $T_1 < c_1(T_2,T_3,T_4)$

is equivalent to $T = \dfrac{\sum_1^n (x_i-\bar{x})^2}{\sum_1^n (x_i-\bar{x})^2 + \sum_1^n (y_i-\bar{y})^2} < c_1^*(T_2,T_3,T$ for some $c_1^*($ for given T_2,T_3,T_4

This is because

$$T = \frac{T_1 - \frac{1}{n}T_3^2}{T_2 - \frac{1}{n}T_3^2 - \frac{1}{n}T_4^2}$$ is increasing in T_1 for every fixe T_2, T_3 & T_4.

Now, $T = \dfrac{\sum_1^n (x_i-\bar{x})^2}{\sum_1^n (x_i-\bar{x})^2+\sum_1^n (y_i-\bar{y})^2} < c_1^*(T_2, T_3, T_4)$

is equivalent to

$$F = \frac{\sum_1^n (x_i-\bar{x})^2/(n-1)}{\sum_1^n (y_i-\bar{y})^2/(n-1)} < c_1^{**}(T_2, T_3, T_4)$$

for some $c^{**}(\cdot)$.

Similarly $T_1 > c_2(T_2, T_3, T_4)$ is equivalent to $F > c_2^{**}(T_2, T_3, T_4)$.
Note that $F \sim F_{n-1, n-1}$ dis$^{\overline{n}}$ under H_0 hence it is ancilliary and hence indept of the sufficient statistics T_2, T_3, T_4.

Thus, the conditional test becomes unconditional as below.

$$\phi(F) = \begin{cases} 1 & \text{if } F < c_1 \text{ or } F > c_2 \\ 0 & \text{otherwise} \end{cases}$$

where c_1 & c_2 are such that

$$E_{\eta_1=0}(\phi(F)) = \alpha \quad \cdots \langle 1 \rangle$$

$$\& \quad E_{\eta_1=0}(F\phi(F)) = \alpha E_{\eta_1=0}(F) = \frac{\alpha(n-1)}{(n-3)} \quad \cdots \langle 2 \rangle$$

Alternatively, we can write the test as

$$\phi(F) = \begin{cases} 1 & \text{if } F < F_{1-\alpha_1; n-1, n-1}, \text{ or } F > F_{\alpha_2; n-1, n-1} \\ 0 & \text{otherwise} \end{cases}$$

where α_1 & α_2 are such that

$$\alpha_1 + \alpha_2 = \alpha$$

$$\& \int_{F_{1-\alpha_1; n-1, n-1}}^{F_{\alpha_2; n-1, n-1}} x f(x)\, dx = 1 - \frac{\alpha(n-1)}{(n-3)}$$

[$f(x)$ is pdf of $F_{n-1, n-1}$ dis$^{\overline{n}}$ & $F_{\beta; n-1, n-1}$ is upper β point of $F_{n-1, n-1}$ dis$^{\overline{n}}$]

Now, note that under $H_0 : \eta_1 = 0$,

$$F \sim F_{n-1, n-1}$$

Also, $1/F \sim F_{n-1, n-1}$.

Since F & $1/F$ have the same distn, thus if we write the above test in terms of $1/F$ as

$$\phi'(\tfrac{1}{F}) = \begin{cases} 1 & \text{if } \frac{1}{F} < c_1' \text{ or } \frac{1}{F} > c_2' \\ 0 & \text{otherwise} \end{cases}$$

then the corresponding eqns of (1) & (2) will yield the same soln for (c_1', c_2') as ~~they~~ (1) & (2) yield for (c_1, c_2), ie

$$c_1' = c_1 \,,\; c_2' = c_2$$

But since $\phi(F)$ & $\phi'(1/F)$ are equivalent, we must have $c_2 = 1/c_1'$ & $c_1 = 1/c_2'$.

Hence $c_1 c_2 = 1$.

Then, $P_{H_0}(F > c_2) + P_{H_0}(F < c_1) = \alpha$

$$\Rightarrow \quad P_{H_0}(F > c_2) + P_{H_0}(1/F > 1/c_1) = \alpha$$

$$\Rightarrow \quad P(F_{n-1,n-1} > c_2) + P(F_{n-1,n-1} > c_2) = \alpha \quad [\because c_2 = 1/c_1]$$

$$\Rightarrow \quad P(F_{n-1,n-1} > c_2) = \alpha/2 \;\Rightarrow\; c_2 = F_{\alpha/2; n-1, n-1}$$

Hence the UMPU test in simplest form is

$$\phi(F) = \begin{cases} 1 & \text{if } F > F_{\alpha/2; n-1, n-1} \text{ or } F < \frac{1}{F_{\alpha/2; n-1, n-1}} = F_{1-\alpha/2; n-1, n-1} \\ 0 & \text{otherwise} \end{cases}$$

(3) (b) When $\sigma_{12}=0$, the log-likelihood is

$$l_n(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2)$$

$$= \text{constant} - \frac{n}{2}\log(\sigma_1^2) - \frac{n}{2}\log(\sigma_2^2) - \frac{1}{2\sigma_1^2}\sum_1^n (x_i-\mu_1)^2 - \frac{1}{2\sigma_2^2}\sum_1^n (y_i-\mu_2)^2$$

Under $H_0: \sigma_1^2=\sigma_2^2$ (Again we choose WLOG $\Delta_0=1$), the mle of (μ_1,μ_2,σ^2), where σ^2 is the common value of $\sigma_1^2=\sigma_2^2$, would be given by

$$\tilde{\mu}_1=\bar{x}$$
$$\tilde{\mu}_2=\bar{y}$$
$$\tilde{\sigma}^2=\frac{1}{2n}\left[\sum_1^n (x_i-\bar{x})^2+\sum_1^n (y_i-\bar{y})^2\right]$$

In the unrestricted case, mle of $(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2)$ is given by

$$\hat{\mu}_1=\bar{x} \qquad \hat{\sigma}_1^2=\frac{1}{n}\sum_1^n (x_i-\bar{x})^2$$
$$\hat{\mu}_2=\bar{y} \qquad \hat{\sigma}_2^2=\frac{1}{n}\sum_1^n (y_i-\bar{y})^2$$

Thus, the LRT statistic will be

$$\Delta = \exp\{ l_n(\tilde{\mu}_1,\tilde{\mu}_2,\tilde{\sigma}^2) - l_n(\hat{\mu}_1,\hat{\mu}_2,\hat{\sigma}_1^2,\hat{\sigma}_2^2)\}$$

$$= \left\{ \frac{\hat{\sigma}_1^2\hat{\sigma}_2^2}{(\tilde{\sigma}^2)^2} \right\}^{n/2}$$

$$= \left[\frac{\left\{\sum_1^n (x_i-\bar{x})^2\right\}\left\{\sum_1^n (y_i-\bar{y})^2\right\}}{\left(\frac{1}{2}\left\{\sum_1^n (x_i-\bar{x})^2+\sum_1^n (y_i-\bar{y})^2\right\}\right)^2} \right]^{n/2}$$

Thus, the LRT would be given by reject H_0 iff

$$\Lambda < c$$

or, $$\frac{F}{(1+F)^2} < K = c^{2/n} \quad \left[\text{where } F = \frac{\sum_1^n (x_i-\bar{x})^2}{\sum_1^n (y_i-\bar{y})^2}\right]$$

or, $$\left(\sqrt{F} + \frac{1}{\sqrt{F}}\right)^2 > \frac{1}{K}$$

or, $$F > K_1 \text{ or } F < \frac{1}{K_1} \quad \left[\text{where } \frac{K_1}{(1+K_1)^2} = K\right.$$

Now,
$$P_{H_0}(F > K_1) + P_{H_0}\left(F < \frac{1}{K_1}\right) = \alpha$$
$$\Rightarrow \quad P_{H_0}(F > K_1) + P_{H_0}\left(\frac{1}{F} > K_1\right) = \alpha$$

Now, under H_0, $F \sim F_{n-1,n-1}$ distn & $\frac{1}{F} \sim F_{n-1,n-1}$ di

Thus,
$$2 \cdot P(F_{n-1,n-1} > K) = \alpha$$
$$\Rightarrow \quad K_1 = F_{\alpha/2; n-1, n-1}$$

Hence LRT is given by

$$\phi(F) = \begin{cases} 1 & \text{if } F > F_{\alpha/2; n-1, n-1} \text{ or } F < \dfrac{1}{F_{\alpha/2; n-1, n-1}} = F_{1-\alpha/2; n-1,} \\ 0 & \text{otherwise} \end{cases}$$

Note that the LRT has the same form as the UMPU test.

Note that under H_1, the likelihood is maximized at

$$\left[\frac{1}{2\pi\hat{\sigma}_{x.1}\hat{\sigma}_{y.1}\sqrt{1-R^2}}\right]^n \exp\left[-\frac{1}{2(1-R^2)}\left[\frac{1}{\hat{\sigma}_{x.1}^2}\sum_{i=1}^{n}(x_i-\bar{x})^2+\frac{1}{\hat{\sigma}_{x.1}^2}\sum_{i=1}^{n}(x_i-\bar{x})^2-\frac{2R}{\hat{\sigma}_{x.1}\hat{\sigma}_{x.1}}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})\right]\right]$$

$$\left[\frac{1}{2\pi\hat{\sigma}_{x.1}\hat{\sigma}_{y.1}\sqrt{1-R^2}}\right]^n \exp\left[-\frac{1}{2(1-R^2)}\left[n+n-2n\right]\right]$$

$$\left[\frac{1}{2\pi\hat{\sigma}_{x.1}\hat{\sigma}_{y.1}\sqrt{1-R^2}}\right]^n$$

Hence the likelihood ratio will be

$$\Lambda=\left[\frac{2\pi\hat{\sigma}_{x.1}\hat{\sigma}_{y.1}\sqrt{1-R^2}}{2\pi\hat{\sigma}_{x.1}\hat{\sigma}_{y.1}}\right]^n=\left(\sqrt{1-R^2}\right)^n$$

Hence the test will accept H_0 when

$$c_1<\left(\sqrt{1-R^2}\right)^n<c_2$$

or, equivalently,

$$c_1^*<R<c_2^*$$

We note that under the null, R is distributed symmetrically about 0, and hence the test rejects when

$$|R|>c$$

ii)
We want to study the distribution of

$$R=\frac{\sum_{i=1}^{n}(x_i)(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2\sum_{i=1}^{n}(y_i-\bar{y})^2}}$$

First, we note that

$$\frac{\sqrt{n-2}R}{\sqrt{1-R^2}}=\frac{s_{xy}}{s_ys_x}\frac{s_ys_x}{\sqrt{s_y^2s_x^2-s_{xy}^2}}=\frac{s_{xy}}{\sqrt{s_y^2s_x^2-s_{xy}^2}}=\frac{s_{xy}}{\sqrt{\sigma_1\sigma_2}}\sqrt{\frac{\sigma_1\sigma_2}{s_y^2s_x^2-s_{xy}^2}}$$

We note that, conditional on $x_1,\dots,x_n$, we have

$$\frac{s_{xy}}{\sqrt{\sigma_1\sigma_2}}=\frac{\sum_{i=1}^{n}x_iy_i-n\bar{x}\bar{y}}{\sqrt{\sigma_1\sigma_2}}$$

where

$$\sum_{i=1}^{n}x_iy_i\sim N\left(n\bar{x}\mu_y,\sigma_y^2\sum_{i=1}^{n}x^2\right)$$

$$n\bar{x}\bar{y}\sim N\left(n\bar{x}\mu_y,\sigma_y^2\bar{x}^2\right)$$

$$s_{xy}\sum_{i=1}^{n}x_iy_i-n\bar{x}\bar{y}\sim N(0,\sigma_y^2s_x^2)$$

Similarly,

$$\frac{s_y^2 s_x^2 - s_{xy}^2}{\sigma_1\sigma_2}$$

$$s_y^2 s_x^2 = s_x^2\chi^2(n-1)$$
$$s_{xy}^2 = \chi^2(n-1)$$

Hence,

$$\frac{\sqrt{n-2}R}{\sqrt{1-R^2}} \sim T(n-2)$$

Furthermore, this distribution doesn't change if we do not condition on x. Hence we have

$$F_R(k) = P(R < k) = P\left(\frac{\sqrt{n-2}R}{\sqrt{1-R^2}} < k\sqrt{\frac{n-2}{1-k^2}}\right) = F_{t(n-2)}\left(k\sqrt{\frac{n-2}{1-k^2}}\right)$$

Taking the derivative with respect to k, we obtain

$$\begin{aligned}
f_R(k) &= f_{t(n-2)}\left(k\sqrt{\frac{n-2}{1-k^2}}\right)\frac{\sqrt{n-2}}{(1-k^2)^{3/2}} \\
&= \frac{\Gamma((n-1)/2)}{\sqrt{\pi}\Gamma((n-2)/2)}\left(1+k^2\frac{n-2}{1-k^2}/(n-2)\right)^{-(n-1)/2}\frac{\sqrt{n-2}}{(1-k^2)^{3/2}} \\
&= \frac{\Gamma((n-1)/2)}{\sqrt{\pi}\Gamma((n-2)/2)}\left(\frac{1}{1-k^2}\right)^{-(n-1)/2}\left[\frac{1}{1-k^2}\right]^{3/2} \\
&= \frac{\Gamma((n-1)/2)}{\sqrt{\pi}\Gamma((n-2)/2)}(1-k^2)^{(n-1)/2}\left[1-k^2\right]^{-3/2} \\
&= \frac{1}{\sqrt{\pi}}\frac{\Gamma((n-1)/2)}{\Gamma((n-2)/2)}(1-k^2)^{(n-4)/2}
\end{aligned}$$

Finding constant c so that

$$F_{t(n-2)}(c) = \alpha/2$$
$$c = F_{t(n-2)}^{-1}(\alpha/2)$$

Which implies that

$$F_R\left(F_{t(n-2)}^{-1}(\alpha/2)\sqrt{\frac{n-2}{1-F_{t(n-2)}^{-2}(\alpha/2)}}\right) = \alpha/2$$

We note that the quantile function of the t-distribution is hard to track in closed form. Alternatively, we can integrate such that

$$\alpha/2 = \int_{-1}^{c}\frac{1}{\sqrt{\pi}}\frac{\Gamma((n-1)/2)}{\Gamma((n-2)/2)}(1-r^2)^{(n-4)/2}dr$$
$$\frac{\alpha\sqrt{\pi}}{2}\frac{\Gamma((n-2)/2)}{\Gamma((n-1)/2)} = \int_{-1}^{c}(1-r^2)^{(n-4)/2}dr$$

iii)

We know that $R\sqrt{(n-2)/(1-R^2)} \sim T_{n-2}$. Furthermore, the t-distribution converges to the normal distribution. Hence, using the delta method with $g(x) = x(x^2+1)^{-1/2}$, $g'(x) = (x^2+1)^{-3/2}$.

$$\sqrt{n-2}\left[R\sqrt{\frac{1}{1-R^2}} - 0\right] \to_d N(0,1)$$
$$\sqrt{n}\left[g\left(R\sqrt{\frac{1}{1-R^2}}\right) - g(0)\right] \to_d N(0, g'(0)^2)$$
$$\sqrt{n}R \to_d N(0,1)$$

2010 THEORY I QUESTION 2

$X_1, \ldots, X_n$ iid $U(\theta, \theta+1)$ θ unknown

(a) $L(\theta|\underset{\sim}{X}) = 1^n \cdot I\{X_{(1)} > \theta,\ X_{(n)} < \theta+1\}$

$= I\{\theta > X_{(n)} - 1,\ \theta < X_{(1)}\}$

$\Rightarrow \hat\theta \in (X_{(n)}-1, X_{(1)})$ are MLEs of θ

(b) Under absolute error loss, the Bayes estimator of θ is the posterior median

$$f(\theta|\underset{\sim}{X}) = \frac{f(\underset{\sim}{X}|\theta)f(\theta)}{\int_\theta f(\underset{\sim}{X}|\theta)f(\theta)\,d\theta} = \frac{I\{\theta > X_{(n)}-1,\ \theta < X_{(1)}\}(2\pi\sigma_0^2)^{-\frac{1}{2}}\exp\left\{\frac{-1}{2\sigma_0^2}(\theta-\mu_0)^2\right\}}{\int_{X_{(n)}-1}^{X_{(1)}}(2\pi\sigma_0^2)^{-\frac{1}{2}}\exp\left\{\frac{-1}{2\sigma_0^2}(\theta-\mu_0)^2\right\}}$$

$$\Rightarrow \frac{1}{2} = P(\theta|\underset{\sim}{X} < \hat\theta_B) \Rightarrow \frac{1}{2} = \int_{X_{(n)}-1}^{\hat\theta_B} f(\theta|\underset{\sim}{X}) \Rightarrow \frac{1}{2} = \frac{\Phi\left(\frac{\hat\theta_B-\mu_0}{\sigma_0}\right) - \Phi\left(\frac{X_{(n)}-1-\mu_0}{\sigma_0}\right)}{\Phi\left(\frac{X_{(1)}-\mu_0}{\sigma_0}\right) - \Phi\left(\frac{X_{(n)}-1-\mu_0}{\sigma_0}\right)}$$

$$\Rightarrow \Phi\left(\frac{\hat\theta_B-\mu_0}{\sigma_0}\right) = \frac{1}{2}\left[\Phi\left(\frac{X_{(1)}-\mu_0}{\sigma_0}\right) + \Phi\left(\frac{X_{(n)}-1-\mu_0}{\sigma_0}\right)\right]$$

$$\Rightarrow \hat\theta_B = \mu_0 + \sigma_0\,\Phi^{-1}\left\{\frac{1}{2}\left[\Phi\left(\frac{X_{(1)}-\mu_0}{\sigma_0}\right) + \Phi\left(\frac{X_{(n)}-1-\mu_0}{\sigma_0}\right)\right]\right\}$$

(c) $d(X) = aX_{(1)} + bX_{(n)} + c$

Let $Z_{(1)} = X_{(1)} - \theta$, $Z_{(n)} = X_{(n)} - \theta$

$\Rightarrow f_{Z_{(1)},Z_{(n)}}(s,t) = n(n-1)(t-s)^{n-2},\quad 0<s<t<1$

so that

$E[Z_{(1)}] = \frac{1}{n+1}$, $E[Z_{(1)}^2] = \frac{2}{(n+1)(n+2)}$ $\Rightarrow Var(Z_{(1)}) = \frac{n}{(n+2)(n+1)^2}$

$E[Z_{(n)}] = \frac{n}{n+1}$, $E[Z_{(n)}^2] = \frac{n}{n+2}$ $\Rightarrow Var(Z_{(n)}) = \frac{n}{(n+2)(n+1)^2}$

$E[Z_{(1)}Z_{(n)}] = \frac{1}{n+2}$ $\Rightarrow cov[Z_{(1)}, Z_{(n)}] = \frac{1}{(n+2)(n+1)^2}$

so that

$E[X_{(1)}] = \frac{1}{n+1} + \theta$, $E[X_{(n)}] = \frac{n}{n+1} + \theta$ $\Rightarrow Bias[d(X)] = E[d(X)] - \theta$

$= a\left[\frac{1}{n+1}+\theta\right] + b\left[\frac{n}{n+1}+\theta\right] + c - \theta$

$= (a+b-1)\theta + \frac{a+bn}{n+1} + c$

$Var[d(X)] = a^2 Var[X_{(1)}] + b^2 Var[X_{(n)}] + 2ab\,Cov[X_{(1)}, X_{(n)}]$

$= \frac{(a^2+b^2)n}{(n+2)(n+1)^2} + \frac{2ab}{(n+2)(n+1)^2}$

Frequentist risk of $d(X)$ is $Var[d(X)] + Bias^2[d(X)]$

$= \frac{1}{(n+2)(n+1)^2}\left[(a^2+b^2)n + 2ab\right] + \left[(a+b-1)\theta + \frac{a+bn}{n+1} + c\right]^2$

minimax frequentist risk:

frequentist risk constant if $a+b=1$. Find "best" choice of a, b, c subject to this constraint. constrained minimization ⇒ Lagrange.

minimize (for a, b, c)

$$f(a,b,c) = \frac{1}{(n+1)^2(n+2)}\left[(a^2+b^2)n+2ab\right] + (a+b-1)^2\theta^2 + \frac{(a+bn)^2}{(n+1)^2} + c^2 + 2c(a+b-1)\theta + \frac{2c(a+bn)}{n+1}$$

$$\quad + \frac{2(a+b-1)\theta(a+bn)}{n+1} + \lambda(a+b-1)$$

$$\frac{\partial f}{\partial a} = \frac{2an+2b}{(n+1)^2(n+2)} + 2(a+b-1)\theta^2 + \frac{2(a+bn)}{(n+1)^2} + 2c\theta + \frac{2c}{n+1} + \frac{2\theta}{(n+1)}\left(2a+bn+b-1\right) + \lambda = 0$$

$$\frac{\partial f}{\partial b} = \frac{2bn+2a}{(n+1)^2(n+2)} + 2(a+b-1)\theta^2 + \frac{2n(a+bn)}{(n+1)^2} + 2c\theta + \frac{2cn}{n+1} + \frac{2\theta}{n+1}\left(an+a+2bn-n\right) + \lambda = 0$$

$$\frac{\partial f}{\partial c} = 2c + 2\theta(a+b-1) + \frac{2(a+bn)}{n+1} = 0 \Rightarrow 2c = -2\theta(a+b-1) - \frac{2(a+bn)}{n+1}$$

$[an+a+2bn-n]$

solving for $\lambda \Rightarrow \frac{2an+2b}{(n+1)^2(n+2)} + \frac{2(a+bn)}{(n+1)^2} + \frac{2c}{n+1} + \frac{2\theta}{(n+1)}\left(2a+bn+b-1\right) = \frac{2bn+2a}{(n+1)^2(n+2)} + \frac{2n(a+bn)}{(n+1)^2} + \frac{2cn}{n+1} + \frac{2\theta}{n+1}[\dots]$

$$\Rightarrow \frac{1}{(n+1)^2(n+2)}\left[2an+2b-2bn-2a\right] + \frac{1}{(n+1)^2}\left[2a+2bn-2an-2bn^2\right] + \frac{1}{(n+1)}\left(2c-2cn\right) + \frac{2\theta}{(n+1)}\left(2a+bn+b-1 \; -an-a-2bn+n\right) = 0$$

$$\Rightarrow \frac{2}{(n+1)^2(n+2)}\left[an-bn+b-a+(a+bn-an-bn^2)(n+2)\right] + \frac{2}{(n+1)}\left[-\theta(a+b-1)-\frac{(a+bn)}{n+1}\right](1-n) + 2\theta\left[\begin{matrix}a+b-1+n\\ -an-bn\end{matrix}\right] = 0$$

$$\Rightarrow n(a-b)+b-a+\left(a+n(b-a)-bn^2\right)(n+2) - (n^2+1)(n+2)\left[\theta(a+b-1)-\frac{a+bn}{n+1}\right] + \theta(n+1)^2(n+2)\left(\begin{matrix}a+b-1+n\\ -n(a+b)\end{matrix}\right) = 0$$

$$\Rightarrow (n-1)(a-b) + (n+2)\left[a+n(b-a)-bn^2\right] + \theta(n+1)^2(n+2)\left[(a+b)(1-n)-1+n\right] = (n^2+1)(n+2)\left[\theta(a+b+1)-\frac{(a+bn)}{n+1}\right]$$

optimal choice is $a = b = \frac{1}{2}$, $c = \frac{-1}{2}$. Call this rule $d^*(X)$.

$d^*(X)$ is admissible within these class of estimators because for any biased rule $R(d(X),\theta) \to \infty$ as $\theta \to \infty$, and $d^*(X)$ has the minimum frequentist risk among all unbiased rules within this class.

(d) $d^*(X)$ from (c) is minimax as it has constant frequentist risk, and it is Bayes with respect to some prior (because it is admissible, as shown in (c)).

(e) As $f_{X_{(1)},X_{(n)}}(s,t) = n(n-1)(t-s)^{n-2}$, $0<s<t<1$

and $R_n = X_{(n)} - X_{(1)}$, [illegible]

[illegible]

$$f_R(r) = n(n-1)\, r^{n-2} \int_0^{1-r} j\, dj$$
$$= n(n-1)\, r^{n-2} (1-r), \quad \text{or} \sim Beta(n-1, 2)$$

Consider $n(1-R)$.

$$P[n(1-R) < t] = P[1-R < \tfrac{t}{n}]$$
$$= P[R > 1-\tfrac{t}{n}]$$
$$= 1 - P[R < 1-\tfrac{t}{n}]$$
$$= 1 - \int_0^{1-\frac{t}{n}} n(n-1)\, r^{n-2} (1-r)$$
$$= 1 - n(n-1) \left\{ \int_0^{1-\frac{t}{n}} r^{n-2}\, dr - \int_0^{1-\frac{t}{n}} r^{n-1}\, dr \right\}$$
$$= 1 - n(n-1) \left\{ \tfrac{1}{n-1} [1-\tfrac{t}{n}]^{n-1} - \tfrac{1}{n} [1-\tfrac{t}{n}]^{n} \right\}$$
$$= 1 - n(1-\tfrac{t}{n})^{n-1} + (n-1)(1-\tfrac{t}{n})^{n}$$
$$= 1 - (1-\tfrac{t}{n})^{n-1} \left[n - (n-1)(1-\tfrac{t}{n}) \right]$$
$$= 1 - (1-\tfrac{t}{n})^{n-1} \left[n - (n - t - 1 + \tfrac{t}{n}) \right]$$
$$= 1 - (1-\tfrac{t}{n})^{n-1} \left[t + 1 - \tfrac{t}{n} \right]$$
$$= 1 - t(1-\tfrac{t}{n})^{n-1} - (1-\tfrac{t}{n})^{n}$$
$$\longrightarrow \quad 1 - te^{-t} - e^{-t}$$

Take derivatives to get density

$$= f_{n(1-R)}(t) \longrightarrow \tfrac{\partial}{\partial t}\left[1 - te^{-t} - e^{-t}\right] + c$$
$$\cancel{-e^{-t}} + te^{-t} + \cancel{e^{-t}} + c \quad = te^{-t}, \text{ a Gamma}(2,1) \text{ density}$$

2010 THEORY 1 QUESTION 3

$X_1, \dots, X_n$ iid $E(X_1) = 0$, $0 < Var(X_1) = \sigma^2 < \infty$.

(a) show for X close to 0, $e^X - 1 - X = \frac{X^2}{2} + o(X^2)$

let $f(x) = e^x - 1 - X$. taylor expansion of $f(x)$ near 0 leads to

$$f(x) = \underbrace{f(0)}_{=0} + \underbrace{f'(0)}_{=0}(X-0) + \frac{\overbrace{f''(0)}^{=1}(X-0)^2}{2} + o(X^2)$$

$$= \frac{X^2}{2} + o(X^2)$$

(b) From WLLN, $\bar{X}_n \xrightarrow{p} 0$. Use continuous map with $f(x) = e^x - 1 - X$

to conclude $f(\bar{X}_n) = e^{\bar{X}_n} - 1 - \bar{X}_n \xrightarrow{p} f(0) = 0$.

consider $\frac{e^{\bar{X}_n} - \bar{X}_n - 1}{\bar{X}_n} = \frac{1}{\bar{X}_n}\left(\frac{\bar{X}^2}{2} + o(\bar{X}^2)\right)$ by taylor expansion

$$= \frac{\bar{X}}{2} + \frac{o(\bar{X}^2)}{\bar{X}} \xrightarrow{p} 0 \text{ as } \bar{X} \xrightarrow{p} 0.$$

consider $\frac{e^{\bar{X}_n} - \bar{X}_n - 1}{\bar{X}_n^2} = \frac{1}{\bar{X}_n^2}\left(\frac{\bar{X}^2}{2} + o(\bar{X}^2)\right) = \frac{1}{2} + \frac{o(\bar{X}^2)}{\bar{X}^2}$

$$= \frac{1}{2} + o(1) \xrightarrow{p} \frac{1}{2}.$$

(c) $\frac{2n}{S_n^2}\left(e^{\bar{X}_n} - 1 - \bar{X}_n\right) = \frac{2n}{S_n^2}\left(\frac{\bar{X}_n^2}{2} + o(\bar{X}^2)\right)$

(x) $S_n^2 \xrightarrow{p} \sigma^2$

$\therefore \frac{\sigma^2}{S_n^2} \xrightarrow{p} 1$

$$= \frac{n\bar{X}_n^2}{S_n^2} + \frac{2n\, o(\bar{X}^2)}{S_n^2}$$

$$= \underbrace{\left[\frac{\sqrt{n}(\bar{X}_n)}{\sigma}\right]^2}_{\xrightarrow{d} N(0,1)} \cdot \underbrace{\frac{\sigma^2}{S_n^2}}_{\xrightarrow{p} 1} + \underbrace{\frac{2n\, o(\bar{X}^2)}{S_n^2}}_{\xrightarrow{p} 0} \xrightarrow{d} \chi_1^2 \cdot 1$$

(d) $\frac{2\sqrt{n}}{S_n}\left(\frac{e^{\bar{X}_n} - 1 - \bar{X}_n}{\bar{X}_n}\right) = \frac{2\sqrt{n}}{S_n}\left(\frac{\bar{X}}{2} + \frac{o(\bar{X}^2)}{\bar{X}}\right) = \underbrace{\frac{\sqrt{n}\bar{X}}{\sigma}}_{\xrightarrow{d} N(0,1)} \cdot \underbrace{\frac{\sigma}{S_n}}_{\xrightarrow{p} 1} + \underbrace{\frac{2\sqrt{n}\, o(\bar{X}^2)}{S_n \bar{X}}}_{\xrightarrow{p} 0}$

$$\xrightarrow{d} N(0,1) \cdot 1$$

(e) $\underbrace{\frac{2n}{S_n^2}(e^{\bar{X}_n} - 1 - \bar{X}_n)}_{\xrightarrow{d} \chi_1^2} \underbrace{\tan \bar{X}_n}_{\substack{\xrightarrow{p} \tan 0 = 0 \\ \text{(continuous map)}}} \xrightarrow{d} 0 \cdot \chi_1^2 = 0 \Rightarrow \xrightarrow{p} 0.$

let $g(x) = \begin{cases} \frac{\tan x}{x} & x \neq 0 \\ 1 & x = 0 \end{cases}$ then by continuous map,

as $\bar{X}_n \xrightarrow{p} 0$

$$\frac{\tan \bar{X}_n}{\bar{X}_n} \to \frac{0}{0} \text{ (l'hopital's) } = \frac{\sec^2 \bar{X}_n}{1} \xrightarrow{p} 1$$

(f) $$\underbrace{\frac{2\sqrt{n}\left(e^{\bar{X}_n} - 1 - \bar{X}_n\right)}{S_n \bar{X}_n}}_{\xrightarrow{d} N(0,1)} \cdot \underbrace{\frac{\tan \bar{X}_n}{\bar{X}_n}}_{\xrightarrow{p} 1} \xrightarrow{d} N(0,1) \cdot 1$$

(slutsky's theorem used throughout these problems)

2010 THEORY 1 QUESTION 1

$(X_1, Y_1), \ldots, (X_n, Y_n)$ independent. $X_i \mid Y_i = m \sim N(\mu_m, \sigma^2)$, $P(Y_i = m) = \pi_m$, $m = 0, 1$

$$= \pi_0^{(1-y)} \pi_1^{y}, \quad y = 0, 1$$

(a) $$P(Y_i = m \mid X_i) = \frac{P(X_i \mid Y_i = m)\, P(Y_i = m)}{\sum_{m=0}^{1} P(X_i \mid Y_i = m)\, P(Y_i = m)} = \frac{\pi_m N(\mu_m, \sigma^2)}{\left[\pi_0 N(\mu_0, \sigma^2) + \pi_1 N(\mu_1, \sigma^2)\right]}, \quad m = 0, 1$$

$$= \frac{\left[\pi_0 (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left\{\frac{-1}{2\sigma^2}(X_i - \mu_0)^2\right\}\right]^{1-y} \left[\pi_1 (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left\{\frac{-1}{2\sigma^2}(X_i - \mu_1)^2\right\}\right]^{y}}{(2\pi\sigma^2)^{-\frac{1}{2}} \left[\pi_0 \exp\left\{\frac{-1}{2\sigma^2}(X_i - \mu_0)^2\right\} + \pi_1 \exp\left\{\frac{-1}{2\sigma^2}(X_i - \mu_1)^2\right\}\right]}$$

$$= \frac{\left[\pi_0 \exp\left\{\frac{-1}{2\sigma^2}(X_i - \mu_0)^2\right\}\right]^{1-y} \left[\pi_1 \exp\left\{\frac{-1}{2\sigma^2}(X_i - \mu_1)^2\right\}\right]^{y}}{\pi_0 \exp\left\{\frac{-1}{2\sigma^2}(X_i - \mu_0)^2\right\} + \pi_1 \exp\left\{\frac{-1}{2\sigma^2}(X_i - \mu_1)^2\right\}}, \quad y = 0, 1$$

$$\Rightarrow P(Y_i = 1 \mid X_i) = \frac{\pi_1 \exp\left\{\frac{-1}{2\sigma^2}(X_i - \mu_1)^2\right\}}{\pi_0 \exp\left\{\frac{-1}{2\sigma^2}(X_i - \mu_0)^2\right\} + \pi_1 \exp\left\{\frac{-1}{2\sigma^2}(X_i - \mu_1)^2\right\}}$$

$$\Rightarrow \operatorname{logit}\left[P(Y_i = 1 \mid X_i)\right] = \log\left[\frac{P(Y_i = 1 \mid X_i)}{1 - P(Y_i = 1 \mid X_i)}\right] = \log\left[\frac{P(Y_i = 1 \mid X_i)}{P(Y_i = 0 \mid X_i)}\right]$$

$$= \log\left[\frac{\pi_1 \exp\left\{\frac{-1}{2\sigma^2}(X_i - \mu_1)^2\right\}}{\pi_0 \exp\left\{\frac{-1}{2\sigma^2}(X_i - \mu_0)^2\right\}}\right]$$

$$= \log\left(\frac{\pi_1}{\pi_0}\right) + \frac{1}{2\sigma^2}\left[(X_i - \mu_0)^2 - (X_i - \mu_1)^2\right]$$

$$= X_i^2 - 2X_i\mu_0 + \mu_0^2 - (X_i^2 - 2X_i\mu_1 + \mu_1^2)$$

$$= 2X_i(\mu_1 - \mu_0) + \mu_0^2 - \mu_1^2$$

$$= \underbrace{\log\left(\frac{\pi_1}{1-\pi_1}\right) + \frac{1}{2\sigma^2}(\mu_0^2 - \mu_1^2)}_{\alpha_0} + \underbrace{\frac{(\mu_1 - \mu_0)}{\sigma^2}}_{\alpha_1} X_i$$

$$\theta = (\pi_1, \mu_0, \mu_1, \sigma^2), \quad \alpha = (\alpha_0, \alpha_1) = \left(\log\left(\frac{\pi_1}{1-\pi_1}\right) + \frac{1}{2\sigma^2}(\mu_0^2 - \mu_1^2), \; \frac{(\mu_1 - \mu_0)}{\sigma^2}\right)$$

(b) $$\log\left[\frac{P(Y_i = 1 \mid X_i, \alpha)}{1 - P(Y_i = 1 \mid X_i, \alpha)}\right] = \alpha_0 + \alpha_1 X_i \Rightarrow P(Y_i = 1 \mid X_i; \alpha) = \frac{e^{\alpha_0 + \alpha_1 x}}{1 + e^{\alpha_0 + \alpha_1 x}}, \text{ and}$$

$$P(Y_i = 0 \mid X_i; \alpha) = \frac{1}{1 + e^{\alpha_0 + \alpha_1 x}}$$

$$P(Y_i = y \mid X_i; \alpha) = \left[\frac{e^{\alpha_0 + \alpha_1 x}}{[illegible]}\right]^{y} \left[\frac{1}{[illegible]}\right]^{1-y} \quad [illegible]$$

Can use Newton Raphson to get MLEs of $\alpha = (\alpha_0, \alpha_1)$ as follows, the log-likelihood in all n observations is given as

$$\ell(\alpha_0, \alpha_1 \mid \underset{\sim}{X}, \underset{\sim}{Y}) = \sum_{i=1}^{n} [\alpha_0 y_i + \alpha_1 y_i x_i] - \log[1 + e^{\alpha_0 + \alpha_1 x_i}]$$

$$= \alpha_0 \sum_{i=1}^{n} y_i + \alpha_1 \sum_{i=1}^{n} y_i x_i - \sum_{i=1}^{n} \log[1 + e^{\alpha_0 + \alpha_1 x_i}]$$

$$\Rightarrow \dot{\ell}(\alpha) = \begin{bmatrix} \dot{\ell}(\alpha_0) \\ \dot{\ell}(\alpha_1) \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} \left(y_i - \frac{e^{\alpha_0 + \alpha_1 x_i}}{1 + e^{\alpha_0 + \alpha_1 x_i}} \right) \\ \sum_{i=1}^{n} x_i \left(y_i - \frac{e^{\alpha_0 + \alpha_1 x_i}}{1 + e^{\alpha_0 + \alpha_1 x_i}} \right) \end{bmatrix}$$

$$\Rightarrow -\ddot{\ell}(\alpha) = \begin{bmatrix} -\ddot{\ell}(\alpha_0) & -\ddot{\ell}(\alpha_0, \alpha_1) \\ -\ddot{\ell}(\alpha_1, \alpha_0) & -\ddot{\ell}(\alpha_1) \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} \frac{e^{\alpha_0 + \alpha_1 x_i}}{(1 + e^{\alpha_0 + \alpha_1 x_i})^2} & \sum_{i=1}^{n} x_i \frac{e^{\alpha_0 + \alpha_1 x_i}}{(1 + e^{\alpha_0 + \alpha_1 x_i})^2} \\ & \sum_{i=1}^{n} x_i^2 \frac{e^{\alpha_0 + \alpha_1 x_i}}{(1 + e^{\alpha_0 + \alpha_1 x_i})^2} \end{bmatrix} = I_n(\alpha)$$

with $\hat{\alpha}^{(t)}$ as the value of $\hat{\alpha} = \begin{bmatrix} \hat{\alpha}_0 \\ \hat{\alpha}_1 \end{bmatrix}$ in the t^{th} iteration (where α^0 must be specified),

the Newton-Raphson algorithm iterates the following equation until convergence:

$$\hat{\alpha}^{(t+1)} = \hat{\alpha}^{(t)} - \dot{\ell}(\alpha) \left[\ddot{\ell}(\alpha)^{-1} \right]_{\alpha = \hat{\alpha}^{(t)}} \qquad \alpha^{t+1} = \alpha^t + \left[-\ddot{\ell}(\alpha) \right]^{-1} \dot{\ell}(\alpha)$$

The asymptotic covariance of $\hat{\alpha}$ is given as $\left[\lim_{n \to \infty} \frac{1}{n} I_n(\alpha) \right]^{-1}$, where $I_n(\alpha)$ is given above.

(c) $f(x_i, y_i) = f(x_i \mid y_i) f(y_i) = (2\pi\sigma^2)^{-\frac{1}{2}} \left\{ (1-\pi_1) \exp\left[\frac{-1}{2\sigma^2} (x_i - \mu_0)^2 \right] \right\}^{1-y_i} \left\{ \pi_1 \exp\left[\frac{-1}{2\sigma^2} (x_i - \mu_1)^2 \right] \right\}^{y_i}, \quad y_i = 0, 1$ otherwise

$\Rightarrow$ with $\theta = (\pi_1, \mu_0, \mu_1, \sigma^2)$, the log-likelihood in n observations is proportional to:

$$\ell(\theta \mid \underset{\sim}{X}, \underset{\sim}{Y}) \propto -\frac{n}{2} \log \sigma^2 + \sum_{i=1}^{n} (1-y_i) \log(1-\pi_1) - \frac{(1-y_i)}{2\sigma^2} (x_i - \mu_0)^2 + y_i \log \pi_1 - \frac{y_i}{2\sigma^2} (x_i - \mu_1)^2$$

$$= -\frac{n}{2} \log \sigma^2 + \log(1-\pi_1) \sum_{i=1}^{n} (1-y_i) + \log \pi_1 \sum_{i=1}^{n} y_i - \frac{1}{2\sigma^2} \sum_{i=1}^{n} \left[(x_i - \mu_0)^2 (1-y_i) + (x_i - \mu_1)^2 y_i \right]$$

$$\Rightarrow \dot{\ell}(\pi_1) = \frac{-1}{1-\pi_1} \left(n - \sum y_i \right) + \frac{1}{\pi_1} \left(\sum y_i \right) = 0 \Rightarrow (1-\pi_1) \sum y_i - \pi_1 (n - \sum y_i) = 0 \Rightarrow \sum y_i = \pi_1 n \Rightarrow \hat{\pi}_1 = \frac{1}{n} \sum y_i$$

$$\dot{\ell}(\mu_0) = \frac{1}{\sigma^2} \sum_{y_i = 0} (x_i - \mu_0) = 0 \Rightarrow n_0 \mu_0 = \sum_{y_i = 0} x_i \Rightarrow \hat{\mu}_0 = \frac{1}{n_0} \sum_{y_i = 0} x_i$$

let this mean "all i where $y_i = 0$"

$$\dot{\ell}(\mu_1) = \frac{1}{\sigma^2} \sum_{y_i = 1} (x_i - \mu_1) = 0 \Rightarrow \hat{\mu}_1 = \frac{1}{n_1} \sum_{y_i = 1} x_i$$

n_0 of them, n_1 where $y_i = 1$ $\}$ $n_0 + n_1 = n$

$$\dot{\ell}(\sigma^2) = \frac{-n}{2\sigma^2} + \frac{1}{2\sigma^4} \left\{ \sum_{y_i = 0} (x_i - \mu_0)^2 + \sum_{y_i = 1} (x_i - \mu_1)^2 \right\} = 0 \Rightarrow \hat{\sigma}^2 = \frac{1}{n} \left\{ \sum_{y_i = 0} (x_i - \hat{\mu}_0)^2 + \sum_{y_i = 1} (x_i - \hat{\mu}_1)^2 \right\}$$

Then, $I_1(\theta) = \lim_{n\to\infty} \frac{-E}{n} \begin{bmatrix} \frac{-1}{(1-\pi_1)^2}(n-\sum y_i) - \frac{1}{\pi_1^2}\sum y_i & 0 & 0 & 0 \\ 0 & \frac{-n_0}{\sigma^2} & 0 & \frac{-1}{\sigma^4}\sum_{y_i=0}(x_i-\mu_0) \\ 0 & 0 & \frac{-n_1}{\sigma^2} & \frac{-1}{\sigma^4}\sum_{y_i=1}(x_i-\mu_1) \\ 0 & \frac{-1}{\sigma^4}\sum(x_i-\mu_0) & \frac{-1}{\sigma^4}\sum_{y_i=1}(x_i-\mu_1) & \frac{n}{2\sigma^4} - \frac{1}{\sigma^6}\left\{\sum_{y_i=0}(x_i-\mu_0)^2 + \sum_{y_i=1}(x_i-\mu_1)^2\right\} \end{bmatrix}$

$$= \lim_{n\to\infty} \frac{1}{n} \begin{bmatrix} \frac{n}{1-\pi_1} + \frac{n}{\pi_1} & 0 & 0 & 0 \\ 0 & \frac{n\pi_0}{\sigma^2} & 0 & 0 \\ 0 & 0 & \frac{n\pi_1}{\sigma^2} & 0 \\ 0 & 0 & 0 & \frac{n}{2\sigma^4} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{\pi_1(1-\pi_1)} & 0 & 0 & 0 \\ 0 & \frac{\pi_0}{\sigma^2} & 0 & 0 \\ 0 & 0 & \frac{\pi_1}{\sigma^2} & 0 \\ 0 & 0 & 0 & \frac{1}{2\sigma^4} \end{bmatrix} \Rightarrow I^{-1}(\theta) = \begin{bmatrix} \pi_1(1-\pi_1) & 0 & 0 & 0 \\ 0 & \frac{\sigma^2}{\pi_0} & 0 & 0 \\ 0 & 0 & \frac{\sigma^2}{\pi_1} & 0 \\ 0 & 0 & 0 & 2\sigma^4 \end{bmatrix}$$

$$\Rightarrow \sqrt{n}\left(\underbrace{\begin{bmatrix} \hat\pi_1 \\ \hat\mu_0 \\ \hat\mu_1 \\ \hat\sigma^2 \end{bmatrix}}_{\hat\theta_F} - \underbrace{\begin{bmatrix} \pi_1 \\ \mu_0 \\ \mu_1 \\ \sigma^2 \end{bmatrix}}_{\theta} \right) \xrightarrow{d} N_4\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \underbrace{I^{-1}(\theta)}_{\text{as given above}} \right)$$

d)

$$g(\hat\theta_F) = \left(\log\left(\frac{\hat\pi_1}{1-\hat\pi_1}\right) + \frac{1}{2\hat\sigma^2}(\hat\mu_0^2 - \hat\mu_1^2),\ \frac{(\hat\mu_1 - \hat\mu_0)}{\hat\sigma^2} \right)$$

$$\nabla g(\theta) = \begin{bmatrix} \frac{1}{\pi_1(1-\pi_1)}, & \frac{\mu_0}{\sigma^2}, & \frac{\mu_1}{\sigma^2}, & \frac{-1}{2\sigma^4}(\mu_0^2 - \mu_1^2) \\ 0 & & & \frac{\cdot}{\sigma^4}(\mu_1 - \mu_0) \end{bmatrix}$$

given above

By delta,

$$\sqrt{n}\left(g(\hat\theta_F) - g(\theta)\right) \xrightarrow{d} N_2\left(0,\ \nabla g(\theta)\, I^{-1}(\theta)\, \nabla g(\theta)'\right)$$

[illegible]

$$\Sigma^* = \begin{bmatrix} \frac{1}{\pi_1(1-\pi_1)}, & \frac{\mu_0}{\sigma^2}, & \frac{-\mu_1}{\sigma^2}, & \frac{-1}{2\sigma^4}(\mu_0^2-\mu_1^2) \\ 0, & \frac{-1}{\sigma^2}, & \frac{1}{\sigma^2}, & \frac{-1}{\sigma^4}(\mu_1-\mu_0) \end{bmatrix} \begin{bmatrix} \pi_1(1-\pi_1) & 0 & 0 & 0 \\ 0 & \frac{\sigma^2}{\pi_0} & 0 & 0 \\ 0 & 0 & \frac{\sigma^2}{\pi_1} & 0 \\ 0 & 0 & 0 & 2\sigma^4 \end{bmatrix} \begin{bmatrix} \frac{1}{\pi_1(1-\pi_1)} & 0 \\ \frac{\mu_0}{\sigma^2} & \frac{-1}{\sigma^2} \\ \frac{-\mu_1}{\sigma^2} & \frac{1}{\sigma^2} \\ \frac{-1}{2\sigma^4}(\mu_0^2-\mu_1^2) & \frac{-1}{\sigma^4}(\mu_1-\mu_0) \end{bmatrix}$$

$\pi_0 = (1-\pi_1)$

$$= \begin{bmatrix} 1 & \frac{\mu_0}{\pi_0} & \frac{-\mu_1}{\pi_1} & -(\mu_0^2-\mu_1^2) \\ 0 & \frac{-1}{\pi_0} & \frac{1}{\pi_1} & -2(\mu_1-\mu_0) \end{bmatrix} \begin{bmatrix} \frac{1}{\pi_1(1-\pi_1)} & 0 \\ \frac{\mu_0}{\sigma^2} & \frac{-1}{\sigma^2} \\ \frac{-\mu_1}{\sigma^2} & \frac{1}{\sigma^2} \\ \frac{-1}{2\sigma^4}(\mu_0^2-\mu_1^2) & \frac{-1}{\sigma^4}(\mu_1-\mu_0) \end{bmatrix}$$

$$\mathrm{cov}[g(\hat{\theta}_F)] = \begin{bmatrix} \frac{1}{\pi_1(1-\pi_1)} + \frac{\mu_0^2}{\pi_0\sigma^2} + \frac{\mu_1^2}{\pi_1\sigma^2} + \frac{1}{2\sigma^4}(\mu_0^2-\mu_1^2)^2, & \frac{-\mu_0}{\pi_0\sigma^2} - \frac{\mu_1}{\pi_1\sigma^2} + \frac{1}{\sigma^4}(\mu_0^2-\mu_1^2)(\mu_1-\mu_0) \\ \frac{-\mu_0}{\pi_0\sigma^2} - \frac{\mu_1}{\pi_1\sigma^2} + \frac{1}{\sigma^4}(\mu_0^2-\mu_1^2)(\mu_1-\mu_0), & \frac{1}{\sigma^2}\left[\frac{1}{\pi_0}+\frac{1}{\pi_1}\right] + \frac{2}{\sigma^4}(\mu_1-\mu_0)^2 \end{bmatrix}$$

SEE NOTEBOOK PAPER

(e) suppose $\mu_0 = \mu_1 = \mu$ then,

$$\mathrm{cov}[g(\hat{\theta}_F)] = \begin{bmatrix} \frac{1}{\pi_1(1-\pi_1)} + \frac{\mu}{\sigma^2}\left[\frac{1}{\pi_1}+\frac{1}{1-\pi_1}\right], & -\frac{1}{\sigma^2}\left[\frac{1}{\pi_1}+\frac{1}{1-\pi_1}\right] \\ -\frac{1}{\sigma^2}\left[\frac{1}{\pi_1}+\frac{1}{1-\pi_1}\right], & \frac{1}{\mu\sigma^2}\left[\frac{1}{\pi_1}+\frac{1}{1-\pi_1}\right] \end{bmatrix}$$

Recall $\frac{1}{\pi_1}+\frac{1}{1-\pi_1} = \frac{1}{\pi_1(1-\pi_1)}$

$$\frac{1}{\pi_1(1-\pi_1)} \begin{bmatrix} 1+\frac{\mu}{\sigma^2}, & -\frac{1}{\sigma^2} \\ -\frac{1}{\sigma^2}, & \frac{1}{\mu\sigma^2} \end{bmatrix}$$

$\mathrm{cov}(\hat{\alpha})^{-1}$ from (a) is equal to $\pi_1(1-\pi_1)\begin{bmatrix} 1 & \mu \\ \mu & \mu^2+\sigma^2 \end{bmatrix}$

$$\therefore \mathrm{cov}(\hat{\alpha})^{-1}\mathrm{cov}[g(\hat{\theta}_F)] = \begin{bmatrix} 1 & \mu \\ \mu & \mu^2+\sigma^2 \end{bmatrix}\begin{bmatrix} 1+\frac{\mu}{\sigma^2} & \frac{-1}{\sigma^2} \\ \frac{-1}{\sigma^2} & \frac{1}{\mu\sigma^2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \mu & \frac{1}{\mu} \end{bmatrix}$$

... estimated using ... doesn't depend on ...

... ...

2010 THEORY 2 QUESTION 1.

(e) $\mu_1 = \mu_0 = \mu$. Must adjust $\text{cov}[\hat{\alpha}]$ and $\text{cov}[g(\hat{\theta}_F)]$.

Under this constraint, the model in (a) becomes

$$\text{logit}[P(Y_i=1|X_i)] = \log\left(\frac{\pi_1}{1-\pi_1}\right) = \alpha_0 \quad \Rightarrow e^{\alpha_0} = \frac{\pi_1}{1-\pi_1}$$

Then, a likelihood is given as:

$$P[Y_i = y_i | x_i, \alpha] = L(\alpha|X, Y) = \left[\frac{e^{\alpha_0}}{1+e^{\alpha_0}}\right]^{y_i}\left[\frac{1}{1+e^{\alpha_0}}\right]^{1-y_i}$$

$$\therefore \ell(\alpha|x_i, y_i) = y_i \cdot \alpha_0 - \log(1+e^{\alpha_0})$$

$$\Rightarrow \dot{\ell}(\alpha|x_i,y_i) = y_i - \frac{e^{\alpha_0}}{1+e^{\alpha_0}} \quad \Rightarrow -\ddot{\ell}(\alpha|x_i,y_i) = \frac{e^{\alpha_0}}{(1+e^{\alpha_0})^2}$$

so that
$$\text{cov}(\hat{\alpha}) = \frac{(1+e^{\alpha_0})^2}{e^{\alpha_0}} = \frac{\left(\frac{1}{1-\pi_1}\right)^2}{\frac{\pi_1}{(1-\pi_1)}} = \frac{1}{\pi_1(1-\pi_1)}$$

Under this constraint, the log-likelihood in (c) is given as:

$$\ell(\theta|X, Y) \propto -\frac{n}{2}\log\sigma^2 + n_0\log(1-\pi_1) + n_1\log\pi_1 - \frac{1}{2\sigma^2}\left\{\sum(x_i-\mu)^2\right\}$$

$$\Rightarrow \dot{\ell}(\pi_1) = \frac{-n_0}{1-\pi_1} + \frac{n_1}{\pi_1} \Rightarrow -n_0\pi_1 + n_1 - n_1\pi_1 = 0 \Rightarrow n\pi_1 = n_1 \Rightarrow \hat{\pi}_1 = \frac{n_1}{n}$$

$$\dot{\ell}(\mu) = \frac{1}{\sigma^2}\sum(x_i-\mu) \Rightarrow \hat{\mu} = \frac{1}{n}\sum x_i = \bar{x}$$

$$\dot{\ell}(\sigma^2) = \frac{-n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum(x_i-\mu)^2 \Rightarrow \hat{\sigma}^2 = \frac{1}{n}\sum(x_i-\bar{x})^2$$

$$\therefore -\ddot{\ell} = \begin{bmatrix} \frac{n_0}{(1-\pi_1)^2}+\frac{n_1}{\pi_1^2} & 0 & 0 \\ 0 & \frac{n}{\sigma^2} & 0 \\ 0 & 0 & \frac{n}{2\sigma^4}\end{bmatrix}$$

$$\Rightarrow I^{-1}(\theta) = \begin{bmatrix} \pi_1(1-\pi_1) & 0 & 0 \\ 0 & \sigma^2 & 0 \\ 0 & 0 & 2\sigma^4\end{bmatrix} \qquad \frac{1-\pi_1}{\pi_1}\cdot\frac{1}{(1-\pi_1)^2} = \frac{1}{\pi_1(1-\pi_1)}$$

$$g(\theta_F) = \log\left(\frac{\pi_1}{1-\pi_1}\right) \Rightarrow \nabla g = \left[\frac{1}{\pi_1(1-\pi_1)}, 0, 0\right]$$

$$\Rightarrow \text{cov}[g(\hat{\theta}_F)] = \frac{1}{\pi_1(1-\pi_1)}, \text{ and } \text{cov}(\hat{\alpha})^{-1}\,\text{cov}[g(\hat{\theta}_F)] = 1$$

estimating same thing under same parameterization

$\alpha_0 = \log\left[\frac{\pi_1}{1-\pi_1}\right] = g(\theta)$

(f) π_1 known

(i) no change

(ii) results about (μ_0, μ_1, σ^2) do not change

(iii) [illegible]

Then, $g(\theta) = \left[\log\left(\frac{\pi_1}{1-\pi_1}\right) + \frac{1}{2\hat{\sigma}^2}\left(\hat{\mu}_0^2 - \hat{\mu}_1^2\right), \quad \frac{\hat{\mu}_1 - \hat{\mu}_0}{\hat{\sigma}^2}\right]$

$\Rightarrow \nabla g(\theta) = \begin{bmatrix} \frac{\mu_0}{\sigma^2}, & \frac{-\mu_1}{\sigma^2}, & \frac{-1}{2\sigma^4}(\mu_0^2 - \mu_1^2) \\ \frac{-1}{\sigma^2}, & \frac{1}{\sigma^2}, & \frac{-1}{\sigma^4}(\mu_1 - \mu_0) \end{bmatrix}$

$\Rightarrow \operatorname{Cov}[g(\hat{\theta}_F)] = \nabla g(\theta)\, I^{-1}(\theta)\, \nabla g'(\theta)$

$$= \begin{bmatrix} \frac{\mu_0}{\sigma^2}, & \frac{\mu_1}{\sigma^2}, & \frac{-1}{2\sigma^4}(\mu_0^2 - \mu_1^2) \\ \frac{-1}{\sigma^2}, & \frac{1}{\sigma^2}, & \frac{-1}{\sigma^4}(\mu_1 - \mu_0) \end{bmatrix} \begin{bmatrix} \frac{\sigma^2}{1-\pi_1} & 0 & 0 \\ 0 & \frac{\sigma^2}{\pi_1} & 0 \\ 0 & 0 & 2\sigma^4 \end{bmatrix}$$

$$\begin{bmatrix} \frac{\mu_0}{1-\pi_1}, & \frac{-\mu_1}{\pi_1}, & -(\mu_0^2 - \mu_1^2) \\ \frac{-1}{1-\pi_1}, & \frac{1}{\pi_1}, & -2(\mu_1 - \mu_0) \end{bmatrix} \begin{bmatrix} \frac{\mu_0}{\sigma^2}, & \frac{-1}{\sigma^2} \\ \frac{-\mu_1}{\sigma^2}, & \frac{1}{\sigma^2} \\ \frac{-1}{2\sigma^4}(\mu_0^2 - \mu_1^2), & \frac{-1}{\sigma^4}(\mu_1 - \mu_0) \end{bmatrix}$$

$$= \begin{bmatrix} \frac{\mu_0^2}{\sigma^2(1-\pi_1)} + \frac{\mu_1^2}{\sigma^2\pi_1} + \frac{1}{2\sigma^4}(\mu_0^2 - \mu_1^2)^2, & \frac{-\mu_0}{(1-\pi_1)\sigma^2} - \frac{\mu_1}{\pi_1\sigma^2} - \frac{1}{\sigma^4}(\mu_1 - \mu_0)(\mu_0^2 - \mu_1^2) \\ \text{''}, & \frac{1}{\sigma^2(1-\pi_1)} + \frac{1}{\pi_1\sigma^2} + \frac{2}{\sigma^4}(\mu_1 - \mu_0)^2 \end{bmatrix}$$

(e) $\operatorname{logit}[P(Y_i = 1 \mid X_i, q)] = \log\left(\frac{\pi_1}{1-\pi_1}\right)$, a known quantity. Thus it is meaningless to talk about MLE of $q_0 = \log\left(\frac{\pi_1}{1-\pi_1}\right)$, and its asymptotic covariance matrix

2010 THEORY 2 QUESTION 2

$$\begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} X_{11} & \cdots & X_{1p} \\ \vdots & & \\ X_{n1} & \cdots & X_{np} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} + \underbrace{\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} U}_{J_n U} + \underbrace{\begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}}_{\underset{\sim}{\varepsilon}}$$

(a) as $U \sim N(\alpha, k\sigma^2)$, $J_n U \sim N_n(J_n\alpha, k\sigma^2 \underbrace{J_n J_n'}_{J_n^n})$

+ of $\underset{\sim}{\varepsilon} \sim N_n(0, \sigma^2 I)$ $\Rightarrow$ $Y = X\underset{\sim}{\beta} + J_n U + \varepsilon \sim N_n\left(X\beta + J_n\alpha,\ \sigma^2\left[I + kJ_n^n\right]\right)$

the variance-covariance matrix is always positive semi-definite or positive definite. If we can show inverse exists, then $\Rightarrow$ positive definite

$[I + kJ][I + dJ] = I$

$\Rightarrow I + dJ + kJ + kdJ^2 = I \quad \Rightarrow dknJ = -(d+k)J$

$\Rightarrow dkn = -d-k \Rightarrow d(kn+1) = -k$

$\Rightarrow d = \frac{-k}{kn+1}$

$\Rightarrow$ inverse exists $\Rightarrow$ positive definite.

(b) to show estimability of $\theta = (\alpha, \beta_1, \dots, \beta_p)$, write

$$\begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & X_{11} & \cdots & X_{1p} \\ \vdots & \vdots & & \\ 1 & X_{n1} & \cdots & X_{np} \end{bmatrix}}_{\text{must show full rank } = p+1} \begin{bmatrix} \alpha \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} e_1 \\ \vdots \\ e_n \end{bmatrix} \quad (*)$$

$\Rightarrow Y = X_* \theta + e$

where $Y \sim N_n(X_*\theta, \sigma^2 V)$

$V = I + kJ$

Add a row to X_* as follows

$$\begin{bmatrix} 1 & X_{11} & \cdots & X_{1p} \\ \vdots & & & \vdots \\ 1 & X_{n1} & \cdots & X_{np} \\ 1 & \bar{X}_{\cdot 1} & \cdots & \bar{X}_{\cdot p} \end{bmatrix} \xrightarrow{\text{subtract last row from all previous}} \begin{bmatrix} 0 & X_{11} - \bar{X}_{\cdot 1} & \cdots & X_{1p} - \bar{X}_{\cdot p} \\ \vdots & & & \vdots \\ 0 & X_{n1} - \bar{X}_{\cdot 1} & \cdots & X_{np} - \bar{X}_{\cdot p} \\ 1 & \bar{X}_{\cdot 1} & \cdots & \bar{X}_{\cdot p} \end{bmatrix}$$

assumed to have rank p

last row must be ind of above due to a 1.

$\Rightarrow \operatorname{rank}(X_*) = p+1 \Rightarrow \theta = (\alpha, \beta_1, \dots, \beta_p)$ estimable.

(c) $Y \sim N_n(X_* \theta, \sigma^2 V)$

$\Rightarrow L(\theta, \sigma^2 | Y) \propto |\sigma^2 V| \exp\{-\frac{1}{2\sigma^2}(Y - X_*\theta)' V^{-1}(Y - X_*\theta)\}$

$\Rightarrow \ell(\theta | Y) \propto -\frac{1}{2\sigma^2}[Y'V^{-1}Y - 2Y'V^{-1}X_*\theta + \theta' X_*' V^{-1} X_* \theta]$

$\Rightarrow \dot{\ell}(\theta | Y) = -\frac{1}{2\sigma^2}[-2X_*'V^{-1}Y + 2X_*'V^{-1}X_*\theta] = 0$

$\Rightarrow X_*'V^{-1}Y = X_*'V^{-1}X_*\theta \Rightarrow \hat{\theta} = (X_*'V^{-1}X_*)^{-1}X_*'V^{-1}Y$

$\hat{\theta} \sim N_{p+1}(\theta, \sigma^2(X_*'V^{-1}X_*)^{-1})$

(d) $\tilde{\theta}$ minimizes $(Y - X_*\theta)'(Y - X_*\theta) = f(\theta)$

$f(\theta) = Y'Y - 2Y'X_*\theta + \theta'X_*'X_*\theta$

$\Rightarrow \dot{f}(\theta) = -2X_*'Y + 2X_*'X_*\theta = 0$

$\Rightarrow X_*'Y = X_*'X_*\theta \Rightarrow \tilde{\theta} = (X_*'X_*)^{-1}X_*'Y$

$\tilde{\theta} \sim N_{p+1}(\theta, \sigma^2(X_*'X_*)^{-1}X_*'VX_*(X_*'X_*)^{-1})$

$= Q'Q$ as $V = V'$

As V is positive definite, $V = QQ'$ for some Q. We can rewrite (*) as

$Q^{-1}Y = Q^{-1}X_*\theta + Q^{-1}e$, where $Q^{-1}e \sim N(0, \sigma^2 I)$

From Gauss-Markov, the least squares estimator

$\hat{\theta}_{LS} = ((Q^{-1}X_*)'(Q^{-1}X_*))^{-1}(Q^{-1}X_*)'(Q^{-1}Y)$

$= (X_*'Q^{-1'}Q^{-1}X_*)^{-1}X_*'Q^{-1'}Q^{-1}Y$

$= (X_*'V^{-1}X_*)^{-1}X_*'V^{-1}Y = \hat{\theta}$ from (c)

is the BLUE. thus, $Var(\hat{\theta}) < Var(\tilde{\theta})$ as $\hat{\theta}$ and $\tilde{\theta}$ are both unbiased.

2010 THEORY 2 QUESTION 3

$X_1 = \begin{bmatrix} X_{11} \\ X_{12} \end{bmatrix}, \ldots, X_n = \begin{bmatrix} X_{n1} \\ X_{n2} \end{bmatrix}$ iid $N_2\left(\mu_1 = \begin{bmatrix} \mu_{11} \\ \mu_{21} \end{bmatrix}, \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}\right)$

⊥ of $Y_1 = \begin{bmatrix} Y_{11} \\ Y_{12} \end{bmatrix}, \ldots, Y_m = \begin{bmatrix} Y_{m1} \\ Y_{m2} \end{bmatrix}$ iid $N_2\left(\mu_2 = \begin{bmatrix} \mu_{12} \\ \mu_{22} \end{bmatrix}, \Sigma \text{ same}\right)$

Σ positive definite (inverse exists), $\Sigma = QQ'$, $m = \gamma n$.

(a) $AUC(\beta) = P(\beta' X_1 \geq \beta' Y_1) = P(\beta'(X_1 - Y_1) \geq 0)$

$\beta'(X_1 - Y_1) \sim N_1(\beta'(\mu_1 - \mu_2), 2\beta'\Sigma\beta)$

$\Rightarrow AUC(\beta) = P(\beta'(X_1 - Y_1) \geq 0)$ $\quad Z \sim N(0,1)$ $\quad$ let $\Phi(x) = P(Z \leq x)$

$$= P\left[\frac{\beta'(X_1 - Y_1) - \beta'(\mu_1 - \mu_2)}{\sqrt{2\beta'\Sigma\beta}} \geq \frac{-\beta'(\mu_1 - \mu_2)}{\sqrt{2\beta'\Sigma\beta}}\right]$$

$$= 1 - \Phi\left[\frac{-\beta'(\mu_1 - \mu_2)}{\sqrt{2\beta'\Sigma\beta}}\right] = \Phi\left[\frac{\beta'(\mu_1 - \mu_2)}{\sqrt{2\beta'\Sigma\beta}}\right], \text{ as } 1 - \Phi(X) = \Phi(-X)$$

(b) $AUC(\beta)$ is increasing in its argument $\Rightarrow$ Maximize $AUC(\beta)$ by maximizing argument $= \dfrac{\beta'(\mu_1 - \mu_2)}{\sqrt{2\beta'\Sigma\beta}}$

$$\frac{\beta'(\mu_1 - \mu_2)}{\sqrt{2\beta'\Sigma\beta}} = \sqrt{\frac{\beta'(\mu_1 - \mu_2)(\mu_1 - \mu_2)'\beta}{2\beta'\Sigma\beta}} = \sqrt{\frac{b'c\,c'b}{2b'b}} \qquad \begin{aligned} b &= \Sigma^{\frac{1}{2}}\beta \\ c &= \Sigma^{-\frac{1}{2}}(\mu_1 - \mu_2) \end{aligned}$$

$$= \sqrt{\frac{(b'c)^2}{2b'b}}$$

$$(\text{Cauchy-Schwartz}) \leq \sqrt{\frac{(b'b)(c'c)}{2(b'b)}} = \sqrt{\frac{c'c}{2}} = \sqrt{\frac{(\mu_1 - \mu_2)'\Sigma^{-1}(\mu_1 - \mu_2)}{2}}$$

(c) $A^{optimal} = \Phi\left([(\mu_1 - \mu_2)'\Sigma^{-1}(\mu_1 - \mu_2)/2]^{\frac{1}{2}}\right)$. By invariance of MLEs,

$\hat{A} = \Phi\left\{[(\hat{\mu}_1 - \hat{\mu}_2)'\hat{\Sigma}^{-1}(\hat{\mu}_1 - \hat{\mu}_2)/2]^{\frac{1}{2}}\right\}$, where $\hat{\mu}_1$, $\hat{\mu}_2$, and $\hat{\Sigma}$ are the MLEs of μ_1, μ_2, Σ respectively.

If we let $g(z) = \Phi\left(\sqrt{\dfrac{z'z}{2}}\right)$, then $g(z) = \hat{A}$ when $z = \hat{\Sigma}^{-\frac{1}{2}}[\bar{X} - \bar{Y}]$

$g[E(z)] = A^{optimal}$, when $E[z] = \Sigma^{-\frac{1}{2}}(\mu_1 - \mu_2)$

$\sqrt{n}(\bar{X} - \mu_1) \sim N_2(0, \Sigma)$ ⊥ $\sqrt{m}(\bar{Y} - \mu_2) \sim N_2(0, \Sigma)$

$\Rightarrow \sqrt{n}(\bar{X} - \mu_1) - \sqrt{m}(\bar{Y} - \mu_2) \sim N_2(0, 2\Sigma)$

$\Rightarrow \sqrt{2}\,\Sigma^{-\frac{1}{2}}[\sqrt{n}(\bar{X} - \mu_1) - \sqrt{m}(\bar{Y} - \mu_2)] \xrightarrow{d} N_2(0, I)$, $\quad \hat{\Sigma}^{-\frac{1}{2}} \xrightarrow{P} \Sigma^{-\frac{1}{2}}$

(slutsky) $\Rightarrow \sqrt{2}\,\hat{\Sigma}^{-\frac{1}{2}}[\sqrt{n}(\bar{X} - \mu_1) - \sqrt{m}(\bar{Y} - \mu_2)] \xrightarrow{d} N_2(0, I)$?

[MORE DETAILS HERE] . Use MLE theory or reasoning on previous page?

asymptotic distribution of

(d) Can use delta method to get $\sqrt{n}(\hat{A} - A^{\text{optimal}}) \xrightarrow{d} N_1(0, \sigma^2)$

(e) Under H_0, $\sqrt{n}(\hat{A} - \frac{1}{2}) \xrightarrow{d} N(0, \sigma^2) \Rightarrow \frac{\sqrt{n}}{\sigma}(\hat{A} - \frac{1}{2}) \xrightarrow{d} N(0,1)$.

(i) $P_{H_0}(\hat{A} > c_n) = \alpha$

$\Rightarrow P_{H_0}\left[\frac{\sqrt{n}}{\sigma}(\hat{A} - \frac{1}{2}) > \frac{\sqrt{n}}{\sigma}(c_n - \frac{1}{2})\right] = \alpha$

$\Rightarrow 1 - \Phi\left[\frac{\sqrt{n}}{\sigma}(c_n - \frac{1}{2})\right] = \alpha$

$\Rightarrow \Phi\left[\frac{\sqrt{n}}{\sigma}(c_n - \frac{1}{2})\right] = 1 - \alpha \Rightarrow c_n = \frac{1}{2} + \frac{\sigma}{\sqrt{n}}\Phi^{-1}(1-\alpha)$

(ii) Under H_1, $\frac{\sqrt{n}}{\sigma}\left[\hat{A} - (\frac{1}{2} + \frac{\delta}{\sqrt{n}})\right] \xrightarrow{d} N(0,1)$

$\Rightarrow$ asymptotic power at local alternative given as:

$$P_{H_1}(\hat{A} > c_n) = P\left\{\frac{\sqrt{n}}{\sigma}\left[\hat{A} - (\tfrac{1}{2} + \tfrac{\delta}{\sqrt{n}})\right] > \frac{\sqrt{n}}{\sigma}\left[c_n - (\tfrac{1}{2} + \tfrac{\delta}{\sqrt{n}})\right]\right\}$$
$$= 1 - \Phi\left\{\frac{\sqrt{n}}{\sigma}\left[\frac{\sigma}{\sqrt{n}}\Phi^{-1}(1-\alpha) - \frac{\delta}{\sqrt{n}}\right]\right\}$$
$$= 1 - \Phi\left\{\Phi^{-1}(1-\alpha) - \frac{\delta}{\sigma}\right\}$$

($= \alpha$ when $\delta = 0$, $\to 0$ as $\delta \to \infty$).