BASIC PHD WRITTEN EXAMINATION IN BIOSTATISTICS

THEORY, SECTION 2

(9:00 AM–1:00 PM, July 27, 2016)

INSTRUCTIONS:

(a) This is a **CLOSED-BOOK** examination.

(b) The time limit for this examination is four hours.

(c) Answer any TWO (2) (BUT ONLY TWO) of the THREE (3) questions that follow.

(d) Put the answers to different questions on separate sets of paper.

(e) Put your exam code, **NOT YOUR NAME**, on each page. The same code will be used for Section 1 and Section 2 of the PhD Theory Exam. Please keep the code confidential and do not share this information with any students or faculty. Sharing your code with either students or faculty is viewed as a violation of the UNC honor code.

(f) Return the examination with a signed statement of the UNC honor pledge, separately from your answers. The pledge statement is given on the last page of the exam handout.

(g) In the questions to follow, you are required to answer only what is asked, and not to tell all you know about the topics involved.

1. (25 points) Suppose that $y_1, \ldots, y_n$ are positive and independent random variables, where

$$p(y_i|\mu_i) = \frac{1}{\mu_i} \exp(-y_i/\mu_i), \quad \mu_i > 0, \tag{1}$$

where $E(y_i|\mu_i) = \mu_i$, $i = 1, \ldots n$. Let $\theta_i = 1/\mu_i$.

(a) (3 points) Suppose that θ_i is random with $\theta_i \sim \text{Gamma}(a_i, b_i)$, where $a_i/b_i = \exp(-x_i'\beta)$ and $a_i = 3$. Further assume $\text{Var}(\theta_i) = \tau \exp(x_i'\beta)$. Here, x_i is a $p \times 1$ vector of covariates and β is a $p \times 1$ vector of regression coefficients, and β is unknown. Derive the **marginal** mean and variance of y_i, that is, compute $E(y_i)$ and $\text{Var}(y_i)$.

(b) (3 points) Under the same assumptions as part (a), derive the marginal distribution of y_i.

(c) (7 points) Under the same assumptions as part (a), derive the score test for testing $H_0 : \tau = 0$ and give its asymptotic distribution under the null hypothesis.

(d) Now suppose we take μ_i to be a **fixed and unknown parameter** and we incorporate over-dispersion by taking $\text{Var}(y_i) = \sigma^2(v_i + \mu_i)$ where v_i is the variance function of the GLM in (1). Let $\mu_i = \exp\{x_i'\beta\}$.

(i) (5 points) Derive the quasi-likelihood score equations for β and a moment estimator for σ^2.

(ii) (7 points) Let $\hat{\beta}_P$ denotes the quasi-likelihood estimate of β. Derive the asymptotic covariance matrix for $\hat{\beta}_P$.

2. (25 points) Suppose that Y is a 4×1 vector with $E(Y) = \mu$, $\mu \in E$, where E is the set $E = \{u : u' = (\beta_1 + \beta_2 - \beta_3, \beta_2 + \beta_3, -\beta_2 - \beta_3, -\beta_1 - \beta_2 + \beta_3)\}$, where the β_i are real numbers, $i = 1, 2, 3$ and a $'$ denotes matrix (vector) transposition. Further assume that $\text{Cov}(Y) = \sigma^2 I_{4\times 4}$, where σ^2 is unknown.

   (a) (5 points) Derive $\hat{\mu}$, the ordinary least squares estimate of μ, by carrying out the appropriate projection.

   (b) (4 points) Find the BLUE of $\beta_2 - \beta_3$ or show that it is nonestimable.

   (c) (4 points) Consider testing $H_0 : \beta_2 + \beta_3 = 0$ versus $H_1 : \beta_2 + \beta_3 \neq 0$. Let E_0 denote the set E assuming that H_0 is true. Explicitly give the sets E_0 and $E \cap E_0^{\perp}$, where $E_0^{\perp}$ denotes the orthogonal complement of E_0.

   (d) (6 points) Assuming normality for Y, construct the simplest possible expression for the F statistic for the hypothesis $H_0 : \mu \in E_0$ versus $H_1 : \mu \notin E_0$, where E_0 is specified in part (c), and give the distribution of the F statistic under the null and alternative hypotheses.

   (e) (6 points) Assuming normality for Y, construct an exact 95% confidence interval for $\beta_2 + \beta_3$.

3. (25 points) Consider n independent observations $(\mathbf{y}_i, \mathbf{x}_i)$ satisfying a Multivariate Linear Model (MLM) given by

$$\mathbf{y}_i = \mathbf{B}^T \mathbf{x}_i + \mathbf{e}_i, \tag{2}$$

where $\mathbf{y}_i$ is a $q \times 1$ response vector, $\mathbf{x}_i$ is a $p \times 1$ vector of covariates, and $\mathbf{B} = (\beta_{jl})$ is a $p \times q$ coefficient matrix with $\text{rank}(\mathbf{B}) = r^\star \leq \min(p, q)$. Moreover, the error term $\mathbf{e}_i \sim N(\mathbf{0}, \Sigma_R)$ for all i, where Σ_R is a $q \times q$ positive definite matrix, and the $\mathbf{x}_i$ are independently and identically distributed (i.i.d) with $E(\mathbf{x}_i) = \mu_x$ and $\text{Cov}(\mathbf{x}_i) = \Sigma_X$. Our problem of interest is to perform hypothesis testing on $\mathbf{B}$ as follows:

$$H_0 : \mathbf{CB} = \mathbf{B}_0 \quad \text{v.s.} \quad H_1 : \mathbf{CB} \neq \mathbf{B}_0, \tag{3}$$

where $\mathbf{C}$ is an $r \times p$ matrix and $\mathbf{B}_0$ is an $r \times q$ matrix. For simplicity, Σ_R is assumed to be known.

(a) (3 points) Consider a Projection Regression Modeling (PRM) given by

$$\mathbf{w}^T \mathbf{y}_i = (\mathbf{B}\mathbf{w})^T \mathbf{x}_i + \mathbf{w}^T \mathbf{e}_i = \boldsymbol{\beta}_{\mathbf{w}}^T \mathbf{x}_i + \boldsymbol{\varepsilon}_i, \tag{4}$$

where $\mathbf{w}$ is a $q \times 1$ direction vector such that $\mathbf{w}^T \mathbf{w} = 1$. For a fixed vector $\mathbf{w}$, that is independent of data, please derive the maximum likelihood estimate of $\boldsymbol{\beta}_{\mathbf{w}}$, denoted as $\widehat{\boldsymbol{\beta}}_{\mathbf{w}}$ and its distribution.

(b) (3 points) Consider the following hypotheses:

$$H_{0W} : \mathbf{C}\boldsymbol{\beta}_{\mathbf{w}} = \mathbf{b}_0 \quad \text{v.s.} \quad H_{1W} : \mathbf{C}\boldsymbol{\beta}_{\mathbf{w}} \neq \mathbf{b}_0, \tag{5}$$

where $\mathbf{C}\boldsymbol{\beta}_{\mathbf{w}} = \mathbf{CBw}$ and $\mathbf{b}_0 = \mathbf{B}_0\mathbf{w}$. We define four spaces associated with the null and alternative hypotheses of (3) and (5) as follows:

$$S_{H_0} = \{\mathbf{B} : \mathbf{CB} = \mathbf{B}_0\}, \qquad S_{H_{0W}} = \{\mathbf{B} : \mathbf{C}\boldsymbol{\beta}_{\mathbf{w}} = \mathbf{b}_0\},$$
$$S_{H_1} = \{\mathbf{B} : \mathbf{CB} \neq \mathbf{B}_0\}, \qquad S_{H_{1W}} = \{\mathbf{B} : \mathbf{C}\boldsymbol{\beta}_{\mathbf{w}} \neq \mathbf{b}_0\}.$$

Show $S_{H_0} \subset S_{H_{0W}}$ and $S_{H_{1W}} \subset S_{H_1}$ for any $\mathbf{w}$ with unit norm.

(c) (5 points) For a given $\mathbf{w}$, derive the Wald test statistic $T_n(\mathbf{w})$, its null distribution, and its mean and variance under H_{1W} conditional on $\mathbf{x}_i$s, based on model (5). Hint: for $\mathbf{u} \sim N(\mu, \Sigma_0)$, the mean and variance of $\mathbf{u}^T \Lambda \mathbf{u}$ are, respectively, given by $\text{tr}[\Lambda\Sigma_0] + \mu^T \Lambda \mu$ and $2\text{tr}[\Lambda\Sigma_0\Lambda\Sigma_0] + 4\mu^T \Lambda\Sigma_0\Lambda\mu$, where Λ is a symmetric matrix.

(d) (5 points) Show that conditional on $\mathbf{x}_i$s,

$$\text{SNR}(\mathbf{w}) = \{E_{H_1}[T_n(\mathbf{w})] - E_{H_0}[T_n(\mathbf{w})]\}/\sqrt{\text{Var}_{H_0}[T_n(\mathbf{w})]}$$

is an increasing function of $\text{HR}(\mathbf{w}) = \mathbf{w}^T\widehat{\Sigma}_C\mathbf{w}/\mathbf{w}^T\Sigma_R\mathbf{w}$. Please derive the explicit form of $\widehat{\Sigma}_C$ and its limit.

(e) (5 points) For $r = 1$, derive $\widehat{\mathbf{w}} = \text{argmax}_{\mathbf{w},\mathbf{w}^T\mathbf{w}=1}\text{HR}(\mathbf{w})$ and its limit.

(f) (4 points) Calculate $T_n(\widehat{\mathbf{w}})$ and simplify its expression as much as possible.

Statement of the UNC honor pledge:

"In recognition of and in the spirit of the honor code, I certify that I have neither given nor received aid on this examination and that I will report all Honor Code violations observed by me."

(Signed) ________________________________

NAME

(Printed) ________________________________

NAME

2016 S2 Q1.

$p(y_i|\mu_i) = \frac{1}{\mu_i} \exp(-y_i/\mu_i)$, $\mu_i > 0$. $\mathbb{E}(y_i|\mu_i) = \mu_i$. $\theta_i = \frac{1}{\mu_i}$

(a) $p(y_i|\theta_i) = \exp(-\frac{1}{\mu_i} y_i - \log \mu_i)$

$$= \exp(\theta_i(-y_i) - (-\log\theta_i))$$

Let $z_i = -y_i$

$p(z_i|\theta_i) = \exp(\theta_i z_i - \underbrace{(-\log\theta_i)}_{b(\theta_i)})$: exp'l family.

$$\mathbb{E} y_i = \mathbb{E}[-z_i] = -\mathbb{E}[\mathbb{E}(z_i|\theta_i)] = -\mathbb{E}[\dot{b}(\theta_i)] = -\mathbb{E}[-\theta_i^{-1}] = \mathbb{E}[\theta_i^{-1}]$$

$$= \int_0^\infty \theta_i^{-1} \cdot \frac{1}{\Gamma(a_i) b_i^{a_i}} \theta_i^{a_i-1} e^{-\theta_i/b_i} d\theta_i = \frac{1}{\Gamma(a_i) b_i^{a_i}} \cdot \Gamma(a_i-1) \cdot b_i^{a_i-1} = \frac{1}{(a_i-1)b_i} = \frac{1}{2b_i}$$

$$\mathrm{Var}(y_i) = \mathrm{Var}(z_i) = \mathbb{E}[\mathrm{Var}(z_i|\theta_i)] + \mathrm{Var}(\mathbb{E}[z_i|\theta_i])$$

$$= \mathbb{E}[\ddot{b}(\theta_i)] + \mathrm{Var}(\dot{b}(\theta_i))$$

$$= \mathbb{E}[\theta_i^{-2}] + \mathrm{Var}(-\theta_i^{-1})$$

$$= \mathbb{E}\theta_i^{-2} + \mathrm{Var}(\theta_i^{-1}) = 2\mathbb{E}\theta_i^{-2} - (\mathbb{E}\theta_i^{-1})^2.$$

$\mathbb{E}\theta_i^{-2} = \frac{\Gamma(a_i-2) b_i^{a_i-2}}{\Gamma(a_i) b_i^{a_i}} = \frac{1}{(a_i-1)(a_i-2) b_i^2}$ $\therefore \mathrm{Var}(y_i) = \frac{2}{(a_i-1)(a_i-2)b_i^2} - \frac{1}{(a_i-1)^2 b_i^2} = \frac{1}{b_i^2} - \frac{1}{4b_i^2}$

$$= \frac{3}{4b_i^2}$$

(b) $p(y_i) = \int p(y_i, \theta_i) d\theta_i$

$$= \int p(y_i|\theta_i) p(\theta_i) d\theta_i$$

$$= \int \theta_i \exp(-\theta_i y_i) \frac{1}{\Gamma(a_i) b_i^{a_i}} \theta_i^{a_i-1} e^{-\theta_i/b_i} d\theta_i$$

$$= \frac{1}{\Gamma(a_i) b_i^{a_i}} \int \theta_i^{a_i} \exp(-\theta_i(y_i + \frac{1}{b_i})) d\theta_i$$

$$= \frac{1}{\Gamma(a_i) b_i^{a_i}} \Gamma(a_i+1) (y_i + \frac{1}{b_i})^{-(a_i+1)} = \frac{3}{b_i^3 (y_i + \frac{1}{b_i})^4}$$

(or $\frac{3b_i^3}{(y_i+b_i)^4}$ if $Ga(a_i, \frac{1}{b_i})$ notation.)

(c) Score test for two level hierarchical overdispersion model

- $z_i|\theta_i \sim P(\theta_i,1) \Rightarrow p(z_i|\theta_i) = \exp(z_i\theta_i - b(\theta_i))$ $\qquad b(\theta_i) = -\log\theta_i$

- $E\theta_i = k(x_i^T\beta) = \exp(-x_i^T\beta)$

  $var\,\theta_i = \tau f_i(x_i^T\beta) = \tau\cdot\exp(x_i^T\beta)$. $f_i(\eta) = e^{\eta}$.

- $E(\theta_i - E\theta_i)^r = o(\tau)$, $r \geq 3$ assumption.

- $\alpha = (\beta,\tau)$.

  expectation over θ_i

  $$l_n(\alpha) = \sum_{i=1}^n \log p(z_i;\alpha) = \sum_{i=1}^n \log\left(E[p(z_i|\theta_i;\alpha)]\right)$$

- $H_0: \tau=0$ vs $H_1: \tau>0$

  Score statistic

  $$S_\tau = \partial_\alpha l_n(\tilde\alpha)^T [I_n(\tilde\alpha)]^{-1}\partial_\alpha l_n(\tilde\alpha) = (\partial_\tau l_n(\tilde\alpha))^2/\sigma_\tau^2 \xrightarrow[d]{H_0} \frac{1}{2}\chi_0^2 + \frac{1}{2}\chi_1^2$$

  $$\sigma_\tau^2 = I_{\tau\tau} - I_{\tau\beta}I_{\beta\beta}^{-1}I_{\beta\tau}\Big|_{\alpha=\tilde\alpha}.$$

- $$\partial_\tau l_n(\tilde\alpha) = \sum_i \frac{1}{2} f_i\{(y_i - \dot b(\theta_i))^2 - \ddot b(\theta_i)\}$$

  $$I_{\tau\tau} = \sum_i \frac{1}{4} f_i^2\{2\ddot b(\theta_i)^2 + b^{(4)}(\theta_i)\}$$

  $$I_{\beta\tau} = \sum_i \frac{1}{2} f_i\, b^{(3)}(\theta_i)\frac{\partial\theta_i}{\partial\beta}$$

  $$I_{\beta\beta} = \sum_i \ddot b(\theta_i)\left(\frac{\partial\theta_i}{\partial\beta}\right)^{\otimes 2}$$

- $\boxed{\text{Note}}$

$$p(z_i|\theta_i) = p(z_i|E\theta_i) + \partial_\theta p(z_i|\theta)\big|_{\theta=E\theta_i}\cdot(\theta_i - E\theta_i) + \frac{1}{2!}\partial_\theta^2 p(z_i|\theta)\big|_{\theta=E\theta_i}(\theta_i - E\theta_i)^2 + \sum_{r\geq 3}\frac{1}{r!}\partial_\theta^r p(z_i|\theta)\big|_{\theta=E\theta_i}(\theta_i - E\theta_i)^r$$

$$\Rightarrow E[p(z_i|\theta_i)] = p(z_i|E\theta_i) + \frac{1}{2!}p(z_i|\theta_i)\{(z_i - \dot b(\theta_i))^2 - \ddot b(\theta_i)\}\big|_{\theta_i=E\theta_i} var(\theta_i)$$

$$\Rightarrow \log p(z_i) = \log p(z_i|E\theta_i) + \log(1 + \Delta\tau + o(\tau)) \qquad \Delta = \frac{1}{2}\{(z_i - \dot b(\theta_i))^2 - \ddot b(\theta_i)\}\big|_{\theta_i=E\theta_i}\cdot f_i$$

$$= \begin{cases} \log p(z_i|E\theta_i) + \Delta\tau + o(\tau) & (\text{for } \beta \text{ argument}) \\ \text{or} & \\ \log p(z_i|E\theta_i) + \Delta\tau - \frac{1}{2}\Delta^2\tau^2 + o(\tau) & (\text{for } \tau \text{ argument}) \end{cases}$$

$$\log p(z_i|E\theta_i) = z_i\cdot E\theta_i - b(E\theta_i) \overset{H_0}{=} z_i\theta_i - b(\theta_i) \quad (\text{GLM setting})$$

$$\partial_\beta \log p(z_i|E\theta_i) \overset{H_0}{=} \partial_\beta(z_i\theta_i - b(\theta_i)) = z_i\frac{\partial\theta_i}{\partial\beta} - \dot b(\theta_i)\frac{\partial\theta_i}{\partial\beta} = (z_i - \dot b(\theta_i))\frac{\partial\theta_i}{\partial\beta}$$

$$\partial_\beta\Delta \overset{H_0}{=} \partial_\beta\left(\frac{1}{2}\{(z_i - \dot b(\theta_i))^2 - \ddot b(\theta_i)\} f_i\right)$$

$$= \frac{1}{2}\{2(z_i - \dot b)(-\ddot b\,\partial_\beta\theta_i) - b^{(3)}\partial_\beta\theta_i\} f_i + \frac{1}{2}\{(z_i - \dot b)^2 - \ddot b\}\dot f_i$$

$$\partial_\beta\Delta^2 = 2\Delta\,\partial_\beta\Delta$$

$$\therefore\ \partial_\tau l_i(\tilde\alpha) \overset{H_0}{=} \Delta - \Delta^2\tau\big|_{\tau=0} = \tfrac{1}{2} f_i \{(z_i - \dot b(\theta_i))^2 - \ddot b(\theta_i)\}$$

$$\partial^2_{\tau\tau} l_i(\tilde\alpha) \overset{H_0}{=} -\Delta^2 = -\tfrac{1}{4} f_i^2 \{\quad\}^2$$

$$\hookrightarrow I_{\tau\tau}(\tilde\alpha) = \mathbb{E}[-\partial^2_{\tau\tau} l_i(\alpha)]\big|_{\alpha=\tilde\alpha} = \tfrac{1}{4} f_i^2\ \mathbb{E}[((z_i - \dot b(\theta_i))^2 - \ddot b(\theta_i))^2]$$

$$= \tfrac{1}{4} f_i^2\ \mathrm{Var}((z_i - \dot b(\theta_i))^2)$$

$$= \tfrac{1}{4} f_i^2 \left(2\ddot b(\theta_i)^2 + b^{(4)}(\theta_i)\right)$$

$$\partial^2_{\beta\tau} l_i(\tilde\alpha) \overset{H_0}{=} \partial_\beta \Delta = \tfrac{1}{2}\dot f_i\{\qquad\} + \tfrac{1}{2} f_i \{2(z_i - \dot b)(-\ddot b)\partial_\beta\theta_i - b^{(3)}\partial_\beta\theta_i\}$$

$$\hookrightarrow I_{\beta\tau}(\tilde\alpha) = \tfrac{1}{2} b^{(3)}(\theta_i) f_i\, \partial_\beta \theta_i.$$

$$\mathbb{E}[\text{Blah blah}] \overset{H_0}{=} 0$$

$$\partial_\beta l_i = (z_i - \dot b(\theta_i))\partial_\beta\theta_i + \partial_\beta\Delta\cdot\tau - \tfrac{1}{2}(\partial_\beta\Delta^2)\cdot\tau^2$$

$$= \quad '' \quad + \partial_\beta\Delta\cdot\tau - \Delta\tau^2\cdot\partial_\beta\Delta$$

$$\Rightarrow \partial^2_{\beta\beta} l_i = -\ddot b(\theta_i)(\partial_\beta\theta_i)^{\otimes 2} + (z_i - \dot b(\theta_i))\frac{\partial^2\theta_i}{\partial\beta\partial\beta^T} + (\partial^2_\beta\Delta)\tau - (\partial_\beta\Delta)\tau^2(\partial_\beta\Delta) - \Delta\tau^2(\partial^2_\beta\Delta)$$

$$\hookrightarrow I_{\beta\beta}(\tilde\alpha) = \mathbb{E}[-\partial^2_{\beta\beta} l_i]\big|_{(\tilde\beta,0)} = \sum_i \ddot b(\theta_i)(\partial_\beta\theta_i)^{\otimes 2}$$

Back to the problem,

$$b(\theta_i) = -\log\theta_i \to \dot b(\theta_i) = -\theta_i^{-1} \quad \ddot b(\theta_i) = \theta_i^{-2} \quad b^{(3)}(\theta_i) = -2\theta_i^{-3} \quad b^{(4)}(\theta_i) = 6\theta_i^{-4}$$

$$\mathbb{E}\theta_i = \exp(-x_i^T\beta)$$

$$\mathrm{Var}\,\theta_i = \tau\cdot\exp(x_i^T\beta) \Rightarrow f_i(\eta) = e^\eta$$

$\tilde\alpha = (\tilde\beta, 0)^T$, where $\hat\beta$: MLE of GLM

$$S_\tau = (\partial_\tau l_n(\tilde\alpha))^2 / (I_{\tau\tau} - I_{\tau\beta} I_{\beta\beta}^{-1} I_{\beta\tau}\big|_{\alpha=\tilde\alpha})$$

$$\partial_\tau l_n(\tilde\alpha) = \sum_i \tfrac{1}{2} f_i\{(z_i - \dot b(\theta_i))^2 - \ddot b(\theta_i)\}$$

$$= \sum_i \tfrac{1}{2}\exp(x_i^T\beta)\{(-y_i + (\mathbb{E}\theta_i)^{-1})^2 - (\mathbb{E}\theta_i)^{-2}\}$$

$$= \sum_i \tfrac{1}{2}\exp(x_i^T\beta)\{(y_i - \exp(x_i^T\beta))^2 - \exp(2x_i^T\beta)\}.$$

- $$I_{\tau\tau} = \sum_i \tfrac{1}{4} f_i^2\{2\ddot b(\theta_i)^2 + b^{(4)}(\theta_i)\}$$

$$= \sum_i \tfrac{1}{4}\exp(2x_i^T\beta)\{2(\mathbb{E}\theta_i)^{-2} + 6(\mathbb{E}\theta_i)^{-4}\}$$

$$= \sum_i \tfrac{1}{4}\exp(2x_i^T\beta)\{2\exp(2x_i^T\beta) + 6\exp(4x_i^T\beta)\}.$$

- $$I_{\beta\beta} = \sum_i \ddot b(\theta_i)(\partial_\beta\theta_i)^{\otimes 2}$$

$$= \sum_i (\mathbb{E}\theta_i)^{-2}(\exp(-x_i^T\beta)\cdot(-x_i))^{\otimes 2}$$

$$= \sum_i x_i^{\otimes 2} = X^TX.$$

- $$I_{\beta\tau} = \sum_i \tfrac{1}{2} f_i\, b^{(3)}(\theta_i)\,\partial_\beta\theta_i$$

$$= \sum_i \cancel{\tfrac{1}{2}}\exp(x_i^T\beta)(\cancel{-2}(\mathbb{E}\theta_i)^{-3})\exp(-x_i^T\beta)(\cancel{-}x_i)$$

$$= \sum_i \exp(x_i^T\beta)\exp(3x_i^T\beta)\exp(-x_i^T\beta)x_i$$

$$= \sum_i \exp(3x_i^T\beta)x_i$$

$$\therefore\ S_\tau = \frac{\left(\sum_i \tfrac{1}{2}\exp(x_i^T\beta)\{(y_i - \exp(x_i^T\beta))^2 - \exp(2x_i^T\beta)\}\right)^2}{\sum_i \tfrac{1}{2}\exp(4x_i^T\beta)\{1 + 3\exp(2x_i^T\beta)\} - \underbrace{(\sum_i \exp(3x_i^T\beta)x_i)^T(X^TX)^{-1}(\sum_i\exp(3x_i^T\beta)x_i)}_{(X^T\omega)^T(X^TX)^{-1}(X^T\omega) = \omega^T X(X^TX)^{-1}X^T\omega}}$$

$$\omega = (\exp(3x_i^T\beta))^T_{1\le i\le n}$$

(d) $\mathbb{E}(y_i|\mu_i) = \mu_i = \exp(x_i^T\beta)$

$var(y_i|\mu_i) = \sigma^2(v_i + \mu_i)$

(i) Quasi-likelihood score equation

$$\sum_{i=1}^{n}\left(\frac{\partial \mathbb{E}Y_i}{\partial\beta^T}\right)^T var(Y_i)^{-1}(Y_i - \mathbb{E}Y_i) = 0$$

$$\Rightarrow \sum_{i=1}^{n}\left(\exp(x_i^T\beta)x_i\right)^T \frac{1}{\sigma^2(v_i+\mu_i)}(Y_i-\mu_i)=0$$

$$\Rightarrow \sum_{i=1}^{n} x_i \frac{\exp(x_i^T\beta)}{\sigma^2(\mu_i^2+\mu_i)}(Y_i-\mu_i) = 0$$

$$\Rightarrow \sum_{i=1}^{n} x_i \frac{1}{\exp(x_i^T\beta)+1}\left(Y_i - \exp(x_i^T\beta)\right) = 0$$

The solution β is Quasi-likelihood ~~maximum~~ estimator.

Moment estimation

$$\mathbb{E}\left[\sum \frac{(y_i - \mathbb{E}y_i)^2}{(v_i+\mu_i)}\right] = \sum \frac{\sigma^2(v_i+\mu_i)}{v_i+\mu_i} = n\sigma^2$$

$$\Rightarrow \hat{\sigma}^2 = \frac{1}{n}\sum\frac{(Y_i-\hat{\mu}_i)^2}{(\hat{\mu}_i^2+\hat{\mu}_i)} = \frac{1}{n}\sum\frac{(Y_i - \exp(x_i^T\hat{\beta}))^2}{(\exp(2x_i^T\hat{\beta}) + \exp(x_i^T\hat{\beta}))}$$

(ii) $$S_n(\beta) = \sum_{i=1}^{n}\left(\frac{\partial\mathbb{E}Y_i}{\partial\beta^T}\right)^T var(Y_i)^{-1}(Y_i - \mathbb{E}Y_i)$$

$$= \sum_{i=1}^{n} x_i \exp(x_i^T\beta)\frac{1}{\sigma^2(\mu_i^2+\mu_i)}(Y_i - \exp(x_i^T\beta))$$

$$= \frac{1}{\sigma^2}\cdot\sum_{i=1}^{n}\frac{1}{1+\exp(x_i^T\beta)}(Y_i - \exp(x_i^T\beta))x_i$$

$$\partial_\beta S_n(\beta) = \frac{1}{\sigma^2}\sum_{i=1}^{n}\frac{1}{(1+\exp(x_i^T\beta))^2}\left\{-\exp(x_i^T\beta)x_i\cdot(1+\exp(x_i^T\beta)) + (Y_i - \exp(x_i^T\beta))\exp(x_i^T\beta)x_i\right\}x_i^T$$

$$I_n(\beta) = \mathbb{E}[-\partial_\beta S_n(\beta)] = \frac{1}{\sigma^2}\sum_{i=1}^{n}\frac{1}{(1+\exp(x_i^T\beta))^2}\left\{\exp(x_i^T\beta) + \exp(2x_i^T\beta)\right\}x_i x_i^T$$

$$= \frac{1}{\sigma^2}\sum_{i=1}^{n}\frac{\exp(x_i^T\beta)}{1+\exp(x_i^T\beta)}x_i x_i^T$$

$$= \frac{1}{\sigma^2}X^T \operatorname{diag}\left(\frac{\exp(x_i^T\beta)}{1+\exp(x_i^T\beta)}\right)X$$

$$\therefore Var(\hat{\beta}_p) \simeq (I_n(\hat{\beta}_p))^{-1} = \sigma^2(X^T \quad X)^{-1}$$

Q2.

$$E = \left\{ \mu : \mu = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 1 \\ 0 & -1 & -1 \\ -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}, \beta_i \in \mathbb{R} \right\} = \operatorname{col}(X), \quad V = [\ \cdots\ ]$$

(a) We can model it by

$Y = X\beta + \varepsilon$, $\exists \beta \in \mathbb{R}^3$ s.t. $\mu = X\beta$, ε : random vector s.t. $\mathbb{E}\varepsilon = 0$, $Var(\varepsilon) = \sigma^2 I_{4\times 4}$.

Then, $\mu = X\beta$ is always estimable, and $\hat{\mu} = X\hat{\beta} = MY$, where M: o.p.o onto $Col(X)$.

Two independent columns of $X = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} \Rightarrow M =$ o.p.o onto $Col\left(\underbrace{\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & -1 \\ -1 & -1 \end{bmatrix}}_{=X_0} \right)$

$$M = X_0 (X_0^T X_0)^{-1} X_0^T$$

$$= \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & -1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ 2 & 4 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 & 0 & -1 \\ 1 & 1 & -1 & -1 \end{bmatrix} = [\quad] \frac{1}{8-4} \begin{bmatrix} 4 & -2 \\ -2 & 2 \end{bmatrix} [\quad]$$

$$= [\quad] \frac{1}{2} \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} [\quad] = [\quad] \frac{1}{2} \begin{bmatrix} 1 & -1 & 1 & -1 \\ 0 & 1 & -1 & 0 \end{bmatrix}$$

$$= \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}$$

$$\therefore \hat{\mu} = MY = \frac{1}{2} \begin{bmatrix} Y_1 - Y_4 \\ Y_2 - Y_3 \\ -Y_2 + Y_3 \\ -Y_1 + Y_4 \end{bmatrix}$$

(b) $\beta_2 - \beta_3 = \lambda^T \beta = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}^T \beta$

To check whether this is estimable, check whether $\lambda = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \in C(X^T) = C\left(\begin{bmatrix} 1 & 0 & 0 & -1 \\ 1 & 1 & -1 & -1 \\ -1 & 1 & -1 & 1 \end{bmatrix} \right)$

$$= C\left(\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ -1 & 1 \end{bmatrix} \right).$$

This does not hold. $\therefore \beta_2 - \beta_3$: <u>NOT</u> estimable.

(c) $\beta_2 + \beta_3 = \lambda^T \beta = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}^T \beta$. $\lambda = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \in C(X^T)$ holds $\therefore \beta_2 + \beta_3$: estimable.

Indeed, $\underset{1\times 3}{\lambda^T} = \underset{1\times 4}{p^T} \underset{4\times 3}{X}$, where $p = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \to \lambda = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = X^T p = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 1 & 1 & -1 & -1 \\ -1 & 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$.

$a=0, b=1, c=0, d=0$ gives it.

Thus, $\lambda^T = p^T X$, where $p = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$.

$H_0: \beta_2 + \beta_3 = 0 \Leftrightarrow H_0: \lambda^T \beta = 0 \Leftrightarrow H_0: p^T X \beta = 0 \Leftrightarrow H_0: \mathbb{E}Y \in C(p)^\perp \cap C(X)$

$\Leftrightarrow H_0: \mathbb{E}Y \in C(M) \cap C(Mp)^\perp = C(M - M_{Mp})$

Here, $C(p)^\perp \cap C(X) = C(M) \cap C(Mp)^\perp$ holds b/c $v \in LHS \Rightarrow p^T v = 0, v = Xb, \exists b \Rightarrow p^T M v = p^T v = 0$

$\Rightarrow v \in C(M) (= C(X)), v \in C(Mp)^\perp \Rightarrow v \in RHS$.

and Since $M_{Mp} \subset M$, both are o.p.o, $C(M) \cap C(Mp)^\perp = C(M - M_{Mp})$ holds.

$\therefore E_0 = C(M - M_{Mp})$. In specific, $Mp = \frac{1}{2} \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix} \Rightarrow M_{Mp} = \frac{1}{2} \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix} \left(\frac{1}{4} \cdot 2 \right)^{-1} \frac{1}{2} \begin{bmatrix} 0 & 1 & -1 & 0 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

$M - M_{MP} = \frac{1}{2}\begin{bmatrix}1&0&0&-1\\0&0&0&0\\0&0&0&0\\-1&0&0&1\end{bmatrix}$: rank 1

$\Rightarrow E_0 = C(M - M_{MP}) = C\left(\begin{bmatrix}1\\0\\0\\-1\end{bmatrix}\right) = \{u : u' = (r_1, 0, 0, -r_1),\ r_1 \in \mathbb{R}\}$

$E \cap E_0^\perp = C(M_{MP})$ $(\because)$ $E_0 = C(M) \cap C(M_{MP})^\perp$, $E = C(M) \Rightarrow E \cap E_0^\perp = C(M_{MP})$.

$= C\left(\begin{bmatrix}0\\1\\-1\\0\end{bmatrix}\right) = \{u : u' = (0, \eta_1, -\eta_1, 0),\ \eta_1 \in \mathbb{R}\}$.

(d) $H_0 : \mu \in E_0$ vs $H_1 : \mu \notin E_0$

$\Leftrightarrow H_0 : \mathbb{E}Y \in \underbrace{C(M - M_{MP})}_{= C(X_0)}$ vs $H_1 : \mathbb{E}Y \notin C(M - M_{MP}) \cap C(X)$

where $X_0 = \begin{bmatrix}1\\0\\0\\-1\end{bmatrix}$, $M_0 = M - M_{MP} = \frac{1}{2}\begin{bmatrix}1&0&0&-1\\0&0&0&0\\0&0&0&0\\-1&0&0&1\end{bmatrix}$: rank 1

If H_0 is true, then we can expect that estimation of $\mathbb{E}Y$ using projection onto $C(X)$ and $C(X_0)$ should be similar that is, $\|(M - M_0)Y\|^2$ should be small, where M_0 is o.p.o onto $C(X_0)$

Indeed, $\mathbb{E}\|(M - M_0)Y\|^2 = tr((M - M_0)\sigma^2 I_4) + \mu^T(M - M_0)\mu$

$= \sigma^2(1) + 0$ (under $H_0 : \mu \in C(X_0) \Rightarrow M\mu = M_0\mu = \mu$)

Thus, $\dfrac{\mathbb{E}\|(M - M_0)Y\|^2}{\sigma^2} = 1$ if H_0 is true.

Then, large value of $\dfrac{\|(M - M_0)Y\|^2}{\sigma^2}$ indicates H_0 is not true.

Since we don't know σ^2, substitute it with $\hat{\sigma}^2 = MSE = \|(I - M)Y\|^2 / (4 - 2)$

$\therefore F = \dfrac{\|(M - M_0)Y\|^2 / (2 - 1)}{\|(I - M)Y\|^2 / (4 - 2)}$.

To get its distribution. first note that $Y \sim N(X\beta, \sigma^2 I_4)$

$\Rightarrow Y^T(M - M_0)Y \sim \chi^2(tr(M - M_0), f_1)$, $f_1 = \frac{1}{2\sigma^2}\mu^T(M - M_0)\mu = \frac{1}{2\sigma^2}\mu^T \frac{1}{2}\begin{bmatrix}0&0&0&0\\0&1&-1&0\\0&-1&1&0\\0&0&0&0\end{bmatrix}\mu = \frac{1}{4\sigma^2}(\mu_2 - \mu_3)^2$

$Y^T(I - M)Y \sim \chi^2(tr(I - M), f_2)$, $f_2 = \frac{1}{2\sigma^2}\mu^T(I - M)\mu = 0$ $(\because)$ $\mu \in C(M)$. $= \frac{1}{\sigma^2}(\beta_2 + \beta_3)^2$

Also, $(M - M_0)Y \perp\!\!\!\perp (I - M)Y$. b/c $(M - M_0)(I - M)^T = 0$

$\therefore F \sim F(1, 2, f_1)$.

In specific. $(M - M_0)Y = \frac{1}{2}\begin{bmatrix}0&0&0&0\\0&1&-1&0\\0&-1&1&0\\0&0&0&0\end{bmatrix}Y = \frac{1}{2}\begin{bmatrix}0\\Y_2 - Y_3\\-Y_2 + Y_3\\0\end{bmatrix} \Rightarrow \|(M - M_0)Y\|^2 = \frac{1}{2}(Y_2 - Y_3)^2$

$(I - M)Y = \left(I - \frac{1}{2}\begin{bmatrix}1&0&0&-1\\0&1&-1&0\\0&-1&1&0\\-1&0&0&1\end{bmatrix}\right)Y = \begin{bmatrix}\frac{1}{2}&0&0&\frac{1}{2}\\0&\frac{1}{2}&\frac{1}{2}&0\\0&\frac{1}{2}&\frac{1}{2}&0\\\frac{1}{2}&0&0&\frac{1}{2}\end{bmatrix}Y = \frac{1}{2}\begin{bmatrix}Y_1 + Y_4\\Y_2 + Y_3\\Y_2 + Y_3\\Y_1 + Y_4\end{bmatrix}$

$\Rightarrow \|(I - M)Y\|^2 = \frac{1}{4}((Y_1 + Y_4)^2 + (Y_2 + Y_3)^2) \cdot 2$.

$\therefore F = \dfrac{2(Y_2 - Y_3)^2}{(Y_1 + Y_4)^2 + (Y_2 + Y_3)^2} \sim F\left(1, 2, \frac{1}{\sigma^2}(\beta_2 + \beta_3)^2\right)$ under after. under H_0: $f_1 = 0$.

(e) $Y \sim N(X\beta, \sigma^2 I_4) \Rightarrow Y - X\beta \sim N(0, \sigma^2 I_4)$

$$\Rightarrow (Y-X\beta)^T (M-M_0)(Y-X\beta) \sim \chi^2(1)$$

$$(Y-X\beta)^T (I-M)(Y-X\beta) \sim \chi^2(2)$$

Note that

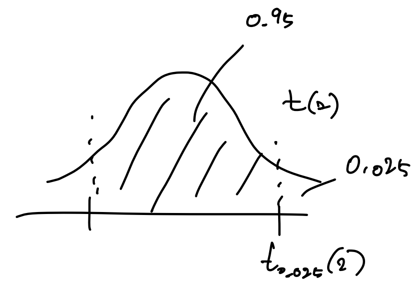
$$(M-M_0)X\beta = M_{MP}X\beta = \frac{1}{2}\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} \beta_1+\beta_2-\beta_3 \\ \beta_2+\beta_3 \\ -\beta_2-\beta_3 \\ -\beta_1-\beta_2+\beta_3 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 0 \\ 2(\beta_2+\beta_3) \\ -2(\beta_2+\beta_3) \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ \beta_2+\beta_3 \\ -\beta_2-\beta_3 \\ 0 \end{bmatrix}$$

$$(M-M_0)Y = M_{MP}Y = \quad '' \quad \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 0 \\ Y_2-Y_3 \\ -Y_2+Y_3 \\ 0 \end{bmatrix}$$

$$(I-M)X\beta = 0$$

$$\therefore F = \frac{\left\| \frac{1}{2}\begin{bmatrix} 0 \\ Y_2-Y_3 \\ -Y_2+Y_3 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ \beta_2+\beta_3 \\ -\beta_2-\beta_3 \\ 0 \end{bmatrix} \right\|^2 / 1}{\| (I-M)Y \|^2 / 2} \sim F(1,2)$$

$$= \frac{2\left((Y_2-Y_3) - 2(\beta_2+\beta_3)\right)^2}{(Y_1+Y_4)^2 + (Y_2+Y_3)^2} \sim F(1,2) = t^2(2)$$



Thus, 95% C.I for $\beta_2+\beta_3$ is given by

$$P\left\{ \frac{\sqrt{2}\,|(Y_2-Y_3) - 2(\beta_2+\beta_3)|}{\sqrt{(Y_1+Y_4)^2 + (Y_2+Y_3)^2}} \le t_{0.025}(2) \right\} = 0.95$$

$$\Rightarrow (\beta_2+\beta_3) \in \left[\frac{1}{2}(Y_2-Y_3) - \frac{t_{0.025}(2)}{2\sqrt{2}}\sqrt{(Y_1+Y_4)^2+(Y_2+Y_3)^2}, \quad + \cdots \right]$$