BASIC PHD WRITTEN EXAMINATION

THEORY, SECTION 2

(9:00 AM–1:00 PM, July 31, 2020)

INSTRUCTIONS:

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

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

(c) Answer both 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 is 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) Consider the linear model

$$Y = X\beta + Z\gamma + \epsilon, \tag{1}$$

where Y is $n \times 1$, X is and $n \times p$ of rank $r \leq p$, , Z is $n \times q$ of rank $s \leq q$, β is $p \times 1$, and γ is $q \times 1$. Further, assume that ϵ is $n \times 1$, where $E(\epsilon) = 0$ and $\text{Cov}(\epsilon) = \sigma^2 I$, where I denotes the $n \times n$ identity matrix and $(\beta, \gamma, \sigma^2)$ are all unknown parameters. Furthermore, let M denote the orthogonal projection operator onto $C(X)$, where $C(X)$ denotes the column space of X.

(a) Show that $C(X, Z) = C(X, (I - M)Z)$.

(b) Let SSE denote the error sum of squares from the model in (1). Show that

$$\text{SSE} = Y'(I - M)Y - Y'(I - M)Z[Z'(I - M)Z]^{-}Z'(I - M)Y,$$

where G^{-} denotes the generalized inverse of an arbitrary matrix G.

(c) Suppose $Z'(I - M)Z$ is non-singular. Show that the least squares estimate of γ is given by

$$\hat{\gamma} = [Z'(I - M)Z]^{-1}Z'(I - M)Y.$$

(d) Show that

(i) $\lambda'\gamma$ is estimable if and only if $\lambda' = \rho'(I - M)Z$ for some vector ρ in R^n, where λ is $q \times 1$.

(ii) if $Z'(I - M)Z$ is singular, neither γ nor $X\beta$ are estimable.

(e) Suppose γ is estimable, and let $(\hat{\beta}, \hat{\gamma})$ denote the least squares estimates of (β, γ), and assume $Z'(I - M)Z$ is non-singular. Derive the simplest possible expression for the $(n + q) \times (n + q)$ covariance matrix of $(X\hat{\beta}, \hat{\gamma})$.

(f) Now suppose $\epsilon \sim N_n(0, \sigma^2 V)$ for the model in (1), where V is positive definite and known. Further, suppose that γ is estimable, $X'V^{-1}Z = 0$, and we wish to test $H_0 : \gamma = 0$ versus $H_1 : \gamma \neq 0$. Under H_0, explicitly find projection matrices A_1 and A_2 in the simplest possible form to show that the statistic

$$B = \frac{Y'A_1Y}{Y'A_2Y}$$

has a beta distribution. Also find the distribution of B under the alternative hypothesis.

Points: (a) 3; (b) 4; (c) 4; (d) 4; (e) 5; (f) 5.

2. (25 points) Suppose that $Y_1, \cdots, Y_n$ are independent random variables and each Y_i is distributed as normal with mean $\mu > 0$ and variance μ. Define $\bar{Y}_n = (1/n)\sum_{i=1}^n Y_i$ and let $\hat{\mu}$ denote the maximum-liklelihood estimator of μ. Note: Asymptotics in this problem are in terms of $n \to \infty$. In what follows, obtain explicit expressions and simplify them as much as possible. Define any new notation you introduce.

   (a) Show that the distribution of Y_1 belong to the exponential family of distributions. Identify the canonical parameter, the canonical link function, and the cumulant function (a function of the canonical parameter).

   (b) Find the asymptotic variance of $\hat{\mu}$, that is, the limiting variance of $\sqrt{n}(\hat{\mu} - \mu)$ as $n \to \infty$. Sketch the asymptotic relative efficiency of $\bar{Y}$ relative to $\hat{\mu}$. Hint: Obtaining an explicit form for $\hat{\mu}$ is not necessary.

   (c) Now suppose Y_i is normally distributed with mean $\mu_i > 0$ and variance μ_i, and $\log(\mu_i + \mu_i^2) = x_i^\top \beta = \eta_i$, where x_i is a $p \times 1$ vector of covariates, $i = 1, \ldots, n$, and β is a vector of unknown parameters. Obtain expressions for the score function and the expected (Fisher) information matrix.

   (d) Develop a quasi-score equation that is a linear combination of the residuals $Y_i - \mu_i$ for estimating β in the regression model in part (c). Let $\tilde{\beta}$ denote the resulting estimator. Obtain an expression for the asymptotic covariance matrix of $\sqrt{n}(\tilde{\beta} - \beta)$. Develop an expression for the "sandwich" (robust, empirical) covariance (covariance of $\tilde{\beta}$) estimator.

   (e) In the context of the regression model part in (c), suppose that $x_i^\top = (1, 0)$ for $i = 1, \ldots, n_0$ and $x_i^\top = (1, 1)$ for $i = 1 + n_0, \ldots, n$, and $n_0/n \to \gamma$ as $n \to \infty$, where $\gamma \in (0, 1)$ is a known constant. The parameter vector is $(\beta_1, \beta_2)^\top$. Derive the score test for testing the hypothesis $H_0 : \beta_2 = 0$ against $H_1 : \beta_2 \neq 0$.

Points: (a) 5; (b) 5; (c) 5; (d) 5; (e) 5.

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

2020 S2 Q1.

(a) $v \in C(X,Z) \iff v = Xc_1 + Zc_2 \;,\; \exists c_1, c_2$

$\iff v = X\left(c_1 + (X^TX)^- X^T Z c_2\right) + (I-M)Zc_2 \;,$ where $c_2 = d_2,\ c_1 = d_1 - (X^TX)^- X^T Z d_2$

$\iff v = Xd_1 + (I-M)Zd_2 \;,\; \exists d_1, d_2$

$\iff v \in C(X, (I-M)Z)$

(b) $Z_0 := (I-M)Z$

$C(X) \perp C(Z_0) \quad \because X^T Z_0 = 0$

$$C(X,Z) = C(X, Z_0) \underset{(a)}{=} C(X) \oplus C(Z_0) \;:\; \text{orthogonal decomposition}$$

o.p.o $\downarrow \qquad \downarrow \quad \downarrow$

$\Rightarrow \quad M_F \quad = \quad M \; + \; M_0$

$\therefore SSE = Y^T(I-M_F)Y$

$= Y^T(I-M)Y - Y^T M_0 Y = Y^T(I-M)Y - Y^T(I-M)Z\left[Z^T(I-M)Z\right]^- Z^T(I-M)Y$

(c) $Y = X\beta + Z\gamma + \epsilon \iff Y = X\eta + Z_0\gamma + \epsilon \,,$ where $\eta = \beta + (X^TX)^- X^T Z\gamma$

<u>Gauss–Markov Thm</u> : $\left[X \,\vdots\, Z_0\right] \begin{bmatrix} \eta \\ \gamma \end{bmatrix}^{\wedge \text{BLUE}} = M_F Y$

<u>Normal Equation</u> : $\begin{bmatrix} X^T \\ Z_0^T \end{bmatrix} \begin{bmatrix} X & Z_0 \end{bmatrix} \begin{bmatrix} \hat\eta \\ \hat\gamma \end{bmatrix} = \begin{bmatrix} X^T \\ Z_0^T \end{bmatrix} Y$

Note

$X\hat\beta = MY \iff X^TX\hat\beta = X^TY$

$\Rightarrow \begin{bmatrix} X^TX & 0 \\ 0 & Z_0^TZ_0 \end{bmatrix} \begin{bmatrix} \hat\eta \\ \hat\gamma \end{bmatrix} = \quad // \qquad (\because)\ X^T Z_0 = 0$

$\Rightarrow \hat\gamma = (Z_0^TZ_0)^{-1} Z_0^T Y$

$= \left[Z^T(I-M)Z\right]^{-1} Z^T(I-M)Y.$

(d)

(i) $\lambda^T\gamma$: estimable $\iff \lambda^T\gamma = \mathbb{E}(\rho^T Y) \,,\ \forall \beta, \gamma \,,\ \exists \rho$

$\iff \lambda^T\gamma = \rho^T X\beta + \rho^T Z\gamma \,,\ \forall \beta, \gamma, \exists \rho$

$\iff \lambda^T = \rho^T Z \,,\quad \rho^T X = 0$

$\iff \lambda^T = \rho^T Z \,,\quad \rho \in C(X)^\perp = C(I-M)$

$\iff \lambda^T = \rho^T(I-M)Z \,,\ \exists \rho$ s.t $\rho = (I-M)\rho.$

(d)

(ii) From (c), $\hat{\gamma}^{BLUE}$ is given by any γ s.t $(Z_0^T Z_0)\gamma = Z_0^T Y$

If $Z^T(I-M)Z = Z_0^T Z_0$ is singular, then the above equation does not have a unique BLUE.

However, if γ is estimable, then by <u>Gauss-Markov Thm</u>, $\hat{\gamma}^{BLUE}$ is unique.

Thus, γ is <u>NOT</u> estimable.

From (c), $\widehat{X\beta}^{BLUE}$ is given by any $X\beta$ s.t $X^T X\eta = X^T Y$

$$\Leftrightarrow X^T X\beta + X^T Z\hat{\gamma} = X^T Y.$$

$$\Leftrightarrow X^T \cdot X\beta = X^T Y - X^T Z\hat{\gamma}$$

$$\Leftrightarrow X(X^T X)^- X^T \cdot X\beta = X(X^T X)^- (X^T Y - X^T Z\hat{\gamma})$$

$$\Leftrightarrow X\beta = X(X^T X)^- X^T Y - X(X^T X)^- X^T Z\hat{\gamma}$$

$$= M(Y - Z\hat{\gamma}).$$

Since $\hat{\gamma}^{BLUE}$ is <u>NOT</u> unique when $Z^T(I-M)Z$ is singular, the above equation does not have a unique solution.

Thus, by <u>Gauss-Markov Thm</u>, $X\beta$ is <u>NOT</u> estimable.

(e) $\begin{cases} X^T X\hat{\eta} = X^T Y \\ Z_0^T Z_0 \hat{\gamma} = Z_0^T Y \end{cases} \Rightarrow \begin{cases} X\hat{\beta} = M(Y - Z\hat{\gamma}) \\ \hat{\gamma} = (Z_0^T Z_0)^{-1} Z_0^T Y \end{cases} \Rightarrow \begin{cases} X\hat{\beta} = M(I - Z(Z_0^T Z_0)^{-1} Z_0^T)Y \\ \hat{\gamma} = (Z_0^T Z_0)^{-1} Z_0^T Y. \end{cases}$

$$Cov\left(\begin{bmatrix} X\hat{\beta} \\ \hat{\gamma} \end{bmatrix}\right) = \begin{bmatrix} Var(X\hat{\beta}) & Cov(X\hat{\beta}, \hat{\gamma}) \\ Cov(\hat{\gamma}, X\hat{\beta}) & Var(\hat{\gamma}) \end{bmatrix}$$

$$= \sigma^2 \cdot \begin{bmatrix} M(I - Z(Z_0^T Z_0)^{-1} Z_0^T)(I - Z_0(Z_0^T Z_0)^{-1} Z)M & M(I - Z(Z_0^T Z_0)^{-1} Z_0^T) Z_0 (Z_0^T Z_0)^{-1} \\ \text{trans pose.} & (Z_0^T Z_0)^{-1} \end{bmatrix}$$

$$= \sigma^2 \begin{bmatrix} M + MZ(Z^T(I-M)Z)^{-1} Z^T M & -MZ(Z^T(I-M)Z)^{-1} \\ -(Z^T(I-M)Z)^{-1} Z^T M & (Z^T(I-M)Z)^{-1} \end{bmatrix}$$

(ii) $MZ_0 = 0$

(f) $Y = X\beta + Z\gamma + \varepsilon$

$\Rightarrow V^{-1/2}Y = V^{-1/2}X\beta + V^{-1/2}Z\gamma + V^{-1/2}\varepsilon$

$\Rightarrow \quad Y_* = X_*\beta + Z_*\gamma + \varepsilon_*, \quad \varepsilon_* \sim N(0, \sigma^2 I)$

$H_0: Y_* = X_*\beta + \varepsilon_* \iff H_0: \mathbb{E}Y_* \in C(X_*)$

Full model : $\mathbb{E}Y_* \in C(X_*, Z_*) = C(X_*) \oplus C(Z_*)$ $\quad (\because)\ X_*^T Z_* = 0 \Rightarrow$ orthogonal decomposition

Reduced " : $\mathbb{E}Y_* \in C(X_*)$

Test statistic :
$$F = \frac{Y_*^T M_{Z_*} Y_* / r(M_{Z_*})}{Y_*^T (I - M_{F_*}) Y_* / r(I - M_{F_*})} \overset{H_0}{\sim} F(r(M_{Z_*}), r(I - M_{F_*}))$$

Here, $Y_*^T M_{Z_*} Y_* / \sigma^2 \sim \chi^2(r(M_{Z_*}), \delta) \overset{d}{=} \chi^2\Big(s, \underbrace{\frac{1}{2\sigma^2}\|M_{Z_*}(X_*\beta + Z_*\gamma)\|^2}_{= \frac{1}{2\sigma^2}\gamma^T Z_*^T Z_* \gamma}\Big)$

$Y_*^T (I - M_{F_*}) Y_* / \sigma^2 \sim \chi^2(r(I - M_{F_*})) \overset{d}{=} \chi^2(n - s - r)$

$Y_*^T M_{Z_*} Y_* \perp\!\!\!\perp Y_*^T (I - M_{F_*}) Y_*$

$$\Rightarrow B = \frac{Y_*^T M_{Z_*} Y_*}{Y_*^T M_{Z_*} Y_* + Y_*^T (I - M_{F_*}) Y_*} \overset{d}{=} \frac{\chi^2(s, \delta)}{\chi^2(s, \delta) + \chi^2(n - s - r)} \overset{H_0}{\sim} \text{Beta}\left(\frac{s}{2}, \frac{n-s-r}{2}\right)$$

Indeed, $$B = \frac{Y_*^T M_{Z_*} Y_*}{Y_*^T (I - M_{F_*} + M_{Z_*}) Y_*} = \frac{Y_*^T M_{Z_*} Y_*}{Y_*^T (I - M_{X_*}) Y_*}$$ $\quad (\because)\ M_{F_*} = M_{X_*} + M_{Z_*}$ by ortho. decomp.

$$= \frac{Y^T V^{-1} (Z^T V^{-1} Z)^{-1} V^{-1} Y}{Y^T (V^{-1} - V^{-1} (X^T V^{-1} X)^{-1} V^{-1}) Y} = \frac{Y^T A_1 Y}{Y^T A_2 Y} \text{ "}$$

Under $H_1: \gamma \neq 0, \quad B \sim \text{Beta}\left(\frac{s}{2}, \frac{n-s-r}{2}; \delta\right)$

Indeed, A_1 & A_2 are NOT projection matrices. The problem is wrong.

2020 S2 Q2

(a) $p(y|\mu) = \exp\left(\left(-\frac{1}{2\mu}\right) y^2 - \left(\frac{1}{2}\mu + \frac{1}{2}\log\mu\right) - \left(\frac{1}{2}\log 2\pi - y\right)\right)$

$= \exp\left(\theta \cdot T(y) - b(\theta) - c(y) \right)$: Exp'l family.

Canonical parameter $\theta = -\frac{1}{2\mu}$

Canonical link $g(\mu) = -\frac{1}{2\mu}$

Cumulant ft^n $K_Y(t) = \log \mathbb{E} e^{tY} = \log \left(\int_{\mathbb{R}} e^{ty} \cdot (2\pi\mu)^{-1/2} \exp\left(-\frac{1}{2\mu}(y-\mu)^2\right) dy \right)$

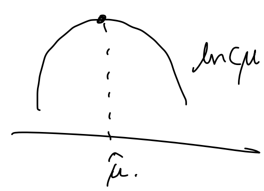
$= \log \left((2\pi\mu)^{-1/2} \int_{\mathbb{R}} \exp\left[-\frac{1}{2\mu}\left(y - \mu(1+t)\right)^2 + \frac{\mu}{2}(1+t)^2 - \frac{\mu}{2} \right] dy \right)$

$= \frac{\mu}{2}(2t + t^2) = \mu t + \frac{\mu}{2} t^2$

(b) $l_n(\mu) = \sum_i \left(-\frac{1}{2\mu}\right) y_i^2 - \left(\frac{1}{2}\mu + \frac{1}{2}\log\mu\right) - \left(\frac{1}{2}\log 2\pi - y_i\right)$

$\dot{l}_n(\mu) = \frac{1}{2}\mu^{-2} \sum_i y_i^2 - \left(\frac{1}{2} + \frac{1}{2}\mu^{-1}\right) \cdot n \overset{set}{=} 0 \Rightarrow \mu^2 + \mu - \frac{1}{n}\sum_i y_i^2 \overset{set}{=} 0$

$\Rightarrow \hat{\mu} = -\frac{1}{2} + \sqrt{\overline{y^2} + \frac{1}{4}}$ ($\because \hat{\mu} > 0$)



$\ddot{l}_n(\mu) = n\mu^{-3}\left(-\overline{y^2} + \frac{1}{2}\mu\right) \Rightarrow \ddot{l}_n(\hat{\mu}) < 0$ implies that $\hat{\mu}$ is MLE.

$I_n(\mu) = \mathbb{E}\left[-\ddot{l}_n(\mu)\right] = n\mu^{-3}\left(\mathbb{E}Y_1^2 - \frac{1}{2}\mu\right)$

$= n\mu^{-3}\left(\mu + \mu^2 - \frac{1}{2}\mu\right) = n\mu^{-2}\left(\frac{1}{2} + \mu\right)$

$I(\mu) = \lim_{n\to\infty} \frac{1}{n} I_n(\mu) = \mu^{-2}\left(\frac{1}{2} + \mu\right)$

<u>Asymptotic Normality of MLE in Exp'l family</u> : $\sqrt{n}(\hat{\mu} - \mu) \to_d N(0, I(\mu)^{-1})$

$\to_d N\left(0, \frac{\mu^2}{\frac{1}{2} + \mu}\right)$

CLT : $\sqrt{n}(\bar{Y} - \mathbb{E}Y_1) \to_d N(0, \mathrm{Var}\, Y_1) \Rightarrow \sqrt{n}(\bar{Y} - \mu) \to_d N(0, \mu)$

$\mathrm{ARE}(\bar{Y}, \hat{\mu}) = \frac{\mathrm{Asymp.Var}\left(\sqrt{n}(\bar{Y}-\mu)\right)^{-1}}{\mathrm{Asymp.Var}\left(\sqrt{n}(\hat{\mu}-\mu)\right)^{-1}} = \frac{\mu^{-1}}{\left(\frac{\mu^2}{\frac{1}{2}+\mu}\right)^{-1}} = \frac{\mu}{\frac{1}{2} + \mu} < 1$

(c) $g(\mu) = \log(\mu + \mu^2)$. $\Rightarrow g(\mu_i) = x_i^T\beta \Rightarrow \mu_i = \left(\exp(x_i^T\beta) + \frac{1}{4}\right)^{1/2} - \frac{1}{2}$ ($\because \mu_i > 0$)

<Score>

$l_n(\beta) = \sum_i \left(-\frac{1}{2\mu_i(\beta)}\right) y_i^2 - \left(\frac{1}{2}\mu_i(\beta) + \frac{1}{2}\log \mu_i(\beta)\right) - \left(\frac{1}{2}\log 2\pi - y_i\right)$

$\frac{\partial l_n}{\partial \beta} = \sum_i \frac{1}{2}\mu_i^{-2} \partial_\beta \mu_i \cdot y_i^2 - \frac{1}{2}\partial_\beta \mu_i + \frac{1}{2}\mu_i^{-1} \partial_\beta \mu_i$

Scalar form

$= \sum_i \frac{1}{2}\mu_i^{-2}\left(y_i^2 - (\mu_i + \mu_i^2)\right) \partial_\beta \mu_i = \sum_i \frac{1}{2}\left[\left(\exp(x_i^T\beta) + \frac{1}{4}\right)^{1/2} - \frac{1}{2}\right]^{-2} \left(y_i^2 - \exp(x_i^T\beta)\right) \cdot \frac{e^{x_i^T\beta}}{2\left(\exp(x_i^T\beta) + \frac{1}{4}\right)^{1/2}} \cdot x_i$

(c) (cont'ed)

$$\frac{\partial l_n}{\partial \beta} = \left[\partial_\beta \mu_1 \vdots \cdots \vdots \partial_\beta \mu_n \right]_{p\times n} \begin{bmatrix} \frac{1}{2}\mu_1^{-2} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{2}\mu_n^{-2} \end{bmatrix}_{n\times n} \begin{bmatrix} y_1^2 - (\mu_1+\mu_1^2) \\ \vdots \\ y_n^2 - (\mu_n+\mu_n^2) \end{bmatrix}_{n\times 1}$$

$$= \quad D(\beta)^T \quad V(\beta)^{-1} \quad e(\beta)$$: matrix form

where $D(\beta)^T = \left[\frac{\partial \mu^T}{\partial \beta}\right]_{p\times n} = \left[\frac{e^{x_i^T\beta}}{2(e^{x_i^T\beta}+\frac{1}{4})^{1/2}} x_i \right]_{p\times n}$

$V(\beta) = \mathrm{diag}(2\mu_i^2) = \mathrm{diag}\left(2\left[(e^{x_i^T\beta}+\frac{1}{4})^{1/2} - \frac{1}{2}\right]^2 \right)$

$e(\beta) = (y_i^2 - (\mu_i+\mu_i^2))_{n\times 1} = (y_i^2 - e^{x_i^T\beta})_{n\times 1}$

<Fisher Information>

$$\frac{\partial^2 l_n}{\partial\beta\partial\beta^T} = \sum_i -\mu_i^{-3}(y_i^2-(\mu_i+\mu_i^2))\left(\frac{\partial\mu_i}{\partial\beta}\right)^{\otimes 2} + \frac{1}{2}\mu_i^{-2}(-(1+2\mu_i))\left(\frac{\partial\mu_i}{\partial\beta}\right)^{\otimes 2} + \frac{1}{2}\mu_i^{-2}(y_i^2-(\mu_i+\mu_i^2))\frac{\partial^2\mu_i}{\partial\beta\partial\beta^T}$$

Using the fact that $\mathbb{E}(y_i^2 - (\mu_i+\mu_i^2)) = 0$,

$$I_n(\beta) = \mathbb{E}\left[-\frac{\partial^2 l_n}{\partial\beta\partial\beta^T}\right] = \sum_i \frac{1}{2}\frac{1+2\mu_i}{\mu_i^2}\left(\frac{\partial\mu_i}{\partial\beta}\right)^{\otimes 2}$$

$$= \sum_i \frac{(e^{x_i^T\beta}+\frac{1}{4})^{1/2}}{\left[(e^{x_i^T\beta}+\frac{1}{4})^{1/2}-\frac{1}{2}\right]^2} \cdot \frac{e^{2x_i^T\beta}}{4(e^{x_i^T\beta}+\frac{1}{4})} x_i x_i^T$$: scalar form

$$= \left[\partial_\beta \mu_1 \vdots \cdots \vdots \partial_\beta \mu_n \right]_{p\times n} \begin{bmatrix} \frac{1+2\mu_1}{2\mu_1^2} & & 0 \\ & \ddots & \\ 0 & & \frac{1+2\mu_n}{2\mu_n^2} \end{bmatrix}_{n\times n} \begin{bmatrix} \partial_\beta \mu_1^T \\ \vdots \\ \partial_\beta \mu_n^T \end{bmatrix}_{n\times p}$$

$$= D(\beta)^T W(\beta) D(\beta)$$: matrix form

(d) <Quasi-Score equation>

$$S_n(\beta) = \sum_i \frac{(\partial_\beta \mu_i)(y_i-\mu_i)}{var(y_i)} = 0$$

$$\Leftrightarrow \quad S_n(\beta) = \sum_i \frac{\partial_\beta \mu_i}{\mu_i} e_i = 0$$

$$\Rightarrow \quad \frac{\partial S_n}{\partial \beta} = \sum_i -\frac{e_i}{\mu_i^2}\partial_\beta \mu_i^{\otimes 2} + \frac{e_i}{\mu_i}\frac{\partial^2 \mu_i}{\partial\beta\partial\beta^T} - \frac{1}{\mu_i}\partial_\beta \mu_i^{\otimes 2}$$

$$\Rightarrow \quad E\left[-\frac{\partial S_n}{\partial \beta}\right] = \sum_i \frac{1}{\mu_i}\partial_\beta \mu_i^{\otimes 2} \qquad (\because)\ E e_i = 0.$$

<Taylor Expansion> $0 = S_n(\tilde\beta) = S_n(\beta) + \frac{\partial S_n}{\partial \beta}(\beta^*)(\tilde\beta-\beta)$, $\exists\, \beta^* = t\tilde\beta + (1-t)\beta$.

$$\Rightarrow \quad \sqrt{n}(\tilde\beta-\beta) = \left[-\frac{1}{n}\frac{\partial S_n}{\partial\beta}(\beta^*)\right]^{-1}\frac{1}{\sqrt n}S_n(\beta)$$

$$= \left(\frac{1}{n}E\left[-\frac{\partial S_n}{\partial\beta}\right]\right)^{-1}\cdot\frac{1}{\sqrt n}S_n(\beta) + o_p(1) \qquad (\because)\ \text{WLLN}$$

$$\Rightarrow \quad Cov(\sqrt n(\tilde\beta-\beta)) \simeq \left(\frac{1}{n}E\left[-\frac{\partial S_n}{\partial\beta}\right]\right)^{-1}\cdot Cov\left(\frac{1}{\sqrt n}S_n(\beta)\right)\cdot\left(\frac{1}{n}E\left[-\frac{\partial S_n}{\partial\beta}\right]\right)^{-1}$$

$$= n\cdot\left(\sum_i \frac{1}{\mu_i}\partial_\beta\mu_i^{\otimes 2}\right)^{-1}\cdot\sum_i\frac{(\partial_\beta\mu_i)^{\otimes 2}}{var(y_i)^2}var(y_i)\cdot\left(\sum_i\frac{1}{\mu_i}\partial_\beta\mu_i^{\otimes 2}\right)^{-1}$$

$$= n\cdot\left(\sum_i\frac{1}{\mu_i}\partial_\beta\mu_i^{\otimes 2}\right)^{-1}$$: asymptotic covariance matrix of $\sqrt n(\tilde\beta-\beta)$.

$$= n\cdot(D^TV^{-1}D)^{-1}$$

<Sandwich covariance estimator>

$$\widehat{Cov}(\tilde\beta) = \left(\hat{E}\left[-\frac{\partial S_n}{\partial\beta}\right]\right)^{-1}\widehat{Cov}(S_n(\beta))\left(\hat{E}\left[-\frac{\partial S_n}{\partial\beta}\right]\right)^{-1}\Big|_{\beta=\tilde\beta}$$

$$= \left(\sum_i\frac{1}{\mu_i}\partial_\beta\mu_i^{\otimes 2}\right)^{-1}\left(\sum_i\frac{(\partial_\beta\mu_i)^2}{var(y_i)^2}(y_i-\mu_i)^2\right)\left(\sum_i\frac{1}{\mu_i}\partial_\beta\mu_i^{\otimes 2}\right)^{-1}\Big|_{\beta=\tilde\beta}$$

Note: $cov(S_n(\beta)) = E S_n(\beta)^{\otimes 2} - (E S_n(\beta))^{\otimes 2} = E S_n(\beta)^{\otimes 2} = E\sum_i\frac{(\partial_\beta\mu_i)^2}{var(y_i)^2}(y_i-\mu_i)^2$

$\Rightarrow \widehat{Cov}(S_n(\beta)) = \sum_i\frac{(\partial_\beta\mu_i)^2}{var(y_i)^2}\hat{E}(y_i-\mu_i)^2 \simeq \sum_i\frac{(\partial_\beta\mu_i)^2}{var(y_i)^2}(y_i-\mu_i)^2$

$$= (D^TV^{-1}D)^{-1}(D^TV^{-1}EV^{-1}D)(D^TV^{-1}D)^{-1}\Big|_{\beta=\tilde\beta}$$

, where $E = diag((y_i-\mu_i)^2)$

Finally, plug $\mu_i = (e^{x_i^T\beta}+\frac14)^{1/2}-\frac12$, $\partial_\beta\mu_i = \frac{e^{x_i^T\beta}}{2(e^{x_i^T\beta}+\frac14)^{1/2}}x_i$ into the equations

(e) Let $n_1 = n - n_0$

$i \in A,\ |A| = n_0 \Rightarrow x_i = (1,0)^T \Rightarrow \mu_0 : \log(\mu_0 + \mu_0^2) = x_i^T\beta = \beta_1 \quad \Rightarrow \partial_\beta \mu_0 = \dfrac{e^{\beta_1}}{2(e^{\beta_1}+\frac{1}{4})^{1/2}} \begin{bmatrix}1\\0\end{bmatrix}$

$j \in B,\ |B| = n_1 \Rightarrow x_j = (1,1)^T \Rightarrow \mu_1 : \log(\mu_1 + \mu_1^2) = x_j^T\beta = \beta_1 + \beta_2 \Rightarrow \partial_\beta \mu_1 = \dfrac{e^{\beta_1+\beta_2}}{2(e^{\beta_1+\beta_2}+\frac{1}{4})^{1/2}} \begin{bmatrix}1\\1\end{bmatrix}$

$$l_n(\beta) = \sum_i -\frac{1}{2\mu_i} y_i^2 - \left(\frac{1}{2}\mu_i + \frac{1}{2}\log\mu_i\right) - \left(\frac{1}{2}\log 2\pi - y_i\right)$$

$$\partial_\beta l_n = \sum_i \frac{1}{2\mu_i^2}\left(y_i^2 - (\mu_i + \mu_i^2)\right)\partial_\beta \mu_i$$

$$= \sum_{i\in A} \frac{1}{2\mu_0^2}(y_i^2 - e^{\beta_1}) \frac{\mu_0 + \mu_0^2}{1+2\mu_0}\begin{bmatrix}1\\0\end{bmatrix}$$

$\boxed{\text{Note}}$ $\quad \partial_\beta \mu_i = \dfrac{\mu_i + \mu_i^2}{1+2\mu_i} x_i$

$$+ \sum_{j\in B} \frac{1}{2\mu_1^2}(y_j^2 - e^{\beta_1+\beta_2}) \frac{\mu_1 + \mu_1^2}{1+2\mu_1}\begin{bmatrix}1\\1\end{bmatrix}$$

Let $\tilde{\beta} = (\tilde{\beta}_1, 0)$: MLE under $H_0 : \beta_2 = 0$

Then, $l_n^{H_0}(\beta) = l_n(\beta_1, 0) = \sum_i -\frac{1}{2\mu_0} y_i^2 - \left(\frac{1}{2}\mu_0 + \frac{1}{2}\log\mu_0\right) - \left(\frac{1}{2}\log 2\pi - y_i\right)$ $\quad (\because)$ all same mean "μ_0"

$$\Rightarrow \frac{\partial}{\partial\beta_1} l_n^{H_0} = \sum_i \frac{1}{2\mu_0^2}\left(y_i^2 - (\mu_0+\mu_0^2)\right)\partial_{\beta_1}\mu_0 \overset{\text{set}}{=} 0$$

$$\Leftrightarrow \sum_i \frac{1}{2\mu_0^2}(y_i^2 - e^{\beta_1}) \cdot \frac{e^{\beta_1}}{2(e^{\beta_1}+\frac{1}{4})^{1/2}} \overset{\text{set}}{=} 0 \quad \Leftrightarrow \quad \sum_i (y_i^2 - e^{\beta_1}) = 0$$

$$\Leftrightarrow \tilde{\beta}_1 = \log\left(\overline{y^2}\right)$$

Then,

$$\partial_\beta l_n(\tilde{\beta}) = \sum_{i\in A} \frac{1}{2\tilde{\mu}_0^2}(y_i^2 - e^{\tilde{\beta}_1}) \frac{\tilde{\mu}_0 + \tilde{\mu}_0^2}{1+2\tilde{\mu}_0}\begin{bmatrix}1\\0\end{bmatrix} + \sum_{j\in B} \frac{1}{2\tilde{\mu}_1^2}(y_j^2 - e^{\tilde{\beta}_1+0}) \frac{\tilde{\mu}_1 + \tilde{\mu}_1^2}{1+2\tilde{\mu}_1}\begin{bmatrix}1\\1\end{bmatrix}$$

where $\tilde{\mu}_0 : \log(\tilde{\mu}_0 + \tilde{\mu}_0^2) = \tilde{\beta}_1$ & $\tilde{\mu}_1 : \log(\tilde{\mu}_1 + \tilde{\mu}_1^2) = \tilde{\beta}_1 + 0 \Rightarrow \tilde{\mu}_0 = \tilde{\mu}_1 = \left(e^{\tilde{\beta}_1} + \frac{1}{4}\right)^{1/2} - \frac{1}{2}$

$$\Rightarrow \partial_\beta l_n(\tilde{\beta}) = \begin{bmatrix} 0 \\ \sum_{j\in B} \dfrac{1+\tilde{\mu}_1}{2\tilde{\mu}_1 + 4\tilde{\mu}_1^2}(y_j^2 - e^{\tilde{\beta}_1}) \end{bmatrix}$$

$$I_n(\beta) = E\left[-\frac{\partial^2 l_n}{\partial\beta\,\partial\beta^T}\right] = \sum_i \frac{1}{2}\frac{1+2\mu_i}{\mu_i^2}\left(\frac{\partial\mu_i}{\partial\beta}\right)^{\otimes 2}$$

Then,

$$I_n(\tilde{\beta}) = \sum_{i\in A} \frac{1}{2}\frac{1+2\tilde{\mu}_0}{\tilde{\mu}_0^2}\left(\frac{\tilde{\mu}_0+\tilde{\mu}_0^2}{1+2\tilde{\mu}_0}\right)^2 x_i^{\otimes 2} + \sum_{j\in B} \frac{1}{2}\frac{1+2\tilde{\mu}_1}{\tilde{\mu}_1^2}\left(\frac{\tilde{\mu}_1+\tilde{\mu}_1^2}{1+2\tilde{\mu}_1}\right)^2 x_j^{\otimes 2}$$

$$= \sum_{i\in A} \underbrace{\frac{(1+\tilde{\mu}_0)^2}{2+4\tilde{\mu}_0}}_{\overset{\text{let}}{=}\alpha} \begin{bmatrix}1 & 0\\0 & 0\end{bmatrix} + \sum_{j\in B} \frac{(1+\tilde{\mu}_1)^2}{2+4\tilde{\mu}_1}\begin{bmatrix}1 & 1\\1 & 1\end{bmatrix} = \begin{bmatrix}(n_0+n_1)\alpha & n_1\alpha \\ n_1\alpha & n_1\alpha\end{bmatrix}$$

$$I_n(\tilde{\beta})^{-1} = \begin{bmatrix} n\alpha & n_1\alpha \\ n_1\alpha & n_1\alpha \end{bmatrix}^{-1} = \frac{1}{\alpha} \frac{1}{n \cdot n_1 - n_1^2} \begin{bmatrix} n_1 & -n_1 \\ -n_1 & n \end{bmatrix} = \frac{1}{\alpha} \frac{1}{n_0 n_1} \begin{bmatrix} n_1 & -n_1 \\ -n_1 & n \end{bmatrix}$$

$\therefore$ Score test for $H_0: \beta_2 = 0$ vs $H_1: \beta_2 \neq 0$ is

$$SC_n = (\partial_\beta l_n)^T [I_n(\beta)]^{-1} (\partial_\beta l_n) \Big|_{\beta = \tilde{\beta}} \xrightarrow{d} \chi^2_1$$

$$= \begin{bmatrix} 0 \\ \sum_{j \in B} \frac{1+\tilde{\mu}_1}{2\tilde{\mu}_1 + 4\tilde{\mu}_1^2} (y_j^2 - e^{\tilde{\beta}_1}) \end{bmatrix}^T \frac{1}{\alpha n_0 n_1} \begin{bmatrix} n_1 & -n_1 \\ -n_1 & n \end{bmatrix} \begin{bmatrix} 0 \\ \sum_{j \in B} \frac{1+\tilde{\mu}_1}{2\tilde{\mu}_1 + 4\tilde{\mu}_1^2} (y_j^2 - e^{\tilde{\beta}_1}) \end{bmatrix}$$

$$= \frac{n}{\alpha n_0 n_1} \left(\sum_{j \in B} \frac{1+\tilde{\mu}_1}{2\tilde{\mu}_1 + 4\tilde{\mu}_1^2} (y_j^2 - e^{\tilde{\beta}_1}) \right)^2$$

$$= \frac{2+4\tilde{\mu}_0}{(1+\tilde{\mu}_0)^2} \frac{n}{n_0 n_1} \left(\frac{1+\tilde{\mu}_1}{2\tilde{\mu}_1 + 4\tilde{\mu}_1^2} \right)^2 \Big(\underbrace{\sum_{j \in B} (y_j^2 - e^{\tilde{\beta}_1})}_{= n_1 (\overline{y_B^2} - \overline{y^2})} \Big)^2$$

Note that $\tilde{\mu}_0 = \tilde{\mu}_1 = \left(e^{\tilde{\beta}_1} + \frac{1}{4}\right)^{1/2} - \frac{1}{2}$

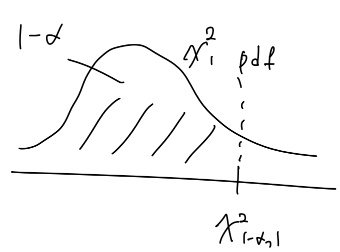
$= \left(\overline{y^2} + \frac{1}{4}\right)^{1/2} - \frac{1}{2}$

$$= \frac{n}{n_0} \cdot \frac{1}{2\tilde{\mu}_1^2 + 4\tilde{\mu}_1^3} n_1 \left(\overline{y_B^2} - \overline{y^2} \right)^2$$

$$= \frac{n}{n_0} \cdot \frac{1}{2\left[\left(\overline{y^2} + \frac{1}{4}\right)^{1/2} - \frac{1}{2}\right]^2 + 4\left[\left(\overline{y^2} + \frac{1}{4}\right)^{1/2} - \frac{1}{2}\right]^3} \cdot n_1 \left(\overline{y_B^2} - \overline{y^2} \right)^2$$

$$\xrightarrow{d} \chi^2_1 \,_{//}$$

Thus, reject H_0 iff $SC_n > \chi^2_{1-\alpha, 1}$



$$\sqrt{n_0}\left(\frac{1}{n_0}\sum_i y_i^2 - (\mu+\mu^2)\right) \to_d N(0, f(\mu)) \qquad \left(\begin{array}{l} \mathbb{E}y_i^2 = \mu+\mu^2 \\ \mathrm{Var}\, y_i^2 = f(\mu) \end{array}\right)$$

$$\sqrt{n_1}\left(\frac{1}{n_1}\sum_j y_j^2 - (\mu+\mu^2)\right) \to_d N(0, f(\mu))$$

$$\sqrt{n_1}\left(\frac{1}{n_1}\sum_j y_j^2 - \frac{1}{n_0}\sum_i y_i^2\right) \to_d N(\quad , \quad)$$