Theory Section II Exam 2015

1). $Y \sim N(\mu, \Sigma)$

Σ symmetric & full rank.

A = symmetric matrix

(a) Show that the quadratic form $Y^T A Y$ can be represented as

$Y^T A Y = \sum_{i=1}^{k} \lambda_i W_i$

W_i = indep. noncentral Chi-square

w/ df = d_i & δ_i = noncentrality param.

$\sim \chi^2_{d_i}(\delta_i)$ $i = 1, \dots, k$

Indicate what λ_i, d_i, & δ_i are equal to

~~Since A is symmetric, use spectral decomposition to decompose it to $P'\Lambda P$~~

~~P = eigenvectors~~
~~Λ = diag (eigenvalues)~~ of A

Σ = positive semidefinite

$Y'AY = Y'\Sigma^{-1/2}\Sigma^{1/2}A\Sigma^{1/2}\Sigma^{-1/2}Y$

If Σ positive semidefinite

$\Rightarrow \Sigma^{1/2}$ positive semidefinite

$\Rightarrow \Sigma^{1/2}A\Sigma^{1/2}$ positive semidefinite

At the very least, Symmetric, which is all that is necessary for spectral decomp.

let $Q = \Sigma^{1/2}A\Sigma^{1/2}$

By spectral decomposition, can write

$Q = P'\Lambda P$

P = eigenvector of $\Sigma^{1/2}A\Sigma^{1/2}$

Λ = diag(eigenvalues of $\Sigma^{1/2}A\Sigma^{1/2}$)

$Y'\Sigma^{-1/2} Q \Sigma^{-1/2} Y$

Let $B = \Sigma^{-1/2} Y$

$B = \Sigma^{-1/2} Y \sim N(\Sigma^{-1/2}\mu, \Sigma^{-1/2}\Sigma\Sigma^{-1/2})$

$\equiv N(\Sigma^{-1/2}\mu, I)$

$B'QB$

$= B'P'\Lambda PB$ by spectral decomp

Let $Z = PB$

$Z = PB \sim N(P\Sigma^{-1/2}\mu, P'IP')$

Note: Since P = matrix of eigenvectors

$\Rightarrow P$ is orthogonal matrix (orthonormal columns)

$\Rightarrow P'P = I$

$\Rightarrow Z \sim N(P\Sigma^{-1/2}\mu, I)$

Therefore, $Z'\Lambda Z$

$= \sum_{i=1}^{k} z_i^2(\lambda_i)$ λ_i = eigenvalue of $\Sigma^{-1/2} A \Sigma^{-1/2}$

$z_i^2 = w_i$

Since each $z_i \sim N([P\Sigma^{-1/2}\mu]_i, 1)$

i^{th} value of $P\Sigma^{-1/2}\mu$ matrix

w_i

$\Rightarrow z_i^2 \sim \chi^2(df=1, \delta_i)$

- noncentral

where $\delta_i = \dfrac{[P\Sigma^{-1/2}\mu]_i^2}{2}$

each z_i indep since Z vector has cov matrix of I

$\Rightarrow z_i$ have covariance of 0 for $\mathrm{Cov}(z_i, z_i')$

$\Rightarrow z_i$ indep (thm: normal rv's indep if cov = 0)

Since $Y'AY = Z'\Lambda Z$

$\Rightarrow Y'AY = \sum_{i=1}^{k} \lambda_i Z_i^2 \equiv \sum_{i=1}^{k} \lambda_i W_i$

where $W_i \sim \chi^2(d_i = 1, \delta_i)$

+ W_i's indep for $i = 1, \ldots, n$.

+ λ_i = eigenvalue of $\Sigma^{1/2} A \Sigma^{1/2}$

(b) Use part (a) to derive the MGF of $Y'AY$

Let $M(t)$ = MGF

Show $M(t)$ exists in small neighborhood of $t=0$

Find maximal area of $|t| < \underline{t_0}$.

Find MGF of $Z_i^2 \sim \chi^2(1, \delta_i)$

$\Rightarrow M(t) = \prod_{i=1}^{k} M_{Z_i}(t)$

$\chi^2(1, 0) \equiv \text{Gamma}(1/2, 2)$

Let $X \sim \chi^2(1)$

$$f_X(x) = \frac{1}{\Gamma(1/2)\, 2^{1/2}} x^{-1/2} \exp(-x/2)$$

$$E[e^{tX}] = \int_0^\infty e^{tx} \frac{1}{\sqrt{\pi}\cdot 2^{1/2}} x^{-1/2} \exp(-x/2)\, dx$$

$$= \int_0^\infty \frac{1}{\Gamma(1/2)\ldots^{1/2}} x^{-1/2} \exp(-x(1/2 - t))\, dx$$

$$= \frac{(1/2 - t)^{1/2}}{2^{1/2}} \int_0^\infty \frac{1}{\Gamma(1/2)(1/2 - t)^{1/2}} x^{-1/2} \exp(-x(1/2 - t))\, dx$$

$$= (1/2)^{1/2} (1/2 - t)^{1/2}$$

$$= \left(\frac{1}{2}\left(\frac{1}{2} - t\right)\right)^{1/2}$$

$$= \left(\frac{1}{2^2}\left(1 - t/2\right)\right)^{1/2} = \frac{(1 - t/2)^{1/2}}{2}$$

$\Rightarrow Z = X + r \quad , \quad X \sim N(0,1)$

Find dist of $V = Z^2 = (X+r)^2$

$\Rightarrow V^{1/2} = X + r$

$\Rightarrow X = V^{1/2} - r$

$f_V(v) = f_X(x = v^{1/2} - r)\, |J|$

$|J| = |\partial x / \partial v| = \frac{1}{2} |v|^{-1/2}$

$$= \frac{1}{\sqrt{2\pi}} \left(\frac{1}{2}\right) |v|^{-1/2} \exp\left(-\frac{(|v|^{1/2} - r)^2}{2}\right)$$

$$E[e^{tv}] = \int_0^\infty \frac{(2)}{2} \frac{1}{\sqrt{2\pi}} v^{-1/2} \exp\left(\frac{-(v - 2rv^{1/2} + r^2)}{2}\right) e^{tv}$$

$$= \int_0^\infty \frac{1}{\sqrt{2\pi}} v^{-1/2} \exp\left(-\frac{1}{2}(v - 2tv - 2rv^{1/2} + r^2)\right)$$

$$= \int_0^\infty \frac{1}{\Gamma(1/2) 2^{1/2}} v^{-1/2} \exp\left(-\frac{1}{2}(1-2t)\left(v - \frac{2rv^{1/2}}{(1-2t)} + \frac{r^2}{(1-2t)}\right)\right)$$

$(v^{1/2} + c)^2 = v + 2cv^{1/2} + c^2$

$c = \frac{r}{(1-2t)}$

$$= \exp\left(-\frac{1}{2}(1-2t)\frac{r^2}{(1-2t)} - \left(-\frac{1}{2}\right)(1-2t)\frac{r^2}{(1-2t)^2}\right)$$

$$\int_0^\infty \frac{1}{\Gamma(1/2) \cdot 2^{1/2}} v^{-1/2} \exp\left(-\frac{1}{2}(1-2t)\left(v - \frac{2rv^{1/2}}{(1-2t)} + \frac{r^2}{(1-2t)^2}\right)\right) dv$$

$$= \exp\left(-\cancel{\frac{1}{2}}^{k} r^2\right) \int_0^\infty \frac{1}{\Gamma(1/2) 2^{1/2}} v^{-1/2} \exp\left(-\frac{1}{2}(1-2t)\left(v^{1/2} - \frac{r}{1-2t}\right)^2\right) dv$$

let $y = v^{1/2} - \frac{r}{1-2t}$

$\Rightarrow v^{1/2} = y + \frac{r}{(1-2t)}$

$$y = v^{1/2} - \frac{r}{1-2t}$$

$$dy = \frac{1}{2} v^{-1/2} dv$$

$$\Rightarrow 2v^{1/2} dy = dv$$

$$= \exp\left(-\frac{1}{2}\cancel{r^2}\right) \int_0^\infty \frac{1}{\Gamma(1/2) 2^{1/2}} \cancel{(2v^{1/2}) v^{-1/2}} \exp\left(-\frac{1}{2}(1-2t) y^2\right) dy$$

$$= \sqrt{\frac{2}{\pi}} \exp\left(-\delta + \frac{\delta}{1-2t}\right) \int_0^\infty \exp\left(-\frac{(1-2t)}{2} y^2\right) dy$$

$$= \exp\left(-\delta\left(1 - \frac{1}{1-2t}\right)\right) \frac{2^{1/2}}{(1-2t)^{1/2}} \int_0^\infty \frac{1}{\sqrt{2\pi}} \frac{1}{(1-2t)^{1/2}} \exp\left(-\frac{1}{2}(1-2t) y^2\right) dy$$

$$= \frac{1}{2^{1/2}(1-2t)^{1/2}} \exp\left(-\delta\left(\frac{-2t}{1-2t}\right)\right)$$

$$= \frac{1}{2^{1/2}(1-2t)^{1/2}} \exp\left(\frac{2\delta t}{1-2t}\right)$$

not right

$$\Rightarrow M_v(t) = \frac{1}{(1-2t)^{1/2}} \exp\left(\frac{2\delta t}{1-2t}\right)$$

$$M(t) = \prod_{i=1}^{k} M_{v_i}(t) = \frac{1}{(1-2t)^{+k/2}} \exp\left(\frac{2t \sum_{i=1}^{k} \delta_i}{1-2t}\right)$$

This exists for $1 - 2t > 0$

$\Rightarrow 1 > 2t$

$\Rightarrow t < 1/2$

$\Rightarrow \boxed{t_0 = 1/2}$

(c) Use part (a) to show that if

$tr((A\Sigma)^2) = tr(A\Sigma) = r$

$r = rank(A)$

$Y'AY$ = chisquare dist (find df & noncentrality param)

$$Y'AY = Y^T A^{1/2} A^{1/2} Y$$

$\llcorner$ by spectral decomp (?)

$$\Rightarrow A^{1/2} Y \sim N(\mu, A^{1/2} \Sigma A^{1/2})$$

$$tr((A\Sigma)^2) = tr(A\Sigma A\Sigma)$$

by cyclic permutation property

$$\Rightarrow tr(A\Sigma A\Sigma) = tr(\Sigma A \Sigma A)$$

Consequently

$$tr((A\Sigma)^2) = tr((\Sigma A)^2) = tr(A\Sigma) = tr(\Sigma A)$$

$Y'AY$ has a non-central χ^2 dist if

$Y'AY = \Sigma_i W_i$ where $W_i \sim \chi^2(1, \delta_i)$

$\Rightarrow$ Show all $\lambda_i = 0$ or 1

(i) Show eigenvalues of $A\Sigma$ = eigenvalues of $\Sigma^{1/2} A \Sigma^{1/2}$

(ii) Show if $tr((A\Sigma)^2) = tr(A\Sigma) = r$

$\Rightarrow \lambda_i = \{0, 1\}$ for all $i = 1, \dots, K$

(i) Def^n of eigenvalue: $A\underset{\sim}{x} = \lambda \underset{\sim}{x}$ for all non-zero $\underset{\sim}{x}$

$\Leftrightarrow |A - \lambda I| = 0$

$$|\Sigma^{1/2} A \Sigma^{1/2} - \lambda I| = 0$$

$$\Rightarrow |\Sigma^{1/2} A \Sigma^{1/2} - \lambda \Sigma^{1/2} \Sigma^{-1/2}| = 0$$

$$\Rightarrow |\Sigma^{1/2}| \, |A \Sigma^{1/2} - \lambda \Sigma^{-1/2}| = 0$$

$$\Rightarrow |A \Sigma^{1/2} - \lambda \Sigma^{-1/2}| = 0$$

$$\Rightarrow |A \Sigma^{1/2} - \lambda \Sigma^{-1/2}| \, |\Sigma^{1/2}| = 0$$

$\Rightarrow |A\Sigma - \lambda I| = 0$

$\Rightarrow$ if λ = eigenvalue of $\Sigma^{1/2} A \Sigma^{1/2}$

$\Rightarrow \lambda$ eigenvalue of $A\Sigma$ as well ✓

(ii) Recall:

$tr(B) = \sum_{i=1}^{m}$ eigenvalues of B

$\Rightarrow tr(A\Sigma) = \sum_{i=1}^{m} \lambda_i$

$\Rightarrow tr((A\Sigma)^2) = \sum_{i=1}^{m} \lambda_i^2$

equal

$\Rightarrow \sum_{i=1}^{m} \lambda_i = \sum_{i=1}^{m} \lambda_i^2$

$\Rightarrow \lambda_i = 0, 1$ for all $i = 1, \dots, m$.

W/o loss of generality, let

$\lambda_1 = \dots = \lambda_k = 1, \quad \lambda_{k+1} = \dots = \lambda_m = 0$

$\Rightarrow \sum_{i=1}^{m} \lambda_i W_i = \sum_{i=1}^{k} \lambda_i W_i + \cancel{\sum_{i=k+1}^{m} \lambda_i W_i}^{\,0}$

$= \sum_{i=1}^{k} (1) W_i$

$= \sum_{i=1}^{k} W_i$

If $W_i \sim \chi^2(1, \delta_i)$

$\Rightarrow \sum_{i=1}^{k} W_i \sim \chi^2(k, \sum_{i=1}^{k} \delta_i)$

Note: if λ_i of $(A\Sigma)$ only 0 or 1 $\Rightarrow$

λ_i of $\Sigma^{1/2} A \Sigma^{1/2}$ only 0 or 1

(d) Show that $Y^T A Y$ has a noncentral chi square dist if & only if $A\Sigma$ is idempotent

$$Y^T A Y = Y^T \Sigma^{-1/2} \Sigma^{1/2} A \Sigma^{1/2} \Sigma^{-1/2} Y$$

$A\Sigma$ idempotent

$\Rightarrow A\Sigma A\Sigma = A\Sigma$

Again, show $\underset{\lambda}{\overset{\text{of } \Sigma^{1/2} A \Sigma^{1/2}}{\lambda}} = \{0, 1\}$ if this is the case

Have shown previously λ of $\Sigma^{1/2} A \Sigma^{1/2} = A\Sigma$

Since both A & Σ symmetric $\Rightarrow A\Sigma$ also symmetric

If $A\Sigma$ also idempotent $\Rightarrow A\Sigma$ is an orthogonal projection operator

If $A\Sigma$ is an (orthogonal) projection operator, then all $\lambda_i = 0$ or 1 for $i = 1, \dots, k$

↑ result just needs idempotent ⇒ projection operator

WLOG, let $\lambda_1 = \dots = \lambda_r = 1$, $\lambda_{r+1} = \dots = \lambda_k = 0$

(i) $A\Sigma$ idempotent $\Rightarrow Y'AY$ noncentral χ^2

$A\Sigma$ idempotent $\Rightarrow$ (see work above)

$\Rightarrow \lambda_1 = \dots = \lambda_r = 1$, $\lambda_{r+1} = \dots = \lambda_k = 0$

From (a) we know $Y'AY = \sum_{i=1}^{k} \lambda_i W_i$

$$\Rightarrow Y'AY = \sum_{i=1}^{r} (1) W_i + \sum_{i=r+1}^{k} (0) W_i$$
$$= \sum_{i=1}^{r} (W_i \sim \chi^2(1, \delta_i))$$
$$\sim \chi^2(r, \textstyle\sum_{i=1}^{r} \delta_i) \checkmark$$

(ii) $Y'AY$ noncentral $\chi^2 \Rightarrow A\Sigma$ idempotent

$Y'AY$ noncentral χ^2

$\Rightarrow$ all $\lambda_i = 0, 1$

$\Rightarrow \sum_{i=1}^{r} W_i = Y'AY$

WLOG, $\lambda_1 = \cdots = \lambda_r = 1, \quad \lambda_{r+1} = \cdots = \lambda_k = 0$

Have shown λ of $\Sigma^{1/2} A \Sigma^{1/2} = \lambda$ of $A\Sigma$

$(A\Sigma)\underset{\sim}{x} = \lambda \underset{\sim}{x} \quad \forall \underset{\sim}{x} \in \mathbb{R}^k$

since $\lambda = \{0, 1\}$

$\Rightarrow (A\Sigma)\underset{\sim}{x} = \lambda^2 \underset{\sim}{x}$

We also know $(A\Sigma)^2 \underset{\sim}{x} = \lambda^2 \underset{\sim}{x}$ if $A\Sigma \underset{\sim}{x} = \lambda \underset{\sim}{x}$

$\Rightarrow (A\Sigma)^2 = A\Sigma$

$\Rightarrow A\Sigma$ idempotent

2). $Y = XB + \varepsilon$

$\varepsilon \sim N(0, \sigma^2 I)$

X full rank $n \times p$

B $p \times 1$

(B, σ^2) unknown

$H = X(X'X)^{-1}X'$

h_{ii} = i^{th} diag element of H

$\hat{\sigma}^2 = \dfrac{Y'(I-H)Y}{n-p}$ $(H = M)$

$\hat{\varepsilon} = Y - X\hat{B} = Y - HY = (I-H)Y$

$\hat{\varepsilon}_i = d_i^T(I-H)Y$

d_i = vector w/ 1 in i^{th} position, 0 elsewhere

$A_i = \dfrac{\hat{\varepsilon}_i^2}{\sigma^2(1-h_{ii})}$; $B_i = \dfrac{(n-p)\hat{\sigma}^2}{\sigma^2} - \dfrac{\hat{\varepsilon}_i^2}{\sigma^2(1-h_{ii})}$

(a) Show $B_i \sim \chi^2(n-p-1)$

$\checkmark$ $V^2 = V'V$

$B_i = \dfrac{1}{\sigma^2}\left(Y'(I-H)Y - \dfrac{(d_i^T(I-H)Y)^2}{1-h_{ii}} \right)$

$= \dfrac{1}{\sigma^2}\left(Y'(I-H)Y - \dfrac{Y'(I-H)d_i d_i^T(I-H)Y}{1-h_{ii}} \right)$

$= \dfrac{1}{\sigma^2} Y'(I-H)\left(I - \dfrac{d_i d_i^T}{1-h_{ii}} \right)(I-H)Y$

By thm, if M = orthog proj matrix

$\Rightarrow \dfrac{Y'MY}{\sigma^2} \sim \chi^2(\text{rank}(M), \delta)$

$\delta = \dfrac{E(Y)'ME(Y)}{2\sigma^2}$ $E(Y) = XB$

since $ME(Y) = (\ldots)(I-H)XB = 0$

$\Rightarrow \delta = 0$

Show $\text{rank}(M) = n-p-1$

If we can show

$$(I-H)\left(I - \frac{d_i d_i^T}{1-h_{ii}}\right)(I-H) \text{ is an orthog. proj. matrix}$$

of rank $n-p-1$, then ✓

Symmetric:

$(I-H)$ symmetric ✓

$$\left(I - \frac{d_i d_i^T}{1-h_{ii}}\right)^T = I - \frac{d_i d_i^T}{1-h_{ii}} \checkmark$$

Idempotent

$$(I-H)\left(I - \frac{d_i d_i^T}{1-h_{ii}}\right)\overbrace{(I-H)(I-H)}^{I-H}\left(I - \frac{d_i d_i^T}{1-h_{ii}}\right)(I-H)$$

$$= (I-H)\left[(I-H) - \frac{d_i d_i^T}{1-h_{ii}}(I-H)\right]\left(I - \frac{d_i d_i^T}{1-h_{ii}}\right)(I-H)$$

$$= (I-H)\left[(I-H) - \frac{d_i d_i^T}{1-h_{ii}}(I-H) - (I-H)\frac{d_i d_i^T}{1-h_{ii}} + \frac{d_i \overbrace{d_i^T(I-H)d_i}^{(1-h_{ii})} d_i^T}{(1-h_{ii})^2}\right](I-H)$$

$$= (I-H)\left[(I-H) - \frac{d_i d_i^T \cancel{(I-H)}}{\cancel{1-h_{ii}}} - \frac{(I-H)d_i d_i^T}{1-h_{ii}}\frac{(\cancel{I-H} + \cancel{d_i d_i^T})}{(1-h_{ii})}(I-H)\right]$$

$$= (I-H)\left[(I-H) - \frac{(I-H)d_i d_i^T(I-H)}{1-h_{ii}}\right]$$

$$= (I-H)\left[I - \frac{d_i d_i^T}{1-h_{ii}}\right](I-H) \checkmark$$

We don't care what it projects onto or along ✓

rank (orthog proj matrix) = tr (orthog proj matrix)

$$= \operatorname{tr}\left((I-H)\left[I - \frac{d_i d_i^T}{1-h_{ii}}\right](I-H)\right)$$

→

by cyclic permutation property

$$= tr\left((I-H)\left(I - \frac{d_i d_i^T}{1-h_{ii}}\right)\right)$$

$$= tr(I-H) - tr\left((I-H) d_i d_i^T / (1-h_{ii})\right)$$

$$= n-p - tr\left(d_i^T (I-H) d_i / 1-h_{ii}\right)$$

$$= n-p - tr\left(1-h_{ii} / 1-h_{ii}\right)$$

$$= n-p-1 \checkmark$$

$$\Rightarrow \quad B_i \sim \chi^2(n-p-1) \checkmark$$

(b) Show A_i & B_i indep.

$$A_i = Y'AY$$
$$B_i = Y'BY$$

$Y'AY$ & $Y'BY$ indep iff AY & BY indep

$\Rightarrow$ if $AB = 0$

$$A_i = \frac{\hat{\varepsilon}_i^2}{\sigma^2(1-h_{ii})} = Y'(I-H)d_id_i^T(I-H)Y$$

$$B_i = Y'\left[(I-H)\left(I - \frac{d_id_i^T}{1-h_{ii}}\right)(I-H)\right]Y$$

$$AB = (I-H)d_id_i^T\overbrace{(I-H)(I-H)}^{I-H}\left(I - \frac{d_id_i^T}{1-h_{ii}}\right)(I-H)$$

$$= (I-H)d_id_i^T(I-H) - (I-H)d_i\overbrace{d_i^T(I-H)d_i}^{\cancel{1-h_{ii}}}\left(\frac{d_i^T}{\cancel{1-h_{ii}}}\right)(I-H)$$

$$= (I-H)d_id_i^T(I-H) - (I-H)d_id_i^T(I-H)$$

$$= 0 \checkmark$$

(c) $r_i = \dfrac{\hat{\varepsilon}_i}{\hat{\sigma}\sqrt{1-h_{ii}}}$

Find exact dist of $\dfrac{r_i^2}{n-p}$ using (a) & (b)

$$r_i^2 = \frac{\hat{\varepsilon}_i^2}{\hat{\sigma}^2\left(\sqrt{1-h_{ii}}\right)^2}$$

$$\frac{r_i^2}{n-p} = \left(\frac{\hat{\varepsilon}_i^2}{\sigma^2(1-h_{ii})}\right)\left(\frac{\sigma^2}{(n-p)\hat{\sigma}^2}\right)$$

$$= \frac{A_i}{B_i + A_i}$$

$$B_i \sim \chi^2(n-p-1)$$

$$A_i = Y'\overbrace{(I-H)\frac{d_i d_i^T}{1-h_{ii}}(I-H)}^{M}Y$$

Show M = orthog proj matrix
- symmetric & idempotent

symmetric ✓

$$(I-H)\left(\frac{d_i d_i^T}{1-h_{ii}}\right)\overbrace{(I-H)(I-H)\ \frac{d_i d_i^T}{\cancel{1-h_{ii}}}}^{\cancel{1-h_{ii}}\ \ I-H}(I-H)$$

$$= (I-H)\frac{d_i d_i^T}{1-h_{ii}}(I-H) \checkmark$$

$$A_i \sim \chi^2(\text{rank}(M), \delta) \Rightarrow \chi^2(\text{rank}(M))$$

$$\delta = \frac{E(Y)'(M)E(Y)}{2\sigma^2}$$

$E(Y) = XB$

since $(I-H)XB = 0 \Rightarrow \delta = 0$

rank(M) = tr(M) since orthog proj. matrix

tr(M) due to cyclic permutation properties of tr

$= tr((I-H) d_i d_i^T / 1-h_{ii})$

$= tr(d_i^T (I-H) d_i / 1-h_{ii})$

$= tr(1-h_{ii} / 1-h_{ii})$

$= 1$

We know that

$$X = \frac{\chi^2(p)}{\chi^2(p) + \chi^2(q)} \sim Beta(p/2, q/2)$$

$p = 1, \; q = n-p-1$

$$\Rightarrow \frac{r_i^2}{n-p} \sim Beta(1/2, n-p-1/2) \checkmark$$

(d) i^{th} case = outlier

Mean shift outlier model:

$$Y = XB + d_i\phi + \varepsilon$$

$$\varepsilon \sim N(0, \sigma^2 I)$$

ϕ unknown scalar, d_i $n\times 1$

$d_i = 1$ in i^{th} position, 0 elsewhere

(i) Derive the MLE of ϕ

$$L(B,\phi) \propto \exp\left(-\frac{1}{2\sigma^2}(Y - XB - d_i\phi)'(Y - XB - d_i\phi)\right)$$

$$\ell(B,\phi) = \log L(B,\phi)$$

$$= -\frac{1}{2\sigma^2}(Y - XB - d_i\phi)'(Y - XB - d_i\phi) + C$$

C = constant not including B or ϕ

$$= \frac{-1}{2\sigma^2}\left[(Y - XB)'(Y - XB) - 2(Y - XB)'d_i\phi + d_i'd_i\phi^2\right]$$

$$\frac{\partial}{\partial\phi}(\text{above}) \overset{set}{=} 0$$

$$\Rightarrow -2(Y - XB)'d_i + d_i'd_i(2\phi) = 0$$

$$\Rightarrow d_i'd_i\phi = (Y - XB)'d_i$$

$$\Rightarrow \phi = \frac{(Y - XB)'d_i}{d_i'd_i} \quad (d_i^T d_i = 1)$$

$$= (Y - XB)'d_i$$

MLE of $XB = HY$

$$\Rightarrow \hat{\phi} = (I - H)Y'd_i$$

(ii) Suppose we wish to test $H_0: \phi = 0$

- Derive the test statistic for this hypothesis & derive its exact dist under H_0.

Wald test & statistic

$$I(\phi) = E[-\partial^2/\partial\phi^2 \, l(B,\phi)]$$
$$= E\left[-\left(\frac{-2}{2\sigma^2} d_i'd_i\right)\right]$$
$$= E[1/\sigma^2]$$
$$= \frac{1}{\sigma^2}$$

MLE of σ^2:

$$\frac{\partial}{\partial\sigma^2} l(B,\phi) = \frac{1}{2\sigma^4}(Y - XB - d_i\phi)'(Y - XB - d_i\phi) + \frac{\partial}{\partial\sigma^2}\overbrace{\log((\sigma^2)^{1/2})}^{-\frac{1}{2}\log\sigma^2}$$

$$= \frac{1}{\cancel{2}\sigma^4}\|Y - XB - d_i\phi\|^2 - \frac{1}{\cancel{2}}\frac{1}{\sigma^2} \overset{\text{set}}{=} 0$$

$$\Rightarrow \|Y - XB - d_i\phi\|^2 - \sigma^2 = 0$$
$$\Rightarrow \sigma^2 = \|Y - XB - d_i\phi\|^2$$

MLE of $XB = MY$

MLE of $\phi = \hat{\phi}$

$$\hat{\sigma}^2 = \|(I-H)Y - d_i\hat{\phi}\|^2$$

Wald statistic

$$= (\hat{\phi} - \phi_0)^T I_n(\phi)(\hat{\phi} - \phi_0)$$
$$= d_i^T Y(I-H)\left(\frac{n}{\|(I-H)Y + d_i\hat{\phi}\|^2}\right)(I-H)Y'd_i$$

$$\left(\frac{n}{\hat{\sigma}^2}\right)$$

Under $H_0 \Rightarrow \sim \chi^2(1)$

(e) $I = \{1, \dots, m\}$

- subset of the first m cases in the dataset

D_I = cooks Distance based on simultaneously deleting m cases from the dataset

$$D_I = \frac{(\hat{\beta} - \hat{\beta}_I)'(X'X)(\hat{\beta} - \beta_I)}{p\hat{\sigma}^2}$$

Show $D_I = \frac{1}{p} \sum_{i=1}^{m} h_i^2 \left(\frac{\lambda_i}{1-\lambda_i}\right)$

λ_i = eigenvalues of $P_I = X_I (X^T X)^{-1} X_I^T$

- based on spectral decomp of P_I

$h_i^2 = \frac{(\gamma_i^T \hat{\varepsilon}_I)^2}{\hat{\sigma}^2 (1-\lambda_i)}$ γ_i = eigenvector corresponding w/ λ_i

$\hat{\varepsilon}_I = y_I - X_I \hat{\beta}$

$$\sum_{i=1}^{m} h_i^2 \left(\frac{\lambda_i}{1-\lambda_i}\right) = \sum_{i=1}^{m} \frac{(\gamma_i^T \hat{\varepsilon}_I)^2}{\hat{\sigma}^2 (1-\lambda_i)} \left(\frac{\lambda_i}{1-\lambda_i}\right)$$

$$= \sum_{i=1}^{m} \left(\frac{\lambda_i}{(1-\lambda_i)^2}\right) \frac{(\gamma_i^T (y_I - X_I \hat{\beta}))^2}{\hat{\sigma}^2}$$

(*) Show $\sum_{i=1}^{m} \left(\frac{\lambda_i}{(1-\lambda_i)^2}\right) (\gamma_i^T (\overbrace{y_I - X_I \hat{\beta}}^{\varepsilon_I}))^2 = (\hat{\beta} - \hat{\beta}_I)'(X'X)(\hat{\beta} - \hat{\beta}_I)$

$(\hat{\beta} - \hat{\beta}_I)'(X'X)(\hat{\beta} - \hat{\beta}_I)$

$= [X(\hat{\beta} - \hat{\beta}_I)]'[X(\hat{\beta} - \hat{\beta}_I)]$

Note: X full rank p (in original description)

$\hat{\beta} = (X'X)^{-1} X' Y$

$\hat{\beta}_I = (X_I' X)^{-1} X_I' Y$?

We know for a single deletion case,

$$\hat{\beta}_{(i)} = \hat{\beta} - \frac{(X'X)^{-1}x_i'\hat{\varepsilon}_i}{1-h_{ii}}$$

$$\Rightarrow \hat{\beta}_I = \hat{\beta} - (X'X)^{-1}X_I'[(I-H)_I]^{-1}\hat{\varepsilon}_I$$

$d_I = n \times m$ Like before, let d_I = matrix of 1's in the I diag positions + 0 elsewhere (partial identity matrix)

$$= \hat{\beta} - (X'X)^{-1}(d_I'X)'[d_I'(I-H)d_I]^{-1}(d_I\hat{\varepsilon})$$

$$= \hat{\beta} - (X'X)^{-1}(d_I'X)'[d_I'd_I - d_I'Hd_I]^{-1}\underbrace{(d_I\hat{\varepsilon})}_{\hat{\varepsilon}_I}$$

Note:

$$d_I'Hd_I = d_I'X(X'X)^{-1}X'd_I$$

$$= X_I(X'X)^{-1}X_I'$$

$$= P_I$$

$$= \hat{\beta} - (X'X)^{-1}(X_I)'[\underbrace{d_I'd_I}_{I_m \text{ (identity matrix } m\times m)} - P_I]^{-1}\hat{\varepsilon}_I$$

$$(\hat{\beta} - \hat{\beta}_I) = (X'X)^{-1}X_I'(I-P_I)^{-1}\hat{\varepsilon}_I$$

$$(\hat{\beta} - \hat{\beta}_I)'(X'X)(\hat{\beta} - \hat{\beta}_I)$$

$$= \hat{\varepsilon}_I'(I-P_I)^{-1}\underbrace{X_I(X'X)^{-1}\cancel{(X'X)}\cancel{(X'X)^{-1}}X_I'}_{P_I}(I-P_I)^{-1}\hat{\varepsilon}_I$$

$$= \hat{\varepsilon}_I'(I-P_I)^{-1}P_I(I-P_I)^{-1}\hat{\varepsilon}_I$$

Since we want to transition this into terms of eigenvectors & eigenvalues of P_I, perform spectral decomposition of P_I

$P_I = B'\Lambda B$

B = matix of eigenvectors (matrix of γ_i as rows)

$\Lambda = \text{diag}(\text{eigenvalues} = \lambda_i)$

Note: since B = orthogonal matrix (columns orthonormal), then $B'B = BB' = I$

$\longrightarrow$

$\Rightarrow I = B'B \quad + \quad P_I = B'\Lambda B$

$\Rightarrow I - P_I = B'B - B'\Lambda B = B'(I-\Lambda)B$

$[I - P_I]^{-1} = [B'(I-\Lambda)B]^{-1}$

since $B'B = BB' = I$

$\Rightarrow B^{-1} = B'$ and $(B')^{-1} = B$

$= B^{-1}(I-\Lambda)^{-1}(B')^{-1}$

$= B'(I-\Lambda)^{-1}B$

By rules of spectral decomp,
If $A = P'\Lambda P \Rightarrow A^{-1} = P'\Lambda^{-1}P$
Consequently, if
$I - P_I = B'(I-\Lambda)B$
$\Rightarrow [I - P_I]^{-1} = B'(I-\Lambda)^{-1}B$

$\hat{\varepsilon}_I'(I-P_I)^{-1}P_I(I-P_I)^{-1}\hat{\varepsilon}_I$

$= \hat{\varepsilon}_I' B'(I-\Lambda)^{-1}\underbrace{B(B'}_{I}\Lambda\underbrace{B)B'}_{I}(I-\Lambda)^{-1}B\hat{\varepsilon}_I$

$= \hat{\varepsilon}_I' B'(I-\Lambda)^{-1}\Lambda(I-\Lambda)^{-1}B\hat{\varepsilon}_I$

$(I-\Lambda)^{-1} = \text{diag}(1-\lambda_i)$

$\Lambda = \text{diag}(\lambda_i)$

$\Rightarrow (I-\Lambda)^{-1}\Lambda(I-\Lambda)^{-1} = \text{diag}\left(\frac{\lambda_i}{(1-\lambda_i)^2}\right)$

$\Rightarrow \hat{\varepsilon}_I' B' \,\text{diag}\left(\frac{\lambda_i}{(1-\lambda_i)^2}\right) B\hat{\varepsilon}_I$

$= \left(\text{diag}\left(\frac{\lambda_i^{1/2}}{(1-\lambda_i)}\right)B\hat{\varepsilon}_I\right)'\left(\text{diag}\left(\frac{\lambda_i}{1-\lambda_i}\right)B\hat{\varepsilon}_I\right)$

$= \sum_{i=1}^{m}\left(\frac{\lambda_i^{1/2}}{1-\lambda_i}\right)^2(\gamma_i'\hat{\varepsilon}_I)^2$ since rows of $B = \gamma_i$

$= \sum_{i=1}^{m}(\gamma_i'\hat{\varepsilon}_I)^2\left(\frac{\lambda_i}{(1-\lambda_i)^2}\right)$ ✓

Note: Can write $A = P\Lambda P'$ or $A = P'\Lambda P$

in $P\Lambda P'$ form, columns of P = eigenvectors

in $P'\Lambda P$ form, rows of P = eigenvectors

→ First matrix (P first or P' first) has columns as the eigenvectors.

3) (a) $\mu(x) = [1 \ x^T]\underset{\sim}{\beta}_0$

$E[Y|x] = \mu(x)$

$E[Y|x,D] = \tilde{\mu}(x,D)$

Show $\mu(x) = \tilde{\mu}(x,1)P(D=1|x) + \tilde{\mu}(x,0)P(D=0|x)$

and

$\tilde{\mu}(x,D) = E[Y|x,D,S=1] = \mu(x) + [D - P(D=1|x)]\gamma(x)$

$\gamma(x)$ = association between Y & D.

- Derive explicit form of $\gamma(x)$ in terms of $\tilde{\mu}(x,1)$ & $\tilde{\mu}(x,0)$

(i) $E[Y|x] = E[E[Y|x,D]|x]$

Since $D \sim$ Bern

$= E[Y|x,D=1]P(D=1|x) + E[Y|x,D=0]P(D=0|x)$

$= \tilde{\mu}(x,1)P(D=1|x) + \tilde{\mu}(x,0)P(D=0|x)$ ✓

(ii) $\tilde{\mu}(x,D) = E[Y|x,D]$

$E[E[Y|x,D,S]|x,D]$

$S=1 \Rightarrow$ subject sampled

$S=0 \Rightarrow$ subject not sampled

$= E[Y|x,D,S=1]P(S=1|x,D) + E[Y|x,D,S=0]P(S=0|x,D)$

If D is measured at all $(D=0,1)$ then individual is in study

$\Rightarrow P(S=1|x,D=\{0,1\}) = 1$

$P(S=0|x,D=\{0,1\}) = 0$

$= E[Y|x,D,S=1]$ ✓

(iii) $\mu(x) = \tilde{\mu}(x,1)P(D=1|x) + \tilde{\mu}(x,0)P(D=0|x)$

Since $D \sim$ Bernoulli,

$P(D=0|x) = 1 - P(D=1|x)$

$= \underbrace{\tilde{\mu}(x,1)}_{E[Y|x,D=1]} P(D=1|x) + \underbrace{\tilde{\mu}(x,0)}_{E[Y|x,D=0]}(1-P(D=1|x))$

$E[Y|x,D=1]P(D=1|x) + E[Y|x,D=0](1-P(D=1|x)) +$

$+ D\gamma(x) - P(D=1|x)\gamma(x)$