Ex. 1.

2.1. $E\left(E\left(E(y \mid x_1, x_2, x_3) \mid x_1, x_2\right) \mid x_1\right)$

$E\left(E(y \mid x_1, x_2) \mid x_1\right) = E(y \mid x_1).$

2.4.

$$Pr(y=0 \mid x=0) = \frac{Pr(x=0, y=0)}{Pr(x=0)}$$

$$= \frac{0.1}{0.1+0.4} = 0.2.$$

$$Pr(y=1 \mid x=0) = \frac{0.4}{0.5} = 0.8$$

$$Pr(y=0 \mid x=1) = \frac{0.2}{0.5} = 0.4.$$

$$Pr(y=1 \mid x=1) = \frac{0.3}{0.5} = 0.6.$$

$$E(f(y) \mid x) = \Sigma f(y) \cdot Pr(Y=y \mid X=x).$$

2.6. $Var(y) = var(m(x) + e)$

$= Var(m(x)) + var(e) + 2\underbrace{cov(m(x), e)}_{= 0 \text{ by exogeneity.}}$

$= Var(m(x)) + \sigma^2.$

2.10. True.

$y = x\beta + e, \quad E(e|x) = 0.$

$E(x^2 e) = E(E(x^2 e | x)) = E(x^2 E(e|x)) = E(x^2 \cdot 0) = 0$

2.11. False.

$\cancel{E(ex)}$ $E(xe) = 0.$

$x \sim N(0,1).$

$e = x^2 - 1.$

$E(xe) = E(x(x^2-1)) = E(x^3 - x)$

$= E(x^3) - E(x)$

$= 0.$

$E(x^2 e) = E(x^4 - x^2) \neq 0.$

Ex 2.

3.2. $\underset{n\times 1}{y} = \underset{n\times k}{X}\underset{k\times 1}{\beta} + \underset{n\times 1}{\varepsilon}$.

$\hat{\beta} = (X'X)^{-1}X'y$ $\qquad\qquad$ $\hat{\varepsilon} = y - X\hat{\beta} = [\underbrace{I - X(X'X)^{-1}X'}_{M_X.}]\, y$

$\underset{n\times 1}{y} = \underset{n\times k}{Z}\underset{k\times 1}{\delta} + \underset{n\times 1}{\eta}$

$\hat{\delta} = (Z'Z)^{-1} Z'y$

Given $Z = XC$

$$\delta = (C'X'XC)^{-1}C'X'y$$

$$= C^{-1}(X'X)^{-1}\cancel{(C')^{-1}}\,\cancel{C'}X'y \qquad \because C,\ X'X \text{ are non-singular}$$

$$= C^{-1}(X'X)^{-1}X'y$$

$$= C^{-1}\beta$$

$$\hat{\eta} = y - Z\hat{\delta} = y - XCC^{-1}\hat{\beta} = y - X\hat{\beta} = \hat{\varepsilon}.$$

3.4 $\hat{e} = y - X\hat{\beta}_{OLS} = y - X(X'X)^{-1}X'y = [I - X(X'X)^{-1}X']\,y$

$X = [\underset{n\times k_1}{X_1} \quad \underset{n\times k_2}{X_2}]_{n\times k}.$

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

$$X'\hat{e} = X'[I - X(X'X)^{-1}X']\,y = [X' - X']\,y = 0_{k\times 1}$$

$$X'\hat{e} = \begin{bmatrix} X_1' \\ X_2' \end{bmatrix}\hat{e} = \begin{bmatrix} X_1'\hat{e} \\ X_2'\hat{e} \end{bmatrix} = 0_{k\times 1}$$

$\therefore\ X_1'\hat{e} = 0_{k_1\times 1} \qquad X_2'\hat{e} = 0_{k_2\times 1}$

3.7. $X = [X_1 \quad X_2]$ $\qquad P = X(X'X)^{-1}X'$

$$M = I - X(X'X)^{-1}X' = I_{n\times n} - P$$

$$\underset{(n\times n)(n\times k)}{PX} = \underset{(n\times k)}{X} \qquad \underset{(n\times n)(n\times k)}{MX} = \underset{(n\times k)}{0}.$$

$$\underset{n\times n}{P}[\underset{n\times k_1}{X_1}, \underset{n\times k_2}{X_2}] = [X_1 \quad X_2]$$

$$[\underset{n\times k_1}{PX_1} \quad \underset{n\times k_2}{PX_2}] = [\underset{n\times k_1}{X_1} \quad \underset{n\times k_2}{X_2}]. \quad \Rightarrow \quad PX_J = X_J, \quad J = 1, 2.$$

$$MX_J = (I_{n\times n} - P)X_J = X_J - PX_J = 0. \quad J = 1, 2.$$

3.9. trace = the sum of diagonal elements.

$$tr(M) = tr(I_n - P) = tr(I_n) - tr(P)$$

$$tr(I_n) = n.$$

$$tr(P) = tr(X(X'X)^{-1}X')$$

$$= tr((X'X)^{-1}X'X) \qquad tr(AB) = tr(BA).$$

(Given AB & BA are productable)

$$= tr(I_k)$$

$$= k$$

$$\therefore tr(M) = n - k.$$

3.12. $d_{1i} = \begin{cases} 1, & \text{male} \\ 0, & \text{female} \end{cases}$ $d_{2i} = \begin{cases} 0 & \text{male} \\ 1 & \text{female} \end{cases}$

① $y = \mu + d_1\alpha_1 + d_2\alpha_2 + e$.

② $y = d_1\alpha_1 + d_2\alpha_2 + e$

③ $y = \mu + d_1\phi + e$.

① can't, $\because d_1 + d_2 = 1_{n\times 1}$

Multicolinearity ~~with~~ between $1_{n\times1}$ and $1_{n\times1}$, one of which is the original constant regressor.

$$y = \mu + (d_1 + d_2)\alpha_1 + d_2(\alpha_2 - \alpha_1) + e.$$

$$= (\mu + \alpha_1) + d_2(\alpha_2 - \alpha_1) + e.$$

(a) ② : $y = d_1(\alpha_1 - \alpha_2) + (d_1 + d_2)\alpha_2 + e$

$$= \alpha_2 + d_1(\alpha_1 - \alpha_2) + e.$$

$\alpha_2 = \mu$, $\alpha_1 - \alpha_2 = \phi \Rightarrow \alpha_1 = \phi + \mu$

(b). $\iota' = \underbrace{(1, 1, \cdots, 1)}_{n}$

$$\iota' d_1 = (1, \cdots, 1)\begin{pmatrix} 1 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}\begin{matrix} \}\, n_1 \\ \\ \}\, n_2 \end{matrix} = n_1$$

$\iota' d_2 = n_2$.

(c). $\alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}$ $\quad X = \begin{pmatrix} d_1 & d_2 \end{pmatrix}_{n\times 2}$.

$\phantom{(c). \alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \quad X = (} n\times 1 \quad n\times 1$.

$y = d_1\alpha_1 + d_2\alpha_2 + e = X\alpha + e$.

$x_i = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ for male, $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ for female.

$E(x_i e_i) = E\left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} e_i\right] = E\begin{pmatrix} e_i \\ 0 \end{pmatrix} = \begin{pmatrix} E(e_i) \\ 0 \end{pmatrix}$

With constant regressor, we can always make $E(e_i) = 0$.

$\therefore E(x_i e_i) = 0$

3.16. $y = X_1 \tilde{\beta}_1 + \tilde{e}$

$$y = X_1\hat{\beta}_1 + X_2\hat{\beta}_2 + \hat{e}$$
$$= [X_1 \quad X_2] \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} + \hat{e}.$$
$$= X\hat{\beta} + \hat{e}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} \hat{\varepsilon}_i^2}{\sum_{i=1}^{n} (y_i - \bar{y}_i)}$$

$$\hat{\varepsilon} = \begin{pmatrix} \hat{\varepsilon}_1 \\ \vdots \\ \hat{\varepsilon}_n \end{pmatrix}.$$

$$\sum_{i=1}^{n} \hat{\varepsilon}_i^2 = \hat{\varepsilon}'\hat{\varepsilon}.$$

To show $R_2^2 \geq R_1^2$, only need $\hat{e}'\hat{e} \leq \tilde{e}'\tilde{e}$

① $X_2 \neq C X_1$

$\hat{e} = (I - P)y$, $P = X(X'X)^{-1}X'$

$\tilde{e} = (I - P_1)y$, $P_1 = \underline{X_1} (X_1'X_1)^{-1} \underline{X_1'}$

$P X_1 = X_1$

$\therefore P_1 = \underline{P\, X_1} (X_1'X_1)^{-1} \underline{X_1'\, P} = P P_1 P.$

$\hat{e}'\hat{e} = y'(I-P)y$

$\tilde{e}'\tilde{e} = y'(I-P_1)y$

$\tilde{e}'\tilde{e} - \hat{e}'\hat{e} = y'(P-P_1)y.$

$P - P_1 = PP - \underline{PP_1P} = P\underbrace{(I - P_1)}_{M_1}P.$ is p.d

$\therefore \tilde{e}'\tilde{e} - \hat{e}'\hat{e} \geq 0$

when $R_2^2 = R_1^2$. ?

$X_2 = C X_1.$

$$R_2^2 \geq R_1^2.$$

3.13. (a). $y = d_1\hat{\gamma}_1 + d_2\hat{\gamma}_2 + \hat{u}$.

P93.

For men, $d_1 = 1$, $d_2 = 0$.

$$\Leftrightarrow \quad y_1 = 1 \cdot \hat{\gamma}_1 + \hat{u}_1$$

$$\hat{\gamma}_1 = \bar{y}_1$$

For women, $d_1 = 0$, $d_2 = 1$.

$$\Leftrightarrow \quad y_2 = 1 \cdot \hat{\gamma}_2 + \hat{u}_2$$

$$\hat{\gamma}_2 = \bar{y}_2$$

(b). Frisch - Waugh - Lovell theorem.

$$y^* = y - d_1\bar{y}_1 - d_2\bar{y}_2.$$

y^* is the residual of regressing y on d_1, d_2.

X^* - - - - - X - - -

(c). Define $d = (d_1, d_2)$, $M_d = I - d(d'd)^{-1}d'$

then $y^* = M_d y$ $\quad X^* = M_d X$.

$$y^* = X^*\tilde{\beta} + \tilde{e}$$

$$\tilde{\beta} = (X'M_d X)^{-1} X'M_d y.$$

$$y = d_1\hat{\alpha}_1 + d_2\hat{\alpha}_2 + X\hat{\beta} + \hat{e}$$

$$\hat{\beta} = (X'M_d X)^{-1} X'M_d y.$$

$$\tilde{\beta} = \hat{\beta}.$$

4.3. Heteroskedasticity. BLUE. P.119.

~~[illegible]~~

$$Var(e) = D = \begin{pmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2 \end{pmatrix}_{n \times n}.$$

$$\tilde{\beta} = (X'D^{-1}X)^{-1} X'D^{-1} y$$

$$E(\tilde{\beta}|X) = \beta$$

$$\tilde{\beta} - \beta = (X'D^{-1}X)^{-1} X'D^{-1}(X\beta + e)$$

$$= (X'D^{-1}X)^{-1} X'D^{-1}e.$$

$$E[\tilde{\beta} - \beta | X] = 0.$$

$$Var(\tilde{\beta} - \beta|X) = E[(\tilde{\beta}-\beta)(\tilde{\beta}-\beta)'|X].$$

$$= E\left[(X'D^{-1}X)^{-1} X'D^{-1} ee' D^{-1} X (X'D^{-1}X)^{-1} | X\right]$$

$$= (X'D^{-1}X)^{-1} X'D^{-1} \underbrace{E(ee'|X)}_{= D} D^{-1} X (X'D^{-1}X)^{-1}$$

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

For any unbiased linear estimator $\hat{\beta} = A'y$.

where $A'X = I$.

$$\operatorname{Var}(\hat{\beta} \mid X) = E[A'ee'A \mid X] = A'DA.$$

To prove $\operatorname{Var}(\tilde{\beta} \mid X) \leq \operatorname{Var}(\hat{\beta} \mid X)$.

i.e. $(X'D^{-1}X)^{-1} \leq A'DA$.

$$\cancel{X'D^{-1}X \geq (A'DA)^{-1}}$$

Write $A = C + D^{-1}X(X'D^{-1}X)^{-1}$.

$$A'X = I \Rightarrow X'C = 0$$

$$A'DA = [(X'D^{-1}X)^{-1}X'D^{-1} + C']D[C + D^{-1}X(X'D^{-1}X)^{-1}]$$

$$= (X'D^{-1}X)^{-1}X'C + (X'D^{-1}X)^{-1} + C'DC$$

$$+ C'X(X'D^{-1}X)^{-1}$$

$$= (X'D^{-1}X)^{-1} + C'DC.$$

$A'DA - (X'D^{-1}X)^{-1} = C'DC$ is p.s.d.

4.4. GLS.

$$y = X\beta + e \qquad E(e|X) = 0 \qquad Var(e|X) = \sigma^2\Omega$$

$$\tilde{\beta} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y.$$

$$\hat{e} = y - X\tilde{\beta}$$

$$s^2 = \frac{1}{n-k}\hat{e}'\Omega^{-1}\hat{e}.$$

(a). $E(\tilde{\beta}|X) = E\left((X'\Omega^{-1}X)^{-1}X'\Omega^{-1}(X\beta + e)|X\right)$

$$= E\left(\beta + (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}e \,|X\right) = \beta.$$

(b) $\tilde{\beta} - \beta = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}e.$

$$Var(\tilde{\beta} - \beta|X) = E\left[(X'\Omega^{-1}X)^{-1}X'\Omega^{-1}ee'\Omega^{-1}X(X'\Omega^{-1}X)^{-1} \,\middle|\, X\right].$$

$$= \sigma^2(X'\Omega^{-1}X)^{-1}$$

(c). $\hat{e} = y - X\tilde{\beta}$

$$= y - X(X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y$$

$$= \left[I - X(X'\Omega^{-1}X)^{-1}X'\Omega^{-1}\right][X\beta + e].$$

$$= X\beta - X\beta + M_1 e$$

$$= M_1 e.$$

(d). $M_1' \Omega^{-1} M_1$

$$= \left[I - \Omega^{-1} X (X'\Omega^{-1}X)^{-1} X'\right] \Omega^{-1} \left[I - X(X'\Omega X)^{-1} X'\Omega^{-1}\right].$$

$$= \Omega^{-1} - \Omega^{-1} X (X'\Omega^{-1}X)^{-1} X'\Omega^{-1}$$

(e). $s^2 = \frac{1}{n-k} \hat{e}' \Omega^{-1} \hat{e}$

$$= \frac{1}{n-k} (M_1 e)' \Omega^{-1} (M_1 e).$$

$$= \frac{1}{n-k} e' M_1' \Omega^{-1} M_1 e.$$

$$E(e' M_1' \Omega^{-1} M_1 e \mid X)$$

$$= E\left(\operatorname{trace}(e' M_1' \Omega^{-1} M_1 e) \mid X\right)$$

$$= E\left(\operatorname{trace}(M_1' \Omega^{-1} M_1 ee') \mid X\right)$$

$$= \operatorname{trace}\left(E(M_1' \Omega^{-1} M_1 ee' \mid X)\right)$$

$$= \operatorname{trace}\left(\sigma^2 M_1' \Omega^{-1} M_1 \Omega\right)$$

$$= \sigma^2 \operatorname{trace}\left(\left(\Omega^{-1} - \Omega^{-1} X (X'\Omega^{-1}X)^{-1} X'\Omega^{-1}\right)\Omega\right)$$

$$= \sigma^2 \operatorname{trace}\left(I_n - \Omega^{-1} X (X'\Omega^{-1}X)^{-1} X'\right).$$

$$= \sigma^2 \left[n - \operatorname{trace}\left(\Omega^{-1} X (X'\Omega^{-1}X)^{-1} X'\right)\right]$$

$$= \sigma^2 \left[n - \operatorname{trace}\left((X'\Omega^{-1}X)^{-1} X'\Omega^{-1}X\right)\right]$$

$$= \sigma^2 [n-k].$$

$$\therefore E[s^2 \mid X] = \sigma^2.$$

5.1. $\liminf_{n\to\infty} X_n = \lim_n \left(\inf_{m\geq n} X_m \right).$

P.143.

$\limsup_{n\to\infty} X_n = \lim_n \left(\sup_{m\geq n} X_m \right).$

(1) $a_n = 1/n.$ $\lim = \liminf = \limsup = 0.$

(2) $\liminf_{n\to\infty} \sin\left(\frac{\pi}{2}n\right) = \lim_n \left(\inf_{m\geq n} \sin\left(\frac{\pi}{2}m\right)\right)$

$= \lim (-1) = -1.$

$\limsup_{n\to\infty} \sin\left(\frac{\pi}{2}n\right) = 1.$

$\lim$ $\nexists$

(3). $a_n = \frac{1}{n}\sin\left(\frac{\pi}{2}n\right).$

$\limsup_{n\to\infty} a_n = \liminf_{n\to\infty} a_n = \lim_{n\to\infty} a_n = 0.$

5.2. (1). $E(\bar{y}^*)$

$= E\left(\frac{1}{n}\sum_{i=1}^{n} w_i y_i\right)$

$= \frac{1}{n}\left(w_1 E(y_1) + \cdots + w_n E(y_n)\right)$

$= \frac{1}{n}\sum_{i=1}^{n} w_i \; E(y_i)$ $\because$ y_i is i.i.d.

$= E(y_i)$

$= \mu.$

(2) $E(\bar{y}^{*2})$

$= E\left[\left(\frac{1}{n}\sum_{i=1}^{n} w_i y_i\right)^2\right]$

$= \frac{1}{n^2} E\left[\sum_{i=1}^{n} w_i^2 y_i^2 + \sum_{i=1}^{n}\sum_{j\neq i} w_i w_j y_i y_j\right]$

$= \frac{1}{n^2}\sum_{i=1}^{n} w_i^2 E(y_i^2) + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j\neq i} w_i w_j \; E(y_i y_j) = \frac{1}{n^2}\sum_{i=1}^{n} w_i^2 E(y_i^2)$

$$Var(\bar{y}^*) = E(\bar{y}^{*2}) - E(\bar{y}^*)^2$$

$$= \frac{1}{n^2}\sum_{i=1}^{n} w_i^2 E(y_i^2) - \mu^2.$$

(3). $\frac{1}{n^2}\sum_{i=1}^{n} w_i^2 \to 0.$

$Var(\bar{y}^*) < \infty.$

by LLN, $\bar{y}^* \xrightarrow{p} \mu.$

(4). $\frac{1}{n^2}\sum_{i=1}^{n} w_i^2 \leq \frac{1}{n^2}\cdot n \cdot (\max_{i\leq n} w_i^2)$

$$= \frac{1}{n}\cdot O(n)$$

$$= o(1)$$

$$\to 0.$$

Convergence in Mean Square $\Rightarrow$ Convergence in Probability.

$\lim_{n\to\infty} E((Z_n - Z)^2) = 0.$

By chebyshev's inequality.

$$Pr(|Z_n - Z| \geq \varepsilon) \leq \frac{E(Z_n - Z)^2}{\varepsilon^2}.$$

$$\lim_{n\to\infty} Pr(|Z_n - Z| \geq \varepsilon) \leq \frac{1}{\varepsilon^2}\lim_{n\to\infty} E[(Z_n - Z)^2]$$

$$= 0.$$

$\therefore$ m.s $\Rightarrow$ P.

6.1. (P175)

True Model : $y_i = x_{1i}'\beta_1 + x_{2i}'\beta_2 + e_i \qquad E(x_i e_i) = 0.$

x_{2i}' are omitted.

$$\hat{\beta}_1 = \left(\frac{1}{n}\sum x_{1i}x_{1i}'\right)^{-1}\left(\frac{1}{n}\sum x_{1i}y_i\right)$$

$$= \left(\frac{1}{n}\sum x_{1i}x_{1i}'\right)^{-1}\left(\frac{1}{n}\sum x_{1i}\left(x_{1i}'\beta_1 + x_{2i}'\beta_2 + e_i\right)\right)$$

$$= \beta_1 + \Sigma_{11}^{-1}\Sigma_{12}\beta_2 + \Sigma_{11}^{-1}\left(\frac{1}{n}\sum x_{1i}e_i\right)$$

where $\Sigma_{11} = \frac{1}{n}\sum x_{1i}x_{1i}' \qquad \Sigma_{12} = \frac{1}{n}\sum x_{1i}x_{2i}'$

If Σ_{11} is non-singular and finite.

$$\hat{\beta}_1 \xrightarrow{p} \beta_1 + Q_{11}^{-1}Q_{12}\beta_2$$

$Q_{11} = E(x_{1i}x_{1i}') \qquad Q_{12} = E(x_{1i}x_{2i}')$

Generally, $\hat{\beta}_1$ is inconsistent.

When could it be consistent ?

① $Q_{12} = 0$

$\forall\ i, \quad E(x_{1i}x_{2i}) = 0$

$\forall\ i \neq j \quad E(x_{1i}x_{2j}) = 0$

② $\beta_2 = 0.$

6.2. $\underset{n\times 1}{y} = \underset{n\times k}{X}\ \underset{k\times 1}{\beta} + \underset{n\times 1}{e}$ long solution, for shorter answer, refer to tutorial

Define $\underset{k\times k}{\Sigma} = \frac{1}{n} X'X \xrightarrow[p]{} Q = E(x_i x_i')$

$\text{rank}(X) = k \Rightarrow \Sigma$ is full rank, rank $= k$.

Σ has k eigenvalues and corresponding eigenvectors.

$\cancel{\Sigma \lambda_j = \lambda_j r_j}$

$\Sigma r_j = \lambda_j r_j \quad , \quad j = 1, \cdots, k$

$$\Sigma \underbrace{(r_1 \ r_2 \ \cdots \ r_k)}_{R} = (r_1 \ r_2 \ \cdots \ r_k) \underbrace{\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_k \end{pmatrix}}_{D}$$

$$\Sigma R = RD.$$

$$\Sigma = RDR^{-1}$$

$$\hat{\beta} = \left(\frac{1}{n}\sum_{i=1}^{n} x_i x_i' + \frac{\lambda}{n} I_k\right)^{-1} \left(\frac{1}{n}\sum_{i=1}^{n} x_i y_i\right).$$

$$= \left(RDR^{-1} + R\Lambda R^{-1}\right)^{-1} \left(RDR^{-1}\beta + \frac{1}{n}\sum_{i=1}^{n} x_i e_i\right).$$

$$= \left(R(D+\Lambda)R^{-1}\right)^{-1} \left(RDR^{-1}\beta + \frac{1}{n}\sum_{i=1}^{n} x_i e_i\right).$$

$$= (RBR^{-1})^{-1} RDR^{-1}\beta + \underbrace{(RBR)^{-1}}_{\text{finite}} \underbrace{\left(\frac{1}{n}\sum_{i=1}^{n} x_i e_i\right)}_{\xrightarrow{p}\ 0}$$

$$= RB^{-1}DR^{-1}\beta + o_p(1).$$

$$B = \begin{pmatrix} \lambda_1 + \frac{1}{n}\lambda & & \\ & \ddots & \\ & & \lambda_k + \frac{1}{n}\lambda \end{pmatrix}$$

$$B^{-1} = \begin{pmatrix} \frac{1}{\lambda_1 + \frac{1}{n}\lambda} & & \\ & \ddots & \\ & & \frac{1}{\lambda_k + \frac{1}{n}\lambda} \end{pmatrix}$$

$$B^{-1}D = \begin{pmatrix} \frac{\lambda_1}{\lambda_1 + \frac{1}{n}\lambda} & & \\ & \ddots & \\ & & \frac{\lambda_k}{\lambda_k + \frac{1}{n}\lambda} \end{pmatrix} \xrightarrow[p]{} I_K.$$

$\therefore \quad R B^{-1} D R^{-1} \xrightarrow[p]{} I_K$

$\therefore \quad \hat{\beta} \xrightarrow[p]{} \beta.$

6.7. Measurement Error.

$y_i^* = x_i'\beta + e_i \qquad y_i = y_i^* + u_i \;,\; E(y_i^* u_i) = 0$

$y_i = x_i'\beta + e_i + u_i \overset{\text{def}}{=} x_i'\beta + \eta_i$

with $E(x_i e_i) = 0 \quad E(x_i u_i) = 0$

Linear Projection coefficient of y_i on x_i

$E(x_i x_i')^{-1} E(x_i y_i)$

$= E(x_i x_i')^{-1} E(x_i x_i'\beta + x_i \eta_i)$

$= E(x_i x_i')^{-1} E(x_i x_i')\beta + E(x_i x_i')^{-1} E(x_i \eta_i)$

$= \beta.$

(b). $\hat{\beta} = \left(\frac{1}{n}\sum_{i=1}^{n} x_i x_i'\right)^{-1} \left(\frac{1}{n}\sum_{i=1}^{n} x_i y_i\right)$

$$= \left(\frac{1}{n}\sum_{i=1}^{n} x_i x_i'\right)^{-1} \left(\frac{1}{n}\sum_{i=1}^{n} x_i (x_i'\beta + \eta_i)\right)$$

$$= \beta + \underbrace{\left(\frac{1}{n}\sum_{i=1}^{n} x_i x_i'\right)^{-1}}_{\downarrow p \atop E(x_i x_i')^{-1}} \underbrace{\left(\frac{1}{n}\sum_{i=1}^{n} x_i \eta_i\right)}_{\downarrow p \atop E(x_i \eta_i) = 0}$$

$\xrightarrow{p} \beta$.

It's consistent.

(c). $\sqrt{n}(\hat{\beta} - \beta) =$

$E(x_i \eta_i) = E(x_i(e_i + u_i)) = 0$.

$Var(x_i \eta_i) = E(x_i \eta_i \eta_i x_i')$

$E(\eta_i^2 | X) = E((u_i + e_i)^2 | X)$

$= E(u_i^2 + e_i^2 + 2u_i e_i | X) \equiv \Omega$

(c). $\sqrt{n}(\hat{\beta} - \beta) = \underbrace{\left(\frac{1}{n}\sum_{i=1}^{n} x_i x_i'\right)^{-1}}_{\downarrow p \atop Q^{-1} \equiv E(x_i x_i')^{-1}} \underbrace{\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} x_i \eta_i\right)}_{\downarrow d \atop N(0,\ V)}$.

where $V = Var(x_i \eta_i) = E(x_i x_i' \eta_i^2) = E(x_i x_i' (u_i + e_i)^2)$

$\therefore \sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} N(0,\ Q^{-1} V Q^{-1})$.

6.11. (P176).

$$\widehat{\log}(w) = \beta_1 edu + \beta_2 exp + \beta_3 \frac{exp^2}{100} + \beta_4$$

$$X = (edu \quad exp \quad exp^2/100 \quad 1). \qquad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix}$$

$$\hat{\beta} - \beta \xrightarrow{d} N(0, V).$$

The estimated variance - covariance matrix of $\hat{\beta}$ is denoted as $\hat{V}_{\hat{\beta}}$.

For a certain point x_0 (4×1), estimated $\widehat{\log}(w) = x_0'\hat{\beta}$.

$$Var(x_0'\hat{\beta}) = x_0'\hat{V}_{\hat{\beta}}x_0$$

The $(1-\alpha)$ confidence interval for $\log(w)(x_0)$ is.

$$\left[\ x_0'\hat{\beta} - \Phi^{-1}(97.5\%)\sqrt{x_0'\hat{V}_{\hat{\beta}}x_0}\ ,\quad x_0'\hat{\beta} + \Phi^{-1}(97.5\%)\sqrt{x_0'\hat{V}_{\hat{\beta}}x_0}\ \right].$$

(b) - (d) refer to your notes.

P222

8.2. Two independent sample.

$$y_1 = X_1\beta_1 + e_1 , \quad E(x_{1i} e_{1i}) = 0.$$

$$y_2 = X_2\beta_2 + e_2.$$

OLS estimator $\hat{\beta}_1 = (X_1'X_1)^{-1}X_1'y_1$, $E(\hat{\beta}_1) = \beta_1$

$\hat{\beta}_2 = (X_2'X_2)^{-1}X_2'y_2$, $E(\hat{\beta}_2) = \beta_2$.

$$\sqrt{n}(\hat{\beta}_1 - \beta_1) \xrightarrow{d} N(0, V_1)$$

$$\sqrt{n}(\hat{\beta}_2 - \beta_2) \xrightarrow{d} N(0, V_2).$$

$$\text{cov}[(\hat{\beta}_1 - \beta_1), (\hat{\beta}_2 - \beta_2)]$$

$$= E[(\hat{\beta}_1 - \beta_1)(\hat{\beta}_2 - \beta_2)']$$

$$= E[(X_1'X_1)^{-1}X_1e_1, (X_2'X_2)^{-1}X_2e_2]$$

$$= 0 \quad (\because \text{independent sampling}).$$

$$\therefore \begin{pmatrix} \sqrt{n}(\hat{\beta}_1 - \beta_1) \\ \sqrt{n}(\hat{\beta}_2 - \beta_2) \end{pmatrix} \xrightarrow{d} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \underbrace{\begin{pmatrix} V_1 & 0 \\ 0 & V_2 \end{pmatrix}}_{= V}\right).$$

(a) $A = (-1, 1)$.

$$\sqrt{n}[-(\hat{\beta}_1 - \beta_1) + (\hat{\beta}_2 - \beta_2)] \xrightarrow{d} N(0, AVA').$$

b) & (c). $H_0: \beta_2 = \beta_1 \iff \beta_2 - \beta_1 = 0$.

$$\sqrt{n}\left[(\hat{\beta_2} - \hat{\beta_1}) + (\beta_1 - \beta_2)\right] = \sqrt{n}\,\underset{k\times 1}{(\hat{\beta_2} - \hat{\beta_1})} \xrightarrow{d} N(0, \underset{k\times k}{AVA'})$$

$$n(\hat{\beta_2} - \hat{\beta_1})'(AVA')^{-1}(\hat{\beta_2} - \hat{\beta_1}) \xrightarrow{d} \chi^2(2)$$

8.3. $I_i = \beta_1 Q_i + \beta_2 C_i + \beta_3 D_i + \varepsilon_i$

(c). ① $H_0: \beta_2 = 0, \ \beta_3 = 0$.

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$A\beta = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

② $H_0: \beta_1 = 0$.

$$A = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}$$

(d). $I_i = \gamma_1 Q_i + \gamma_2 C_i + \gamma_3 D_i + \gamma_4 Q_i^2 + \gamma_5 C_i^2 + \gamma_6 D_i^2$
$+ \gamma_7 Q_i C_i + \gamma_8 Q_i D_i + \gamma_9 C_i D_i$

$H_0: \gamma_4 = \gamma_5 = \gamma_6 = \gamma_7 = \gamma_8 = \gamma_9 = 0$.

A is a 6×9 matrix.

For $\forall A$.

$$T_n = n(A\hat{\beta})'[A\,\widehat{Var}(\hat{\beta})A']^{-1}(A\hat{\beta}) \sim \chi^2(r)$$

10.1. $F_n(x)$ is E.D.F.

(251).

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} 1(X_i \le x).$$

Define $Y_i := 1(X_i \le x)$.

$\{X_i\}$ is i.i.d, so is $\{Y_i\}$. and

$$E(Y_i) = E\left(1(X_i \le x)\right) = P_r(X_i \le x) = F_0(x).$$

$$Y_i = \begin{cases} 1, & \text{with } P_r \ F_0(x) \\ 0, & 1-F_0(x). \end{cases}$$

$$Var(Y_i) = (1-F_0(x))^2 F_0(x) + (0-F_0(x))^2(1-F_0(x))$$

$$= F_0(x)(1-F_0(x)).$$

i.i.d + finite mean + finite variance.

By Lindeberg - Levy CLT.

$$\sqrt{n}\left(\frac{1}{\cancel{\sqrt{n}}\, n}\sum Y_i - F_0(x)\right) \xrightarrow{d} N\left(0,\ F_0(x)(1-F_0(x))\right).$$

i.e. $\sqrt{n}\left(F_n(x) - F_0(x)\right) \xrightarrow{d} N\left(0,\ F_0(x)(1-F_0(x))\right).$

10.2. $E y_i = \mu$ $\quad Var(y_i) = \sigma^2$.

sample. $T_n = \bar{y}_n = \frac{1}{n}\sum_{i=1}^{n} y_i$

$$E(T_n) = E\left(\frac{1}{n}\sum y_i\right) = \frac{1}{n}\sum E(y_i) = \mu.$$

$$Var(T_n) = E\left((T_n-\mu)^2\right)$$

$$= E\left(\left(\frac{1}{n}\sum_{i=1}^{n}(y_i-\mu)\right)^2\right)$$

$$= \frac{1}{n^2}\sum_{i=1}^{n} E(y_i-\mu)^2$$

$$= \frac{1}{n}\sigma^2$$

bootstrap sample.

$T_n^* = \frac{1}{n}\sum_{i=1}^{n} y_i^*$.

Note that $y_i^* \in \{y_1, \cdots, y_n\}$

$\therefore E y_i^* = \mu, \quad Var(y_i^*) = \sigma^2$.

$\therefore E(T_n^*) = \mu$

$Var(T_n^*) = \frac{\sigma^2}{n}$

10.4. $T_n^* = \dfrac{\hat{\theta}^* - \hat{\theta}}{s(\hat{\theta})}$.

$q_n^*(\alpha)$ is the α-th quantile of T_n^*.

Bootstrap Confidence Interval.

$C = [\hat{\theta} - s(\hat{\theta})\, q_n^*(0.95),\ \hat{\theta} - s(\hat{\theta})\, q_n^*(0.05)]$.

Alternative percentile interval:

$B_n^* = \hat{\theta}^* - \hat{\theta}$.

b_n^* is quantile of B_n^*.

$C_1 = [\hat{\theta} - b_n^*(0.95),\ \hat{\theta} - b_n^*(0.05)]$.

$B_n^* = T_n^* \cdot s(\hat{\theta})$.

$b_n^*(\cdot) = q_n^*(\cdot) \cdot s(\hat{\theta})$.

$C = C_1$.

10.6. $\hat{\theta} = 1.2$ $\quad s(\hat{\theta}) = 0.2$.

$\hat{\theta}^*$ $\quad$ 2.5% percentile $\quad 0.75$.

$\quad$ 97.5% percentile. $\quad 1.3$.

Efron. $\quad T_n^* = \hat{\theta}^*$. $\qquad C = [0.75, \ 1.3]$.

Alternative. $\quad T_n^* = \hat{\theta}^* - \hat{\theta}$.

$$\cancel{C = [0.75 - 1.2, \ 1.3 - 1.2]}$$

$$q_n^*(0.025) = 0.75 - 1.2 = -0.45.$$

$$q_n^*(0.975) = 1.3 - 1.2 = 0.1.$$

$$C = [\hat{\theta} - q_n^*(0.975), \ \hat{\theta} - q_n^*(0.025)]$$

$$= [1.1, \ 1.65].$$

Percentile. $\quad T_n^* = \dfrac{\hat{\theta}^* - \hat{\theta}}{s(\hat{\theta}^*)}$. $\leftarrow$ we don't know.

13.1. $y_i = x_i'\beta + e_i$ $\quad E(x_i e_i) = 0.$

$e_i^2 = z_i'\gamma + \eta_i$ $\quad E(z_i \eta_i) = 0.$

Method of moments.

① $E(x_i e_i) = 0.$

$\Rightarrow E\left(x_i(y_i - x_i'\beta)\right) = 0.$

$E(x_i y_i) - E(x_i x_i')\beta = 0.$

if $E(x_i x_i')$ is inventable.

$\beta = E(x_i x_i')^{-1} E(x_i y_i).$

$\hat{\beta} = \left(\frac{1}{n}\sum_{i=1}^{n} x_i x_i'\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} x_i y_i\right).$

② $E(z_i \eta_i) = 0.$

$\Rightarrow E\left(z_i(e_i^2 - z_i'\gamma)\right) = 0.$

$\gamma = E(z_i z_i')^{-1} E(z_i e_i^2).$

$= E(z_i z_i')^{-1} E\left(z_i(y_i - x_i'\beta)^2\right).$

$\hat{\gamma} = \left(\frac{1}{n}\sum_{i=1}^{n} z_i z_i'\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} z_i(y_i - x_i'\hat{\beta})^2\right).$

13.2. GMM. with $W_n = (Z'Z)^{-1}$.

$$\hat{\beta} = (X'Z W_n Z'W)^{-1} X'Z W_n Z'y$$

$$= \beta + (X'Z W_n Z'W)^{-1} X'Z W_n Z'e.$$

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} N(0, \Omega).$$

$$\frac{1}{\sqrt{n}} Z'e \xrightarrow{d} N(0, \sigma^2 M).$$

$$\frac{1}{n} X'Z \xrightarrow{p} Q'$$

$$n W_n \xrightarrow{p} M^{-1}.$$

$$\Omega = (Q'M^{-1}Q)^{-1} \cancel{Q'M^{-1}} \sigma^2 \cancel{M M^{-1} Q (Q'M^{-1}Q)^{-1}}.$$

$$= \sigma^2 (Q'M^{-1}Q)^{-1}.$$

13.3. To simplify, consider $K = L = 1$.

Want to show. $\frac{1}{n}\sum_{i=1}^{n} z_i^2 \hat{e}_i^2 \xrightarrow{p} E(z_i^2 e_i^2)$.

$$\hat{e}_i^2 = (y_i - x_i\hat{\beta})^2$$

$$= e_i^2 - 2(\hat{\beta}-\beta)x_i e_i + (\hat{\beta}-\beta)^2 x_i^2.$$

$$\frac{1}{n}\sum_{i=1}^{n} z_i^2\hat{e}_i^2 = \underbrace{\frac{1}{n}\sum_{i=1}^{n} z_i^2 e_i^2}_{\downarrow p \;\; E(z_i^2 e_i^2)} - 2\underbrace{(\hat{\beta}-\beta)}_{\downarrow p \;\; 0}\frac{1}{n}\sum z_i^2 x_i e_i + \underbrace{(\hat{\beta}-\beta)^2}_{\downarrow p \;\; 0.}\frac{1}{n}\sum x_i^2 z_i^2$$

What ~~tetf~~ left to be shown:

① $\frac{1}{n}\sum z_i^2 x_i e_i$ is finite.

w.t.s. $E(z_i^2 x_i e_i)$ exist & finite.

$$E(z_i^2 x_i e_i) \leq E|z_i^2 x_i e_i|$$

$$= E|z_i x_i \cdot z_i e_i|.$$

$$\leq \sqrt{E(z_i^2 x_i^2)\cdot E(z_i^2 e_i^2)}$$

② $\frac{1}{n}\sum_{i=1}^{n} x_i^2 z_i^2 \xrightarrow{p} E(x_i^2 z_i^2)$.

We need the assumption: $E(x_i^2 z_i^2)$ exists & finite.

or a lower level condition:

$E x_i^4 < \infty$ and $E z_i^4 < \infty$

Remark; if $K>1$, $L>1$, the assumption becomes to.

$E[x_{ik}^2 z_{il}^2]$ exists & finite for all $l = 1, \cdots, L$, $k = 1, \cdots, K$.

or, $\forall\ k, l$, $E x_{ik}^4 < \infty$, $E z_{il}^4 < \infty$.