

1 (a) We have  $\hat{\mu}_{\text{gls}} = (X^T V^{-1} X)^{-1} X^T V^{-1} y$ , where  $X = \frac{1}{\sqrt{n}} \mathbf{1}_n$  and  $V = \text{diag}(\sigma_1^2, \dots, \sigma_n^2)$ .

Since

$$X^T V^{-1} X = \frac{1}{\sqrt{n}}^T \begin{pmatrix} \frac{1}{\sigma_1^2} \\ \vdots \\ \frac{1}{\sigma_n^2} \end{pmatrix} \frac{1}{\sqrt{n}} = \sum_{i=1}^n \frac{1}{\sigma_i^2}$$

and

$$X^T V^{-1} y = \frac{1}{\sqrt{n}}^T \begin{pmatrix} \frac{1}{\sigma_1^2} \\ \vdots \\ \frac{1}{\sigma_n^2} \end{pmatrix} y = \sum_{i=1}^n \frac{y_i}{\sigma_i^2},$$

we obtain

$$\hat{\mu}_{\text{gls}} = \frac{\sum_{i=1}^n y_i / \sigma_i^2}{\sum_{i=1}^n 1 / \sigma_i^2}.$$

(b) We have  $\hat{\mu}_{\text{ols}} = (X^T X)^{-1} X^T y = \left( \frac{1}{\sqrt{n}}^T \frac{1}{\sqrt{n}} \right)^{-1} \frac{1}{\sqrt{n}}^T y = \frac{1}{n} \sum_{i=1}^n y_i$ , so

$$\hat{\mu}_{\text{ols}} = \bar{y}.$$

(c) We have

$$\begin{aligned} \text{Var } \hat{\mu}_{\text{gls}} &= \text{Var} \left( \frac{\sum_{i=1}^n y_i / \sigma_i^2}{\sum_{i=1}^n 1 / \sigma_i^2} \right) \\ &= \left( \frac{1}{\sum_{i=1}^n 1 / \sigma_i^2} \right)^2 \sum_{i=1}^n \frac{\text{Var } y_i}{\sigma_i^4} \quad \leftarrow \sigma_i^2 \\ &= \frac{1}{\sum_{i=1}^n 1 / \sigma_i^2}. \end{aligned}$$

(d) We have  $\text{Var } \hat{\mu}_{\text{ols}} = \text{Var } \bar{y} = \frac{1}{n^2} \text{Var} \left( \sum_{i=1}^n y_i \right) = \frac{1}{n^2} \sum_{i=1}^n \sigma_i^2$ .

(e) If  $\sigma_i^2 = \sigma^2$  for all  $i$  then

$$\text{Var } \hat{\mu}_{\text{MLE}} = \frac{\sigma^2}{n} \quad \text{and} \quad \text{Var } \hat{\mu}_{\text{OLS}} = \frac{\sigma^2}{n},$$

so the two become equal.

(f) We have

$$\begin{aligned} \sum_{i=1}^n \left( \frac{y_i}{\sigma_i} \right)^2 &= \sum_{i=1}^n \left( \frac{\mu + \varepsilon_i}{\sigma_i} \right)^2 \\ &= \sum_{i=1}^n \left( \frac{\mu}{\sigma_i} + z_i \right)^2, \quad \text{where } z_i = \frac{\varepsilon_i}{\sigma_i} \\ &\sim \chi_n^2 \left( \phi = \sum_{i=1}^n \frac{\mu^2}{\sigma_i^2} \right). \end{aligned}$$

So for  $\mu=0$ ,  $\sum_{i=1}^n \frac{y_i^2}{\sigma_i^2} \sim \chi_n^2$ .

2 (a)

$$X = \begin{bmatrix} 1 & & & x_{11} \\ \vdots & & & \vdots \\ 1 & & & x_{1n} \\ & 1 & & x_{21} \\ & \vdots & & \vdots \\ & 1 & & x_{2n} \\ & & 1 & x_{31} \\ & & \vdots & \vdots \\ & & 1 & x_{3n} \end{bmatrix} \quad \tilde{b} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}$$

(b) We need  $n \geq 2$ . If  $n=1$  the columns will be linearly dependent.

(c) The assumption, along with requiring  $n \geq 2$ , gives  $X$  full column rank. This makes each of the parameters estimable.

Intuitively, we could not estimate the "slope" parameters  $\beta_1, \beta_2, \beta_3$  if in a group we observed only a single  $(x, y)$  pair.

(d) Reformulate the hypothesis as  $H_0: \beta_1 = 0$  &  $\beta_2 = 0$  &  $\beta_3 = 0$ .

Then set  $K^T = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$  and  $w = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

(e) The numerator df is  $\text{rank } K = 3$ .

The denominator df is the total # obs minus rank  $X$ . This is

$$3n - 6$$

(f) This corresponds to testing  $H_0: \mu_1 - \mu_2 = 0$  &  $\mu_2 - \mu_3 = 0$   
&  $\beta_1 - \beta_2 = 0$  &  $\beta_2 - \beta_3 = 0$ .

set  $K^T = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}$  and  $w = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$

(g) Numerator df =  $\text{rank } K = 4$  and denominator df =  $3n - 6$ .

**3** (a) let  $\tilde{P}_X = X(X^T V^{-1} X)^{-1} X^T V^{-1}$ .

(i) We have  $X \underbrace{(X^T V^{-1} X)^{-1} X^T V^{-1} X}_{I: n \times n} (X^T V^{-1} X)^{-1} X^T V^{-1} = X (X^T V^{-1} X)^{-1} X^T V^{-1}$ , so  $\tilde{P}_X$  is idempotent.

(ii) For any  $\tilde{y}$ ,  $X (X^T V^{-1} X)^{-1} X^T V^{-1} \tilde{y} \in \text{Col } X$ .

(iii) For  $\tilde{z} \in \text{Col } X$ ,  $\tilde{z} = X \tilde{z}_2$  for some  $\tilde{z}_2$ , so

$$X (X^T V^{-1} X)^{-1} X^T V^{-1} \tilde{z} = X (X^T V^{-1} X)^{-1} X^T V^{-1} X \tilde{z}_2 = X \tilde{z}_2 = \tilde{z}$$

(b) For any  $\tilde{y} \in \mathbb{R}^n$ , check whether

$$\tilde{y} = \tilde{P}_X \tilde{y} + (\mathbf{I} - \tilde{P}_X) \tilde{y}$$

is an orthogonal decomposition.

We have

$$\begin{aligned} (\tilde{P}_X \tilde{y}) \cdot (\mathbf{I} - \tilde{P}_X) \tilde{y} &= \tilde{y}^T \tilde{P}_X^T (\mathbf{I} - \tilde{P}_X) \tilde{y} \\ &= \tilde{y}^T \left[ V^{-1} X (X^T V^{-1} X)^{-1} X^T (\mathbf{I} - X (X^T V^{-1} X)^{-1} X^T V^{-1}) \right] \tilde{y} \\ &= \tilde{y}^T \left[ V^{-1} X (X^T V^{-1} X)^{-1} X^T - V^{-1} X (X^T V^{-1} X)^{-1} X^T X (X^T V^{-1} X)^{-1} X^T V^{-1} \right] \tilde{y} \end{aligned}$$

We see that this is 0 if  $V = \mathbf{I} \sigma^2$  for any  $\sigma^2$ , but it is not in general equal to 0.

So  $\tilde{P}_X$  is not an orthogonal projection.

One can also make a result which says that a projection matrix is the matrix of an orthogonal projection iff it is symmetric.

Since  $\tilde{P}_X$  is not symmetric, it is not an orthogonal projection.