THE LINEAR MODEL

• GENERALIZED t- AND F-TESTS

It is now assumed that the random vector $\boldsymbol{y}$ in the linear model has the $N(\boldsymbol{\eta}, \boldsymbol{\sigma^2 I})$ distribution.

Theorem: Let $\boldsymbol{T'y}$ be the best estimator of $\boldsymbol{Q'b}$. Then, the error sum of squares, SSE, has the $\sigma^2\chi^2(n-r)$ distribution, independent of $\boldsymbol{T'y}$.

Proof:

Suppose that $\boldsymbol{V_e}$ is generated by the columns of $\boldsymbol{E}: n \times (n-r)$. Then $\boldsymbol{E'y}$ has the $N(\boldsymbol{0}, \sigma^2\boldsymbol{E'E})$ distribution. It follows that $SSE = \boldsymbol{y'E(E'E)^{-1}E'y}$ has the $\sigma^2\chi^2(n-r)$ distribution. Since $\boldsymbol{T'y}$ and $\boldsymbol{E'y}$ are independent $(\boldsymbol{Cov}\ (\boldsymbol{T'y}, \boldsymbol{E'y}) = \boldsymbol{T'E\sigma^2} = \boldsymbol{0})$, $\boldsymbol{T'y}$ and SSE are independent.

The Generalized $\boldsymbol{t}$-Test

Suppose that $\boldsymbol{q'b}$ is an estimable function with best estimator $\boldsymbol{t'y}$. Therefore

$\boldsymbol{t'y} \sim N(\boldsymbol{q'b}, \sigma^2\boldsymbol{t't})$, independent of SSE/σ^2, which has the $\chi^2(n-r)$ distribution. Under the null hypothesis:

$$H_0: \boldsymbol{q'b} = m$$

for a given value of m, it follows that

$$\begin{aligned} t &= \frac{\boldsymbol{t'y} - m}{\sqrt{\boldsymbol{t't}\sigma^2}} / \sqrt{MSE/\sigma^2} \\ &= \frac{\boldsymbol{t'y} - m}{\sqrt{\boldsymbol{t't}MSE}} \\ &= \frac{\boldsymbol{t'y} - m}{\sqrt{\text{estimated var } (\boldsymbol{t'y})}} \end{aligned}$$

has a $\boldsymbol{t}$ distribution with $n-r$ degrees of freedom under H_0.

A $(1-\epsilon)$ confidence interval for $\boldsymbol{q'b}$ is:

$$\boldsymbol{t'y} \pm t_{\frac{\epsilon}{2}}(n-r)\sqrt{\text{estimated variance of } \boldsymbol{t'y}}$$

where (estimated variance of $\boldsymbol{t'y}$) $= \boldsymbol{t't}MSE = \boldsymbol{q'}(\boldsymbol{X'X})^*\boldsymbol{q}MSE$.

Examples

$E(y) = Xb \qquad Var(y) = Cov(y) = \sigma^2 I$

1. Suppose

$r\left(\begin{pmatrix} q^T \\ X \end{pmatrix}\right) = \text{rank}(X)$

$$\begin{pmatrix} \mu+\alpha+\beta+\gamma \\ \mu+\alpha+\beta+\gamma \\ \mu-\alpha-\beta+\gamma \\ \mu-\alpha-\beta+\gamma \\ \mu+\alpha+\beta-\gamma \\ \mu+\alpha+\beta-\gamma \\ \mu-\alpha-\beta-\gamma \\ \mu-\alpha-\beta-\gamma \end{pmatrix} = \overbrace{\begin{pmatrix} 1 & +1 & +1 & +1 \\ 1 & +1 & +1 & +1 \\ 1 & -1 & -1 & +1 \\ 1 & -1 & -1 & +1 \\ 1 & +1 & +1 & -1 \\ 1 & +1 & +1 & -1 \\ 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 \end{pmatrix}}^{X} \underbrace{\begin{pmatrix} \mu \\ \alpha \\ \beta \\ \gamma \end{pmatrix}}_{}\!{}^{\underset{\sim}{b}}$$

$q = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{pmatrix}$

and $cov\,(\boldsymbol{y}) = \sigma^2\boldsymbol{I}$. Consider the linear function $\boldsymbol{q}'\boldsymbol{b} = q_1\mu + q_2\alpha + q_3\beta + q_4\gamma$. Further, assume $\boldsymbol{y}$ has a normal distribution. Then, variance of best estimator of $\boldsymbol{q}'\boldsymbol{b}$ is given by (independent?) $Cov(q^T\hat{b}) = q^T Cov(\hat{b})\, q$ $\quad q^T\hat{b}$

$$\longrightarrow \text{var}(q_1\hat{\mu} + q_2\hat{\alpha} + q_3\hat{\beta} + q_4\hat{\gamma}) = q_1^2\text{var}(\hat{\mu}) + q_2^2\text{var}(\hat{\alpha}) + q_3^2\text{var}(\hat{\beta}) + q_4^2\text{var}(\hat{\gamma}).$$

Thus, we need to compute $\text{Cov}(\widehat{\boldsymbol{b}})$, which is $\text{Cov}(\widehat{\boldsymbol{b}}) = (\boldsymbol{X}'\boldsymbol{X})^*\sigma^2$. How can we calculate $(\boldsymbol{X}'\boldsymbol{X})^*$?

First, $\boldsymbol{q}'\boldsymbol{b}$ is estimable iff $q_2 = q_3$. Using the fact that $\widehat{\boldsymbol{b}} = (\boldsymbol{X}'\boldsymbol{X})^*\boldsymbol{X}'\boldsymbol{y}$, where

$$\boldsymbol{X}'\boldsymbol{X} = \begin{pmatrix} 8 & 0 & 0 & 0 \\ 0 & 8 & 8 & 0 \\ 0 & 8 & 8 & 0 \\ 0 & 0 & 0 & 8 \end{pmatrix}.$$

and $$\boldsymbol{X}'\boldsymbol{y} = \begin{pmatrix} y_1+y_2+y_3+y_4+y_5+y_6+y_7+y_8 \\ y_1+y_2-y_3-y_4+y_5+y_6-y_7-y_8 \\ y_1+y_2-y_3-y_4+y_5+y_6-y_7-y_8 \\ y_1+y_2+y_3+y_4-y_5-y_6-y_7-y_8 \end{pmatrix} = \begin{pmatrix} Y_1 \\ Y_2 \\ Y_2 \\ Y_3 \end{pmatrix}$$

we have $$\begin{pmatrix} Y_1/8 \\ Y_2/8 \\ 0 \\ Y_3/8 \end{pmatrix} = \begin{pmatrix} ? & ? & ? & ? \\ ? & ? & ? & ? \\ ? & ? & ? & ? \\ ? & ? & ? & ? \end{pmatrix}\begin{pmatrix} Y_1 \\ Y_2 \\ Y_2 \\ Y_3 \end{pmatrix}$$

that gives

$$(\boldsymbol{X}'\boldsymbol{X})^* = \begin{pmatrix} 1/8 & 0 & 0 & 0 \\ 0 & 1/8 & 8 & 0 \\ 0 & 8 & 8 & 0 \\ 0 & 0 & 0 & 1/8 \end{pmatrix}$$

Thus, $\hat{\mu} = Y_1/8$, $\hat{\alpha} = Y_2/8$, $\hat{\beta} = 0$, $\hat{\gamma} = Y_3/8$.

Therefore, $\mathrm{Cov}(\hat{\boldsymbol{b}}) = \begin{pmatrix} 1/8 & 0 & 0 & 0 \\ 0 & 1/8 & 8 & 0 \\ 0 & 8 & 8 & 0 \\ 0 & 0 & 0 & 1/8 \end{pmatrix} \sigma^2$ and under normality assumption, the estimators are independent. Furthermore,

$\widehat{q^T b}$

$$\mathrm{var}\big(q_1\hat{\mu} + q_2\hat{\alpha} + q_3\hat{\beta} + q_4\hat{\gamma}\big) = \mathrm{var}(q_1Y_1/8 + q_2Y_2/8 + q_4Y_3/8) = \frac{1}{8}(q_1^2 + q_2^2 + q_4^2)\sigma^2.$$

For the null hypothesis $H_0\colon \gamma = 0$, it follows that $\boldsymbol{q} = (0,0,0,1)'$ and

$q^T b = \gamma$

$$\mathrm{var}\ (\hat{\gamma}) = \frac{1}{8}\sigma^2$$

and (Recall $t = \frac{\overset{\widehat{q^T b}}{t'y} - \overset{m}{q'b}}{\sqrt{t't\sigma^2}} / \sqrt{MSE/\sigma^2}$)

$$t_5 = \hat{\gamma} \Big/ \sqrt{\frac{1}{8} MSE}.$$

$q^T b$ $\quad q = \begin{pmatrix} 0 \\ 1 \\ 1 \\ -1 \end{pmatrix}$

Also, to construct A $(1-\epsilon)$ confidence interval for $\alpha + \beta - \gamma$, we have

$$\hat{\alpha} + \hat{\beta} - \hat{\gamma} = \frac{1}{4}(y_5 + y_6 - y_3 - y_4)$$

and

$$\mathrm{var}\ (\hat{\alpha} + \hat{\beta} - \hat{\gamma}) = \frac{1}{4}\sigma^2$$

Noting that $q = (0,1,1,-1)'$, a $(1-\epsilon)$ confidence interval for $\alpha + \beta - \gamma$ is

$$\hat{\alpha} + \hat{\beta} - \hat{\gamma} \pm \frac{1}{2} t_{\epsilon/2}(\underset{n-r}{5})\sqrt{MSE}.$$

2. Suppose that $\overbrace{y_{11}, y_{12}, \cdots, y_{1n_1}}^{n_1}$ is a random sample from $N(\mu_1, \sigma^2)$ distribution and that $\overbrace{y_{21}, y_{22}, \cdots, y_{2n_2}}^{n_2}$ is an independent sample from a $N(\mu_2, \sigma^2)$ distribution. The linear model can be expressed as:

$$Xb = E(\mathbf{y}) = E\begin{pmatrix} y_{11} \\ y_{12} \\ \cdots \\ y_{1n_1} \\ y_{21} \\ y_{22} \\ \cdots \\ y_{2n_2} \end{pmatrix} = \overbrace{\begin{pmatrix} 1 & 0 \\ 1 & 0 \\ \cdots & \cdots \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ \cdots & \cdots \\ 0 & 1 \end{pmatrix}}^{X} \underbrace{\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}}_{b}. = \overbrace{\begin{pmatrix} 1_{n_1} & 0_{n_1} \\ 0_{n_2} & 1_{n_2} \end{pmatrix}}^{X} \underbrace{\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}}_{b}$$

$\text{Rank}(X) = 2$

$\begin{pmatrix} \hat\mu_1 \\ \hat\mu_2 \end{pmatrix}$

It follows that

$$\begin{pmatrix} 1_{n_1}^T 1_{n_1} & 0 \\ 0 & 1_{n_2}^T 1_{n_2} \end{pmatrix} = \boldsymbol{X'X} = \begin{pmatrix} n_1 & 0 \\ 0 & n_2 \end{pmatrix} \qquad X^Ty = \begin{pmatrix} 1_{n_1}^T y \\ 1_{n_2}^T y \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{n_1} y_{1j} \\ \sum_{j=1}^{n_2} y_{2j} \end{pmatrix}$$

$$\hat b = (X^TX)^{-1} X^T y = (X^TX)^{-1} X^T y = \begin{pmatrix} \frac{1}{n_1} & 0 \\ 0 & \frac{1}{n_2} \end{pmatrix} \begin{pmatrix} \sum^{n_1} y_{1j} \\ \sum^{n_2} y_{2j} \end{pmatrix} = \begin{pmatrix} \bar y_1 \\ \bar y_2 \end{pmatrix}$$

$$\hat\mu_i = \bar{y}_i$$

$$SSE = \sum_{i=1}^{2} \sum_{j=1}^{n_i} y_{ij}^2 - n_1 \bar{y}_1^2 - n_2 \bar{y}_2^2 = \sum_{i=1}^{2} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2$$

and

$$SSE = y^Ty - \hat b^T X^T y = \sum_j \sum_i y_{ij}^2 - (\bar y_1 \;\; \bar y_2) \begin{pmatrix} n \bar y_1 \\ n \bar y_2 \end{pmatrix}$$

$$MSE = SSE/(\underbrace{n_1 + n_2 - 2}_{n-r}).$$

Since

$$\text{var}\,(\bar{y}_1 - \bar{y}_2) = \left(\frac{1}{n_1} + \frac{1}{n_2}\right) \sigma^2$$

$= \text{Var}(\bar y_1) + \text{Var}(\bar y_2)$

It follows that the t-statistic for the hypothesis:

$q^Tb \qquad b = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \Rightarrow q = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

$H_0: \underset{\mu_1 - \mu_2}{q^Tb} = \underset{0}{m}$

$$H_0: \mu_1 = \mu_2$$

$\mu_1 - \mu_2 = 0$

is given by

$$\frac{q^T\hat b - m}{\sqrt{\text{est var}(q^T\hat b)}} = \frac{\hat\mu_1 - \hat\mu_2}{\sqrt{\text{est var}(\bar y_1 - \bar y_2)}}$$

$$t_{n_1+n_2-2} = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right) MSE}}.$$

$T^T y = \begin{pmatrix} t_1^T y \\ t_2^T y \\ \vdots \\ t_s^T y \end{pmatrix}$ $\quad Q^T b = \begin{pmatrix} q_1^T b \\ q_2^T b \\ \vdots \\ q_s^T b \end{pmatrix}$

The Generalized $\boldsymbol{F}$-Test

Suppose that the elements of $\boldsymbol{Q}'\boldsymbol{b}$, are s linearly independent estimable functions with best estimators the elements of $\boldsymbol{T}'\boldsymbol{y}$.

Then $\boldsymbol{T}'\boldsymbol{y} \sim N(\boldsymbol{Q}'\boldsymbol{b}, \sigma^2\boldsymbol{T}'\boldsymbol{T})$, independent of $\frac{SSE}{\sigma^2}$, which has the $\chi^2(n-r)$ distribution. Therefore

$y^T T(T^T T)^{-1} T^T y \sim \sigma^2 \chi^2_{(s)}$

$$\boldsymbol{y}'\boldsymbol{T}(\boldsymbol{T}'\boldsymbol{T})^{-1}\boldsymbol{T}'\boldsymbol{y}/\sigma^2$$

has the $\chi^2(s)$ distribution under the hypothesis $\boxed{\boldsymbol{Q}'\boldsymbol{b} = \boldsymbol{0}}$. Suppose that V_s and L_s are the vector space and linear set generated by the columns of $\boldsymbol{T}$ and the elements of $\boldsymbol{T}'\boldsymbol{y}$. Then

$$SSH = \boldsymbol{y}'\boldsymbol{T}(\boldsymbol{T}'\boldsymbol{T})^{-1}\boldsymbol{T}'\boldsymbol{y}$$

is the sum of squares for L_s.

Under the hypothesis:

$u \sim \chi^2_{(p)}$ $\perp$ $v \sim \chi^2_{(q)}$ $\Rightarrow W = \frac{u/p}{v/q} \sim F(p,q)$

$$H_0: \boldsymbol{Q}'\boldsymbol{b} = \boldsymbol{0}$$

it follows that

$F = \frac{SSH/s}{SSE/n-r}$

$$F = \frac{SSH}{SSE}\frac{n-r}{s}$$

has an F distribution with s and $n-r$ degrees of freedom.

$$F \sim F(s, n-r).$$

Note that $\underline{V_s \subset V_r}$, $L_s \subset L_r$ and that SSH is the square of the length of the projection of $\boldsymbol{y}$ on V_s.

V_e generated $\quad E: n\times(n-r)$ $\quad y \sim N(Xb, \sigma^2 I)$

$E^T y \sim N(0, \sigma^2 E^T E)$

$z = (\sigma^2 E^T E)^{-\frac{1}{2}} E^T y \sim N(0, I)$

$y^T E(E^T E)^{-1} E^T y \sim \sigma^2 \chi^2_{(n-r)}$

$Cov(T^T y, E^T y) = T^T E \sigma^2 = 0$ $\quad T^T y \perp E^T y$

$T^T y \perp SSE$

es — Directly / PCE

PCE

The Principle of Conditional Error

$t_i \perp t_j \Rightarrow t_i^T t_j = 0$

$SSH = \sum_{i=1}^{s} \frac{(t_i^T y)^2}{t_i^T t_i}$

not always $t_i^T t_j = 0$

Instead of calculating the hypothesis sum of squares

$$SSH = \boldsymbol{y}'\boldsymbol{T}(\boldsymbol{T}'\boldsymbol{T})^{-1}\boldsymbol{T}'\boldsymbol{y}$$

directly, it can be determined more conveniently by using the conditional error sum of squares. The model under the hypothesis is

$E(y) = Xb$

Under $H_0 : Q^T b = 0$

$\Rightarrow E(y) = X_0 b_0$

fewer design   fewer par.

$$E(\boldsymbol{y}) = \boldsymbol{X_0}\boldsymbol{b_0}$$

with the new estimation space V_{r-s} such that

$$V_r = V_{r-s} \oplus V_s$$

with V_{r-s} and V_s mutually orthogonal. A necessary and sufficient set of conditions are:

1. $E(\boldsymbol{T}'\boldsymbol{y}) = 0$ under this model,
2. $V_{r-s} \subset V_r$ and
3. $\text{rank}\ (\boldsymbol{X_0}) = r - s$.

V_{r-s}

rank $= r-s$

This implies that the new error space (the conditional error space)

$$V_c = V_s \oplus V_e$$

and the new error set (the conditional error set)

$$L_c = L_s \oplus L_e.$$

The error sum of squares under the new model (the conditional error sum of squares) is

$$SSC = SSE + SSH$$

and the hypothesis sum of squares:

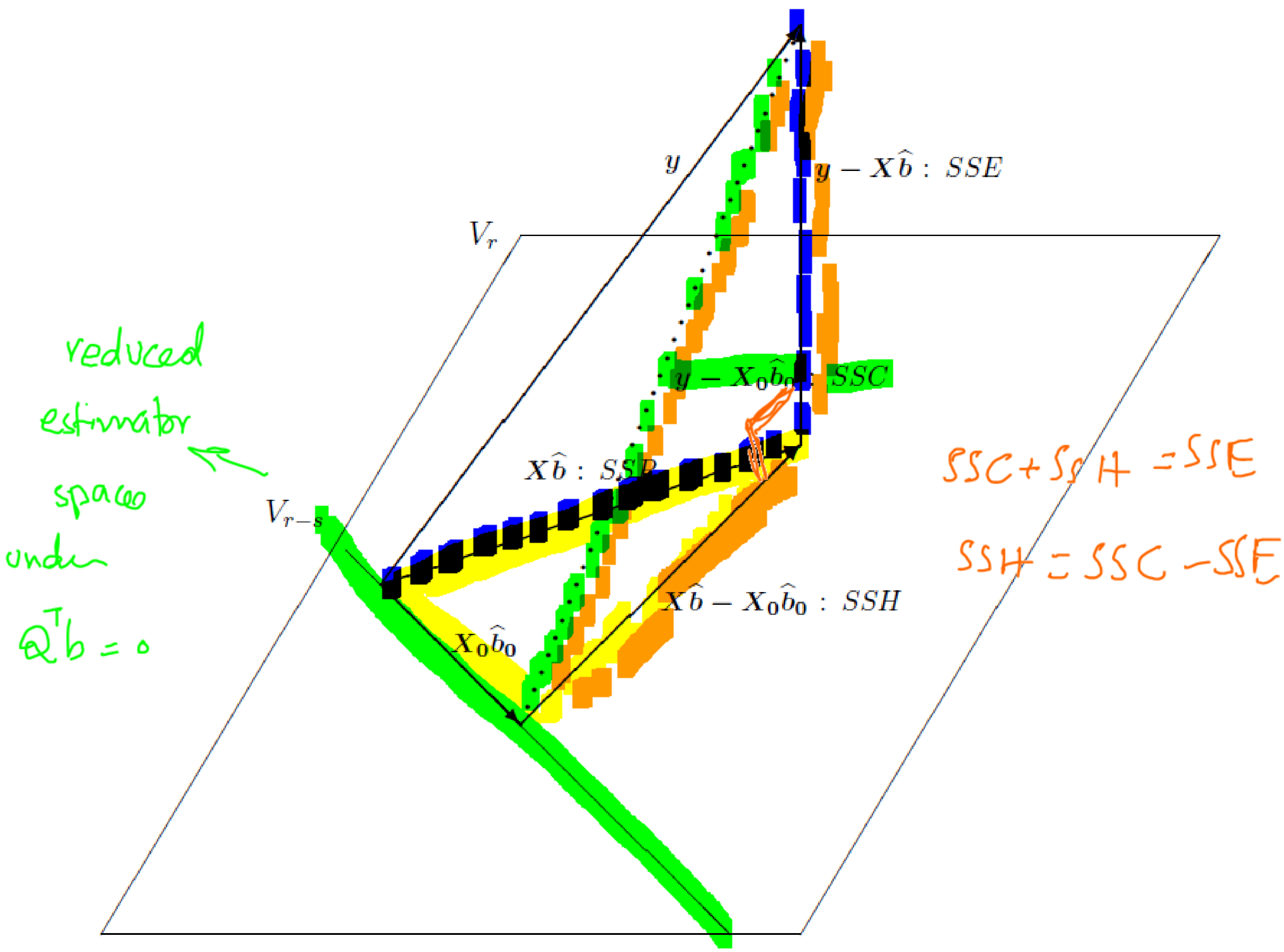
$$SSH = SSC - SSE.$$

$= SSR - SSR|_{H_0}$

Also:

$$L_r = L_s \oplus L_{r-s}.$$
$$SSR = SSH + SS_{r-s}.$$

Schematic Representation



Examples

$\hat{b} = (X^TX)^* X^T y = (X^TX)^{-1} X^T y$

$X^TXb = \begin{pmatrix}4&0&0\\0&2&0\\0&0&2\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\\ \gamma\end{pmatrix}$

$X^Ty = \begin{pmatrix}\sum_1^4 y_i\\ y_1-y_2\\ y_3-y_4\end{pmatrix}$

1. The random vector $\boldsymbol{y}$ has the $N(\boldsymbol{\eta}, \sigma^2\boldsymbol{I})$ distribution with

$$\boldsymbol{\eta} = \begin{pmatrix}\alpha+\beta\\ \alpha-\beta\\ \alpha+\gamma\\ \alpha-\gamma\end{pmatrix} = \underbrace{\begin{array}{ccc} x_1 & x_2 & x_3\\ \end{array}\begin{pmatrix}1&1&0\\1&-1&0\\1&0&1\\1&0&-1\end{pmatrix}}_{X}\underbrace{\begin{pmatrix}\alpha\\ \beta\\ \gamma\end{pmatrix}}_{b}.$$

Rank $(X) = 3$

$(X^TX)^* = (X^TX)^{-1}$

The parameters are estimable with best estimators:

All q^Tb are estimable

$$\hat{\alpha} = \frac{1}{4}(y_1+y_2+y_3+y_4)$$
$$\hat{\beta} = \frac{1}{2}(y_1-y_2)$$
$$\hat{\gamma} = \frac{1}{2}(y_3-y_4).$$

The error sum of squares is $SSE = \boldsymbol{y}'\boldsymbol{y} - 4\hat{\alpha}^2 - 2\hat{\beta}^2 - 2\hat{\gamma}^2$.

Consider the hypothesis $H_0\!: \beta = \gamma = 0$. The hypothesis can be expressed as $H_0\!: \boldsymbol{Q}'\boldsymbol{b} = \begin{pmatrix}\beta\\ \gamma\end{pmatrix} = 0$.

$T^Ty = \begin{pmatrix}\hat\beta\\ \hat\gamma\end{pmatrix}$

$\beta = (0\ 1\ 0)b = 0 = q_1^Tb$

With $\boldsymbol{T}'\boldsymbol{y} = \begin{pmatrix}(y_1-y_2)/2\\ (y_3-y_4)/2\end{pmatrix}$. Since the elements of $\boldsymbol{T}'\boldsymbol{y}$ are mutually orthogonal, it follows that

$\gamma = (0\ 0\ 1)b = 0 = q_2^Tb$

two function $\Rightarrow s = 2$

$n - r = 4 - 3 = 1$

$$SSH = (y_1-y_2)^2/2 + (y_3-y_4)^2/2$$

$Q^T = \begin{pmatrix}q_1^T\\ q_2^T\end{pmatrix}$

and the F-test for the hypothesis is

$F = \frac{SSH/s}{SSE/_{n-r}}$, MSE

$$F = \frac{SSH}{SSE}\frac{1}{2}. \sim F(2,1)$$

By using the principle of conditional error, it follows that under the hypothesis

Method of PCE

Under $Q^Tb = 0$

$X_0^TX_0\, b = 1_4^T1_4\,\alpha$

$$E(\boldsymbol{y}) = \underbrace{\begin{pmatrix}1\\1\\1\\1\end{pmatrix}}_{X_0}\underbrace{\alpha}_{b_0}$$

$\eta = \begin{pmatrix}\alpha+\beta\\ \alpha-\beta\\ \alpha+\gamma\\ \alpha-\gamma\end{pmatrix}$, $\beta = 0$, $\gamma = 0$

$b_0 = \alpha$

$X_0 = \begin{pmatrix}1\\1\\1\\1\end{pmatrix}$

and $y_0^Ty = \sum_{i=1}^4 y_i$

$SSC = y^Ty - \hat{b}_0X_0^Ty$

$$SSC = \boldsymbol{y}'\boldsymbol{y} - (y_1+y_2+y_3+y_4)^2/4.$$

The hypothesis sum of squares is therefore

$= y^Ty - 4\hat\alpha^2$

$b_0 = \hat\alpha$ $\quad X_0^Ty = 4\hat\alpha$

$$SSH = SSC - SSE = 2\hat{\beta}^2 + 2\hat{\gamma}^2.$$

$\cancel{y^Ty} - \cancel{4\hat\alpha^2} - (\cancel{y^Ty} - \cancel{4\hat\alpha^2} - 2\hat\beta^2 - 2\hat\gamma^2)$

$=$ gives the same answer

$$SSE = y^Ty - \hat{b}X^Ty$$

$$= y^Te(e^Te)^{-1}e^Ty$$

$$= \frac{(e^Ty)^2}{e^Te} \qquad e^Te = 4$$

$$= \tfrac{1}{4}(y_1 + y_2 - y_3 - y_4)^2$$

$rank(V_e) = 4 - 3 = 1$

1 vector $\underline{\underline{e}}$ can generate V_e

basis

$x_i^Te = 0 \quad i = 1, 2, 3$

$x_1^Te = 0 \implies e_1 + e_2 + e_3 + e_4 = 0$

$x_2^Te = 0 \implies e_1 - e_2 = 0 \implies e_1 = e_2$

$x_3^Te = 0 \implies e_3 - e_4 = 0 \implies e_3 = e_4$

$$\implies e = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}$$

$$SSH = y^TT(T^TT)^{-1}T^Ty$$

$$= \sum_{i=1}^{2} \frac{(t_i^Ty)^2}{t_i^Tt_i}$$

$t_1 = \tfrac{1}{2}(1 \ \ -1 \ \ 0 \ \ 0)$

$t_2 = \tfrac{1}{2}(0 \ \ 0 \ \ 1 \ \ -1)$

2. Suppose

$$\begin{pmatrix} \mu+\alpha+\beta+\gamma \\ \mu+\alpha+\beta+\gamma \\ \mu-\alpha-\beta+\gamma \\ \mu-\alpha-\beta+\gamma \\ \mu+\alpha+\beta-\gamma \\ \mu+\alpha+\beta-\gamma \\ \mu-\alpha-\beta-\gamma \\ \mu-\alpha-\beta-\gamma \end{pmatrix} = \overbrace{\begin{pmatrix} 1 & +1 & +1 & +1 \\ 1 & +1 & +1 & +1 \\ 1 & -1 & -1 & +1 \\ 1 & -1 & -1 & +1 \\ 1 & +1 & +1 & -1 \\ 1 & +1 & +1 & -1 \\ 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 \end{pmatrix}}^{X} \underset{\sim}{b} \begin{pmatrix} \mu \\ \alpha \\ \beta \\ \gamma \end{pmatrix}$$

$\hat{\mu} = \frac{1}{8}Y_1$

$\hat{\alpha} = \frac{1}{8}Y_2$

$\hat{\beta} = 0$

$\hat{\gamma} = \frac{1}{8}Y_3$

and $cov\,(\boldsymbol{y}) = \sigma^2\boldsymbol{I}$. Further, suppose that the random vector $\boldsymbol{y}$ has a normal distribution and consider the hypothesis

$\Rightarrow SSE = y^Ty - \frac{1}{8}Y_1^2 - \frac{1}{8}Y_2^2 - \frac{1}{8}Y_3^2$

$$H_0: \alpha + \beta = 0, \quad \gamma = 0.$$

In this case

$SSH = y^T T(T^T T)^{-1} T^T y$ $\quad \begin{pmatrix} q_1^T b \\ q_2^T b \end{pmatrix} = \boldsymbol{Q}'\boldsymbol{b} = \begin{pmatrix} \alpha+\beta \\ \gamma \end{pmatrix} \Rightarrow H_0: Q^T b = 0$

and $= \sum_{i=1}^{2} \frac{(t_i^T y)^2}{t_i^T t_i}$

$$\boldsymbol{T}'\boldsymbol{y} = \begin{pmatrix} Y_2/8 \\ Y_3/8 \end{pmatrix} = \begin{pmatrix} (y_1+y_2-y_3-y_4+y_5+y_6-y_7-y_8)/8 \\ (y_1+y_2+y_3+y_4-y_5-y_6-y_7-y_8)/8 \end{pmatrix} = \begin{pmatrix} \hat{\alpha}+\hat{\beta} \\ \hat{\gamma} \end{pmatrix}$$

Since Y_2 and Y_3 are mutually orthogonal, it follows that SSH is the total of the individual sums of squares of Y_2 and Y_3, namely

$F = \frac{SSH/s}{SSE/n-r}$ $\quad s=2 \quad n=8 \quad r=3 \quad n-r=5$

$$SSH = Y_2^2/8 + Y_3^2/8. = \begin{pmatrix} \frac{1}{8}(1\ 1\ -1\ -1\ 1\ 1\ -1\ -1)y \\ \frac{1}{8}(1\ 1\ 1\ 1\ -1\ -1\ -1\ -1)y \end{pmatrix}$$

If the principle of conditional error is used, it follows that under the null hypothesis

$F = \frac{SSH/2}{SSE/5} \sim F(2,5)$

$$E(\boldsymbol{y}) = \underbrace{\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}}_{X_0} \underbrace{\mu}_{b_0}.$$

$X_0 = 1_8$

$b_0 = \mu$

Therefore

$X_0^T X_0 b = 1_8^T 1_8 \mu = 8\mu$

$X_0^T y = \sum y_i$

$$SSC = \boldsymbol{y}'\boldsymbol{y} - \frac{1}{8}Y_1^2$$

$\hat{\mu} = \bar{y}$

and

$$SSE = \boldsymbol{y}'\boldsymbol{y} - \frac{1}{8}Y_1^2 - \frac{1}{8}Y_2^2 - \frac{1}{8}Y_3^2$$

such that

$$SSH = SSC - SSE = \frac{1}{8}Y_2^2 + \frac{1}{8}Y_3^2.$$

The F-test for the hypothesis is

$$F = \frac{SSH}{SSE}\frac{5}{2}.$$

In the case of the hypothesis

$$H_0: \mu = \alpha + \beta = \gamma.$$

it follows, however, that the hypothesis can be given as

$$SSC = y^T y - \hat{b}_0^T X_0^T y \qquad H_0: \boldsymbol{Q}'\boldsymbol{b} = \begin{pmatrix} \mu - \alpha - \beta \\ \mu - \gamma \end{pmatrix} = 0. \qquad \hat{b}_0 = \hat{\mu}$$

In this case

$$= y^T y - 8\hat{\mu}^2 \qquad X_0^T y = 8\hat{\mu}$$

$$\boldsymbol{T}'\boldsymbol{y} = \begin{pmatrix} Y_1/8 - Y_2/8 \\ Y_1/8 - Y_3/8 \end{pmatrix} = \begin{pmatrix} (y_3 + y_4 + y_7 + y_8)/4 \\ (y_5 + y_6 + y_7 + y_8)/4 \end{pmatrix}.$$

Again SSH can be directly determined from the expression

$$SSH = \boldsymbol{y}'\boldsymbol{T}(\boldsymbol{T}'\boldsymbol{T})^{-1}\boldsymbol{T}'\boldsymbol{y}$$

It, however, follows easier that the elements of

$$\begin{pmatrix} (y_3 + y_4 + y_5 + y_6 + 2y_7 + 2y_8)/4 \\ (y_3 + y_4 - y_5 - y_6)/4 \end{pmatrix}$$

form a mutually orthogonal basis for L_s. The hypothesis sum of squares is:

$$SSH = (y_3 + y_4 + y_5 + y_6 + 2y_7 + 2y_8)^2/12 + (y_3 + y_4 - y_5 - y_6)^2/4.$$

By using the principle of conditional error, the model under the hypothesis is

$$SSH = SSC - SSE$$

$$= y^T y - 8\hat{\mu}^2 - \left(y^T y - \tfrac{1}{8}Y_1^2 - \tfrac{1}{8}Y_2^2 - \tfrac{1}{8}Y_3^2\right)$$

$$E(\boldsymbol{y}) = \begin{pmatrix} 3 \\ 3 \\ 1 \\ 1 \\ 1 \\ 1 \\ -1 \\ -1 \end{pmatrix} \mu.$$

Under this model

$$\hat{\mu} = (3y_1 + 3y_2 + y_3 + y_4 + y_5 + y_6 - y_7 - y_8)/24 = Y_0/24, \text{ say }.$$

Therefore

$$SSC = \boldsymbol{y}'\boldsymbol{y} - Y_0^2/24$$

and the hypothesis sum of squares is

$$SSH = SSC - SSE = Y_1^2/8 + Y_2^2/8 + Y_3^2/8 - Y_0^2/24.$$

It must be shown algebraically that the two expressions for SSH are the same.

The F-test for the hypothesis is

$$F = \frac{SSH}{SSE}\frac{5}{2}.$$

To see explanatory variable significantly contribute to the explanation of variation in dependent variable (y)

- ## Multiple Correlation

Method PCE

Consider, under the assumption: $\mathbf{1} \in V_r$, with $\mathbf{1}' = (1,1,\cdots,1)$, the null hypothesis

$$H_0: E(\boldsymbol{y}) = \mathbf{1}\mu \,.$$

It implies that the expected values of all the elements of $\boldsymbol{y}$ are the same. The conditional error sum of squares is $r - s = 1 \Rightarrow s = r - 1$

$X_0 b_0 = 1\mu$

$X_0^T X_0 b_0 = 1_n^T 1_n \mu = n\mu$

$X_0^T y = 1_n^T y = \sum y_i \Rightarrow \hat{\mu} = \bar{y}$

$$SSC = \sum_{i=1}^{n} y_i^2 - n\bar{y}^2 = \sum_{i=1}^{n} (y_i - \bar{y})^2.$$

$SSC = y^T y - \hat{b}_0^T X_0^T y = y^T y - \bar{y}(n\bar{y})$

A measure of how well the model fits, is the ratio of the hypothesis sum of squares and the conditional error sum of squares:

$SSH \sim \chi^2_{(r-1)}$

$SSE \sim \chi^2_{(n-r)}$

$n\bar{y}$ $\frac{R^2}{1-R^2} = \frac{SSH}{\cancel{SSC}} \times \frac{\cancel{SSC}}{SSE} = \frac{SSH}{SSE}$

$$R^2 = \frac{SSC - SSE}{SSC} \,, = \frac{SSH}{SSC}$$

which is the coefficient of determination. The positive square root, R, is the multiple correlation coefficient. The F-test for the hypothesis is:

$$F_{r-1,n-r} = \frac{R^2}{1-R^2}\frac{n-r}{r-1}\,.$$

Note that $0 \leq R \leq 1$. Furthermore

Unexplained variation $\Big\{$ $$1 - R^2 = \frac{SSE}{SSC}$$

the ratio of the total variation of the observations which is not explained by the model, while R^2 is that ratio of the total variation of the observations which is explained by the model.

Since $\mathbf{1} \in V_r$, is $\mathbf{1}'\boldsymbol{P} = \mathbf{1}'$, with $\boldsymbol{P} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^*\boldsymbol{X}'$. Then it follows for the estimated y-values, $\hat{\boldsymbol{y}} = \boldsymbol{X}\hat{\boldsymbol{b}}$, that

$P1 = 1$

$$\sum \hat{y}_i = 1^T \hat{y} = \mathbf{1}'\boldsymbol{X}\hat{\boldsymbol{b}} = \mathbf{1}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^*\boldsymbol{X}'\boldsymbol{y} = \mathbf{1}'\boldsymbol{y} = \sum y_i$$

and

$$\sum_{i=1}^{n} \hat{y}_i = \sum_{i=1}^{n} y_i \text{ or } \bar{\hat{y}} = \bar{y}\,.$$

Consequently, the mean of the estimated y-values is the same as the mean y-value (provided that $\mathbf{1} \in V_r$).

Then it follows that: SSR − SSR$|_{H_0}$

$\hat{\mu} = \frac{1}{n}\sum y_i = \bar{y}$

$\hat{b}_0^T X_0^T y = \hat{\mu}(n\hat{\mu})$

$= n\hat{\mu}^2 = n\bar{y}^2$

$$SSH = \widehat{\boldsymbol{b}}'\boldsymbol{X}'\boldsymbol{X}\widehat{\boldsymbol{b}} - n\bar{\hat{y}}^2 = \widehat{\boldsymbol{b}}'\boldsymbol{X}'\boldsymbol{y} - n\bar{y}^2$$
$$= \sum_{i=1}^{n} \hat{y}_i^2 - n\bar{\hat{y}}^2 = \sum_{i=1}^{n} \hat{y}_i y_i - n\bar{y}\bar{\hat{y}}$$
$$= \sum_{i=1}^{n} \left(\hat{y}_i - \bar{\hat{y}}\right)^2 = \sum_{i=1}^{n} \left(\hat{y}_i - \bar{\hat{y}}\right)(y_i - \bar{y}). \quad (*)$$

Consequently, the coefficient of determination is

$$R^2 = \frac{SSH}{SSC}$$
$$= \frac{\sum_{i=1}^{n} \left(\hat{y}_i - \bar{\hat{y}}\right)(y_i - \bar{y})}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \times \frac{\sum(\hat{y}_i - \bar{\hat{y}})^2}{\sum(\hat{y}_i - \bar{\hat{y}})^2}$$
$$\overset{(*)}{=} \frac{\left\{\sum_{i=1}^{n} \left(\hat{y}_i - \bar{\hat{y}}\right)(y_i - \bar{y})\right\}^2}{\left\{\sum_{i=1}^{n} (y_i - \bar{y})^2\right\}\left\{\sum_{i=1}^{n} \left(\hat{y}_i - \bar{\hat{y}}\right)^2\right\}}$$

i.e. the square of the correlation coefficient between the y-values and the estimated y-values.

$CD = CC^2$

$$\left(\frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}\right)^2 = \frac{S_{xy}^2}{S_{xx}S_{yy}}$$