pdf

$\hat{}$ 0.140880 sd

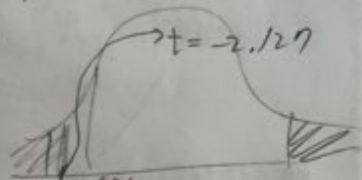
$\hat{\beta}_1 - \hat{\beta}_2 - \hat{\beta}_3 + \hat{\beta}$

#1.

$F_{2, n-p-1}$

$n-2-1 = 99$

$\hookrightarrow n = 102$



$t = -2.127$

0.98 0.998

$t_{n-p-1} \Rightarrow 99$ as well

① $\hookrightarrow$ p-value will be $0.02 - 0.08$

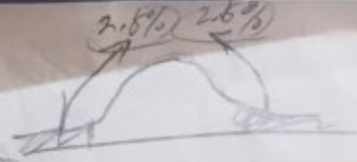
③ $RSE = \sqrt{\dfrac{RSS}{n-p-1}}$

$= \sqrt{\dfrac{\sum \hat{e}_i^2}{n-p-1}}$

$\boxed{99} \rightarrow$ (?) I'm curious why RSE has d.f. even though it is not test statistic?

② $\dfrac{\hat{\beta} - \beta}{SE(\hat{\beta})} = 11.858$ under the $H_0: \beta = 0$.

$\hookrightarrow \boxed{\hat{\beta} = (11.858)(0.3489120)} \approx \boxed{4.137}$



2.5% 2.5%

#2.

① $\dfrac{\hat{\beta}_0 - \beta_0}{s.e.(\hat{\beta}_0)} \sim t(0.025)$

$95\% = 1 - 2P\left(\dfrac{\hat{\beta}_0 - \beta_0}{SE(\hat{\beta}_0)} \geq t(0.025)\right)$

$\hookrightarrow \hat{\beta}_0 - \beta_0 \sim SE(\hat{\beta}_0) \cdot t(0.025)$

$\hookrightarrow \beta_0 \sim \hat{\beta}_0 \pm SE(\hat{\beta}_0) \cdot t(0.025)$

β: $-6.8477 \pm 3.2189 \cdot 1.984$

$4.137 \pm 0.3489 \cdot 1.984$

$0.00136 \pm 0.00022 \cdot 1.984$

(?) And why $TSS_0 = TSS_1$?

$\sum (y_i - \bar{y})^2$

$R^2 = 1 - \dfrac{RSS}{TSS}$ $(M) SS(m) - RegSS(m) - TSS(M) + RegSS(M)$

$RSS(m) - RSS(M)$

$1 - R^2 = \dfrac{RSS}{TSS}$

$TSS = RSS + RegSS$

$RSS(m)$

$\sum (y_i - \bar{y})$

$F = \dfrac{R^2 / p}{(1 - R^2)/n-p-1}$ $\dfrac{RSS(M)/p}{RSS(M)/n-p-1}$

$RSS0 - RSS1$

$RSS1 - RegSS0$

(?) So for lm() $F = \dfrac{RegSS(M)/p}{RSS(M)/n-p-1}$

but this is not equal to $\dfrac{RSS(m)/p}{RSS(M)/n-p-1}$.

#3.

1) $\begin{bmatrix}\beta_0\\ \beta_1\end{bmatrix} = \Lambda^T\beta = \rho^T X\beta$ for Λ, $\beta \in R^p$ for some matrix ρ $\in R^{n\times p}$

False

2) ⇒ contrapositive

↳ If β_1 is estimable, then (β_0, β_1) is estimable.

$\boxed{\beta_1 = \Lambda^T\beta = \rho^T X\beta}$ ⇒ satisfied (Assumption)

$\begin{bmatrix}\beta_0\\ \beta_1\end{bmatrix}$

So, $\Lambda = \begin{bmatrix}0\\1\end{bmatrix}$ $\quad [\rho_1 \cdots \rho_n]\begin{bmatrix}x_{11} & x_{12}\\ \vdots & \\ x_{n1} & x_{n2}\end{bmatrix}\begin{bmatrix}\beta_0\\ \beta_1\end{bmatrix}$

$\to [\rho_1 \cdots \rho_n]\begin{bmatrix}\beta_0 x_{11} + \beta_1 x_{12}\\ \vdots\\ \beta_0 x_{n1} + \beta_1 x_{n2}\end{bmatrix}$

$\begin{cases}\beta_0 x_{11}\rho_1 + \cdots + \beta_0 x_{n1}\rho_n = 0\\ \beta_1 x_{12}\rho_1 + \cdots + \beta_1 x_{n2}\rho_n = 1\end{cases}$

① when β_1 is estimable

$\begin{cases}\beta_0(\underbrace{x_{11}\rho_1 + \cdots + x_{n1}\rho_n}_{=0}) = 0, \text{ for any } \beta_0\\ \beta_1(x_{12}\rho_1 + \cdots + x_{n2}\rho_n) = 1, \text{ for any } \beta_1\end{cases}$

$\begin{bmatrix}\beta_0\\ \beta_1\end{bmatrix} = \underset{2\times 2}{\Lambda^T}\underset{2\times 1}{\begin{bmatrix}\beta_0\\ \beta_1\end{bmatrix}} = \underset{2\times n}{\rho^T}\underset{n\times 2}{X}\underset{2\times 1}{\beta}$

$\Lambda^T = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} = [\rho^T][X][\beta]$

$\begin{bmatrix}\beta_0\\ \beta_1\end{bmatrix} = [\rho^T]\begin{bmatrix}x_{11} & x_{12}\\ \vdots & \\ x_{n1} & x_{n2}\end{bmatrix}\begin{bmatrix}\beta_0\\ \beta_1\end{bmatrix} =$

$= [\rho^T]\begin{bmatrix}\beta_0 x_{11} + \beta_1 x_{12}\\ \vdots\\ \beta_0 x_{n1} + \beta_1 x_{n2}\end{bmatrix}$

$\begin{bmatrix}\rho_{11} & \cdots & \rho_{1n}\\ \rho_{21} & \cdots & \rho_{2n}\end{bmatrix}$

② when (β_0, β_1) is estimable

$\rho_{11}(\beta_0 x_{11} + \beta_1 x_{12}) + \cdots + \rho_{1n}(\beta_0 x_{n1} + \beta_1 x_{n2}) = \beta_0$

↳ $\begin{cases}\beta_0(x_{11}\rho_{11} + \cdots + x_{n1}\rho_{1n}) = \beta_0 \to = 1\\ \beta_1(x_{12}\rho_{11} + \cdots + x_{n2}\rho_{1n}) = 0 \to = 0\end{cases}$

$\rho_{21}(\beta_0 x_{11} + \beta_1 x_{12}) + \cdots + \rho_{2n}(\beta_0 x_{n1} + \beta_1 x_{n2}) = \beta_1$

↳ $\begin{cases}\beta_0(\underbrace{x_{11}\rho_{21} + \cdots + x_{n1}\rho_{2n}}_{=0}) = 0\\ \beta_1(\underbrace{x_{12}\rho_{21} + \cdots + x_{n2}\rho_{2n}}_{=1}) = \beta_1\end{cases}$

But for (β_0, β_1) is estimable

$\begin{cases}x_{11}\rho_{11} + \cdots + x_{n1}\rho_{1n} = 1\\ x_{11}\rho_{21} + \cdots + x_{n1}\rho_{2n} = 0\end{cases}$ is given!!

let's say $x_{11} = \cdots = x_{n1} = 0$, then the $x_{11}\rho_{11} + \cdots + x_{n1}\rho_{1n}$ cannot be 1

Counter example

⇒ As we assumed β_1 is estimable, for any β_0 and β_1, $x_{11}\rho_1 + \cdots + x_{n1}\rho_n$ should be 0 always. That's the condition I got from β_1 is estimable.

b)

If errors are correlated
$e_i \perp\!\!\!\perp e_j$, for $i \neq j$

$$E(\hat\beta) = E((X^TX)^{-1}X^TY)$$
$$= (X^TX)^{-1}X^T E(Y)$$
$$= (X^TX)^{-1}X^T E(X\beta + e)$$
$$= (X^TX)^{-1}X^TX\beta + (X^TX)^{-1}X^T E(e)$$
$$= \beta + (X^TX)^{-1}X^T E(e)$$

As long as $E(e)=0$, correlation does not matter...???

(?) Is it correct?

False

#4 (?) Need to check my work.

$$\hat\mu = \arg\min_{\mu} \|y - X\mu\|$$

$$S(\mu) = \sum_{i=1}^{J} (y_i - \mu_1 x_{1i} - \cdots - \mu_J x_{Ji})^2$$

$$= \sum (Y_i - x_i^T\mu)^2 = \|y - X\mu\|^2$$

$$\Rightarrow \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_J \end{bmatrix} = \begin{bmatrix} x_{11} & \cdots & x_{J1} \\ x_{12} & \cdots & x_{J2} \\ \vdots & & \vdots \\ x_{1J} & \cdots & x_{JJ} \end{bmatrix} \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_J \end{bmatrix} + \begin{bmatrix} e_1 \\ \vdots \\ e_J \end{bmatrix}$$

$$\sum_i e_i^2 \Leftarrow e = y - X\mu \rightarrow (?)$$

$$\|e\|^2 = e^Te = (y - X\mu)^T(y - X\mu)$$
$$= (y^T - (X\mu)^T)(y - X\mu)$$
$$= y^Ty - y^T(X\mu) - \mu^TX^Ty + (X\mu)^T(X\mu) = S(\mu)$$

$$\Rightarrow \frac{dS(\mu)}{d\mu} = \frac{d[y^Ty - 2\mu^TX^Ty + \mu^TX^TX\mu]}{d\mu}$$
$$= -2X^Ty + 2X^TX\mu = 0$$
$$\boxed{X^TX\hat\mu = X^Ty}$$

(?) This only holds for Real vector space. Then what will be OLS estimator for complex?

So, normal equation still holds and I can get the OLS least square estimator $\hat\mu$ even if we don't assume noise is normally distributed?

But the problem is that estimator is the best unbiased estimator if we don't assume normality. Right?

$X^T X \hat{\mu} = X^T Y$

so no intercept !!!

$$X = \begin{bmatrix} X_{11} & \cdots & X_{J1} \\ X_{12} & & \vdots \\ \vdots & & \vdots \\ X_{1J} & \cdots & X_{JJ} \end{bmatrix}$$

$$X^T X = \underset{(J \times J)}{\begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1J} \\ \vdots & & & \vdots \\ X_{J1} & \cdots & & X_{JJ} \end{bmatrix}} \underset{(J \times J)}{\begin{bmatrix} X_{11} & X_{21} & \cdots & X_{J1} \\ \vdots & & & \vdots \\ X_{1J} & X_{2J} & \cdots & X_{JJ} \end{bmatrix}}$$

$$= \begin{bmatrix} \Sigma X_{1i}^2 & \cdots \\ & \\ & \end{bmatrix}$$

Assume this.
- Each i belongs $\in$ ith group for $1 \le i \le J$.

$$X = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & 0 & 1 \end{bmatrix} = I_{J \times J}$$

$\Rightarrow$

$X^T X = I_{J \times J} \cdot I_{J \times J} = I_{J \times J}$

$I_{J \times J}^{-1} = I_{J \times J}$ (X is a full rank)

$\hat{\mu} = (X^T X)^{-1} X^T Y$

$= I \cdot I \cdot Y$

$$= \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} \begin{bmatrix} Y_1 \\ \vdots \\ Y_J \end{bmatrix} = \begin{bmatrix} \hat{\mu}_1 \\ \vdots \\ \hat{\mu}_J \end{bmatrix}$$

proved for one example!!

$\hat{\mu}_1 = \underbrace{1}_{\frac{1}{n_j}} \cdot (Y_1 \cdot 1 + Y_2 \cdot 0 + \cdots + Y_J \cdot 0) = \sum_{i \in j\text{th group}} Y_i$

$\frac{1}{n_j}$ as # there is only one observation in group j

$$X^T y = \begin{bmatrix} X_{11} & \cdots & X_{1J} \\ \vdots & & \vdots \\ X_{J1} & \cdots & X_{JJ} \end{bmatrix} \begin{bmatrix} Y_1 \\ \vdots \\ Y_J \end{bmatrix} = \begin{bmatrix} X_{11} \cdot Y_1 + \cdots + X_{1J} \cdot Y_J \\ \vdots \\ X_{J1} \cdot Y_1 + \cdots + X_{JJ} \cdot Y_J \end{bmatrix}$$

$$\begin{bmatrix} Y_1 \\ \vdots \\ Y_J \end{bmatrix} = \begin{bmatrix} X_{11} & \cdots & X_{J1} \\ \vdots & & \vdots \\ X_{1J} & \cdots & X_{JJ} \end{bmatrix} \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_J \end{bmatrix} + \begin{bmatrix} e_1 \\ \vdots \\ e_J \end{bmatrix}$$

each row should have 1 at max of one. They cannot have two 1s in the row. $\Rightarrow$ each observation will belong $\in$ at most one group

Let's say we have another design matrix

$$X = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 1 \\ \vdots \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

1 ∈ 2nd group
2 ∈ 4th group
3 ∈ 3rd group
4 ∈ 1st group

⇒ each row has only one 1.

$X^T X \hat{\mu} = X^T y$

↳ normal equation

$$X^T = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & 0 & \cdots & 0 \\ \vdots \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \quad \cancel{X = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & \cdots \\ 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}}$$

↳ so, each column has only one 1.

↳ By seeing the pattern (and we know X^TX is symmetric), X^TX is identity matrix always.

so, $\boxed{\hat{\mu} = X^T y}$ always

~~then~~ ~~what does~~

What if some rows belong to no group, or when we have more than two rows have 1 in the same column ?? (X full rank)

$$X = \begin{bmatrix} 0 & \cdots & 0 \\ 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 1 \\ 0 & 0 & 1 & 0 & \cdots \\ 0 & \cdots \end{bmatrix}$$

still, X^TX is symmetric also ~~positive~~ non-negative, non-diagonal

so, this is the key-point.

~~if we have n_j rows~~

Let's say n_j rows that belongs to the group j, then X^TX will have n_j on the diagonal (and we always have a diagonal ~~matrix~~)

↳ so, when we have only one non-zero entry in each row, it will be diagonal matrix (diagonal mat ⊂ symmetric mat)

$$\boxed{\text{so, } (n_j)\cdot(\hat{\mu}_j) = \sum_{i \in j\text{th}} y_j}$$

Then, if $\begin{bmatrix} \hat{\mu}_4 \\ \hat{\mu}_1 \\ \hat{\mu}_3 \\ \vdots \end{bmatrix} = \begin{bmatrix} y_4 \\ y_1 \\ y_3 \\ y_2 \\ \vdots \\ y_J \end{bmatrix}$

#5.

$\boxed{\langle 5.3 \rangle}$

$\sum (y_i - \bar{y})^2 \Rightarrow TSS$

$\sum_{i=1}^{n} (y_i - A')^2 = (y_1 - A')^2 + (y_2 - A')^2 + \cdots (y_n - A')^2$

$\frac{dS(A')}{dA'} = -2(y_1 - A') - 2(y_2 - A') \cdots - 2(y_n - A') = 0$

$\hookrightarrow (y_1 + \cdots + y_n) = A' + \cdots + A' = nA'$

$\frac{d^2S(A')}{d(A')^2} = 2 + 2 + \cdots + 2 > 0$

$A' = \frac{1}{n} \sum y_i = \bar{y}$

$= (y - A')^T (y - A') = (y^T - (A')^T)(y - A')$

fixed

$= y^T y - y^T A' - (A')^T y + (A')^T A'$

$= y^T y - 2y^T A' + \cancel{(A')^T A'}$

$d\,S(A')$

hard time differentiate

$\sum (y_i - \bar{y})^2 = \sum (y_i - \hat{y}_i)^2 + \sum (\hat{y}_i - \bar{y})^2$

② $X' = 10X$ ⇒ Both slope & intercept will change!!!

$y = \alpha + \beta(10X)$

$= \alpha + 10\beta X$

② $X' = 10X$ ⇒ Both slope & intercept will change!!!

$y = \alpha + \beta \cdot (10X)$

$y = \alpha + 10\beta X$

$\frac{1}{10}y - \frac{1}{10}\alpha = \beta X$

~~RSS, Regss & TSS might change?~~

~~So TSS will be the same~~

~~Regss might change?~~

$\hookrightarrow y = \alpha + (10\beta)X$

↳ Think $10\beta = B'$ and $\alpha = A' = A$.

$\sum(y_i - \bar{y})^2 = \sum(y_i - \hat{y}_i)^2 + \sum(\hat{y}_i - \bar{y})^2$

① y_i & $\bar{y}$ & $\hat{y}_i$ change

None of

② y_i does change

& $\bar{y}$

$\hat{y}_i$ ~~change~~

None of

③ y_i, $\hat{y}_i$, $\bar{y}$ change.

None of

<5.4>

$\frac{S_y}{S_x} r = \hat{B}$

$Y = A + BX$

$S_E = \sqrt{\frac{RSS}{n-p-1}} = \sqrt{\frac{RSS}{n-2}}$

$r = \frac{S_{xy}}{S_x S_y} = \hat{B}\frac{S_x}{S_y}$

a) $Y = \alpha + \beta(x - 10)$ ~~little bit confused....~~

different so in b) — ② ③

$B' = \frac{1}{10}B$

$A' = A$

① $Y + 10\beta = \alpha + \beta X$

$Y = (\alpha - 10\beta) + \beta X$

$\alpha - 10\beta = A'$

$A' - 10\beta = A$

$A' = A + 10B$ → correct (checked in R)

$B' = B$

$S_E' = S_E$ → (?) I can intuitively get it, since data just moved to the left so we should get $y - \hat{y} = \hat{e}$ should be the same, but can it be the logical reasoning for this?

$r' = r$ → (?) so, I can say $r^2 = R^2$, and $R^2 = 1 - \frac{RSS}{TSS}$, and I know RSS stays the same and $TSS = \sum(y_i - \bar{y})^2$ does not change...?

(?) when does S_E and r change?

(?) Does TSS, Regss, RSS ever change if we make either X or Y transformation? what if both X and Y transform? → cuz S_E is a function of RSS, and r is a function of RSS and TSS?

$S_E' = S_E$

$r' = r$

(?) What's the best way to think of these logically? I solved it intuitively...

$R^2 = \frac{TSS - RSS}{TSS} = 1 - \frac{RSS}{TSS} \Rightarrow r^2$

$\sum(y_i - \bar{y})^2$

$\frac{Cov(X,Y)}{sd(x)sd(y)}$

→ function of RSS and TSS!!

Since $y = \alpha + \beta X$ we think $\beta = B$ then $A' - 10B = A$, so $\boxed{A' = A + 10B}$

③ $X' = 10(X-1) = 10X - 10$

$y = \alpha + \beta X'$

$y = \alpha + \beta(10X - 10)$

$y = \alpha + (10\beta)X - 10\beta$

$y = (\alpha - 10\beta) + (10\beta)X$

So, $10B' = B$

$A' - 10B' = A \rightarrow A' = A + 10B'$

$\rightarrow A' = A + B$

Good way to approach it.

$A' = A + B$

$B' = \frac{1}{10}B$ → correct

$SE' = SE$

$r' = r$

③ $Y'' = 5(Y+2) = 5Y + 10$

$5Y + 10 = A'' + B''X$

$\hookrightarrow 5Y = (A'' - 10) + B''X$

$\hookrightarrow Y = \left(\frac{A'' - 10}{5}\right) + \left(\frac{B''}{5}\right)X$

$B = \frac{B''}{5} \rightarrow \boxed{B'' = 5B}$ → correct

$A = \frac{A'' - 10}{5} \rightarrow \boxed{A'' = 5A + 10}$

$\boxed{r'' = r}$

$\boxed{S_E'' \approx 5SE}$

$\star$ I found the pattern that if $y'' = aY$ then $SE'' = a \cdot SE$ (?) but don't know why…?

8.00183911

b) ① $Y'' = Y + 10$

$(Y + 10) = A'' + B''X$

$\hookrightarrow Y = (A'' - 10) + B''X$

$B'' = B$, as Y regress on X (as X does not transformed)

$A'' - 10 = A$, as Y regress on X, so intercept should be the same as the original.

$\boxed{A'' = A + 10}$ → correct

$\boxed{B'' = B}$

$\boxed{SE'' = SE}$

$\boxed{r'' = r}$

② $Y'' = 5Y$

$5Y = A'' + B''X$

$\hookrightarrow y = \frac{1}{5}A'' + \left(\frac{1}{5}B''\right)X$

$\frac{1}{5}B'' = B \rightarrow \boxed{B'' = 5B}$

$\frac{1}{5}A'' = A \rightarrow \boxed{A'' = 5A}$

correct

Regss'' =

RSS'' =

$SE'' = \sqrt{\frac{RSS}{n-p-1}} \approx 5SE$

$\boxed{r'' = r}$

(?) Different… but don't know why…

C) <sum>

a) ① $A' = A + 1 \cdot B$, $B' = B$, $SE' = SE$, $r' = r$

③ $A' = A + B$, $B' = 1/_{10}B$, $SE' = SE$, $r' = r$

② $A' = A$, $B' = 1/_{10}B$, $SE' = SE$, $r' = r$

b) ① $A'' = A + 10$, $B'' = B$, $SE'' = SE$, $r'' = r$

③ $A'' = 5A + 10$, $B'' = 5B$, $SE'' \approx 5SE$, $r'' = r$

② $A'' = 5A$, $B'' = 5B$, $SE'' \approx 5SE$, $r'' = r$

whenever ~~[illegible]~~ X' can be factored into X − α, then $\boxed{A' = A + \alpha B}$ and coefficient of that factor goes to reciprocal of B. SE & r X changed due to the transformation of X.

A'' follows exactly the transformation of Y. X'' multiplies by the coefficient of Y. S.E'' follows the coefficient of Y when it's get mul

$\boxed{6.6}$

a) $Y^{*} = A' + B'(10x - 10)$

$Y^{*} = (A' - 10B') + (10B')X$

$A' - 10B' = A$

$10B' = B$

$\hookrightarrow B' = \frac{1}{10}B$

$A' = A + B$

$SE(B') = SE(\frac{1}{10}B) = \frac{1}{10}SE(B)$

$SE(A') = SE(A+B) = \sqrt{Var(A+B)}$

$t'_0 = \frac{B'}{SE(B')} = \frac{\frac{1}{10}B}{\frac{1}{10}SE(B)} = \frac{B}{SE(B)} = t_0$

b) $Y'' = A'' + B''X$

$5Y + 10 = A'' + B''X$

$\hookrightarrow 5Y = (A'' - 10) + B''X$

$\hookrightarrow Y = \frac{A'' - 10}{5} + \frac{B''}{5}X$

$B = \frac{B''}{5} \rightarrow \boxed{B'' = 5B}$

$A = \frac{A'' - 10}{5} \rightarrow \boxed{5A + 10 = A''}$

$SE(B'') = SE(5B) = \boxed{5\,SE(B)}$

$t''_0 = \frac{B''}{SE(B'')} = \frac{5B}{SE(5B)} = \frac{B}{SE(B)} = t_0$

(?) Is this what they want?

☆) Although we make linear combination of X and Y (add & multiplication), t-test for (B_1 or) B with B' (transformed) is alway equal !!!!

Also, when we construct CI for (B_1 or) B, the transformed variables standard derivation will be inverse proportional of factor coefficients of transformation of X and proportional of transformation of Y.

☆ Intuitively, when I transform X by $\alpha > 1$, then the data gets wider, and standard error of it becomes relatively small compared to original. And, when I extend Y by $\alpha > 1$, it will be other way around.

$\langle 9.8 \rangle$

$$t \sim \frac{Z}{\sqrt{\frac{V}{d}}}$$

$$\frac{X^2_{n-p-1}}{n-p-1} \quad \frac{RSS}{\sigma^2}$$

$$\frac{\hat{\beta}_j - \beta_j}{SE(\hat{\beta}_j)} = \frac{\hat{\beta}_j - \beta_j}{\sqrt{\sigma^2 (X^TX)^{-1}_{jj}}} = \frac{\hat{\beta}_j - \beta_j}{\sigma \sqrt{(X^TX)^{-1}_{jj}}}$$

$$\approx \frac{\hat{\beta}_j - \beta_j}{S_E \sqrt{(X^TX)^{-1}_{jj}}}$$

$\hat{\beta} \sim N(\beta, \sigma^2 (X^TX)^{-1})$

$\hat{e} \sim N(0, \sigma^2 I_{n \times n})$

$\sqrt{V_{jj}} \Rightarrow$ constant

$$\hat{\beta} = \begin{bmatrix} \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_n \end{bmatrix} \quad S_E = \sqrt{\frac{e^Te}{n-k-1}} \to RSS$$

Trying to prove they are ind

$$cov(X,Y) = E[(X-E(X))(Y-E(Y))^T] = E(XY^T) - E(X)E(Y)^T$$

$cov(Y,X)^T$

$$cov(\hat{e}, \hat{\beta}) = E(\hat{e} - E(\hat{e}))\, E(\hat{\beta} - E(\hat{\beta}))$$
$$= E(\hat{e}) - E(\hat{e})$$

$$cov(\hat{e}, \hat{\beta}) = E[\hat{e}(\hat{\beta} - \beta)^T]$$

why is that true?

$$\hat{\beta} - \beta = (X^TX)^{-1}X^Ty - \beta = (X^TX)^{-1}X^T(X\beta + e) - \beta$$
$$= \beta + (X^TX)^{-1}X^Te - \beta$$
$$= (X^TX)^{-1}X^Te$$

$$cov(\hat{e}, \hat{\beta}) = E[\hat{e}\, [(X^TX)^{-1}X^Te]^T]$$
$$= E[\hat{e}\, e^TX(X^TX)^{-1}]$$
$$= E[(y - X(X^TX)^{-1}X^Ty)e^TX(X^TX)^{-1}]$$
$$= E[ye^TX(X^TX)^{-1} - X(X^TX)^{-1}X^Tye^TX(X^TX)^{-1}]$$
$$= X(X^TX)^{-1}E(ye^T) - X(X^TX)^{-1}X^TE(e^Ty)X(X^TX)^{-1}$$
$$= X(X^TX)^{-1}E(ye^T) - X(X^TX)^{-1}X^TX(X^TX)^{-1}E(e^Ty)$$
$$= X(X^TX)^{-1}E(ye^T) - X(X^TX)^{-1}E(e^Ty)$$
$$= X(X^TX)^{-1}E(ye^T) - X(X^TX)^{-1}E(e^Ty)$$
$$= X(X^TX)^{-1}E[ye^T - e^Ty]$$
$$= X(X^TX)^{-1} \cdot 0$$
$$= 0$$

(?)

$E[B\Sigma C] = BE[\Sigma]C = BCE[\Sigma]$ where B and C are fixed and Σ are random?

Can we say that?

$\hat{\beta}$ and $\hat{e}$ are jointly normally distributed as $\hat{\beta} = (X^TX)^{-1}X^Ty$, $(X^TX)^{-1}X^T(\hat{y} + \hat{e})$

(?) Why if two vectors are normally distributed, it is enough to prove their covariance is zero to prove they are independence?

$\to f_{XY} = f_X f_Y$ by Bivariate normal

$cov(X,Y) = 0 \not\longrightarrow X \perp Y$

$cov(X,Y) = 0 \longleftarrow$

but for normal dist, uncorrelatedness implies independence

(?) only true for Real space?

Mathematics/Statistics Circ Desk

① <version 1>

β_1 is estimable $\Rightarrow$

$$\begin{cases} \beta_0 (p_1 + \cdots + p_n) = 0 \\ \beta_1 (p_1 X_{12} + \cdots + X_{n2} p_n) = \beta_1 \end{cases}$$

so, from this result, I can say that $X_{12}, \ldots, X_{n2}$ should not have the same values.

s.t. $\begin{bmatrix} 1 & 2 \\ \vdots & 2 \\ \vdots & \vdots \\ 1 & 2 \end{bmatrix}$.

Now check, $\begin{bmatrix} \beta_0 & \beta_1 \end{bmatrix}$ is estimable.

① $\begin{cases} \beta_0 (\underbrace{p_{11} + \cdots + p_{1n}}_{=1}) = \beta_0 \\ \beta_1 (\underbrace{p_{11} X_{12} + \cdots + p_{1n} X_{n2}}_{=0}) = 0 \end{cases}$

② $\begin{cases} \beta_0 (\underbrace{p_{21} + \cdots + p_{2n}}_{=0}) = 0 \\ \beta_1 (\underbrace{p_{21} X_{12} + \cdots + p_{2n} X_{n2}}_{=1}) = \beta_1 \end{cases}$

4 conditions

by saying $\begin{cases} p_{21} = -p_{22} - \cdots - p_{2n} \\ p_{11} = 1 - p_{12} - \cdots - p_{1n} \end{cases}$

I can get

$$\begin{bmatrix} p_{12}(X_{22} - X_{12}) \\ + \\ \vdots \\ p_{1n}(X_{n2} - X_{12}) \\ + \\ X_{12} \end{bmatrix} = 0$$

$$\begin{bmatrix} p_{22}(X_{22} - X_{12}) \\ + \\ \vdots \\ + \\ p_{2n}(X_{n2} - X_{12}) \end{bmatrix} = 1$$

As we proved, $X_{12}, \ldots, X_{n2}$ should have the same value

So one of elements in ① and ② should not be 0. Thus, ② version 2

$\Lambda^T \beta = p^T X \beta$ for all β

so, $\Lambda^T = p^T X$

$\Lambda = X^T p$

1) ~~Prove~~ β_1 is estimable

$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is in $C\left(\begin{bmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_n \end{bmatrix}\right)$.

so, $x_1, \cdots, x_n$ should not have the same constant.

2) prove $\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$ is estimable.

If $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ or $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is not in $C\left(\begin{bmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_n \end{bmatrix}\right)$

then $\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$ is not estimable.

And, since $x_1, \cdots, x_n$ ~~are not~~ are not all the same as we proved,

two column should be lin ind.

so, $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ are in the $C\left(\begin{bmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_n \end{bmatrix}\right)$