$(X\hat{\beta})^T e = \hat{\beta}^T X^T e = \hat{\beta}^T X^T (y - X\hat{\beta}) = \hat{\beta}^T X^T y - \hat{\beta}^T X^T X \hat{\beta}$

$= \hat{\beta}^T (X^T X)(X^T X)^{-1} X^T y - \hat{\beta}^T X^T X \hat{\beta}$

$= \hat{\beta}^T (X^T X)\hat{\beta} - \hat{\beta}^T X^T X \hat{\beta} = 0$

$\sum Y_i = \sum \hat{Y}_i$, $\sum \hat{e}_i = 0$, $\sum X_i \hat{e}_i = 0$.

⓪ $\sum \hat{e}_i \neq 0$, $\sum X_i \hat{e}_i = 0$

$\sum \hat{Y}_i = \sum(\hat{a} + \hat{b}X_i) = \sum(\bar{Y} - \hat{b}\bar{X} + \hat{b}X_i) = \sum Y_i - \hat{b}n\bar{X} + \hat{b}\sum X_i = \sum Y_i$

① $\sum Y_i = \sum \hat{Y}_i$, $\sum \hat{e}_i = 0$, $\sum X_i \hat{e}_i = 0$

$\sum X_i \hat{e}_i = \sum X_i (Y_i - \hat{Y}_i) = \sum X_i (Y_i - \hat{a} - \hat{b}X_i) = \sum (X_i - \bar{X})[(Y_i - \bar{Y}) - \hat{b}(X_i - \bar{X})]$

$= \sum (X_i - \bar{X})(Y_i - \bar{Y}) - \frac{\sum (Y_i - \bar{Y})(X_i - \bar{X})}{\sum (X_i - \bar{X})^2}\sum (X_i - \bar{X})^2 = 0$

$\sum (Y_i - \hat{Y}_i)(\hat{Y}_i - \bar{Y}) = \sum (Y_i - \hat{Y}_i)\hat{Y}_i - \sum (Y_i - \hat{Y}_i)\bar{Y} = \sum \hat{e}_i(\hat{a} + \hat{b}X_i) - \sum \hat{e}_i \bar{Y} = \hat{a}\sum \hat{e}_i + \hat{b}\sum \hat{e}_i X_i - \bar{Y}\sum \hat{e}_i = 0$

② $\sum Y_i = \sum \hat{Y}_i$, $\sum \hat{e}_i = 0$, $\sum X_{1i}\hat{e}_i = 0$, $\sum X_{2i}\hat{e}_i = 0$

$\bar{Y} - \hat{b}_1 \bar{X}_1 - \hat{b}_2 \bar{X}_2$

$\sum (X_{1i} - \bar{X}_1)\hat{e}_i = \sum (X_{1i} - \bar{X}_1)[Y_i - \hat{a} - \hat{b}_1 X_{1i} - \hat{b}_2 X_{2i}]$

$= \sum (X_{1i} - \bar{X}_1)[(Y_i - \bar{Y}) - \hat{b}_1(X_{1i} - \bar{X}_1) - \hat{b}_2(X_{2i} - \bar{X}_2)]$

$= \sum (X_{1i} - \bar{X}_1)(Y_i - \bar{Y}) - \hat{b}_1 \sum (X_{1i} - \bar{X}_1)(X_{1i} - \bar{X}_1) - \hat{b}_2 \sum (X_{1i} - \bar{X}_1)(X_{2i} - \bar{X}_2)$

$= S_{1y} - \hat{b}_1 S_{11} - \hat{b}_2 S_{12}$

$= S_{1y} - S_{1y} = 0$.

③ $\sum Y_i = \sum \hat{Y}_i$, $\sum \hat{e}_i = 0$, $\sum X_{ji}\hat{e}_i = 0$

$SS(\beta) = \sum (Y_i - \beta_0 - \beta_1 X_i - \dots - \beta_k X_k)^2$

$\frac{\partial SS(\beta)}{\partial \beta_0} = -\sum 2(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_1 - \dots - \hat{\beta}_k X_k) = 0 \Rightarrow \sum e_i = 0$

$SS(\beta) = (Y - X\beta)^T (Y - X\beta) = \beta^T X^T X \beta - 2Y^T X\beta + Y^T Y$

$\frac{\partial SS(\beta)}{\partial \beta}\Big|_{\beta = \hat{\beta}} = 2X^T X\hat{\beta} - 2X^T Y = 0$ $\quad X^T \hat{e} = X^T(Y - X\hat{\beta}) = 0$

$SS_{tot} = SS_{res} + SS_{reg}$

$SS_{tot} = \sum (Y_i - \bar{Y})^2 = \sum (Y_i - \hat{Y}_i + \hat{Y}_i - \bar{Y})^2$

$= \sum (Y_i - \hat{Y}_i)^2 + \sum (\hat{Y}_i - \bar{Y})^2 + 2\sum (Y_i - \hat{Y}_i)(\hat{Y}_i - \bar{Y})$

$= SS_{res} + SS_{reg}$

$\sum (Y_i - \hat{Y}_i)\hat{Y}_i = \sum \hat{e}_i(\hat{a} + \hat{b}X_i) = \hat{a}\sum e_i + \hat{b}\sum e_i X_i = 0$.

$\sum (Y_i - \hat{Y}_i)\bar{Y} = \bar{Y}\sum \hat{e}_i = 0$

$Y_i = a + bX_i + e_i$ $\quad a, b, X_i$ $\quad$ Residual standard error

$Y_i \sim N(a + bX_i, \sigma^2)$ $\quad Var(Y_i) = \sigma^2$, $Cov(Y_i, Y_j) = 0$

$Var(Y_i) = Var(e_i) = \sigma^2$

⓪ $S^2_{Y|0} = \frac{1}{n-1}\sum (Y_i - \hat{Y}_i)^2$ $\quad \epsilon \sim N(0, \sigma^2 I_{n\times n})$

② $S^2_{Y|X_1 \dots X_k} = \frac{1}{n-k-1}\sum (Y_i - \hat{Y}_i)^2$

$S_{Y|X}$ ① $\hat{\sigma}^2 = \frac{1}{n-2}\sum (Y_i - \hat{Y}_i)^2 = S^2_{Y|X}$. ② $S_{Y|X_1 X_2} = \sqrt{\frac{1}{n-3}\sum (Y_i - \hat{Y}_i)^2}$

$SS(b) = \sum (Y_i - bX_i)^2$, $\frac{\partial SS(b)}{\partial b} = -2\sum X_i(Y_i - bX_i) = 0$.

$\hat{b}$

① $\hat{b} = \frac{\sum (Y_i - \bar{Y})(X_i - \bar{X})}{\sum (X_i - \bar{X})^2} = \sum C_i(Y_i - \bar{Y}) = \sum C_i Y_i$, $C_i = \frac{X_i - \bar{X}}{\sum (X_i - \bar{X})^2}$.

⓪ $\hat{b} = \frac{\sum X_i Y_i}{\sum X_i^2}$, $\sum \frac{X_i}{\sum X_i^2} Y_i$

$E(\hat{b}) = b$, $Var(\hat{b}) = \frac{\sigma^2}{\sum X_i^2}$, $Var(\hat{b}) = \sum \left(\frac{X_i}{\sum X_i^2}\right)^2 \sigma^2$

$\frac{\hat{b} - b}{S_{Y|X}/\sqrt{\sum X_i^2}} \sim t_{n-1}$

$\sum C_i = 0$. $\quad \sum (X_i - \bar{X}) = \sum X_i - n\bar{X} = 0$

$\sum C_i X_i = 1$ $\quad \sum C_i X_i - \sum C_i (X_i - \bar{X}) = \frac{\sum (X_i - \bar{X})^2}{\sum (X_i - \bar{X})^2} = 1$

$\sum C_i^2 = \frac{1}{\sum (X_i - \bar{X})^2}$

$E(\hat{b}) = \sum C_i E(Y_i) = a\sum C_i + b\sum C_i X_i = b$

$Var(\hat{b}) = Var(\sum C_i Y_i) = \sum C_i^2 Var(Y_i) + \sum_{i \neq j} C_i C_j Cov(Y_i, Y_j)$

$= \sum C_i^2 Var(Y_i) = Var(Y_i)\sum C_i^2 = \frac{\sigma^2}{\sum (X_i - \bar{X})^2}$

$\frac{\hat{b} - b}{S_{Y|X}/\sqrt{\sum (X_i - \bar{X})^2}} \sim t_{n-2}$

$\hat{b} \sim N\left(b, \frac{\sigma^2}{\sum (X_i - \bar{X})^2}\right)$ $\quad \hat{b} \pm t_{n-2}\left(\frac{\alpha}{2}\right)\frac{S_{Y|X}}{\sqrt{\sum (X_i - \bar{X})^2}}$

② $b_1 = \frac{S_{1y}S_{22} - S_{2y}S_{12}}{S_{11}S_{22} - S_{12}^2} = \sum C_{1i} Y_i$ $\quad C_{1i} = \frac{S_{22}(X_{1i} - \bar{X}_1) - S_{12}(X_{2i} - \bar{X}_2)}{S_{11}S_{22} - S_{12}^2}$

$S_{12} = \sum (X_{1i} - \bar{X}_1)(X_{2i} - \bar{X}_2)$, $S_{1y} = \sum (X_{1i} - \bar{X}_1)(Y_i - \bar{Y}) = \sum (X_{1i} - \bar{X}_1) Y_i$, $S_{2y} = \sum (X_{2i} - \bar{X}_2) Y_i$

$b_2 = \frac{S_{2y}S_{11} - S_{1y}S_{12}}{S_{11}S_{22} - S_{12}^2} = \sum C_{2i} Y_i$ $\quad C_{2i} = \frac{S_{11}(X_{2i} - \bar{X}_2) - S_{12}(X_{1i} - \bar{X}_1)}{S_{11}S_{22} - S_{12}^2}$

$\sum C_{1i} = 0$. $\sum C_{2i} = 0$ $\quad \sum (X_{1i} - \bar{X}_1) = 0$

$\sum C_{1i} X_{1i} = 1$. $\sum C_{2i} X_{2i} = 1$ $\quad \sum C_{1i} X_{1i} = \sum C_{1i}(X_{1i} - \bar{X}_1) = \frac{\sum S_{22}(X_{1i} - \bar{X}_1)(X_{1i} - \bar{X}_1) - \sum S_{12}(X_{2i} - \bar{X}_2)(X_{1i} - \bar{X}_1)}{S_{11}S_{22} - S_{12}^2}$

$\sum C_{1i} X_{2i} = 0$ $\quad \sum C_{2i} X_{1i} = 0$ $\quad = \frac{S_{22}\sum (\dots) - S_{12}\sum (\dots)}{S_{11}S_{22} - S_{12}^2} = \frac{S_{11}S_{22} - S_{12}S_{12}}{S_{11}S_{22} - S_{12}^2} = 1$

$\sum C_{1i}^2 = \frac{S_{22}}{S_{11}S_{22} - S_{12}^2}$ $\quad \sum C_{2i}^2 = \frac{S_{11}}{S_{11}S_{22} - S_{12}^2}$

$\sum C_{1i} C_{2i} = -\frac{S_{12}}{S_{11}S_{22} - S_{12}^2}$

$se(\hat{b}_1) = S_{Y|X_1 X_2}\sqrt{\frac{S_{22}}{S_{11}S_{22} - S_{12}^2}}$, $t = \frac{\hat{b}_1}{se(\hat{b}_1)}$

$[\hat{b}_1 - se(\hat{b}_1)\cdot t_{n-3}(\frac{\alpha}{2}), \hat{b}_1 + se(\hat{b}_1)\cdot t_{n-3}(\frac{\alpha}{2})]$

$Var(\hat{b}_1) = \sigma^2 \sum C_{1i}^2 = \frac{S_{22}}{S_{11}S_{22} - S_{12}^2}\sigma^2$ $\quad Var(\hat{b}_2) = \sigma^2 \sum C_{2i}^2 = \frac{S_{11}}{S_{11}S_{22} - S_{12}^2}\sigma^2$.

$Cov(\hat{b}_1, \hat{b}_2) = \sum\sum C_{1i} C_{2j} Cov(Y_i, Y_j) = \sum C_{1i}C_{2i} Cov(Y_i, Y_i) + \sum_{i \neq j} C_{1i}C_{2j} Cov(Y_i, Y_j) = \sigma^2 \sum C_{1i} C_{2i} = \dots$

$r_{12} = \frac{S_{12}}{\sqrt{S_{11}S_{22}}}$

$\frac{\hat{b}_1 - b_1}{S_{Y|X_1 X_2}\sqrt{\frac{S_{22}}{S_{11}S_{22} - S_{12}^2}}} \sim t_{n-3}$ $\quad \frac{\hat{b}_2 - b_2}{S_{Y|X_1 X_2}\sqrt{\frac{S_{11}}{S_{11}S_{22} - S_{12}^2}}} \sim t_{n-3}$

$\hat{b}_1 = \frac{S_{1y}S_{22} - S_{2y}S_{12}}{S_{11}S_{22} - S_{12}^2} = \sqrt{\frac{S_{yy}}{S_{11}}}\frac{r_{1y} - r_{2y}r_{12}}{1 - r_{12}^2}$

$Var(\hat{b}_1) = \frac{S_{22}}{S_{11}S_{22} - S_{12}^2}\sigma^2 = \frac{\sigma^2}{S_{11}(1 - r_{12}^2)}$

③ $\hat{\beta} = (X^T X)^{-1} X^T Y$: $C = (X^T X)^{-1} X^T$

$\hat{b}_j = \sum C_{j+1, i} Y_i$ $\quad Var(\hat{b}_j) = \sigma^2 \sum C_{j+1, i}^2$

$\hat{\beta} = (X^T X)^{-1} X^T Y = (X^T X)^{-1} X^T (X\beta + \epsilon) = \beta + (X^T X)^{-1} X^T \epsilon \sim N(\beta, \sigma^2 (X^T X)^{-1})$

$\frac{\hat{b}_j - b_j}{S_{Y|X_1 \dots X_k}\sqrt{A_{j+1, j+1}}} \sim t_{n-k-1}$

$Cov(B\vec{Z}) = B\, Cov(\vec{Z})B^T$

$Cov(\hat{\beta}) = B\, Cov(Y)\, B^T$.

$\hat{a}$

a. ① $\hat{a} = \bar{Y} - \hat{b}\bar{X} = \sum \frac{1}{n}Y_i - \sum C_i Y_i \bar{X} = \sum d_i Y_i$ $\quad d_i = \frac{1}{n} - \bar{X}C_i$ $\quad \sum d_i = 1$

$E(\hat{a}) = \sum d_i (a + bX_i) = a + b\sum d_i X_i = a + b\bar{X} - b\bar{X}\sum C_i X_i = a$

$Var(\hat{a}) = \sum d_i^2 \sigma^2 = \left(\frac{1}{n} + \frac{\bar{X}^2}{\sum (X_i - \bar{X})^2}\right)\sigma^2$

$\hat{a} \sim N\left(a, \left[\frac{1}{n} + \frac{\bar{X}^2}{\sum (X_i - \bar{X})^2}\right]\sigma^2\right)$ $\quad \frac{\hat{a} - a}{S_{Y|X}\sqrt{\frac{1}{n} + \frac{\bar{X}^2}{\sum (X_i - \bar{X})^2}}} \sim t_{n-2}$. $\quad \hat{a} \pm t_{n-2}(\frac{\alpha}{2})\cdot se(\hat{a})$

$Cov(\hat{a}, \hat{b}) = Cov(\bar{Y} - \hat{b}\bar{X}, \hat{b})$

$= Cov(\bar{Y}, \hat{b}) - Cov(\hat{b}\bar{X}, \hat{b}) = Cov\left(\frac{1}{n}\sum Y_i, \sum C_i Y_i\right) - \bar{X}\, Var(\hat{b})$

$= \frac{1}{n}\sum C_i \sigma^2 - \bar{X}\, Var(\hat{b}) = \frac{\sigma^2}{n}\sum C_i - \bar{X}\, Var(\hat{b}) = -\bar{X}\, Var(\hat{b})$

$= -\frac{\bar{X}\sigma^2}{\sum (X_i - \bar{X})^2}$

② $\hat{a} = \bar{Y} - \hat{b}_1 \bar{X}_1 - \hat{b}_2 \bar{X}_2$ $\quad \hat{a} = \sum \left(\frac{1}{n} - \bar{X}_1 C_{1i} - \bar{X}_2 C_{2i}\right) Y_i$

③ $\hat{a} = \sum C_{1i} Y_i$. $\quad C: (X^T X)^{-1} X^T$ $\quad Var(\hat{a}) = \left(\frac{1}{n} + \frac{\bar{X}_1^2 S_{22} + \bar{X}_2^2 S_{11} - 2\bar{X}_1\bar{X}_2 S_{12}}{S_{11}S_{22} - S_{12}^2}\right)\sigma^2$

$\frac{\hat{a} - a}{S_{Y|X_1 X_2}\sqrt{\frac{1}{n} + \frac{\bar{X}_1^2 S_{22} + \bar{X}_2^2 S_{11} - 2\bar{X}_1\bar{X}_2 S_{12}}{S_{11}S_{22} - S_{12}^2}}} \sim t_{n-3}$

$Var(\hat{a}) = \sigma^2 \sum C_{1i}^2$

$\frac{\hat{a} - a}{S_{Y|X_1 \dots X_k}\sqrt{A_{11}}} \sim t_{n-k-1}$. $\quad A = (X^T X)^{-1}$

$Y^* - \hat{Y}^*$

① $\hat{a} = \sum \left(\frac{1}{n} - C_i\bar{X}\right) Y_i$ $\quad \hat{b} = \sum C_i Y_i$

$\hat{Y}^* = \hat{a} + \hat{b}X^* = \sum \left(\frac{1}{n} + \sum C_i\bar{X} + \sum C_i X^*\right) Y_i = \sum \left(\frac{1}{n} + (X^* - \bar{X})C_i\right) Y_i$

$Y^* \sim N(a + bX^*, \sigma^2)$ $\quad var(\hat{Y}^*) = Var(\hat{a} + \hat{b}X^*) = Var(\hat{a}) + X^{*2} Var(\hat{b}) - 2X^* Cov(\hat{a}, \hat{b})$

$E(Y^* - \hat{Y}^*) = 0$.

$Var(Y^* - \hat{Y}^*) = Var(Y^*) + Var(\hat{Y}^*) = \sigma^2 + \left(\frac{1}{n} + (X^* - \bar{X})C_i\right)^2 \sigma^2 = \sigma^2\left(1 + \frac{1}{n} + \frac{(X^* - \bar{X})^2}{\sum (X_i - \bar{X})^2}\right)$.

$\frac{Y^* - \hat{Y}^*}{S_{Y|X}\sqrt{1 + \frac{1}{n} + \frac{(X^* - \bar{X})^2}{\sum (X_i - \bar{X})^2}}} \sim t_{n-2}$ $\quad \hat{Y}^* \mp t_{n-2, \frac{\alpha}{2}} \cdot S_{Y|X}\sqrt{1 + \frac{1}{n} + \frac{(X^* - \bar{X})^2}{\sum (X_i - \bar{X})^2}}$

② $Var(Y^* - \hat{Y}^*) = \sigma^2 + Var(\hat{Y}^*)$

$Var(\hat{Y}^*) = Var(\hat{a} + \hat{b}_1 X_1^* + \hat{b}_2 X_2^*) =$

$= Var(\bar{Y} + \hat{b}_1(X_1^* - \bar{X}_1) + \hat{b}_2(X_2^* - \bar{X}_2)) = Var\left(\frac{1}{n}\sum Y_i + \dots\right)$.

$= Var\left(\sum \left[\frac{1}{n} - C_{1i}\bar{X}_1 - C_{2i}\bar{X}_2\right] + C_{1i}X_1^* + C_{2i}X_2^*\right) Y_i$

$= \frac{\sigma^2}{n} + (X_1^* - \bar{X}_1)^2 Var(\hat{b}_1) + (X_2^* - \bar{X}_2)^2 Var(\hat{b}_2) + 2(X_1^* - \bar{X}_1)(X_2^* - \bar{X}_2) Cov(\hat{b}_1, \hat{b}_2)$

$\hat{Y}^* \mp t_{\alpha/2, n-3}\sqrt{1 + \frac{1}{n} + V} \cdot S_{Y|X_1 X_2}$

$V = \frac{(X_1^* - \bar{X}_1)^2 S_{22} + (X_2^* - \bar{X}_2)^2 S_{11} - 2(X_1^* - \bar{X}_1)(X_2^* - \bar{X}_2) S_{12}}{S_{11}S_{22} - S_{12}^2}$

③ $Y^* - \hat{Y}^* = [1\ X^*](\beta - \hat{\beta}) + \epsilon^* \sim N(0, \sigma^2(1 + [1\ X^*](X^T X)^{-1}[1\ X^*]^T))$

$\hat{Y}^* \mp t_{\alpha/2, n-k-1}\sqrt{V} \cdot S_{Y|X_1 \dots X_k}$

$V = 1 + [1\ X^*](X^T X)^{-1}\begin{bmatrix}1 \\ X^*\end{bmatrix}$

Y^* and $\hat{Y}^*$ independent since ϵ^*, ϵ_i independent.

H_0

① $H_0: b = 0$.

$F = \frac{SS_{reg}/1}{SS_{res}/(n-2)} \sim F_{1, n-2}$

② $H_0: b_1 = b_2 = 0$

$F = \frac{SS_{reg}/2}{SS_{res}/(n-3)} \sim F_{2, n-3}$

③ $H_0: b_1 = \dots = b_k = 0$

$F = \frac{SS_{reg}/k}{SS_{res}/(n-k-1)} \sim F_{k, n-k-1}$

$R^2 = \frac{SS_{reg}}{SS_{tot}}$ ~~[illegible]~~ coef of determination

① $R_{adj} = 1 - \frac{SS_{res}/(n-2)}{SS_{tot}/(n-1)}$

③ $R^2_{adj} = 1 - \frac{SS_{res}/(n-k-1)}{SS_{tot}/(n-1)}$

$y = a + \epsilon$, $\Rightarrow \hat{a} = \bar{y}$, $SS_{reg}(1) = \sum (\hat{a} - \bar{y})^2 = 0$

add X_k. $SS_{reg}(X_k | X_1 \dots X_{k-1})$ big. X_k important

$SS_{reg}(X_2 | X_1, 1) = SS_{reg}(1, X_1, X_2) - SS_{reg}(1, X_1)$.

$SS_{reg}(X_3, X_4 | 1, X_1, X_2) = SS_{reg}(1, X_1, X_2, X_3, X_4) - SS_{reg}(1, X_1, X_2)$

$= SS_{reg}(1, X_1, X_2, X_3, X_4) - SS_{reg}(1, X_1, X_2, X_3) + SS_{reg}(1, X_1, X_2, X_3) - SS_{reg}(1, X_1, X_2)$

$= SS_{reg}(X_4 | 1, X_1, X_2, X_3) - SS_{reg}(X_3 | 1, X_1, X_2)$

Sum of Sq / Seq SS

Seq. Sum of squares

$H_0: b_{m+1} = b_{m+2} = \dots = b_k = 0$

$F = \frac{(SS_{reg}(FM) - SS_{reg}(RM))/(k-m)}{SS_{res}(FM)/(n-k-1)} = \frac{(R^2(FM) - R^2(RM))/(k-m)}{(1 - R^2(FM))/(n-k-1)} \sim F_{k-m, n-k-1}$

$\hat{\sigma}$

① $E(S^2_{Y|X}) = \sigma^2$

$Var(S^2_{Y|X}) = \frac{2\sigma^4}{n-2}$

$S^2_{Y|X} = \frac{SS_{res}}{n-2}$, $\frac{SS_{res}}{\sigma^2} = \frac{(n-2)S^2_{Y|X}}{\sigma^2} \sim \chi^2_{n-2}$ $\quad W \sim \chi^2_k$, $E(W) = k$, $Var(W) = 2k$

$H_0: \dots, H_1 \dots \frac{(n-2)S^2_{Y|X}}{\sigma^2} < \chi^2(\frac{\alpha}{2})$ or $> \chi^2(1-\frac{\alpha}{2})$

$H_0: \dots, H_1 > \dots > \chi^2(\alpha)$

CI of σ^2: $\left[\frac{(n-2)S^2_{Y|X}}{\chi^2_{n-2}(\frac{\alpha}{2})}, \frac{(n-2)S^2_{Y|X}}{\chi^2_{n-2}(1 - \frac{\alpha}{2})}\right]$

$\hat{e}_i$

$E(\hat{e}_i) = 0$

① $Cov(\hat{e}_i, \hat{e}_j) = \left(-\frac{1}{n} - \frac{(X_i - \bar{X})(X_j - \bar{X})}{\sum (X_k - \bar{X})^2}\right)\sigma^2$

$= Cov(Y_i - \hat{Y}_i, Y_j - \hat{Y}_j)$

$= Cov(Y_i, Y_j) - Cov(Y_i, \hat{Y}_j) - Cov(\hat{Y}_i, Y_j) + Cov(\hat{Y}_i, \hat{Y}_j)$

$Cov(Y_i, \hat{a} + \hat{b}X_j)$

$= Cov(Y_i, \hat{a}) + X_j Cov(Y_i, \hat{b})$

$= Cov(Y_i, \frac{1}{n}\sum Y_k - \bar{X}\sum C_k Y_k) + X_j Cov(Y_i, \sum C_k Y_k)$

$= \frac{1}{n}\sigma^2 - \bar{X}C_i\sigma^2 + X_j C_i \sigma^2$

$Cov(\hat{a} + \hat{b}X_i, \hat{a} + \hat{b}X_j) = Var(\hat{a}) + (X_i + X_j)Cov(\hat{a}, \hat{b}) + X_i X_j Var(\hat{b})$

$\sum C_k Cov(Y_i, Y_k) = C_i \sigma^2$

$Var(\hat{e}_i) = \left(1 - \frac{1}{n} - \frac{(X_i - \bar{X})^2}{\sum (X_i - \bar{X})^2}\right)\sigma^2$

$\hat{b} = \sum \frac{X_i - \bar{X}}{\sum (X_i - \bar{X})^2} Y_i$ $\quad X_i \updownarrow$

$Var(\hat{e}_i) = \left[1 - \frac{1}{n} - \frac{(X_i - \bar{X})^2}{\sum (X_i - \bar{X})^2}\right]\sigma^2$ $\quad X_i \updownarrow$

$= var(Y_i - \hat{Y}_i)$

$= var(Y_i - \bar{Y} - \hat{b}(X_i - \bar{X}))$

$= var(Y_i) + var(\bar{Y}) + (X_i - \bar{X})^2 Var(\hat{b})$

$- 2Cov(Y_i, \bar{Y}) - 2(X_i - \bar{X})Cov(Y_i, \hat{b}) + 2(X_i - \bar{X})Cov(\bar{Y}, \hat{b})$

$\tilde{e}_i = \frac{\hat{e}_i}{\sqrt{1 - \frac{1}{n} - \frac{(X_i - \bar{X})^2}{\sum (X_j - \bar{X})^2}}\cdot S_{Y|X}}$ $\quad \frac{\hat{e}_i - 0}{\sqrt{1 - \frac{1}{n} - \frac{(X_i - \bar{X})^2}{\sum (X_j - \bar{X})^2}}\cdot S_{Y|X}} \sim t_{n-2}$

$Cov(Y_i, \frac{1}{n}\sum Y_j) = \frac{1}{n}\sum_j Cov(Y_i, Y_j) = \frac{1}{n}\sigma^2$

$Cov(Y_i, \sum C_j Y_j) = \sum C_j Cov(Y_i, Y_j) = C_i \sigma^2$

$Var(\frac{1}{n}\sum Y_i) = \frac{1}{n^2}\sum Var(Y_i) = \frac{\sigma^2}{n}$.

$Cov(\frac{1}{n}\sum Y_i, \sum C_j Y_j) = \frac{1}{n}\sum\sum C_j Cov(Y_i, Y_j) = \frac{1}{n}\sum C_i \sigma^2 = 0$.

$n_1 = n_2$

$H_0: \sigma_1^2 = \sigma_2^2$ $\quad \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} = \frac{S_1^2}{S_2^2} \sim F$

$\frac{S_1^2}{S_2^2} \sim F_{n_1-2, n_2-2}$

$\rho_{ij} = \frac{Cov(\hat{e}_i, \hat{e}_j)}{\sqrt{Var(\hat{e}_i)Var(\hat{e}_j)}}$

$Cov(\hat{e}_i, \hat{Y}_j) = Cov(Y_i, \hat{Y}_j) - Cov(\hat{Y}_i, \hat{Y}_j) = 0$.

① $SS_{res} = (n-2) S^2_{Y|X}$
$SS_{tot} = (n-1) S^2_Y$
$SS_{reg} = SS_{tot} - SS_{res}$.

	SS	df	MS	①
Regression	$\Sigma(\hat{Y}_i-\bar{Y})^2$	k	SS_{reg}/K	1
Residual	$\Sigma(Y_i-\hat{Y})^2$	$n-k-1$	$SS_{res}/(n-k-1)$	$n-2$
Total	$\Sigma(Y_i-\bar{Y})^2$	$n-1$		$n-1$

seek a measure . not unit dependent . .

$W_i = \frac{Y_i-\bar{Y}}{S_Y}$ $\quad Z_i = \frac{X_i-\bar{X}}{S_X}$

sample correlation coef $r_{X,Y} = \hat{b}_{W,Z} = \frac{\Sigma(Z_i-\bar{Z})(W_i-\bar{W})}{\Sigma(Z_i-\bar{Z})^2} = \frac{\Sigma Z_i W_i}{(n-1)S_Z^2} = \frac{\Sigma Z_i W_i}{n-1} = \frac{\Sigma\left(\frac{X_i-\bar{X}}{S_X}\right)\left(\frac{Y_i-\bar{Y}}{S_Y}\right)}{n-1}$.

$= \frac{\Sigma(X_i-\bar{X})(Y_i-\bar{Y})/(n-1)}{S_X \cdot S_Y} = \frac{C_{XY}}{S_X S_Y}$.

r_{YX} $r_{X,W}$.

$\hat{b}_{X,Y} = \frac{C_{X,Y}}{S_X^2} = r_{X,Y} \cdot \frac{S_Y}{S_X}$

$\hat{b}_{Y,X} = \frac{C_{X,Y}}{S_Y^2} = r_{X,Y} \cdot \frac{S_X}{S_Y}$

$R^2 = r^2_{X,Y}$ $\quad R^2_{Y|X}$

correlation coefficient.
$\rho(X,Y) = \frac{Cov(X,Y)}{SD(X)\cdot SD(Y)}$.

$\frac{SS_{reg}}{SS_{tot}} = \frac{\Sigma(\hat{Y}_i-\bar{Y})^2}{\Sigma(Y_i-\bar{Y})^2} = \frac{\Sigma[\hat{b}(X_i-\bar{X})]^2}{\Sigma(Y_i-\bar{Y})^2} = \hat{b}^2 \frac{\Sigma(X_i-\bar{X})^2}{\Sigma(Y_i-\bar{Y})^2}$

$= \left[\frac{\Sigma(X_i-\bar{X})(Y_i-\bar{Y})}{\Sigma(X_i-\bar{X})^2}\right]^2 \frac{\Sigma(X_i-\bar{X})^2}{\Sigma(Y_i-\bar{Y})^2}$

$= \frac{(\Sigma(X_i-\bar{X})(Y_i-\bar{Y}))^2}{\Sigma(X_i-\bar{X})^2\,\Sigma(Y_i-\bar{Y})^2} = \frac{(n-1)^2 C^2_{X,Y}}{(n-1)^2 S_X^2 \cdot S_Y^2} = r^2_{X,Y}$.

Added variable

① $Y \sim X_1 \Rightarrow \hat{e}(Y,X_1)$ — $\hat{e}_i(Y,X_1) = y_i - \hat{y}_i$, $[\hat{y}_i = \hat{u}_1 + \hat{v}_1 X_{1i}][\hat{v}_1 = \frac{S_{1Y}}{S_{11}} \quad \hat{u}_1 = \bar{y} - \hat{v}_1\bar{X}_1]$

$X_2 \sim X_1 \Rightarrow \hat{e}(X_2,X)$ — $\hat{e}_i(X_2,X_1) = X_{2i} - \hat{X}_{2i}$, $[\hat{X}_{2i} = \hat{u}_2 + \hat{v}_2 X_{1i}][\hat{v}_2 = \frac{S_{12}}{S_{11}} \quad \hat{u}_2 = \bar{X}_2 - \hat{v}_2\bar{X}_1]$.

$\hat{e}(Y,X_1) \sim \hat{e}(X_2,X_1)$ — slope $= \frac{\Sigma \hat{e}_i(Y,X_1)\cdot\hat{e}_i(X_2,X_1)}{\Sigma[\hat{e}(X_2,X_1)]^2} = \frac{\Sigma[(y_i-\bar{y})-\hat{v}_1(X_{1i}-\bar{X}_1)][(X_{2i}-\bar{X}_2)-\hat{v}_2(X_{1i}-\bar{X}_1)]}{\Sigma[(X_{2i}-\bar{X}_2)-\hat{v}_2(X_{1i}-\bar{X}_1)]^2}$

(0,0)

$= \frac{\Sigma(y-)(X_2-) - \hat{v}_2\Sigma(y-)(X_1-) - \hat{v}_1(X_1-)(X_2-) + \hat{v}_1\hat{v}_2(X_1-)^2}{\Sigma(X_2-)^2 - 2\hat{v}_2\Sigma(X_2-)(X_1-) + \hat{v}_2^2(X_1-)^2}$

② $y \sim X_1 \sim X_{k-1} \Rightarrow \hat{e}(Y)$

$X_k \sim X_1 \sim X_{k-1} \Rightarrow \hat{e}(X_k)$. $\quad = \frac{S_{2y} - \frac{S_{12}}{S_{11}}S_{1y} - \frac{S_{1y}}{S_{11}}S_{12} + \frac{S_{1y}}{S_{11}}\frac{S_{12}}{S_{11}}\cdot S_{11}}{S_{22} - 2\frac{S_{12}}{S_{11}}S_{12} + \frac{S_{12}}{S_{11}}\frac{S_{12}}{S_{11}}S_{11}} = \frac{S_{2y}S_{11} - S_{1y}S_{12}}{S_{11}S_{22} - S_{12}^2} = \hat{b}_2$.

$\hat{e}(Y) \sim \hat{e}(X_k)$ close to a single straight line

①② $b_1^{slm} = \Sigma c_i Y_i \quad c_i = \frac{X_{1i}-\bar{X}_1}{S_{11}}$

$E b_1^{slm} = \Sigma c_i E[a + b_1 X_{1i} + b_2 X_{2i} + \epsilon_i] = a\underbrace{\Sigma c_i}_{0} + b_1^{biv}\underbrace{\Sigma c_i X_{1i}}_{\frac{\Sigma(X_{1i}-\bar{X}_1)}{S_{11}}(X_{1i}-\bar{X}_1)} + b_2^{biv}\Sigma c_i X_{2i}$ $\quad \frac{\Sigma(X_{1i}-\bar{X}_1)}{S_{11}}(X_{2i}-\bar{X}_2)$

$= b_1^{biv} + b_2^{biv}\frac{S_{12}}{S_{11}}$ (bias)

$var(\hat{b}_1^{biv}) = \frac{S_{22}}{S_{11}S_{22}-S_{12}^2}\sigma^2 \geqslant var(\hat{b}_1^{slm}) = \frac{1}{S_{11}}\sigma^2$

Collinearity.

large correlation of X_1, X_2. $\Rightarrow S_{11}S_{22} - S_{12}^2$ small

$\Rightarrow \frac{1}{S_{11}S_{22}-S_{12}^2}$ more sensitive to change than S_{22} and S_{2y}

$\hat{b}_1 = \frac{S_{1y}S_{22} - S_{2y}S_{12}}{S_{11}S_{22} - S_{12}^2}$ sensitive to change .

$var(\hat{b}_1) = \frac{1}{S_{11}S_{22}-S_{12}^2}\sigma^2 \nearrow$ $\quad SD = \sqrt{\frac{1}{S_{11}S_{22}-S_{12}^2}\cdot\sigma^2} \nearrow$. $\quad \hat{b}_2 \sim$

$\begin{cases} var(Y^* - \hat{Y}^*) \\ var(Y - \hat{Y}) \end{cases}$.

$Var(y^* - \hat{y}^*) = Var(y^*) + var(\hat{y}^*)$

$Var(y - \hat{y}) = var(y) + var(\hat{y}) - 2cov(y, \hat{y})$.

$Cov(y_i, y_j) = 0 \quad Var(y_i) = \sigma^2$
$cov(y_i, \hat{y}_j) \neq 0$
$Cov(\hat{y}_i, \hat{y}_j) \neq 0$. $Var(\hat{y}_i) \neq 0$
$var(y^*) = \sigma^2$
$Cov(y^*, \hat{y}^*) = 0$

① $\hat{y}^* = \hat{a} + \hat{b}x^* = \Sigma(\frac{1}{n} + (x^*-\bar{x})c_i)Y_i$

$var(Y^* - \hat{Y}^*) = Var(Y^*) + Var(\hat{Y}^*) = \sigma^2 + \Sigma(\frac{1}{n} + (x^*-\bar{x})c_i)^2\sigma^2$

$= \sigma^2\left(1 + \frac{1}{n} + \frac{(x^*-\bar{x})^2}{\Sigma(x_i-\bar{x})^2}\right)$ $\boxed{cov(y^*, \hat{y}^*) = 0}$

$Var(Y - \hat{Y}) = Var(Y_i - \bar{Y} - \hat{b}(x_i - \bar{x})) = var(y_i) + var(\bar{y}) + (x_i - \bar{x})^2 Var(\hat{b})$

$-2cov(y_i, \bar{y}) - 2(x_i - \bar{x})cov(y_i, \hat{b}) + 2(x_i - \bar{x})cov(\bar{y}, \hat{b})$

$\| =$

$\left(1 - \frac{1}{n} - \frac{(x_i-\bar{x})^2}{\Sigma(x_k-\bar{x})^2}\right)\sigma^2$

$cov(y_i, \frac{1}{n}\Sigma_j y_j) = \frac{1}{n}\Sigma cov(y_i, y_j) = \frac{\sigma^2}{n}$ $\quad cov(y_i, \Sigma_j c_j y_j) = \Sigma_j c_j cov(y_i, y_j) = c_i\sigma^2$ $\quad cov(\frac{1}{n}\Sigma_i y_i, \Sigma_j c_j) = \frac{1}{n}\Sigma_i\Sigma_j cov(y_i, y_j) = \frac{1}{n}\Sigma_i c_i\sigma^2 = 0$

③ $Var(Y^* - \hat{Y}^*) = Var(Y^* - [1\ x^*]\hat{\beta})$

$= Var(Y^*) + [1\ x^*]Var(\hat{\beta})\begin{bmatrix}1\\x^*\end{bmatrix}\sigma^2 = \sigma^2(1 + [1\ x^*](X^TX)^{-1}\begin{bmatrix}1\\x^*\end{bmatrix})$.

$Var(Y - \hat{Y}) = Var(Y - X\hat{\beta}) = Var[(I - X(X^TX)^{-1}X^T)Y]$

$= (I - X(X^TX)^{-1}X^T)\sigma^2$ $\quad$ since $I - X(X^TX)^{-1}X^T$ ...

$Var(Y - X\hat{\beta}) = Var(Y) + Var(X\hat{\beta}) - 2Cov(Y, X\hat{\beta})$

$= Var(Y) + X Var(\hat{\beta})X^T - 2X(X^TX)^{-1}X^T Var(Y)$

$= (I - X(X^TX)^{-1}X^T)\sigma^2$

$\begin{cases} \hat{e}_i = \frac{\hat{e}_i}{\hat{\sigma}(1-h_{ii})} \\ t_i = \frac{d_i}{SD(d_i)} = \hat{e}_i\left(\frac{n-k-1}{n-k-\hat{e}_i^2}\right)^{\frac{1}{2}} ; \quad d_i = Y_i - \hat{Y}_{(i)} \end{cases}$

$logit(p) = log(\frac{p}{1-p})$

$g^{-1}(x) = \frac{e^x}{1+e^x} = \frac{1}{1+e^{-x}}$

$\begin{pmatrix} x = log(\frac{p}{1-p}) & \\ e^x = \frac{p}{1-p} & e^x = (e^x+1)p \\ e^x - e^x p = p & p = \frac{e^x}{e^x+1} \end{pmatrix}$

$L(a, b_1 \ldots b_k) = \prod_{i=1}^n p_i^{Y_i}(1-p_i)^{1-Y_i} = \prod_{i=1}^n (1-p_i)\left(\frac{p_i}{1-p_i}\right)^{Y_i}$

$l(\beta_0 \ldots \beta_k) = \sum_{i=1}^n Y_i \log\frac{p_i}{1-p_i} + \log(1-p_i)$

$= \sum_{i=1}^n (Y_i \sum_{j=0}^k \beta_j x_{ji} - \log(1 + e^{\sum_{j=0}^k \beta_j x_{ji}}))$

$y_i \sim Bern(p_i)$	$y_i \sim Bern(p_i)$
$\eta_i = a + bX_i$	$\eta_i = a + b_1 X_{1i} + \ldots + b_k X_{ki}$
$logit(p_i) = \eta_i$	$logit(p_i) = \eta_i$
$p_i = \frac{e^{a+bX_i}}{1+e^{a+bX_i}}$	$p_i = \frac{e^{a+b_1X_{1i}+\ldots+b_kX_{ki}}}{1+e^{a+b_1X_{1i}+\ldots+b_kX_{ki}}}$

$odds(E) = \frac{P(E)}{1-P(E)} = e^{a+b_1x_1+\ldots+b_kx_k}$ $\quad$ | odds ratio : $e^{b_j\cdot\Delta}$.

$g_\lambda(Y) = \begin{cases} \frac{Y^\lambda - 1}{\lambda} & \lambda \neq 0 \\ \ln Y. & \lambda = 0 \end{cases}$ . . $\quad$ ($\hat{\lambda} = 0.46 : \sqrt{Y}$; 95% includes 1 : Y ; 95% includes 0 : $\log Y$)

$L(\lambda) = -\frac{n}{2}\ln(SS_{res}(\lambda)/n) + (\lambda-1)\sum_{i=1}^n \ln Y_i$

$AIC_p = n\ln(SS_{res}/n) + 2p$

Best subset selection : AIC . BIC ..

Backward : FM . remove 1/step . (p value)

Forward : (p value)

forward. $+X_2$ $+X_4$ $+ \ldots$ (not drop that already entered,

AIC : ① X_2

② not necessarily include X_2 .

①③ $X = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}$ $\quad (X^TX)^{-1} = \begin{pmatrix} n & \Sigma x_i \\ \Sigma x_i & \Sigma x_i^2 \end{pmatrix}^{-1} = \frac{1}{n\Sigma x_i^2 - (\Sigma x_i)^2}\begin{pmatrix} \Sigma x_i^2 & -\Sigma x_i \\ -\Sigma x_i & n \end{pmatrix}$.

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

$\hat{a}$ $\quad Var(\hat{a}) = V_{11} = \frac{\Sigma x_i^2}{n\Sigma x_i^2 - (\Sigma x_i)^2} = \frac{\Sigma x_i^2 - \frac{1}{n}(\Sigma x_i)^2 + \frac{1}{n}(\Sigma x_i)^2}{n\Sigma x_i^2 - (\Sigma x_i)^2} = \frac{1}{n} + \frac{\bar{x}^2}{\Sigma(x_i-\bar{x})^2}$

$\hat{b}$ $\quad Var(\hat{b}) = V_{22} = \frac{n}{n\Sigma x_i^2 - (\Sigma x_i)^2} = \frac{1}{\Sigma(x_i-\bar{x})^2}$

$\hat{a}, \hat{b}$ $\quad cov(\hat{a}, \hat{b}) = V_{12} = \frac{-\Sigma x_i}{n\Sigma x_i^2 - (\Sigma x_i)^2} = \frac{-\Sigma x_i/n}{\Sigma x_i^2 - \frac{1}{n}(\Sigma x_i)^2} = \frac{-\bar{x}}{\Sigma(x_i-\bar{x})^2}$.

$\begin{cases} x_i \text{ is known. constant} \\ E(\epsilon_i) = 0 \quad Var(\epsilon_i) = \sigma^2 \\ \epsilon_i \text{ mutually ind} \\ \epsilon_i \text{ from normal} \end{cases}$

① $Var(y_i) = \sigma^2$. $Cov(y_i, y_j) = 0$
$cov(y_i, \hat{y}_j) \neq 0$.
$var(\hat{y}_i) \neq 0 \quad cov(\hat{y}_i, \hat{y}_j) \neq 0$
$var(y^*) = \sigma^2$
$cov(y^*, \hat{y}^*) = 0$

② $Var(\epsilon_i) = \sigma^2 \quad Cov(\epsilon_i, \epsilon_j) = 0$
$Var(\hat{e}_i) = (\ldots)\sigma^2 \quad cov(\hat{e}_i, \hat{e}_j) = (\ldots)\sigma^2$
$Cov(\hat{e}_i, \hat{Y}_j) = 0$.

$cov(y_i, \bar{Y}) = \frac{1}{n}\sigma^2$
$Cov(Y_i, \hat{b}) = \ldots$
$Cov(\bar{Y}, \hat{b}) = 0$

$Cov(\hat{e}_i, \hat{Y}_j) = cov(Y_i, \hat{Y}_j) - cov(\hat{Y}_i, \hat{Y}_j)$

$cov(y_i, \frac{1}{n}\Sigma y_k + (x_j - \bar{x})\Sigma c_k y_k)$

$= \frac{1}{n}cov(y_i, y_i) + c_i(x_j - \bar{x})cov(y_i, y_i)$

$= \frac{1}{n}\sigma^2 + c_i(x_i - \bar{x})\sigma^2$

$= Var(\hat{Y}_i) = var(\hat{a} + \hat{b}x_i)$

$= var(\hat{a}) + x_i^2 Var(\hat{b}) + 2x_i Cov(\hat{a}, \hat{b})$

$= \sigma^2\left[\frac{1}{n} + \frac{\bar{x}^2}{\Sigma(x_i-\bar{x})^2} + \frac{x_i^2}{\Sigma(x_i-\bar{x})^2} - \frac{2x_i\bar{x}}{\Sigma(x_i-\bar{x})^2}\right] = \sigma^2\left[\frac{1}{n} + \frac{(x_i-\bar{x})^2}{\Sigma(x_i-\bar{x})^2}\right]$

$Var[\bar{Y} + \hat{b}(x_i - \bar{x})]$

$= var(\bar{Y}) + (x_i - \bar{x})^2 var(\hat{b}) + 2(x_i - \bar{x})cov(\bar{Y}, \hat{b})$

residual plots : non-linearity. non-constant variance. serial correlation outliers
(residual vs. fitted)

Normal QQ plot : $P(X_1 \leq x_1) = P(X_2 \leq x_2) \Rightarrow x_2 = \frac{\sigma_2}{\sigma_1}x_1 - \frac{\sigma_2}{\sigma_1}\mu_1 + \mu_2$.

lower quartile. Q_1 $\begin{cases} \text{if } k = 0.25 \times n \text{ is int} : Q_1 = X_{(k)} + 0.25 \times (X_{(k+1)} - X_{(k)}) \\ \text{if } k = 0.25 \cdot n \text{ not int} : Q_1 = X_{(\lceil k \rceil)} \end{cases}$

sample median. $\tilde{x}$ $\begin{cases} \text{if } n = 2k+1 . & \tilde{x} = X_{(k+1)} \\ \text{if } n = 2k . & \tilde{x} = (X_{(k)} + X_{(k+1)})/2 . \end{cases}$

upper quartile Q_3 $\begin{cases} \text{if } k = 0.75 \times n \text{ is int} : Q_3 = X_{(k)} + 0.75 \times (X_{(k+1)} - X_{(k)}) \\ \text{if } k = 0.75 \times n \text{ not int} \cdot Q_3 = X_{(\lceil k \rceil)} \end{cases}$

interquartile range : $IQR = Q_3 - Q_1$. $\quad$ ($\pm S$, $\pm\frac{IQR}{2}$)

$SS(a,b) = \Sigma(Y_i - (a + bx_i)^2)$

$LAD(a,b) = \Sigma|Y_i - (a + bx_i)|$ $\longrightarrow \begin{cases} \tilde{b} = \text{median}\left\{\frac{Y_i - \tilde{Y}}{x_i - \tilde{x}} : 1 \leq i \leq n ; x_i \neq \tilde{x}\right\} . \\ \tilde{a} = \tilde{Y} - \tilde{b}\tilde{x} \end{cases}$

$na + b\Sigma x_i = \Sigma Y_i$

$a\Sigma x_i + b\Sigma x_i^2 = \Sigma x_i Y_i$ $\leftarrow$ $\begin{pmatrix} \hat{b}(\Sigma x_i^2 - n\bar{x}^2) = \Sigma x_i Y_i - n\bar{x}\bar{Y} \\ \hat{b}\bar{x} = \bar{Y} - \hat{a} \\ \hat{b}(\Sigma x_i^2 - n\bar{x}^2) = \hat{b}\Sigma x_i^2 - n\bar{x}(\bar{Y} - \hat{a}) = \hat{b}\Sigma x_i^2 - n\bar{x}\bar{Y} + n\bar{x}\hat{a} \\ \hat{b}\Sigma x_i^2 + \hat{a}\Sigma x_i = \Sigma x_i Y_i \end{pmatrix}$

$\begin{pmatrix} n & \Sigma x_i \\ \Sigma x_i & \Sigma x_i^2 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \Sigma Y_i \\ \Sigma x_i Y_i \end{pmatrix}$.

$y_i = \beta_1 x_i + \beta_2 i + \epsilon_i$

$e_i = y_i - \beta_1 x_i - \beta_2 i$. $SS(\beta_1, \beta_2) = \Sigma e_i^2 = \Sigma(y_i - \beta_1 x_i - \beta_2 i)^2$

$\begin{cases} \frac{\partial S(\beta_1,\beta_2)}{\partial\beta_1} = -2\Sigma(y_i - \beta_1 x_i - \beta_2 i)x_i = 0 \\ \frac{\partial S(\ )}{\partial\beta_2} = -2\Sigma(y_i - \beta_1 x_i - \beta_2 i)\,i = 0 \end{cases}$

$\Rightarrow \Sigma x_i y_i - \hat{\beta}_1\Sigma x^2 - \hat{\beta}_2\Sigma x_i i = 0 \Rightarrow \hat{\beta}_1 = \frac{\Sigma x_i y_i - \hat{\beta}_2\Sigma x_i i}{\Sigma x_i^2}$

$\Rightarrow \Sigma i y_i - \hat{\beta}_1\Sigma i x_i - \hat{\beta}_2\Sigma i^2 = 0 . \Rightarrow \hat{\beta}_2 = \frac{\Sigma i y_i - \hat{\beta}_1\Sigma x_i i}{\Sigma i^2}$

$\beta_1 = \frac{\Sigma i^2\Sigma x_i y_i - (\Sigma i x_i)(\Sigma i y_i)}{\Sigma i^2\Sigma x_i^2 - (\Sigma i x_i)^2}$.

$\beta_2 = \frac{\Sigma x_i^2\Sigma i y_i - (\Sigma x_i y_i)(\Sigma i x_i)}{\Sigma i^2\Sigma x_i^2 - (\Sigma i x_i)^2}$ $\quad$ not well def . $\Sigma i^2\Sigma x_i^2 = (\Sigma i x_i)^2$

$N(\mu, \sigma^2) \xrightarrow{20} N(20\mu, 20\sigma^2)$.

$W_i = c + dY_i$

$\bar{W}_n = c + d\bar{Y}$

$SD(W)^2 = \frac{1}{n-1}\Sigma(W_i - \bar{W}_n)^2 = \frac{1}{n-1}\Sigma(c + dY_i - c - d\bar{Y})^2 = d^2 S_Y^2$.

$S_Y = \sqrt{\Sigma(Y_i - \bar{Y})^2/(n-1)}$. $\quad SD(W) = 1$. $\bar{W} = 0$. $\quad d = \frac{1}{S_Y}$. $c = -\frac{\bar{Y}}{S_Y}$