14.8: Residual Analysis: validating Model Assumption.

Def: Residual for observation i:

$y_i - \hat{y}_i$ where y_i: observed value of Y at x_i

$\hat{y}_i$: estimated value of Y at x_i.

→ Model Assumptions:

$Y = \beta_0 + \beta_1 X$

(A_1) $E(\varepsilon) = 0$

(A_2) $var(\varepsilon) = \sigma^2$ constant.

(A_3) ε independent.

(A_4) ε Normal.

→ Residual Analysis:

1. plot of Residual against X.
2. Plot of Residual against $\hat{y}$.
3. Plot of standardized Residual.
4. Normal probality Plot.

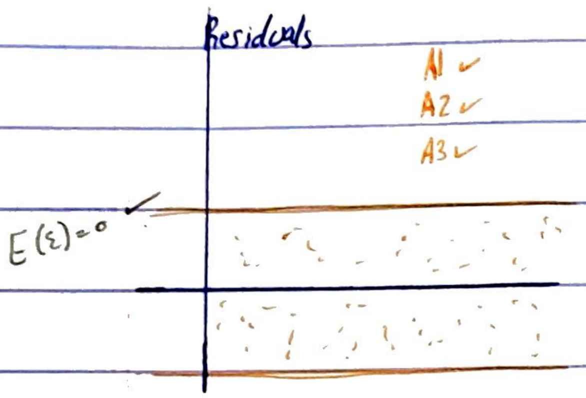
Def: standardized Residual for observation i:

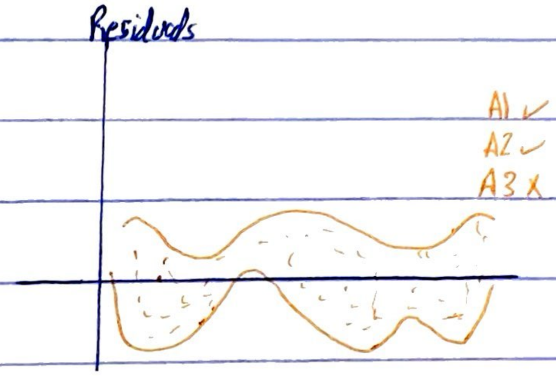
$\dfrac{y_i - \hat{y}_i}{S_{y_i - \hat{y}_i}}$ where $S_{y_i - \hat{y}_i}$ standard deviation for residual i.

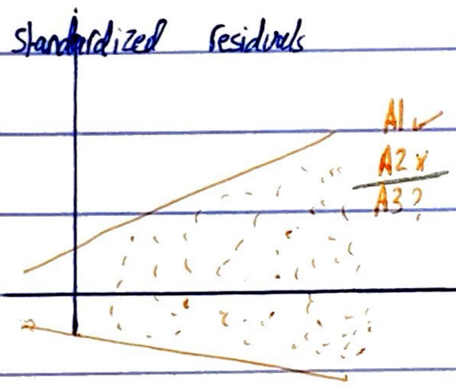
→ $S_{y_i - \hat{y}_i} = S\sqrt{1 - h_i}$

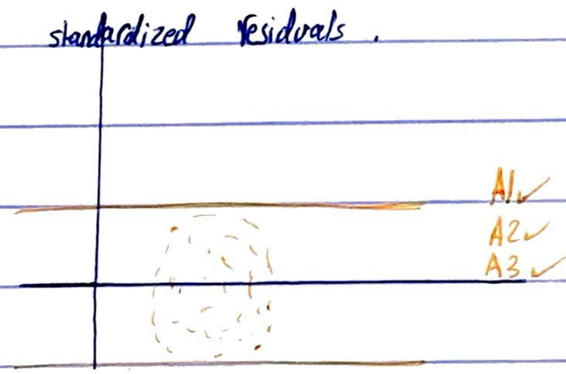
→ $h_i = \dfrac{1}{n} + \dfrac{(x_i - \bar{x})^2}{\sum (x_j - \bar{x})^2}$.

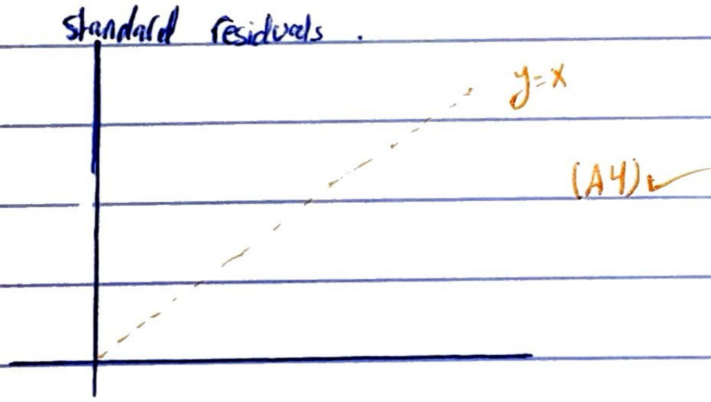
Notes:











Normal Prob. plot.

→ on Exel:

Residual

Residual

A1 : لازم Residual تكون مبعثرة حوالين 0 (موزعين بشكل شبه متساوي حوالين الصفر)

A2 : Variance = σ انو يكونو حواليهم بشريط ثابت.

A3 : ولازم مايكون في نمط ثابت إلهم (عشان independent).

A4 : انو اقدر اعمل خط مستقيم بشكل تقريبي.

from excel sheet.

Example: (A1) ✓

(A2) ✓ $\hat{y} = 0.2 + 2.6\,x$. ANOVA / t-test

(A3) ✓ $r^2 = 0.845$. Model X

(A4) ✓ p-value = 0.03.

Model Adjustmints: (A1) ✓ (A2) ✓ (A3) ✓ (A4) ✓.