~ Tests for Constancy of error Variance ~

Besides graphical methods ....

P. 116

1. Modified Levene test (Brown-Forsythe test)

Applicable to simple linear reg. when $\sigma^2\{\varepsilon_i\}$ changes monotonically w/ predictor x (*)

Need : $n \gg p$ s.t. dependence among e_i's $\to$ ignored

data (X_i, Y_i), $i = 1, \cdots, n$ < one group w/ low x / " " " high x

$\hookrightarrow e_i$ < e_{i1}, $i = 1, \cdots, n_1$ / e_{i2} $i = 1, \cdots, n_2$ $\qquad n_1 + n_2 = n$.

Let $\tilde{e}_j = \text{median}\{e_{ij} : i = 1, \cdots, n_j\}$ $\quad j = 1, 2$

If (*), then $|e_{i1} - \tilde{e}_1|$'s tend to be smaller/larger than $|e_{i2} - \tilde{e}_2|$'s

Let $\quad d_{ij} = |e_{i1} - \tilde{e}_j|$, $i = 1, \cdots n_j$, $j = 1, 2$

$$\bar{d}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} d_{ij}.$$

$$t_L^* = \frac{\bar{d}_1 - \bar{d}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad \text{w/ } s^2 = \frac{\Sigma(d_{i1} - \bar{d}_1)^2 + \Sigma(d_{i2} - \bar{d}_2)^2}{n-2}$$

Result : Under $H_0 : \sigma^2\{\varepsilon_i\} = \sigma^2 \ \forall i$

If n_1, n_2 are not extremely small

then $t_L^* \sim t_{(n-2)}$

Hence, reject H_0 at level α

$$\text{if } |t_L^*| > t(1-\tfrac{\alpha}{2}; n-2)$$

c.f. ex. on p. 117 - 118

2. Breusch - Pagan Test. (p. 118)

Applicable to large sample and $\log \sigma_i^2 = \gamma_0 + \gamma_1 x_i$

i.e. $H_0: \sigma^2\{\varepsilon_i\} = \sigma^2 \ \forall i \Rightarrow H_0: \gamma_1 = 0$

Reg. e_i^2 on x_i $\Rightarrow$ obtain SSR^*

Result:

$$X_{BP}^2 = \frac{\frac{SSR^*}{2}}{\left(\frac{SSE}{n}\right)^2} \xrightarrow{n \to \infty} \chi_1^2$$

w) SSE is from ANOVA by reg. Y on x

under $H_0: \gamma_1 = 0$

Hence, reject $H_0: \sigma^2\{\varepsilon_i\} = \sigma^2 \ \forall i$

if $X_{BP}^2 > \chi^2(1-\alpha; 1)$

such test has level $\approx \alpha$.

c.f. p. 119 for example.

~ Goodness - of - fit ~ (3.7)

— F test for lack of fit —

One predictor : X

S-L-R model : $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$, $i=1 \cdots n$

w/ ε_i iid $N(0, \sigma^2)$ (*)

$\Leftrightarrow$ $E(Y \mid X = X_i) = \beta_0 + \beta_1 X_i$, $i=1 \cdots n$

$Y \mid X = X_i \sim N(\beta_0 + \beta_1 X_i, \sigma^2)$,

$Y \mid X \sim N(\beta_0 + \beta_1 X, \sigma^2)$

Goodness-of-fit test is to test

$H_0 : E(Y \mid X) = \beta_0 + \beta_1 X$ v.s. H_1: Not H_0

assuming $Y_i = Y \mid X = X_i$, $i = 1 \cdots n$, indep. normal w/ common var. σ^2.

I.e. test the appropriateness of a linear reg. fit.

$\Rightarrow$ Assume $\underset{\sim}{Y} \sim N_n(\underset{\sim}{\theta}, \sigma^2 I)$; Ω = span by $\{\underset{\sim}{1}, \underset{\sim}{X}\}$

to test $H_0 : \theta_i = \beta_0 + \beta_1 X_i$, $i=1 \cdots n$ $\Leftrightarrow$ $H_0 : \underset{\sim}{\theta} \in \Omega$

v.s. H_1 : NOT H_0

Note : Let H : hat matrix , projection matrix onto Ω.

$$\hat{\underset{\sim}{Y}} = H \underset{\sim}{Y} , \quad \underset{\sim}{e} = \underset{\sim}{Y} - \hat{\underset{\sim}{Y}}$$

IF $\underset{\sim}{\theta} \in \Omega$,

$$E(\underset{\sim}{e}) = 0 \Leftrightarrow E(e_i) = 0 \quad \forall i = 1 \cdots n$$

O/W. ($\underset{\sim}{\theta} \notin \Omega$)

$$E(\underset{\sim}{e}) = E(\underset{\sim}{Y}) - E(\hat{\underset{\sim}{Y}}) = \underset{\sim}{\theta} - H\underset{\sim}{\theta} \neq \underset{\sim}{0}$$

$H_0 : \underset{\sim}{\theta} \in \Omega \Leftrightarrow H_0 : B_i = 0, \forall i.$ $\Leftarrow$ $E(e_i) = B_i$, $i=1 \cdots n$, not all B_i's are 0 but $\quad \sigma^2\{\underset{\sim}{e}\} = (I-H)\,\sigma^2 \qquad H = (h_{ij})_{n\times n}.$

$$\sigma^2\{e_i\} = (1-h_{ii})\,\sigma^2 \quad , i=1\cdots n$$

$$\Rightarrow \quad E(SSE) = E(\Sigma e_i^2) = \Sigma E(e_i^2)$$
$$= \Sigma\left(\sigma^2\{e_i\} + (E(e_i))^2\right)$$
$$= \sigma^2\Sigma(1-h_{ii}) + \Sigma B_i^2$$
$$= \sigma^2(n-\Sigma h_{ii}) + \Sigma B_i^2$$
$$= (n-2)\,\sigma^2 + \Sigma B_i^2 \qquad \Sigma h_{ii} = 2$$

$$\Rightarrow \quad E(MSE) = E\left(\frac{SSE}{n-2}\right)$$
$$= \sigma^2 + \frac{1}{n-2}\Sigma B_i^2 \quad \cdots\cdots (\Delta 1)$$

note : $\quad I-H$: sym. idempotent w/ rank $= n-2$

$$SSE = \underset{\sim}{e}^t\underset{\sim}{e} = \underset{\sim}{Y}^t(I-H)\underset{\sim}{Y}$$

thm. 1 $\Rightarrow \quad \dfrac{SSE}{\sigma^2} \sim \chi^2_{n-2,\,\delta} \quad$ w/ $\delta = \underset{\sim}{\theta}^t(I-H)\underset{\sim}{\theta}/\sigma^2$ $\quad$ (*1)

$$\Rightarrow \quad E\left(\frac{SSE}{\sigma^2}\right) = n-2+\delta$$

$$\Rightarrow \quad E(MSE) = \sigma^2 + \sigma^2\cdot\delta/(n-2) \quad \cdots (\Delta 2)$$

$$(\Delta 1) = (\Delta 2) \quad \Rightarrow \quad \boxed{\delta = \frac{1}{\sigma^2}\sum_1^n B_i^2}$$

Hence , under H_0 : $\underset{\sim}{\theta} \in \Omega$ i.e. all B_i's $= 0$

$\delta = 0 \quad \Rightarrow \quad \dfrac{SSE}{\sigma^2} \sim \chi^2_{n-2}$

$(\Delta 1) = \sigma^2$

under H_1 : $\underset{\sim}{\theta} \notin \Omega$ i.e. some B_i's $\neq 0$

$\delta \neq 0 \quad \dfrac{SSE}{\sigma^2} \sim \chi^2_{n-2,\,\delta} \quad \cdots\cdots$ (*1) $\quad \delta = \dfrac{1}{\sigma^2}\Sigma B_i^2$

$(\Delta 1) > \sigma^2$

$E\left(\dfrac{SSE}{\sigma^2}\right)$ larger under H_1 than under H_0.

$\Rightarrow$ (I) If σ^2 = known

$\Rightarrow$ rej. $H_0: \underset{\sim}{\theta} \in \Omega$ if $\frac{SSE}{\sigma^2} > \chi^2_{n-2, 1-\alpha}$

is a level α test.

(II) σ^2 = unknown (typical case.)

need an est. for σ^2 ....

($\Delta 1$) : MSE is unbiased for σ^2 <u>under</u> H_0

But H_0 is now under testing....

need an est. for σ^2 when

$\underset{\sim}{Y} \sim N_n(\underset{\sim}{\theta}, \sigma^2 \cdot I)$ ($\underset{\sim}{\theta}$ may not in Ω.)

↳ require replications .. at some values of the predictor.

Let $x_1, \ldots, x_c$ be the different values (levels) of the predictor X in the data.

i.e. $\forall i = 1 \ldots n$. $X_i \in \{x_1, \ldots, x_c\}$

let Y_{ij} = the response of the repeated trials, $i = 1 \cdots n_j$, when $X = x_j$, $j = 1 \cdots c$.

$\sum_{j=1}^{c} n_j = n$.

$E(Y_{ij}) = \underbrace{E(Y \mid X = x_j)}_{= \mu_j}$, $\forall i = 1 \ldots n_j$, $j = 1 \ldots c$.

$\underset{\sim}{Y} = \left(\begin{array}{c|c|c} Y_{11} & Y_{12} & \\ \vdots & \vdots & \cdots \\ Y_{n_1 1} & Y_{n_2 2} & \end{array} \right)_{n \times}$

$$\underset{\sim}{\theta} = E(\underset{\sim}{Y}) = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_1 \\ \hline \vdots \\ \mu_c \\ \vdots \end{pmatrix}$$

↳ Rewrite model as: $Y_{ij} = \mu_j + \varepsilon_{ij}$. $j = 1 \cdots c$, $i = 1 \cdots n_j$

w/ $\varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma^2)$

$$\left.\begin{array}{l} Y_{ij} = \mu_j + \varepsilon_{ij}, \quad i = 1 \cdots n_j, \; j = 1 \cdots c \\ \qquad\qquad\qquad\qquad (\sum_{j=1}^{c} n_j = n) \\ \varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma^2) \end{array}\right\} (**)$$

$$\longleftrightarrow \quad \underset{\sim}{Y} = \begin{pmatrix} Y_{11} \\ Y_{21} \\ \vdots \\ Y_{n_1 1} \\ Y_{12} \\ \vdots \\ Y_{n_2 2} \\ \vdots \\ Y_{n_c c} \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_1 \\ \mu_2 \\ \vdots \\ \mu_2 \\ \vdots \\ \mu_c \end{pmatrix} + \begin{pmatrix} \varepsilon_{11} \\ \varepsilon_{21} \\ \vdots \\ \varepsilon_{n_1 1} \\ \vdots \\ \vdots \\ \varepsilon_{n_c c} \end{pmatrix}$$

$$= \underbrace{\begin{bmatrix} 1 & 0 & & 0 \\ \vdots & \vdots & & \vdots \\ 1 & 0 & \cdots & \vdots \\ 0 & 1 & & \\ \vdots & \vdots & & 0 \\ 0 & 0 & & \vdots \\ & & & 1 \end{bmatrix}_{n \times c}}_{\triangleq D} \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_c \end{pmatrix}_{c \times 1} + \underset{\sim}{\varepsilon}$$

(also a multiple linear reg. model)
w/ special design matrix of rank = c

I.e. (**)

$$\Longleftrightarrow \underset{\sim}{Y}_{n \times 1} = D \cdot \underset{\sim}{\mu}_{n \times c} + \underset{\sim}{\varepsilon} \qquad \text{w/} \quad \underset{\sim}{\varepsilon} \sim N_n(0, \sigma^2 \cdot I)$$

Reduced vs Full model

$H_0: \theta_i = \beta_0 + \beta_1 X_i, \; i = 1 \cdots n \quad \Leftrightarrow \quad H_0: \mu_j = \beta_0 + \beta_1 x_j, \; j = 1 \cdots c$

vs H_a: Not $H_0 \quad \Leftrightarrow \quad H_a: \mu_j \neq \beta_0 + \beta_1 x_j$, for some j.

$\Rightarrow$ L.S.E. of $\underset{\sim}{\mu}$: $\hat{\underset{\sim}{\mu}} = (D^t D)^{-1} D^t \underset{\sim}{Y}$

$\hookrightarrow$ SSE(F) and SSE(R)

<u>note</u>

$$D^t D = \begin{bmatrix} n_1 & & 0 \\ & n_2 & \\ 0 & & \ddots \; n_c \end{bmatrix} \Rightarrow (D^t D)^{-1} = \begin{bmatrix} \frac{1}{n_1} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{n_c} \end{bmatrix}$$

$$D^t \underset{\sim}{Y} = \left(\sum_{i=1}^{n_1} Y_{i1}, \; \sum_{i=1}^{n_2} Y_{i2}, \; \cdots, \; \sum_{i=1}^{n_c} Y_{ic} \right)^t$$

$$\Rightarrow \hat{\underset{\sim}{\mu}} = \left(\frac{1}{n_1} \sum_{1}^{n_1} Y_{i1}, \; \frac{1}{n_2} \sum_{i=1}^{n_2} Y_{i2}, \; \cdots, \; \frac{1}{n_c} \sum_{i=1}^{n_c} Y_{ic} \right)^t$$

i.e. $\hat{\mu}_j, \; j = 1 \cdots, c, \; = \frac{1}{n_j} \sum_{1}^{n_j} Y_{ij}$: sample mean at $X = x_j$

$\hat{\mu}_j = \frac{1}{n_j}\sum_{i=1}^{n_j} Y_{ij} \triangleq \bar{Y}_j$: sample mean of Y's at $X = X_j$.

under $(**)$ $\Rightarrow \forall i_{=1\cdots n_j}: \hat{Y}_{ij} = \hat{\mu}_j = \bar{Y}_j$, $j = 1 \cdots c$,

$H_0: \underline{\theta} \in \Omega$ v.s. $H_1: \underline{\theta} \notin \Omega$

$\Leftrightarrow$ $H_0: \mu_j = \beta_0 + \beta_1 X_j$, $j = 1 \cdots c$ $\leftarrow$ Reduced model

v.s.

$H_1: \mu_j \neq \beta_0 + \beta_1 X_j$, some j. $\leftarrow$ Full model

$$\Rightarrow SSE(F) = \sum_{j=1}^{c}\sum_{i=1}^{n_j}(Y_{ij} - \hat{Y}_{ij})^2$$

$$= \sum_{j=1}^{c}\sum_{i=1}^{n_j}(Y_{ij} - \bar{Y}_j)^2$$

: the pure error sum of squares

$\triangleq$ SSPE ---- ($\Delta 3$)

<u>note</u> $(**)$: Y_{ij}, $i = 1 \cdots n_j$, $\overset{iid}{\sim} N(\mu_j, \sigma^2)$ for each given j.

$$\Rightarrow \frac{1}{\sigma^2}\sum_{i=1}^{n_j}(Y_{ij} - \bar{Y}_j)^2 \sim \chi^2_{n_j - 1}$$

for each $j = 1 \cdots c$

$$\Rightarrow \frac{1}{\sigma^2}\sum_{j=1}^{c}\sum_{i=1}^{n_j}(Y_{ij} - \bar{Y}_j)^2$$

$$\overset{d}{=} \sum_{j=1}^{c}\chi^2_{n_j - 1} \quad \text{indep. sum.}$$

$$= \chi^2_{\Sigma n_j - c} = \chi^2_{n-c} \quad (c < n)$$

$df(F) = n - c$.

I.e. $SSPE/\sigma^2 \sim \chi^2_{n-c}$ ....($\Delta 4$)

$SSE(R) = SSE$ under ~~(*)~~ S-L-R

$$= \sum_{j=1}^{c} \sum_{i=1}^{n_j} (Y_{ij} - \hat{Y}_{ij})^2$$

↑ fitted value of Y_{ij} under (*)

$df(R) = n-2.$ $\quad \dfrac{SSE(R)}{\sigma^2} \overset{H_0}{\sim} \chi^2_{n-2}.$

i.e. $\hat{Y}_{ij} = b_0 + b_1 X_j$, $j = 1 \cdots c$, $i = 1 \cdots n_j$.

w/ b_0, b_1 : L.S.E. of β_0, β_1 under (*)

note:
$$b_1 = \frac{\sum_{k=1}^{n} (X_k - \bar{X})(Y_k - \bar{Y})}{\sum_{k=1}^{n} (X_k - \bar{X})^2} \qquad \bar{Y} = \frac{1}{n} \sum\sum Y_{ij}$$

$$= \frac{\sum_{j=1}^{c} \sum_{i=1}^{n_j} (x_j - \bar{x})(Y_{ij} - \bar{Y})}{\sum_{j=1}^{c} \sum_{i=1}^{n_j} (x_j - \bar{x})^2}$$

$$= \frac{\sum_{j=1}^{c} n_j (x_j - \bar{x})(\bar{Y}_j - \bar{Y})}{\sum_{j=1}^{c} n_j (x_j - \bar{x})^2}$$

: ft of Y_{ij}'s only through $\underline{\bar{Y}_j\text{'s}}$.

$b_0 = \bar{Y} - b_1 \bar{X}.$ $\quad \cdots$ $\bar{Y}_j$'s ft.

$\Rightarrow \quad \hat{Y}_{ij} = b_0 + b_1 x_j = \underline{\bar{Y} + b_1 (x_j - \bar{x})}$

$\forall j = 1 \cdots c, \; i = 1 \cdots n_j.$

$\triangleq \hat{Y}_j.$

$$\Rightarrow \quad SSE(R) = \sum\sum (Y_{ij} - \hat{Y}_{ij})^2$$
$$= \sum\sum (Y_{ij} - \hat{Y}_j)^2$$
$$= \sum\sum (Y_{ij} - \bar{Y}_j + \bar{Y}_j - \hat{Y}_j)^2$$
$$(\triangle 3) = SSE(F) + \sum\sum (\bar{Y}_j - \hat{Y}_j)^2 \nearrow \text{the lack of fit SS.}$$

p.93

I.e. $SSE(R) = SSE(F) + SSLF$ $\Rightarrow F^* = \dfrac{SSE(R) - SSE(F)}{df_R - df_F} \Big/ \dfrac{SSE(F)}{df_F}$

I.e. $SSE = SSPE + SSLF$ $= \dfrac{SSLF}{c-2} \Big/ \dfrac{SSE(F)}{df_F}$

$\int (\star 1)$ $\int (\Delta 4)$ .... $(\Delta 5)$

$\sigma^2 \cdot \chi^2_{n-2, \delta} = df(R)$ $\quad \sigma^2 \chi^2_{n-c} = df(F)$ under H_1 $= \dfrac{SSLF}{c-2} \Big/ \dfrac{SSPE}{n-c}$

<u>Note</u> : $SSLF = \sum_{j=1}^{c} \sum_{i=1}^{n_j} (\bar{Y}_j - \hat{Y}_j)^2$

depends on Y_{ij}'s only through $\bar{Y}_j$'s ..

$(\because \bar{Y} = \frac{1}{n} \sum_{j=1}^{n} n_j \bar{Y}_j)$

$SSPE = \sum \sum (Y_{ij} - \bar{Y}_j)^2$

depends on $Y_{ij} - \bar{Y}_j$'s .. $j = 1 \cdots c$.

$i = 1 \cdots n_j$

But $Y_{ij} \overset{iid}{\sim} N(\mu_j, \sigma^2)$, $i = 1 \cdots n_j$.

$\hookrightarrow$ $Y_{ij} - \bar{Y}_j$, $i = 1 \cdots n_j$, $\perp$ $\bar{Y}_j$.

$\forall\ j = 1 \cdots c$

$\Rightarrow$ $SSLF \perp SSPE$

plus $(\Delta 5)$ $\Rightarrow$ $SSLF \sim \sigma^2 \cdot \chi^2_{c-2, \delta}$

$\Rightarrow$ $E(SSLF) = \sigma^2 (c - 2 + \delta)$

$= \sigma^2 (c-2) + \underbrace{\sigma^2 \delta}$

$\sum B_i^2 = 0$ under H_0

$E(SSPE) = \sigma^2(n-c)$ $\hookrightarrow E\left(\dfrac{SSPE}{n-c}\right) = \sigma^2$ $(c > 2,$

what call it pure error $\leftarrow$ always unbiased for σ^2 no matter what the rest

$\Rightarrow$ $E(SSLF/c-2) = E[SSPE/n-c) = \sigma^2$ under

$MSLF = SSLF/c-2$ tends to be larger than $\frac{SSPE}{n-c} = MSPE$ under H_1 P.74

Hence

$$\frac{SSLF}{c-2} \Big/ \frac{SSPE}{n-c} \triangleq F^*. \text{ (two)}$$

under H_0

$$\frac{\sigma^2 \chi^2_{c-2}/c-2}{\sigma^2 \chi^2_{n-c}/n-c} \quad (\because ①)$$

$$\sim F(c-2, n-c)$$

To test $H_0: \underset{\sim}{\theta} \in \Omega$ i.e. $E(Y|X=x) = \beta_0 + \beta_1 x$

v.s. NOT H_1

(assuming indep normality w/ common σ^2 ...)

rej. H_0: if $F^* = \frac{SSLF}{c-2} \Big/ \frac{SSPE}{n-c}$

$$> F(1-\alpha;\ c-2, n-c)$$

Such test has level α.

c.f. p.123 ~ 124. read!

(NOTE: Hope NOT To Reject H_0 ...

i.e. large p-value ---)

In ANOVA Table

sources	SS	df	MS	F-value	
Regression	SSR	1	MSR	$\frac{MSR}{MSE}$	$H_0: \beta_1=0$
error	SSE	$n-2$	MSE		
lack of fit	SSLF	$c-2$	MSLF	$\frac{MSLF}{MSPE}$	$H_0: E(Y\|X) = \beta_0 + \beta_1 X$
pure error	SSPE	$n-c$	MSPE		
Total	SSTO	$n-1$			

Similarly, for multiple reg. to test ($K = p-1$ predictors)

H_0 : $E(Y) = \beta_0 + \beta_1 X_1 + \cdots + \beta_K X_K$ v.s. H_1: NOT H_0

Goodness-of-fit.

c : # of groups w/ distinct sets of levels of the K predictors ($c > p$)

$\bar{Y}_j$: j-th group's group mean

$$SSPE = \sum_{j=1}^{c} \sum_{i=1}^{n_j} (Y_{ij} - \bar{Y}_j)^2$$

$$SSE = \sum \sum (Y_{ij} - \hat{Y}_{ij})^2, \quad \hat{\underset{\sim}{Y}} = X\underset{\sim}{b} = H\underset{\sim}{Y}$$

$$\Rightarrow \underset{(n-p)}{SSE} = SSPE + \underset{(c-p)}{SSLF}$$

i.e. $SSLF = SSE - SSPE$

$$MSPE = SSPE/n-c$$

$$MSLF = SSLF/c-p$$

$$F^* = MSLF / MSPE \overset{H_0}{\sim} F(c-p, n-c)$$

rej. H_0 if $F^* > F(1-\alpha; c-p, n-c)$

: level α test.

Such test : F-test for lack of fit.