

#1

a) Linearly dependent

b) Yes, when weight increase 1 unit, body fat increases 0.2434/unit

c) RSS should be still

$$(5.142)^2 \cdot 249$$

Interpretation: If we did not do anything in previous model, so when you took this out, nothing would've changed...

H_0 : Height coef = weight coef = 0.
↳ dropped two variables...

so need to use F-stat

$$\frac{RSS(M) - RSS(M)/2}{RSS(M)/249} = \frac{(5.142)^2 \cdot 249}{(5.696)^2 \cdot 249}$$

But how to get a p-value?
F-table

2)

Malb's Cp for m

$$\Rightarrow \frac{RSS(M)}{\hat{\sigma}^2} - (n - 2 - 2p(m))$$

$$= \frac{(5.696)^2 \cdot 249}{(5.142)^2} - (252 - 2 - 4)$$

$$RSS(M) = 59.545$$

$$n - 2 - 2p(m) = 252 - 2 - 4 = 246$$

Malb's Cp for M → choose still model
 $\Rightarrow \frac{RSS(M)}{\frac{RSS(M)}{n-p-1}} - (n - 2p - 2)$

$$= n - p - 1 - n + 2p + 2 = p + 1 = 5$$

M is better
↳ also adj R² is higher

#2

a) ① $\frac{0.23705}{0.02146} \approx 11.05$

② $R^2 = 1 - \frac{RSS/n-1}{TSS/n}$

$= 1 - \frac{RSS}{TSS} \cdot \frac{n}{n-1}$

$= 1 - \frac{RSS}{TSS} \cdot \frac{251}{245}$

$= 1 - \frac{(5.155)^2 \cdot 245}{15077.891} \cdot \frac{251}{245}$
 $= 1 - 10.6054 \approx 0.554$

$R^2 = 1 - \frac{RSS}{TSS} = 1 - \frac{(5.155)^2 \cdot 245}{TSS} = 0.5682$

$\hookrightarrow 1 - 0.5682 = \frac{(5.155)^2 \cdot 245}{TSS}$

$\hookrightarrow TSS = \frac{(5.155)^2 \cdot 245}{1 - 0.5682} = 15077.891$

③ d.f. \Rightarrow 6 variables

④ $n-1 = 245$

⑤ $\frac{R^2/6}{(1-R^2)/245}$
 $= 53.732$

b) standardized residual

$\frac{e_i}{\hat{\sigma} \sqrt{1-h_{ii}}}$

It seems like it will be both an outlier & influential (High leverage & outlier) \uparrow weight

Height is too short

compare to weight \hookrightarrow how to know if it's not a data giver

d) adj $R^2 \uparrow$

F-stat also goes up meaning that ~~more variables~~ \hookrightarrow cannot false out all variables

p-value for ~~height~~ goes down

Call number: QA241 .M66 1994
 Item barcode: C044434789
 Item status: IN TRANSIT
 Terminal note: 332

and harmonic analysis / Hugh L.
 Ten lectures on the interface
 between analytic number theory

Checkin Date: Fri Aug 18 2017 09:50AM
 Title:

istics

Mathematics/Stat

$\Delta e_{adj} = \frac{1}{2} - \frac{1}{2} \frac{RSS}{TSS}$
 $\hookrightarrow e_{adj} = \frac{1}{2} (1 - \frac{RSS}{TSS})$

calculate p-value for where

we know it

picture

#3

1) No Z 's biased.

$$E[\hat{\epsilon}^T \hat{\epsilon}]$$

$$\Rightarrow \text{Use } E[Z^T A Z] = \text{tr}[A \Sigma] + \mu^T A \mu$$

$$\Rightarrow \text{tr}[I \cdot \text{cov}(\hat{\epsilon})] + \mu^T A \mu$$

So, $\text{cov}(\hat{\epsilon}) = \text{cov}((I-n)y) = \sigma^2(I-n)$

$$E(\hat{\epsilon}) = E((I-n)y) = (I-n)E(y)$$

$$= (I-n)(X\beta + rZ)$$

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

$$= (X\beta - X\beta) + (I-n)rZ$$

$$= (I-n)rZ$$

$$\Rightarrow \text{tr}(\sigma^2(I-n)) + r^2 Z^T(I-n)Z$$

$$= \text{tr}(\sigma^2(I-n)) + r^2 Z^T(I-n)Z$$

$$= \sigma^2(n-p-1) + r^2 Z^T(I-n)Z$$

bias will be $\frac{r^2 Z^T(I-n)Z}{n-p-1}$

b)

$$\sum \hat{\epsilon}_i = 0 ?$$

$$X^T X \beta = X^T y$$

$$X^T (y - X\beta) = 0$$

☆ BUT NOT ZERO

$$\langle X, \hat{\epsilon} \rangle = 0$$

$$E(\hat{\epsilon}) \neq 0$$

it implies sum of residuals will not be zero

$$\frac{1}{n} \sum \hat{\epsilon}_i \neq 0$$

$$E(\hat{\epsilon}) = (I-n)rZ$$

#4
- 2l(max value of like in m)

$$y_k \sim N(X\beta, \sigma^2)$$

$$\propto \left(\prod_{k=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(y_k - X\beta)^2} \right)$$

$$\propto \left(\frac{RSS(m)}{n} \right)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2} \|y - X\beta\|^2}$$

$$\Rightarrow -\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \|y - X\beta\|^2$$

$$\Rightarrow X \rightarrow \text{then } n \log(2\pi) + n \log(\sigma^2) + \frac{1}{\sigma^2} \|y - X\beta\|^2$$

$$\Rightarrow -2l(\beta) = \frac{RSS}{\sigma^2} + n \log(2\pi\sigma^2)$$

$$= \frac{RSS}{\sigma^2} + n \log(2\pi\sigma^2)$$

And

$$-2l(\hat{\beta}, \hat{\sigma}^2) = n + n \log(2\pi) + n \log\left(\frac{RSS}{n}\right)$$

$$= n \log(e) + n \log(2\pi) + n \log\left(\frac{RSS}{n}\right)$$

$$= n \log(2\pi e) + n \log\left(\frac{RSS}{n}\right)$$

$$y = X\beta$$

#5.

$$Q(\beta) = Y^T(Y - P)$$

$$= \sum (y_i - \hat{y}_i)(1, \dots, x_{ip})$$

$$\langle X, Y - \hat{P} \rangle = 0$$



inner product = 0
means

$$X \perp Y - \hat{P}$$

Fri Aug 18 2017 09:50AM
Monographies internationales de
mathématiques modernes.

Mathematics/Statistics Periodicals

Checkin Date: Title:

to: send

QA3 .M54 v.8
C036452585
IN TRANSIT
332

Call number:
Item barcode:
Item status:
Terminal:
Hold note:

#6 -

a) 212 obs $\begin{cases} 1 \Rightarrow 19 \\ 0 \Rightarrow 133 \end{cases}$

0.288928

$(0.3131)^2$

$\text{cov}(\hat{\beta}) = (X^T W X)^{-1}$

$\frac{0.6864}{xxx} = 0.313$

① $\frac{0.6864}{0.313} \approx 2.19297$

② $\frac{-0.9050}{-4.349} \approx 0.20809$

with d.f. $n-1 = 211$

so $219.981 \Rightarrow \chi^2_{NM}$

and $\bar{y} = \frac{19}{212}$

$-2n[\bar{y} \ln \bar{y} + (1-\bar{y}) \ln (1-\bar{y})]$

③ residual dev $\Rightarrow \sum [y_i \ln p_i + (1-y_i) \ln (1-p_i)]$
 $= -2 \sum [y_i \ln \hat{p}_i + (1-y_i) \ln (1-\hat{p}_i)]$

$|AIC - 2(4p \text{ km})|$

$222.18 - 8 = 214.18$

with d.f. $n-p-1 = 212 - 4 = 208$

null dev $\Rightarrow -2 \sum [y_i \ln p_i + (1-y_i) \ln (1-p_i)] = -2 \sum [y_i \ln p + (1-y_i) \ln (1-p)]$

and $y_i \sim \text{Ber}(p)$
 and $E(y_i) = p = \bar{y}$

$-2 \sum [y_i \ln \bar{y} + (1-y_i) \ln (1-\bar{y})]$
 //

send
to: **Mathematics/Statistics**

#7

Checkin Date:

Thu Aug 17 2017 03:13PM

Title:

Naked statistics : stripping the
dread from the data / Charles
Wheelan.

Call number:

QA276 .W458 2013

Item barcode:

C110098572

Item status:

IN TRANSIT

Terminal:

332

Hold note:

a) $\beta^{(m+1)} = \beta^{(m)} + (X^T W X)^{-1} X^T (y - p)$

$\beta^{(m+1)} = (X^T W X)^{-1} X^T W z$

where ~~$\beta^{(m)} = X(y-p)$~~

$z = X\beta^{(m)} + W^{-1}(y-p)$

b) $\hat{p}_i = \frac{e^{x_i^T \beta}}{1 + e^{x_i^T \beta}}$

$\Rightarrow x_0 = \begin{bmatrix} 1 \\ \log(265) \\ \log(26) \\ 3.5 \end{bmatrix}$

$\hat{\beta} = \begin{bmatrix} 0.6864 \\ -0.9050 \\ 0.8027 \\ 1.1153 \end{bmatrix}$

Then $\hat{p}_i \approx 0.770195$

c) Residual dev \Rightarrow Decrease / stay the same
 $W = 2 \sum [y_i \hat{p}_i + (1-y_i) \hat{p}_i]$

num dev \Rightarrow same \Rightarrow only intercept

b) ~~$\hat{p}_i = \frac{e^{x_i^T \beta}}{1 + e^{x_i^T \beta}}$~~

$\sum [y_i \hat{p}_i + (1-y_i) \hat{p}_i]$

~~$= \sum [y_i \hat{p}_i] + (1-y_i) \hat{p}_i = \sum [y_i \hat{p}_i]$~~

$= \sum [y_i (x_i^T \beta)] - \ln [1 + e^{x_i^T \beta}]$

$\hat{p}_i = \frac{e^{x_i^T \beta}}{1 + e^{x_i^T \beta}}$

c) #5

d) ~~$X^T (y - p) = 0$~~

$\sum y_i = \sum \hat{p}_i$
 $\sum I(y_i = 1)$

$$l(\beta) = \sum [y_i \ln p_i + (1-y_i) \ln (1-p_i)]$$

$$\uparrow = \sum [y_i \ln(x_i^T \beta) + (1-y_i) \ln(1 + e^{x_i^T \beta})]$$

c) $-2 \ln(\text{max value in likelihood in model})$

$$= -2 \sum [y_i \ln \hat{p}_i + (1-y_i) \ln (1-\hat{p}_i)]$$

f)

pre \ obs	1	0
1	a	b
0	c	d

precision $\Rightarrow \frac{a}{a+c}$
 (= PPV)
 Recall $\Rightarrow \frac{a}{a+b}$
 (= sensitivity)

$$\Rightarrow \frac{\sum I(\hat{y}_i = 1 \cap y_i = 1)}{\sum I(\hat{y}_i = 1)}$$

$$\Rightarrow \frac{\sum I(\hat{y}_i = 1 \cap y_i = 1)}{\sum I(y_i = 1)}$$

$$\frac{\sum \hat{p}_i y_i}{\sum \hat{p}_i y_i + \sum \hat{p}_i (1-y_i)}$$

send
to:

Mathematics/Statistics Circ Desk

Checkin Date:

Thu Aug 17 2017 01:57PM

Title:

International economics : theory and policy / Paul R. Krugman, Princeton University, Maurice Obstfeld, University of California,

Call number:

HF1359 .K78 2015

Item barcode:

C111643432

Item status:

IN TRANSIT

Terminal:

330

Hold note:

full model
 $E(e) = 0$
 $Var(e) = \sigma^2 I_n$
 \Rightarrow Gauss-Markov

#8

a) $\hat{\beta} = X\beta$

$\hat{\beta}$ TRUE

b) False \Rightarrow bias \uparrow but variance \downarrow

c) False \Rightarrow we can still do permutation with no normality assumption

but p-value will be calculated based on properties of T

d) True

IWLS (= IRLS) for GLM is the same as ~~any other~~ Newton Raphson method with canonical link

e) $\hat{\beta}$

~~RSS~~

$$\hat{e}_{OLS} = \frac{1}{n-p-2} \sqrt{\frac{RSS_{OLS}}{n-p-2}}$$

$\hat{\sigma}^2$

$$= \frac{RSS_{OLS}}{n-p-2}$$

$$\hat{e}_{OLS} = y_i - x_i^T \hat{\beta}_{OLS} \perp RSS_{OLS}$$

~~OLS~~

$$\hat{e}_{OLS} = \frac{\hat{e}_i}{\sqrt{h_{ii}}} \perp RSS_{OLS}$$

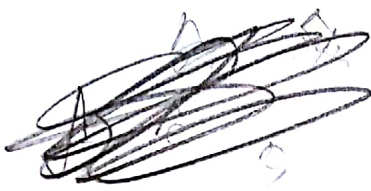
$$\hat{e}_i \perp RSS_{OLS}$$

True

$$\eta_{binomial} \rightarrow \logit$$

$$\frac{\eta_{binomial}}{\eta_{binomial}} = \eta_{binomial}$$





1) $\sum \hat{e}_i^2 = RSS$

$\hat{e}^T \hat{e}$

~~False~~

\Rightarrow always goes down when add variables...

2)

pre obs	T	F
T	a	b
F	c	d

precision $\Rightarrow \frac{a}{a+c}$

Recall $\Rightarrow \frac{a}{a+b}$

3) $\sum \hat{e}_{cv}^2$ PRESS

$\hat{e}_{cv} = y_i - x_i^T \hat{\beta}_{cv}$

\hookrightarrow leave one out CV

$\sum \left(\frac{\hat{e}_i}{1 - \frac{p+1}{n}} \right)^2$

$= \sum \left(\frac{\hat{e}_i}{1 - \frac{p+1}{n}} \right)^2$

$= \left(1 - \frac{p+1}{n} \right)^{-2} \sum \hat{e}_i^2$

$= \left(1 + 2 \frac{p(m)+1}{n} \right) RSS = RSS + 2(1+p(m)) \frac{RSS}{n}$

\hookrightarrow MLE of $\hat{\sigma}^2$

$a+b$ stays the same although cutoff changes...

PNR, TNR \uparrow
FPR, TPR \downarrow

cutoff $\uparrow \Rightarrow b, d \uparrow, a, c \downarrow$

cutoff $\downarrow \Rightarrow b, d \downarrow, a, c \uparrow$

True

$\frac{a}{a+c} + \frac{a}{a+b} \Rightarrow$ False

$\frac{RSS}{\hat{\sigma}^2} - (n-2-2p(m))$

$\frac{RSS}{\hat{\sigma}^2} = \frac{RSS - \hat{\sigma}^2(n-2-2p(m))}{\hat{\sigma}^2}$

$\approx \frac{RSS + 2\hat{\sigma}^2(p(m)+1)}{\hat{\sigma}^2}$