

Suppose that the following sample of size n-1 has been observed

1 α) ψ(β)= ||y-Bβ||2, β ∈ R°, Ψ(β)=11y-Bβ112+211β112, βεR, 220 = 119-BB112, BERP arg max (4(B))= \hat{\beta} = (\hat{\beta} \hat{\beta}) \heta y = \hat{\beta} y (\*1) => argmax (\$(B)) = (\$\bar{G} \bar{G})^{\bar{G}} \bar{G} \bar{\bar{G}}  $\bullet \ \widetilde{\mathcal{Q}} \ \widetilde{\mathcal{Q}} = \left( \mathcal{C}_{\beta} \ \nabla \ \mathcal{I}_{\gamma} \right) \left( \begin{array}{c} \mathcal{C}_{\gamma} \\ \mathcal{C}_{\gamma} \end{array} \right) = \left( \begin{array}{c} \mathcal{C}_{\gamma} \\ \mathcal{C}_{\gamma} \end{array} \right) = \left( \begin{array}{c} \mathcal{C}_{\gamma} \\ \mathcal{C}_{\gamma} \end{array} \right) + \left( \begin{array}{c} \mathcal{C}_{\gamma} \\ \mathcal{C}_{\gamma} \end{array}$ .(8°0) = 1+2 (100) = 1+2 Ir (\*2) · (B' B) B' = AT IP (B' TIP) = AT (B' TIP) with (\*1): B's = B'y agree  $(\hat{\varphi}(\rho)) = \frac{1}{412} \hat{\beta}^{LS}$ => (1)= # / Per 7= 4 ~ = ((1)  $((\lambda) = \frac{1}{4^2} = \frac{2}{5}, for \lambda = \frac{1}{2}$ cov(\(\hat{\beta}^{\end{a}}\)= 02 (\(\hat{\beta}\)\(\beta^{\sigma}\)

 $=9.5I_2=(60)$ 

I.4.8 Definition  $\label{eq:Leta} \mbox{Let } Y = B\beta + \epsilon \mbox{ be a LM and } y \mbox{ be a realisation of } Y. \mbox{ TI}$  $\psi(\beta) = \|y - B\beta\|^2$ 

is called least-squares-estimate.

- β\*(Y) is called least-squares-estimator (LSE) for
- For  $A \in \mathbb{R}^{k \times p}$ , a LSE of  $A\beta$  is defined as  $A\beta^*(Y)$ .

### I.4.9 Theorem (LSE in LM)

Given a LM with design matrix  $B \in \mathbb{R}^{n \times p}$  , a LSE  $\widehat{\beta} = \widehat{\beta}$ 

The set of all LSEs is given by  $\{\widehat{\beta} = \widehat{\beta}^+ + (I_p - B^+B)z \mid$ 

■ I.4.12 Corollary (LSE in LM)

Given a LM with design matrix B satisfying de

In case the LSE is unique, we write for short  $\widehat{\beta}$ 

## Properties of the LSE $\hat{\beta}^+$

Given a LM with design matrix B, the LSE  $\widehat{\beta}^+$  has the fol  $\mathbb{E}\widehat{\beta}^+ = \mathbb{B}^+\mathbb{B}\beta$ , Cov  $(\widehat{\beta}^+) = \sigma^2(\mathbb{B}'\mathbb{B})^+$  $@ \ \, \text{Under a NoLM}, \ \widehat{\beta}^+ \sim N_p(B^+B\beta, \sigma^2(B'B)^+) \\$ 

If the design matrix B has rank(B) = p, then the (unique)

# BLUEs and MLEs

In the situation of Theorem I.4.28, the LSE  $\hat{\beta}$  is called best linear unblased estimator (BLUE) of  $\beta$ . For  $c \in \mathbb{R}^p$ ,  $c'\hat{\beta}$  is called BLUE of  $c'\beta$ .

= 2 = B(a)

= (-4 x 5 + 12 x 6)e-12/k

= ~ + exp {-20 + 2} = ~ + exp ( 2) }

Notice that the derivation of LSEs and BLUEs as well as their means and variances do not depend on the particular distributional assumption. However, their distributions depend on the distribution of the error term r.

I.4.31 Theorem

Given a NoLM with a regular matrix B'B and unknown variance parameter  $\sigma^2>0$ ,  $\widetilde{\beta}$  is also the Maximum-Likelihood-Estimator (MLE) of  $\beta$ .

With  $P = I_n - B(B'B)^{-1}B'$ , the MLE of  $\sigma^2$  is given by  $\overline{\sigma^2} = \frac{1}{n}Y'PY$ .

moxikik,21) -> VL(d,2) = 0 loglikly hood will have Jd ~4 = 12 = -4 ~5 = 12/2 + ~4 Da2 = 12/2 same result, since loglikely hood will have the logarithm is a mondonic function ~ max (ln(L(x))=max(L(x))  $\longrightarrow \min_{\beta \in \mathbb{R}^{n}}$  (1.1) β. y) has to satisfy the normal equations 3 and B'B, respectively, a solution is given  $=y'(I_n-B(B'B)^+B')y.$   $z\in \mathbb{R}^p\}.$ 

instead of  $\widehat{\beta}^+$  for the LSE based on the Moore-Penrose

 $\operatorname{et}(B'B) > 0$ , the unique LSE is given by

CESW 9000, Cramer, Kaloni 16

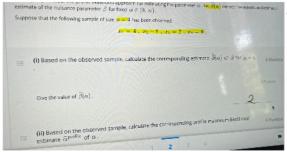
owing properties:

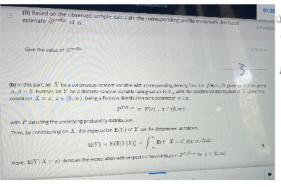
 $(B'B)^{-1}B'Y$ .

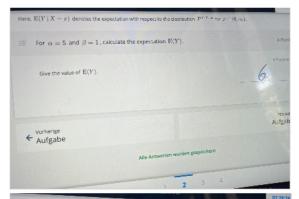
LSE  $\widehat{\beta}^+$  is an unbiased estimator for  $\beta,$ 

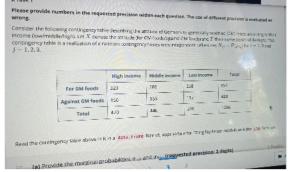
quals the inverse matrix  $(B^rB)^{-1}$ .

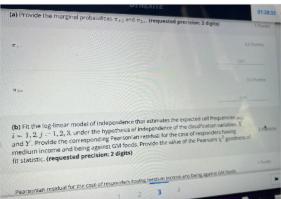
SPW 2020, Chanse, Hand ES







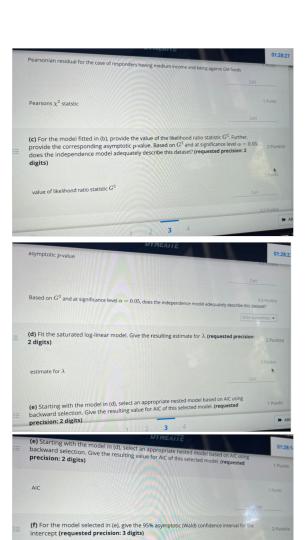


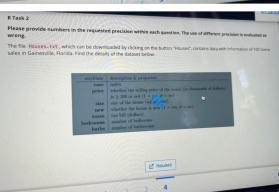


 $\int_{\mathbb{R}^{2}} \frac{1}{x^{2}} = -4 x^{2} e^{-\frac{\pi x}{2}} + x^{2} \mathbb{R}^{2} e^{-\frac{\pi x}{2}}$   $= (-4 x^{2} + 12 x^{6})e^{-\frac{\pi x}{2}} = 0$   $= (-4 x^{2} + 12 x^{6})$ 

 $= \sqrt[2]{\left[2 \times c \times b \left(\frac{2}{4} - \frac{2}{x}\right)\right] + \left(\frac{2}{x} - c \times b \left(-\frac{2}{x - x}\right)qx\right)}$ 

 $=\frac{4}{5}\left(5+\left(-25\,\text{eV}\left(\frac{4}{5}-\frac{\times}{5}\right)\right)^{2}\right)$  =30/5=(





lower bound

upper bound

(a) Load the data file into your workspace and transform the attribute new into a factor variable.

Obtain the number of houses in the dataset for which the tax bills are greater than 3000 dollars

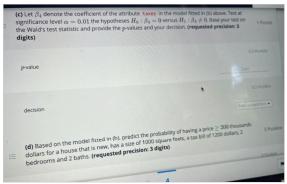
1 Punits and the number of bedrooms are at least equal to 2.

Value

(b) Fit a GLM using the canonical link from predicting the value of price using the variables size, new, taxes, bedrooms, baths as explanatory variables. Provide the resulting value for the estimated coefficient of the attribute bedrooms. (requested precision: 3 digits)

value for estimated coefficient of attribute bedrooms.

2 | Punits |
2 | Punits |
2 | Punits |
2 | Punits |
3 | Punits |
3 | Punits |
4 | Punits |
3 | Punits |
4 | Punits |
3 | Punits |
4 | Punits |
5 | Punits |
5 | Punits |
6 | Punits |
6 | Punits |
6 | Punits |
7 | Pun



(d) Based on the model fitted in (b), predict the probability of having a price $\geq$ 200 thou dollars for a house that is new, has a size of 1000 square feets, a tax bill of 1200 dollars, bedrooms and 2 baths. (requested precision: 3 digits)	sands 2 3 Punkte
	3 Punkte
predicted probability	
(e) For the model fitted in (b), compute 95% profile likelihood confidence interval for (coefficient of baths). (requested precision: 3 dig/ts)*	β <sub>6</sub> 2 Punkt
(coefficient of bactis) (Coeff	1 Punkt
t and	
lower bound	1 Punk
	Zahl
upper bound	

