

NB10 case study (the Gaussian sampling) Stat 206  
18 Feb 20

model is a 2nd fit to the data (the 0 tails are not heavy enough); let's try the  $t_r(\mu, \sigma^2)$  sampling model instead (r small: heavy tails)

① the  $t_r(\mu, \sigma^2)$

marginal sampling distribution (Gelman et al. Appendix A)

$(Z_i | \mu, \sigma, r, \theta) \stackrel{IID}{\sim} t_r(\mu, \sigma^2)$  is  
( $i = 1, \dots, n$ )  $\theta = (\mu, \sigma, r)$   $k=3$

$Z = (Z_1, \dots, Z_n)$   
 $Y = (Y_1, \dots, Y_n)$

$$P_{Z_i}(y_i | \mu, \sigma, r, \theta) = \frac{\Gamma(\frac{r+1}{2})}{\sigma \Gamma(\frac{r}{2}) \sqrt{\pi r}} \left[ 1 + \frac{1}{r} \left( \frac{y_i - \mu}{\sigma} \right)^2 \right]^{-\frac{r+1}{2}} \text{ for } y_i \in \mathbb{R}$$

② the joint sampling distribution is ②  
 therefore  $P_{\Sigma}(z | \mu, \sigma, r, T, B) = \prod_{i=1}^n P_{\Sigma_i}(y_i | \mu, \sigma, r, T, B)$

$$= \prod_{i=1}^n \frac{\Gamma(\frac{r+1}{2})}{\sigma \Gamma(\frac{r}{2}) \sqrt{\pi r}} \left[ 1 + \frac{1}{r} \left( \frac{y_i - \mu}{\sigma} \right)^2 \right]^{-\frac{r+1}{2}} \quad (*)$$

③ the likelihood function is therefore

$$L(\mu, \sigma, r | z, T, B) = c \cdot (*), \text{ and the}$$

log likelihood function is therefore

$$\ell(\mu, \sigma, r | z, T, B) = n \log \Gamma(\frac{r+1}{2})$$

$$- n \log \sigma - n \log \Gamma(\frac{r}{2}) - \frac{n}{2} \log r$$

$$- \left(\frac{r+1}{2}\right) \sum_{i=1}^n \log \left[ 1 + \frac{1}{r} \left( \frac{y_i - \mu}{\sigma} \right)^2 \right]$$

(see R code for the rest of the story)

$$p(\theta | \mathcal{B}) = \theta^s (1-\theta)^{n-s} \quad (2)$$

$(0 < \theta < 1)$   
 $k=1$

1-dim exp. family

$$p(\theta | \mathcal{B}) = f(\gamma) g(\theta)$$

$$s = \sum_{i=1}^n \gamma_i$$

$\exp \left[ \sum_{i=1}^n \eta_i(\gamma_i) \right]$   
 $\phi(\theta)$

$h_i(\gamma) = \gamma$

$$p(\theta | \mathcal{B}) = \theta^s (1-\theta)^{n-s}$$

$$= \theta^s (1-\theta)^n (1-\theta)^{-s}$$

$$= \underbrace{1 \cdot (1-\theta)^n}_{f(\gamma)} \underbrace{\left( \frac{\theta}{1-\theta} \right)^s}_{g(\theta)}$$

for any  $x > 0$   
 $x = \exp[\log x]$

$$= \underbrace{1 \cdot (1-\theta)^n}_{f(\gamma)} \underbrace{\exp \left( \sum_{i=1}^n \eta_i(\gamma_i) \right)}_{\phi(\theta)} \underbrace{\log \left( \frac{\theta}{1-\theta} \right)}_{\text{logit}(\theta)}$$

ML with  $k > 1$

$$l(\theta \sim | \mathcal{Z} \dots \mathcal{B}) \quad \checkmark \quad (4)$$

how maximize?

$$\theta \sim = (\theta_1, \dots, \theta_k)$$

$$ll(\theta \sim | \mathcal{Z} \dots \mathcal{B}) \quad \checkmark$$

$$\frac{d}{d\theta_1} ll(\theta \sim | \mathcal{Z} \dots \mathcal{B}) = 0$$

$$\frac{d}{d\theta_2} ll(\theta \sim | \mathcal{Z} \dots \mathcal{B}) = 0$$

$\vdots$

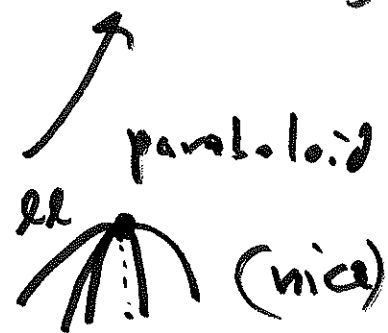
$$\frac{d}{d\theta_k} ll(\theta \sim | \mathcal{Z} \dots \mathcal{B}) = 0$$

$k$  equations

$\vdots$

$k$   $\oplus$   
unkn. how many

$\hat{\theta}_{MLE}$  is solution to  
under some conditions



often need numerical methods, <sup>(3)</sup>  
to find  $\hat{\theta}_{MLE}$ ; 2 approaches:

(1) solve system  $\odot$ , or

(2) directly (numerically)

maximize  $l(\theta | \text{---}) \sim -\ell(\theta | \text{---})$

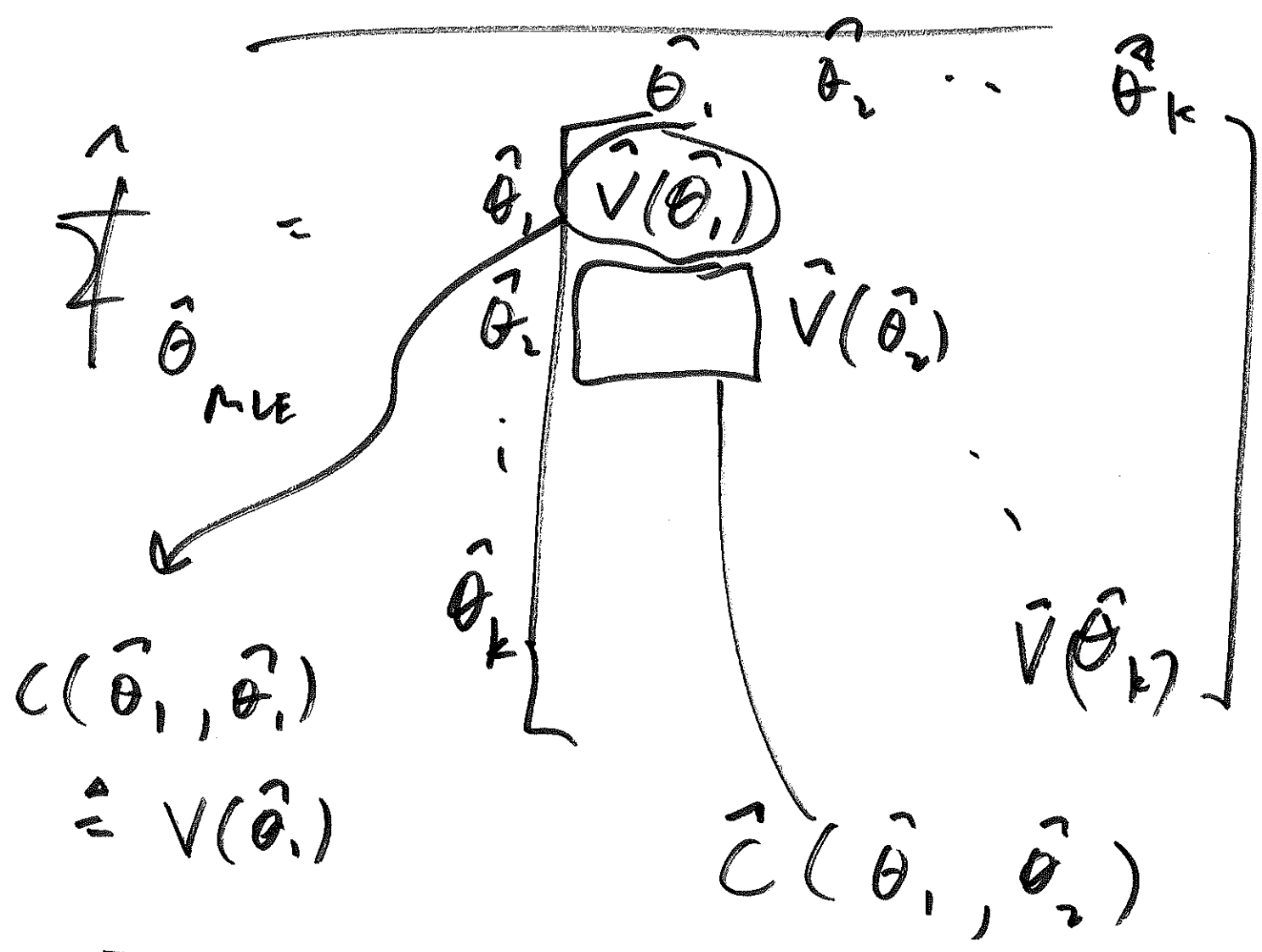
$$k=1 \quad \left( \hat{\theta}_{MLE} \right), \quad SE_{RS}^{\rightarrow} \left( \hat{\theta}_{MLE} \right) = \sqrt{\bar{V}_{RS} \left( \hat{\theta}_{MLE} \right)}$$
$$\& \quad \hat{\theta}_{MLE} \pm \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) SE_{RS}^{\rightarrow} \left( \hat{\theta}_{MLE} \right)$$

100(1- $\alpha$ )% CI (large n)

$$(k > 1) \quad \left( \hat{\theta}_{MLE} \right)_k = \left[ \left( \hat{\theta}_{MLE} \right)_1, \dots, \left( \hat{\theta}_{MLE} \right)_k \right]$$

generalization of  $\hat{V}_{RS}(\hat{\theta}_1)$   $\hat{\theta}_1$

is covariance matrix :



---


$$SE_{RS}(\hat{\theta}_j) = \sqrt{V(\hat{\theta}_j)} \quad \checkmark$$


---

$$\textcircled{k=1} \quad \hat{V}_{R5}(\hat{\theta}) = \hat{I}^{-1}(\hat{\theta}), \quad \textcircled{7}$$

$$\hat{I}(\hat{\theta}) = \left[ - \frac{d^2}{d\theta} \ell(\theta | y, \mathcal{B}) \right]_{\theta = \hat{\theta}}$$

$\textcircled{k>1}$  Fisher information matrix:

$\left( \hat{I}(\hat{\theta}) \right)_{k, k} =$  matrix of  $-2^{nd}$  partials  
eval. at MLE

$= - \hat{H}(\ell)$  otto  
Hesse  
← Hessian

est.  
 Cov. matrix of MLE  $\begin{pmatrix} \hat{\theta} \\ \hat{\sigma}^2 \end{pmatrix}$  is  $\begin{pmatrix} \hat{\theta} \\ \hat{\sigma}^2 \end{pmatrix}$

$$\text{Cov} \begin{pmatrix} \hat{\theta} \\ \hat{\sigma}^2 \end{pmatrix} = \begin{pmatrix} \hat{I}^{-1}(\hat{\theta}) \\ \hat{\sigma}^2 \end{pmatrix}$$

$$\hat{I}(\hat{\theta}) = -H(\theta) \quad \theta = \hat{\theta}_{MLE}$$

