NB10 case study (the Gaussian sampling model is a bad fit to the data (the tails are not heavy enough); let's try the $t_\nu(\mu, \sigma^2)$ sampling model instead (ν small: heavy tails)

Stat 206
18 Feb 20

① the $t_\nu(\mu, \sigma^2)$ marginal sampling distribution (Gelman et al. Appendix A)

$$(Y_i \mid \mu\, \sigma\, \nu\, \mathcal{T}\, \mathcal{B}) \stackrel{IID}{\sim} t_\nu(\mu, \sigma^2) \quad (i = 1, \dots, n) \quad \text{is}$$

$\underset{\sim}{\theta} = (\mu, \sigma, \nu)$ $k=3$

$\underset{\sim}{Y} = (Y_1, \dots, Y_n)$

$\underset{\sim}{y} = (y_1, \dots, y_n)$

$$p_{Y_i}(y_i \mid \mu\, \sigma\, \nu\, \mathcal{T}\, \mathcal{B}) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sigma\, \Gamma\left(\frac{\nu}{2}\right) \sqrt{\pi \nu}} \left[1 + \frac{1}{\nu} \left(\frac{y_i - \mu}{\sigma} \right)^2 \right]^{-\frac{\nu+1}{2}} \quad \text{for } y_i \in \mathbb{R}$$

(2) the joint sampling distribution is (2)

therefore $p_{\underset{\sim}{Y}}(\underset{\sim}{y} \mid \mu\, \sigma\, \nu\, \mathcal{T}\, \mathcal{B}) = \prod_{i=1}^{n} p_{Y_i}(y_i \mid \mu\, \sigma\, \nu\, \mathcal{T}\, \mathcal{B})$

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

---

(3) the likelihood function is therefore

$\ell(\mu, \sigma, \nu \mid \underset{\sim}{y}\, \mathcal{T}\, \mathcal{B}) = c \cdot (*)$, and the

log likelihood function is therefore

$$\ell\ell(\mu\, \sigma\, \nu \mid \underset{\sim}{y}\, \mathcal{T}\, \mathcal{B}) = n \log \Gamma\left(\frac{\nu+1}{2}\right)$$

$$- n \log \sigma - n \log \Gamma\left(\frac{\nu}{2}\right) - \frac{n}{2} \log \nu$$

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

(see R code for the rest of the story)

③

$$\ell(\theta \mid y\, B) = \underline{\underline{\theta^{s}(1-\theta)^{n-s}}} \quad \begin{pmatrix} 0<\theta<1 \\ k=1 \end{pmatrix}$$

1-dim exp. family

$$\ell(\theta \mid y \cdots B) = f(y)\, g(\theta) \cdot \exp\left[\underbrace{\phi_1(\theta)}\sum_{i=1}^{n} \underline{h_1}(y_i) \right]$$

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

$h_1(y) = y$

$$\ell(\theta \mid y\, B) = \theta^{s}(1-\theta)^{n-s}$$

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

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

for any $x > 0$

$x = \exp[\log x)$

$$= \underbrace{1}_{f(y)} \cdot \underbrace{(1-\theta)^{n}}_{g(\theta)} \exp\left[\underbrace{s}_{\sum_{i=1}^{n} h_1(y_i)} \underbrace{\overbrace{\log\left(\frac{\theta}{1-\theta}\right)}^{\text{logit}(\theta)}}_{\phi_1(\theta)} \right]$$

4

ML with $k > 1$

$\underset{\sim}{\theta} = (\theta_1, \dots, \theta_k)$

$\ell(\underset{\sim}{\theta} \mid Z \dots B)$ ✓

how maximize?

$\ell\ell(\underset{\sim}{\theta} \mid Z \dots B)$ ✓

$$\left\{\begin{aligned} \frac{\partial}{\partial \theta_1} \ell\ell(\underset{\sim}{\theta} \mid Z \dots B) &= 0 \\ \frac{\partial}{\partial \theta_2} \ell\ell(\underset{\sim}{\theta} \mid Z \dots B) &= 0 \\ &\vdots \\ \frac{\partial}{\partial \theta_k} \ell\ell(\underset{\sim}{\theta} \mid Z \dots B) &= 0 \end{aligned}\right\}$$

k equations in k unknowns (*)

$\hat{\theta}_{MLE}$ is solution to

under some conditions

paraboloid

(nice)

often need numerical methods to find $\hat{\theta}_{MLE}$; 2 approaches:

① solve system (*), or

② directly (~~num~~erically) maximize $\ell(\theta \mid —)$ or $\ell\ell(\theta \mid —)$

$\boxed{k=1}$ $(\hat{\theta}_{MLE})$, $\hat{SE}_{RS}(\hat{\theta}_{MLE}) = \sqrt{\hat{V}_{RS}(\hat{\theta}_{MLE})}$

& $\hat{\theta}_{MLE} \pm \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\hat{SE}_{RS}(\hat{\theta}_{MLE})$

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

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

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

is covariance matrix:

$$
\hat{\Sigma}_{\hat{\theta}_{MLE}} = \begin{array}{c|c}
 & \begin{array}{cccc} \hat{\theta}_1 & \hat{\theta}_2 & \cdots & \hat{\theta}_k \end{array} \\
\hline
\begin{array}{c} \hat{\theta}_1 \\ \hat{\theta}_2 \\ \vdots \\ \hat{\theta}_k \end{array} & \begin{bmatrix} \hat{V}(\hat{\theta}_1) & \boxed{} & & \\ & \hat{V}(\hat{\theta}_2) & & \\ & & \ddots & \\ & & & \hat{V}(\hat{\theta}_k) \end{bmatrix}
\end{array}
$$

$C(\hat{\theta}_1, \hat{\theta}_1) \doteq V(\hat{\theta}_1)$

$\hat{C}(\hat{\theta}_1, \hat{\theta}_2)$

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

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

$$\hat{I}(\hat{\theta}) = \left[-\frac{\partial^2}{\partial \theta^2} \, \ell\ell(\theta \mid \underset{\sim}{y} \cdot \mathcal{B}) \right]_{\theta = \hat{\theta}}$$

$k > 1$ Fisher information <u>matrix</u>:

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

$$= -\hat{H}(\ell\ell)$$

$\leftarrow$ Hessian

Otto Hesse

est.
cov. mx. of MLE $\underset{1 \times k}{(\hat{\theta})}$ is

$$\underset{k \times k}{\hat{\Sigma}_{\hat{\theta}}} = \underset{k \times k}{\left(\hat{I}^{-1}(\hat{\theta})\right)}$$

$$\hat{I}(\hat{\theta}) = \left[-H(\ell\ell)\right]_{\theta = \hat{\theta}_{MLE}}$$