AMS 206
7 Feb 19

7

this time: exchangeability & next time: conjugate modeling

$\boxed{2}$ hypothesis / significance testing (soon)

Frequentist inference

$\boxed{1}$ (point & interval estimates)

	Neyman	Fisher/...
small-sample	Quiz (~~CS~~ 3) exact CIs	? ← bootstrap
large-sample	(CS 2) Confidence intervals (CIs) based on CLT & "conf. trick"	intervals based on MLE & Fisher information

CS 1

Frank Ramsey (– 1927) died age 26

$\underline{P}(A \mid B) \in (0,1)$ — T/F prop.

$$\boxed{P(\theta = 1 \mid y_1 = 1, B)} = \frac{\underline{P}(\theta = 1 \mid B) \cdot \underline{P}(y_1 = 1 \mid \theta = 1, B)}{\underline{P}(y_1 = 1 \mid B)}$$

cs 2) $\theta \in (0,1)$ cont. r.v.

Θ space of possible θ

$p(\theta \mid y\, \mathcal{B})$ = post. dist. for θ given $y\, \mathcal{B}$

$p(\theta \mid \mathcal{B})\, p(y \mid \theta\, \mathcal{B})$

$p(y \mid \mathcal{B})$

$y = (y_1, \ldots, y_n)$

$y_i \in \{0, 1\}$

$\mathcal{B} = \{ B_1, \ldots, B_b \}$

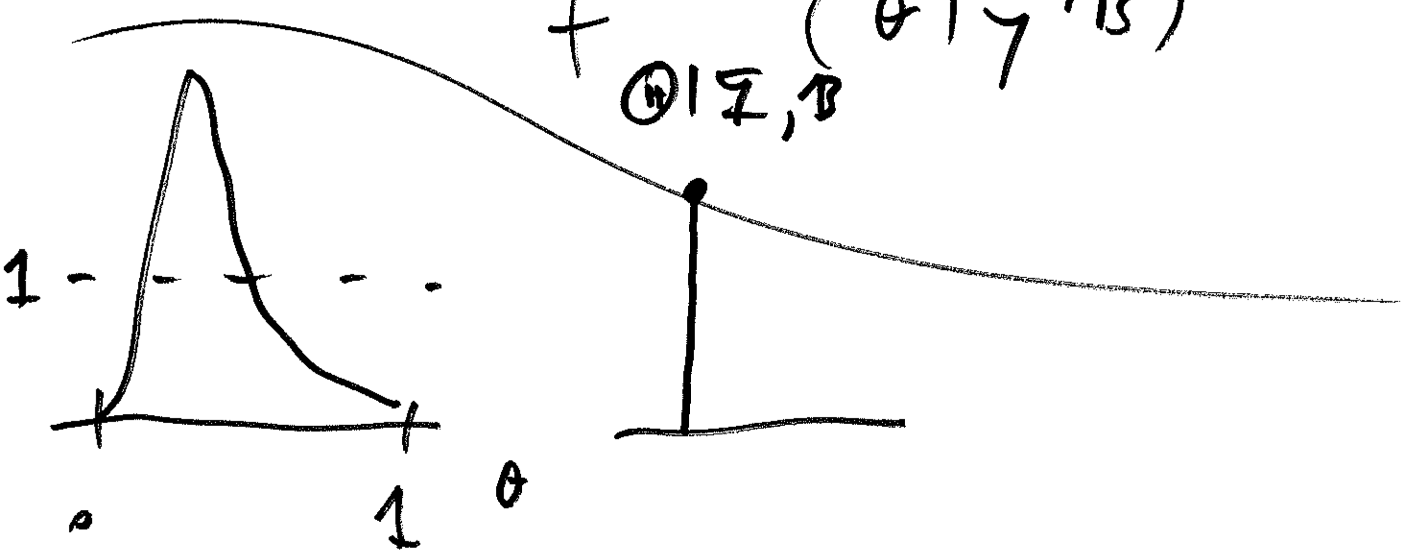
$f_{Y \mid \mathcal{B}}(y \mid \mathcal{B})$

$Y = (Y_1, \ldots, Y_n)$

$f_{Y \mid \Theta, \mathcal{B}}(y \mid \theta\, \mathcal{B})$

$f_{\Theta \mid \mathcal{B}}(\theta \mid \mathcal{B})$

$f_{\Theta \mid Y, \mathcal{B}}(\theta \mid y\, \mathcal{B})$



$$p(\theta \mid y \, \mathcal{B}) = \frac{p(\theta \mid \mathcal{B}) \cdot p(y \mid \theta \, \mathcal{B})}{p(y \mid \mathcal{B})}$$

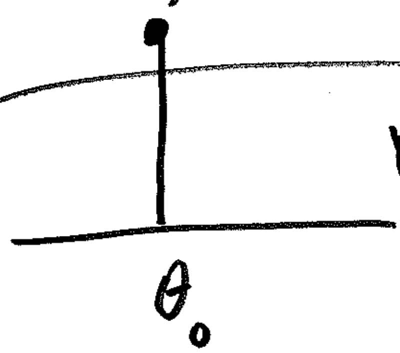
post. dist for θ given y & $\mathcal{B}$

(thm) Math fact:

under plausible Axioms (assumptions), the optimal action a^* maximizes

$$E_{(\theta \mid y \mathcal{B})} \, U(a, \theta \mid \mathcal{B}) = \text{expected utility}$$

MEU



$p(\theta \mid y \mathcal{B})$

find a to max. $U(a, \theta_0 \mid \mathcal{B})$

$$\hat{SE}(\bar{Y}) = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}}$$

④

$$(Y_i \mid \mu\,\sigma\,\mathcal{B}) \stackrel{IID}{\sim} N(\mu, \sigma^2)$$
$$(i = 1, \dots, n)$$

$y = (y_1, \dots, y_n)$

$\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i$

$$\hat{\sigma}^2_{MLE} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \bar{y})^2$$

$$\hat{\sigma}^2_{unbiased} = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})^2$$

$(\theta = \sigma^2)$

$$\text{bias}(\hat{\theta}_{\underline{MLE}}) = E(\hat{\theta}_{MLE}) - \theta$$

$$= \underline{\underline{O}}\left(\frac{1}{n}\right)$$

a function is $\underline{\underline{O}}\left(\frac{1}{n}\right)$ if when $n \uparrow \infty$, the function goes to 0 at a $\frac{1}{n}$ rate

$$\hat{SE}(\hat{\theta}_{MLE}) = \underline{\underline{O}}\left(\frac{1}{\sqrt{n}}\right)$$

$n = 10$

$s = 5$

$$
\begin{array}{cl}
\times & 1\ 1\ 1\ 1\ 1\ 0\ 0\ 0\ 0\ 0 \\
\times & 1\ 0\ 1\ 0\ 1\ 0\ 1\ 0\ 1\ 0 \\
\checkmark & 1\ 0\ 1\ 1\ 0\ 1\ 0\ 0\ 1\ 0
\end{array}
$$

$(Y_i \mid \theta\, \mathcal{B}) \stackrel{IID}{\sim} \text{Bernoulli}(\theta)$

$(i = 1, \dots, n)$

total info

info relevant to model-checking

info about θ

similarity

↓

exchangeability

$\{\underbrace{\underline{Y_1, \dots, Y_n}, Y_{n+1}, Y_{n+2}, \dots}_{\mathcal{P}}\}$

extending judgment of exch. from $(Y_1, \dots, Y_n)$ to $(Y_1, Y_2, \dots)$ is exactly like thinking of $(Y_1, \dots, Y_n)$ as an IID sample from $\mathcal{P} = (Y_1, Y_2, \dots)$

$$(Y_1, \dots, Y_n) \xrightarrow{\text{exch}} (Y_1, Y_2, \dots)$$

binary

I judge my uncertainty exchangeable

$$\to \begin{cases} \text{prior} & (\theta \mid B) \sim p(\theta \mid B) \\ \text{lik} & (Y_i \mid \theta\, B) \overset{IID}{\sim} \text{Bernoulli}(\theta) \quad (i = 1, \dots, n) \end{cases}$$

(hierarchical model)