this time: Bayesian inference with $\theta \in (0, 1)$
next time:

read: (D) ch 1–(4), App A, B; (G) ch 1, 2

AMS 206

6 Feb 18

1

de Finetti's prior

$p(Y_1 = y_1, \dots, Y_n = y_n \mid \mathcal{B})$

predictive dist.

outcomes

$y_1, y_2, \dots, y_n \mid y_{n+1}, y_{n+2}, \dots$

what we're going to see

mean θ

this is de Finetti's population p

if $\{Y_1, Y_2, \dots\}$ are exch. for Inu,
all logically internally consistent
$p(y_1, \dots, y_n \mid \mathcal{B})$ ← predictive dist.
have to look like:

$$\boxed{p(y_1, \dots, y_n \mid \mathcal{B})} = \int_0^1 \theta^s (1-\theta)^{n-s} \, p(\theta) \, d\theta,$$

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

Bernoulli sampling dist.

mixture dist

prob dist on θ

Axms 1–3

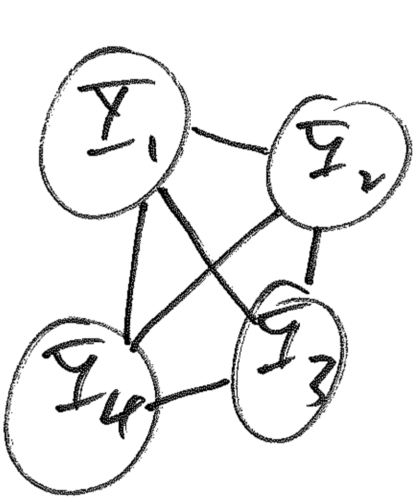
(LTP) part 1

$$p(y_1, \dots, y_n \mid \mathcal{B}) = \int_0^1 p(y_1, \dots, y_n, \theta \mid \mathcal{B}) \, d\theta$$

prior dist. (LTP) part 2

$$= \int_0^1 \underbrace{p(y_1, \dots, y_n \mid \theta \, \mathcal{B})}_{n \text{ IID Bernoulli}(\theta) \text{ draws}} \cdot \boxed{p(\theta \mid \mathcal{B})} \, d\theta$$

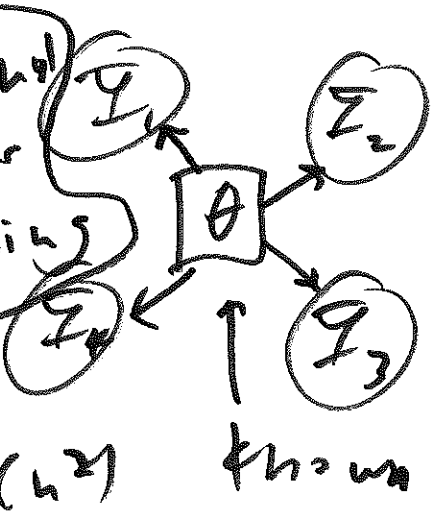
(sampling dist.)



$\binom{4}{2} = 6$

$\binom{n}{2} = \frac{n(n-1)}{2} = O(n^2)$

(n=4) conditional IID is hugely simplifying



$O(n)$

(3)

punchline

$Y_1, Y_2, \ldots$ (Bernoulli) exch.

$$\leftrightarrow \left\{ \begin{array}{l} (\theta \mid \mathcal{B}) \sim \overset{\text{prior}}{p(\theta \mid \mathcal{B})} \\ (Y_i \mid \theta \mathcal{B}) \overset{\text{IID}}{\sim} \text{Bernoulli}(\theta) \\ \quad i = 1, \ldots, n \end{array} \right\}$$

(likelihood)

Bayesian hierarchical model

$$Y \leftrightarrow \left\{ \begin{array}{c} X \\ Y \mid X \end{array} \right\}$$

Bayes's Thm. for (T/F) propositions

$$\underset{\text{unknown}}{P(A} \mid \underset{\text{data}}{D)} = \frac{P(A) \cdot P(D \mid A)}{P(D)}$$

(PDF or PMF)

more general Bayes's Thm

θ unknown $\in (0,1)$

$y = D$
$= (y_1, \ldots, y_n)$
1/0

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

Ans 131: $p_\Theta(\theta) = f_\Theta(\theta)$

④

$$p(y) \longleftrightarrow \overset{f}{p}_{\boxed{Y_1,\dots,Y_n}}(y_1,\dots,y_n) \quad \left(\begin{matrix}\text{Ans}\\ 131\end{matrix}\right)$$

$$\mathbb{P}(Y_1 = y_1, \dots, Y_n = y_n)$$

PMF because Y_i are discrete

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

more carefully

(and)

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

$$p(\theta \mid y) = \frac{p(\theta) \cdot \boxed{p(y \mid \theta)}}{p(y)} = \propto p(\theta)\, p(y \mid \theta)$$

is proportional to

f'n of θ for fixed y

$$p(\theta|y) \propto p(\theta) \cdot p(y|\theta) \propto p(\theta)\,\ell(\theta|y)$$

prior density

likelihood density

$$\boxed{\ell(\theta|y) = c\, p(y|\theta)}$$

posterior density

AMI case study

$$(\theta|\mathcal{B}) \sim \boxed{p(\theta|\mathcal{B})} \leftarrow ?$$

$$(Y_i|\theta\,\mathcal{B}) \overset{IID}{\sim} \text{Bernoulli}(\theta) \leftarrow \text{lik}$$

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

(exch. / Linearizations) + unique sampling dist. / likelihood

$$p(\theta|y) \propto p(\theta)\, \theta^{s}(1-\theta)^{n-s} \qquad s = \sum_{i=1}^{n} y_i$$

$$c\,\theta^{\alpha-1}(1-\theta)^{\beta-1} \quad (\alpha, \beta > 0)$$

$$\left(\begin{array}{l} n = 400 \\ s = 72 \end{array}\right)$$