Single - Parameter Models

Previous Example: ① parameter Space is Discrete with finite possibilities
(generic status; spell correction) ② observations . . . . Discrete . . . .

e.g. parameter $\{$ woman carrier, not $\}$

observation $\{$ son healthy, disease $\}$

$$P(\theta|y) = \frac{P(\theta)\, P(y|\theta)}{\sum_{\theta^*} P(\theta^*) P(y|\theta^*)}$$

Now: the parameter space can be continuous
observation can be anything

Example 1: estimating the bias of a coin:

- Observation: $\{ \overset{y_1}{1}, \overset{y_2}{0}, \overset{y_3}{0}, \overset{y_4}{1}, 1, 1, \cdots \}$ $y_i \in \{0, 1\}$, $i = 1, \cdots, n$

  a sequence of binary variable.

- quantity of interest (parameters): bias of a coin θ $\in (0,1)$
- inference objective: $P(\theta | y_1, y_2, \cdots, y_n)$

  (posterior distribution)

$$= \frac{P(\theta)\, P(y_1, y_2, \cdots, y_n \mid \theta)}{P(y_1, y_2, \cdots, y_n)}$$

① Specify a Model $\longrightarrow$ Likelihood

$$Y_i \overset{iid}{\sim} \text{Bernoulli}(\theta) \quad (\text{Model})$$

$$i = 1, \cdots, n$$

So $$P(y_1, y_2, \cdots, y_n \mid \theta) \overset{iid}{=} \prod_{i=1}^{n} P(y_i \mid \theta)$$

Bernouli probability

$$= \prod_{i=1}^{n} \theta^{y_i} (1-\theta)^{(1-y_i)}$$

$$= \theta^{\underbrace{\boxed{\sum_{i=1}^{n} y_i}}_{\#\text{head}}} (1-\theta)^{\underbrace{\boxed{n - \sum_{i=1}^{n} y_i}}_{\#\text{tail}}}$$

② prior distribution:

$$\theta \sim \text{Uniform}(0, 1)$$

$$P(\theta) = 1 \in \theta \in (0, 1)$$

③* $$P(y_1, y_2, \ldots, y_n) = \int_0^1 P(y_1, y_2, \ldots, y_n \mid \theta)\, P(\theta)\, d\theta$$

$$= \int_0^1 \theta^{\sum_{i=1}^n y_i} (1-\theta)^{n - \sum_{i=1}^n y_i} \, d\theta$$

$\because$ θ now is a continuous variable.

$$\int e^{-x^2/2} \, dx = \sqrt{2\pi}$$

$\downarrow$

standard normal

$$\int p(x)\,dx = 1$$

$x \sim N(0,1)$

$\downarrow$

$$\int \frac{e^{-x^2/2}}{\sqrt{2\pi}} \, dx = 1$$

$$= \mathrm{Beta}\left(\sum_{i=1}^n y_i + 1,\; n - \sum_{i=1}^n y_i + 1\right)$$

$\theta \sim \mathrm{Beta}(\alpha, \beta)$

$$p(\theta) = \frac{\theta^{\alpha-1}(1-\theta)^{\beta-1}}{\mathrm{Beta}(\alpha, \beta)}$$

④ posterior:

$$P(\theta \mid y_1, y_2, \ldots, y_n) = \frac{\theta^{\sum_{i=1}^n y_i} (1-\theta)^{n - \sum_{i=1}^n y_i}}{\mathrm{Beta}\left(\sum_{i=1}^n y_i + 1,\; n - \sum_{i=1}^n y_i + 1\right)}$$

变量 $B \Longleftarrow \mathrm{Beta}$

$$[\theta \mid y_1, \ldots, y_n] \sim \mathrm{Beta}\left(\sum_{i=1}^n y_i + 1,\; n - \sum_{i=1}^n y_i + 1\right)$$

distribution

$$\mathbb{E}[\theta \mid y_1, \ldots, y_n] = \frac{\sum_{i=1}^{n} y_i + 1}{n+2} \quad \frac{p}{p+q}$$

(posterior mean)

(in future might need to do approximation here)

$$= \underbrace{\left(\frac{\sum_{i=1}^{n} y_i}{n}\right)}_{\text{sample mean (data only)}} * \boxed{\left(\frac{n}{n+2}\right)} + \underbrace{\left(\frac{1}{2}\right)}_{\text{prior mean}} * \boxed{\left(\frac{2}{n+2}\right)}$$

Since $p(\theta) = 1$

uniform distribution

(prior only)

So posterior mean

$$= \frac{n}{n+2} * \text{Data part}$$

$$+ \frac{2}{n+2} * \text{prior part}$$

what does "2" here?

→ degree of freedom from Beta

if prior info is sufficient ⇒ assign a large weight to prior part

$$p(\theta) = 1 = \theta^{1-1}(1-\theta)^{1-1}$$

$$\sim \text{Beta}(1, 1)$$

(Specify how strong our prior is)

uniform
$\Downarrow$
Beta(1,1)

Review: what is Beta Distribution?

$\triangle$ What Happen when $n \to \infty$? $\frac{n}{n+2} \approx 1$

$\longrightarrow$ effect from prior will disappear.

only the data part explains

$$\begin{cases} \mathbb{E}(\theta \mid y_1, \cdots, y_n) \approx \dfrac{\sum y_i}{n} \\ Var(\theta \mid y_1, \cdots, y_n) = \dfrac{(\overbrace{\sum_{i=1}^{n} y_i + 1}^{a})(\overbrace{n - \sum_{i=1}^{n} y_i + 1}^{b})}{(\underbrace{n+2}_{a+b})^2 (\underbrace{n+3}_{a+b+1})} \end{cases}$$

$$\leq \frac{\frac{1}{4}(\cancel{\sum_{i=1}^{n} y_i} + 1 + n - \cancel{\sum_{i=1}^{n} y_i} + 1)^2}{(n+2)^2 (n+3)}$$

$$= \frac{1}{4} * \frac{1}{n+3} < \frac{1}{12}$$

Variance of uniform dis'

prior variance

posterior $\Rightarrow$ variance reduction to the prior.

Beta distribution:

$$f(x) = \frac{(x-a)^{p-1}(b-x)^{q-1}}{B(p,q)\,(b-a)^{p+q-1}}, \quad a \le x \le b;\ p,q > 0$$

$$B(\alpha,\beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}\,dt$$

$$\text{mean} = \frac{p}{p+q}, \quad \text{standard deviation} = \sqrt{\frac{pq}{(p+q)^2(p+q+1)}}$$