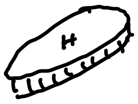


$y_1, \dots, y_n$ $\qquad$ $y_i \in \{H, T\}$ $\qquad$ $H = 1$, $T = 0$

y = # Heads
n = # flips

$\Theta = [0, 1]$ parameter space

$$\mathcal{F} := \left\{ \binom{n}{y} \theta^{y} (1-\theta)^{n-y} \;\middle|\; \theta \in \Theta \right\}$$

$\binom{n}{y} \frac{1}{2}^{y} \frac{1}{2}^{n-y}$ $\qquad$ $\binom{n}{y} \frac{1}{3}^{y} \frac{2}{3}^{n-y}$

<u>Case 1:</u> $\quad f(y) \in \mathcal{F} \quad \exists\, \theta_0 \in \Theta$ s.t. $f(y) = \binom{n}{y} \theta_0^{y} (1-\theta_0)^{n-y}$

<u>Case 2:</u> Misspecification of the model

$f(y) \notin \mathcal{F}$ $\quad$ but then there is some function $p(\theta_0 | y)$ that's the closest to $f(y)$.

Sketch Proof 1:

Assumptions:

not necessarily the smallest

1. Θ is finite $\quad \Theta = \{\theta_0, \theta_1, \dots, \theta_k\}$
2. θ_0 is the $KL(\theta)$ minimizer
3. $p(\theta = \theta_0) > 0$ it's within our prior

Aim: show $p(\theta = \theta_0 \mid y) \to 1$ as $n \to \infty$

Let $\theta \neq \theta_0$

$$\log\left(\frac{p(\theta \mid y)}{p(\theta_0 \mid y)}\right) = \log\left(\frac{p(\theta)\prod_{i=1}^{n} p(y_i \mid \theta)}{p(\theta_0)\prod_{i=1}^{n} p(y_i \mid \theta)}\right)$$

$$= \log\left(\frac{p(\theta)}{p(\theta_0)}\right) + \sum_{i=1}^{n} \log\left(\frac{p(y_i \mid \theta)}{p(y_i \mid \theta_0)}\right)$$

Observe:

$$\sum_{i=1}^{n} \log\left(\frac{p(y_i \mid \theta)}{p(y_i \mid \theta_0)}\right) = \log\big(p(y_i \mid \theta)\big) - \log\big(p(y_i \mid \theta_0)\big)$$

$$= \log\big(p(y_i \mid \theta)\big) - \log\big(f(y_i)\big) + \log\big(f(y_i)\big) - \log\big(p(y_i \mid \theta_0)\big)$$

$$= \log\left(\frac{f(y_i)}{p(y_i \mid \theta_0)}\right) - \log\left(\frac{f(y_i)}{p(y_i \mid \theta)}\right)$$

Recall y_i are i.i.d. So by the law of large numbers

$$\log\left(\frac{p(\theta|y)}{p(\theta_0|y)}\right) = \log\left(\frac{p(\theta)}{p(\theta_0)}\right) + \sum_{i=1}^{n}\left[\log\left(\frac{f(y_i)}{p(y_i|\theta_0)}\right) - \log\left(\frac{f(y_i)}{p(y_i|\theta)}\right)\right]$$

$$= \log\left(\frac{p(\theta)}{p(\theta_0)}\right) + nE\left(\log\left(\frac{f(y_i)}{p(y_i|\theta_0)}\right)\right) - nE\left(\log\left(\frac{f(y_i)}{p(y_i|\theta)}\right)\right)$$

$$= \underbrace{\log\left(\frac{p(\theta)}{p(\theta_0)}\right)}_{\text{constant}} + n\Big[\underbrace{KL(\theta_0)}_{\text{minimizer}} - KL(\theta)\Big]$$

Since $\theta \neq \theta_0$ and θ_0 is a minimizer of $KL(\cdot)$

$$\log\left(\frac{p(\theta|y)}{p(\theta_0|y)}\right) \to -\infty \quad \text{as} \quad n \to \infty$$

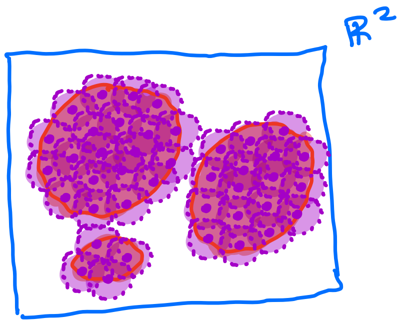
$$\Rightarrow \frac{p(\theta|y)}{p(\theta_0|y)} \to 0 \qquad \Rightarrow \quad p(\theta|y) \to 0$$

Because probabilities sum to 1, $\boxed{p(\theta_0|y) \to 1}$.

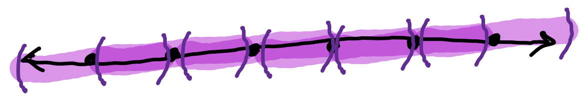
Hence, the posterior distribution converges.

A set Θ is compact if every open cover contains a finite subcover.

Since $\Theta \subseteq \mathbb{R}^n$, we picture this as closed and bounded.



Counter $\mathbb{Z}$ in $\mathbb{R}$



$\mathcal{A} = \{ (z-1, z+1) \mid z \in \mathbb{Z} \}$

I cannot delete any set in this open cover because <u>only one</u> covers each point.

<u>Sketch Proof for Theorem 2</u>

Let Θ be a compact set and define $\mathcal{A}$ to be an open cover such that <u>only one</u> set $A_0 \in \mathcal{A}$ contains θ_0.

Since Θ is compact, there exists a finite subcover. $\{A_0, A_1, \dots, A_k\}$ where A_0 is the set previously specified.

Use theorem 1's arguement.

Let $A \neq A_0$

$$\log\left(\frac{p(\theta \in A \mid y)}{p(\theta \in A_0 \mid y)}\right) \approx \log\left(\frac{p(\theta \in A)}{p(\theta \in A_0)}\right) + n\, E\left(\frac{p(y_i \mid \theta \in A)}{p(y_i \mid \theta \in A_0)}\right)$$

$p(\theta \in A \mid y) \to 0$ as $n \to \infty$

$p(\theta \in A_0 \mid y) \to 1$.