$X_i = \begin{cases} 1 & (H) \\ 0 & (T) \end{cases}$

$X_1 \perp X_2$?

$p(X_{100} = 1 \mid X_1 = 0, X_2 = 0, \ldots, X_{99} = 0)$ $>$ $<$ $=$ $.5$?

$X_i \not\perp X_j$

But exchangeable seems plausible.

de Finetti $\Rightarrow$ $X_i \perp X_j \mid \theta$ & identically distr.

$p(x_1, \ldots, x_n \mid \theta) = p(x_1 \mid \theta) p(x_2 \mid \theta) \cdots p(x_n \mid \theta)$

What could θ be here?

Ex: $p(X_i = 1) = \theta$
$p(X_i = 0) = 1 - \theta$

Together $p(X_i = x_i) = \boxed{\theta^{x_i} (1-\theta)^{1-x_i}}$

Exercise: write the joint density

* $p(x_1, \ldots, x_n \mid \theta)$

Sol'n $= \prod_{i=1}^{n} p(x_i \mid \theta)$

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

* is called (1) the joint density of the data
(2) the "data gen. model"
(3) the likelihood function

Let $y = \Sigma x_i$

$$p(Y = y \mid \theta) = \binom{n}{y} \theta^y (1-\theta)^{n-y}$$

Bayes' thm tells us

$$p(\theta) \longrightarrow p(\theta \mid y)$$

θ is called a parameter of the data generative model.

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

posterior

prior

likelihood

normalizing constant

$$\int p(y, \theta)\, d\theta$$

$$= p(y)$$

not a function of θ!

ex

$$\int_0^1 \theta x^2 \, dx$$

$$\theta \left(\frac{1}{3} x^3 \Big|_0^1 \right)$$

not a function of x!

What is a suitable prior? $p(\theta)$

Note: $0 < \theta < 1$

$\theta \sim \text{beta}(a, b)$
$\theta \sim \text{uniform}(0, 1)$

Let's examine uniform first.

Bayesian model*

$$p(y|\theta) = \binom{n}{y} \theta^y (1-\theta)^{n-y}$$

$$p(\theta) = \begin{cases} 1 & \text{if } \theta \in [0,1] \\ 0 & \text{o.w.} \end{cases}$$

* a Bayesian model is composed of a
① Likelihood (DGM)
② prior

$$p(\theta|y) = \frac{\binom{n}{y} \theta^y (1-\theta)^{n-y} \cdot 1}{C}$$

$p(\Theta = \theta | Y = y)$

$$\propto \boxed{\theta^y (1-\theta)^{n-y}}$$

$$RV^y (1-RV)^{n-y}$$

$$x^y (1-x)^{n-y}$$

$$p(\theta|y) \propto \underbrace{p(y|\theta)}_{\text{binomial}}\underbrace{p(\theta)}_{\text{beta}}$$

$$\propto \theta^{y}(1-\theta)^{n-y}\,\theta^{a-1}\cdot(1-\theta)^{b-1}$$

$$\propto \theta^{y+a-1}(1-\theta)^{n-y+b-1}$$

this is the kernel of a beta(α, β)

$$\alpha - 1 = y + a - 1$$
$$\alpha = y + a$$
$$\beta - 1 = n + b - y - 1$$
$$\beta = n + b - y$$

$$E(\theta|y) = \frac{\alpha}{\alpha+\beta} = \frac{a+y}{a+b+n}$$

$$y = \sum_{i=1}^{n} x_i \qquad \bar{x} = \frac{1}{n}\sum x_i$$

$$y = n\bar{x}$$

$$\lim_{n\to\infty} \quad \frac{a}{a+b+n} + \frac{n\bar{x}}{a+b+n}\cdot\frac{\frac{1}{n}}{\frac{1}{n}} = \frac{\bar{x}}{\frac{a}{n}+\frac{b}{n}+1}$$

$$0$$

$$\frac{\bar{x}}{1} = \bar{x}$$