Exchangability

$Y_1,\dots,Y_N$ ~~is~~ are exchangable random variables if $\pi(y_1,\dots,y_N)=\pi(y_{\sigma_1},\dots y_{\sigma_N})$ for any permutation σ of $\{1,\dots,N\}$. (Def^n 1.8)

Thm 1.1 (de Finetti) If a sequence of random variables $(Y_1,\dots Y_N)$ is exchangeable, then its joint distribution can be written as.

$$\pi(y_1,\dots,y_N)=\int\left\{\prod_{i=1}^{N}\pi(y_i|\theta)\right\}\pi(\theta)\,d\theta.$$

for some parameter θ ~~and~~, some distribution on θ, and some sampling model $\pi(y_i|\theta)$.

Questions.

1. What does exchangability mean?
   Labels of random variables do not tell us anything about the outcomes.

2. What does de Finetti's theorem tell us about Bayesian inference?
   It is a kind of existence theorem for Bayesian inference. It says if we have exchangable random variables, then a parameter θ must exist and ~~we can put on~~ a subjective distribution must also exist for θ.

3. Let $x_1,\dots,x_N\sim \mathrm{Ber}(p)$ be independent RVs. Show they are exchangeable.
   We have ~~x_i iid~~ $\pi(x_i|p)=p^{x_i}(1-p)^{1-x_i}$. Let σ be a permutation of $\{1,\dots,N\}$. The joint distribution of $x_1,\dots x_N$ is

$$\pi(x_1,\dots,x_N)=\prod_{i=1}^{N}p^{x_i}(1-p)^{1-x_i}$$
$$=p^{\sum_{i=1}^{N}x_i}(1-p)^{\sum_{i=1}^{N}1-x_i}$$
$$=p^{\sum_{i=1}^{N}x_i}(1-p)^{N-\sum_{i=1}^{N}x_i}$$
$$=p^{\sum_{i\in\sigma}x_i}(1-p)^{N-\sum_{i\in\sigma}x_i}\qquad\text{(as addition is invariant to permutations).}$$
$$=p^{\sum_{i\in\sigma}x_i}(1-p)^{\sum_{i\in\sigma}(1-x_i)}$$
$$=\prod_{i\in\sigma}p^{x_i}(1-p)^{1-x_i}$$
$$=\pi(x_{\sigma_1},\dots,x_{\sigma_N})\qquad\square$$

Priors Let $x_1, \ldots, x_N \sim Bin(M, p)$ How could we perform inference on p: MLE. Conjugate prior. Non-informative prior. Prior ellicitation. ~~Conjugate~~ Invariant prior distribution.

1. Show $p \sim Beta(\alpha, \beta)$ is conjugate.

By Bayes' theorem $\pi(p \mid \underline{x}) \propto \pi(\underline{x} \mid p)\pi(p)$

Likelihood $\pi(\underline{x} \mid p) \propto \prod_{i=1}^{N} \binom{M}{x_i} p^{x_i}(1-p)^{M-x_i}$

$$\propto p^{\sum x_i}(1-p)^{NM - \sum x_i}$$

Prior. $\pi(p) \propto p^{\alpha-1}(1-p)^{\beta-1}$

$\Rightarrow$ Posterior: $\pi(p \mid x) \propto p^{\sum x_i}(1-p)^{NM - \sum x_i} p^{\alpha-1}(1-p)^{\beta-1}$

$$\propto p^{\sum x_i + \alpha - 1}(1-p)^{NM - \sum x_i + \beta - 1}$$

$$\Rightarrow p \mid x \sim Beta(\sum x_i + \alpha,\ NM - \sum x_i + \beta).$$

As both prior and posterior distributions have same functional form. $p \sim Bet(\alpha, \beta)$ is conjugate.

2. Derive the posterior distribution for p using a non-informative prior distribution.

A non-informative prior distribution is $p \sim U[0,1]$. $\pi(p) = \begin{cases} 1 & p \in [0,1] \\ 0 & \text{otherwise} \end{cases}$

$\pi(p \mid x) \propto p^{\sum x_i}(1-p)^{NM - \sum x_i}$

$$\sim Beta(\sum x_i + 1,\ NM - \sum x_i + 1).$$

3. Derive an invariant prior distribution.

By Jeffrey's theorem $\pi(\theta) \propto \sqrt{I_\theta(y)}$ is an invariant prior distribution. We have $I_\theta(y) = -\mathbb{E}\left[\frac{d^2}{d\theta^2} \log \pi(y \mid \theta)\right]$. For the binomial distribution, we have.

$\log \pi(\underline{x} \mid p) \propto \sum x_i \log p + (NM - \sum x_i) \log(1-p)$.

$\Rightarrow \frac{d\pi(\underline{x} \mid p)}{dp} \cancel{\propto} \frac{\sum x_i}{p} - \frac{NM - \sum x_i}{1-p}$

$\Rightarrow \frac{d^2\pi(\underline{x} \mid p)}{dp^2} \cancel{\propto} -\frac{\sum x_i}{p^2} + \frac{NM - \sum x_i}{(1-p)^2}$

$\Rightarrow \mathbb{E}\left(\frac{d^2\pi(\underline{x} \mid p)}{dp^2}\right) = -\frac{\mathbb{E}(\sum x_i)}{p^2} + \frac{NM - \mathbb{E}(\sum x_i)}{(1-p)^2}$

$= -\frac{\sum \mathbb{E}(x_i)}{p^2} + \frac{NM - \sum \mathbb{E}(x_i)}{(1-p)^2}$

$= -\frac{NMp}{p^2} + \frac{NM - NMp}{(1-p)^2}$

$= NM\left(\frac{-1}{p} + \frac{1}{1-p}\right)$

$= \frac{NM}{p(1-p)}$.

$\Rightarrow \pi(\theta) \propto p^{-1/2}(1-p)^{-1/2}$