STAT 206B
Topic: Exchangeability & de Fenetti's Theorem

Winter 2022

† Exchangeability

- PH Section 2.7 (CR 3.8.2 or JB 3.5.7)
- So far we have been taking the notion of the model $y_i \stackrel{iid}{\sim} p(y \mid \theta)$ for observations $y_1, \ldots, y_n$.
- For this topic, let's focus on the actual distribution of the outcomes $(y_1, \ldots, y_n)$.
- Conditional Independence and Exchangeability.
- de Fenetti's Theorem.

PH Example p27 Participants in the 1998 General Social Survey were asked whether or not they were generally happy. Let Y_i be the random variable associated with this questions, so that

$$Y_i = \begin{cases} 1 & \text{if participant } i \text{ says that they are generally happy,} \\ 0 & \textit{otherwise}. \end{cases}$$

- Suppose someone told you the numerical value of θ, the rate of happiness among the 1272 respondents to the question.

- We assume 10 participants are sampled randomly ($10 \ll 1271$) and build a probability model for $p(y_1, \ldots, y_n)$ with $n = 10$.

PH Example p27 (contd) We may build a model as follow;

- There is some common numerical value of θ, the rate of happiness among the respondents to the question.

- It is reasonable to assume the Y_i's as conditionally independent and identically distributed given θ (or at least approximately so due to the finite population).

$$
\begin{aligned}
\Pr(Y_i = y_i \mid \theta, Y_j = y_j, i \neq j) &= \theta^{y_i}(1-\theta)^{1-y_i} \\
\Pr(Y_1 = y_1, \ldots, Y_{10} = y_{10} \mid \theta) &= \prod_{i=1}^{10} \theta^{y_i}(1-\theta)^{1-y_i} \\
&= \theta^{\sum_{i=1}^{10} y_i}(1-\theta)^{\sum_{i=1}^{10}(1-y_i)}
\end{aligned}
$$

- Let's think about a distribution $p(y_1, \ldots, y_n)$ describing the actual observables.

† PH Def 3 (or Def 3.8.2): Finite Exchangeability

Let $p(y_1, \ldots, y_n)$ be the joint density of $Y_1, \ldots, Y_n$. If

$$p(y_1, \ldots, y_n) = p(y_{\pi_1}, \ldots, y_{\pi_n})$$

for all permutations π of $\{1, \ldots, n\}$, then $Y_1, \ldots, Y_n$ are exchangeable.

- Exchangeable if reordering of $(Y_1, \ldots, Y_n)$ does not change the joint distribution of $(Y_1, \ldots, Y_n)$.

  $\Leftrightarrow$ Roughly speaking, $Y_1, \ldots, Y_n$ are exchangeable if the subscripts (the "labels" identifying the individual random quantities) are **uninformative**

- Exchangeable random variables are not necessarily independent.

† Conditional Independence (continue PH Example)

- If θ is uncertain, we describe our belief about it with $\pi(\theta)$. The marginal joint distribution of $Y_1, \ldots, Y_n$ is then

$$\begin{aligned} p(y_1, \ldots, y_{10}) &= \int_0^1 p(y_1, \ldots, y_{10} \mid \theta)\pi(\theta)d\theta \\ &= \int_0^1 \theta^{\sum_{i=1}^{10} y_i}(1-\theta)^{\sum_{i=1}^{10}(1-y_i)}\pi(\theta)d\theta \end{aligned}$$

- For example, we have

$$\begin{aligned} p(1,0,0,1,0,1,1,0,1,1) &= \int_0^1 \theta^6(1-\theta)^4\pi(\theta)d\theta \\ p(1,1,0,0,1,1,0,0,1,1) &= \int_0^1 \theta^6(1-\theta)^4\pi(\theta)d\theta \end{aligned}$$

$\Leftrightarrow$ $Y_1, \ldots, Y_n$ **are exchangeable under this model of beliefs!**

† Summary!

- **Claim** If $\theta \sim \pi(\theta)$ and $Y_1, \ldots, Y_n$ are conditionally i.i.d given θ, then marginally (unconditionally on θ), $Y_1, \ldots, Y_n$ are exchangeable.

  Proof Suppose $Y_1, \ldots, Y_n$ are conditionally iid given some unknown parameter θ. Then for any permutation π of $\{1, \ldots, n\}$ and any set of values $(y_1, \ldots, y_n) \in \mathcal{Y}^n$,

$$
\begin{aligned}
p(y_1, \ldots, y_n) &= \int p(y_1, \ldots, y_n \mid \theta)\pi(\theta)d\theta \\
&= \int \prod_{i=1}^{n} p(y_i \mid \theta)\pi(\theta)d\theta \\
&= \int \prod_{i=1}^{n} p(y_{\pi_i} \mid \theta)\pi(\theta)d\theta \\
&= p(y_{\pi_1}, \ldots, y_{\pi_n})
\end{aligned}
$$

† So far we have

$$\left.\begin{array}{l} Y_1, \ldots, Y_n \mid \theta \quad \text{i.i.d} \\ \theta \sim \pi(\theta) \end{array}\right\} \Rightarrow Y_1, \ldots, Y_n \quad \text{are exchangeable!} \tag{1}$$

Then how about the opposite direction?

† **Infinite Exchangeability** The infinite sequence of random quantities $y_1, y_2, \ldots$ is said to be judged (infinitely) exchangeable if every finite subsequence is judged exchangeable in the sense of Def 1 Finite exchangeability.

† PH Thorem 1: de Finetti's Theorem

Let $Y_i \in \mathcal{Y}$ for all $i \in \{1, 2, \ldots\}$. Suppose that for any n, our belief model for $Y_1, \ldots, Y_n$ is exchangeable:

$$p(y_1, \ldots, y_n) = p(y_{\pi_1}, \ldots, y_{\pi_n})$$

for all permutations π of $\{1, \ldots, n\}$. Then our model can be written as

$$p(y_1, \ldots, y_n) = \int p(y_1, \ldots, y_n \mid \theta)\pi(\theta) d\theta$$

for some parameter θ, some prior distribution on θ and some sampling model $p(y \mid \theta)$. The prior and sampling model depend on the form of the belief model $p(y_1, \ldots, y_n)$.

† **Continue PH example** de Finetti's Theorem says that

- There exists a probability distribution $\pi(\theta)$ such that, for every n, the joint distribution of $(Y_1, \ldots, Y_n)$ writes down as

$$p(y_1, \ldots, y_n) = \int_0^1 \prod_{i=1}^{n} \theta^{y_i}(1-\theta)^{(1-y_i)}\pi(\theta)d\theta,$$

that is, conditional on θ, the y_i's iid Beroulli $\mathrm{Ber}(\theta)$ random variables where $\pi(\theta)$ represents our beliefs about $\lim_{n\to n}\sum Y_i/n$.

$\dagger$ **For more general cases** where the y_i's are real valued and infinitely exchangeable, **de Finetti's Theorem** says that there exists an interesting representation, under the form

$$p(y_1, \ldots, y_n) = \int p(y_1, \ldots, y_n \mid \theta)\pi(\theta)d\theta.$$

This implies

- The assumption of exchangeability for every n

$$\Updownarrow$$

- There exists a likelihood $p(y_1, \ldots, y_n \mid \theta) = \prod_{i=1}^{n} p(y_i \mid \theta)$ for $y_1, \ldots, y_n$ (that is, conditionally independent).
- There exist a prior distribution on θ, $\pi(\theta)$ that represents our beliefs about the outcomes of $\{Y_1, Y_2, \ldots\}$ induced by our belief model $p(y_1, y_2, \ldots)$.

- Is infinite exchangeability reasonable?

  Possibly good approximation in many cases, but sometimes too restrictive. To be less restrictive, may consider partial exchangeability...

- Exchangeability will hold if the labels convey no information.

- When is the condition "$Y_1, \ldots, Y_n$ are exchangeable for all n" reasonable?

  For this condition to hold, we must have exchangeability and repeatability.