Notes on Foundations of Bayesian Analysis

Adam N. Smith
July 31, 2021

This note reviews some of the probabilistic foundations of the Bayesian paradigm. We focus on somewhat existential questions about prior distributions through the lens of de Finetti's representation theorem. The goal is to show that a Bayesian analysis – or more generally, the subjectivist view of probability – can be both motivated and justified by the simple belief of exchangeability.

1 Exchangeability

Exchangeability is a statement of symmetry. We can permute the "labels" indexing a sequence of random variables and the joint distribution of the data remains unchanged. In other words, the labeling and sequencing of observations do not provide additional information for the analysis. Formally, a finite sequence of random variables $X_1, \dots, X_n$ is **exchangeable** if, for any permutation of indices σ,

$$(X_1, \dots, X_n) \stackrel{d}{=} (X_{\sigma(1)}, \dots, X_{\sigma(n)}). \tag{1}$$

An infinite sequence $\{X_i\}_{i=1}^{\infty}$ is exchangeable if every finite subset of random variables is exchangeable.

While the assumption of exchangeability may seem innocuous, it is likely too restrictive for observed data in many empirical contexts. For example, observations are unlikely to be exchangeable in time series data, panel data, or cross-sectional data where differences across observations can be explained by observable covariates. Additional examples are provided in Section 4.2.2 of Bernardo and Smith (1994). That said, assumptions of exchangeability are likely to be invoked at *some stage* in a statistical analysis, perhaps conditional on a set of observed covariates.

2 de Finetti's Representation Theorem

It turns out that the belief of exchangeability encodes a lot of information about the structure of the underlying data generating processes, which also motivates a subjectivist view of probability as well as the specific "ingredients" of a Bayesian model. This is formalized in the following theorem.

Theorem 1 (de Finetti). *An infinite sequence* $\{X_i\}_{i=1}^{\infty}$ *is exchangeable if and only if there exists a random variable* $\theta \in \Theta$ *with probability distribution* $\Pi(\theta)$ *such that:*

$$p(x_1, \ldots, x_n) = \int_{\Theta} \prod_{i=1}^{n} p(x_i|\theta) d\Pi(\theta).$$

This theorem is referred to as de Finetti's representation theorem[1] and has profound implications. If the data are exchangeable, then (i) a parameter θ must exist; (ii) a likelihood must exist; (iii) a prior distribution on θ must exist. In other words, if the data are believed to be exchangeable, then the data must be a random sample from some probability model and there must exist a prior distribution over the parameters of that model. For this reason, de Finetti's representation theorem is often taken as the *raison d'être* of Bayesian analysis.

de Finetti's representation theorem can also be interpreted within the context of hierarchical models. For example, consider a set of random variables X_{ij} observed for $i = 1, \ldots, n$ sampling periods and $j = 1, \ldots, m$ cross-sectional units. While the complete collection $\{X_{ij}\}$ is unlikely to be exchangeable, we may reasonably assume exchangeability across cross-sectional units $\mathbf{X}_1, \ldots, \mathbf{X}_m$.

$$p(\mathbf{x}_1, \ldots, \mathbf{x}_m) = \int_{\Theta} \prod_{i=1}^{n} \prod_{j=1}^{m} p(x_{ij}|\theta_j) d\Pi(\theta_1, \ldots, \theta_m)$$

Assumptions of exchangeability can then be carried through to the parameters $\theta_1, \ldots, \theta_m$:

$$\pi(\theta_1, \ldots, \theta_m) = \int_{\Phi} \prod_{j=1}^{m} p(\theta_j|\phi) d\bar{\Pi}(\phi)$$

which implies that population-level parameters ϕ, which govern the probability models of the unit-level parameters, must exist and also have a prior $\bar{\Pi}(\phi)$.

[1] The first version of Theorem 1 was proven for binary sequences of random variables (de Finetti, 1930). The theorem was then extended to extended to any real-valued sequences (de Finetti, 1937), finite sequences (Diaconis and Freedman, 1980), and more general measure spaces (Hewitt and Savage, 1955).

References

Bernardo, J. M. and Smith, A. F. M. (1994). *Bayesian Theory*. John Wiley & Sons.

de Finetti, B. (1930). Funzione caratteristica di un fenomeno aleatorio. *Memorie della R. Accademia dei Lincei*, 4(5):86–133.

de Finetti, B. (1937). La prévision: ses lois logiques, ses sources subjectives. In *Annales de l'institut Henri Poincaré*, volume 7, pages 1–68.

Diaconis, P. and Freedman, D. (1980). Finite exchangeable sequences. *The Annals of Probability*, 8(4):745–764.

Hewitt, E. and Savage, L. J. (1955). Symmetric measures on cartesian products. *Transactions of the American Mathematical Society*, 80:470–501.

Schervish, M. J. (1995). *Theory of Statistics*. Springer.