STATS 370: Bayesian Statistics
Stanford University, Winter 2016

Problem Set 5

Solutions provided by Ahmed Bou-Rabee

Problem 1. Let X_i for $1\leq i < \infty$, Y_j for $1 \leq j < \infty$ be binary random variables. Suppose that for fixed values of the X_i, the Y_i are exchangeable and similarly, for fixed values of the Y_i, the X_i are exchangeable. Using theorems stated in class (or otherwise), show that for all n, m, a_i, b_i,

$$P(X_1 = a_1, \ldots, X_n = a_n, Y_1 = b_1, \ldots, Y_m = b_m) = \int_0^1 \int_0^1 p_1^S (1-p_1)^{n-S} p_2^T (1-p_2)^{m-T} \mu(d_{p_1}, d_{p_2}),$$

where $S = a_1 + \ldots + a_n$ and $T = b_1 + \ldots + b_m$.

Solution 1. We follow the steps of De Finetti's original proof.
[See, e.g., `http://wwwf.imperial.ac.uk/~das01/MyWeb/M3S3/Handouts/DeFinetti.pdf`]
Only providing details where this proof differs from the original. Let n, m, a_i, b_i be given. Let $T = \sum X_i$ and $S = \sum Y_j$. We claim that

$$P(X_1, \ldots, X_n, Y_1, \ldots, Y_m | S = s, T = t) = \frac{1}{\binom{n}{s}\binom{m}{t}}$$

Indeed, using the fact that $P(X_1, \ldots, X_n | S = s) = \frac{1}{\binom{n}{s}}$ and $P(Y_1, \ldots, Y_m | T = t) = \frac{1}{\binom{m}{t}}$

$$\begin{aligned}
P(X_1, \ldots, X_n, Y_1, \ldots, Y_m | S = s, T = t) &= P(X_1, \ldots X_n | Y_1, \ldots, Y_m, S = s, T = t) P(Y_1, \ldots, Y_m | S = s, T = t) \\
&= \frac{1}{\binom{n}{s}} \sum_X P(Y_1, \ldots, Y_m | S = s, T = t, X_1, \ldots, X_n) P(X_1, \ldots, X_n | S = s, T = \\
&= \frac{1}{\binom{n}{s}} \frac{1}{\binom{m}{t}} \sum P(X_1, \ldots, X_n | S = s, T = t) \\
&= \frac{1}{\binom{n}{s}\binom{m}{t}}
\end{aligned}$$

Let $T_n = \sum_{i=1}^n X_i$ and $S_m = \sum_{i=1}^n Y_i$ and let $N > n$ and $M > m$. Then, by partial exchangeability, we have,

$$\begin{aligned}
P(T = t, S = s) &= \sum_{s_m, t_n} p(T = t, S = s | S_m = s_m, T_n = t_n) P(S_m = s_m, T_n = t_n) \\
&= \binom{n}{s}\binom{m}{t} \sum_{s_m, t_n} \frac{(T_N)_{t_n} (S_M)_{s_m} (M - s_M)_{m - s_n}}{(N)_n M_m} p(S_m = s_m, T_n = t_n) \\
&= \binom{n}{s}\binom{m}{t} \int_0^1 \int_0^1 (\theta_1 N)_{T_n} (\theta_2 M)_{S_m} ((1-\theta)N)_{n - T_n} ((1-\theta_2)M)_{m - T_m} dQ_{M,N}(\theta_1, \theta_2)
\end{aligned}$$

Where we define $Q_{N,M}(\theta_1, \theta_2)$ on $\mathbb{R}^2$ to be the natural extension of the step function in the original proof.

Then, observe that using what we showed above

$$\begin{aligned}
P(X_1, \dots, X_n, Y_1, \dots, Y_m) &= P(X_1, \dots, X_n, Y_1, \dots, Y_m | S = s, T = t) P(S = s, T = t) \\
&= \frac{\binom{n}{s}\binom{m}{t}}{\binom{n}{s}\binom{m}{t}} \int_0^1 \int_0^1 (\theta_1 N)_{T_n} (\theta_2 M)_{S_m} ((1-\theta)N)_{n-T_n} ((1-\theta_2)M)_{m-T_m} dQ_{M,N}(\theta_1, \theta_2) \\
&= \int_0^1 \int_0^1 (\theta_1 N)_{T_n} (\theta_2 M)_{S_m} ((1-\theta)N)_{n-T_n} ((1-\theta_2)M)_{m-T_m} dQ_{M,N}(\theta_1, \theta_2)
\end{aligned}$$

Then, taking $N, M \to \infty$ and invoking Helly's theorem yields the desired claim.

Problem 2. Show that partial exchangeability described above is very different than marginal exchangeability. This is easy, we are asking for a simple counter example.

Solution 2. Consider flipping a coin n times. Let $Y_i = X_i = 1$ if coin lands heads and 0 otherwise. The sequences are marginally exchangeable because X_i, Y_i are both exchangeable sequences (as shown in class). However, $P(X_1, \dots, X_n | Y_1, \dots, Y_n) = \{0, 1\}$. Therefore, this is not conditionally exchangeable.

Problem 3. A finite version of Laplace's Law of Succession

Suppose first that $X_1, \dots, X_n$ are exchangeable binary variables that are the start of an infinite exchangeable sequence. Observing s successes out of n, show that, with a uniform prior on p, the probability of success on trial $n+1$ is $\frac{s+1}{n+2}$. Now suppose that the sample can only be extended to $n+k$ for $k \geq 1$ fixed. We showed that the law of X_i for $1 \leq i \leq n+k$ can be represented as a mixture of $n+k+1$ urn measures. Put a uniform prior over the urns, i.e., $(1/(n+k+1))$. Show that, given s successes out of the first n, the chance of success is $\frac{s+1}{n+2}$, no matter what k is.

Solution 3. We first do the infinitely exchangeable case. Suppose we observe s successes in the first n trials. Then, by the definition of conditional probability and de Finetti's theorem, then, plugging in the uniform prior we get

$$\begin{aligned}
P(X_{n+1} = 1 | X_1, \dots, X_n) &= \frac{P(X_1, \dots, X_{n+1})}{P(X_1, \dots X_n)} \\
&= \frac{\int_0^1 p^{s+1}(1-p)^{n-s} dp}{\int_0^1 p^s (1-p)^{n-s} dp} \\
&= \frac{(s+1)!(n-s)!}{(n+2)!} \frac{(n+1)!}{s!(n-s)!} \\
&= \frac{s+1}{n+2}
\end{aligned}$$

It suffices to show that in the finite exchangeable case, the joint density $P(X_1, \dots, X_{n+k})$ is the same as in the infinite exchangeable case. Indeed, if so, then the above calculation would be identical. We know that the law of X_i can be represented as a uniform mixture of urn measures. To generate a uniform mixture of $n+k+1$ possible urns, choose a probability $p \sim U(0,1)$ and draw a sample of size $n+k$ from a binomial random variable with this parameter p. Then, as Thomas Bayes showed in his original memoir, the number of successes is uniformly distributed between 0 and $n+k$. Hence,

$$P(X_1 = x_1, \dots, X_{n+k} = x_{n+k}) = \frac{1}{n+k+1} (\binom{n+k}{s})^{-1},$$

where $s = \sum_{i=1}^{n+k} x_i$, which is the same distribution as the first part.