where the summation is over $\Sigma_{i=0}^{g-1} a_i = (y-1) - (g-1) = y - g$. (Note that in this case the number of balls in the urn does not remain constant, but decreases by one each time a black ball is drawn.) The sum in the right hand side of (8) is the coefficient of t^{y-g} in the expansion of

$$\prod_{i=0}^{g-1} \left\{1 - \frac{wt}{b+w-i}\right\}^{-1} \tag{9}$$

(cf. (3)).

We find, after some straightforward calculations, that

$$\prod_{i=0}^{g-1} \left(1 - \frac{wt}{b+w-i}\right)^{-1} = \sum_{i=0}^{g-1} A_i \left(1 - \frac{wt}{b+w-i}\right)^{-1},$$

where

$$A_i = \frac{(b+w)^{(g)}}{(g-1)!} (-1)^{g-1-i} \binom{g-1}{i} \frac{1}{b+w-i}. \tag{10}$$

Hence in view of (8)

$$P[Y+g=y] = \frac{b^{(g)}}{(g-1)!} \sum_{i=0}^{g-1} (-1)^{g-i-1} \times \binom{g-1}{i} (b+w-i)^{-1} \times \left(\frac{w}{b+w-i}\right)^{y-g}$$

$$= w^{y-g} \frac{b^{(g)}}{(g-1)!} \sum_{i=0}^{g-1} (-1)^{g-i-1} \times \binom{g-1}{i} (b+w-i)^{-(y-g+1)},$$

or using finite difference notation:

$$P[Y+g=y] = w^{y-g} \frac{b^{(g)}}{(g-1)!} \Delta^{g-1}(b+w+0)^{-(y-g+1)},$$

(Δ operates on the 0).

For the case $\beta_w = \omega_w = \omega_b = 0$; $\beta_b = -s$ we obtain

$$P[Y+g=y] = \frac{\left(\frac{w}{s}\right)^{y-g} \left(\frac{b}{s}\right)^{(g)}}{(g-1)!} \Delta^{g-1}\left(\frac{b+w}{s} - 0\right)^{-(y-g+1)}$$

REFERENCES

[1] Eggenberger, F. and Pólya, G.: Über die Statistik verketteter Vorgänge, *Z. für Ang. Math. und Mech.* **3** (1923), 279–289.

[2] Johnson, N. L. and Kotz, S.: On a multivariate generalized occupancy model, *J. Appl. Prob.*, **13**, (1976), 392–399.

[3] Kaminsky, K. S., Luks, E. M. and Nelson, P. I.: An urn model and the prediction of order statistics, *Comm. Statist.* **4** (1975), 245–250.

[4] Styve, B.: A variant of Pólya urn model, *Nordisk. Mat. Tidsk.* **13** (1965), 147–152 (in Norwegian).

[5] Uppuluri, V. R. R. and Carpenter, J. A.: A generalization of the classical occupancy problem, *J. Math. Anal. Appl.*, **34** (1971), 316–324.

De Finetti's Theorem on Exchangeable Variables

DAVID HEATH* AND WILLIAM SUDDERTH**

A simple proof is given for de Finetti's theorem that every sequence of exchangeable 0–1 random variables is a probability mixture of sequences of independent, identically distributed variables. The proof can easily be presented to seniors or first year graduate students of mathematical statistics and should aid them in understanding the relationship between the classical and the Bayesian point of view.

For identically distributed random variables, exchangeability (see the definition below) is a weaker condition than independence. For example, suppose that an urn contains m balls, r of which have a "1" written on them and $m - r$ of which have a "0". If these balls are drawn from the urn, one at a time at random without replacement, and X_k denotes the digit on the k^{th} ball drawn, then the sequence $X_1, \ldots, X_m$ is an exchangeable sequence; however, the random variables are not independent.

Definition. The random variables $X_1, \ldots, X_n$ are *exchangeable* if the $n!$ permutations $(X_{k_1}, \ldots, X_{k_n})$ have the same n-dimensional probability distribution. The variables of an infinite sequence (X_n) are *exchangeable* if $X_1, \ldots, X_m$ are exchangeable for each m.

Consider a sequence of experiments (say, coin tosses) whose outcomes are the Bernoulli variables $X_1, X_2, \ldots$. The classical assumption is that the X_n's are independent with a common distribution. However, for a Bayesian statistician, the observation of X_j will quite likely result in a change of opinion about the distribution of X_k. Thus the natural assumption for a Bayesian is not independence, but

* Dept. of Operations Research, Cornell Univ., Ithaca, NY 14850. Research supported by National Science Foundation Grant NSF GP 43085.

** Dept. of Theoretical Statistics, Univ. of Minnesota, 270 Vincent Hall, Minneapolis, MN 55455. Research supported by National Science Foundation Grant MPS 75-06173.

exchangeability. These ideas were explored in depth by Bruno de Finetti [1], who discovered the following beautiful theorem.

Theorem. To every infinite sequence of exchangeable random variables (X_n) having values in $\{0, 1\}$, there corresponds a probability distribution F concentrated on $[0, 1]$ such that

$$\begin{aligned} P\{X_1 = 1, \ldots, X_k = 1,& \\ X_{k+1} = 0, \ldots, X_n = 0\}& \qquad (1) \\ = \int_0^1 \theta^k(1-\theta)^{n-k}F(d\theta)& \\ \text{for all } n \text{ and } 0 \le k \le n.& \end{aligned}$$

The distribution F may be regarded as the prior for the random parameter Θ. In fact, if Θ has distribution F and if, given $\Theta = \theta$, $X_1, X_2, \ldots$ are independent Bernoulli variables with parameter θ, then (1) holds. For further discussion of the theorem, see Savage [5, section 3.7].

Several proofs of the theorem are already known, but existing proofs use concepts more complicated than exchangeability or provide little probabilistic intuition. De Finetti [1] and Feller [3] give proofs which rely on the consideration of moment sequences while Hewitt and Savage [4] prove more general results using the notions of convexity and extreme points. The proof given below is essentially the one sketched by de Finetti in [2, pp. 215–216] except for the limit argument.

One counter-intuitive observation is that, if one deletes the word "infinite", the statement of the theorem becomes false. (A simple example is provided by the exchangeable variables X_1, X_2 where $P\{X_1 = 1, X_2 = 0\} = \frac{1}{2} = P\{X_1 = 0, X_2 = 1\}$.) However, a pretty result *is* true for finite exchangeable sequences. It is presented in the lemma below which shows that every such sequence is a mixture of urn sequences of the type described in the first paragraph. By applying the lemma to truncations of an infinite sequence and passing to the limit, one can obtain the theorem.

Lemma. Let $p_{k,n}$ denote the quantity on the left side of equation (1). Suppose $X_1, \ldots, X_m$ is a sequence of exchangeable random variables with values in $\{0, 1\}$ and set $q_r = P\{\sum_{j=1}^m X_j = r\}$. Then

$$p_{k,n} = \sum_{r=0}^{m} \frac{(r)_k(m-r)_{n-k}}{(m)_n} q_r \text{ for } 0 \le k \le n \le m, \qquad (2)$$

where $(x)_k = \prod_{j=0}^{k-1}(x-j)$.

Proof. It follows from the assumption of exchangeability that, given the event $[\sum_{j=1}^m X_j = r]$, all possible arrangements of the r ones among the m places are equally likely. In other words, given that there are r ones, the distribution of $X_1, \ldots, X_m$ is the same as that obtained by drawing from an urn containing r ones and $m - r$ zeroes. Thus the r^{th} term of the series on the right side of (2) is simply

$$P\left\{X_1 = 1, \ldots, X_r = 1, \; X_{r+1} = 0, \ldots, X_m = 0 \mid \sum_{j=1}^m X_j = r\right\} \times P\left\{\sum_{j=1}^m X_j = r\right\},$$

and hence, (2) is clear.

It remains to be shown that the theorem is simply the limiting case (as m tends to infinity) of the lemma:

Let (X_n) be as in the statement of the theorem; apply the lemma to $X_1, \ldots, X_m$ and rewrite equation (2) to get

$$p_{k,n} = \int_0^1 \frac{(\theta m)_k((1-\theta)m)_{n-k}}{(m)_n} F_m(d\theta) \qquad (3)$$

where F_m is the distribution function concentrated on $\left\{\frac{r}{m}: 0 \le r \le m\right\}$ whose jump at $\frac{r}{m}$ is q_r.

Now apply Helly's theorem [3, Theorem VIII.6.1] to get a subsequence F_{m_j} which converges in distribution to a limit, say F. Since, as m tends to infinity, the integrand in (3) tends uniformly in θ to the integrand on the right side of (1), equation (1) holds for this F.

The advanced reader may wish to check that very similar methods will prove the more general result that every sequence of exchangeable (not necessarily 0–1) variables is a mixture of sequences of independent, identically distributed variables. To get started, first check that every finite sequence of exchangeable variables is a mixture of urn sequences by conditioning on the value of the order statistic.

Acknowledgement. We were motivated to look for a simple proof of de Finetti's theorem by the remarks of Dennis Lindley that such a proof was desirable and should be possible.

REFERENCES

1. Bruno de Finetti, Foresight: its logical laws, its subjective sources, in *Studies in Subjective Probability* (Kyburg, H. E., Jr., and Smokler, H. E., Eds.), Wiley, New York (1964) pp. 93–158. (First appeared as La prevision: ses lois logiques, ses sources subjectives, *Ann. Inst. H. Poincaré,* (1937) 7, pp. 1–68).
2. Bruno de Finetti, *Probability, Induction, and Statistics,* Wiley, New York (1972).
3. W. Feller, *An Introduction to Probability Theory and Its Applications,* v. II, Wiley, New York (1966).
4. E. Hewitt and L. J. Savage, Symmetric measures on Cartesian products, *Trans. Amer. Math. Soc.*, v. 80, (1956) pp. 470–501.
5. Leonard J. Savage, *The Foundations of Statistics,* Dover, New York (1972).