THE FREQUENCY COUNT OF A MARKOV CHAIN AND THE TRANSITION TO CONTINUOUS TIME

BY I. J. GOOD

Admiralty Research Laboratory, Teddington, England

1. Introduction. Consider a chain, N letters long, generated by a discrete-time Markov process that has a finite number, t, of available states. Each state will be called a "letter", and the set of t states the "alphabet". I shall discuss the joint probability distribution of the frequencies of the t letters of the alphabet, in other words the probability distribution of the "frequency count" of the chain, by making use of what may be called a "pseudo probability generating function". The discussion makes use of the interesting method of multiple contour integration, previously used by Whittle for another problem concerning Markov chains. I shall then apply a transition to continuous time. For the case $t = 2$, the result for continuous time is already known, but our result is more general; and it is of interest to relate the theories of discrete and continuous time.

The main results are given by formulae (3), (8), (9), and (10). Formula (8), for example, gives the covariance between the frequencies of any pair of letters when the chain is ergodic and is in its stable state; formula (9) gives a neat expression for the variance of the number of 0's when $t = 2$, and shows clearly how it differs from the familiar result for binomial sampling; and formula (10) provides, in principle, the joint density function for the durations of the t states when time is continuous, and where the chain is not necessarily in a stable state.

I believe this paper is of interest largely for its methods. I have not found it convenient to present it in the conventional theorem-proof form.

2. Frequency Counts of a Markov Chain. Let the matrix of transition probabilities be $Q = (q_{\mu,\nu})$ $(\mu, \nu = 0, 1, \cdots, t - 1)$. Let p_r be the probability that the first letter of the chain is r $(r = 0, 1, \cdots, t - 1)$. These need not be stable-state probabilities. Let $p(\mathbf{n})$ be the probability that the letter frequency count will be $\mathbf{n} = (n_0, n_1, \cdots, n_{t-1})$, where $n_0 + n_1 + \cdots + n_{t-1} = N$. The probability generating function (P.G.F.) of the frequency count is (cf., [1])

$$\sum p(\mathbf{n})\mathbf{x}^{\mathbf{n}} = \sum p(n_0, n_1, \cdots, n_{t-1}) x_0^{n_0} x_1^{n_1} \cdots x_{t-1}^{n_{t-1}}$$

summed over all $\mathbf{n}$ for which $n_0 + \cdots n_{t-1} = N$. If however the summation is over all $\mathbf{n}$ for which $n_0 + \cdots + n_{t-1}$ is positive, then the result may be called the "universal" P.G.F., and it serves for all positive values of N simultaneously.

Let $\mathbf{e}$ be the column vector consisting of t 1's, and let $\mathbf{X}$ be the diagonal matrix diag $(x_0, \cdots, x_{t-1})$. Then it is easy to check that the P.G.F. is (cf., [1]

$$(1) \qquad (p_0x_0, p_1x_1, \cdots, p_{t-1}x_{t-1})(\mathbf{QX})^{N-1}\mathbf{e},$$

Received September 29, 1959; revised October 29, 1960.