SOME ASPECTS OF THE EMIGRATION-IMMIGRATION PROCESS[1a]

By Harold Ruben[1b]

Columbia University

1. Introduction and summary. A multivariate emigration-immigration or a Poisson-Markoff process (Bartlett [2], Ruben and Rothschild [9], Patil [8]; cf. also Bartlett [3], p. 78) is a vector stochastic process $\mathbf{n}(t) = (n_1(t), n_2(t), \cdots, n_m(t))$ in continuous time t which is described by the following properties: There exists a complex of non-negative time-independent parameters, $\lambda_{rs}(r, s = 1, 2, \cdots, m, r \neq s)$, λ_r^* and $\mu_r(r = 1, 2, \cdots, m)$ such that the probability of a change of $\mathbf{n}(t)$ to

$$(n_1(t), \cdots, n_r(t) - 1, \cdots, n_s(t) + 1, n_{s+1}(t), \cdots, n_m(t))$$

in the time interval $(t, t + h)$ is $\lambda_{rs} n_r(t)h + o(h)$, for h sufficiently small; the probability of a change of $\mathbf{n}(t)$ to

$$(n_1(t), \cdots, n_r(t) + 1, \cdots, n_s(t) - 1, n_{s+1}(t), \cdots, n_m(t))$$

is $\lambda_{sr} n_s(t)h + o(h)$; the probability of a change of $\mathbf{n}(t)$ to

$$(n_1(t), \cdots, n_r(t) - 1, \cdots, n_m(t))$$

is $\lambda_r^* n_r(t)h + o(h)$; the probability of a change of $\mathbf{n}(t)$ to

$$(n_1(t), \cdots, n_r(t) + 1, \cdots, n_m(t))$$

is $\mu_r h + o(h)$; the probability of no change is

$$1 - \{\sum_r \mu_r + \sum_r (\lambda_r^* + \sum_{s \neq r} \lambda_{rs}) n_r(t)\}h + o(h).$$

Finally, it is assumed that the above probabilities are independent of the past realization of the process. It is readily seen that these assumptions imply that the process is Markovian and strongly stationary.

It will be convenient for our purposes to visualize the Poisson-Markoff process more concretely in the following manner: Consider a system consisting of $m + 1$ states $\{E_1, E_2, \cdots, E_m, E^*\}$ such that at any point in time an "individual" is in one of these $m + 1$ states. Thus, $\{E_1, \cdots, E_m\}$ may represent m stages of a certain disease while E^* represents the healthy state.[2] A population of individuals is being studied such that the number of individuals in E^* at any instant of time is effectively infinite. Let $\lambda_{rs}h + o(h)$ be the probability that an individual in

---

Received February 20, 1961; revised September 25, 1961.

[1a] Research sponsored in part under Contract AF 41(657)-214, Teachers College, Columbia University.

[1b] Present address: Department of Statistics, The University, Sheffield, England.

[2] It is likely that the Poisson-Markoff process may find application in telephone traffic problems and in genetics (cf., Feller [4], pp. 413–414, and Malécot [6]).

state E_r at time t shall be in E_s at time $t + h$, $\lambda_r^* h + o(h)$ the probability that an individual in E_r at time t shall be in E^* at time $t + h$, while $\mu_r h + o(h)$ is the probability that *precisely one* of the (infinitely many) individuals in E^* at time t shall be in E_r at time $t + h$; further, assume that the probability of more than one interchange (of an individual from any state to any other state) in $(t, t + h)$ is $o(h)$ and that there is no interaction between the individuals. Let $n_r(t)$ be the number of individuals in E_r at time t. Then clearly the various probabilities enumerated in the first paragraph are precisely the probabilities of the various transitions of $\mathbf{n}(t)$ in the infinitesimal interval $(t, t + h)$. In this connexion, it should be noted that the emigration-immigration process is a natural generalization of the time-homogeneous birth and death process, in the sense that when $m = 1$, the former process reduces to a birth and death process, with λ^* and μ representing the death and birth parameters. For $m > 1$, the λ's and μ's may be regarded as the instantaneous mean interaction-rates between the states $E_1, \cdots, E_m, E^*$.

Since, as remarked previously, the process $\{\mathbf{n}(t)\}$ is stationary, the expectation of $\mathbf{n}(t)$ is independent of t. Let, then, $\boldsymbol{\nu} \equiv (\nu_1, \cdots, \nu_m) = E\mathbf{n}(t)$. Then the vector $\boldsymbol{\nu}$ together with the interaction parameters ((2.1) and (2.2)) serve to specify the system completely. (Observe that $\boldsymbol{\nu}$ is functionally related to the interaction parameters through (2.2).) From (2.8), $\boldsymbol{\nu}$ may also be regarded as a limiting vector $(t \to \infty)$ which represents the "steady state" configuration of the system in the specific stochastic equilibrium sense implied by this equation (a configuration is defined by the vector $\mathbf{n}(t)$), while the parameters determine the approach of the system, again in a probabilistic sense, from an initial configuration to the equilibrium configuration.

In Section 2, some basic properties of the Poisson-Markoff process are listed, including the mean lifetime and recurrence time of any configuration in both discrete and continuous time. Section 3 contains the main result of this paper, viz., the joint distribution of a finite set of observations of $\mathbf{n}(t)$ in discrete time from the point of view of the joint factorial moment generating function (f.m.g.f.)[3] of this distribution. These results are of intrinsic theoretical interest and will be utilized in a subsequent paper dealing with the estimation of the fundamental interaction parameter in a singly-infinite class of emigration-immigration processes.

2. Some basis properties of the emigration-immigration process. Define the $m \times m$ matrix $\boldsymbol{\Lambda} = ((\Lambda_{ij}))$ by

$$\begin{aligned} \Lambda_{ij} &= -\lambda_{ij}, & j \neq i, \\ &= \lambda_i^* + \sum_{l \neq i} \lambda_{il}, & j = i \end{aligned} \tag{2.1}$$

($\boldsymbol{\Lambda}$ may be regarded as a fundamental probability interaction-rate matrix for

[3] The f.m.g.f. is here the most convenient transform. Other transforms, such as the probability generating and characteristic functions, can, of course, be obtained directly from the f.m.g.f. by simple transformations of the dummy variables.

the process), and the row-vector $\mathbf{v}$ by

$$\boldsymbol{\mu} = \mathbf{v}\boldsymbol{\Lambda}. \tag{2.2}$$

(It will be shown presently that $\mathbf{v}$ as defined by (2.2) is consistent with its previous definition as the expected value of $E\mathbf{n}(t)$.)
Then the conditional f.m.g.f. of $\mathbf{n}(t_2)$ for given $\mathbf{n}(t_1)$, $t_1 < t_2$, is

$$\begin{aligned} \phi(\boldsymbol{\alpha}_2 \mid \mathbf{n}(t_1)) &\equiv E\left\{\prod_{j=1}^{m} (1 + \alpha_{2j})^{n_j(t_2)} \mid \mathbf{n}(t_1)\right\} \\ &= \exp\,[\mathbf{v}(\mathbf{I} - \mathbf{P}(t_2 - t_1))\boldsymbol{\alpha}_2'] \prod_{j=1}^{m} \{(1 + \mathbf{p}_j(t_2 - t_1)\boldsymbol{\alpha}_2')^{n_j(t_1)}\}, \end{aligned} \tag{2.3}$$

where $\boldsymbol{\alpha}_2 = (\alpha_{21}, \alpha_{22}, \cdots, \alpha_{2m})$, $\mathbf{I}$ is the unit $m \times m$ matrix, $\mathbf{p}_j(t_2 - t_1)$ is the jth row-vector of the $m \times m$ matrix $\mathbf{P}(t_2 - t_1) \equiv ((P_{ij}(t_2 - t_1)))$ which is defined as follows:

$$\mathbf{P}(\tau) = \boldsymbol{\Theta}^{-1}\mathbf{K}(\tau)\boldsymbol{\Theta}. \tag{2.4}$$

Here $\boldsymbol{\Theta}$ is the normalizing matrix which reduces $\boldsymbol{\Lambda}$ to diagonal, canonical form, while $\mathbf{K}(\tau)$ is the diagonal matrix with diagonal elements $\exp(-\kappa_1\tau), \cdots, \exp(-\kappa_m\tau)$, the κ_i being the eigenvalues of $\boldsymbol{\Lambda}$; these m eigenvalues are all real and positive.

On letting $t_2 \rightarrow \infty$ in (2.3), the f.m.g.f. of $\mathbf{n}(t_2)$ is obtained as $\exp(\mathbf{v}\boldsymbol{\alpha}_2')$, i.e. the $n_i(t_1)$ $(i = 1, 2, \cdots, m)$ are independent Poisson variables with means ν_i. Hence on multiplying $\phi(\boldsymbol{\alpha}_2 \mid \mathbf{n}(t_1))$ by $\prod_{i=1}^{m}(1 + \alpha_{1i})^{n_i(t_1)}$ and taking expectations with respect to $\mathbf{n}(t_1)$ the joint f.m.g.f. of $\mathbf{n}(t_1)$ and $\mathbf{n}(t_2)$ is obtained as

$$\begin{aligned} \psi(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2) &\equiv E\left\{\prod_{i=1}^{m} (1 + \alpha_{1i})^{n_i(t_1)} \prod_{j=1}^{m} (1 + \alpha_{2j})^{n_j(t_2)}\right\} \\ &= \exp\,[\mathbf{v}(\boldsymbol{\alpha}_1' + \boldsymbol{\alpha}_2' + \mathbf{A}_1\mathbf{P}(t_2 - t_1)\boldsymbol{\alpha}_2')], \end{aligned} \tag{2.5}$$

where $\boldsymbol{\alpha}_1 = (\alpha_{11}, \alpha_{12}, \cdots, \alpha_{1m})$ and $\mathbf{A}_1$ is the $m \times m$ diagonal matrix with diagonal elements $\alpha_{11}, \alpha_{12}, \cdots, \alpha_{1m}$.

From (2.5), the expectation and autocovariance function of the process are obtained as

$$E\mathbf{n}(t) = \mathbf{v} \tag{2.6}$$

and

$$E\{(\mathbf{n}(t) - \mathbf{v})(\mathbf{n}(t + \tau) - \mathbf{v})'\} = \mathbf{N}\mathbf{P}(\tau), \tag{2.7}$$

where $\mathbf{N}$ is the $m \times m$ diagonal matrix with diagonal elements $\nu_1, \nu_2, \cdots, \nu_m$. In particular, if the ν_i are equal ($\mathbf{N} = \nu\mathbf{I}$, where ν is the common value of the ν_i), the autocorrelation function of the process is $\mathbf{P}(\tau)$.

In view of the physical interpretation of $\mathbf{v}$ provided by (2.6) ($\mathbf{v}$ was originally defined quite formally in terms of the fundamental interaction parameters through (2.2)), (2.2) can in its turn be interpreted as a relation which reflects the property of statistical equilibrium; specifically, in terms of the example of a system of individuals discussed in the introductory section, (2.2) shows that the instantaneous mean rate of entry of individuals into the state $E_i(i = 1, 2, \cdots, m)$ at time t is equal to the instantaneous mean rate of exit of individuals from E_i for all t. Again, from the Markovian nature of the process and with the aid of (2.3), we have, for $t_1 < t_2 < \cdots < t_k$,

$$E\{(\mathbf{n}(t_k) - \boldsymbol{\nu}) \mid \mathbf{n}(t_1), \mathbf{n}(t_2), \cdots, \mathbf{n}(t_{k-1})\} \tag{2.8}$$
$$= E\{(\mathbf{n}(t_k) - \boldsymbol{\nu}) \mid \mathbf{n}(t_{k-1})\} = (\mathbf{n}(t_{k-1}) - \boldsymbol{\nu})\mathbf{P}(t_k - t_{k-1}),$$

i.e., the regression of $\mathbf{n}(t_k)$ on $\mathbf{n}(t_{k-1})$ is linear. Equation (2.8) indicates that $\mathbf{P}(\cdot)$ may be interpreted further as a mean subsidence matrix, inasmuch as it determines the manner in which an arbitrary initial configuration subsides towards the equilibrium or mean configuration with lapse of time. (Observe that $\lim_{t\to\infty} \mathbf{P}(t) = \mathbf{0}$.)

The mean lifetimes and recurrence times of the process $\mathbf{n}(t)$ in both discrete and continuous time may be obtained from known formulae. In fact, (see e.g., Bartlett [3], p. 182 and Kac [5], p. 171) the mean lifetime and mean recurrence time, $T_\tau(\mathbf{n})$ and $\Theta_\tau(\mathbf{n})$, respectively, of a specified configuration $\mathbf{n} \equiv \mathbf{n}(t)$ with $t = \cdots, -2\tau, -\tau, 0, \tau, 2\tau, \cdots$, are given by

$$T_\tau(\mathbf{n}) = \tau[1 - p_\tau(\mathbf{n} \mid \mathbf{n})]^{-1} \tag{2.9}$$

and

$$\Theta_\tau(\mathbf{n}) = \tau(1 - p(\mathbf{n}))[p(\mathbf{n})(1 - p_\tau(\mathbf{n} \mid \mathbf{n}))]^{-1}, \tag{2.10}$$

where $p(\mathbf{n})$ is the probability of $\mathbf{n}$ (the product of m Poisson probabilities with parameters $\nu_1, \nu_2, \cdots, \nu_m$), while $p_\tau(\mathbf{n} \mid \mathbf{n})$ is the transition probability from $\mathbf{n}$ to $\mathbf{n}$ in a time interval of width τ. The latter probability may be expressed as a complicated series via the conditional probability generating function directly implied by (2.3) (such a series is a generalization of the series obtained by Patankar [7] for $m = 1$; see also Aitken [1], pp. 94–95). However, this will not be obtained here since the mean lifetime and recurrence time of the process $\mathbf{n}(t)$ in continuous time $(-\infty < t < \infty)$ are of greater interest, these being obtained by letting $\tau \to 0$ in (2.9) and (2.10), respectively. Thus

$$T_0(\mathbf{n}) \equiv \lim_{\tau\to 0} T_\tau(\mathbf{n}) = [\sum_r \mu_r + \sum_r n_r\lambda_r^* + \sum_{r\neq s}\sum n_r\lambda_{rs}]^{-1} \tag{2.11}$$

and

$$\Theta_0(\mathbf{n}) \equiv \lim_{\tau\to 0} \Theta_\tau(\mathbf{n}) = T_0(\mathbf{n}) \left\{\left[\prod_{i=1}^{m} e^{-\nu_i}\nu_i^{n_i}/n_i\right]^{-1} - 1\right\}. \tag{2.12}$$

We conclude with some additional properties of the matrix $\mathbf{P}(\cdot)$. Observe first that $\mathbf{P}(\cdot)$ satisfies the functional relationship

$$\mathbf{P}(t_1 + t_2) = \mathbf{P}(t_1)\mathbf{P}(t_2) \tag{2.13}$$

for all t_1 and t_2, whence for any polynomial $U(\cdot)$ with scalar coefficients

$$U(\mathbf{P}(\tau)) = \mathbf{\Theta}^{-1}U(\mathbf{K}(\tau))\mathbf{\Theta}, \tag{2.14}$$

$$[U(\mathbf{P}(\tau))]^{-1} = \mathbf{\Theta}^{-1}[U(\mathbf{K}(\tau))]^{-1}\mathbf{\Theta}, \tag{2.15}$$

and, more generally, for any non-singular rational function $V(\cdot)$ with scalar coefficients,

$$V(\mathbf{P}(\tau)) = \mathbf{\Theta}^{-1}V(\mathbf{P}(\tau))\mathbf{\Theta}, \tag{2.16}$$

while for any non-singular rational functions $V_1(\cdot)$, $V_2(\cdot)$ with scalar coefficients,

$$V_1(\mathbf{P}(\tau))V_2(\mathbf{P}(\tau)) = \mathbf{\Theta}^{-1}V_1(\mathbf{K}(\tau))V_2(\mathbf{K}(\tau))\mathbf{\Theta} = V_2(\mathbf{P}(\tau))V_1(\mathbf{P}(\tau)). \tag{2.17}$$

Finally, the trace of $V(\mathbf{P}(\tau))$, $\mathrm{tr}\,[V(\mathbf{P}(\tau))]$, and the determinant of $V(\mathbf{P}(\tau))$ are given by

$$\mathrm{tr}\,[V(\mathbf{P}(\tau))] = \sum_{i=1}^{m} V(e^{-\kappa_i\tau}), \tag{2.18}$$

$$\| V(\mathbf{P}(\tau)) \| = \prod_{i=1}^{m} V(e^{-\kappa_i\tau}). \tag{2.19}$$

It will be noted that by (2.16) any element of $V(\mathbf{P}(\tau))$ is obtained by substituting $V(\exp(-\kappa_i\tau))$, $i = 1, 2, \cdots, m$, in the corresponding element of $\mathbf{P}(\tau)$. From (2.17), a similar rule holds for the general element of the product matrix $V_1(\mathbf{P}(\tau))V_2(\mathbf{P}(\tau))$.

3. The joint f.m.g.f. and the moments of a finite set of observations on $\mathbf{n}(t)$. We shall prove the following

THEOREM. *Let $\mathbf{n}(t_i) \equiv \mathbf{n}_i$ be k consecutive observations of $\mathbf{n}(t)$ with $t_1 < t_2 < \cdots < t_k$. Then the joint* f.m.g.f. *of $\mathbf{n}_1, \mathbf{n}_2, \cdots, \mathbf{n}_k$ is*

$$\begin{aligned}\psi(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_k) = \exp\Bigg[\mathbf{v}\Bigg\{\sum_{r=1}^{k}\sum_{\gamma} \mathbf{A}_{\gamma_1}\mathbf{P}(t_{\gamma_2} - t_{\gamma_1})\mathbf{A}_{\gamma_2}\mathbf{P}(t_{\gamma_3} - t_{\gamma_2}) \\ \cdots \mathbf{A}_{\gamma_{r-2}}\mathbf{P}(t_{\gamma_{r-1}} - t_{\gamma_{r-2}})\mathbf{A}_{\gamma_{r-1}}\mathbf{P}(t_{\gamma_r} - t_{\gamma_{r-1}})\boldsymbol{\alpha}'_{\gamma_r}\Bigg\}\Bigg],\end{aligned} \tag{3.1}$$

where $\boldsymbol{\alpha}_i = (\alpha_{i1}, \alpha_{i2}, \cdots, \alpha_{im})$, $\mathbf{A}_i$ is the $m \times m$ diagonal matrix with diagonal elements $\alpha_{i1}, \alpha_{i2}, \cdots, \alpha_{im}$, $i = 1, 2, \cdots, k$, $\gamma_1, \gamma_2, \cdots, \gamma_r$ is a selection of distinct integers from the set $\{1, 2, \cdots, k\}$ and $\sum_\gamma$ denotes summation over all distinct r-selections.

PROOF. The proof is by induction. Let $\mathbf{n}_{k+1}$ be a further observed value of $\mathbf{n}(t)$ at $t = t_{k+1} > t_k$. Then

$$\begin{aligned}\psi(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_{k+1}) &= \underset{(k+1)}{E}\left[\prod_{i=1}^{k+1}\prod_{j_i=1}^{m}(1 + \alpha_{ij_i})^{n_{ij_i}}\right] \\ &= \underset{(k)}{E}\left[\prod_{i=1}^{k}\prod_{j_i=1}^{m}(1 + \alpha_{ij_i})^{n_{ij_i}}\phi(\boldsymbol{\alpha}_{k+1} \mid \mathbf{n}_k)\right],\end{aligned} \tag{3.2}$$

where $E_{(k)}$ and $E_{(k+1)}$ are expectation operators with respect to $\mathbf{n}_1, \cdots, \mathbf{n}_k$ and $\mathbf{n}_1, \cdots, \mathbf{n}_{k+1}$, respectively, use having been made of the Markovian property of $\{\mathbf{n}_i\}$. Also, $\phi(\alpha_{k+1} \mid \mathbf{n}_k)$ is the conditional f.m.g.f. of $\mathbf{n}_{k+1}$ for fixed $\mathbf{n}_k$ and is given by (2.3). Thus,

$$\psi(\alpha_1, \cdots, \alpha_{k+1}) = \underset{(k)}{E}\left\{\prod_{i=1}^{k}\prod_{j_i=1}^{m}(1+\alpha_{ij_i})^{n_{ij_i}} \exp[\mathbf{v}(\mathbf{I} - \mathbf{P}(t_{k+1}-t_k)\alpha'_{k+1}] \cdot \prod_{j=1}^{m}(1+\mathbf{p}_j(t_{k+1}-t_k)\alpha'_{k+1})^{n_{kj}}\right\}$$

$$(3.3) \qquad = \exp[\mathbf{v}(\mathbf{I} - \mathbf{P}(t_{k+1}-t_k))\alpha'_{k+1}] \underset{(k)}{E} \cdot\left\{\prod_{i=1}^{k-1}\prod_{j_i=1}^{m}(1+\alpha_{ij_i})^{n_{ij_i}}\prod_{j=1}^{m}(1+\alpha_{kj}+\mathbf{p}_j(t_{k+1}-t_k)\alpha'_{k+1} + \alpha_{kj}\mathbf{p}_j(t_{k+1}-t_k)\alpha'_{k+1})^{n_{kj}}\right\}.$$

Since the second term of the right-hand member of (3.3) is equal to the joint f.m.g.f. of $\mathbf{n}_1, \cdots, \mathbf{n}_k$, except that α'_k is replaced by

$$\alpha'_k + \mathbf{P}(t_{k+1}-t_k)\alpha'_{k+1} + \mathbf{A}_k\mathbf{P}(t_{k+1}-t_k)\alpha'_{k+1},$$

the following recursion relationship is obtained:

$$(3.4) \qquad \psi(\alpha_1, \cdots, \alpha_{k+1}) = \exp[\mathbf{v}(\mathbf{I} - \mathbf{P}(t_{k+1}-t_k)\alpha'_{k+1}]\psi(\alpha_1, \cdots, \alpha_{k-1}, \alpha_k + \alpha_{k+1}\mathbf{P}(t_{k+1}-t_k) + \alpha_{k+1}\mathbf{P}(t_{k+1}-t_k)\mathbf{A}_k).$$

Assume now that (3.1) is true for some integral k. Then

$$(3.5) \qquad \psi(\alpha_1, \cdots, \alpha_k) = \psi(\alpha_1, \cdots \alpha_{k-1}) \exp[\mathbf{v}\{1 + \sum_{(k-1)}\mathbf{A}_\beta\mathbf{P}(t_k - t_\beta) + \sum_{(k-1)}\sum\mathbf{A}_\beta\mathbf{P}(t_\gamma - t_\beta)\mathbf{A}_\gamma\mathbf{P}(t_k - t_\gamma) + \cdots + \mathbf{A}_1\mathbf{P}(t_2-t_1)\mathbf{A}_2\mathbf{P}(t_3-t_2)\cdots\mathbf{A}_{k-1}\mathbf{P}(t_k - t_{k-1})\}\alpha'_k],$$

where $(k-1)$ means that the various subscripts $\beta, \gamma, \cdots$ are positive integers which are members of the set $\{1, 2, \cdots, k-1\}$ and $\sum \cdots \sum$ means that all distinct selections of integers are to be included, it being assumed that the subscripts have, for convenience, been arranged after selection in increasing order of magnitude. On applying (3.5) in (3.4),

$$(3.6) \qquad \psi(\alpha_1, \cdots, \alpha_{k+1}) = \psi(\alpha_1, \cdots, \alpha_{k-1})\exp[\mathbf{v}(\mathbf{I} - \mathbf{P}(t_{k+1}-t_k))\alpha'_{k+1}] \exp\{\mathbf{v}[(1 + \sum_{(k-1)}A_\beta\mathbf{P}(t_k - t_\beta) + \sum_{(k-1)}\sum\mathbf{A}_\beta\mathbf{P}(t_\gamma - t_\beta)\mathbf{A}_\gamma\mathbf{P}(t_k - t_\gamma) + \cdots + \mathbf{A}_1\mathbf{P}(t_2-t_1)\cdots\mathbf{A}_{k-1}\mathbf{P}(t_k-t_{k-1})) (\alpha'_k + \mathbf{P}(t_{k+1}-t_k)\alpha'_{k+1} + \mathbf{A}_k\mathbf{P}(t_{k+1}-t_k)\alpha'_{k+1})]\}.$$

The general term of order $r + 1$ (containing r $\mathbf{A}$'s and 1 α'), $r = 1, 2, \cdots, k - 1$, in the expression within the second square bracket on the right-hand side of Equation (3.6) is

$$
\begin{aligned}
&\sum_{(k-1)} \mathbf{A}_{\beta_1}\mathbf{P}(t_{\beta_2} - t_{\beta_1})\mathbf{A}_{\beta_2}\mathbf{P}(t_{\beta_3} - t_{\beta_2}) \cdots \mathbf{A}_{\beta_r}\mathbf{P}(t_k - t_{\beta_r})\alpha_k' \\
&+ \sum_{(k-1)} \mathbf{A}_{\beta_1}\mathbf{P}(t_{\beta_2} - t_{\beta_1})\mathbf{A}_{\beta_2}\mathbf{P}(t_{\beta_3} - t_{\beta_2}) \cdots \\
(3.7) \quad &\mathbf{A}_{\beta_r}\mathbf{P}(t_k - t_{\beta_r})\mathbf{P}(t_{k+1} - t_k)\alpha_{k+1}' \\
&+ \sum_{(k-1)} \mathbf{A}_{\gamma_1}\mathbf{P}(t_{\gamma_2} - t_{\gamma_1})\mathbf{A}_{\gamma_2}\mathbf{P}(t_{\gamma_3} - t_{\gamma_2}) \cdots \\
&\mathbf{A}_{\gamma r-1}\,\mathbf{P}(t_k - t_{\gamma_r})\mathbf{A}_k\mathbf{P}(t_{k+1} - t_k)\alpha_{k+1}'.
\end{aligned}
$$

Note that the second component of (3.7) may be simplified by setting

$$\mathbf{P}(t_k - t_{\beta_r})\mathbf{P}(t_{k+1} - t_k) = \mathbf{P}(t_{k+1} - t_{\beta_r})$$

with the aid of (2.13).

Now the first component of (3.7) gives the sum of terms, each of which contains r $\mathbf{A}$'s, chosen from the set $\{\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_{k-1}\}$, as well as α_k'; the second component gives the sum of all terms, each of which contains r $\mathbf{A}$'s, chosen from the set $\{\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_{k-1}\}$, as well as α_{k+1}'; the third component gives the sum of terms, each of which contains $r - 1$ $\mathbf{A}$'s, chosen from the set $\{\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_{k-1}\}$, as well as $\mathbf{A}_k$ and α_{k+1}'. Combining these terms with the terms of order $r + 1$ from $\psi(\alpha_1, \alpha_2, \cdots, \alpha_{k-1})$, each of which contains r $\mathbf{A}$'s chosen from the set $\{\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_{k-1}\}$, as well as one α' from the set $\{\alpha_1', \alpha_2', \cdots, \alpha_{k-1}'\}$, it is seen that the right-hand member of (3.6) includes all terms of order $r + 1$ such that each of them contains r $\mathbf{A}$'s from the set $\{\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_{k+1}\}$ together with one α' from the set $\{\alpha_1', \alpha_2', \cdots, \alpha_{k+1}'\}$.

There are only two terms of order 1 in the expression within the second square bracket in Equation (3.6). These are α_k' and $\mathbf{P}(t_{k+1} - t_k)\alpha_{k+1}'$ which contribute a portion $\exp\{\mathbf{v}(\alpha_k' + \mathbf{P}(t_{k+1} - t_k)\alpha_{k+1}')\}$. This combined with

$$\exp\{\mathbf{v}(\mathbf{I} - \mathbf{P}(t_{k+1} - t_k))\alpha_{k+1}'\}$$

gives $\exp\{\mathbf{v}(\alpha_k' + \alpha_{k+1}')\}$. Combining this term, in its turn, with the terms of order 1 from $\psi(\alpha_1, \alpha_2, \cdots, \alpha_{k-1})$, the right-hand member of (3.6) will be seen to include all terms of order 1 when the set of subscripts from which selection is permitted is $\{1, 2, \cdots, k + 1\}$.

Finally, the expression within the square bracket contains a single term of order $k + 1$, viz.,

$$\mathbf{A}_1\mathbf{P}(t_2 - t_1)\mathbf{A}_2\mathbf{P}(t_3 - t_2) \cdots \mathbf{A}_{k-1}\mathbf{P}(t_k - t_{k-1})\mathbf{A}_k\mathbf{P}(t_{k+1} - t_k)\alpha_{k+1}',$$

while $\psi(\alpha_1, \alpha_2, \cdots, \alpha_{k-1})$ contains no such term.

Combining these results, we deduce that the right-hand member of (3.6) includes all possible terms of order $1, 2, \cdots, k + 1$, when the set of subscripts from which selection is permitted is $\{1, 2, \cdots, k + 1\}$. Thus (3.1) is true for $k + 1$

provided it is true for k, and since (3.1) is valid for $k = 2$ (see (2.5)), it is valid for all positive integral k. This completes the proof of the theorem.

We shall now discuss the nature of the joint distribution of $\mathbf{n}_1, \mathbf{n}_2, \cdots, \mathbf{n}_k$ and at the same time derive expressions for the moments of the distribution. The function $\psi(\alpha_1, \alpha_2, \cdots, \alpha_k)$ may in fact be identified as the f.m.g.f. of a multivariate correlated Poisson distribution of order mk. A distribution of this type may be regarded as a limiting case of the sampling distribution from a multiple dichotomy. Consider an h-dimensional random vector $\boldsymbol{\xi} = (\xi_1, \xi_2, \cdots, \xi_h)$, whose components can assume the values 0 and 1 only, and where, for any selection of r distinct integers $\{\gamma_1, \gamma_2, \cdots, \gamma_r\}$, $r = 0, 1, \cdots, h$, from the set $\{1, 2, \cdots, h\}$,

$$(3.8) \qquad \text{Prob}\begin{Bmatrix} \xi_j = 1, & j = \gamma_1, \gamma_2, \cdots, \gamma_r \\ \quad = 0, & \text{otherwise} \end{Bmatrix} = p_{(\gamma_1, \gamma_2, \cdots, \gamma_r)},$$

the latter quantities being the probabilities associated with the 2^h possible values of $\boldsymbol{\xi}$. The joint f.m.g.f. for N independent observations on $\boldsymbol{\xi}$ is

$$\phi(\alpha_1, \alpha_2, \cdots, \alpha_h) = \left[\sum_{r=0}^{h} \sum_{\gamma} p_{(\gamma_1, \gamma_2, \cdots, \gamma_r)}(1 + \alpha_{\gamma_1})(1 + \alpha_{\gamma_2}) \cdots (1 + \alpha_{\gamma_r})\right]^N,$$

$\sum_\gamma$ indicating summation over all possible distinct r-selections. Thus

$$\phi(\alpha_1, \alpha_2, \cdots, \alpha_h) = (g_0 + \sum_i g_i\alpha_i + \sum_{i<j} g_{ij}\alpha_i\alpha_j + \cdots + g_{123\cdots h}\alpha_1\alpha_2 \cdots \alpha_h)^N,$$

where for a selection $\{i_1, i_2, \cdots, i_\beta\}$ from the set $\{1, 2, \cdots, h\}$, $\beta = 0, 1, \cdots, h$,

$$g_{i_1 i_2 \cdots i_\beta} = p_{(i_1, i_2, \cdots, i_\beta)} + \sum_{l_1}{}' p_{(i_1, i_2, \cdots, i_\beta, l_1)}$$
$$+ \sum_{l_1}{}' \sum_{l_2}{}' p_{(i_1, i_2, \cdots, i_\beta, l_1, l_2)} + \cdots + p_{(1,2,3,\cdots,h)}, \quad \beta = 0, 1, 2, \cdots h,$$

the l's denoting distinct members from the complementary selection $\{l_1, l_2, \cdots, l_{h-\beta}\}$ and $\sum'_{l_1} \sum'_{l_2} \cdots \sum'_{l_j} p_{(i_1, i_2, \cdots, i_\beta, l_1, l_2, \cdots, l_j)}$ being the sum of probabilities effected over all possible j-combinations from the complementary selection, $j = 1, 2, \cdots, h - \beta$. Consequently, if the marginal probabilities be denoted by $p^{(i_1, i_2, \cdots, i_\beta)}$,

$$(3.9) \qquad p^{(i_1, i_2, \cdots, i_\beta)} = \text{Prob}\{\xi_{i_1} = 1, \xi_{i_2} = 1, \cdots, \xi_{i_\beta} = 1\},$$

the factorial cumulant generating function for the sample of N observations is given by

$$\log \phi(\alpha_1, \alpha_2, \cdots, \alpha_h)$$
$$= N \log (1 + \sum_i p^{(i)}\alpha_i + \sum_{i<j} p^{(i,j)}\alpha_i\alpha_j + \cdots + p^{(1,2,\cdots,h)}\alpha_1\alpha_2 \cdots \alpha_h)$$
$$= N[(\sum_i p^{(i)}\alpha_i + \sum_{i<j} p^{(i,j)}\alpha_i\alpha_j + \cdots + p^{(1,2,\cdots,h)}\alpha_1\alpha_2 \cdots \alpha_h)$$
$$- \tfrac{1}{2}(\cdots)^2 + \tfrac{1}{3}(\cdots)^3 - \cdots].$$

If now $p^{(i_1, i_2, \cdots, i_\beta)} = O(N^{-1})$ as $N \to \infty$, for $\beta = 1, 2, \cdots, h$, then

$$\lim_{N\to\infty} \log \phi(\alpha_1, \alpha_2, \cdots, \alpha_h) = \sum_i c_i\alpha_i + \sum_{i<j} c_{ij}\alpha_i\alpha_j + \cdots + c_{12\cdots h}\alpha_1\alpha_2 \cdots \alpha_h, \tag{3.10}$$

where

$$\begin{aligned} c_{i_1 i_2 \cdots i_\beta} = \lim_{N\to\infty} N[&S_1(i_1, i_2, \cdots, i_\beta) \\ &- S_2(i_1, i_2, \cdots, i_\beta) + 2!S_3(i_1, i_2, \cdots, i_\beta) \\ &- \cdots + (-)^{\beta-1}(\beta - 1)!S_\beta(i_1, i_2, \cdots, i_\beta)], \end{aligned} \tag{3.11}$$

$S_j(i_1, i_2, \cdots, i_\beta)$ denoting the sum of all terms, each of which is a product of p's with superscripts such that these superscripts form a complete partitioning of the set $\{i_1, i_2, \cdots, i_\beta\}$ into j distinct subsets, $j = 1, 2, \cdots, \beta$. We have thus shown that the limiting form of the f.m.g.f., $\phi^*(\alpha_1, \alpha_2, \cdots, \alpha_h)$, say, is given by

$$\phi^*(\alpha_1, \alpha_2, \cdots, \alpha_h) = \exp\{\sum_i c_i\alpha_i + \sum_{i<j} c_{ij}\alpha_i\alpha_j + \cdots + c_{12\cdots h}\alpha_1\alpha_2 \cdots \alpha_h\}, \tag{3.12}$$

and this is precisely the form of

$$\psi(\alpha_1, \alpha_2, \cdots, \alpha_k) \equiv \psi(\alpha_{11}, \alpha_{12}, \cdots, \alpha_{1m}, \cdots, \alpha_{k1}, \alpha_{k2}, \cdots, \alpha_{km}),$$

with $h = km$, in (3.1).

We now derive the moments of the distribution represented by (3.12). Denoting the random variables of this distribution by $x_1, x_2, \cdots, x_h$, it is seen that $E(x_{i_1}x_{i_2} \cdots x_{i_\beta})$ is given by the coefficient of $\alpha_{i_1}\alpha_{i_2} \cdots \alpha_{i_\beta}$ in (3.12). Hence

$$E(x_{i_1}x_{i_2} \cdots x_{i_\beta}) = \sum_{j=1}^{\beta} S_j'(i_1, i_2, \cdots, i_\beta), \qquad \beta = 1, 2, \cdots, h, \tag{3.13}$$

where $S_j'(i_1, i_2, \cdots, i_\beta)$ denotes the sum of all terms, each of which is a product of c's, such that the subscripts in these c's form a complete partitioning of the set $\{i_1, i_2, \cdots, i_\beta\}$ into j distinct non-empty subsets.[4] This result may further be used to express the c's in terms of simple product moments of the x's when the latter are measured from their means. In particular,

$$c_\alpha = Ex_\alpha, \tag{3.14}$$

$$c_{\alpha\beta} = E(x_\alpha'x_\beta'), \tag{3.15}$$

$$c_{\alpha\beta\gamma} = E(x_\alpha'x_\beta'x_\gamma'), \tag{3.16}$$

$$\begin{aligned} c_{\alpha\beta\gamma\delta} = &E(x_\alpha'x_\beta'x_\gamma'x_\delta') - E(x_\alpha'x_\beta')E(x_\gamma'x_\delta') \\ &- E(x_\alpha'x_\gamma')E(x_\beta'x_\delta') - E(x_\alpha'x_\delta')E(x_\beta'x_\gamma'), \end{aligned} \tag{3.17}$$

where $x_i' = x_i - Ex_i$.

[4] As an example, $S_3'(1, 2, 3, 4) = c_{12}c_3c_4 + c_{13}c_2c_4 + c_{14}c_2c_3 + c_{23}c_1c_4 + c_{24}c_1c_3 + c_{34}c_1c_2$.

Reverting now to (3.1), the coefficient of $\alpha_{i_1 j_1}\alpha_{i_2 j_2} \cdots \alpha_{i_r j_r}$ in the latter expression, where $\{i_1, i_2, \cdots, i_r\}$ is a subset of $\{1, 2, \cdots, k\}$ with $i_1 < i_2 < \cdots < i_r$ while $j_1, j_2, \cdots, j_r = 1, 2, \cdots, m$, is $\nu_{j_1} P_{j_1 j_2}(t_{i_2} - t_{i_1}) P_{j_2 j_3}(t_{i_3} - t_{i_2}) \cdots P_{j_{r-1} j_r}(t_{i_r} - t_{i_{r-1}})$. It then follows from (3.14)–(3.17) that

$$En_{i\alpha} = \nu_\alpha, \tag{3.18}$$

$$E\{(n_{i\alpha} - \nu_\alpha)(n_{j\beta} - \nu_\beta)\} = \nu_\alpha P_{\alpha\beta}(t_j - t_i), \qquad i < j, \tag{3.19}$$

$$\begin{aligned} &E\{(n_{i\alpha} - \nu_\alpha)(n_{j\beta} - \nu_\beta)(n_{l\gamma} - \nu_\gamma)\} \\ &\qquad = \nu_\alpha P_{\alpha\beta}(t_j - t_i) P_{\beta\gamma}(t_l - t_j), \quad i < j < l, \end{aligned} \tag{3.20}$$

$$\begin{aligned} &E\{(n_{i\alpha} - \nu_\alpha)(n_{j\beta} - \nu_\beta)(n_{l\gamma} - \nu_\gamma)(n_{r\delta} - \nu_\delta)\} \\ &= \nu_\alpha P_{\alpha\beta}(t_j - t_i) P_{\beta\gamma}(t_l - t_j) P_{\gamma\delta}(t_r - t_l) + \nu_\alpha \nu_\gamma P_{\alpha\beta}(t_j - t_i) P_{\gamma\delta}(t_r - t_l) \\ &\quad + \nu_\alpha \nu_\beta P_{\alpha\gamma}(t_l - t_i) P_{\beta\delta}(t_r - t_j) + \nu_\alpha \nu_\beta P_{\alpha\delta}(t_r - t_i) P_{\beta\gamma}(t_l - t_j), \\ &\qquad\qquad i < j < l < r \end{aligned} \tag{3.21}$$

for $\alpha, \beta, \gamma, \delta = 1, 2, \cdots, m$. ((3.18) and (3.19) are restatements of (2.6) and (2.7), respectively.) Equation (3.13) may be used to determine moments of *all* order without cumbrous differentiation (or expansion in series) of $\phi^*(\alpha_1, \alpha_2, \cdots, \alpha_h)$. For example, $E(x_1 x_2^2)$ or $E(x_1' x_2'^2)$ may be evaluated by taking $h = 3$, $x_3 \equiv x_2$ in (3.12) and evaluating $E(x_1 x_2 x_3)$. This means that such expressions as $E(n_{1\alpha} n_{2\beta}^2)$ or $E\{(n_{1\alpha} - \nu_\alpha)(n_{2\beta} - \nu_\beta)^2\}$ may be evaluated from $\psi(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_k)$ for all $k \geqq 3$ by contraction (i.e., by letting the time interval between the appropriate vector observations approach zero).

Finally, it follows readily from the joint characteristic function of $\mathbf{n}_1, \cdots, \mathbf{n}_k$ implied directly by (3.1) that the limiting distribution, $\nu_j \to \infty$ for $j = 1, 2, \cdots, m$, of the mk standardized random variables $(n_{ij} - \nu_j)/\nu_j^{1/2}$ $(i = 1, 2, \cdots, k;$ $j = 1, 2, \cdots, m)$ is an mk-dimensional Normal distribution with zero expectation vector and variance-covariance matrix $\mathbf{W}$, where $\mathbf{W}$ may be partitioned into k^2 submatrices, each of order $m \times m$, such that the (i, j)th submatrix is $\mathbf{P}((j - i)\tau) \equiv [\mathbf{P}(\tau)]^{j-i}$ or $\mathbf{P}((j - i)\tau) \equiv [\mathbf{P}(\tau)]^{i-j}$, according as to whether $j \geqq i$ or $j < i$.

REFERENCES

[1] A. C. Aitken, *Statistical Mathematics*, 2nd ed., Interscience Pub., New York, 1942.
[2] M. S. Bartlett, "Some evolutionary stochastic processes," *J. Roy. Stat. Soc.*, Ser. B, Vol. 11 (1949), pp. 211–229.
[3] M. S. Bartlett, *An Introduction to Stochastic Processes*, 2nd ed., Cambridge University Press, 1956.
[4] William Feller, *An Introduction to Probability Theory and its Applications*, Vol. 1, 2nd ed., John Wiley and Sons, New York, 1957.
[5] Marc Kac, *Probability and Related Topics in Physical Sciences*, Vol. 1, Interscience Publishers, New York, 1959.
[6] G. Malécot, "Un traitement stochastique des problèmes linéaries (mutations, linkage,

migration) en génétique de population," *Ann. Univ. Lyon*, Vol. 14 (1951), pp. 79–117.

[7] V. N. PATANKAR, "The goodness of fit of frequency distributions obtained from stochastic processes," Ph.D. Thesis, University of Manchester, 1953.

[8] V. T. PATIL, "The consistency and adequacy of the Poisson-Markoff model for density fluctuations," *Biometrika*, Vol. 44 (1951), pp. 43–56.

[9] HAROLD RUBEN AND LORD ROTHSCHILD, "Estimation of mean speeds of organisms and particles by counting," unpublished document, 1953.