The Annals of Probability
1976, Vol. 4, No. 2, 326–334

INFINITE F-DIVISIBILITY OF INTEGER-VALUED RANDOM VARIABLES

BY IAN R. JAMES

CSIRO Division of Mathematics and Statistics, Melbourne

Consider $m \geqq 2$ nonnegative, integer-valued random variables $X_1^{(n)}, \cdots, X_m^{(n)}$ satisfying $\sum_{j=1}^m X_j^{(n)} \leqq n$. If $(X_1^{(n)}, \cdots, X_m^{(n)})$ is one member of a family of random vectors, indexed by different values of the bound n, Darroch has proposed a definition of "independence except for the constraint," termed F-independence, which relates members of the family through their conditional distributions. In this paper we study the limit theory for sums of nonnegative, integer-valued variables, when the sums are bounded and the variables F-independent. The F-independence analogue of infinite divisibility, termed infinite F-divisibility, is defined and characterized, and it is shown that limit distributions of sums of F-independent, asymptotically negligible variables are infinitely F-divisible. Conditions under which the limit is binomial are given. Our results apply to families of random variables, induced by the F-independence definition, and their role in the theory is discussed.

1. Introduction. Let $X_1^{(n)}, \cdots, X_m^{(n)}$ be $m \geqq 2$ nonnegative, integer-valued random variables which satisfy the constraint $X_1^{(n)} + \cdots + X_m^{(n)} \leqq n$. Under the usual definitions of independence, the constraint by itself is sufficient to make $X_1^{(n)}, \cdots, X_m^{(n)}$ dependent, apart from exceptional cases, and to answer the question of whether they are associated in any way other than by competition for space we require a modified definition of "independence" which takes the constraint into account. Darroch (1971) has proposed one such definition, termed F-independence, which is appropriate when $(X_1^{(n)}, \cdots, X_m^{(n)})$ can be considered as one member of a family of related vectors indexed by different values of the bound n. More precisely, let $\Psi_m(I) = \{(X_1^{(n)}, \cdots, X_m^{(n)}); n \in I\}$ be a family of vectors of nonnegative, integer-valued random variables, each vector satisfying $X_1^{(n)} + \cdots + X_m^{(n)} \leqq n$, where $I = \{N_1, N_1 + 1, \cdots, N_2\}$, N_1 integral, $1 \leqq N_1 < N_2$. (If $N_2 = \infty$, I is the set of integers greater than $N_1 - 1$.) Denote by $g_j^{(n)}(i_j \mid i_k;\, k \neq j)$ the conditional probability that $X_j^{(n)} = i_j$ given $X_k^{(n)} = i_k$, $k \neq j$, which we assume is defined and positive for $i_1, \cdots, i_m \geqq 0$, $i_1 + \cdots + i_m \leqq n$, $n \in I$. F-independence then relates members of $\Psi_m(I)$ by the conditions

$$g_j^{(n)}(i_j \mid i_k;\, k \neq j) = g_j^{(n')}(i_j \mid i_k';\, k \neq j)\,, \qquad j = 1, 2, \cdots, m \tag{1.1}$$

whenever $n - \sum_{k\neq j} i_k = n' - \sum_{k \neq j} i_k'$, $n, n' \in I$, i.e. the conditional distribution of $X_j^{(n)}$ depends on the size of the bound n and the values of the other variables only through its conditional range.

Received September 3, 1974; revised August 29, 1975.

AMS 1970 *subject classification.* Primary 60G50.

Key words and phrases. Dependence due to a constraint, F-independence, infinite divisibility, bounded-sum variables, binomial, beta-binomial, limits of sums of random variables.

Since one of the important areas of probability theory is the study of limit distributions of sums of independent variables, it is natural now to look at the corresponding theory when the variables are nonnegative, integer-valued and bounded in sum, and F-independence replaces independence as the criterion of "no-association." In Section 2 some further preliminary results on F-independence are stated, and we discuss the necessity of introducing the family $\Psi_m(I)$. The concept of infinite F-divisibility, which is the F-independence analogue of infinite divisibility, is defined and characterized in Section 3, while in Section 4 it is shown that limits of sums of asymptotically negligible, F-independent, integer-valued variables belong to infinitely F-divisible families. As a by-product of this result we deduce that limit distributions of sums of independent, nonnegative, integer-valued variables satisfying some unusual nonnegligibility conditions, are infinitely divisible. Finally, in Section 5, the binomial distribution is derived from assumptions similar to the well-known Poisson process axioms.

Although the theory of F-independence for positive, continuous, bounded-sum variables has been given in [2], the theory of infinite F-divisibility for continuous variables apparently requires a number of regularity conditions not necessary in the discrete case, and it is not presented here.

2. Preliminary F-independence results. As in the previous section let $I = \{N_1, N_1 + 1, \cdots, N_2\}$ and $\Psi_m(I) = \{(X_1^{(n)}, \cdots, X_m^{(n)}); n \in I\}$, where $X_1^{(n)}, \cdots, X_m^{(n)}$ are nonnegative, integer-valued with $X_1^{(n)} + \cdots + X_m^{(n)} \leqq n$. Denote by $w^{(n)}(i_1, \cdots, i_m)$ the probability that $X_j^{(n)} = i_j$, $j = 1, \cdots, m$, and suppose that this is positive for $i_j \geqq 0$, $i_1 + \cdots + i_m \leqq n$, $n \in I$.

DEFINITION. $X_1^{(n)}, \cdots, X_m^{(n)}$ are F-independent in the family $\Psi_m(I)$ for each $n \in I$ if for some positive functions $a, b_1, \cdots, b_m, c$,

$$(2.1) \qquad w^{(n)}(i_1, \cdots, i_m) = a(n)b_1(i_1) \cdots b_m(i_m)c(n - i_1 - \cdots - i_m)\,,$$

$i_j \geqq 0$, $i_1 + \cdots + i_m \leqq n$, $n \in I$.

The main result in [1] establishes the equivalence of (2.1) and (1.1) when $N_1 = 1$, while the same result for general N_1 can be proved similarly and is established in [6]. In many cases of interest the property (2.1) holds with I the set of positive integers. This is true for example in the multinomial family with parameters $(n, q_1, \cdots, q_m)$, $q_j > 0$, $q_1 + \cdots + q_m < 1$, where $w^{(n)}(i_1, \cdots, i_m) = n! \prod_{j=1}^{m+1} (q_j^{i_j}/i_j!)$, $q_{m+1} = 1 - \sum_{j=1}^m q_j$, $i_{m+1} = n - \sum_{j=1}^m i_j$; and in the Dirichlet compound multinomial family (see [7], page 309 for alternate names) with parameters $(n, \alpha_1, \cdots, \alpha_{m+1})$, $\alpha_j > 0, j = 1, \cdots, m+1$, where $w^{(n)}(i_1, \cdots, i_m) = n!(\prod_{j=1}^{m+1} \alpha_j^{[i_j]}/i_j!)/(\alpha_1 + \cdots + \alpha_{m+1})^{[n]}$, with $i_{m+1} = n - \sum_{j=1}^m i_j$ and $\alpha^{[j]} = \alpha(\alpha + 1) \cdots (\alpha + j - 1)$. However, there exist important families, such as the hypergeometric families, for which (2.1) holds only for a finite N_2 (see [1]). In any case, the added generality of arbitrary N_1 and N_2 causes no extra complications.

If (2.1) holds we shall simply say that $X_1^{(n)}, \cdots, X_m^{(n)}$ are F-independent, on the understanding that $(X_1^{(n)}, \cdots, X_m^{(n)})$ belongs to an F-independent family

$\Psi_m(I)$. Note that the property defined by (2.1) is termed "complete F-independence" in [1] (definitions of F-independence less than complete are obtained for example by imposing (1.1) only for $j = 1, \cdots, k < m$) but we shall not be considering anything less than "complete" F-independence in this paper.

Further discussion regarding interpretation of the model (2.1) and tests of F-independence hypotheses may be found in [1], [3].

Under the concept of F-independence, $w^{(n)}(i_1, \cdots, i_m)$ is a function of $i_1, \cdots, i_m$ and also the bound n. For fixed $n = n_0$, say, if we omit the index n_0, (2.1) gives the functional form

$$w(i_1, \cdots, i_m) = d(i_1 + \cdots + i_m) \prod_{j=1}^{m} b_j(i_j)\,, \qquad \text{say.} \tag{2.2}$$

For $m \geqq 3$, (2.2) is equivalent to the conditions (1.1) considered only for $n = n' = n_0$, i.e. the conditional distribution of each $X_j^{(n_0)}$ depends on the remaining variables only through their sum. Thus property (2.2) may be termed *sum-dependence* (the consequences of which in a more general setting are being investigated separately). For $m = 2$ of course the equivalence does not hold since (1.1) is vacuous for $n = n' = n_0$.

Suppose that $X_1^{(n_0)}, \cdots, X_m^{(n_0)}$ are sum-dependent with probability function (pf) given by (2.2). Then $S^{(n_0)} = X_1^{(n_0)} + \cdots + X_m^{(n_0)}$ has pf $p(i) = d(i)b(i)$, where for $0 \leqq i \leqq n_0$,

$$b(i) = b_1(i) * \cdots * b_m(i) \equiv \sum b_1(i_1) \cdots b_m(i_m)\,, \tag{2.3}$$

summation being over the set $\{i_1, \cdots, i_m; i_j \geqq 0, i_1 + \cdots + i_m = i\}$. Since $d(i)$ is arbitrary, it follows that *any* bounded, nonnegative integer-valued variable with positive pf is the sum of any number of sum-dependent variables. On the other hand, if other members of $\Psi_m(I)$ are also considered, and $X_1^{(n)}, \cdots, X_m^{(n)}$ are F-independent, then $S^{(n)} = X_1^{(n)} + \cdots + X_m^{(n)}$ has pf of the form

$$p^{(n)}(i) = a(n)b(i)c(n-i)\,, \qquad 0 \leqq i \leqq n,\ n \in I\,, \tag{2.4}$$

where $b(i)$ again is the convolution (2.3) for $0 \leqq i \leqq N_2$. Members of the family $\{S^{(n)}; n \in I\}$ are related by (2.4) in that specifying $b(i)$ and $c(i)$ for $i = 0, 1, \cdots, n_0$ completely determines $p^{(n)}(i)$ for $n \leqq n_0$ and partly determines $p^{(n)}(i)$ for $n > n_0$. Thus not every *family* of nonnegative, bounded variables has members which are sums of F-independent variables from some family $\Psi_m(I)$.

Note that if $p^{(n)}(i)$ has the functional form (2.4), its factorization into functions of n, i and $n - i$ is not unique. According to our next result, however, if b has the convolution property (2.3), then no matter what factorization is used, the function of i must always be a convolution.

LEMMA 2.1. *Suppose $a(n)b(i)c(n-i) = a'(n)b'(i)c'(n-i)$ for $0 \leqq i \leqq n$, $n \in I$, where a, b, c, a', b', c' are positive functions. Then there exist positive constants α_1, α_2, λ such that $a(n) = \alpha_1\alpha_2\lambda^n a'(n)$, $b'(i) = \alpha_1\lambda^i b(i)$, $c'(i) = \alpha_2\lambda^i c(i)$. The converse is also true.*

PROOF. A proof for the case $I = \{1, 2, \cdots, N_2\}$ is in [1], and this carries over to general I. □

It follows from the lemma that if $b(i) = b_1(i) * \cdots * b_m(i)$ for $0 \leqq i \leqq N_2$, then $b'(i) = [\alpha_1^{1/m}\lambda^i b_1(i)] * \cdots * [\alpha_1^{1/m}\lambda^i b_m(i)]$.

We conclude this preliminary section with a result which will be required in the later sections.

LEMMA 2.2. *Let* $p_k^{(n)}(i) = a_k(n)b_k(i)c_k(n-i) > 0$ *for* $0 \leqq i \leqq n$, $n \in I$, $k = 1, 2, \cdots$. *Then if* $p_k^{(n)}(i) \to p^{(n)}(i) > 0$ *as* $k \to \infty$, *there exist sequences* $\{\theta_k\}$, $\{\mu_k\}$, $\{\lambda_k\}$ *of positive numbers such that* $[\theta_k \mu_k \lambda_k^n]^{-1}\ a_k(n) \to a(n)$, $\theta_k \lambda_k^i b_k(i) \to b(i)$, $\mu_k \lambda_k^i c_k(i) \to c(i)$, $0 \leqq i \leqq n$, $n \in I$, *for some* $a, b, c > 0$ (*so that* $p^{(n)}(i) = a(n)b(i)c(n-i)$).

PROOF. By considering the ratio $p_k^{(n+1)}(i)p_k^{(n)}(i-1)/(p_k^{(n+1)}(i-1) \cdot p_k^{(n)}(i))$ for $1 \leqq i \leqq n < N_2$, it follows that $c_k(i)\lambda_k/c_k(i-1)$ converges to a positive limit with $\lambda_k = c_k(0)/c_k(1)$. Thus $c_k(i)\lambda_k^i/c_k(0)$ also converges to a positive limit, and the same result for $b_k(i)\lambda_k^i/b_k(0)$ follows by considering the ratio $p_k^{(n)}(i)/p_k^{(n)}(i-1)$. □

In view of Lemma 2.2 we may assume without loss of generality that $b_k(i)$ and $c_k(i)$ themselves converge whenever $p_k^{(n)}(i)$ does.

3. Infinite F-divisibility of integer-valued variables. Let $\Psi(I) = \{X^{(n)}; n \in I\}$ be a family of nonnegative, integer-valued variables with $0 \leqq X^{(n)} \leqq n$. Let $p^{(n)}(i)$ be the pf of $X^{(n)}$, assumed to be positive for $0 \leqq i \leqq n$, $n \in I$, and put $P(I) = \{p^{(n)}; n \in I\}$.

DEFINITION. $\Psi(I)$ (or $P(I)$) is infinitely F-divisible (inf F-div) if for each integer $m \geqq 2$ there exist families $\Psi_m(I)$ of vectors of exchangeably distributed, integer-valued, F-independent variables $(X_{m1}^{(n)}, \cdots, X_{mm}^{(n)})$ such that $X^{(n)} = X_{m1}^{(n)} + \cdots + X_{mm}^{(n)}$, $n \in I$.

We shall say that $X^{(n)}$ is inf F-div if the family to which it belongs is. One can easily verify that an equivalent definition is obtained if "exchangeably distributed" is replaced by "identically distributed."

THEOREM 3.1. $\Psi(I)$ *is* inf F-div *if and only if there exist positive functions* a, b, c *and* β_m, $m \geqq 2$, *such that* $p^{(n)}(i) = a(n)b(i)c(n-i)$, $0 \leqq i \leqq n$, $n \in I$, *where for each* m,

$$b(i) = \beta_m(i) * \cdots * \beta_m(i) \equiv_d \beta_m{}^{m*}(i)\,, \qquad i = 0, 1, \cdots, N_2\,, \tag{3.1}$$

the m-fold convolution as defined in Section 2.

PROOF. The proof of sufficiency is trivial, while necessity follows immediately from Lemma 2.1 and the observation that if $X_{m1}^{(n)}, \cdots, X_{mm}^{(n)}$ are exchangeably distributed and F-independent, then they have a pf of the form $w_m^{(n)}(i_1, \cdots, i_m) = a_m(n)\gamma_m(i_1) \cdots \gamma_m(i_m)c_m(n - i_1 - \cdots - i_m)$. □

Note again that however $p^{(n)}(i)$ is factorized into functions of n, i, $n-i$, the function of i must be a convolution of the type (3.1).

Infinite F-divisibility does not impose restrictions on the function c, and therefore does not on $p^{(n)}(i)$ when considered only for a *single* value of n, say $n = n_0$.

However, if both $\Psi(I)$ and $\{n - X^{(n)}; n \in I\}$ are inf F-div, then $b(i) = \beta_m{}^{m*}(i)$ and $c(i) = \rho_m{}^{m*}(i)$, $i = 0, 1, \cdots, N_2$, for some $\beta_m, \rho_m, m \geqq 2$. This follows since $n - X^{(n)}$ has pf of the form $s^{(n)}(i) = a(n)c(i)b(n-i)$. Thus each $p^{(n)}$ is, apart from a normalizing constant, a product of two functions, one of i and one of $n - i$, both satisfying convolution properties of the type (3.1). Two important examples of such families are the binomial and binomial-beta families (see [7], page 189 and page 231 for alternate names) with $I = \{1, 2, \cdots\}$. If $X^{(n)}$ is binomial with parameters (n, q), $0 < q < 1$, then for any $m \geqq 2$ it can be expressed as the sum of m variables having a multinomial distribution with parameters $(n, q/m, \cdots, q/m)$ (Section 2), while if $X^{(n)}$ is beta-binomial with parameters (n, α, β), the variables in the sum are Dirichlet compound multinomial with parameters $(n, \alpha/m, \cdots, \alpha/m, \beta)$ (Section 2). The variables $n - X^{(n)}$ are also respectively binomial and beta-binomial.

The property $b(i) = \beta_m{}^{m*}(i)$ for $m \geqq 2$ is of course similar to the infinite divisibility property for pf's on the nonnegative integers. It differs in that if I is finite we require it to hold only on a finite subset of the nonnegative integers, while even if I is infinite, $b(i)$ (or $\lambda^i b(i)$ for any $\lambda > 0$) need not be summable. Conversely, any positive infinitely divisible pf possesses the convolution property (3.1) and thus generates a class of inf F-div families. For example, the Poisson and negative binomial distributions generate the families with pf's $p^{(n)}(i) = a(n)c(n-i)/i!$ and $p^{(n)}(i) = a(n)c(n-i)\alpha^{[i]}/i!$ respectively. The binomial and beta-binomial families are special cases of these.

A more convenient characterization of inf F-div is obtained by modifying the proof of a result of Katti (1967), which characterizes infinitely divisible pf's on the nonnegative integers.

THEOREM 3.2. *$b(i) = \beta_m{}^{m*}(i)$ for $m \geqq 2$, $i = 0, 1, \cdots, N_2$ if and only if there exist nonnegative numbers $\tau_1, \cdots, \tau_{N_2}$ with $\tau_1 > 0$, such that*

$$ib(i) = \sum_{k=0}^{i-1} b(k)\tau_{i-k}, \qquad i = 1, 2, \cdots, N_2. \tag{3.2}$$

PROOF. Suppose first $N_2 < \infty$. To prove the necessity half, put $\beta_m(i) = 0$ for $i > N_2$ and $b_m(i) = \beta_m{}^{m*}(i)$, $i = 0, 1, \cdots$, so that $b_m(i) = b(i)$ for $i \leqq N_2$. We then mimic the proof in [8]. Namely, if $f_m(z)$, $h_m(z)$ are the generating functions of $b_m(i)$, $\beta_m(i)$ respectively, then $f_m(z) = h_m{}^m(z)$, which gives on differentiating and rearranging, $f_m'(z)h_m(z) = f_m(z)h_m'(z)$. Equating coefficients of z^k for $k \leqq N_2 - 1$, and letting $m \to \infty$, gives (3.2) with $\tau_k = k \lim_{m\to\infty} m\beta_m(k)$. $\tau_1 = b(1)/b(0) > 0$. Conversely, put $h_m(z) = b^{1/m}(0) \exp[(1/m) \sum_{j=1}^{N_2} \tau_j z^j/j]$. Then the coefficients of the powers of z in the power series expansion of $h_m(z)$ are all positive, and $b(k)$ is the coefficient of z^k in $h_m{}^m(z)$, $k \leqq N_2$. This completes the proof for N_2 finite. If N_2 is infinite, the above proof can be applied for $i = 0, 1, \cdots, N < \infty$, and since τ_k and $\beta_m(k)$ are invariant with respect to N for $N \geqq k$, the result follows. If $b(i)$ is summable, the proof of Katti (1967) may be used directly. □

Using (3.2) one obtains immediately that the hypergeometric families are not inf F-div for any N_2, since τ_2 is negative.

We may note here two other properties which characterize infinitely divisible, nonnegative, integer-valued variables [4]:

(A) X has a compound Poisson distribution (and is integer-valued);

(B) X is distributed as $\sum_{i=1}^{\infty} iZ_i$, where the Z_i's are independent Poisson variables with parameters λ_i, $\sum_{i=1}^{\infty} \lambda_i < \infty$.

A characterization of inf F-div corresponding to (A) follows by observing from the proof of Theorem 3.2 that (3.2) is equivalent to $b(i) = \sum_{k=0}^{\infty} \theta^{k*}(i)/k!$, where $\theta(0) = b(0)$, $\theta(i) = \tau_i/i$, $i = 1, 2, \cdots, N_2$. That is, if $Y_{m1}^{(n)}, \cdots, Y_{mm}^{(n)}$ have a pf of the form

$$w_m^{(n)}(i_1, \cdots, i_m) = a(n)[\textstyle\prod_{j=1}^{m} \phi(i_j)]c(n - i_1 - \cdots - i_m)\,,$$

($Y_{mi}^{(n)}$ and $Y_{ri}^{(n)}$ are not identically distributed for $r \neq m$), then an inf F-div distribution corresponds to the distribution of $X_m^{(n)} = Y_{m1}^{(n)} + \cdots + Y_{mm}^{(n)}$ compounded with a Poisson distribution for m. If both $X^{(n)}$ and $n - X^{(n)}$ are inf F-div, then $X^{(n)}$ has a pf of the form

$$p^{(n)}(i) = a(n) \textstyle\sum_{m,k=0}^{\infty} \theta^{m*}(i)\mu^{k*}(n - i)/m!\, k!\,.$$

This corresponds to the pf of $X_{mk}^{(n)} = Y_{mk1}^{(n)} + \cdots + Y_{mkm}^{(n)}$, compounded with independent Poisson distributions for m and k, where $Y_{mk1}^{(n)}, \cdots, Y_{mkm}^{(n)}$ have pf of the form

$$w_{mk}^{(n)}(i_1, \cdots, i_m) = a(n)[\textstyle\prod_{j=1}^{m} \phi(i_j)]\eta^{k*}(n - i)\,.$$

The F-independence analogue of (B) does not appear to have a clear interpretation.

The final results for this section have well-known counterparts in infinite divisibility theory, and can be easily proved using the inf F-div characterizations above.

THEOREM 3.3. *Let* $\Psi_j(I) = \{X_j^{(n)}; n \in I\}$, $j = 1, 2, \cdots$, *be a sequence of* inf F-div *families.*

(a) *If* $X_1^{(n)}, \cdots, X_k^{(n)}$ *are F-independent*, $n \in I$, *then* $\{S_k^{(n)}; S_k^{(n)} = X_1^{(n)} + \cdots + X_k^{(n)}, n \in I\}$ *is* inf F-div.

(b) *If* $X_j^{(n)} \to X^{(n)}$ *as* $j \to \infty$, *where* $X^{(n)}$ *has* pf $p^{(n)}(i) > 0$ *for* $0 \leqq i \leqq n$, $n \in I$, *then* $\{X^{(n)}; n \in I\}$ *is* inf F-div.

4. Sums of F-independent, integer-valued variables. Suppose we have a family of double sequences $X_{mj}^{(n)}$ of nonnegative, integer-valued variables, indexed by $n \in I$, with $j = 1, 2, \cdots, v(m)$, $v(m) \geqq 2$, $m = 1, 2, \cdots$, where $X_{m1}^{(n)} + \cdots + X_{mv(m)}^{(n)} \leqq n$ for each m.

DEFINITION. If the variables in each row, $X_{m1}^{(n)}, \cdots, X_{mv(m)}^{(n)}$, are F-independent for $n \in I$, and $v(m) \to \infty$ as $m \to \infty$, the family will be called a *triangular F-array*.

If in addition $\max_{1 \leqq j \leqq v(m)} \Pr[X_{mj}^{(n)} = k] \to 0$ as $m \to \infty$ for $k \geqq 1$, the triangular F-array will be called *null* (cf. [5], page 174).

Let $w_m^{(n)}(i_1, \cdots, i_m) = a_m(n)[\prod_{j=1}^{v(m)} \beta_{mj}(i_j)]c_m(n - i_1 - \cdots - i_m)$ be the positive pf of $X_{m1}^{(n)}, \cdots, X_{mv(m)}^{(n)}$, so that the row sum $S_m^{(n)} = X_{m1}^{(n)} + \cdots + X_{mv(m)}^{(n)}$ has pf $p_m^{(n)}(i) = a_m(n)b_m(i)c_m(n - i)$, with $b_m(i) = \beta_{m1}(i) * \cdots * \beta_{mv(m)}(i)$. According to Lemma 2.2 we may assume that if $p_m^{(n)}(i) \to p^{(n)}(i) = a(n)b(i)c(n - i) > 0$ as $m \to \infty$, $0 \leqq i \leqq n$, $n \in I$, then $a_m(n) \to a(n)$, $b_m(i) \to b(i)$ and $c_m(i) \to c(i)$.

LEMMA 4.1. *For a null triangular F-array, if $p_m^{(n)}(i) \to p^{(n)}(i) > 0$ as $m \to \infty$, $0 \leqq i \leqq n$, $n \in I$, then $\max_{1 \leqq j \leqq v(m)} \beta_{mj}(k)/\beta_{mj}(0) \to 0$ as $m \to \infty$ for $k \geqq 1$.*

PROOF. Since $\max_j w_m^{(n)}(0, \cdots, 0, k, 0, \cdots, 0) \to 0$ as $m \to \infty$, where the k corresponds to variable j, $k \geqq 1$, we have

$$\max_j a_m(n)b_m(0)c_m(n - k)\beta_{mj}(k)/\beta_{mj}(0) \to 0 ,$$

and a_m, b_m, c_m converge to positive functions. □

THEOREM 4.1. *For a null triangular F-array, suppose that each row sum $S_m^{(n)} \to X^{(n)}$ as $m \to \infty$, where $X^{(n)}$ has* pf *$p^{(n)}(i) > 0$ for $0 \leqq i \leqq n$, $n \in I$. Then $\{X^{(n)}, n \in I\}$ is* inf *F*-div.

PROOF. Suppose firstly that $N_2 < \infty$, and put $\beta_{mj}(i) = 0$ for $i > N_2$. Put $d_m(i) = \beta_{m1}(i) * \cdots * \beta_{mv(m)}(i)$, $i = 0, 1, 2, \cdots$, so that $d_m(i) = b_m(i)$ for $i \leqq N_2$. Then if $f_m(z)$ and $h_{mj}(z)$ are respectively the generating functions of $d_m(i)$ and $\beta_{mj}(i)$, $f_m(z) = \prod_{j=1}^{v(m)} h_{mj}(z)$, which gives on differentiating and rearranging,

$$f_m'(z) = f_m(z) \sum_{j=1}^{v(m)} h'_{mj}(z)/h_{mj}(z) .$$

Let $\tau_{km}(j)$ denote the coefficient of z^{k-1} in the power series expansion of $h'_{mj}(z)/h_{mj}(z)$, so that on equating coefficients of powers of z we obtain

$$kb_m(k) = \sum_{i=0}^{k-1} b_m(i)\tau_{k-im} , \qquad \tau_{km} = \sum_{j=1}^{v(m)} \tau_{km}(j) .$$

By Theorem 3.2 we need only show that $\tau_{km} \to \tau_k \geqq 0$ as $m \to \infty$. Now it follows that

$$\tau_{km}(j) = k\beta_{mj}(k)/\beta_{mj}(0) - \sum_{i=1}^{k-1} \beta_{mj}(i)\tau_{k-im}(j)/\beta_{mj}(0) ,$$

$k \geqq 2$, and $\tau_{1m}(j) = \beta_{mj}(1)/\beta_{mj}(0)$. From Lemma 4.1, $\max_j \tau_{km}(j) \to 0$ as $m \to \infty$, and since $\sum_{j=1}^{v(m)} \beta_{mj}(i)/\beta_{mj}(0)$ is bounded above by $b(1)/b(0)$ in the limit as $m \to \infty$, we have

$$\tau_k = k \lim_{m \to \infty} \sum_{j=1}^{v(m)} \beta_{mj}(k)/\beta_{mj}(0) \geqq 0 , \qquad \tau_1 = b(1)/b(0) > 0 . \tag{4.1}$$

If N_2 is infinite the proof follows as in Theorem 3.2. □

COROLLARY. *The form of $b(i)$ is determined by* (3.2) *with the τ_k given by* (4.1).

In the next section the corollary is used to derive the binomial distribution. Note that conversely to Theorem 4.1 it is trivially true that an inf F-div family is the limit of row sums in a null triangular F-array.

REMARK. If $b(i)$ and $\beta_{mj}(i)$ are pf's on the nonnegative integers, the proof of Theorem 4.1 implicitly shows the following: let $\{Y_{mj}; j = 1, \cdots, v(m), m = 1, 2, \cdots\}$ be a triangular array of nonnegative, integer-valued variables with $Y_{m1}, \cdots, Y_{mv(m)}$ mutually independent in each row, $v(m) \to \infty$ as $m \to \infty$, and $\max_{1 \leqq j \leqq v(m)} \Pr[Y_{mj} = k] \to 0$ as $m \to \infty$, $k \geqq 1$. Then if $Y_{m1} + \cdots + Y_{mv(m)} \to Y$ as $m \to \infty$, with $\Pr[Y = 0] > 0$, Y is infinitely divisible, and has pf given by (3.2) with $\tau_k = k \lim_{m\to\infty} \sum_{j=1}^{v(m)} \Pr[Y_{mj} = k]/\Pr[Y_{mj} = 0]$. Note that our "negligibility" conditions do not require that $\Pr[Y_{mj} = 0] \to 1$ as $m \to \infty$ (cf. [5], page 174) but if this is not the case, then Y_{mj} cannot converge properly.

5. The binomial distribution. Let $\{X_{mj}^{(n)}, j = 1, \cdots, v(m), m = 1, 2, \cdots; n \in I\}$ and $\{Z_{mj}^{(n)}, j = 1, \cdots, t(m), m = 1, 2, \cdots; n \in I\}$ be two null triangular F-arrays for which $X^{(n)} = \lim_{m\to\infty} X_{m1}^{(n)} + \cdots + X_{mv(m)}^{(n)}$ and $n - X^{(n)} = \lim_{m\to\infty} Z_{m1}^{(n)} + \cdots + Z_{mt(m)}^{(n)}$. Then

THEOREM 5.1. *Provided* $\max_{1 \leqq j \leqq v(m)} v(m) \Pr[X_{mj}^{(n)} = k] \to 0$ *and* $\max_{1 \leqq j \leqq t(m)} t(m) \times \Pr[Z_{mj}^{(n)} = k] \to 0$ *as* $m \to \infty$ *for* $k \geqq 2$, $X^{(n)}$ *(and* $n - X^{(n)}$*) has a binomial distribution.*

PROOF. If we adopt the notation of Theorem 4.1, then it follows that $\max_j v(m)\beta_{mj}(k)/\beta_{mj}(0) \to 0$ as $m \to \infty$, $k \geqq 2$, so that $\tau_k = 0$, $k \geqq 2$. Then $ib(i) = \tau_1 b(i-1)$, or $b(i) = b(0)\tau_1^{\,i}/i!$. A similar argument for $n - X^{(n)}$ gives $c(i) = c(0)\sigma_1^{\,i}/i!$ say, and $X^{(n)}$ has a binomial distribution with parameters $(n, \tau_1/(\tau_1 + \sigma_1))$. □

Theorem 5.1 can be viewed in a point process context as follows: consider arrivals of two types, e.g. males and females, at some location, where at time T a total of n arrivals have occurred, $X^{(n)}$ males and $n - X^{(n)}$ females. Divide $(0, T]$ into m subintervals and let $X_{mj}^{(n)}$, $Z_{mj}^{(n)}$ denote the numbers of male and female arrivals in the jth subinterval. Then if $X_{m1}^{(n)}, \cdots, X_{mm}^{(n)}$ are F-independent, and separately $Z_{m1}^{(n)}, \cdots, Z_{mm}^{(n)}$ are F-independent, for each m, and, further, the arrivals are sparse in the sense of Theorem 5.1, it follows that the male (and female) arrivals have a binomial distribution. In this context the conditions of Theorem 5.1 resemble the well-known Poisson process axioms. (Note that if X, Y are *independent* Poisson variables, then conditional on $X + Y = n$, X is binomial. Our conditions do not explicitly assume any particular interaction between male and female arrivals, although if n is a realization of a random bound, and $p^{(n)}$ is the conditional pf of X given $X + Y = n$, say, it follows implicitly that X, Y have a pf of the form $p(i, j) = \phi_1(i + j)\phi_2(i)\phi_3(j)$.)

General counting processes with F-independent increments can be considered analogously to counting processes with independent increments. For brevity the topic is not pursued here.

Acknowledgments. I wish to thank Professor J. N. Darroch for some useful discussions relating to this work. The constructive advice of an associate editor and referee on a previous draft is also acknowledged.

REFERENCES

[1] Darroch, J. N. (1971). A definition of independence for bounded-sum, nonnegative, integer-valued variables. *Biometrika* **58** 357–368.

[2] Darroch, J. N. and James, I. R. (1974). *F*-independence and null correlation of continuous, bounded-sum, positive variables. *J. Roy. Statist. Soc. Ser. B* **36** 467–483.

[3] Darroch, J. N. and Ratcliff, D. (1973). Tests of *F*-independence with reference to quasi-independence and Waite's fingerprint data. *Biometrika* **60** 395–401.

[4] Feller, William (1968). *An Introduction to Probability Theory and its Applications* **1**, 3rd ed. Wiley, New York.

[5] Feller, William (1966). *An Introduction to Probability Theory and its Applications* **2**. Wiley, New York.

[6] James, I. R. (1973). Concepts of independence for bounded-sum variables. Ph. D. thesis, The Flinders University of South Australia.

[7] Johnson, N. L. and Kotz, S. (1969). *Distributions in Statistics—Discrete Distributions.* Houghton-Mifflin, Boston.

[8] Katti, S. K. (1967). Infinite divisibility of integer-valued random variables. *Ann. Math. Statist.* **38** 1306–1308.

CSIRO, Division of Mathematics and Statistics
P.O. Box 310
69 Yarra Bank Road
South Melbourne, Victoria 3205
Australia