PROBABILISTIC BOUNDS ON THE COEFFICIENTS OF POLYNOMIALS WITH ONLY REAL ZEROS *

Jim Pitman

Technical Report No. 453

Department of Statistics
University of California
367 Evans Hall # 3860
Berkeley, CA 94720-3860

pitman@stat.berkeley.edu

Revised November 12, 1996

Abstract

The work of Harper and subsequent authors has shown that finite sequences $(a_0, \cdots, a_n)$ arising from combinatorial problems are often such that the polynomial $A(z) := \sum_{k=0}^{n} a_k z^k$ has only real zeros. Basic examples include rows from the arrays of binomial coefficients, Stirling numbers of the first and second kinds, and Eulerian numbers. Assuming the a_k are non-negative, $A(1) > 0$ and that $A(z)$ is not constant, it is known that $A(z)$ has only real zeros iff the normalized sequence $(a_0/A(1), \cdots, a_n/A(1))$ is the probability distribution of the

*Research supported in part by N.S.F. Grant MCS9404345

number of successes in n independent trials for some sequence of success probabilities. Such sequences $(a_0, \cdots, a_n)$ are also known to be characterized by total positivity of the infinite matrix (a_{i-j}) indexed by non-negative integers i and j. This papers reviews inequalities and approximations for such sequences, called *Pólya frequency sequences* which follow from their probabilistic representation. In combinatorial examples these inequalities yield a number of improvements of known estimates.

1 Introduction

The work of Harper [58] and subsequent authors [60, 94, 62, 128, 119, 15, 16] has shown that finite sequences $(a_0, \cdots, a_n)$ arising from combinatorial problems are often such that the polynomial $A(z) := \sum_{k=0}^{n} a_k z^k$ has only real zeros. Typically $a_k = a_{nk}$ is the number of elements ω of some finite set Ω_n such that $S_n(\omega) = k$, for some function $S_n : \Omega_n \to \{0, 1, \cdots, n\}$. The *normalized sequence* $(a_0/A(1), \cdots, a_n/A(1))$ then describes the *probability distribution* of $S_n(\omega)$ for ω picked uniformly at random from Ω_n. See Section 4 for examples and further references, and [39] for definition of probabilistic terms.

A sequence of real numbers $(a_k)_{k\in K}$ indexed by a subset K of the non-negative integers is called a *Pólya frequency sequence of order* r, abbreviated PF_r if the infinite matrix $M := (a_{i-j})_{i,j=0,1,2,\cdots}$, where $a_k = 0$ for $k \notin K$, is *totally positive of order* r. That is to say, for each $1 \leq s \leq r$, each $s \times s$ minor of M has a non-negative determinant. The sequence (a_k) is called a *Pólya frequency (PF) sequence* if it is PF_r for every $r = 1, 2, \cdots$. See Karlin [74] and Ando [4] for background on total positivity, and Brenti [16, 17] for recent combinatorial developments of this concept.

Proposition 1 [81, 108] *Let $(a_0, \cdots, a_n)$ be a sequence of non-negative real numbers with associated polynomial $A(z) := \sum_{k=0}^{n} a_k z^k$ such that $A(1) > 0$. The following conditions are equivalent:*

(i) *the polynomial $A(z)$ is either constant or has only real zeros;*

(ii) *$(a_0, \cdots, a_n)$ is a PF sequence;*

(iii) *the normalized sequence $(a_0/A(1), \cdots, a_n/A(1))$ is the distribution of the number S_n of successes in n independent trials with probability p_i of success on the ith trial, for some sequence of probabilities $0 \leq p_i \leq 1$. The roots of $A(z)$ are then given by $-(1-p_i)/p_i$ for i with $p_i > 0$.*

The equivalence (i) $\Leftrightarrow$ (ii) is due to Aissen, Schoenberg and Whitney [1, 108]. This equivalence is a special case of the more general representation of totally positive infinite sequences due to Edrei [33], which is recalled in Section 5. See also Chapter 8 of Karlin [74], Theorem 1.2 and Corollary 3.1. The equivalence (i) $\Leftrightarrow$ (iii), due to P. Lévy [80, 81], follows easily from the interpretation of $A(z)/A(1)$ as a probability generating function. Call an array of real numbers

$$(a_{nk}) = (a_{nk}, 0 \leq k \leq n, n = 1, 2, \cdots)$$

a *PF array* iff $(a_{nk}, 0 \leq k \leq n)$ is a PF sequence for every $n = 1, 2, 3, \cdots$. Basic examples of PF arrays are provided by the binomial coefficients, the Stirling numbers of the first and second kinds, and the Eulerian numbers. Harper [58] and others have exploited the implication (i) $\Rightarrow$ (iii) to deduce normal approximations for the nth row of a PF array from the normal approximation to the distribution of S_n as in (iii). Normal approximations have also been obtained for sequences of combinatorially defined distributions satisfying other conditions [20, 40, 41, 46]. But results in the probability and statistics literature, reviewed in Section 2, show that PF sequences satisfy a number of useful inequalities which do not hold for just any sequence that is approximately normal. As shown in Section 4, even for the two Stirling arrays which have been extensively studied, the probabilistic bounds yield some improvements of known estimates.

The notion of a PF_r sequence was developed early in this century by Fekete, Pólya, Schoenberg and others. See [74] for a survey of this development. Polynomials with real coefficients and only real zeros were the subject of intensive study in the 19th and early 20th century, by Lagrange, Laguerre, Pólya, and many other authors. Much information about such polynomials can be found in [92, 101]. See also [68] and papers cited there. As observed by Schoenberg [108], a sequence of non-negative reals (a_k) is PF_2 iff it is *log-concave* ($a_k^2 \geq a_{k-1}a_{k+1}$) and has *no internal zeros* ($i < j < k$ and $a_i a_k > 0 \Rightarrow a_j > 0$). Sequences with these properties, and the weaker property of *unimodality*, have been extensively studied in probability and statistics[76, 75, 130, 19, 122, 109, 56], and in combinatorics and other fields [116]. The PF_r property for $r \geq 3$ is harder to describe intuitively. But see Brenti [16] for recent combinatorial interpretations of total positivity.

According to *Newton's inequality* ([57], p. 52), if a polynomial $\sum a_k z^k$ with real coefficients has only real roots, and in particular if $(a_0, \cdots, a_n)$ is a

PF sequence, then

$$a_k^2 \geq a_{k-1}a_{k+1}\left(1+\frac{1}{k}\right)\left(1+\frac{1}{n-k}\right) \tag{1}$$

which is stronger than the log-concavity implied by the PF_2 condition. Lévy [80, 81] noted that (1) is a constraint on the probabilities $a_k := P(S_n = k)$ for S_n the number of successes in n independent trials. Lévy also observed that (1) cannot be improved: given non-negative a_{k-1}, a_k, a_{k+1} satisfying (1) for some $1 \leq k \leq n$, there exists a PF sequence $(a_0, \cdots, a_n)$ with these terms at places $k-1, k, k+1$. As shown by Samuels [107], further applications of Newton's inequality imply that for each $r = 1, 2, \cdots$ the sequence of rth order differences derived from a finite PF sequence has at most r strict sign changes.

2 Review of Probabilistic Results.

Let $(a_0, \cdots, a_n)$ be a *frequency sequence*, that is a sequence of non-negative real numbers. Let $A'(z)$ and $A''(z)$ denote the first two derivatives of the polynomial $A(z) = \sum_i a_i z^i$. Abbreviate $A = A(1), A' = A'(1), A'' = A''(1)$, and assume throughout that $A(1) > 0$. Let P denote the *probability distribution* on $\{0, 1, \cdots, n\}$ defined by *normalization* of $(a_0, \cdots, a_n)$. So for example $P(k) := a_k/A$ and for an interval $[b, c]$

$$P[b, c] := \sum_{b \leq j \leq c} P(j) = \frac{1}{A}\sum_{b \leq j \leq c} a_j \tag{2}$$

Let μ and σ denote the *mean* and *standard deviation* of P. That is

$$\mu := \frac{1}{A}\sum_{k=0}^{n} k a_k = \frac{A'}{A} \tag{3}$$

$$\sigma^2 := \frac{1}{A}\sum_{k=0}^{n}(k-\mu)^2 a_k = \frac{A''}{A} + \frac{A'}{A} - \left(\frac{A'}{A}\right)^2 \tag{4}$$

In probabilistic language, if S is a random variable with distribution P, then S has *expectation* μ and *variance* σ^2. If $(a_0, \cdots, a_n)$ is a PF sequence then call P a *PF distribution.* Say a random variable X has *Bernoulli(p) distribution* if X assumes the values 0 and 1 with probabilities $P(X = 1) = p$ and $P(X = 0) = 1 - p$. According to Proposition 1, a probability distribution P on $\{0, 1, \cdots, n\}$ is a PF distribution iff P is the distribution of a sum of independent variables $S_n := X_1 + \cdots + X_n$, where X_i has Bernoulli(p_i) distribution. So for a PF distribution P there are the following standard probabilistic expressions [39], which can also be checked algebraically using the fact that the $-(1 - p_i)/p_i$ are the roots of $A(z)$:

$$\mu = \sum_i p_i \text{ and } \sigma^2 = \sum_{i=1}^n p_i(1 - p_i) \tag{5}$$

History and Terminology. What is called here a *PF distribution* is called in the statistics literature the *distribution of the number of successes in independent trials*. Such trials with two possible outcomes, success and failure, and varying probabilities of success, are known as *Poisson trials* or *Poisson-binomial trials*. The distribution of the number of successes S_n is sometimes called a *Poisson-binomial distribution*, but that term has also acquired other meanings. Study of the distribution of S_n dates back to the 1837 monograph of Poisson [99]. Chebyshev [24] established bounds for tail probabilities and the law of large numbers for the distribution S_n. The work of subsequent authors, reviewed below, has provided sharper bounds for tail probabilities, precise estimates for location of the mode and median, and error bounds for normal and Poisson approximations.

The binomial(n, p) array. The array of binomial coefficients is a PF array due to the factorization of the associated polynomials

$$\sum_{k=0}^n \binom{n}{k} z^k = (1 + z)^n \tag{6}$$

Replace z by pz/q in (6) and normalize to obtain the polynomial associated with the *Binomial(n, p) distribution*. The corresponding PF array with parameter $0 \le p \le 1$, which describes the distribution of the number of successes in n independent trials with constant success probability p, may be presented as follows:

$$P_{n,p}(k) := \binom{n}{k} p^k (1 - p)^{n-k} \qquad (0 \le k \le n) \tag{7}$$

Hoeffding's Inequalities [64]. Let P be a PF distribution on $\{0, 1, \cdots, n\}$ and let $P_{n,p}$ as in (6) denote the binomial(n, p) distribution with the same mean as P, that is with $p = \mu/n$. Then for all integers b and d with $0 \le b \le \mu - 1$ and $\mu + 1 \le d \le n$,

$$P[0, b] \le P_{n,p}[0, b] \text{ and } P[d, n] \le P_{n,p}[d, n] \tag{8}$$

Also, for every convex function g, there is the inequality

$$\sum_{j=0}^{n} g(j)P(j) \le \sum_{j=0}^{n} g(j)P_{n,p}(j) \tag{9}$$

These inequalities make very precise sense of the following idea: amongst all PF distributions on $\{0, 1, \cdots, n\}$ with a given mean μ, the binomial(n, p) distribution for $p = \mu/n$ is the one that is "most spread out". See Gleser [47] for refinements, and Marshall-Olkin [84] for a survey of related inequalities. Hoeffding showed also that for an arbitrary real-valued function g any PF distribution P that maximizes the sum on the left side of (9) over all PF distributions on $\{0, 1, \cdots, n\}$ is necessarily a *shifted binomial distribution.* That is to say $P(k) = P_{m,p}(k - h)$ for all $h \le k \le h + m$ for some integers h and m with $0 \le h \le h + m \le n$ and some p with $0 \le p \le 1$. For $g(j) = 1(j \ge k)$ this result dates back to Chebychev [24], who combined it with bounds on binomial probabilities to obtain a weak law of large numbers.

Large Deviation Bounds. Good bounds for binomial tail probabilities were obtained by Okamoto [93] using the method of Chernoff [25]. Combined with Hoeffding's inequality (8), these bounds show that every PF distribution P on $\{0, 1, \cdots, n\}$ is subject to

$$P[b, n] \le \left(\frac{\mu}{b}\right)^{b} \left(\frac{n-\mu}{n-b}\right)^{n-b} \text{ for } \mu + 1 \le b \le n \tag{10}$$

By an obvious reversal, the same function of (μ, b, n) is an upper bound on $P[0, b]$ for $0 \le b \le \mu - 1$. Numerous other bounds for binomial probabilities are known [39, 83, 112, 66, 14, 86, 69], any of which can be used to bound the tails of a PF sequence via (8). Appendix A of [3] derives the following simpler bounds for all PF distributions P which are adequate for many purposes. For all $c > 0$,

$$P[0, \mu - c] \le \exp\left(-\frac{c^2}{2\mu}\right) \text{ and } P[\mu + c, n] \le \exp\left(-\frac{c^2}{2\mu} + \frac{c^3}{2\mu^2}\right) \tag{11}$$

See [64, 63, 44, 39, 3, 69] for variations, refinements and generalizations of these inequalities, and references to earlier results. If both the variance σ^2 and the mean μ are known or can be bounded, further tail bounds are available for a PF distribution which are sharper than either the above estimates or Chebychev's inequality [12, 69].

Quite a different kind of bound, discovered by Nicholas Bernoulli for binomial probabilities around 1713, is presented in Section 16.3 of Hald [54]. This bound generalizes as follows to any PF_2 distribution on the integers, derived as in (2) by normalization of a summable PF_2 sequence (a_k): for integers $b \leq m \leq c$ with $a_m > 0$,

$$P[b, c] \geq 1 - \max(a_b, a_c)/a_m \tag{12}$$

The bound is non-trivial only if both a_b and a_c are less than a_m, so the best choice of m is a mode of the distribution, as discussed in the next paragraph. Note that the bound can be computed without knowing the constant of normalization $A := \sum_k a_k$ when P is defined via (2). Let $\sum[i, j] = \sum_{i \leq k \leq j} a_k$. By choosing b, c and m so that $\max(a_b, a_c)/a_m < \epsilon$, the probability outside $[b, c]$ is at most ϵ. Probabilities $P[i, j]$ for $b \leq i \leq j \leq c$ are therefore approximated from above by $\sum[i, j]/\sum[b, c]$ with a relative error of at most ϵ.

Darroch's Rule for the Mode [27]. As a well known consequence of Newton's inequality (1), a PF sequence $(a_0, \cdots, a_n)$ has either a unique index m or two consecutive indices m such that $a_m = \max_k a_k$. Darroch showed that such a *mode* m differs from the mean μ by less than 1. This remarkable result seems to be quite unknown to combinatorialists, though it has numerous combinatorial applications indicated in the next section. To be more precise, acording to Theorem 4 of [27], for integer k with $0 \leq k \leq n$

$$\begin{aligned}
m &= k \text{ if} & & k \leq \mu < k + \tfrac{1}{k+2} \\
m &= k \text{ or } k+1 \text{ or both if} & & k + \tfrac{1}{k+2} \leq \mu \leq k + 1 - \tfrac{1}{n-k+1} \\
m &= k+1 \text{ if} & & k + 1 - \tfrac{1}{n-k+1} < \mu \leq k + 1
\end{aligned} \tag{13}$$

Jogdeo-Samuels [70] gave a similar result for the *median* instead of the mode.

Bounds for Consecutive Ratios. By Newton's inequality, the consecutive ratios a_k/a_{k+1} derived from a PF sequence are strictly increasing over the range where they are well defined. Useful bounds for these ratios can be obtained as follows.

Given a frequency sequence (a_k) and $\theta > 0$, consider the *tilted sequence* $(a_k\theta^k)$ associated with $A(\theta z)$. The mean of the probability distribution obtained by normalization of the tilted sequence is

$$\mu(\theta) := \frac{\theta A'(\theta)}{A(\theta)} \tag{14}$$

and its variance is found from (4) to be

$$\sigma^2(\theta) := \frac{\theta^2 A''(\theta)}{A(\theta)} + \frac{\theta A'(\theta)}{A(\theta)} - \left(\frac{\theta A'(\theta)}{A(\theta)}\right)^2 = \theta\,\mu'(\theta) \tag{15}$$

Let $m(\theta)$ be the least m such that $a_m/a_{m+1} \geq \theta$. If (a_k) is a PF sequence, then so is the tilted sequence $(a_k\theta^k)$, and $m(\theta)$ is a mode of this tilted sequence. So Darroch's rule gives

$$|m(\theta) - \mu(\theta)| < 1 \qquad\qquad (\theta > 0) \tag{16}$$

Asuming both $a_k > 0$ and $a_{k+1} > 0$, the more precise version of Daroch's rule stated above gives

$$\frac{a_k}{a_{k+1}} \geq \theta \text{ if } \mu(\theta) \leq k + \frac{1}{k+2} \tag{17}$$

$$\frac{a_k}{a_{k+1}} \leq \theta \text{ if } \mu(\theta) \geq k + 1 - \frac{1}{n-k+1} \tag{18}$$

Let ℓ be the least k and r the greatest k such that $a_k > 0$. For (a_k) with $\ell < r$ it follows from (15) that $\mu'(\theta) > 0$, hence that $\mu(\theta)$ is continuous and strictly increasing from ℓ to r as θ increases from 0 to ∞. For $0 < x < \infty$ let

$$\theta(x) \text{ be the unique positive solution of } \theta A'(\theta)/A(\theta) = x. \tag{19}$$

Then (17) and (18) combine to show that if a polynomial $A(z) := \sum_{k=0}^{n} a_k z^k$ with non-negative coefficients has only real zeros, then

$$\theta\left(k + \frac{1}{k+2}\right) \leq \frac{a_k}{a_{k+1}} \leq \theta\left(k + 1 - \frac{1}{n-k+1}\right) \tag{20}$$

Because $\theta(x)$ is a strictly increasing function of x, (20) implies

$$\theta(k) < \frac{a_k}{a_{k+1}} < \theta(k+1) \tag{21}$$

The Normal Approximation. ([58, 127, 10, 21, 98]). Let

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \tag{22}$$

denote the standard normal density function, and let

$$\Phi(z) = \int_{-\infty}^{z} \phi(x)dx \tag{23}$$

Then for every PF distribution P on $\{0, 1, \cdots, n\}$ with mean μ, variance σ^2,

$$\max_{0 \le k \le n} \left| P[0,k] - \Phi\left(\frac{k-\mu}{\sigma}\right) \right| < \frac{0.7975}{\sigma} \tag{24}$$

and there exists a universal constant C such that

$$\max_{0 \le k \le n} \left| \sigma P(k) - \phi\left(\frac{k-\mu}{\sigma}\right) \right| < \frac{C}{\sigma} \tag{25}$$

The estimate (24) follows from a refinement of the Berry-Esseen theorem [127]. The bound (25) is due to Platonov [98], Theorem 11.2. According to Remark 4 of Vatutin and Mikhailov [128], the more general result claimed by Platonov is false, but his argument is correct for a PF distribution. See also Canfield [21] for a local limit bound of a weaker form with explicit constants that applies to more general sequences. An explicit C in (25) can doubtless be obtained by a more careful analysis using the Fourier method of [98].

As a consequence of the above estimates, if (S_n) is a sequence of random variables such that S_n has a PF distribution with mean μ_n and variance σ_n^2, the asymptotic distribution of $(S_n - \mu_n)/\sigma_n$ is standard normal iff $\sigma_n \to \infty$ as $n \to \infty$. Also, as a consequence of the estimate (24) and a standard weak convergence result([13], Theorem 14.2) no other continuous limit distribution besides the normal can be obtained as the asymptotic distribution of $(S_n - b_n)/c_n$ for a sequence of random variables (S_n), each with a PF distribution with finite range, and sequences of constants (b_n) and (c_n). Note that this result is not true for a sequence of random variables S_n each with a PF distribution over the set of positive integers. See Section 5 for further discussion.

Following the method used by Bender [10] for Stirling numbers, an approximation to any individual term a_k in a finite PF sequence may be obtained as follows. Replace (a_j) by the tilted sequence $(a_j\theta^j)$ and apply (25) to see that for every PF sequence $(a_0, \cdots, a_n)$

$$a_k = \frac{1}{\sqrt{2\pi}\sigma(\theta)} \frac{A(\theta)}{\theta^k}(1+\epsilon_k) \tag{26}$$

for $\theta = \theta(k)$ as in (19) and $\sigma^2(\theta)$ as in (15), and where the error term $\epsilon_k = \epsilon(a_0, \cdots, a_n, k)$ is bounded by

$$|\epsilon_k| \leq C/\sigma(\theta) \tag{27}$$

for C as in (25). The Edgeworth expansion [38, 96] suggests that in (27) the error can be bounded by C/σ^2 rather than C/σ. The asymptotic formula (26) is a close relative of Hayman's [61] generalization of Stirling's formula. The basic method traces back to Laplace [79].

See also Holst [65] for a related probabilistic method applied to occupancy problems and [106, 10, 11, 20, 77, 40, 41, 46] for normal approximations to various other kinds of combinatorial sequences.

The Poisson Approximation. From (2) and (4) the mean μ and variance σ^2 of the probability distribution P derived from a PF distribution on $\{0, 1, \cdots, n\}$ are such that

$$\mu - \sigma^2 = \sum_{i=1}^{n} p_i^2 \geq 0 \tag{28}$$

By formula (1.23) of [9], there is the following bound on the total variation distance between P and the Poisson(μ) distribution:

$$\sum_k \left| P(k) - \frac{e^{-\mu}\mu^k}{k!} \right| \leq (1 - e^{-\mu}) \left(\frac{\mu - \sigma^2}{\mu} \right) \tag{29}$$

So the Poisson approximation to a PF distribution will be good whenever $\mu - \sigma^2 << \mu$. See also Theorems 6.B and 6.H of [9] for other settings in which the same bound applies, and Corollary 3.D.1 of [9] which shows that the bound (29) cannot be improved by much more than a constant factor.

Further Inequalities and Approximations. A number of more refined inequalities and approximations for PF sequences can be read from the probabilistic literature. Typically these involve third and higher order moments of the distribution to obtain sharper approximations. See for instance [95, 7, 29, 120].

3 Operations

The collection of all finite PF sequences is closed under a number of operations which arise naturally in combinatorial applications. In particular, if $(a_0, \cdots, a_n)$ is a PF sequence, then so is the sequence $(b_0, \cdots b_n)$ obtained by each of the following operations. For those operations for which the closure property is not obvious, the works cited provide proofs, references to original sources, and various related results:

Reversal: $b_k = a_{n-k}$

Geometric Tilting: $b_k = \theta^k a_k$ for arbitrary $\theta > 0$.

Factorial Tilting: ([100], [15] Th 2.4.1) $b_k = a_k/k!$

Binomial Moments: ([74], Th. 8.6.2) $b_k = \sum_{i=0}^{n} \binom{i}{k} a_i$

Further, if $(a_0, \cdots, a_n)$ and $(b_0, \cdots, b_n)$ are two PF sequences, then so is each of the sequences $(c_0, \cdots, c_n)$ defined by

Convolution: $c_k = \sum_{j=0}^{k} a_j b_{k-j}$.

Product: ([92] Satz 7.4,[15] Th. 4.7.8) $c_k = a_k b_k k!$ and hence also $c_k = a_k b_k$

Closure under the product operations is useful in combinatorial examples, but not at all obvious probabilistically. An interpretation of the probability distribution derived from $(a_k b_k)$ can be given as follows. Let S_n be the number of successes in some sequence of n independent trials $P(S_n = k) =$

a_k/A, and let T_n be the number of successes in some further sequence of n independent trials with success probabilities arranged so that $P(T_n = k) = b_k/B$. Assuming the two sets of n trials are independent, the conditional distribution of S_n given $S_n = T_n$ is the distribution obtained by normalization of $(a_k b_k)$. But it is not at all apparent probabilistically why this distribution is representable as the distribution of the number of successes in some other set of n trials.

4 Examples

Throughout this section, arrays are indexed by n and k with $n = 1, 2 \cdots$ and $0 \le k \le n$. Call (a_{nk}) a *combinatorial array* if the a_{nk} are non-negative integers. A combinatorial array is usually defined by letting a_{nk} be the number of elements ω in some finite set Ω_n such that $S_n(\omega) = k$ for some function $S_n : \Omega_n \to \{0, 1, \cdots, n\}$. Then S_n may be viewed as a random variable defined on the *combinatorial probability space* defined by Ω_n equipped with the uniform probability distribution. For example, take $\Omega_n = \{0, 1\}^n$ and $S_n(\omega)$ to be the number of 1's in the sequence ω to obtain the array of binomial coefficients $a_{nk} = \binom{n}{k}$.

According to Proposition 1, a combinatorial array is a PF array iff for each n the random variable S_n defined on an associated combinatorial probability space has the same distribution as $\hat{S}_n = \sum_{k=1}^{n} X_{nk}$ where $(X_{nk}, 1 \le k \le n)$ are independent Bernoulli(p_{nk}) random variables defined on some probability space $\hat{\Omega}_n$ for some sequence (p_{nk}). Given a combinatorial PF array, it may or may not be possible to implement this construction of independent X_{nk} on a combinatorial probability space Ω_n equipped with uniform distribution. It is easy to do this for the array of binomial coefficients, and for the array of Stirling numbers of the first kind, as indicated below. But such a construction is impossible for the array of Stirling numbers of the second kind. Still, for any PF array, X_{nk} can be defined as the kth co-ordinate map on $\hat{\Omega}_n := \{0, 1\}^n$ equipped with the product measure P_n determined by the $p_{nk} = P_n(X_{nk} = 1)$.

Let $[n] := \{1, \cdots, n\}$. The notation for Stirling numbers follows [52].

The number of cycles of a random permutation of $[n]$**.** The array of *unsigned Stirling numbers of the first kind* [26] is defined by

$$\begin{bmatrix} n \\ k \end{bmatrix} = \text{ number of permutations of } [n] \text{ with } k \text{ cycles} \tag{30}$$

The associated polynomial admits the elementary factorization

$$\sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix} z^k = z(z+1)\cdots(z+n-1) = \frac{\Gamma(n+z)}{\Gamma(z)} \tag{31}$$

whose probabilistic interpretation is well known [51, 39]. In this case it is easy to construct independent random variables $X_1, \cdots, X_n$ as functions on the set Ω_n of all permutations ω of $[n]$ such that the number of cycles of ω is $X_1(\omega) + \cdots + X_n(\omega)$ and X_i has Bernoulli (i/n) distribution. For instance, write ω in standard cycle notation, and let X_i be the indicator that some cycle ends at the ith place in the cycle notation [39]. Or see [113, 71] for alternatives. Let $A_n(z)$ denote the polynomial associated with the nth row of the array, as displayed in (31). The function $\mu(n, \theta)$ derived from $A_n(z)$ as in (14) is easily calculated as

$$\mu(n, \theta) = \theta[\psi(n+\theta) - \psi(\theta)] = \sum_{j=1}^{n} \frac{\theta}{\theta + j - 1} \tag{32}$$

where $\psi(\theta) := \Gamma'(\theta)/\Gamma(\theta)$ is the digamma function. Erdös [35] showed that for $n \geq 3$ the nth row of these Stirling numbers has a unique mode m_n. According to Darroch's rule, $|m_n - \mu(n, 1)| < 1$, with more precise evaluations for some n. This complements the result of Hammersley [55] that $m_n = \log(n) + O(1)$. For $\theta > 0$ the function $\mu(n, \theta)$ gives the mean exactly, and the mode and median to within 1, for the sequence whose polynomial is $A_n(\theta z)$. This sequence defines the distribution of the number of parts in a random partition of n governed by the *Ewens sampling formula* with parameter $\theta > 0$. See [36, 32, 5, 8, 37]. Bender [10], Example 5.1 shows how the normal approximation (26) in this case yields the leading term of an asymptotic expansion for the Stirling numbers of the first kind due to Moser and Wyman [89]. There is no shortage of asymptotic approximations for these Stirling numbers [88, 124, 131, 67, 125], but little in the way of easily computable bounds. Consider for instance the problem of computing the ratio

$$r(n,k) := \begin{bmatrix} n \\ k \end{bmatrix} \begin{bmatrix} n \\ k+1 \end{bmatrix}^{-1} \tag{33}$$

for large n and k. According to (20),

$$\theta(n,k) < r(n,k) < \theta(n,k+1) \qquad (1 \leq k < n) \tag{34}$$

where $\theta(n,k)$ is the unique root θ of $\mu(n,\theta) = k$. Working in *Mathematica*, the functions $r(n,k), \mu(n,\theta)$ and $\theta(n,k)$ can each be defined as follows by one line programs:

```
r[n_, k_] := Abs[StirlingS1[n, k]/StirlingS1[n, k + 1]]//N
```

```
mu[n_, t_] := t(PolyGamma[n + t] − PolyGamma[t])//N
```

```
theta[n_, k_] := (fr = FindRoot[mu[n, t] == k, t, 1]; fr = fr[[1]]; t/.fr)
```

In my implementation of *Mathematica* the `StirlingS1` routine involved in direct computation of $r(n,k)$ produces the response "Out of memory. Exiting" for $n > 500$. However the routines for computing $\mu(n,k)$ and $\theta(n,k)$ are fast and apparently stable even for very large n. To illustrate, for $n = 10^{10}$ and $k = 10^3$, the bounds (34) so computed are

$$52.4216 < r(10^{10}, 10^3) < 52.477 \tag{35}$$

The number of subsets in a random partition of $[n]$ The array of *Stirling numbers of the second kind*, is defined by

$$\left\{ \begin{matrix} n \\ k \end{matrix} \right\} = \text{number of partitions of } [n] \text{ into } k \text{ subsets} \tag{36}$$

Let $B_n(z)$ denote the associated polynomial:

$$B_n(z) := \sum_{k=0}^{n} \left\{ \begin{matrix} n \\ k \end{matrix} \right\} z^k = \sum_{j=0}^{\infty} \frac{e^{-z} z^j j^n}{j!} \tag{37}$$

The second equality follows from the well known double generating function for the array [26]. For $z \geq 0$ the infinite sum identifies $B_n(z)$ as the nth moment of the Poisson distribution with parameter z, as observed by Riordan [104]. Let $B_n = B_n(1)$, the total number of partions of $[n]$, known as the nth *Bell number.* Even $B_3(z) = z(1 + 3z + z^2)$ does not factor over the rationals, so there no way to represent the number of subsets in a random partition of $\mathbb{N}_3$ as a sum of 3 independent indicator variables defined on a combinatorial probability space with equally likely outcomes. Still, Harper [58] proved that the Stirling numbers of the second kind form a PF array by showing by induction that the associated sequence of polynomials $B_n(z)$ is a *Sturm sequence*, that is to say they have interlaced simple real zeros. Let m_n be the mode and μ_n the mean of the distribution defined by the nth row of Stirling numbers of the second kind. It is known [31] that m_n is unique. Harper used the formula $\mu_n = B_{n+1}/B_n - 1$ to read asymptotics for μ_n from those for B_n due to Szekeres and Binet [123], and Harper gave a crude bound for $|m_n - \mu_n|$ using the normal approximation. The problem of obtaining asymptotics for m_n has been discussed by a number of subsequent authors (see Menon [87] and papers cited there). Darroch's formula $|m_n - \mu_n| < 1$ shows that asymptotics for either sequence can be simply be read from the other.

Consecutive ratios of Stirling numbers of the second kind can be estimated by the method described above for Stirling numbers of the first kind, using the formula $\mu_n(\theta) = B_{n+1}(\theta)/B_n(\theta) - 1$ and approximating $B_{n+1}(\theta)$ and $B_n(\theta)$ either by appropriate truncation of the infinite series expression (37), or by the asymptotic methods of [123]. The discussion of the classical occupancy problem below provides an even simpler approach to the estimation of these ratios for large k. See [53, 106, 45, 6] for further results about uniform random partitions of $[n]$, and see [110] for a survey of inequalities and probabilistic interpretations of Stirling numbers of both kinds.

The Hypergeometric Distribution. Suppose a random sample of size n is taken without replacement from a population of G good and B bad elements. The probability that the sample contains exactly k good elements is

$$P_{n,G,B}(k) = \binom{G}{k}\binom{B}{n-k}\binom{G+B}{n}^{-1} \tag{38}$$

For each fixed pair of non-negative integers G and B each $1 \leq n \leq G + B$, this formula defines a probability distribution $P_{n,G,B}$ on $\{0, 1, \cdots, n\}$, called the *hypergeometric distribution* with parameters (n, G, B). The formula (38) displays $P_{n,G,B}(k)$ as the product of a binomial sequence, a shifted binomial sequence, and a constant. So it follows from the closure of PF sequences under these operations that the hypergeometric distribution is a PF distribution for all (n, G, B). Vatutin-Mikhailov [128] obtained this result by showing directly that the generating polynomial has only real zeros. See also Kou-Ying [132] for another derivation involving Jacobi polynomials, and a statistical application. Many results in the statistics literature concerning the hypergeometric distribution, first obtained by other methods, can be read from the general properties of PF distributions described in Section 2. For example, Hoeffding's inequalities (8) and (9) yield inequalities of Hoeffding [63] and Uhlmann [126] comparing the hypergeometric (n, G, B) distribution for sampling without replacement with the binomial(n, p) distribution for sampling with replacement from the same population, that is with $p = G/(G + B)$. The well known normal and Poisson approximations for the hypergeometric distribution follow similarly.

The Classical Occupancy Problem In the classical occupancy problem [28], n labelled balls are thrown independently at random into N boxes. The probability distribution of the number of occupied boxes $O_{n,N}$ is then given by

$$P[O_{n,N} = k] = \left\{ {n \atop k} \right\} \binom{N}{k} \frac{k!}{N^n} \qquad (0 \leq k \leq n) \tag{39}$$

If balls labelled by $[n]$ are regarded as partitioned by the boxes, this scheme induces a particular non-uniform distribution for a random partition of $[n]$ into $O_{n,N}$ non-empty subsets. Lieb [82] showed that the generating polynomial has only real zeros for $N = n$. Harris and Park [60] showed this for all N and n. So the distribution of $O_{n,N}$ defined by (39) is a PF distribution for every N and n. Since the first two factors in (39) are PF sequences, and the remaining factor is $k!$ times a constant, this result can also be read from product rule of Section 3. The approximate normality of the distribution of $O_{n,N}$, provided the variance of $O_{n,N}$ is sufficiently large, has been known for a long time. Englund [34] obtained the estimate (24) in this case by another method, with a constant of 10.4 instead of 0.7975. See also [102] for similar but weaker bounds in a more general occupancy problem.

Another PF distribution was obtained by Park [94] from the variation of the classical occupancy problem where each ball falls through its box with some constant probability, independent of all other balls. Vatutin-Mikhailov [128] obtained a family of PF distributions from the number of occupied boxes after the following allocation scheme: first N_1 balls are placed one per box in each of N_1 boxes picked at random from N boxes, then independently N_2 balls are placed in N_2 boxes picked at random from the same N boxes, and so on, for some arbitrary finite sequence of positive integers $(N, N_1, \cdots, N_j)$ with $N \geq N_i$ for $1 \leq i \leq j$. The family of PF distributions so defined includes both the clasical occupancy distribution and the hypergeometric distibution as special cases. Another special case is the distribution of the number of occupied boxes amongst M particular boxes when n balls are placed independently at random in N boxes. (Take $N_1 = N - M, N_2 = \cdots = N_{n+1} = 1$). As shown in [128], the approach to normal and Poisson approximations for this family of occupancy distributions via general results for PF distributions is a substantial simplification of earlier approaches. See also [59].

Stam [114] describes a way to construct a uniform random partition of $[n]$ by suitably randomizing N which relates asymptotics for uniform random partitions to those for the classical occupancy scheme. See [97] for further discussion of Stam's scheme, and [73] for recent work on large deviation bounds in the classical occupancy problem. As observed by Janson [69], such bounds follow immediately from the PF representation.

It is known [77, (1.1.4)] that the random variable $O_{n,N}$ with distribution (39) has mean $N - N(1-1/N)^n$. From (39), for $1 \leq k \leq N$,

$$P(O_{n,N} = k)/P(O_{n,N} = k+1) = r_2(n,k)/(N-k) \tag{40}$$

where

$$r_2(n,k) = \left\{ {n \atop k} \right\} \left\{ {n \atop k+1} \right\}^{-1} \tag{41}$$

Darroch's rule applied to the PF distribution of $O_{n,N}$ yields

$$r_2(n,k) \geq N - k \text{ if } N - N(1-1/N)^n \leq k + \frac{1}{k+2} \tag{42}$$

$$r_2(n,k) \leq N - k \text{ if } N - N(1-1/N)^n \geq k + 1 - \frac{1}{n-k+1} \tag{43}$$

For $k \geq 1$ let $x(n,k)$ denote the unique real root in $[1,\infty)$ of the equation $x - x(1-1/x)^n = k$. Then (42) and (43) combined yield

$$\lfloor x(n,k) \rfloor - k < r_2(n,k) < \lceil x(n,k+1) \rceil - k \tag{44}$$

To illustrate, for $n = 1000$ and $k = 700$ these bounds are

$$612 < r_2(1000, 700) < 618 \tag{45}$$

Laplace's asymptotic formula for Stirling numbers of the second kind ([79, 28]), which is similar to (26) but much easier to compute, gives

$$r_2(1000, 700) \sim 614.938... \tag{46}$$

Note that because the bounds in (44) are necessarily integers, these bounds will only be tight for k much larger than the mean, that is $k >> n/\log(n)$.

Leaves of a Random Tree. Let Ω_n be the set of all n^{n-2} trees labelled by $[n]$, and for $\omega \in \Omega_n$ let $L_n(\omega)$ be the number of *leaves* of ω, that is the number of vertices of degree 1. As observed by Rényi [103], the well known Prüfer coding of random trees implies that the distribution of $L_n(\omega)$ for ω picked uniformly at random from Ω_n is identical to the distribution of the number of empty boxes when $n-2$ balls are thrown independently at random into n boxes. That is to say the distribution of L_n is the reversal of the distribution of $O_{n-2,n}$ for $O_{n,N}$ as in the classical occupancy problem. Since $O_{n-2,n}$ has a PF distribution, so does L_n. Steele [119] uses this example to illustrate interpretations of the one-parameter exponential family of Gibbs' distributions on a finite outcome space Ω, with a real parameter β, obtained by tilting the uniform distribution by a factor of $\theta^{S(\omega)}$ for an arbitrary function S defined on Ω and $\theta = e^{-\beta}$. If S has a PF distribution with generating function $A(z)/A(1)$ when ω is assigned uniform distribution, then under the Gibbs' distribution S has the PF distribution with generating function $A(\theta z)/A(\theta)$ as considered in Section 2.

Generalized Stirling Numbers. Various generalizations of both kinds of Stirling numbers are known to define PF arrays. See [18, 23] for background, and [15] for the PF results. The Munch numbers also form a PF array [85].

Random Mappings. Let $M(n,k)$ be the number of mappings from $[n]$ to $[n]$ whose associated digraph has exactly k components. See [78, 90] for background. Brenti [15] obtained the formula

$$M(n,k) = \sum_{i=1}^{n} \binom{n-1}{i-1} n^{n-i} \begin{bmatrix} i \\ k \end{bmatrix} \tag{47}$$

and used it to show that $M(n,k)$ is a PF array. Asymptotic normality in this case is due to Stepanov [121]. See also [40] for another approach, and [2] for related asymptotics and further references.

Matchings. Let G be a graph, with multiple edges allowed, and let a_k be the number of matchings of size k in G. That is, a_k is the number of k-element sets M of edges of G, no two edges in M having a common vertex. Heilmann and Lieb ([62], theorem 4.2) showed using Sturm sequences that $(a_0, \cdots, a_m)$ is a PF sequence. Special cases of this construction include both the binomial coefficients and the Stirling numbers of the second kind. Other special cases include the coefficient sequences of rook polynomials [91], and several of the classical families of orthogonal polynomials. In this connection, see also [50, 48, 129, 116]. The consequent asymptotic normality of various arrays associated with sequences of graphs was treated by Godsil [49]. Godsil's results have recently been refined by Kahn [72].

Partitions of Multisets. Another extension of the PF property of the Stirling numbers of the second kind is the following result obtained by Simion [111], also using Sturm sequences. Call a sequence of non-negative integers $\mathbf{n} := (n_1, n_2, \cdots)$ with $0 < \sum_i n_i < \infty$ a *multiset*. For a multiset $\mathbf{n}$ let $a(\mathbf{n}, k)$ be the number of sequences of multisets $(\mathbf{n}_1, \cdots, \mathbf{n}_k)$ such that $\sum_{i=1}^{k} \mathbf{n}_i = \mathbf{n}$. Then for each multiset $\mathbf{n}$ with $\sum_i n_i = n$ the sequence $(a(\mathbf{n}, k), 0 \leq k \leq n)$ is PF. Take $\mathbf{n}$ to be the sequence $(1, 1, \cdots, 1)$ of length n to deduce that the array $\left\{ {n \atop k} \right\} k!$ is a PF array. As noted by Simion, this is a stronger result than the PF property of the array $\left\{ {n \atop k} \right\}$, due to the general factorial tilting rule of Section 3.

Eulerian Numbers Let Ω_n be the set of permutations of $[n]$. For $\sigma \in \Omega_n$ let

$$D_n(\sigma) = \#\{1 \leq i < n : \sigma_{i+1} < \sigma_i\}$$

be the *number of descents* of σ, and let A_{nk} be the number of $\sigma \in \Omega_n$ such that $D(\sigma) = k$. The numbers A_{nk} form the array of *Eulerian numbers* [43, 42]. It was known already to Frobenius that for each $n \geq 2$ the Eulerian polynomial $\sum_k A_{nk} z^k$ has only real zeros. So the Eulerian numbers form a PF array. The mean and variance of D_n, viewed as a function of a uniformly distributed random permutation σ, are easily shown to be $(n-1)/2$ and $(n-1)/12$. The asymptotic normality of the distribution of D_n was deduced in [22] and [10] from the PF property of Eulerian numbers. There is another probabilistic representation of the Eulerian numbers related to the formula [26, p. 243]

$$A_{nk} = \sum_{i=0}^{k} (-1)^i (k-i)^n \binom{n+1}{i}. \tag{48}$$

Comparison of this formula with the formula of Laplace [79] for the distribution of the sum $U_1 + \cdots + U_n$ of independent uniform $[0,1]$ random variables $U_1, \cdots, U_n$, reviewed in Feller [38] and Diaconis-Efron [30], shows that the probability that a random permutation of $[n]$ has k descents is

$$P(D_n = k) = \frac{A_{nk}}{n!} = P(k-1 \leq U_1 + \cdots + U_n < k) \tag{49}$$

This identity allows exceptionally accurate normal approximations for the Eulerian numbers to be deduced from corresponding approximations for the sum $U_1 + \cdots + U_n$ [38]. Stanley [115] gives a geometric proof of (49) without involving the explicit formula (48). A quick probabilistic proof of (49) can be given as follows. Let $S_n = U_1 + \cdots + U_n$ and let V_n be S_n modulo 1. It is easily verified that $V_1, \cdots, V_n$ are independent and uniform on $[0,1]$ and that

$$S_n = \lfloor S_n \rfloor + V_n \text{ where} \tag{50}$$

$$\lfloor S_n \rfloor = \#\{1 \leq i < n : V_{i+1} < V_i\} = \#\{1 \leq i < n : \sigma_{i+1} < \sigma_i\} \tag{51}$$

where $\sigma_i = \#\{1 \leq j \leq n : V_j \leq V_i\}$, and the possibility of ties among the V_i can be ignored as such ties occur with probability zero. Thus $\lfloor S_n \rfloor = D(\sigma)$ with probability one where σ is a uniformly distributed random permutation of $[n]$, and the formula (49) follows immediately.

Further Examples. See Brenti [15] for extensive discussion of techniques for proving that a sequence is PF and a wealth of further examples. See [105, 116, 16, 117, 118] for still more instances of PF arrays, and further results regarding log-concave and unimodal sequences.

5 Infinite Polya Frequency Sequences

Associate the sequence $(a_0, a_1, \cdots)$ with its generating function

$$A(z) := \sum_{i=1}^{\infty} a_i z^i \tag{52}$$

According to the result of Edrei [33] presented in [74] as Theorem 5.3 of Ch. 8, a sequence $(a_0, a_1, \cdots)$ with $a_0 = 1$ is a PF sequence iff its generating function can be expressed as

$$A(z) = e^{\lambda z} \prod_{i=1}^{\infty} \frac{(1+\alpha_i z)}{(1-\beta_i z)} \tag{53}$$

for some $\lambda \geq 0, \alpha_i \geq 0, \beta_i \geq 0$ where $\sum_i (\alpha_i + \beta_i) < \infty$. To interpret this result probabilistically, let the probability distribution of a non-negative integer valued random variable X be described either by a sequence of probabilities $P(X=0), P(X=1), \cdots$ or by the corresponding probability generating function $\sum_i P(X=i) z^i$. In particular, the probability generating function of X with Bernoulli(p) distribution on $\{0,1\}$ is $(1-p)+pz$. Say X has a *geometric*(p) distribution if

$$\sum_i P(X=i) z^i = \sum_i p q^i z^i = \frac{p}{(1-qz)} \tag{54}$$

In the representation (53), write $(1+\alpha_i z) = (q_i + p_i z)/q_i$ where $p_i + q_i = 1$, and the apply standard properties of probability generating functions [39], to deduce the following proposition. Combined with the representation formula (53) this result contains the equivalence of conditions (ii) and (iii) in Proposition 1 as the special case when the sequence $(a_0, a_1, \cdots)$ has only a finite number of non-zero entries.

Proposition 2 *Let $(a_0, a_1, \cdots)$ be a sequence of non-negative real numbers such that $0 < A < \infty$ where $A := \sum_i a_i$. The sequence is a PF sequence iff the normalized sequence $(a_0/A, a_1/A, \cdots)$ is the distribution of*

$$\sum_j X_j + \sum_j Y_j + Z \tag{55}$$

for independent random variables $X_1, X_2, \cdots$, $Y_1, Y_2, \cdots$, Z, where X_j has Bernoulli(p_j) distribution, Y_j has geometric(β_j) distribution, and Z has Poisson(λ) distribution, for some $0 \le p_i \le 1$ and $0 \le \beta_i < 1$ such that $\sum_i (p_i + \beta_i) < \infty$, and $0 \le \lambda < \infty$.

Conditions for asymptotic normality of a sequence of PF sequences can easily be deduced from Lindeberg's theorem ([13] Th. 27.2). Note that due to the possibility of geometric components, for infinite sequences a large variance alone is not enough to ensure a good normal approximation. Because a PF sequence is obtained by restriction to a lattice of a Pólya frequency function defined on the whole line [74], any probability distribution on $(0, \infty)$ whose density is a given by such a function, for instance an exponential or gamma distribution, may be obtained as a weak limit of some sequence of rescaled PF distributions on $\{0, 1, 2, \cdots\}$. Theorem 9.5 of Ch 8. of Karlin [74] provides the analog of the representation (53) for a PF sequence indexed by the set of integers, which for a summable sequence has a similar probabilistic interpretation.

6 Acknowledgements

Thanks to Lucien Le Cam, Charles Stone, Persi Diaconis and Richard Stanley for help with the literature. Thanks also to Francesco Brenti, Richard Dudley, Svante Janson, Jeff Kahn, Richard Stanley, Charles Stein and Zhiliang Ying for providing preprints.

References

[1] M. Aissen, I.J. Schoenberg, and A.M. Whitney. On the generating function of totally positive sequences I. *J. d'Anal. Math.*, 2:93–103, 1952.

[2] D. Aldous and J. Pitman. Brownian bridge asymptotics for random mappings. *Random Structures and Algorithms*, 5:487–512, 1994.

[3] N. Alon and J.H. Spencer. *The Probabilistic Method*. Wiley, New York, 1992.

[4] T. Ando. Totally positive matrices. *Linear Algebra and Its Applications*, 90:165–219, 1987.

[5] R. Arratia, A.D. Barbour, and S. Tavaré. Poisson process approximations for the Ewens sampling formula. *Ann. Appl. Prob.*, 2:519–535, 1992.

[6] R. Arratia and S. Tavaré. Independent process approximations for random combinatorial structures. *Adv. Math.*, pages 90–154, 1994.

[7] A. D. Barbour. Asymptotic expansions in the Poisson limit theorem. *The Annals of Probability*, 15:748 – 766, 1987.

[8] A. D. Barbour. Refined approximations for the Ewens sampling formula. *Rand. Struct. Alg.*, 3:267–276, 1992.

[9] A. D. Barbour, L. Holst, and S. Janson. *Poisson approximation*. Clarendon Press, 1992.

[10] E. A. Bender. Central and local limit theorems applied to asymptotic enumeration. *J. Combin. Theory Ser. A*, 15:91–111, 1973.

[11] E. A. Bender. Asymptotic methods in enumeration. *SIAM Rev.*, 16:485–515, 1974.

[12] G. Bennett. Probability inequalities for the sum of independent random variables. *J. Amer. Stat. Assoc.*, 57:33 – 45, 1962.

[13] P. Billingsley. *Probability and Measure*. Wiley, New York, 1995. 3rd ed.

[14] B. Bollabas. *Random Graphs*. Academic, New York, 1985.

[15] F. Brenti. Unimodal, Log-Concave, and Pólya Frequency Sequences in Combinatorics. *Mem. Amer. Math. Soc.*, 413, 1989.

[16] F. Brenti. Combinatorics and total positivity. *J. Comb. Theory A*, 71:175–218, 1995.

[17] F. Brenti. The aplication of total positivity to combinatorics, and conversely. In M. Gasca and C. A. Micchelli, editors, *Total Positivity and its Applications*, volume 359 of *Mathematics and its applications*, pages 451–473. Kluwer, 1996.

[18] Z. Broder. The r-Stirling numbers. *Discrete Mathematics*, 49:241–259, 1984.

[19] L.D. Brown, I.M. Johnstone, and K. B. MacGibbon. Variation diminishing transformations: A direct approach to total positivity and its statistical applications. *J. Amer. Stat. Assoc.*, 76:824 – 832, 1981.

[20] E.R. Canfield. Central and local limit theorems for the coefficients of polynomials of binomial type. *J. Comb. Theory A*, 23:275–290, 1977.

[21] E.R. Canfield. Application of the Berry-Esseen inequality. *J. Combin. Theory Ser. A*, 28:17–25, 1980.

[22] L. Carlitz, D. C. Kurtz, R. Scoville, and O. P. Stackelberg. Asymptotic properties of Eulerian numbers. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 23:47 – 54, 1972.

[23] Ch. A. Charalambides and J. Singh. A review of the Stirling numbers, their generalizations and statistical applications. *Commun. Statist.-Theory Meth.*, 17:2533–2595, 1988.

[24] P.L. Chebyshev. Démonstration élémentaire d'une proposition générale de la théorie des probabilités. *J. für die reine und angewandte Math.*, 33:259–267, 1846. P.L. Tchebychef, *Oeuvres*,1899,16-26.

[25] H. Chernoff. A measure of the asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Stat.*, 23:493–509, 1952.

[26] L. Comtet. *Advanced Combinatorics.* D. Reidel Pub. Co., Boston, 1974. (translated from French).

[27] J. N. Darroch. On the distribution of the number of successes in independent trials. *Ann. Math. Stat.*, 35:1317 – 1321, 1964.

[28] F.N. David and D.E. Barton. *Combinatorial Chance*. Griffins, London, 1962.

[29] P. Deheuvels, M.L. Puri, and S.S. Ralescu. Asymptotic expansions for sums of nonidentically distributed Bernoulli random variables. *Journal of Multivariate Analysis*, 28:282 – 303, 1989.

[30] P. Diaconis and B. Efron. Probabilistic-geometric theorems arising from the analysis of contingency tables. In A.E. Gelfand, editor, *Contributions to the Theory and Applications of Statistics*, pages 103–125. Academic Press, 1987.

[31] A.J. Dobson. A note on Stirling numbers of the second kind. *J. Comb. Theory A.*, 5:212–214, 1968.

[32] P. Donnelly, T.G. Kurtz, and S. Tavaré. On the functional central limit theorem for the Ewens sampling formula. *Ann. Appl. Prob.*, 1:539–545, 1991.

[33] A. Edrei. Proof of a conjecture of of Schoenberg on the generating function of totally positive sequences. *Canad. J. Math. Soc*, 5:86–94, 1953.

[34] G. Englund. A remainder term estimate for the normal approximation in classical occupancy. *The Annals of Probability*, 9:684 – 692, 1981.

[35] P. Erdös. On a conjecture of Hammersley. *J. Lond. Math. Soc.*, 28:232–236, 1953.

[36] W.J. Ewens. The sampling theory of selectively neutral alleles. *Theor. Popul. Biol.*, 3:87 – 112, 1972.

[37] W.J. Ewens and S. Tavaré. The Ewens sampling formula. To appear in *Multivariate Discrete Distributions* edited by N.S. Johnson, S. Kotz, and N. Balakrishnan, 1995.

[38] W. Feller. *An Introduction to Probability Theory and its Applications, Vol 2.* Wiley, 1966.

[39] W. Feller. *An Introduction to Probability Theory and its Applications, Vol 1,3rd ed.* Wiley, New York, 1968.

[40] P. Flajolet and M. Soria. Gaussian limiting distributions for the number of components in combinatorial structures. *J. Comb. Theory A.*, 153:165–182, 1990.

[41] P. Flajolet and M. Soria. General combinatorial schemes: Gaussian limit distributions and exponential tails. *Discrete Mathematics*, 114:159–180, 1993.

[42] D. Foata. Distributions eulériennes et mahoniennes sur le groupe des permutations. In M. Aigner, editor, *Higher Combinatorics*. Reidel, Dordrecht, 1977.

[43] D. Foata and M.P. Schützenberger. *Théorie Géometrique des Polynômes Eulériens*. Springer, Lecture Notes in Math. 138, 1970.

[44] D. Freedman. Another note on the Borel-Cantelli lemma and the strong law, with the Poisson approximation as a by-product. *The Annals of Probability*, 1:910 – 925, 1973.

[45] B. Fristedt. The structure of partitions of large sets. Technical Report 86-154, University of Minnesota Mathematics Dept., Minneapolis, Minnesota, 1987.

[46] Z. Gao and L.B. Richmond. Central and local limit theorems applied to asymptotic enumeration IV: multivariate generating functions. *Journal of Computational and Applied Mathematics*, 41:177–186, 1992.

[47] L. J. Gleser. On the distribution of the number of successes in independent trials. *The Annals of Probability*, 3:182 – 188, 1975.

[48] C.D. Godsil. Hermite polynomials and a duality relation for matching polynomials. *Combinatorica*, 1:257–262, 1981.

[49] C.D. Godsil. Matching behaviour is asymptotically normal. *Combinatorica*, 1:369–376, 1981.

[50] C.D. Godsil and I. Gutman. On the theory of matching polynomials. *J. Graph Theory*, 5:137–144, 1981.

[51] V. Gončharov. On the field of combinatory analysis. *Amer. Math. Soc. Transl.*, 19:1–46, 1962.

[52] R. L. Graham, D.E. Knuth, and O. Patashnik. *Concrete Mathematics: a foundation for computer science. 2nd ed.* Addison-Wesley, Reading, Mass., 1989.

[53] J. Haigh. Random equivalence relations. *Journal of Combinatorial Theory*, 13:287–295, 1972.

[54] A. Hald. *A History of Probability and Statistics and their Applications before 1750.* Wiley, New York, 1990.

[55] J. M. Hammersley. The sum of products of the natural numbers. *Proc. Lond. Math. Soc. (3)*, 1:435–452, 1951.

[56] B. G. Hansen. On log-concave and log-convex infinitely divisible sequences and densities. *The Annals of Probability*, 16:1832 – 1839, 1988.

[57] G.H. Hardy, J.E. Littlewood, and G. Pólya. *Inequalities, 2nd Ed.* Cambridge Univ. Press, Cambridge, England, 1952.

[58] L. H. Harper. Stirling behavior is asymptotically normal. *Ann. Math. Stat.*, 38:410–414, 1966.

[59] B. Harris, M. Marden, and C.J. Park. The distribution of the number of empty cells in a generalized random allocation scheme. *Annals of Discrete Mathematics*, 33:77–90, 1987.

[60] B. Harris and C.J. Park. A note on the asymptotic normality of the distribution of the number of cells in occupancy problems. *Ann. Inst. Statist. Math. (Tokyo)*, 23:507–513, 1971.

[61] W. Hayman. A generalization of Stirling's formula. *Journal für die reine und angewandte Mathematik*, 196:67–95, 1956.

[62] O.J. Heilmann and E.H. Lieb. Theory of monomer-dimer systems. *Commun. Math. Phys.*, 25:190–232, 1972.

[63] W. Hoeffding. Probability inequalities for sums of bounded random variables. *J. Amer. Stat. Assoc.*, 58:13 – 30, 1963.

[64] W. Hoeffding. On the distribution of the number of successes in independent trials. *Ann. Math. Stat.*, 35:713 – 721, 1964.

[65] L. Holst. On numbers related to partitions of unlike objects and occupancy problems. *Europ. J. Combinatorics*, 2:231–237, 1981.

[66] R.J. Holt, R.M. Dudley, D.Y. Gao, and L. Pakula. Handbook and tables of classic probabilities. Unpublished, 1994.

[67] H-K Hwang. Asymptotic expansions for the Stirling numbers of the first kind. *J. Combin. Theory A*, 71:343–351, 1995.

[68] E.K. Ifantis and P.N. Panagopoulos. On the zeros of a class of polynomials defined by a three term recurrence relation. *J. Math. Anal. and Appl.*, 182:361–370, 1994.

[69] S. Janson. Large deviation inequalities for sums of indicator variables. Preprint, 1994.

[70] K. Jogdeo and S. M. Samuels. Monotone convergence of binomial probabilities and a generalization of Ramanujan's equation. *The Annals of Mathematical Statistics*, 39:1191 – 1195, 1968.

[71] P. Joyce and S. Tavaré. Cycles, permutations and the structure of the Yule process with immigration. *Stochastic Process. Appl.*, 25:309–314, 1987.

[72] J. Kahn. A normal law for matchings. Preprint, Dept. Math. and RUTCOR, Rutgers University, 1996.

[73] A. Kamath, R. Motwani, K. Palem, and P. Spirakis. Tail bounds for occupancy and the satisfiability threshhold conjecture. *Random Structures and Algorithms*, 7:59–80, 1995.

[74] S. Karlin. *Total Positivity.* Stanford Univ. Press, Stanford, 1968.

[75] J. Keilson. A threshold for log-concavity for probability generating functions and associated moment inequalities. *The Annals of Mathematical Statistics*, 43:1702 – 1708, 1972.

[76] J. Keilson and H. Gerber. Some results for discrete unimodality. *Journal of the American Statistical Association*, 66:386 – 389, 1971.

[77] V. F. Kolchin, B.A. Sevastyanov, and V.P. Christyakov. *Random Allocations*. V.H. Winston & Sons, Washington, D. C., 1978.

[78] V.F. Kolchin. *Random Mappings*. Optimization Software, New York, 1986. (Translation of Russian original).

[79] P.S. Laplace. *Théorie Analytique des Probabilités (2nd Ed.)*. Paris, 1814.

[80] P Lévy. Sur une propriété de la loi de Poisson relative aux petites probabilités. *Soc. Math. de France, Comptes rendus des scéances de l'annee 1936*, page 29, 1937.

[81] P. Lévy. *Théorie de l'Addition des Variables Aléatoires, 2nd ed.* Gauthier-Villars, Paris, 1954.

[82] E.H. Lieb. Concavity properties and a generating function for Stirling numbers. *J. Combin. Theory A*, 5:203–206, 1968.

[83] J. E. Littlewood. On the probability in the tail of a binomial distribution. *Advances in Applied Probability*, 1:43 – 72, 1969.

[84] A. W. Marshall and I. Olkin. *Inequalities – Theory of Majorization and its Applications*. Academic Press, 1979.

[85] J. R. McGregor. Asymptotic normality of the Munch numbers. *Scand. J. Statist.*, 15:299–301, 1988.

[86] B.D. McKay. On Littlewood's estimate for the binomial distribution. *Advances in Applied Probability*, 21:475 – 478, 1989.

[87] V.V. Menon. On the maximum of Stirling numbers of the second kind. *J. Comb. Theory A*, 15:11–24, 1973.

[88] P.A.P. Moran. Asymptotic analysis and local lattice central limit theorems. *J. Appl. Prob.*, 16:541–553, 1979.

[89] L. Moser and M. Wyman. Asymptotic development of the Stirling numbers of the first kind. *Journal London Math. Soc.*, pages 133–146, 1958.

[90] L. Mutafciev. Probability distributions and asymptotics for some characteristics of random mappings. In W. Grossman et al., editor, *Proc. 4th Pannonian Symp. on Math. Statist.*, pages 227–238, 1983.

[91] A. Nijenhuis. On permanents and zeros of rook polynomials. *J. Combin. Theory A*, 21:240–244, 1976.

[92] N. Obreschkov. *Verteilung und Berechnung der Nullstellen reeller Polynome.* Hochschulbücher für Mathematik; Bd. 55. Deutscher Verlag der Wissenschaften, Berlin, 1963.

[93] M. Okamoto. Some inequalities relating to the partial sum of binomial probabilities. *Ann. Inst. Statist. Math. (Tokyo)*, 10:29–35, 1958.

[94] C. J. Park. A note on the classical occupancy problem. *The Annals of Mathematical Statistics*, 43:1698 – 1701, 1972.

[95] O.E. Percus and J.K. Percus. Probability bounds on the sum of independent nonidentically distributed binomial random variables. *SIAM Journal on Applied Mathematics*, 45:621 – 640, 1985.

[96] V.V. Petrov. *Limit Theorems of Probability Theory : Sequences of Independent Random Variables.* Oxford : Clarendon Press ; New York : Oxford University Press, 1995.

[97] J. Pitman. Some probabilistic aspects of set partitions. Technical Report 452, Dept. Statistics, U.C. Berkeley, 1996. To appear in *The American Mathematical Monthly.*

[98] M.L. Platonov. *Combinatorial Numbers of a Class of Mappings and Their Applications (in Russian).* Nauka, Moscow, 1980.

[99] S.D. Poisson. *Recherches sur la probabilité des jugements en matière criminelle et en matière civile.* Bachelier, Paris, 1837.

[100] G. Pólya and J. Schur. Über Zwei Arten von Faktorfolgen. *J. Reine Angew. Math.*, 144:89–133, 1914.

[101] G. Pólya and G. Szegö. *Problems and Theorems in Analysis II.* Springer-Verlag, New York, 1976.

[102] M. P. Quine and J. Robinson. A Berry-Esseen bound for an occupancy problem. *The Annals of Probability*, 10:663 – 671, 1982.

[103] A. Rényi. Some remarks on the theory of trees. *MTA Mat. Kut. Inst. Kozl.*, 4:73–85, 1959.

[104] J. Riordan. Moment recurrence relations for the binomial Poisson and hypergeometric probability distributions. *Ann. Math. Stat.*, 8:103–111, 1937.

[105] A. Ruciński. The behaviour of $\binom{n}{k,\ldots,k,n-ik} c^i/i!$ is asymptotically normal. *Discrete Mathematics*, 49:287–290, 1984.

[106] V. N. Sachkov. Random partitions of sets. *Theory Probab. Appl.*, 19:184–190, 1973.

[107] S. M. Samuels. On the number of successes in independent trials. *The Annals of Mathematical Statistics*, 36:1272 – 1278, 1965.

[108] I.J. Schoenberg. On the zeros of the generating functions of multiply positive sequences and functions. *Ann. Math.*, 62:447–471, 1955.

[109] M. Shaked and J. G. Shanthikumar. Characterization of some first passage times using log-concavity and log-convexity as aging notions. *Probability in the Engineering and Informational Sciences*, 1:279–291, 1987.

[110] M. Sibuya. Log-concavity of Stirling numbers and unimodality of Stirling distributions. *Ann. Inst. Statist. Math.*, 40:693–714, 1988.

[111] R. Simion. A multiindexed Sturm sequence of polynomials and unimodality of certain combinatorial sequences. *J. Combin. Theory A*, 36:15–22, 1984.

[112] E. V. Slud. Distribution inequalities for the binomial law. *The Annals of Probability*, 5:404 – 412, 1977.

[113] A.J. Stam. Cycles of random permutations. *Ars Combinatorica*, 16:43–48, 1983.

[114] A.J. Stam. Generation of a random partition of a set by an urn model. *J. Combin. Theory A*, 35:231–240, 1983.

[115] R. Stanley. Eulerian partitions of a unit hypercube. In M. Aigner, editor, *Higher Combinatorics*. Reidel, Dordrecht, 1977.

[116] R. Stanley. Log-concave and unimodal sequences in Algebra, Combinatorics and Geometry. *Ann. New York Acad. Sci.*, 576:500–534, 1989.

[117] R. Stanley. Graph colorings and related symmetric functions: ideas and applications. Preprint, Dept. Math. MIT, 1995.

[118] R. Stanley. A matrix for counting paths with acyclic digraphs. Preprint, Dept. Math. MIT, 1995.

[119] J. M. Steele. Gibbs' measures on combinatorial objects and the central limit theorem for an exponential family of random trees. *Probability in the Engineering and Informational Sciences*, 1:47–59, 1987.

[120] C. Stein. Application of Newton's identities to a generalized birthday problem and to the Poisson Binomial distribution. Technical Report 354, Dept. Statistics, Stanford University, 1990.

[121] V.E. Stepanov. Limit distributions of certain characteristics of random mappings. *Theory Probab. Appl.*, 14:612–626, 1969.

[122] F. W. Steutel. Log-concave and log-convex distributions. *Encyclopedia of Statistical Sciences*, 9:247, 1985.

[123] G. Szekeres and F.E. Binet. On Borel fields over finite sets. *Ann. Math. Stat.*, 28:494–498, 1957.

[124] N.M. Temme. Asymptotic estimates for Stirling numbers. *Studies in Applied Math.*, 89:233–243, 1993.

[125] E.G. Tsylova. Probabilistic methods for obtaining asymptotic formulas for generalized Stirling numbers. *Journal of Math. Sciences*, 75:1607–1614, 1995.

[126] W. Uhlmann. Vergleich der hypergeometrischen mit der Binomial-Verteilung. *Metrika*, 10:145–158, 1966.

[127] P. van Beek. An application of Fourier methods to the problem of sharpening the Berry-Esseen inequality. *Z. Wahrsch. Verw. Gebiete*, 23:187–196, 1972.

[128] V. A. Vatutin and V. G. Mikhailov. Limit theorems for the number of empty cells in an equiprobable scheme for group allocation of particles. *Theory of Probability and its Applications (Transl. of Teorija Verojatnostei i ee Primenenija)*, 27:734 – 743, 1982.

[129] G. Viennot. Combinatorial theory for general orthogonal polynomials with extensions and applications. In *Polynômes Orthogonaux et Applications*, volume 1171 of *Lecture Notes in Math.*, pages 139–157. Springer, Berlin-New York, 1985.

[130] D. W. Walkup. Pólya sequences, binomial convolution and the union of random sets. *Journal of Applied Probability*, 13:76 – 85, 1976.

[131] H. Wilf. The asymptotic behaviour of the Stirling numbers of the first kind. *J. Combin. Theory Ser. A*, 64:344–349, 1993.

[132] S.G. You and Z. Ying. Asymptotics for a 2×2 table with fixed margins. Preprint, Dept. Statistics, Rutgers Universtity, 1996.