

# THE MAHONIAN PROBABILITY DISTRIBUTION ON WORDS IS ASYMPTOTICALLY NORMAL

RODNEY CANFIELD, SVANTE JANSON, AND DORON ZEILBERGER

ABSTRACT. We study XXX

[ Abstract! Correct date! No showkeys! ]

XXX

S:intro

## 1. INTRODUCTION

The most important discrete probability distribution, by far, is the *Binomial distribution*,  $B(n, p)$  for which we know everything *explicitly*,  $\mathbb{P}(X = i) = \binom{n}{i} p^i (1-p)^{n-i}$ , the probability generating function  $((pt + (1-p))^n)$ , the moment generating function  $((pe^t + 1-p)^n)$ , etc. etc. Most importantly, it is *asymptotically normal*, which means that the normalized random variable

$$Z_n = \frac{X_n - np}{\sqrt{np(1-p)}}$$

tends to the standard Normal distribution  $N(0, 1)$ , as  $n \rightarrow \infty$ .

Another important discrete distribution function is the *Mahonian* distribution, defined on the set of *permutations* on  $n$  objects, and describing, inter-alia, the random variable “number of inversions”. (Recall that an inversion in a permutation  $\pi_1, \dots, \pi_n$  is a pair  $1 \leq i < j \leq n$  such that  $\pi_i > \pi_j$ ). Let us call this random variable  $M_n$ . The probability generating function, due to Netto, is given *explicitly* by:

$$F_n(q) = \frac{1}{n!} \prod_{i=1}^n \frac{1-q^i}{1-q} . \tag{1.1} \quad \text{netto}$$

The formula (1.1) has a simple probabilistic interpretation (see Feller’s account in [3, Section X.6]): If  $Y_j$  is the number of  $i$  with  $1 \leq i < j$  and  $\pi_i > \pi_j$ , then

$$M_n = Y_1 + \dots + Y_n, \tag{1.2} \quad \text{nettoy}$$

and  $Y_1, \dots, Y_n$  are independent random variables and  $Y_j$  is uniformly distributed on  $\{0, \dots, j-1\}$ , as is easily seen by constructing  $\pi$  by inserting

---

*Date:* July 4, 2009 (typeset July 4, 2009 9.24).

Accompanied by Maple package MahonianStat available from <http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/mahon.html>. The work of D. Zeilberger was supported in part by the United States of America National Science Foundation.

$1, \dots, n$  in this order at random positions; thus  $Y_j$  has probability generating function  $(1 - q^j)/(j(1 - q))$ . It follows from (1.1) or (1.2) by simple calculations that the Mahonian distribution has mean and variance

$$\mathbb{E} M_n = \frac{n(n-1)}{4}, \quad (1.3) \quad \boxed{\text{em}}$$

$$\text{Var} M_n = \frac{n(n-1)(2n+5)}{72} = \frac{2n^3 + 3n^2 - 5n}{72}. \quad (1.4) \quad \boxed{\text{varm}}$$

Even though there is no explicit expression for the coefficients themselves (i.e. for the exact probability that a permutation of  $n$  objects would have a certain number of inversions), it is a classical result (see [3, Section X.6]), that follows from an extended form of the Central Limit Theorem, that the normalized version

$$\frac{M_n - n(n-1)/4}{\sqrt{(2n^3 + 3n^2 - 5n)/72}},$$

tends to  $N(0, 1)$ , as  $n \rightarrow \infty$ . So this sequence of probability distributions, too, is asymptotically normal.

But what about *words*, also known as *multi-set permutations*? Permutations on  $n$  objects can be viewed as words in the alphabet  $\{1, 2, \dots, n\}$ , where each letter shows up *exactly* once. But what if we allow *repetitions*? I.e., we consider all words with  $a_1$  occurrences of 1,  $a_2$  occurrences of 2,  $\dots$ ,  $a_m$  occurrences of  $m$ . (We assume throughout that  $m \geq 2$  and each  $a_j \geq 1$ .) We all know that the number of such words is the multinomial coefficient

$$\binom{a_1 + \dots + a_m}{a_1, \dots, a_m}$$

and many of us also know that the number of such words with exactly  $k$  inversions is the coefficient of  $q^k$  in the  $q$ -analog of the multinomial coefficient

$$\binom{a_1 + \dots + a_m}{a_1, \dots, a_m}_q := \frac{[a_1 + \dots + a_m]!}{[a_1]! \dots [a_m]!}, \quad (1.5) \quad \boxed{\text{qmulti}}$$

where  $[n]! := [1][2] \dots [n]$ , and  $[n] := (1 - q^n)/(1 - q)$ . Assuming that all words are equally likely (the uniform distribution), the probability generating function is

$$F_{a_1, \dots, a_m}(q) := \frac{(\prod_{i=1}^m a_i!) \cdot \prod_{i=1}^{a_1 + \dots + a_m} (1 - q^i)}{(a_1 + \dots + a_m)! \prod_{j=1}^m \prod_{i=1}^{a_j} (1 - q^i)} = \frac{F_{a_1 + \dots + a_m}(q)}{F_{a_1}(q) \dots F_{a_m}(q)}. \quad (1.6) \quad \boxed{\text{F}}$$

Indeed, this can be seen as follows. Let  $M_{a_1, \dots, a_m}$  denote the number of inversions in a random word. If we distinguish the  $a_i$  occurrences of  $i$  by adding different fractional parts, in random order, the number of inversions will increase by  $Z_i$ , say, with the same distribution as  $M_{a_i}$ ; further  $M_{a_1, \dots, a_m}$  and  $Z_1, \dots, Z_m$  are independent. On the other hand,  $M_{a_1, \dots, a_m} + Z_1 + \dots + Z_m$  has the same distribution as  $M_{a_1 + \dots + a_m}$ . Hence,

$$F_{a_1, \dots, a_m}(q) F_{a_1}(q) \dots F_{a_m}(q) = F_{a_1 + \dots + a_m}(q), \quad (1.7) \quad \boxed{\text{FF}}$$

which is (1.6).

By (1.6), we further have the factorization

$$F_{a_1, \dots, a_m}(q) = \prod_{j=2}^m F_{A_{j-1}, a_j}(q), \quad (1.8) \quad \boxed{\text{F2}}$$

where  $A_j := a_1 + \dots + a_j$ , which reduces the general case to the two-letter case.

Note that (1.6) shows that the distribution of  $M_{a_1, \dots, a_m}$  is invariant if we permute  $a_1, \dots, a_m$ ; a symmetry which is not obvious from the definition.

Rpart

**Remark 1.1.** The two-letter case is particularly interesting, since the unnormalized generating function

$$\binom{a+b}{a} F_{a,b}(q) = \frac{(1-q^{a+b})(1-q^{a+b-1}) \dots (1-q^{a+1})}{(1-q^b)(1-q^{b-1}) \dots (1-q^1)} = \frac{[a+b]!}{[a]! [b]!},$$

(the  $q$ -binomial coefficient in (1.5)) is the same as the generating function for the set of integer-partitions with largest part  $\leq a$  and  $\leq b$  parts, in other words the set of integer-partitions whose Ferrers diagram lies inside an  $a$  by  $b$  rectangle, where the random variable is the “number of dots” (i.e. the integer being partitioned). In other words, the number of such partitions of an integer  $n$  equal the number of words of  $a$  1’s and  $b$  2’s with  $n$  inversions.

It is easy to see that the *mean* of  $M_{a_1, \dots, a_m}$  is

$$\mu(a_1, \dots, a_m) := \mathbb{E} M_{a_1, \dots, a_m} = e_2(a_1, \dots, a_m)/2$$

(here  $e_k(a_1, \dots, a_m)$  is the degree  $k$  elementary symmetric function), so considering the shifted random variable  $M_{a_1, \dots, a_m} - \mu(a_1, \dots, a_m)$ , “number of inversions minus the mean”, we get that the probability generating function is

$$G_{a_1, \dots, a_m}(q) := q^{-\mu(a_1, \dots, a_m)} F_{a_1, \dots, a_m}(q) = \frac{F_{a_1, \dots, a_m}(q)}{q^{e_2(a_1, \dots, a_k)/2}} \quad (1.9) \quad \boxed{\text{G}}$$

By computing  $(q(qG)')'$  and plugging-in  $q = 1$ , or from (1.7) and (1.3)–(1.4), it is easy to see that the *variance*  $\sigma^2 := \text{Var} M_{a_1, \dots, a_m}$  is

$$\sigma^2 = \frac{(e_1 + 1)e_2 - e_3}{12}. \quad (1.10) \quad \boxed{\text{sigma}}$$

(By  $\sigma$  we mean  $\sigma(a_1, \dots, a_m)$  and we omit the arguments  $(a_1, \dots, a_m)$  from the  $e_i$ ’s.)

Let  $N := e_1 = a_1 + \dots + a_m$ , the length of the random word, and let  $a^* := \max_j a_j$  and  $N_* := N - a^*$ .

One main result of the present article is:

T1

**Theorem 1.2.** Consider the random variable,  $M_{a_1, \dots, a_m}$ , “number of inversions”, on the (uniform) sample space of words with  $a_1$  1’s,  $a_2$  2’s,  $\dots$ ,  $a_m$   $m$ ’s. For any sequence of sequences  $(a_1, \dots, a_m) = (a_1^{(\nu)}, \dots, a_m^{(\nu)})$  such

that  $N_* := N - a^* \rightarrow \infty$ , the sequence of normalized random variables

$$X_{a_1, \dots, a_m} = \frac{M_{a_1, \dots, a_m} - \mu(a_1, \dots, a_m)}{\sigma(a_1, \dots, a_m)} ,$$

tends to the standard normal distribution  $\mathcal{N}(0, 1)$ , as  $\nu \rightarrow \infty$ .

Theorem 1.2 includes both the case when  $m \geq 2$  is fixed, and the case when  $m \rightarrow \infty$ . If  $m$  is fixed and  $a_1 \geq a_2 \geq \dots \geq a_m$ , as may be assumed by symmetry, then the condition  $N_* \rightarrow \infty$  is equivalent to  $a_2 \rightarrow \infty$ . In the case  $m \rightarrow \infty$ , the assumption  $N_* \rightarrow \infty$  is redundant, because  $N_* \geq m - 1$ .

**Remark 1.3.** The condition  $N_* \rightarrow \infty$  is also necessary for asymptotic normality, see Section 6.

We give a short proof of this result using characteristic functions in Section 3. We give first in Section 2 another proof (at least of a special case) that is *computer-assisted*, using the Maple package `MahonianStat` available from the webpage of this article:

<http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/mahon.html>, where one can also find sample input and output. This first proof uses the *method of moments*.

We conjecture that Theorem 1.2 can be refined to a local limit theorem as follows:

Conjlocal

**Conjecture 1.4.** *Uniformly for all  $a_1, \dots, a_m$  and all integers  $k$ ,*

$$\mathbb{P}(M_{a_1, \dots, a_m} = k) = \frac{1}{\sqrt{2\pi\sigma}} \left( e^{-(k-\mu)^2/(2\sigma^2)} + O\left(\frac{1}{N_*}\right) \right). \quad (1.11) \quad \text{local}$$

We have not been able to prove this conjecture in full generality, but we prove it under additional hypotheses on  $a_1, \dots, a_m$  in Section 4.

Spf1

## 2. A COMPUTER-INSPIRED PROOF

We assume for simplicity that  $m$  is fixed, and that  $(a_1, \dots, a_m) = (ta_1^0, \dots, ta_m^0)$  for some fixed  $a_1^0, \dots, a_m^0$  and  $t \rightarrow \infty$ .

We discover and prove the leading term in the asymptotic expansion, in  $t$ , for an *arbitrary*  $2r$ -th moment, for the normalized random variable  $X_{a_1, \dots, a_m} = (M_{a_1, \dots, a_m} - \mu)/\sigma$ , and show that it converges to the moment  $\mu_{2r} = (2r)!/(2^r r!)$  of  $\mathcal{N}(0, 1)$ , for every  $r$ .

For the sake of exposition, we will only treat in detail the two-letter case, where we can find *explicit* expressions for the asymptotics of the  $2r$ -th moment, for symbolic  $a_1, a_2, t$  and  $r$  to *any* desired (specific) order  $s$  (i.e. the leading coefficient  $t^{3r}$  as well as the terms involving  $t^{3r-1}, \dots, t^{3r-s}$ ). A modified argument works for the general case, but we can only find the leading term, i.e. that

$$\alpha_{2r} := \mathbb{E}(X_{a_1, \dots, a_m})^{2r} = \frac{(2r)!}{2^r r!} t^{3r} + O(t^{3r-1}) .$$

Of course the odd moments are all zero, since the distribution of  $M_{a_1, \dots, a_m}$  is symmetric about  $\mu$ .

The mean in the two-letter case is simply  $ab/2$ , so the probability generating function for  $M_{a,b} - \mu$  is

$$G_{a,b}(q) = \frac{F_{a,b}(q)}{q^{ab/2}(a+b)!/(a!b!)} = \frac{a!b!(1-q^{a+b})(1-q^{a+b-1})\cdots(1-q^{a+1})}{q^{ab/2}(a+b)!(1-q^b)(1-q^{b-1})\cdots(1-q^1)} ,$$

Taking ratios, we have:

$$\frac{G_{a,b}(q)}{G_{a-1,b}(q)} = \frac{a(1-q^{a+b})}{q^{b/2}(a+b)(1-q^a)} . \quad (2.1) \quad \boxed{\text{ratio}}$$

Recall that the binomial moments  $A_r := \mathbb{E}[\binom{X}{r}]$  are the Taylor coefficients of the probability generating function (in our case  $G_{a,b}(q)$ ) around  $q = 1$ . Writing  $q = 1 + z$ , we have

$$G_{a,b}(1+z) = \sum_{r=0}^{\infty} A_r(a,b)z^r$$

Note that  $A_0(a,b) = 1$  and  $A_1(a,b) = 0$ . Let us call the expression on the right side of (2.1), with  $q$  replaced by  $1+z$ ,  $P(a,b,z)$ :

$$P(a,b,z) := \frac{a(1-(1+z)^{a+b})}{(1+z)^{b/2}(a+b)(1-(1+z)^a)}$$

Maple can easily expand  $P(a,b,z)$  to any desired power of  $z$ . It starts out with

$$\begin{aligned} P(a,b,z) = & 1 + \frac{1}{24}(2a+b)bz^2 - \frac{1}{24}(2a+b)bz^3 \\ & - \frac{1}{5760}(8a^3 - 8a^2b - 12ab^2 - 3b^3 - 440a - 220b)bz^4 + \dots \end{aligned}$$

note that the coefficients of all the powers of  $z$  are polynomials in  $(a,b)$ .

So let us write

$$P(a,b,z) = \sum_{i=0}^{\infty} p_i(a,b)z^i ,$$

where  $p_i(a,b)$  are certain polynomials that Maple can compute for any  $i$ , no matter how big.

Looking at the recurrence

$$G_{a,b}(1+z) = P(a,b,z)G_{a-1,b}(1+z) ,$$

and comparing coefficients of  $z^r$  on both sides, we get

$$A_r(a,b) - A_r(a-1,b) = \sum_{s=1}^r A_{r-s}(a-1,b)p_s(a,b) . \quad (2.2) \quad \boxed{\text{recurrence}}$$

Assuming that we already know the polynomials  $A_{r-1}(a,b), A_{r-2}(a,b), \dots, A_0(a,b)$ , the left side is a certain specific polynomial in  $a$  and  $b$ , that Maple can easily compute, and then  $A_r(a,b)$  is simply the indefinite sum of that polynomial, that Maple can do just as easily. So (2.2) enables us to get *explicit* expressions for the binomial moments  $A_r(a,b)$  for *any* (numeric)  $r$ .

But what about the general (symbolic)  $r$ ? It is too much to hope for the full expression, **but** we can easily conjecture as many leading terms as we wish. We first conjecture, and then immediately prove by induction, that for  $r \geq 1$

$$A_{2r}(a, b) = \frac{1}{r!} \left( \frac{ab(a+b)}{24} \right)^r + \text{lower order terms}$$

$$A_{2r+1}(a, b) = \frac{-1}{(r-1)!} \left( \frac{ab(a+b)}{24} \right)^r + \text{lower order terms} \quad ,$$

where we can conjecture (by fitting polynomials in  $(a, b)$  to the data obtained from the numerical  $r$ 's) any (finite, specific) number of terms.

Once we have asymptotics, to any desired order, for the binomial moments, we can easily compute the moments  $\mu_r(a, b)$  of  $M_{a,b} - \mu$  themselves, for *any* desired specific  $r$  and asymptotically, to any desired order. We do that by using the expressions of the powers as linear combination of falling-factorials (or equivalently binomials) in terms of Stirling numbers of the second kind,  $S(n, k)$ . Note that for the asymptotic expressions to any desired order, we can still do it symbolically, since for any specific  $m$ ,  $S(n, n-m)$  is a polynomial in  $n$  (that Maple can easily compute, symbolically, as a polynomial in  $n$ ). In particular, the variance is:

$$\sigma^2 = \mu_2(a, b) = \frac{ab(a+b+1)}{12} \quad ,$$

in accordance with (1.10). In general we have  $\mu_{2r+1}(a, b) = 0$ , of course, and the six leading terms of  $\mu_{2r}(at, bt)$  can be found in the webpage of this article. From this, Maple finds that  $\alpha_{2r}(at, bt) := \mu_{2r}(at, bt) / \mu_2(at, bt)^r$  are given asymptotically (for fixed  $a, b$  and  $t \rightarrow \infty$ ) by:

$$\alpha_{2r}(at, bt) = \frac{(2r)!}{2^r r!} \cdot \left( 1 - \frac{r(r-1)(b^2 + ab + a^2)}{5ab(a+b)} \cdot \frac{1}{t} \right) + O(t^{-2}) \quad .$$

In particular, as  $t \rightarrow \infty$ , they converge to the famous moments of  $\mathcal{N}(0, 1)$ . QED.

**2.1. The general case.** To merely prove asymptotic normality, one does not need a computer, since we only need the leading terms. The above proof can be easily adapted to the general case  $(a_1, \dots, a_m) = (ta_1^0, \dots, ta_m^0)$ . One simply uses induction on  $m$ , the number of different letters.

**2.2. The Maple package MahonianStat.** The Maple package `MahonianStat`, accompanying this article, has lots of features, that the readers can explore at their leisure. Once downloaded into a directory, one goes into a maple session, and types `read MahonianStat`. To get a list of the main procedures, type: `ezra()`; . To get help with a specific procedure, type `ezra(ProcedureName)`; . Let us just mention some of the more important procedures.

**AsyAlphaW2tS(r, a, b, t, s)**: inputs symbols  $\mathbf{r}, \mathbf{a}, \mathbf{b}, \mathbf{t}$  and a positive integer  $s$ , and outputs the asymptotic expansion, to order  $s$ , for  $\alpha_{2r} (= \mu_{2r}/\mu_2^r)$

**ithMomWktE(r, e, t)** the  $r$ -th moment about the mean of the number of inversions of  $a_1 t$  1's,  $\dots$ ,  $a_m t$   $n$ 's in terms of the elementary symmetric functions, in  $a_1, \dots, a_m$ . Here  $r$  is a specific (numeric) positive integer, but  $e$  and  $t$  are symbolic.

**AppxWk(L, x)**: Using the asymptotics implied by the asymptotic normality of the (normalized) random variable under consideration, finds an approximate value for the number of words with  $L[1]$  1's,  $L[2]$  2's,  $\dots$ ,  $L[n]$   $n$ 's with exactly  $x$  inversions. For example, try: **AppxWk**( [100, 100, 100] , 15000 );

For the two-lettered case, one can get better approximations, by procedure **BetterAppxW2**, that uses improved limit-distributions, using more terms in the probability density function.

The webpage of this article has some sample input and output.

### 3. A GENERAL PROOF OF THEOREM 1.2

Spf2

We have an exact formula (1.10) for the variance  $\sigma^2$  of  $M_{a_1, \dots, a_m}$ . We first show that  $\sigma^2$  is always of the order  $\Theta(N^2 N_*)$ .

Lsigma

**Lemma 3.1.** *For any  $a_1, \dots, a_m$ ,*

$$\frac{N^2 N_*}{36} \leq \sigma^2 \leq \frac{(N+1)NN_*}{12} \leq \frac{N^2 N_*}{6}.$$

*Proof.* For the upper bounds we assume, by symmetry, that  $a_1 \geq \dots \geq a_m$ . Then  $a^* = a_1$  and

$$e_2 = a_1 \sum_{j=2}^m a_j + a_2 \sum_{j=3}^m a_j + \dots \leq N \sum_{j=2}^m a_j = NN_*.$$

Since  $e_1 = N$ , (1.10) yields the upper bounds.

For the lower bound, we first observe that  $2e_2 e_1 - 6e_3 \geq 0$  (since this difference can be written as a sum of certain  $a_j a_k a_l$ ). Hence  $e_3 \leq e_1 e_2 / 3$  and (1.10) yields

$$12\sigma^2 \geq e_1 e_2 - e_3 \geq \frac{2}{3} e_1 e_2.$$

Further,

$$2e_2 = \sum_{j=1}^m a_j (N - a_j) \geq \sum_{j=1}^m a_j (N - a^*) = NN_*,$$

and the lower bound follows.  $\square$

*Proof of Theorem 1.2.* From (1.6) follows the identity

$$F_{n_1, n_2}(e^{i\theta}) = \prod_{j=1}^{n_2} \frac{(e^{i(n_1+j)\theta} - 1)/(i(n_1+j)\theta)}{(e^{ij\theta} - 1)/(ij\theta)}. \quad (3.1) \quad \text{f2q}$$

By Taylor's series

$$\log \frac{e^z - 1}{z} = z/2 + z^2/24 + O(z^4), \quad |z| \leq 1,$$

and we substitute this expansion into the identity (3.1) to conclude:

$$F_{n_1, n_2}(e^{i\theta}) = \exp\left(in_1 n_2 \theta / 2 - n_1 n_2 (n_1 + n_2 + 1) \theta^2 / 24 + O(n_2 n_1^4 \theta^4)\right), \quad (3.2) \quad \boxed{\text{r3}}$$

uniformly for  $n_1 \geq n_2 \geq 1$  and  $|\theta| \leq (n_1 + n_2)^{-1}$ .

We use the factorization (1.8). By symmetry, we may assume  $a_1 \geq a_2 \geq \dots \geq a_m$ , and then  $A_{j-1} \geq a_{j-1} \geq a_j$  for each  $j$ . Thus (3.2) yields, uniformly for  $q = e^{i\theta}$  with  $|\theta| \leq N^{-1}$ ,

$$\begin{aligned} F_{a_1, \dots, a_m}(q) &= \prod_{j=2}^m F_{A_{j-1}, a_j}(q) \\ &= \exp\left(\sum_{j=2}^m \left(i A_{j-1} a_j \theta / 2 - A_{j-1} a_j (A_j + 1) \theta^2 / 24 + O(a_j A_{j-1}^4 \theta^4)\right)\right). \end{aligned}$$

Here, the sums of the coefficients of  $\theta$  and  $\theta^2$  are easily evaluated, but we do not have to do that since they have to equal  $i\mu$  and  $-\sigma^2/2$ , respectively. Further,

$$\sum_{j=2}^m A_{j-1}^4 a_j \leq N^4 \sum_{j=2}^m a_j = N^4 N_*. \quad (3.3) \quad \boxed{\text{pa1}}$$

Consequently, if  $|\theta| \leq N^{-1}$ ,

$$F_{a_1, \dots, a_m}(e^{i\theta}) = \exp(i\mu\theta - \sigma^2\theta^2/2 + O(N^4 N_* \theta^4)) \quad (3.4) \quad \boxed{\text{pa2}}$$

and, by (1.9),

$$G_{a_1, \dots, a_m}(e^{i\theta}) = \exp(-\sigma^2\theta^2/2 + O(N^4 N_* \theta^4)). \quad (3.5) \quad \boxed{\text{pag}}$$

Let  $\theta = t/\sigma$ . For any fixed  $t$ , by Lemma 3.1,

$$|Nt/\sigma| = O(N_*^{-1/2}) = o(1),$$

so  $|\theta| \leq N^{-1}$  if  $\nu$  is large enough. Hence, by (3.3)–(3.5) and Lemma 3.1,

$$G_{a_1, \dots, a_m}(e^{it/\sigma}) = \exp\left(-\frac{t^2}{2} + O\left(\frac{N^4 N_* t^4}{N^4 N_*^2}\right)\right) = \exp(-t^2/2 + o(1)),$$

and Theorem 1.2 follows by the continuity theorem [4, Theorem XV.3.2].  $\square$

#### Slocal

#### 4. THE LOCAL LIMIT THEOREM

“If one can prove a central limit theorem for a sequence  $a_n(k)$  of numbers arising in enumeration, then one has a qualitative feel for their behavior. A local limit theorem is better because it provides asymptotic information about  $a_n(k) \dots$ ” [2]. In this section we prove that the relation (1.11) holds uniformly over certain very general, albeit not unrestricted, sets of tuples  $\mathbf{a} = (a_1, \dots, a_m)$ . The exact statement is given below in Theorem 4.5.

As explained in Bender [2], there are two standard conditions for passage from a central to a local limit theorem: (1) if the sequence in question is

unimodal, then one has a local limit theorem for  $n$  in the set  $\{|n - \mu| \geq \epsilon\sigma\}$ ,  $\epsilon > 0$ ; (2) if the sequence in question is log-concave, then one has a local limit theorem for all  $n$ . Our sequence, the coefficients of the  $q$ -mutinomial, is in fact unimodal, as first shown by Schur [6] using invariant theory, and later by O'Hara [5] using combinatorics. Unfortunately, the ensuing local limit theorem fails to cover the most interesting coefficients, the largest ones, near the mean  $\mu$ . However, our polynomials are manifestly not log-concave as is seen by inspecting the first three coefficients (assuming  $n_1, n_2 \geq 2$ )

$$\binom{n_1 + n_2}{n_1}_q = 1 + q + 2q^2 + \cdots .$$

The question arises might the coefficients be log-concave near the mean, and here is a small table of empirical values: ( $c[j] = [q^j] \binom{2n}{n}_q$ )

$n$	$(c[n^2/2 - 1])^2 - c[n^2/2] \times c[n^2/2 - 2]$
2	-1
4	-7
6	-165
8	-1529
10	44160
12	7715737
14	905559058
16	101507214165
18	11955335854893
20	1501943866215277

Based on this scant evidence, we speculate that some sort of log-concavity theorem is true, but that its proper statement is complicated by describing the appropriate range of  $\mathbf{a}$  and  $j$ . Thus, we use neither of the two standard methods mentioned above for proving our local limit theorem. (Later, we shall see that our theorem has implications for log-concavity.) Instead, we use another standard method, direct integration (Fourier inversion) of the characteristic function, or equivalently of the probability generating function  $F(q)$  for  $q = e^{i\theta}$  on the unit circle. This method uses the estimates for small  $\theta$  derived in Section 3 together with estimates of  $|F(e^{i\theta})|$  for larger  $\theta$ . We begin with one such estimate for rather small  $\theta$ .

**LG1** **Lemma 4.1.** *There exists a constant  $\tau > 0$  such that for any  $a_1, \dots, a_m$  and  $|\theta| \leq \tau/N$ ,*

$$|F_{a_1, \dots, a_m}(e^{i\theta})| = |G_{a_1, \dots, a_m}(e^{i\theta})| \leq e^{-\sigma^2\theta^2/4}.$$

*Proof.* Suppose that  $0 < |\theta| \leq \tau/N$ . Then, using Lemma 3.1,

$$\frac{N^4 N_* \theta^4}{\sigma^2 \theta^2} \leq \frac{N^2 N_* \tau^2}{\sigma^2} \leq 36\tau^2,$$

so if  $\tau$  is chosen small enough, the error term  $O(N^4 N_* \theta^4)$  in (3.4) and (3.5) is  $\leq \sigma^2 \theta^2/4$ , and thus the result follows from (3.5).  $\square$

We let in the sequel  $\tau$  denote this constant. We may assume  $0 < \tau \leq 1$ .

**Lmagnus**

**Lemma 4.2.** *Uniformly, for all  $a_1, \dots, a_m$  and all integers  $k$ ,*

$$\mathbb{P}(M_{a_1, \dots, a_m} = k) = \frac{1}{\sqrt{2\pi\sigma}} e^{-(k-\mu)^2/(2\sigma^2)} + \int_{\tau/N}^{\pi} |F_{a_1, \dots, a_m}(e^{i\theta})| d\theta + O\left(\frac{1}{\sigma N_*}\right).$$

*Proof.* For any integer  $k$ ,

$$\begin{aligned} \mathbb{P}(M_{a_1, \dots, a_m} = k) &= \frac{1}{\sqrt{2\pi\sigma}} e^{-(k-\mu)^2/(2\sigma^2)} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} F_{a_1, \dots, a_m}(e^{i\theta}) e^{-ik\theta} d\theta - \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\sigma^2\theta^2/2} e^{-i(k-\mu)\theta} d\theta \\ &= \frac{1}{2\pi} \int_{|\theta| \leq \tau/N} \left( G_{a_1, \dots, a_m}(e^{i\theta}) - e^{\sigma^2\theta^2/2} \right) e^{-i(k-\mu)\theta} d\theta \\ &\quad + \frac{1}{2\pi} \int_{\tau/N \leq |\theta| \leq \pi} F_{a_1, \dots, a_m}(e^{i\theta}) e^{-ik\theta} d\theta \\ &\quad - \frac{1}{2\pi} \int_{|\theta| \geq \tau/N} e^{-\sigma^2\theta^2/2} e^{-i(k-\mu)\theta} d\theta \\ &=: I_1 + I_2 + I_3. \end{aligned}$$

For  $|\theta| \leq \tau/N$ ,  $|G_{a_1, \dots, a_m}(e^{i\theta})| \leq e^{-\sigma^2\theta^2/4}$  by Lemma 4.1, so (3.5) and the mean-value theorem applied to  $z \mapsto e^z$  yield

$$\begin{aligned} |G_{a_1, \dots, a_m}(e^{i\theta}) - e^{-\sigma^2\theta^2/2}| &\leq O(N^4 N_* \theta^4) \max(|G_{a_1, \dots, a_m}(e^{i\theta})|, e^{-\sigma^2\theta^2/2}) \\ &= O(N^4 N_* \theta^4 e^{-\sigma^2\theta^2/4}). \end{aligned}$$

Integrating, we find, using Lemma 4.1,

$$\begin{aligned} |I_1| &\leq \int_{|\theta| \leq \tau/N} |G_{a_1, \dots, a_m}(e^{i\theta}) - e^{-\sigma^2\theta^2/2}| d\theta \\ &\leq O(N^4 N_*) \int_{-\infty}^{\infty} \theta^4 e^{-\sigma^2\theta^2/4} d\theta = O(N^4 N_* \sigma^{-5}) = O\left(\frac{1}{\sigma N_*}\right). \end{aligned}$$

Further, again using Lemma 4.1,

$$|I_3| \leq \int_{\tau/N}^{\infty} e^{-\sigma^2\theta^2/2} d\theta \leq 3\sigma^{-1} e^{-(\sigma\tau/N)^2/2} \leq \frac{6}{\sigma(\sigma\tau/N)^2} = O\left(\frac{1}{\sigma N_*}\right).$$

Finally,  $|I_2| \leq \int_{\tau/N}^{\pi} |F_{a_1, \dots, a_m}(e^{i\theta})| d\theta$ . □

In order to verify Conjecture 1.4, it thus suffices to show that the integral  $\int_{\tau/N}^{\pi} |F_{a_1, \dots, a_m}(e^{i\theta})| d\theta$  in Lemma 4.2 is  $O\left(\frac{1}{\sigma N_*}\right)$ .

**Remark 4.3.** For example, an estimate

$$F_{a_1, \dots, a_m}(e^{i\theta}) = O\left(\frac{1}{\sigma^3 \theta^3}\right), \quad 0 < \theta \leq \pi, \quad (4.1)$$

**conj3**

is sufficient for (1.11). We conjecture that this estimate (4.1) holds when  $N_* \geq 6$ , say. Note that it does not hold for very small  $N_*$ : taking  $\theta = \pi$  we

have, for even  $n_1$ ,  $F_{n_1,1}(-1) = 1/(n_1 + 1) = 1/N$ , and the same holds for  $F_{n_1,2}(-1)$ .

Note further that even the weaker estimate

$$F_{a_1, \dots, a_m}(e^{i\theta}) = O\left(\frac{1}{\sigma^2 \theta^2}\right), \quad 0 < \theta \leq \pi, \quad (4.2) \quad \boxed{\text{conj2}}$$

would be enough to prove (1.11) with the weaker error term  $O(N_*^{-1/2})$ .

We obtain a partial proof of Conjecture 1.4 using the following lemma.

**LC1** **Lemma 4.4.** *For a given  $\tau \in (0, 1]$  there exists  $c = c(\tau) > 0$  such that*

$$|F_{n_1, n_2}(e^{i\theta})| \leq e^{-cn_2} \quad (4.3) \quad \boxed{\text{r2}}$$

for  $n_1 \geq n_2 \geq 1$  and  $\tau/(n_1 + n_2) \leq |\theta| \leq \pi$ .

More generally, for any  $a_1, \dots, a_m$  and  $\tau/N \leq |\theta| \leq \pi$ ,

$$|F_{a_1, \dots, a_m}(e^{i\theta})| \leq e^{-cN_*}. \quad (4.4) \quad \boxed{\text{r2m}}$$

*Proof.* We prove first (4.3). For positive integer  $n$  define

$$f_n(y, q) = \prod_{j=0}^{n-1} (1 - yq^j)^{-1}.$$

For  $0 \leq R < 1$ , we have (e.g. by Taylor expansions)  $e^{2R} \leq \frac{1+R}{1-R}$ , and thus  $e^{4R} \leq \frac{(1+R)^2}{(1-R)^2} = 1 + \frac{4R}{(1-R)^2}$ . Hence, by convexity, for any real  $\zeta$ ,

$$e^{2R(1-\cos\zeta)} \leq 1 + \frac{2R(1-\cos\zeta)}{(1-R)^2} = \frac{1+R^2-2R\cos\zeta}{(1-R)^2} = \frac{|1-Re^{i\zeta}|^2}{(1-R)^2},$$

and thus

$$\left| (1 - Re^{i\zeta})^{-1} \right| \leq (1-R)^{-1} \exp(R(1-\cos\zeta)).$$

Consequently, by a simple trigonometric identity, for any real  $\phi$  and  $\theta$ ,

$$\begin{aligned} \left| f_{n_1}(Re^{i\phi}, e^{i\theta}) \right| &\leq (1-R)^{-n_1-1} \\ &\quad \times \exp\left(-R\left(n_1+1-\cos\left(\phi+\frac{n_1\theta}{2}\right)\frac{\sin(n_1+1)\theta/2}{\sin\theta/2}\right)\right) \\ &\leq (1-R)^{-n_1-1} \times \exp\left(R\left(-n_1-1+\frac{\sin(n_1+1)\theta/2}{\sin\theta/2}\right)\right). \end{aligned}$$

The function  $g(\theta) = g_n(\theta) := \frac{\sin n(\theta/2)}{\sin(\theta/2)}$ , where  $n \geq 1$ , is an even function of  $\theta$ ; is decreasing for  $0 \leq \theta \leq \pi/n$ , as can be verified by calculating  $g'$ ; and satisfies  $|g(\theta)| \leq g(\pi/n)$  for  $\pi/n \leq |\theta| \leq \pi$ . Further, for  $n \geq 2$  and  $0 \leq |\theta| \leq \pi/n$ ,

$$\begin{aligned} g_n(\theta) &= 2 \frac{\sin(n\theta/4)}{\sin(\theta/2)} \cos(n\theta/4) = 2g_{n/2}(\theta) \cos(n\theta/4) \leq n \cos(n\theta/4) \\ &\leq n \left(1 - \frac{n^2\theta^2}{40}\right). \end{aligned}$$

Let  $\theta_0 = \tau(n_1 + n_2)^{-1} < \pi/(n_1 + 1)$ . For  $\theta_0 \leq |\theta| \leq \pi$  we thus have

$$|g_{n_1+1}(\theta)| \leq g_{n_1+1}(\theta_0) \leq n_1 + 1 - \frac{n_1^3 \theta_0^2}{40};$$

whence, for  $0 \leq R < 1$ , the estimate above yields

$$\left| f_{n_1}(Re^{i\phi}, e^{i\theta}) \right| \leq (1 - R)^{-n_1-1} \exp(-Rn_1^3 \theta_0^2/40). \quad (4.5) \quad \boxed{\text{tho}}$$

Combinatorially we know that  $[y^\ell q^n] f_{n_1}(y, q)$  is the number of partitions of  $n$  having at most  $\ell$  parts no one of which exceeds  $n_1$ . As said in Remark 1.1, this equals  $[q^n] \binom{n_1+\ell}{n_1} F_{n_1+\ell}(q)$ . Hence, using Cauchy's integral formula, for any  $R > 0$ ,

$$\binom{n_1 + n_2}{n_1} F_{n_1, n_2}(q) = [y^{n_2}] f_{n_1}(y, q) = \frac{1}{2\pi i} \int_{|y|=R} f_{n_1}(y, q) \frac{dy}{y^{n_2+1}}.$$

Consequently, (4.5) implies that for  $\theta_0 \leq |\theta| \leq \pi$  and  $0 < R < 1$ ,

$$\binom{n_1 + n_2}{n_1} |F_{n_1, n_2}(q)| \leq (1 - R)^{-n_1-1} R^{-n_2} \exp(-Rn_1^3 \theta_0^2/40).$$

Now choose  $R = \rho := n_2/(n_1 + n_2) \leq 1/2$ . By Stirling's formula,

$$\binom{n_1 + n_2}{n_1} = \Omega(n_2^{-1/2}) (1 - \rho)^{-n_1-1} \rho^{-n_2}$$

and thus, for  $\theta_0 \leq |\theta| \leq \pi$ ,

$$|F_{n_1, n_2}(q)| \leq O(n_2^{1/2}) \exp(-\rho n_1^3 \theta_0^2/40) = O(n_2^{1/2}) \exp(-\Omega(n_2)).$$

This shows (4.3) for  $n_2$  large enough, say  $n_2 \geq n_0$ . The case  $n_2 \leq n_0$  can be  
XXX

XXX

I hope this case is simple, and perhaps can be dismissed without giving details. If not, I think that we do not really need it.

To prove (4.4), we assume as we may that  $a_1 \geq \dots \geq a_m$  and use the factorization (1.8). Let  $J$  be the first index such that  $a_2 + \dots + a_J \geq N_*/2$ . For  $j \geq J$ , then  $A_{j-1} + a_j = A_j \geq A_J \geq a_1 + N_*/2 \geq N/2$ , and thus  $A_k |\theta| \geq N |\theta|/2 \geq \tau/2$ ; hence (4.3) yields

$$|F_{A_{j-1}, a_j}(e^{i\theta})| \leq e^{-c(\tau/2)a_j}.$$

We thus obtain from (1.8), since each  $F_{n_1, n_2}$  is a probability generating function and thus is bounded by 1 on the unit circle,

$$|F_{a_1, \dots, a_m}(e^{i\theta})| = \prod_{j=2}^m |F_{A_{j-1}, a_j}(e^{i\theta})| \leq \prod_{j=J}^m e^{-c(\tau/2)a_j} \leq e^{-c(\tau/2)N_*/2},$$

because  $\sum_{j=J}^m a_j \geq N_*/2$ . This proves (4.4) (redefining  $c(\tau)$ ).  $\square$

**Tlocal**

**Theorem 4.5.** *There exists a positive constant  $c$  such that for every  $C$ , the following is true. Uniformly for all  $a_1, \dots, a_m$  such that  $a^* \leq Ce^{cN^*}$  and all integers  $k$ ,*

$$\mathbb{P}(M_{a_1, \dots, a_m} = k) = \frac{1}{\sqrt{2\pi\sigma}} \left( e^{-(k-\mu)^2/(2\sigma^2)} + O\left(\frac{1}{N^*}\right) \right). \quad (4.6) \quad \boxed{\text{localxxx}}$$

*Proof.* Let  $c_1 = c(\tau)$  be the constant in Lemma 4.4. Then, Lemmas 4.2 and 4.4 yield

$$\mathbb{P}(M_{a_1, \dots, a_m} = k) = \frac{1}{\sqrt{2\pi}\sigma} \left( e^{-(k-\mu)^2/(2\sigma^2)} + O\left(\frac{1}{N_*} + \sigma e^{-c_1 N_*}\right) \right).$$

For any fixed  $c < c_1$  we have, using Lemma 3.1,  $\sigma N_* e^{-c_1 N_*} = O(N e^{-c N_*})$  and thus

$$\mathbb{P}(M_{a_1, \dots, a_m} = k) = \frac{1}{\sqrt{2\pi}\sigma} \left( e^{-(k-\mu)^2/(2\sigma^2)} + O\left(\frac{1 + N e^{-c N_*}}{N_*}\right) \right).$$

The result follows, since  $N e^{-c N_*} = a^* e^{-c N_*} + N_* e^{-c N_*} = a^* e^{-c N_*} + O(1)$ .  $\square$

[Old stuff to be edited:]

XXX

The combinatorics of the Macmahon  $q$  statistic appear not to have been greatly used, but in fact play an essential, easily overlooked, role. It seems that the Taylor estimate (3.2) is the key for a central limit theorem. As in Bender [2], some kind of smoothness condition is needed to proceed to a local limit theorem (1). Lemma 1, the large  $\theta$  bound, imposes the required smoothness. The proof of Lemma 1 is made possible by the following combinatorial property:

$$[q^n] \binom{n_1 + n_2}{n_1}_q = \sum_{\ell=1}^{n_2} [y^\ell] \prod_{j=1}^{n_1} (1 - yq^j)^{-1}.$$

Thus, our Theorem 4.5 may be the  $h(x) = (1-x)^{-1}$  case of a more general theorem about

$$[q^n] \sum_{\ell=1}^{n_2} [y^\ell] \prod_{j=1}^{n_1} h(yq^j).$$

## 5. LOG-CONCAVITY

[Old stuff to be edited:]

XXX

Let us review the proof of Theorem 4.5 with the intention of greater accuracy. The goal is to prove log-concavity in some range. For concreteness, let  $\mathbf{a} = (n, n)$ . Then  $\sigma^2$  is of order  $n^3$ , and for sufficient accuracy we take the Taylor series (3) out to  $O(\theta^{10})$ . This improves (5) to:

$$\frac{1}{2\pi} \int_{\text{small } \theta} F(e^{i\theta}) e^{-ij\theta} d\theta = \frac{F(1)}{\sigma\sqrt{2\pi}} \left( e^{-x^2/2} \left( 1 + \frac{p_4(x)}{\sigma^4} + \frac{p_6(x)}{\sigma^6} + \frac{p_8(x)}{\sigma^8} \right) + O\left(\frac{1}{n^4}\right) \right).$$

Here,  $p_4(x), \dots$  are certain polynomials and  $x$  is determined by  $j = \mu + x\sigma$ . The sizes of the terms involving the polynomials are  $1, 1/n, 1/n^2, 1/n^3$ . The large  $\theta$  contribution is negligible by comparison. One may now use the formula to calculate the difference  $c_j^2 - c_{j-1}c_{j+1}$  to accuracy  $O(1/n^4)$ . Taking out the obvious common factors, it is found to equal  $1/\sigma^2$ . This gives:

**TC2** **Theorem 5.1** (A log-concavity result). *For each constant  $C$  we have  $n_0$  such that for  $n \geq n_0$*

$$c_j^2 \geq c_{j-1}c_{j+1},$$

for  $|j - \mu| \leq C\sigma$ . Here,

$$c_j = [q^j] \binom{2n}{n}_q.$$

We note that the “mysterious” numbers appearing in our earlier table are asymptotically

$$\frac{1}{2\pi\sigma^4} \binom{2n}{n}^2.$$

## 6. FINAL COMMENTS

**Sfine**

Suppose that  $N_* \not\rightarrow \infty$ . We may, as usual, assume that  $a_1 \geq \dots \geq a_m$ . By considering a subsequence (if necessary), we may assume that  $N_* := N - a^* = a_2 + \dots + a_m$  is a constant; this entails that  $m$  is bounded, so by again considering a subsequence, we may assume that  $m$  and  $a_2, \dots, a_m$  are constant. We thus study the case when  $a_1 \rightarrow \infty$  with fixed  $a_2, \dots, a_m$ .

In this case, the number of inversions between indices  $2, \dots, m$  is  $O(1)$ , which is asymptotically negligible. Ignoring these, we can thus consider the random word as  $N_*$  letters  $2, \dots, m$  inserted in  $a_1$  1’s, and the number of inversions is the sum of their positions, counted from the end. It follows easily, either probabilistically or by calculating the characteristic function from (1.6), that  $M_{a_1, \dots, a_m}/N$ , or equivalently  $M_{a_1, \dots, a_m}/a_1$ , converges in distribution to the sum  $\sum_{j=1}^{N_*} U_j$  of  $N_*$  independent random variables  $U_j$  with the uniform distribution on  $[0, 1]$ . Equivalently, since  $\sigma^2 \sim n_1^2 N_*/12 \sim N^2 N_*/12$ ,

$$\frac{M_{a_1, \dots, a_m} - \mu(a_1, \dots, a_m)}{\sigma(a_1, \dots, a_m)} \xrightarrow{d} \sqrt{\frac{12}{N_*}} \sum_{j=1}^{N_*} (U_j - \frac{1}{2}),$$

where  $\xrightarrow{d}$  denotes convergence in distribution. This limit is clearly not normal for any finite  $N_*$ . (However, its distribution is close to standard normal for large  $N_*$ . Note that it is normalized to mean 0 and variance 1.)

## REFERENCES

- |                 |  |
|-----------------|--|
| <b>Andrews</b>  | [1] G. E. Andrews, <i>The Theory of Partitions</i> , Addison-Wesley, 1976.   |
| <b>Bender</b>   | [2] E. A. Bender, Central and local limit theorems applied to asymptotic enumeration, <i>JCT</i> <b>15</b> (1973) 91–111.                              |
| <b>FellerI</b>  | [3] William Feller, <i>An Introduction to Probability Theory and Its Application</i> , volume I, third edition, John Wiley & Sons, New York, 1968.     |
| <b>FellerII</b> | [4] W. Feller, <i>An Introduction to Probability Theory and its Applications</i> , volume II, 2nd ed., Wiley, New York, 1971.                          |
| <b>Ohara</b>    | [5] K. M. O’Hara, Unimodality of Gaussian coefficients: a constructive proof, <i>Journal of Combinatorial Theory, Series A</i> <b>53</b> (1990) 29–52. |
| <b>Schur</b>    | [6] I. J. Schur, see the explicit reference given on page 48 of [1].   |

COMPUTER SCIENCE DEPARTMENT, UNIVERSITY OF GEORGIA, ATHENS, GA 30602-7404, USA

*E-mail address:* `erc [At] cs [Dot] uga [Dot] edu`

DEPARTMENT OF MATHEMATICS, UPPSALA UNIVERSITY, PO BOX 480, SE-751 06 UPPSALA, SWEDEN

*E-mail address:* `svante.janson [At] math [Dot] uu [Dot] se`

*URL:* `http://www.math.uu.se/~svante/`

MATHEMATICS DEPARTMENT, RUTGERS UNIVERSITY (NEW BRUNSWICK), PISCATAWAY, NJ 08854, USA

*E-mail address:* `zeilberg [At] math [Dot] rutgers [Dot] edu`