

Multivariate asymptotic analysis of set partitions: Focus on blocks of fixed size

Research Article

Guy Louchard

Abstract: Using the Saddle point method and multiserries expansions, we obtain from the exponential formula and Cauchy's integral formula, asymptotic results for the number $T(n, m, k)$ of partitions of n labeled objects with m blocks of fixed size k . We analyze the central and non-central region. In the region $m = n/k - n^\alpha$, $1 > \alpha > 1/2$, we analyze the dependence of $T(n, m, k)$ on α . This paper fits within the framework of Analytic Combinatorics.

2010 MSC: 05A16, 60C05, 60F05

Keywords: Set partitions, Bell numbers, Asymptotics, Saddle point method, Multiserries expansions, Analytic combinatorics

1. Introduction

Set partitions parameters have long been a topic of investigation. See, for example, Graham et al. [7], Knuth [9], Mansour [13], Stanley [15], for other investigations of set partitions. Moreover, set partitions continue to be of interest recently; see Chern et al. [1, 2]. During a talk given by P. Diaconis at the Conference in honour of Svante Janson's 60th birthday, our attention was attracted by a classical parameter of set partitions: the number $T(n, m, k)$ of partitions of n labeled objects with m blocks of fixed size k . The value of k will be fixed in this paper, so we suppress the k and simply write this number as $T(m, n)$.

Our goal is to analyze the asymptotic growth of $T(m, n)$ in several regimes.

Let $\Pi(n)$ be the set of partitions of n labeled objects, with $B_n := |\Pi(n)|$ denoting the n th Bell number.

We define the random variable J_n as the number of blocks of size k in a set partition chosen (uniformly) at random from the class of all set partitions of n labeled objects. Then the distribution of J_n

Guy Louchard; Université Libre de Bruxelles, Département d'Informatique, CP 212, Boulevard du Triomphe, B-1050 Bruxelles, Belgium (email: louchard@ulb.ac.be).

is

$$\mathbb{P}(J_n = m) = \frac{T(m, n)}{B_n}.$$

Fristedt [5] proved that this distribution is asymptotically Gaussian.

Most of our techniques for studying $T(m, n)$ and the distribution of J_n rely on generating functions treated as analytic (complex-valued) functions. Thus, we now introduce some of these functions.

We note that $\exp(e^z - 1)$ is the exponential generating function (GF) of the class of set partitions (see [4, pp. 106–111]). Therefore the B_n 's are exactly the coefficients of this GF:

$$\sum_{n=0}^{\infty} B_n \frac{z^n}{n!} = \exp(e^z - 1).$$

Now we decompose the number of such set partitions.

The celebrated exponential formula (or polymer expansion) is given as follows. Let

$$g(z) := \sum_{k=1}^{\infty} a_k \frac{z^k}{k!}.$$

Then

$$\exp(g(z)) = \sum_{n=0}^{\infty} b_n(a_1, \dots, a_n) \frac{z^n}{n!},$$

where the n th Bell polynomial reads as

$$b_n(a_1, \dots, a_n) = \sum_{\lambda \in \Pi(n)} \prod_{k=1}^n a_k^{X_k(\lambda)},$$

and $X_k(\lambda)$ is the number of blocks in λ of fixed size k , λ denoting a set partition.

Hence, we use

$$f_2(z, y) := \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} T(m, n) y^m \frac{z^n}{n!}$$

to denote the analogous bivariate GF (which is exponential in z and ordinary in y); the variable y is used here to “mark” the blocks of size k . See [4, p. 156] for an introduction to the marking technique. It follows immediately that

$$f_2(z, y) = \exp\left(e^z - 1 + (y - 1) \frac{z^k}{k!}\right). \tag{1}$$

Now we define

$$f_3(z) := [y^m] f_2(z, y) = \sum_{n=0}^{\infty} T(m, n) \frac{z^n}{n!}. \tag{2}$$

To find $f_3(z)$, we take the m th derivative of $f_2(z)$ with respect to y , then divide by $m!$ and evaluate at $y = 1$, so that

$$f_3(z) = \exp\left(e^z - 1 - \frac{z^k}{k!}\right) \frac{(z^k/k!)^m}{m!}. \tag{3}$$

In this paper, we will use multiserries expansions: multiserries are in effect power series (in which the powers may be non-integral but must tend to infinity) and the variables are elements of a scale. The scale is a set of variables of increasing order. The series is computed in terms of the variable of maximum order, the coefficients of which are given in terms of the next-to-maximum order, etc. This is more precise than mixing different terms. This technique was used in the analysis of Stirling numbers (of the first and second kind) and Eulerian numbers in Louchard [10–12].

Our paper is organized as follows: in Section 2, we consider the central region, where we re-derive, with more precision, the asymptotic Gaussian property of J_n . In Section 3, we analyze the large deviation $m = n/k - n^\alpha$, with $1 > \alpha > 1/2$. The appendix provides a brief justification of some integration procedures.

2. Central region

We use several saddle points. To ease this paper’s reading, we summarize the different values we need. We consistently use “root” to be interpreted as “root of smallest modulus.”

In section 2,

ρ is the root of $\rho e^\rho = n$,

$\tilde{\rho}_k$ is the root of $\tilde{\rho}_k \exp(\tilde{\rho}_k) = n - k$,

ρ_k is the root of $\rho_k e^{\rho_k} - \frac{k(\rho_k)^k}{k!} + km - n = 0$,

In section 3,

ρ^* is the root of $\rho^* e^{\rho^*} - \frac{k(\rho^*)^k}{k!} - kn^\alpha = 0$,

ρ is the root of $\rho e^\rho = kn^\alpha$.

2.1. The moments

In this section, we first derive asymptotics for the Bell numbers and then proceed to the analysis of the moments of J_n . For the Bell numbers, we could use Salvy and Shackell [14] or Chern et al. [1]. But the first paper uses n/ρ in the scale and the second paper uses the solution of $ue^u = n + 1$. We prefer to use n and ρ : the root (of smallest modulus) of $\rho e^\rho = n$. The scale of the multiserries is $n \gg \rho^k \gg \rho$ (we assume here $k \geq 2$).

We define $m^\ell := \prod_{j=0}^{\ell-1} (m - j) = (m)(m - 1)(m - 2) \cdots (m - \ell + 1)$ as the ℓ th falling factorial of m . Now we use this notation to study the ℓ th falling factorial of J_n . We have

$$\mathbb{E}((J_n)^\ell) = \sum_{m=0}^{\infty} m^\ell \mathbb{P}(J_n = m) = \sum_{m=0}^{\infty} m^\ell \frac{T(m, n)}{B_n} = \frac{\sum_{m=0}^{\infty} m^\ell T(m, n)}{B_n}.$$

From (1), we have

$$\begin{aligned} \sum_{m=0}^{\infty} m^\ell T(m, n) &= n! [z^n] \left. \frac{\partial^\ell}{\partial y^\ell} f_2(z, y) \right|_{y=1} \\ &= n! [z^n] (z^k/k!)^\ell \exp(e^z - 1) \\ &= \frac{n!}{(k!)^\ell} [z^{n-k\ell}] \exp(e^z - 1) \\ &= \frac{n!}{(k!)^\ell} \frac{B_{n-k\ell}}{(n - k\ell)!} \end{aligned}$$

$$= \binom{n}{\underbrace{k, k, \dots, k}_\ell, n - k\ell} B_{n-k\ell}.$$

In particular, for $\ell = 1$ and $\ell = 2$, respectively, we have

$$\sum_{m=0}^{\infty} mT(m, n) = \binom{n}{k} B_{n-k} \quad \text{and} \quad \sum_{m=0}^{\infty} (m)(m-1)T(m, n) = \binom{n}{k, k, n-2k} B_{n-2k}.$$

Therefore, the first and second falling moments of J_n are

$$\mathbb{E}(J_n) = \binom{n}{k} \frac{B_{n-k}}{B_n} \quad \text{and} \quad \mathbb{E}((J_n)(J_n - 1)) = \binom{n}{k, k, n-2k} \frac{B_{n-2k}}{B_n}.$$

Now we need an asymptotic expansion of B_n . Set

$$f_4(z) := \exp(e^z).$$

Let the saddle point be given by ρ and let Ω denote the circle $\rho e^{i\theta}$. We compute

$$(B_n/n!)(e) = ([z^n] \exp(e^z - 1))(e) = [z^n] \exp(e^z) = [z^n] f_4(z).$$

By Cauchy’s theorem, it follows that

$$\begin{aligned} (B_n/n!)(e) &= \frac{1}{2\pi i} \int_{\Omega} \frac{f_4(z)}{z^{n+1}} dz \\ &= \frac{1}{\rho^n} \frac{1}{2\pi} \int_{-\pi}^{\pi} f_4(\rho e^{i\theta}) e^{-ni\theta} d\theta \quad \text{using } z = \rho e^{i\theta} \\ &= \frac{1}{\rho^n} \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(\ln(f_4(\rho e^{i\theta})) - ni\theta) d\theta \\ &= \frac{1}{\rho^n} \frac{f_4(\rho)}{2\pi} \int_{-\pi}^{\pi} \exp\left[-\frac{1}{2}\kappa_2\theta^2 - \frac{i}{6}\kappa_3\theta^3 + \dots\right] d\theta, \end{aligned} \tag{4}$$

where

$$\kappa_i(\rho) := \left(\frac{\partial}{\partial u}\right)^i \ln(f_4(\rho e^u))|_{u=0}.$$

See Good [6] for a neat description of this technique. We have

$$\rho = W(n),$$

where W is Lambert’s W function (see Corless et al. [3]). We use the principal branch, which is analytic at 0. Let us set $L := \ln(n)$. We have the well-known asymptotic expressions

$$\begin{aligned} \rho &= L - \ln L + \frac{\ln L}{L} + \mathcal{O}\left(\frac{(\ln L)^2}{L^2}\right), \\ \ln(\rho) &= \ln(L) - \frac{\ln(L)}{L} + \mathcal{O}\left(\frac{(\ln L)^2}{L^2}\right). \end{aligned}$$

All expressions involving ρ in the sequel can of course be expanded into powers of L , but this would lead to huge formulae.

The dominant part of (4) gives

$$\frac{f_4(\rho)}{\rho^n} = \exp(n/\rho - n \ln(\rho)).$$

Now we turn to the integral. We have for instance

$$\begin{aligned} \kappa_2 &= n(1 + \rho), \\ \kappa_3 &= n(1 + 3\rho + \rho^2). \end{aligned}$$

We now proceed as in Flajolet and Sedgewick [4, ch. VIII]. Let us choose a splitting value θ_0 such that $\kappa_2\theta_0^2 \rightarrow \infty$, and $\kappa_3\theta_0^3 \rightarrow 0$, as $n \rightarrow \infty$. For instance, we can use $\theta_0 = n^{-5/12}$. We must prove that the integral

$$K_n = \int_{\theta_0}^{2\pi - \theta_0} \exp(\ln(f_4(\rho e^{i\theta})) - ni\theta) d\theta$$

is such that $|K_n|$ is exponentially small. This is done in Appendix 3. Now we use the classical trick of setting

$$-\kappa_2\theta^2/2! + \sum_{\ell=3}^{\infty} \kappa_\ell(i\theta)^\ell/\ell! = -u^2/2.$$

Computing θ as a series in u , this gives, by Lagrange’s inversion,

$$\theta = \frac{1}{\sqrt{n}} \sum_{i=1}^{\infty} a_i u^i,$$

with, for instance,

$$a_1 = \frac{1}{\sqrt{1 + \rho}}.$$

This expansion is valid in the dominant integration domain

$$|u| \leq \frac{\sqrt{n}\theta_0}{a_1} = \sqrt{1 + \rho} n^{1/12}.$$

Setting $d\theta = \frac{d\theta}{du} du$, we integrate on $u = [-\infty, \infty]$. This extension of the range is justified as in Flajolet and Sedgewick [4, ch. VIII]. (From now on, we only provide a few terms in our expansions, but of course we use more terms in our computations. Also, all \mathcal{O} terms in the sequel may depend on k, ρ .) This integration gives

$$\frac{B_n e}{n!} = \frac{f_4(\rho)}{\rho^n} H_1 = \frac{1}{\sqrt{2\pi}\sqrt{1 + 1/\rho}} \exp(E_1) H_2,$$

where

$$E_1 := n/\rho - n \ln(\rho) - L/2 - \ln(\rho)/2; \tag{5}$$

$$H_1 := \frac{H_2}{\sqrt{2\pi n(1 + \rho)}};$$

$$H_2 := 1 - \frac{1}{24} \frac{2\rho^4 + 9\rho^3 + 16\rho^2 + 6\rho + 2}{(1 + \rho)^3 n} + \mathcal{O}\left(\frac{1}{n^2}\right). \tag{6}$$

Now we turn to $B_{n-\ell k}$. We compute

$$\ln(n - \ell k) = L - \frac{\ell k}{n} - \frac{1}{2} \frac{(\ell k)^2}{n^2} + \mathcal{O}\left(\frac{1}{n^3}\right).$$

We will use $\tilde{\rho}_{\ell,k}$ as the root of $\tilde{\rho}_{\ell,k} \exp(\tilde{\rho}_{\ell,k}) = n - \ell k$, i.e., $\tilde{\rho}_{\ell,k} = W(n - \ell k)$. This gives

$$\tilde{\rho}_{\ell,k} = \rho - \frac{\rho \ell k}{n(1 + \rho)} - \frac{1}{2} \frac{\rho^2(2 + \rho)(\ell k)^2}{(1 + \rho)^3 n^2} + \mathcal{O}\left(\frac{1}{n^3}\right),$$

with

$$\ln(\tilde{\rho}_{\ell,k}) = \ln(\rho) - \frac{\ell k}{n(1 + \rho)} - \frac{1}{2} \frac{(\ell k)^2(1 + 3\rho + \rho^2)}{(1 + \rho)^3 n^2} + \mathcal{O}\left(\frac{1}{n^3}\right).$$

Now we specialize to the case $\ell = 1$ to get the mean.

Plugging into E_1 , we obtain

$$\frac{B_{n-k} e}{(n - k)!} = \frac{\exp(E_{1k}) H_{2k}}{\sqrt{2\pi} \sqrt{1 + 1/\tilde{\rho}_k}},$$

where

$$E_{1k} := k \ln(\rho) - \frac{1}{2} L - \frac{1}{2} \ln(\rho) - \frac{1}{2} \frac{k(-\rho + k - 2)}{(1 + \rho)n} + \mathcal{O}\left(\frac{1}{n^2}\right),$$

and

$$H_{2k} := 1 - \frac{1}{24} \frac{2\rho^4 + 9\rho^3 + 16\rho^2 + 6\rho + 2}{(1 + \rho)^3 n} + \mathcal{O}\left(\frac{1}{n^2}\right).$$

We need $\frac{B_{n-k}}{B_n}$. This leads to conclude

$$\frac{B_{n-k}}{B_n} = \frac{(n - k)!}{n!} H_6 \rho^k,$$

with

$$H_6 := H_3 H_4 H_5 = 1 - \frac{1}{2} \frac{k(-\rho^2 + k(1 + \rho) - 3\rho - 1)}{(1 + \rho)^2 n} + \mathcal{O}\left(\frac{1}{n^2}\right),$$

where

$$H_3 := \frac{\sqrt{1 + 1/\rho}}{\sqrt{1 + 1/\tilde{\rho}_k}} = 1 - \frac{1}{2} \frac{k}{(1 + \rho)^2 n} - \frac{1}{8} \frac{k^2(8\rho + 2\rho^2 + 1)}{(1 + \rho)^4 n^2} + \mathcal{O}\left(\frac{1}{n^3}\right),$$

and

$$H_4 := \frac{H_{2k}}{H_2} = 1 - \frac{1}{24} \frac{k(11\rho^5 + 28\rho^4 + 36\rho^3 + 10\rho^2 + 10\rho + 2\rho^6 + 2)}{(1 + \rho)^5 n^2} + \mathcal{O}\left(\frac{1}{n^3}\right),$$

and where H_5 is computed as follows:

$$E_{1k} - E_1 = k \ln(\rho) - \frac{1}{2} \frac{k(-\rho + k - 2)}{(1 + \rho)n} + \mathcal{O}\left(\frac{1}{n^2}\right),$$

and

$$\exp(E_{1k} - E_1) = \rho^k H_5,$$

so

$$H_5 = 1 - \frac{1}{2} \frac{k(-\rho + k - 2)}{n(1 + \rho)} + \frac{1}{24} \frac{(9\rho^3 + 39\rho^2 - 10\rho^2 k + 60\rho + 3k^2(1 + \rho) - 30\rho k + 24 - 16k)k^2}{(1 + \rho)^3 n^2} + \mathcal{O}\left(\frac{1}{n^3}\right).$$

The mean M of J_n is given by

$$M = \frac{B_{n-k}n!}{B_n(n-k)!k!} = \frac{\rho^k}{k!} H_6.$$

Similarly (we omit the details)

$$\mathbb{E}(J_n(J_n - 1)) = \frac{\rho^{2k}}{(k!)^2} H_7,$$

where

$$H_7 := 1 - \frac{k(-\rho^2 + 2k(1 + \rho) - 3\rho - 1)}{(1 + \rho)^2 n} + \mathcal{O}\left(\frac{1}{n^2}\right).$$

Hence the variance σ^2 of J_n is given by

$$\sigma^2 = \mathbb{E}(J_n(J_n - 1)) + M - M^2,$$

and

$$\sigma = \sqrt{\frac{\rho^k}{k!} H_8},$$

with

$$H_8 := 1 - \frac{2k^2(\rho + 2)\rho^k + kk!(-3\rho + k(1 + \rho) - \rho^2 - 1)}{4k!(1 + \rho)^2 n} + \mathcal{O}\left(\frac{1}{n^2}\right).$$

More generally, the ℓ th falling moment is given by

$$\begin{aligned} & \frac{n!}{(n - k\ell)!(k!)^\ell} \frac{B_{n-k\ell}}{B_n} \\ &= \frac{\rho^{k\ell}}{(k!)^\ell} H_{7,\ell}, \\ H_{7,\ell} &= 1 - \frac{1}{2} \frac{k\ell(-\rho^2 + k\ell(1 + \rho) - 3\rho - 1)}{(1 + \rho)^2 n} + \mathcal{O}\left(\frac{1}{n^2}\right). \end{aligned}$$

2.2. Distribution of J_n

Fristedt [5] proved that the distribution of J_n is asymptotically Gaussian. This can also be obtained with Hwang’s techniques: see [8]. We want here to re-derive this property with more precision. The corresponding GF $f_5(z)$ is derived from $f_3(z)$:

$$f_5(z) = \exp\left(e^z - \frac{z^k}{k!} + km \ln(z) - \ln(m!) - m \ln(k!)\right),$$

with

$$m = M + x\sigma, \quad x = \Theta(1).$$

The saddle point equation is now

$$\rho_k e^{\rho_k} - \frac{\rho_k^k k}{k!} + km - n = 0. \tag{7}$$

It is easily seen that we have

$$\rho_k \sim \rho + \delta,$$

with

$$\delta = \sum_{j=1}^{\infty} \frac{\alpha_j}{n^j}.$$

When we write $\rho_k \sim \rho + \delta$, this is simply a statement that $\rho_k - \rho$ can be written as a Taylor series in terms of powers of n^{-1} , and then α_1 is just the first coefficient in this Taylor series.

Solving (7) gives, for instance,

$$\alpha_1 := -kx \sqrt{\frac{\rho^k}{k!}} \rho (1 + \rho)^{-1}.$$

Also, we have the classical result that

$$\ln(m!) = -m + m \ln(m) + \frac{1}{2} \ln(2\pi m) + \frac{1}{12} m^{-1} + \mathcal{O}\left(\frac{1}{m^2}\right).$$

We must analyze

$$\mathbb{P}(J_n = m) = \frac{n!}{eB_n} [z^n] f_5(z).$$

First of all, we must compute the dominant term of $\frac{n! f_5(\rho_k)}{eB_n \rho_k^n}$. We have

$$\frac{n! f_5(\rho_k)}{eB_n \rho_k^n} = \exp(T_1) \frac{1}{H_2}$$

with H_2 given by (6) and, with (5),

$$\begin{aligned} T_1 &= \frac{ne^\delta}{\rho} - \frac{\rho_k^k}{k!} + km \ln(\rho_k) - n \ln(\rho_k) - \ln(m!) - m \ln(k!) \\ &\quad - \left(\frac{n}{\rho} - \frac{1}{2}L - n \ln(\rho) - \frac{1}{2} \ln(\rho) - \frac{1}{2} \ln(1 + \rho^{-1}) - \frac{1}{2} \ln(2\pi) \right). \end{aligned}$$

We compute

$$T_1 = T_0 + \frac{T_4}{n} + \mathcal{O}\left(\frac{1}{n^2}\right).$$

The dominant term of T_0 is computed as

$$T_2 := \frac{1}{2} [L + \ln(1 + \rho) - x^2 - k \ln(\rho) + \ln(k!)].$$

Set now the next term of T_0 as $T_3 := T_0 - T_2$. We have

$$\begin{aligned} \frac{n! f_5(\rho_k)}{e B_n \rho_k^n} &= \exp(T_1) \frac{1}{H_2} \\ &= e^{-x^2/2} \frac{\sqrt{k!}}{\sqrt{\rho^k}} \sqrt{1 + \rho} \sqrt{n} \exp\left(T_3 + T_4/n + \mathcal{O}\left(\frac{1}{n^2}\right)\right) \left(1 + \frac{T_5}{n} + \mathcal{O}\left(\frac{1}{n^2}\right)\right), \end{aligned} \tag{8}$$

and

$$T_3 := \frac{1/6x(x^2 - 3)\sqrt{k!}}{\rho^{k/2}} - \frac{1}{12} \frac{k!(-3x^2 + x^4 + 1)}{\rho^k} + \frac{1}{60} \frac{x(3x^4 + 5 - 10x^2)(k!)^{3/2}}{\rho^{3k/2}} + \mathcal{O}\left(\frac{1}{\rho^{2k}}\right).$$

Later on, we will need

$$T_6 := \exp(T_3) = 1 + \frac{1/6x(x^2 - 3)\sqrt{k!}}{\rho^{k/2}} + \frac{1}{72} \frac{(27x^2 - 12x^4 - 6 + x^6)k!}{\rho^k} + \mathcal{O}\left(\frac{1}{\rho^{3k/2}}\right).$$

Also

$$T_5 := \frac{1}{24} \frac{2\rho^4 + 9\rho^3 + 16\rho^2 + 6\rho + 2}{(1 + \rho)^3}.$$

We now turn to the coefficient of $1/n$ in T_1 .

$$\begin{aligned} T_4 &= \frac{1}{2} \frac{k^3 x \sqrt{(k!)^{-1} \ln(\rho)}}{k!(1 + \rho)} \rho^{3k/2} \\ &\quad + \frac{1}{2} k^2 \ln(\rho) \left(-\rho^2 + \frac{x^2 \rho}{\ln(\rho)} - 3\rho + \frac{x^2}{\ln(\rho)} + k(1 + \rho) - 1\right) \rho^k (k!)^{-1} (1 + \rho)^{-2} \end{aligned}$$

and

$$\exp\left(\frac{T_4}{n}\right) = 1 + \frac{T_7}{n} + \mathcal{O}\left(\frac{1}{n^2}\right),$$

and

$$T_7 \equiv T_4.$$

To compute the integral, we obtain, for instance,

$$\kappa_2 = (1 + \rho)n + \left[-\frac{k^2 \rho^k}{k!} - \frac{x \sqrt{(k!)^{-1}} (1 + 3\rho + \rho^2) k \rho^{k/2}}{(1 + \rho)} + \mathcal{O}\left(\frac{1}{\rho^{k/2}}\right) \right] + \mathcal{O}\left(\frac{1}{n}\right),$$

and the integral leads to

$$I = \frac{1}{\sqrt{n} \sqrt{2\pi} \sqrt{1 + \rho}} \left(1 + \frac{T_8}{n} + \mathcal{O}\left(\frac{1}{n^2}\right)\right),$$

with

$$T_8 := \frac{1}{24} \frac{(24k^2 \rho^2 + 12k^2 \rho^3 + 12k^2 \rho) \rho^k}{(1 + \rho)^3 \rho k!} + \mathcal{O}\left(\rho^{k/2}\right).$$

We compute now

$$T_9 := T_5 + T_7 + T_8 = T_{91} \rho^{3k/2} + T_{92} \rho^k + \mathcal{O}\left(\rho^{k/2}\right),$$

with

$$T_{91} := \frac{1}{2} \frac{k^3 x \ln(\rho)}{(k!)^{3/2} (1 + \rho)},$$

and

$$T_{92} := \frac{1}{2} \frac{k^2 (\rho + x^2(1 + \rho) + 1)}{(1 + \rho)^2 k!} + \frac{1}{2} \frac{k^2 (-\rho^2 + k(1 + \rho) - 3\rho - 1) \ln(\rho)}{(1 + \rho)^2 k!}.$$

Combining the integral I with (8) gives the local limit theorem:

Theorem 2.1. *The asymptotic distribution of J_n is given by the local limit theorem:*

$$\mathbb{P}(J_n = m) = e^{-x^2/2} \frac{\sqrt{k!}}{\sqrt{\rho^k \sqrt{2\pi}}} T_6 \left(1 + \frac{T_9}{n} + \mathcal{O}\left(\frac{1}{n^2}\right) \right). \tag{9}$$

Of course more terms can be mechanically computed, but the expressions become much more intricate.

To check the quality of our asymptotics, we have chosen $k = 2, n = 1000$, the range of interest for m is given by $m \in \left(\frac{\rho^k}{k!} + 2\sqrt{\frac{\rho^k}{k!}}, \frac{\rho^k}{k!} - 2\sqrt{\frac{\rho^k}{k!}} \right) = (6, 21)$. To numerically compute $T(m, n)$, we use (3), which gives

$$\frac{T(m, n)}{n!} = [z^{n-km}] \exp\left(e^z - 1 - \frac{z^k}{k!}\right) \frac{1}{m!(k!)^m}.$$

Figure 1 shows $T(m, n)$ (circle), the asymptotics $e^{-x^2/2} \frac{\sqrt{k!}}{\sqrt{\rho^k \sqrt{2\pi}}}$ (line, of course we use here $\frac{\rho^k}{k!}$ as mean and variance) and Equ. (9) (box).

Figure 2 gives the quotient of Equ. (9) and $T(m, n)$.

Figure 3 gives the quotient of Equ. (9) and $T(m, n)$ (box) and the quotient of $e^{-x^2/2} \frac{\sqrt{k!}}{\sqrt{\rho^k \sqrt{2\pi}}}$ and $T(m, n)$ (line).

Note that the same technique would lead to the joint distribution of $J_n(k_1), J_n(k_2)$ for two (or more) different values of k .

3. Large deviation $m = n/k - n^\alpha, \quad 1 > \alpha > 1/2$

We have a maximum of n/k blocks of size k in a partition. So we use

$$m = \frac{n}{k} - n^\alpha.$$

This can be written as

$$m = \frac{n}{k}(1 - \varepsilon),$$

with $\varepsilon := kn^{\alpha-1}$.

In this section, we choose $\alpha \geq \frac{1}{2}$, but the other case is similarly analyzed. The saddle point equation, from (7) becomes

$$\rho^* e^{\rho^*} - \frac{\rho^{*k} k}{k!} + km - n = \rho^* e^{\rho^*} - \frac{\rho^{*k} k}{k!} - \tilde{n} = 0$$

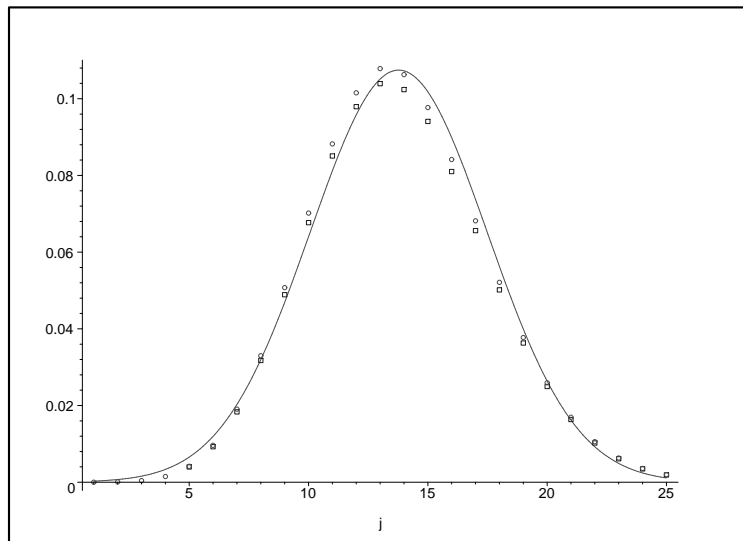


Figure 1. $T(m, n)$ (circle), the asymptotics $e^{-x^2/2} \frac{\sqrt{k!}}{\sqrt{\rho^k \sqrt{2\pi}}}$ (line) and Equ. (9) (box)

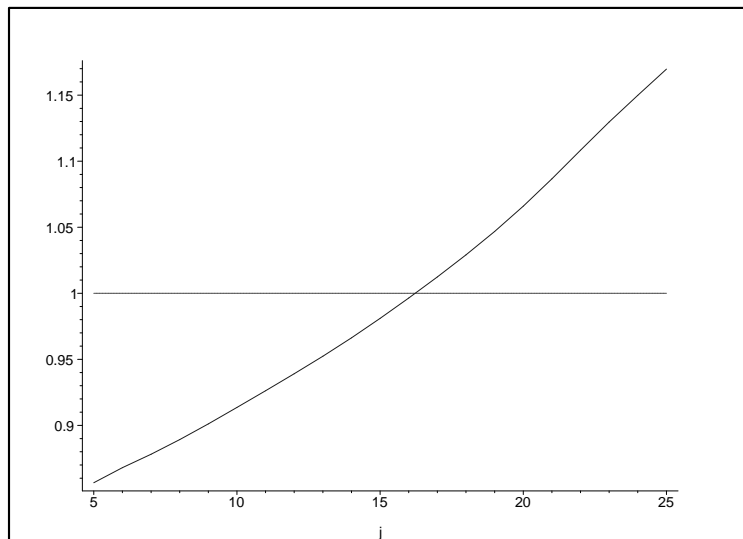


Figure 2. Quotient of Equ. (9) and $T(m, n)$

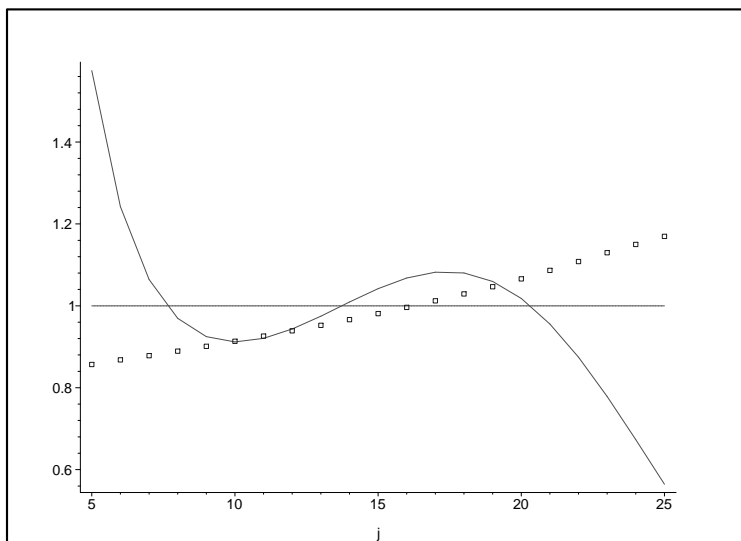


Figure 3. Quotient of Equ. (9) and $T(m, n)$ (box), quotient of $e^{-x^2/2} \frac{\sqrt{k!}}{\sqrt{\rho^k} \sqrt{2\pi}}$ and $T(m, n)$ (line)

with $\tilde{n} := kn^\alpha$. The multiseries' scale is here $n \gg \tilde{n} \gg \frac{1}{\varepsilon} \gg L$. The solution of $\rho e^\rho = \tilde{n}$ is asymptotically given by

$$\rho = \alpha L + \ln(k) - \ln(\alpha) - \ln(L) + \left(-\frac{\ln(k)}{\alpha} + \frac{\ln(\alpha) + \ln(L)}{\alpha} \right) L^{-1} + \mathcal{O}\left(\frac{1}{L^2}\right).$$

As previously, we have

$$\rho^* = \rho + \delta,$$

with, here,

$$\delta = \frac{\rho^k k \rho}{k! (1 + \rho) \tilde{n}} + \frac{1}{2} \frac{(\rho^k)^2 k^2 (-2\rho - \rho^2 + 2k + 2\rho k) \rho}{(k!)^2 (1 + \rho)^3 \tilde{n}^2} + \mathcal{O}\left(\frac{1}{\tilde{n}^3}\right).$$

First of all, we have

$$\frac{n! f_5(\rho^*)}{e^{\rho^* n}} = \frac{n!}{e} \exp(R_0),$$

with

$$R_0 = \frac{e^\delta \tilde{n}}{\rho} - \frac{\rho^{*k}}{k!} + km \ln(\rho^*) - n \ln(\rho^*) - \ln(m!) - m \ln(k!).$$

We have, successively,

$$R_0 = nR_1 + R_2 + \mathcal{O}\left(\frac{1}{n}\right),$$

$$R_1 := R_{10} + \frac{R_{11}}{\tilde{n}} + \frac{R_{12}}{\tilde{n}^2} + \mathcal{O}\left(\frac{1}{\tilde{n}^3}\right),$$

$$R_2 := \frac{\tilde{n}}{\rho} + R_{20} + \frac{R_{21}}{\tilde{n}} + \mathcal{O}\left(\frac{1}{\tilde{n}^2}\right),$$

with

$$R_{10} := \frac{1}{2} \frac{2 - 2 \ln(k!) + 2 \ln(k) - 2L}{k} + \frac{1}{2} \frac{-2 \ln(k) + 2L - 2k \ln(\rho) + 2 \ln(k!)}{k} \varepsilon - \frac{1}{2} k^{-1} \varepsilon^2 + \mathcal{O}(\varepsilon^3),$$

$$R_{11} := -\frac{\rho^k k}{k!(1+\rho)} \varepsilon,$$

$$R_{12} := -\frac{1}{2} \frac{\rho^{2k} (-\rho^2 + 2k(1+\rho) - 3\rho - 1) k^2}{(k!)^2 (1+\rho)^3} \varepsilon,$$

and

$$R_{20} = -\frac{\rho^k}{k!} - \frac{1}{2} L + \frac{\rho^k k}{k!(1+\rho)} - \frac{1}{2} \ln\left(\frac{2\pi}{k}\right) + \frac{1}{2} \varepsilon + \mathcal{O}(\varepsilon^2),$$

and

$$R_{21} = \frac{1}{2} \frac{\rho^{2k} (-2 - 5\rho - 2\rho^2 + 2k(1+\rho))}{(1+\rho)^3 (k-1)!^2}.$$

Computing the integral, we have, for instance,

$$\kappa_2 := (1+\rho) \tilde{n} - \frac{\rho^k k (-\rho^2 + k(1+\rho) - 3\rho - 1)}{k!(1+\rho)} + \mathcal{O}\left(\frac{1}{\tilde{n}}\right),$$

and the integral leads to

$$I = \frac{1}{\sqrt{1+\rho}\sqrt{2\pi}} \frac{1}{\sqrt{\tilde{n}}} \left(1 + \frac{R_5}{\tilde{n}} + \mathcal{O}\left(\frac{1}{\tilde{n}^2}\right)\right),$$

and

$$R_5 := \frac{R_3}{R_4},$$

with

$$R_3 := 12k^2 \rho^{k+1} (1+\rho) - 36\rho^{k+1} k(1+\rho) + 12\rho^k k^2 (1+\rho) - 12\rho^{k+2} k(1+\rho) - 12\rho^k k(1+\rho) - 2k! \rho^4 - 9k! \rho^3 - 16k! \rho^2 - 6k! \rho - 2k!,$$

where

$$R_4 := 24(1+\rho)^3 k!.$$

Finally, we obtain the following asymptotic result

Theorem 3.1. *The asymptotic expression of the $T(m, n)$ for large deviation is given by*

$$T(m, n) = \frac{n!}{e} \exp(R_0) \frac{1}{\sqrt{1 + \rho\sqrt{2\pi}}} \frac{1}{\sqrt{\tilde{n}}} \left(1 + \frac{R_5}{\tilde{n}} + \mathcal{O}\left(\frac{1}{\tilde{n}^2}\right) \right). \tag{10}$$

Let us analyze the importance of our terms. We have two sets: the set A of dominant terms, which stay in the exponent and the set B of small terms, leading to a coefficient of type $(1 + \Delta)$, with Δ small. The property of each term may depend on α . For instance, in R_{10} , the first term leads to an $\mathcal{O}(nL)$ term, the ε term leads to an n^α term, the ε^2 term leads to an $n^{2\alpha-1}$ term, which are all $\in A$. the ε^3 term leads to an $n^{3\alpha-2}$ term which is $\in A$ if $\alpha \geq 2/3$ and $\in B$ otherwise. In R_{11} the ε term leads to a term $\in A$. In R_{12} all terms are $\in B$. In R_2 , the \tilde{n} term is $\in A$, in R_{20} the first term is $\in A$, all other terms are $\in B$, in R_{21} , all terms are $\in B$.

We finally mention that our non-central range is not sacred: other types of ranges can be analyzed with similar methods.

To check the quality of our asymptotics, we have first chosen $k = 2, \alpha = .52$, a range $n \in [10000, 70000]$ and $m = \lfloor \frac{n}{k} - n^\alpha \rfloor$. Figure 4 shows the quotient Equ. (10)/ $T(m, n)$. The fluctuations are due to the fact that m is integer, so the value of α we need is actually the root of $m - (\frac{n}{k} - n^\alpha) = 0$.

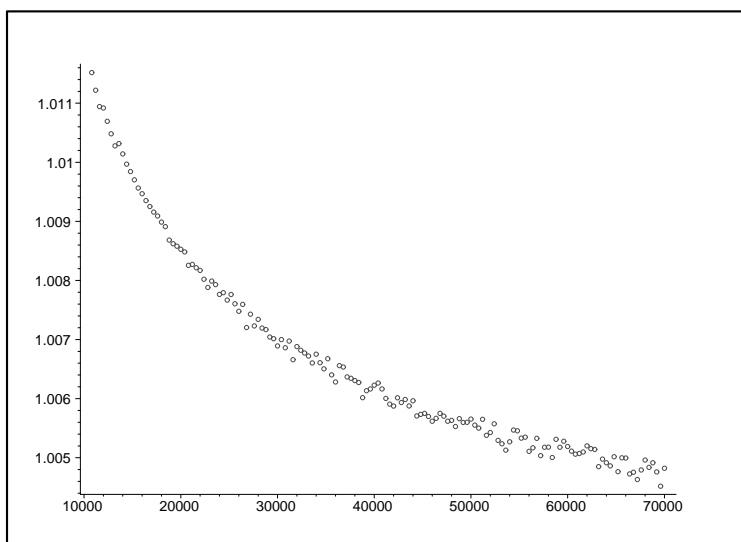


Figure 4. Quotient Equ. (10)/ $T(m, n)$, $\alpha = .52, n \in [10000, 70000]$

For $\alpha = .65$, we choose $n \in [100, 2300]$, this leads to Figure 5. The quality of the asymptotics decreases for $\alpha \geq .7$, more terms would be necessary.

Another way is to fix n . We have chosen $k = 2, n = 10000$ and a range $\alpha \in (.52, .65)$ hence $m \in [4600, 4900]$. Again α is chosen as the root of $m - (\frac{n}{k} - n^\alpha) = 0$. Figure 6 shows the quotient Equ. (10)/ $T(m, n)$.

Acknowledgment: We would like to thank two referees for many useful suggestions that improved the paper.

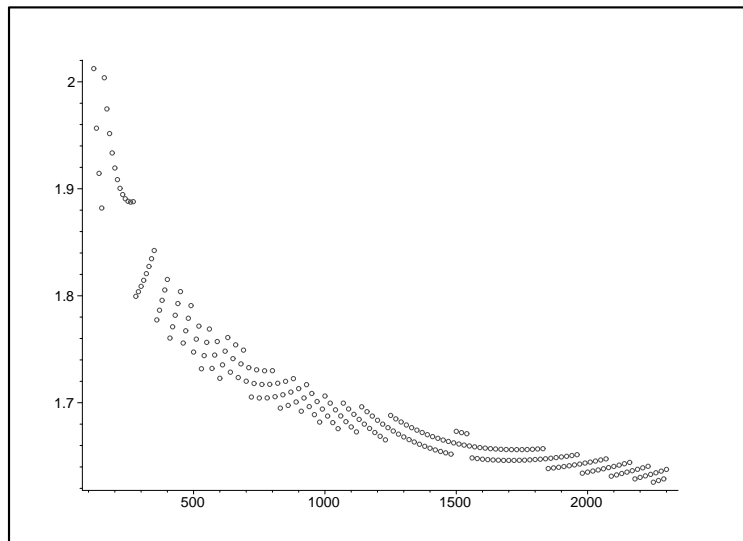


Figure 5. Quotient Equ. (10)/ $T(m, n)$, $\alpha = .65, n \in [100, 2300]$

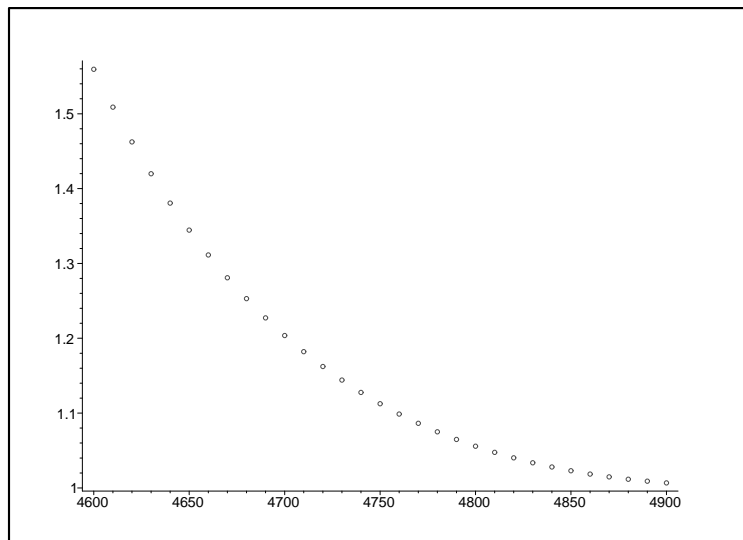


Figure 6. Quotient Equ. (10)/ $T(m, n)$, $n = 10000, \alpha \in (.52, .65), m \in [4600, 4900]$

References

- [1] B. Chern, P. Diaconis, D. M. Kane, R. C. Rhoades, Closed expressions for set partition statistics, *Res. Math. Sci.* 1(2) (2014) 1–32.
- [2] B. Chern, P. Diaconis, D. M. Kane, R. C. Rhoades, Central limit theorems for some set partitions, *Adv. Appl. Math.* 70 (2015) 92–105.
- [3] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, D. E. Knuth, On the Lambert W function, *Adv. Comput. Math.* 5 (1996) 329–359.
- [4] P. Flajolet, R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, 2009.
- [5] B. Fristedt, The structure of random partitions of large sets, Technical report, University of Minnesota, 1987.
- [6] I. J. Good, Saddle-point methods for the multinomial distribution, *Ann. Math. Statist.* 28(4) (1957) 861–881.
- [7] R. L. Graham, D. E. Knuth, O. Patashnik, *Concrete Mathematics*, Second Edition, Addison Wesley, 1994.
- [8] H. K. Hwang, On convergence rates in the central limit theorems for combinatorial structures, *European J. Combin.* 19(3) (1998) 329–343.
- [9] D. E. Knuth, *The Art of Computer Programming*, vol. 4a: Combinatorial Algorithms. Part I, Addison–Wesley, Upper Saddle River, New Jersey, 2011.
- [10] G. Louchard, Asymptotics of the Stirling numbers of the first kind revisited: A saddle point approach, *Discrete Math. Theor. Comput. Sci.* 12(2) (2010) 167–184.
- [11] G. Louchard, Asymptotics of the Stirling numbers of the second kind revisited: A saddle point approach, *Appl. Anal. Discrete Math.* 7(2) (2013) 193–210.
- [12] G. Louchard, Asymptotics of the Eulerian numbers revisited: A large deviation analysis, *Online J. Anal. Comb.* 10 (2015) 1–11.
- [13] T. Mansour, *Combinatorics of Set Partitions*, Discrete Mathematics and Its Applications Series, CRC Press, Boca Raton, FL, 2013.
- [14] B. Salvy, J. Shackell, Symbolic asymptotics: Multiseries of inverse functions, *J. Symbolic Comput.* 20(6) (1999) 543–563.
- [15] R. P. Stanley, *Enumerative Combinatorics*, Volume 1, 2nd edn, Cambridge Studies in Advanced Mathematics, Vol. 49. Cambridge University Press, Cambridge, 2012.

Appendix: Justification of the integration procedure

For the B_n case, we must analyze

$$\Re(\ln(f_4(\rho e^{i\theta}) - ni\theta)).$$

Let us first notice that $ni\theta$ does not contribute to the analysis. Next, we have

$$\Re\left(e^{\rho e^{i\theta}}\right) = e^{\rho \cos(\theta)} \cos(\rho \sin(\theta))$$

which has a dominant peak at 0.

For the Gaussian case, we use $f_5(z)$. We must analyze

$$\Re\left(e^{\rho e^{i\theta}} - \frac{(\rho e^{i\theta})^k}{k!}\right) = e^{\rho \cos(\theta)} \cos(\rho \sin(\theta)) - \frac{\rho^k}{k!} \cos(k\theta)$$

which has a dominant peak at 0.

The non-central region leads to the same analysis.