Probabilistic approach of asymptotics of integer partitions

Jean Peyen
Supervised by Leonid Bogachev and Paul Martin

Department of Statistics – University of Leeds

January 7, 2020

Representation of integer partitions

Definition

An integer partition is an unordered finite sequence of natural numbers. We will denote by N the sum of a partition and by M its length.

Occupation numbers

$\nu_l =$ number of occurences of l.

Young diagram

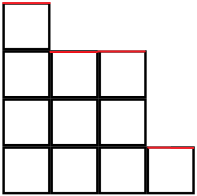
$Y(x) = \sum_{l>x} \nu_l$.

Example

$\lambda = (4,3,3,1)$ is a partition of 11 of length 4. The occupation numbers are:

$$\begin{pmatrix} l & 1 & 2 & 3 & 4 \\ \nu_l & 1 & 0 & 2 & 1 \end{pmatrix}.$$

We can represent its Young diagram by taking the upper boundary of the following graphic:



Asymptotic enumeration

How can we estimate the number of partitions of a large integer ?

Limit shape

For some choice of distribution, up to an appropriate rescaling, what is the average shape of a large partition ?

Examples

Hardy–Ramanujan formula

$$F_n \sim \frac{1}{4\sqrt{3}n} \exp\left(2\pi \cdot \sqrt{\frac{n}{6}}\right)$$

Temperley curve

Let $\langle Y_n(x) \rangle$ be the expected Young diagram of a uniformly distributed partition of n, then:

$$\langle Y_n(\sqrt{n}x)/\sqrt{n} \rangle \longrightarrow -\frac{\sqrt{6}}{\pi} \log(1 - \exp(-\pi x/\sqrt{6}))$$

Interest of studying partitions

Models for gases of identical particles

With various classes of partitions, one can describe various kind of particles.

bosons	partitions
fermions	strict partitions
anyons	continuation of minimal difference partitions
degenerate energy levels	weighted partitions

Random fragmentation ,biodiversity modeling, representation theory etc...

Interest of a probabilistic approach

Simple observation

• If Γ is uniformly distributed over $\mathcal{F}$, then for all possible configuration γ, $\mathbb{P}(\Gamma = \gamma) = 1/\sharp\mathcal{F}$.

• With various distributions, one can describe various physical systems.

How can we prove limit results for integer partitions ?

General scheme for studying a combinatorial structure

1. Specification of the class

2. Ensemble equivalence, construction of the Boltzmann distribution

3. Calibration of the distribution and use of local central limit results

1. Specification

For non-restricted partitions

$$\mathcal{F} = MSet(\mathbb{N}) = \prod_{x \in \mathbb{N}} Seq(\{x\})$$

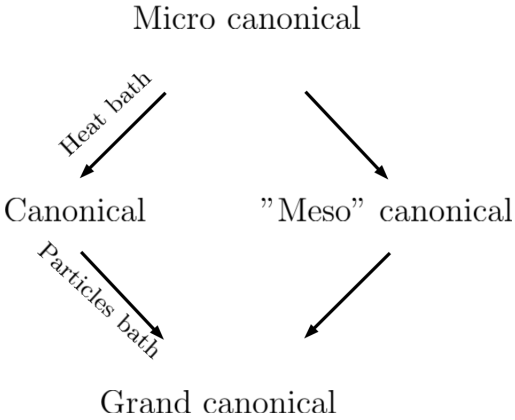
For strict partitions

$$\mathcal{F} = PSet(\mathbb{N}) = \prod_{x \in \mathbb{N}} \{x, \varepsilon\}$$

Restriction of the source of the parts

$\mathbb{N}$ can be replaced by a subset (e.g. we consider an assembly of particles with energy levels that are not equally separated)

2. Grand-canonical ensemble



Instead of studying a distribution of partitions of a given number n ("meso" canonical ensemble) that goes to infinity we study a distribution without restriction of size.

3. Calibration and local central limit results

- The Boltzmann distribution is defined up to a free parameter, the temperature of the system.

- In order to establish results for more restricted ensembles, one has to calibrate that parameter and to compute the asymptotics as the temperature goes to infinity.

Boltzmann distribution

In statistical mechanics

$$\mathbb{P}_\beta(\Gamma = \gamma) = \frac{e^{-\beta H(\gamma)}}{Z(\beta)}$$

where β is the inverse temperature of the system, $H(\gamma)$ the energy of the configuration γ and Z is the partition function of the system.

In combinatorics

$$\mathbb{P}_z(\Gamma = \gamma) = \frac{z^{\mathsf{size}(\gamma)}}{F(z)}$$

z is a free parameter and $F(z) = \sum_\gamma z^{\mathsf{size}(\gamma)} = \sum_n F_n z^n$ is the generating function of the structure that we study.

Why do we use the Boltzmann distribution

- It is uniform up to conditioning i.e. :

$$\mathbb{P}(\Gamma = \gamma \mid \mathsf{size}(\gamma) = n) = \frac{1}{F_n}$$

- z can be calibrated in order to have $\langle \mathsf{size}(\Gamma) \rangle = n$.

- There is a simple correspondence between Boltzmann distribution and constructors operations.

Correspondence between Boltzmann distribution and constructors

$\mathcal{C} = \mathcal{A} \uplus \mathcal{B}$	$\mathbb{P}(\Gamma_{\mathcal{C}} \in \mathcal{A}) = A(z)/(A(z) + B(z))$
$\mathcal{C} = \mathcal{A} \times \mathcal{B}$	$\mathbb{P}(\Gamma_{\mathcal{A}} = (a,b)) = \mathbb{P}(\Gamma_{\mathcal{A}} = a) \cdot \mathbb{P}(\Gamma_{\mathcal{B}} = b)$
$\mathcal{C} = Seq(\mathcal{A})$	$\mathbb{P}(\Gamma_{\mathcal{C}} \in \mathcal{A}^d) = A(z)^d(1 - A(z))$

In the case of partitions we usually use the following constructors:

$MSet(\mathcal{A}) = \prod_{a \in \mathcal{A}} Seq(\{a\})$,
$PSet(\mathcal{A}) = \prod_{a \in \mathcal{A}} \{a, \varepsilon\}$, where size$(\varepsilon) = 0$.

Boltzmann distribution for unrestricted partitions

1. Specification

We consider the structure $MSet(\mathbb{N})$.

2. Boltzmann distribution

The occupation numbers ν_l are independent and follow the Geometric distribution with parameter $1 - z^l$.

3. Calibration of z

Let Γ be a random partition following the Boltzmann distribution, if z is defined by :

$$z = \exp\left(-\frac{\pi}{\sqrt{6n}}\right)$$

then, in the limit $n \to \infty$, we have:

$$\langle \text{size}(\Gamma) \rangle = n - \frac{\sqrt{6n}}{2\pi} + O(1).$$

Sketch of proof

We just need to write $\langle \text{size}(\Gamma) \rangle = \sum_l l \langle \nu_l \rangle$ and to approximate that sum by an integral with the Euler–Maclaurin formula.

Hardy–Ramanujan formula (1918)

$$F_n \sim \frac{1}{4\sqrt{3}n} \exp\left(2\pi \cdot \sqrt{\frac{n}{6}}\right)$$

Sketch of proof

For all choice of parameter z:

$$\begin{aligned}\frac{1}{F_n} &= \mathbb{P}_z(\Gamma = \gamma \mid \mathsf{size}(\Gamma) = n) \\ &= \frac{z^n}{F(z)} \cdot \frac{1}{\mathbb{P}_z(\mathsf{size}(\Gamma) = n)}.\end{aligned}$$

Thus we can estimate F_n by using a local central limit result for size(Γ) and by estimating $F(z)$ for the choice of parameter given previously when $n \to \infty$.

Local central limit result for size(Γ)

$$\mathbb{P}_z(\mathsf{size}(\Gamma) = n) \sim \frac{1}{\sqrt{2\pi V}} \exp\left(-\frac{(n - \langle \mathsf{size}(\Gamma) \rangle)^2}{2V}\right)$$

where the variance verifies:

$$V = Var(\mathsf{size}(\Gamma)) \sim \frac{2\sqrt{6}}{\pi} n^{3/2}.$$

This result can be proved by using the characteristic function of size(Γ).

The Temperley curve (1952)

Let Y be the Young diagram of a random partition Γ following the Boltzmann distribution with the parameter z defined previously, then in the limit $n \to \infty$:

$$\langle Y(\sqrt{n}x)/\sqrt{n}\rangle \longrightarrow -\frac{\sqrt{6}}{\pi}\log(1-\exp(-\pi x/\sqrt{6}))$$

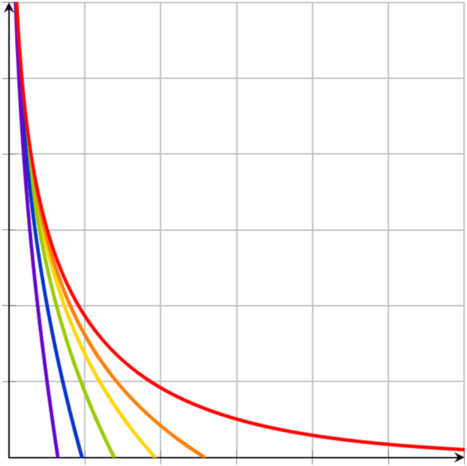
Generalization

Limit shape for Minimal Difference Partitions (Comtet, Majumdar, Sabhapandit, 2008)

The limit shape for uniform integer partitions with a minimal difference of $p \in \mathbb{N}$ between the parts is:

$$y = -px - \frac{\sqrt{6}}{\pi} \log(1 - exp(-\pi x/\sqrt{6}))$$

This result has been further generalized to minimal difference partitions with differences specified by a sequence. It can also be analytically continued to $p \in [0, 1]$ in order to provide models for anyons (Bogachev, Yakubovich, 2019).



Shape for MDP(p) for $p = 0$ (bosons); 0.25; 0.5; 1 (fermions); 2; 4 (resp from red to purple)

Boltzmann distribution with controlled number of parts

In order to approximate the micro canonical ensemble, one can add an extra free parameter for the number of parts.
We define the following distribution:

$$\mathbb{P}_{(z_1,z_2)}(\Gamma = \gamma) = \frac{z_1^n z_2^m}{F(z_1, z_2)},$$

where γ is a partition of n into m parts.

Power restriction and controlled number of parts

Problem

Can we give an asymptotic estimate of the number of partitions of a large integer n in m p^{th} power integers ?

- Subsequently we will consider that Γ follows the distribution defined earlier on strict partitions into p^{th} power integers.

- We will consider the limit $\langle N \rangle \to \infty$, $\langle M \rangle$ fixed.

Calibration of the parameters

In the limit $\langle M\rangle$ constant and $\langle N\rangle \to \infty$:

$$z_1 = \exp\left(\frac{-\langle M\rangle}{p\,\langle N\rangle}\right) + O\left(\frac{1}{\langle N\rangle^{1+1/p}}\right),$$

$$z_2 = \frac{p^{1-1/p}\,\langle M\rangle^{1+1/p}}{\Gamma(1/p)\,\langle N\rangle^{1/p}} + o\left(\frac{1}{\langle N\rangle^{1/p}}\right).$$

Idea of the proof

As previously, we use the occupation numbers. Here the Boltzmann distribution is characterized by their independence and:

$$\nu_{l^p} \sim \text{Bernoulli}\left(\frac{z_1^{l^p} z_2}{1 + z_1^{l^p} z_2}\right).$$

Distribution of N and M

In the limit $\langle M \rangle$ constant and $\langle N \rangle \to \infty$:

$$M \xrightarrow{\text{law}} \text{Poisson}(\langle M \rangle)$$
$$\left(\frac{N}{\langle N \rangle}\right)_{|M=m} \xrightarrow{\text{law}} \text{Gamma}\left(\frac{m}{p}, \frac{p}{\langle M \rangle}\right).$$

Sketch of the proof

For M

We can use the probability generating function:

$$\log \mathsf{PGF}_M(x) = \sum_l \log \mathsf{PGF}_{\nu_{lp}}(x) = \sum_l \log \left\langle x^{\nu_{lp}} \right\rangle$$

$$= \sum_l \log \left(\frac{1 + x z_1^{lp} z_2}{1 + z_1^{lp} z_2} \right)$$

$$= \langle N \rangle (x - 1) + o(1)$$

For $\left(\frac{N}{\langle N\rangle}\right)_{|M=m}$

Using the Bayes formula, we can rewrite the moment generating function for $(N/\langle N\rangle, M)$ as follows:

$$\begin{aligned}\langle \exp((SN/\langle N\rangle) + tM)\rangle = \sum_m \mathbb{P}(M = m)e^{tm} \\ \cdot \sum_n \mathbb{P}\left(E = n\langle E\rangle \,|M = m\right)e^{sn},\end{aligned}$$

then it suffices to estimate the moment generating function and to identify it with the expression above.

Counting partitions?

The number of strict partitions of integers smaller than x into m p^{th} powers verifies:

$$D_{m,p}(x) = \sum_{n \leq x} F_{n,m} = \frac{\Gamma(1/p)^m n^{m/p}}{m! m \Gamma(m/p) p^{m-1}} + o\left(n^{m/p}\right)$$

▶ It is coherent with what we are expecting from the behaviour of the volume of a p-norm ball.

Problem

The rescaling of N makes this approach unable to "see" the gaps. Some numbers, cannot be partitioned in that way (cf. Waring problems).
By differentiating D we only have access to a density function that does not always coincide with the counting sequence.

Hardy–Ramanujan and Temperley

Asymptotic formulæ in combinatory analysis, Hardy and Ramanujan, 1918

Statistical mechanics and the partition of numbers ii. the form of crystal surfaces, Temperley, 1952

Overview of the approach

Statistical mechanics of combinatorial partitions, and their limit shapes, Vershik, 1996

Boltzmann distribution for combinatorics

Boltzmann Samplers for the Random Generation of Combinatorial Structures, Duchon, Flajolet, Louchard, Schaeffer, 2003

Minimal Difference partitions

Limit shape of minimal difference partitions and fractional statistics, Bogachev, Yakubovich, 2019