A PROBABILISTIC PROOF OF TWO COMBINATORIAL IDENTITIES

PATRICK POLO

ABSTRACT. Let n be an integer ≥ 2 and let $X_1, \ldots, X_n$ be denumerable sets, not necessarily disjoint, each X_i being endowed with a probability μ_i given by its discrete density p_i. Let $Y = X_1 \times \cdots \times X_n$, endowed with the product probability μ_Y, and let Y_* be the subset of Y consisting of n-tuples of pairwise distinct elements. In the special case where all X_i equal the same set X, with discrete density p, it is well-known to combinatorists that $\mu_Y(Y_*)$ can be expressed as a certain polynomial in the Newton sums $S_k = \sum_{x \in X} p(x)^k$, for $k = 2, \ldots, n$, and this formula also belongs to the folklore of statistical mechanics, in particular to the combinatorics of cluster expansions, although it seems hard to find in the literature in probability theory. In our slightly more general setting, we give a simple proof of a similar formula, which expresses $\mu_Y(Y_*)$ as a certain polynomial in the sums $S_I = \sum_{x \in \bigcap_{i \in I} X_i} \prod_{i \in I} p_i(x)$, for I varying in the set of subsets of $\{1, \ldots, n\}$ of cardinality ≥ 2. This gives a simple, probabilistic proof of the special case which is worthwile, we believe, to be better known. In the course of our proof, we obtain that the sum of the terms $(-1)^k c_k(n)$, where $c_k(n)$ denotes the number of connected graphs with n vertices and k edges, equals $(-1)^{n-1}(n-1)!$. We also include a purely combinatorial proof of this identity, communicated to us by Jean-Yves Thibon.

INTRODUCTION

During an exercise session for a second year undergraduate course of probability theory, the author was set to explain that if (X, μ) is a denumerable probability space and n an integer ≥ 2, then the experiment consisting of drawing n elements of X with replacement leads to the product space X^n, whereas the drawings without replacement lead to the subset X^n_* consisting of n-tuples of pairwise distinct elements. But he stopped in the middle of it, realising that X^n_* is usually considered only in the case where X is a finite set endowed with the uniform probability, in which case X^n_* is the set of arrangements of n elements of X, endowed with the uniform probability. However it is natural to consider the general case, endowing X^n_* with the probability induced

Date: August 30, 2023.

1991 *Mathematics Subject Classification.* 05C30, 05E05, 60C05.

Key words and phrases. Connected graphs; Symmetric functions; Sampling without replacement;

from the product probability μ_{X^n} on X^n, which leads one to compute the value of $\mu_{X^n}(X^n_*)$.

As, according to the probabilists consulted by the author, this value was not well-known, this led us to write a first version of this Note, requiring no prerequisites beyond a first course in discrete probabilities. A first board of editors kindly replied that the observation and its proof belong to the folklore of statistical mechanics, in particular to the combinatorics of cluster expansions, although they are hard to find in the literature in such a simple form. A second referee, for another generalist journal, kindly replied that the computation is nice but belongs better to a textbook in basic probability and combinatorics. A third referee, for a journal in probability theory, rejected the original Note without adressing the novelty or not of the formula, but commenting on its lack of applications. Finally, Jean-Yves Thibon, acting as an editor of Séminaire Lotharingien de Combinatoire, kindly replied that the formula was well-known to combinatorists as formula (7.23) in [2] (as well as [1], formula (2.14′) p. 25).

This suggested that a short, pedagogical Note could be useful to make a bridge between the probabilistic and combinatorial viewpoints, and this is the aim of the present, expanded Note, where we consider the slightly more general setting of n denumerable probability spaces (X_i, μ_i).

Besides, we obtain in the course of our proof that the sum of the terms $(-1)^k c_k(n)$, where $c_k(n)$ denotes the number of connected graphs with n vertices and k edges, equals $(-1)^{n-1}(n-1)!$. A self-contained, purely combinatorial proof of this was communicated to us by Jean-Yves Thibon and appears in Section 2. We thank him heartily.

1. Probabilistic viewpoint

1.1. **The partial diagonals in the product $Y = X_1 \times \cdots \times X_n$.** Let n be an integer ≥ 2 and let $X_1, \dots, X_n$ be denumerable sets, not necessarily disjoint, each X_i being endowed with a probability μ_i defined by its discrete density $p_i(x) = \mu_i(\{x\})$ for all $x \in X_i$. We endow $Y = X_1 \times \cdots \times X_n$ with the product probability, denoted by μ_Y.

Let $\mathscr{A} = \mathscr{P}_2([n])$ denote the set of 2-element subsets of $[n] = \{1, \dots, n\}$, considered as the set of edges of the complete graph with n vertices G_n. For each $a \in \mathscr{A}$ consider the partial diagonal

$$Y_a = \{(x_1, \dots, x_n) \in Y \mid x_i = x_j \text{ if } i, j \in a\}.$$

Definition 1. For each subset I of $[n]$ of cardinality ≥ 2, set $X_I = \bigcap_{i \in I} X_i$ and $S_I = \sum_{x \in X_I} \prod_{i \in I} p_i(x)$. We also set set $S_I = 1$ if $|I| = 1$.

Then $\mu_Y(Y_a) = S_a$ for all $a \in \mathscr{A}$.

Denote by Y_* the complement of the union of the Y_a's.

1.2. **Connected graphs.** For each non-empty subset A of $\mathscr{A}$, set $Y_A = \bigcap_{a\in A} Y_a$. Set also $Y_\varnothing = Y$. By the inclusion-exclusion formula, one has

$$\mu_Y(Y_*) = \sum_{A\in\mathscr{P}(\mathscr{A})} (-1)^{|A|}\,\mu_Y(Y_A). \tag{1}$$

For each subgraph Γ of G_n, denote by $\mathrm{conn}(\Gamma)$ the set of connected components of Γ. Then, denoting by Γ_A the subgraph of G_n with set of edges A and recalling Definition 1, one has:

$$\mu_Y(Y_A) = \prod_{C\in\mathrm{conn}(\Gamma_A)} S_C.$$

Next, denoting by $\alpha(C)$ the number of edges of each element C of $\mathrm{conn}(\Gamma_A)$, the sum of the $\alpha(C)$'s equals $|A|$ and hence (1) can be rewritten as:

$$\mu_Y(Y_*) = \sum_{A\in\mathscr{P}(\mathscr{A})}\ \prod_{C\in\mathrm{conn}(\Gamma_A)} (-1)^{\alpha(C)} S_C. \tag{2}$$

Each $A \in \mathscr{P}(\mathscr{A})$ gives a partition of the set $[n]$ into the set of connected components of Γ_A, and conversely each partition $(C_1,\dots,C_r)$ of $[n]$ comes from all the $A \in \mathscr{P}(\mathscr{A})$ which are union of subsets $A_i \in \mathscr{P}_2(C_i)$ such that (C_i, A_i) is a connected graph. Denote by $\mathrm{Part}([n])$ the set of partitions of the set $[n]$. Then (2) can be rewritten:

$$\mu_Y(Y_*) = \sum_{\substack{(C_1,\dots,C_r)\\ \in\mathrm{Part}([n])}} \prod_{i=1}^{r} \sigma(|C_i|)\, S_{C_i}, \tag{3}$$

where for each $m \in \mathbb{N}^*$ one denotes by $\sigma(m)$ the sum of the $(-1)^{\alpha(\Gamma)}$'s, for Γ running through all connected graphs on the set of vertices $[m] = \{1,\dots,m\}$. Thus, denoting by $c_k(m)$ the number of connected graphs with set of vertices $[m]$ and k edges, one has

$$\sigma(m) = \sum_{k=m-1}^{m(m-1)/2} (-1)^k c_k(m). \tag{$\star$}$$

1.3. **A special case and the value of $\sigma(n)$.** Let us consider here the case where all X_i contain a common subset X, the $X_i - X$ are pairwise disjoint, and $p_i(x)$ does not depend on i for $x \in X$. In this case, for each $k \in \mathbb{N}^*$, set

$$S_k = \sum_{x\in X} p_1(x)^k.$$

Then, for each subset I of $[n]$ with $|I| \geq 2$, one has $X_I = X$ and $S_I = S_{|I|}$.

Further, for each partition $\lambda = 1^{d_1}2^{d_2}\cdots n^{d_n}$ of the integer n (that is, d_k is the number of parts equal to k), the number of partitions of the set $[n]$ of type λ is

$$\frac{1}{d_1!\cdots d_n!}\,\frac{n!}{\prod_{k=1}^n (k!)^{d_k}}$$

and (3) can be rewritten as:

$$\mu_Y(Y_*) = \sum_{d_1+2d_2+\cdots+nd_n=n} \frac{1}{d_1!\cdots d_n!}\,\frac{n!}{\prod_{k=1}^n (k!)^{d_k}} \prod_{k=1}^n (\sigma(k)S_k)^{d_k}. \tag{4}$$

Next, one has the following:

Lemma 1. *One has* $\sigma(n) = (-1)^{n-1}(n-1)!$.

Proof. We derive this from our previous discussion.[1] Fix an arbitrary integer $N \geq 2$ and consider the special case where all X_i are equal to a set X of cardinality N endowed with the uniform probability; then $Y = X^n$ and $S_k = 1/N^{k-1}$ for each $k \in \mathbb{N}^*$. Further, each term $\prod_{k=1}^n S_k^{d_k}$ in (4) is a monomial in $1/N$ of degree $n-(d_1+\cdots+d_n) \leq n-1$, with equality only in the case where $d_n = 1$. Therefore, $\mu_Y(Y_*)$ is a polynomial in $1/N$ whose leading term is $\sigma(n)$.

On the other hand, the product probability μ_Y on $Y = X^n$ is the uniform probability and therefore the probability of the subset Y_* of n-tuples of distinct elements is:

$$\mu_Y(Y_*) = \frac{|Y_*|}{N^n} = \prod_{k=1}^{n-1}\big(1 - \frac{k}{N}\big)$$

and hence the leading term $\sigma(n)$ equals $(-1)^{n-1}(n-1)!$. □

1.4. **The probability $\mu_Y(Y_*)$.** Let us come back to the general case. Thanks to Lemma 1, one deduces from (3) and (4) the following proposition.

Proposition 1. *One has*

$$\mu_Y(Y_*) = \sum_{\substack{(C_1,\dots,C_r)\\ \in \mathrm{Part}([n])}} (-1)^{n-r} \prod_{i=1}^r (|C_i|-1)!\, S_{C_i}. \tag{5}$$

Further, if $X_1,\dots,X_n$ all contain a common subset X, the $X_i - X$ are pairwise disjoint, and $p_i(x)$ does not depend on i for $x \in X$, then

$$\mu_Y(Y_*) = \sum_{d_1+2d_2+\cdots+nd_n=n} \frac{n!}{d_1!\cdots d_n!} \prod_{k=1}^n \left(\frac{(-1)^{k-1}S_k}{k}\right)^{d_k}. \tag{6}$$

[1] A direct proof, using the exponential formula for generating series, was pointed out by Jean-Yves Thibon and is given in Section 2.

Remark 1. Denote by Q_n the right hand side of (5); it is a polynomial with integral coefficients in the indeterminates S_I, for I a subset of $[n]$ of cardinality ≥ 2. For example, one has has $Q_2 = 1 - S_{12}$ and $Q_3 = 1 - S_{12} - S_{23} - S_{13} + 2S_{123}$, then $Q_4 = 1 - \sum_{1\leq i<j\leq 4} S_{ij} + 2\sum_{1\leq i<j<k\leq 4} S_{ijk} - 6S_{1234}$, and

$$Q_5 = 1 - \sum_{1\leq i<j\leq 5} S_{ij} + \sum_{1\leq i<j<k\leq 5} 2S_{ijk} - \sum_{1\leq i<j<k<\ell\leq 5} 6S_{ijk\ell} + 24S_5 + \sum_{1\leq i<j<k<\ell\leq 5} (S_{ij}S_{k\ell} + S_{ik}S_{j\ell} + S_{i\ell}S_{jk}) - \sum_{1\leq i<j\leq 5} 2S_{ij}S_{abc}$$

where in the last sum one has $[5] - \{i, j\} = \{a, b, c\}$.

Remark 2. For each partition $\lambda = 1^{d_1}2^{d_2}\cdots n^{d_n}$ of n, with $r = \sum_{k=1}^{n} d_k$ parts, set $z_\lambda = \prod_{k=1}^{n} k^{d_k} d_k!$ and $\varepsilon_\lambda = (-1)^{n-r}$, and also $P_\lambda = \prod_{k=1}^{n} S_k^{d_k}$ (recalling that $S_1 = 1$). Then formula (6) can be rewritten as

$$\mu_Y(Y_*) = n! \sum_{\lambda\in\mathrm{Part}(n)} \frac{\varepsilon_\lambda}{z_\lambda} P_\lambda. \tag{6'}$$

When all X_i are equal to the same set X, so that $Y = X^n$, this formula is well-known to combinatorists (see [1], formula (14′) p. 25 or [2], formula (7.23); see also Section 2 for more details) and also belongs to the folklore of statistical mechanics.

2. Combinatorial viewpoint

2.1. **A direct proof of Lemma 1.** We thank heartily Jean-Yves Thibon for communicating to us the following self-contained, purely combinatorial proof of Lemma 1, which is an application of the «exponential formula» in generatingfunctionology, see [2], Cor. 5.1.6 and Example 5.2.1.

We keep the previous notation and, for each $r \in \mathbb{N}^*$, denote by $\mathrm{CG}([r])$ the set of connected graphs with set of vertices $[r]$. Consider the exponential generating series

$$F(x) = \sum_{r\geq 1} \sigma(r)\frac{x^r}{r!} = \sum_{r\geq 1} \sum_{\Gamma\in\mathrm{CG}([r])} (-1)^{\alpha(\Gamma)}\frac{x^r}{r!}.$$

For each integer $m \geq 2$, $\dfrac{F(x)^m}{m!}$ equals

$$\sum_{n\geq m} \frac{x^n}{n!} \left(\frac{1}{m!} \sum_{\substack{r_1,\dots,r_m\geq 1\\ r_1+\dots+r_m=n}} \frac{n!}{r_1!\cdots r_m!} \sum_{\substack{\Gamma_1\in\mathrm{CG}([r_1])\\ \cdots\\ \Gamma_m\in\mathrm{CG}([r_m])}} (-1)^{\alpha(\Gamma_1)+\dots+\alpha(\Gamma_m)} \right).$$

One sees that the expression between parentheses equals the sum of the $(-1)^{\alpha(\Gamma)}$ taken over all subgraphs Γ of G_n with exactly m connected components. It follows that

$$\exp F(x) = 1 + x + \sum_{n\geq 2} \Big(\sum_{\Gamma\subset G_n} (-1)^{\alpha(\Gamma)} \Big) \frac{x^n}{n!}.$$

For each $n \geq 2$, the previous sum equals $\sum_{k=0}^{N} (-1)^k \binom{N}{k} = 0$, where $N = \binom{n}{2}$. Thus $\exp F(x) = 1 + x$ and hence $F(x) = \log(1 + x) = \sum_{n\geq 1} (-1)^{n-1} \dfrac{x}{n}$ and therefore $\sigma(n) = (-1)^{n-1}(n-1)!$ for all $n \geq 1$. □

2.2. **Symmetric functions.** Let $\Lambda = \bigoplus_{n\in\mathbb{N}} \Lambda_n$ be the graded ring of symmetric functions in countably many independent variables $x_1, x_2, \ldots$ (see [1], §I.2 or [2], §7.1). For each $k \in \mathbb{N}^*$, consider the elementary symmetric function

$$e_k = \sum_{i_1<i_2<\cdots<i_k} x_{i_1} \cdots x_{i_k}$$

(with $e_0 = 1$) and the power sum $P_k = \sum_{i\in\mathbb{N}^*} x_i^k$. (It is usually denoted by p_k, but we use here uppercase P_k since we are using lowercase p's to denote probabilities.) More generally, for each partition λ with r parts $\lambda_1 \geq \cdots \geq \lambda_r \geq 1$, consider the power sum

$$P_\lambda = P_{\lambda_1} \cdots P_{\lambda_r};$$

it belongs to Λ_n, where $n = |\lambda| = \lambda_1 + \cdots + \lambda_r$. Then, using the notation z_λ and ε_λ of Remark 2 of Section 1, recall the following identity (see [2], (7.23) or [1], (14′), p. 25).

Proposition 2. *For each* $n \in \mathbb{N}^*$, *one has*

$$e_n = \sum_{\lambda\in\mathrm{Part}(n)} \frac{\varepsilon_\lambda}{z_\lambda} P_\lambda. \tag{9}$$

Proof. For the convenience of the reader, we give the following proof, taken from [1], pp. 23–25. Let t be another indeterminate and consider the generating series

$$E(t) = \sum_{r\geq 0} e_r t^r = \prod_{i\geq 1} (1 + t x_i) \tag{7}$$

and

$$P(t) = \sum_{r\geq 1} P_r (-t)^{r-1} = \sum_{i\geq 1} \sum_{r\geq 1} x_i^r (-t)^{r-1} = \sum_{i\geq 1} \frac{x_i}{1 + t x_i} = \frac{d}{dt} \log E(t).$$

Since $\log E(t)$ is a power series in $\Lambda[[t]]$ with constant term 0, one obtains by formal primitivation that

$$\log E(t) = \sum_{r\geq 1} P_r (-1)^{r-1} \frac{t^r}{r}$$

and hence

$$E(t) = \prod_{r\geq 1} \exp\Big(P_r(-1)^{r-1}\frac{t^r}{r}\Big) = \prod_{r\geq 1}\sum_{d_r\geq 0}\Big(\frac{(-1)^{r-1}P_r}{r}\Big)^{d_r}\frac{t^{rd_r}}{d_r!}. \tag{8}$$

The proposition follows, by comparing the coefficients of t^n in (7) and (8) for each $n \in \mathbb{N}^*$. □

2.3. **Relating Propositions 1 and 2.** Proposition 2 implies formula (6), or equivalently (6′), in the special case where $X_1, \dots, X_n$ all equal the same probability space X, with discrete density p. Indeed, choosing a numbering $\{z_1, z_2, \dots\}$ of X and denoting by x_i the probability $p(z_i)$, one obtains that for each n-tuple of positive integers $i_1 < \cdots < i_n$ and each permutation $\sigma \in S_n$, one has

$$\mu_{X^n}(z_{i_{\sigma(1)}}, \dots, z_{i_{\sigma(n)}}) = x_{i_1}\cdots x_{i_n}$$

and hence $\mu_{X^n}(X^n_*) = n!\,e_n$. This observation was pointed out to me first by Thierry Levy, then again by Jean-Yves Thibon. I thank them both for their comments.

Conversely our formula (6) implies Proposition 2. Indeed, Λ_n is the inverse limit, as $r \to \infty$, of the degree n part of the ring of symmetric functions in the finitely many variables $x_1, \dots, x_r$ (see [1], pp. 18–19). But, for fixed $r \in \mathbb{N}^*$, both sides of (9) are homogeneous polynomials of degree n in the variables $x_1, \dots, x_r$, and the special case of formula (6) where $X_1, \dots, X_n$ are all equal to an arbitrary probability space X of cardinality r tells that these two polynomials agree on the set $\{(x_1, \dots, x_r) \in \mathbb{R}^r_+ \mid x_1 + \cdots + x_r = 1\}$ hence, by homogeneity, on the set $\{(x_1, \dots, x_r) \in \mathbb{R}^r_+ \mid x_1 + \cdots + x_r > 0\}$, hence they are equal.

References

[1] I. G. Macdonald, Symmetric Functions and Hall Polynomials, 2nd. ed, Oxford Univ. Press, 1995.

[2] R. P. Stanley, Enumerative Combinatorics, Vol. 2, Cambridge Univ. Press, 1999.

Sorbonne Université , Institut de Mathématiques de Jussieu, Case 247, 4 place Jussieu, 75252 Paris cedex 05

Email address: patrick.polo@imj-prg.fr