EXPECTED PROPERTIES OF SET PARTITIONS

M.Y. Vardi

Department of Applied Mathematics
The Weizmann Institute of Science
Rehovot, Israel

1. BASIC DEFINITIONS

A <u>sequence</u> is a k-tuple $\langle x_1,\ldots,x_k\rangle$ for some $k > 0$, such that $x_i \neq x_j$ for $1 \leq i < j \leq k$. We define $SET(\langle x_1,\ldots,x_k\rangle) = \{x_1,\ldots,x_k\}$. Let S be a non-empty finite set, $S = \{a_1,\ldots,a_m\}$. A <u>unordered partition</u> of S is a collection of sets $\{S_1,\ldots,S_k\}$ such that $S_i \cap S_j = \emptyset$ for $1 \leq i < j \leq k$, $S_i \subseteq S$ for $1 \leq i \leq k$, and $S_1 \cup \ldots \cup S_k = S$. Without loss of generality assume that $a_1 \in S_1$. A <u>semi-ordered partition</u> of S is a sequence of sets $\langle S_1,\ldots,S_k\rangle$ such that $\{S_1,\ldots,S_k\}$ is a unordered partition of S. An <u>ordered partition</u> of S is a sequence of sequences $\langle s_1,\ldots,s_k\rangle$ such that $\{SET(s_1),\ldots,SET(s_k)\}$ is a unordered partition of S.

Let P be a partition of S (unordered, semi-ordered or ordered). The <u>length</u> of P, denoted L(P), is the number of elements (sets or sequences) in P. The weight of P, denoted W(P), is the number of elements in the first element of P (i.e., S_1 or s_1). Assuming that all partitions of some type are equiprobable, we define $L_U(m)$ as the expected length of all unordered partitions of m-sets, and we define $W_U(m)$ as the expected weight of all unordered partitions of m-sets. Similiarly, we define $L_S(m)$ and $W_S(m)$ for semi-ordered partitions, and we define $L_O(m)$ and $W_O(m)$ for ordered partitions. In what follows we investigate these parameters.

2. ORDERED PARTITIONS

Let h(m) denote the number of ordered partitions of m-sets.

Proposition 2.1: $h(m) = m! \cdot 2^{m-1}$.

Proof: There are $h(m-k) \cdot m!/(m-k)!$ ordered partitions P of m-sets such that W(P) = k. Thus, h(m) is defined by the recurrence relation

$h(m) = \sum_{i=1}^{m} h(m-i) \cdot m!/(m-i)!$, where $h(0) = 1$.

It is easily verified that $h(m) = m! \cdot 2^{m-1}$. □

Proposition 2.2: $W_O(m) = 2 - 2^{1-m}$.

Proof: Since all ordered partitions are equiprobable, we have

$W_O(m) = \sum_{i=1}^{m} i \cdot (m!/(m-i)!) \cdot (h(m-i)/h(m)) = 2 - 2^{1-m}$. □

Corollary: $\lim_{m \to \infty} W_O(m) = 2$. □

Proposition 2.3: $L_O(m) = m/2 + 1/2$.

Proof: There are $m! \cdot C(m-1,k-1)$ (1) ordered partitions P of m-sets such that L(P)=k, hence

$L_O(m) = \sum_{i=1}^{m} i \cdot m! \cdot C(m-1,i-1)/h(m) = m/2 + 1/2.$ □

3. SEMI-ORDERED PARTITIONS

Let g(m) denote the number of semi-ordered partitions of m-sets. There are C(m,k)·f(m-k) semi-ordered partitions P of m-sets such that W(P) = k. Thus, g(m) is defined by the recurrence relation

$g(m) = \sum_{i=1}^{m} C(m,i) \cdot g(m-i)$, where g(0) = 1.

Proposition 3.1: $W_S(m) = 2m \cdot g(m-1)/g(m)$.

Proof: Since all semi-ordered partitions are equiprobable

$W_S(m) = \sum_{i=1}^{m} i \cdot C(m,i) \cdot g(m-i)/g(m) = 2m \cdot g(m-1)/g(m)$. □

Proposition 3.2: $L_S(m) = g(m+1)/2g(m) - 1/2$.

Proof: Let G(m,k) denote the number of semi-ordered partitions P of m-sets such that L(P) = k. Such a partition can be obtained by adding the m-th element in one of k possible ways to a semi-ordered partition P\` of m-1-set such that L(P\`) is either k or k-1. Hence G(m,k) is defined by the recurrence relation

$G(m,k) = k \cdot [G(m-1,k-1) + G(m-1,k)]$, where $G(m,k) = 0$ for $k < 1$ and $k > m$.

Thus, $L_S(m) = \sum_{i=1}^{m} i \cdot G(m,i)/g(m) = g(m+1)/2g(m) - 1/2$. □

To study the asymptotic behaviour of $W_S(m)$ and $L_S(m)$ we study first

---

(1) C(m,n) is the binomial coefficient m!/(n!·(m-n)!).

the asymptotic behaviour of $g(m)$.

Proposition 3.3: $g(m) \sim m!\cdot(\ln 2)^{-m-1}/2$.

Proof: Define the generating function for $g(m)$: $G(x) = \sum_{i=0}^{\infty} x^i g(i)/i!$. Using the recurrence relation for $g(m)$ we get $G(x) = 1 + (e^x - 1)\cdot G(x)$; hence $G(x) = 1/(2 - e^x)$. We use Darboux' method of finding late term coefficients and expand $G(x)$ in the neighbourhood of the singularity point $x = \ln 2$. Thus, using $e^{x-\ln 2} \approx 1 + x - \ln 2$, we get:

$$G(x) = 1/(2-e^x) = 1/2(1-e^{x-\ln 2}) =$$

$$= (1/2)\cdot \sum_{i=0}^{\infty} (x + \ln(e/2))^i = (1/2)\cdot \sum_{i=0}^{\infty} x^i \cdot \sum_{j=0}^{\infty} C(i+j,j)\cdot(\ln(e/2))^j.$$

Hence, $g(m) \sim (m!/2)\cdot \sum_{j=0}^{\infty} C(m+j,j)\cdot(\ln(e/2))^j$, and by Gould (1972, formula 1.3), $g(m) \sim m!\cdot(\ln 2)^{-m-1}/2$. □

Corollary 1: $\lim_{m\to\infty} W_S(m) = \ln 4$. □

Corollary 2: $L_S(m) \sim m/\ln 4$. □

3. UNORDERED PARTITIONS

Let $f(m)$ denote the number of unordered partitions of m-sets. There are $C(m-1,k-1)\cdot f(m-k)$ unordered partitions P of m-sets such that $W(P) = k$. Thus, $f(m)$ is defined by the recurrence relation

$$f(m) = \sum_{i=1}^{m} C(m-1,i-1)\cdot f(m-i), \text{ where } f(0) = 1.$$

Proposition 4.1: $W_U(m) = 1 + (m-1)\cdot g(m-1)/g(m)$.

Proof: Since all unordered partitions are equiprobable

$$W_U(m) = \sum_{i=1}^{m} i\cdot C(m-1,i-1)\cdot f(m-i)/f(m) = 1 + (m-1)\cdot f(m-1)/f(m). \quad \square$$

Proposition 4.2: $L_U(m) = f(m+1)/f(m) - 1$.

Proof: Let $F(m,k)$ denote the number of unordered partitions P of m-sets such that $L(P) = k$. Such a partition can be obtained either by adding the m-th element in one of k possible ways to a unordered partition P` of an m-1-set such that $L(P`)=k$, or by taking a unordered partition P` of an m-1-set such that $L(P`) = k-1$ and adding the m-th element as the k-th set. Hence

$F(m,k) = F(m-1,k-1) + k \cdot F(m-1,k)$, where $F(m,k) = 0$ for $k < 1$ or $k > m$.

Thus, $L_U(m) = \sum_{i=1}^{m} i \cdot F(m,i)/f(m) = f(m+1)/f(m) - 1$. □

Let $u(m)$ be the solution of the transcendental equation $u \cdot e^u = m + 1$, that is $u \sim \ln m - \ln \ln m$.

Proposition 4.3: $W_U(m) \sim u(m)$, $L_U(m) \sim m/u(m)$.

Proof: By de Bruijn (1970, $6.2), the asymptotic approximation of $f(m)$ is

$f(m) \sim m! \cdot \exp(e^u - u \cdot e^u \cdot \ln u - u/2)/(e^2 \cdot 2 \quad \cdot (1+1/u))^{1/2}$.

The claim follows by Propositions 4.1 and 4.2. □

ACKNOWLEDGEMNT

I am grateful to Professor J. Gillis for suggesting the use of Darboux' method to prove Proposition 3.3.

REFERENCES

N.G. de Bruijn (1970): Asymptotic methods in analysis, North-Holland, Amsterdam.

H.W. Gould (1972): Combinatorial identities, Morgantown, W. Va.