2010–05–25 $03^{53}$

# Multinomial expressions summation asymptotic approximations.

by

Jörgen Backelin

**Abstract** The asymptotic behaviour with respect to $n$ of sums of type

$$\sum_{i_1,\dots,i_r} \prod_{j=1}^{r} v_j^{i_j} \binom{n}{i_1,\dots,i_r} \rho^{-f(i_1,\dots,i_r)}$$

is determined, where the $w_j$ are positive weights, and $f(i_1,\dots,i_r)$ is the sum of some of the monomials $i_k i_m$ $(1 \le k < m \le r)$.

**Summary.**

In the study of probabilistic tournaments, questions arise about the asym= ptotic behaviour of the following type:

Consider a (simple and undirected) graph $G = ([r], E)$ (for some positive integer $r$). Define a quadratic form $f(x_1,\dots,x_r)$ as the sum of the monomials $x_j x_l$ $(1 \le j < l \le r)$, such that $\{j,l\} \in E$, and for every positive integer $n$ form the multinomial coefficient sum

$$S(n) = \sum_{i_1,\dots,i_r} \binom{n}{i_1,\dots,i_r} \frac{1}{2}^{-f(i_1,\dots,i_r)}$$

(sum over non-negative integers $i_1$, …, $i_r$, such that the multinomial coefficients are defined). The answer turns out to be that

$$S(n) \sim C \cdot \alpha(G)^n$$

where $\alpha(G)$ is the independence number of $G$, and $C$ is the number of independent sets of size $\alpha(G)$ in $G$.

It turns out that it is easier to prove a generalisation of this, where each term also has a weight factor $\prod_j v_j^{i_j}$, where the $v_i$ are positive weights for the polynomial $p$. In this case, the independence number $\alpha G$ should be replaced by the maximal weight of an independent set in $G$, and $C$ by the number of independent sets of maximal weight. This is formulated as the proposition *infra*, which also encompasses an estimate of the error term.

**Remarks.** The investigation was motivated by combinatorial problems, and the proof methods are elementary and purely combinatoric. However, i am indepted to Anders Martin-Lf and August Tsikh for pointing out that the results seem be related to those studied by analytic means, e.g., by Laplace's method. The results e. g. may be interpreted in terms of the radius of convergence for the analytic function $\sum S(n)x^n$.

Ola Hössjer pointed out that the sum can be interpreted as an expected value, which gives the moment generating function interpretation. $\mathbf{E}(rho^S)$, where $S$ is weighted $U$-statistic, and that this also might lead to some generalisations.

**Notation**. In this note, "natural number" is taken to mean "non-negative integer", and consequently $\mathbf{N} = \{0, 1, 2, \dots\}$. The set of positive integers is denoted $\mathbf{Z}_+$, and likewise $\mathbf{R}_+$ is the set of positive real numbers.

Multinotation will be used in an unstrict manner, with multiitems denoted by lowercase boldface letters, corresponding to slanted letters for their members; e.g., the same $\mathbf{i}$ may stand for either $(i_1, \ldots, i_r)$ or $i_1, \ldots, i_r$, depending on the context, and $\mathbf{v}^{\mathbf{i}}$ then may stand for the product $\prod_j v_j^{i_j}$. Likewise, the range of a sum mostly will be indicated and partly implied by the context; the reader should have no difficulty in filling in the details.

For any set $S$, its *power set* is $\mathcal{P}(S) \stackrel{def}{=} \{D : D \subseteq S\}$. Set difference, denoted $\setminus$, operates on $\mathcal{P}(S)$; $A \setminus B = \{x \in A : x \notin B\}$.

The *support* of any real-valued function $f : S \longrightarrow \mathbf{R}$ is the set

$$\operatorname{supp} F = \{x \in A : f(x) \neq 0\} .$$

Interpreting an $\mathbf{i} \in \mathbf{R}^r$ as a function from $[r]$ to $\mathbf{R}$, naturally we get $\operatorname{supp} \mathbf{i} = \{j : i_j \neq 0\}$.

For any natural number $n$, $[n] = \{i \in \mathbf{N} : 0 < i \leq n\}$. All graphs considered are simple, undirected, and finite. Thus, we may write such a graph $G$ as a pair $(V, E)$, where $|V| \stackrel{def}{=}$ (cardinality of $V$) $< \infty$, and

$$E \subseteq \binom{V}{2} \stackrel{def}{=} \{e \in \mathcal{P}(V) : |e| = 2\} .$$

An *independent set* $S$ in $G$ is a subset of $V$, such that $E \cap \binom{S}{2} = \emptyset$. An independent set which is maximal under set inclusion is called a *basis*. The set of all bases of $G$ is denoted $\mathcal{B}(G)$, and is a subset of $\mathcal{P}(V)$. The *independence number* of $G$ is

$$\alpha(G) = \max_{B \in \mathcal{B}(G)} |B| .$$

**Preliminaries.** We shall investigate some sums involving multinomial coefficients (always assuming that the lower indices, here $i_1, \ldots, i_r$, are natural numbers whose sum is the top one, $n$)

$$\binom{n}{\mathbf{i}} = \binom{n}{i_1, \ldots, i_r} = n! \prod_{j=1}^{r} (i_j!)^{-1} ,$$

and we start by reminding of some well-known properties of these. First, the multinomial theorem states that for any constants $v_1, \ldots, v_r$,

$$\sum_{\mathbf{i}} \binom{n}{\mathbf{i}} \mathbf{v}^{\mathbf{i}} = \sum_{i_1, \ldots, i_r} \binom{n}{i_1, \ldots, i_r} v_1^{i_1} \cdots v_r^{i_r} = (v_1 + \ldots + v_r)^n . \tag{1}$$

Moreover, directly from the definition, for $r \geq 2$ we may express the multinomial coefficient as a product of a simpler one and of a binomial coefficient: putting $\mathbf{i} = \mathbf{i}', i_r$ (i.e., letting $\mathbf{i}' = (i_1, \ldots, i_{r-1})$),

$$\binom{n}{\mathbf{i}} = \binom{n}{i_r} \binom{n - i_r}{\mathbf{i}'} . \tag{2}$$

A *weighted graph* $(V, E, \psi)$ is a graph $G = (V, E)$ together with a *(vertex) weight function* $\psi : V \longrightarrow \mathbf{R}_+$. By abuse of notation, the weighted graph also is called $G$. The weight function extends naturally to $\mathcal{P}(V)$ by

$$\psi(S) = \sum_{v \in S} \psi(v) .$$

Let $K(G)$ be the maximal weight of a basis, and $N(G)$ be the number of times this quantity is reached, i.e.,

$$K(G) = \max_{B \in \mathcal{B}(G)} \psi(B) , \qquad N(G) = |\{B \in \mathcal{B}(G) : \psi(B) = K(G)\}| .$$

For $G = (V, E)$, a subset $C$ of $V$ is a *vertex cover*, if $e \cap C \neq \emptyset$ for each edge $e \in E$. Equivalently, $C$ is a vertex cover if and only if its complement $V \setminus C$ is independent. Let $\mathcal{C}(G)$ be the set of all vertex covers which are minimal under set inclusion. Clearly,

$$C \in \mathcal{C}(G) \iff V \setminus C \in \mathcal{B}(G) .$$

**Main result.** The formulation given here is the one I found easiest to prove directly. Actually, in the original problem, there were no weights (or unit weights) on the vertices; I give this as a corollary. There is also a slight extension, where essentially the uniform $\rho$ is replaced by different $\rho$ for different edges $\{i, j\}$; this is given as a remark, after the proof.

**Lemma.** *Let $G = ([r], E, \psi)$ be a weighted graph, with weights $v_j = \psi(j)$. Furthermore, let $\rho$ be a fixed number strictly between 0 and 1. Finally, for each natural number $n$, put*

$$S(n) = S_{G,\mathbf{v},\rho}(n) = \sum_{\mathbf{i}} \binom{n}{\mathbf{i}} \rho^{f(\mathbf{i})} \mathbf{v}^{\mathbf{i}} ,$$

*where the sum is taken over those $\mathbf{i} = (i_1, \ldots, i_r) \in \mathbf{N}^r$ for which the multinomial coefficient $\binom{n}{\mathbf{i}}$ is defined and non-zero, and*

$$f(\mathbf{i}) = f_G(\mathbf{i}) = \sum_{\{v_j, v_l\} \in E} i_j i_l .$$

*Then $S(n) \sim N(G)K(G)^n$, i.e.,*

$$\lim_{n\to\infty} \frac{S(n)}{K(G)^n} = N(G) .$$

*More precisely, there is some $c > 0$, such that $S(n) - NK(G)^n = o\big((K(G) - c)^n\big)$, i.e., such that*

$$\lim_{n\to\infty} \frac{S(n) - N(G)K(G)^n}{(K(G) - c)^n} = 0 .$$

**Corollary.** *With $G = ([r], E)$, $\rho$, and $f$ as above,*

$$\sum_{\mathbf{i}} \binom{n}{\mathbf{i}} \rho^{f(\mathbf{i})} \sim |\{B \in \mathcal{B}(G) : |B| = \alpha(G)\}| \cdot \alpha(G)^n .$$

*Proof outline.* As seen in the statement, the precise value of $\rho$ is unimportant. Indeed, the idea is to sum the "$\rho$-free" terms separately and to prove the asymptotic behaviour for them; and then to prove that the contribution of the other terms may be ignored. Thus, put $K = K(G)$, $N = N(G)$, and write

$$S(n) = T(n) + U(n) ,$$

where

$$T(n) = \sum_{\mathbf{i}: f(\mathbf{i})=0} \binom{n}{\mathbf{i}} \prod_{j=1}^{r} v_j^{n_j}$$

i.e, "$T(n)$ is $S(n)$ for the case $\rho = 0$". The refined claim is that

$$T(n) \sim NK^n , \tag{3}$$

$$T(n) - NK^n = o\big((K - c)^n\big) , \text{ and} \tag{4}$$

$$U(n) = o\big((K - c)^n\big) , \tag{5}$$

for some suitable constant $c > 0$.

The idea of the proof is first to prove (3) and (4) in full generality directly. With these facts established, clearly, the truth of the lemma for a particular weighted graph is equivalent to the claim (5) for that graph. Moreover, each $U(n) \geq 0$, whence upper estimates will be enough to establish the result for a particular graph, in either the lemma or the (5) formulation. We proceed to do this by induction with respect to $r$, working with the lemma in the induction step proof, but applying (5) in the inductive assumption.

For (3) and (4), note, that $f(\mathbf{i}) = 0$ if and only if $\operatorname{supp} \mathbf{i}$ is an independent set, i.e., the complement of the support is a vertex cover, i.e., if and only if there is some $C \in \mathcal{C}(G)$, such that

$$v_j \in C \implies i_j = 0 .$$

By the multinomial theorem, the sum of the terms in $T(n)$ taken over all $\mathbf{i}$ which are zero on a fixed $C \in \mathcal{C}$ of cardinality $r - t$, whence without loss of generality we may assume $C = \{t+1, \ldots, r\}$, is

$$\sum_{(i_1,\ldots,i_t)} \binom{n}{i_1, \ldots, i_t} \prod_{j=1}^{t} v_j^{i_j} = W^n ,$$

where $W = v_1 + \ldots v_r$ is the weight of the basis $V \setminus C$. The full sum $T(n)$ by means of the principle of inclusion-exclusion is expressible by $2^{|\mathcal{C}|} - 1$ similar terms, whence the growth rate is determined by the $c$ terms corresponding to the $\mathcal{B}(G)$ elements of maximal weight $r - s$. Each one of these terms equals $K^n$; all other terms will be of the form $\pm L^n$ with $L < K$, and thus will not contribute to the limit in (3). For (4), taking any $c$ such that $K - c$ is strictly greater than the *second* largest weight of independent sets will do.

Thus, indeed, (3) and (4) hold in full generality.

We now prove the lemma and equivalently (5) by induction with respect to $r$.

For $r = 1$, $E = \emptyset$, $\mathcal{B}(G) = \{[1]\}$, $K = v_1$, $N = 1$, and indeed $S(n) = T(n) = v_1^n = 1 \cdot K^n$.

Now, assume that $r \geq 2$, and that the lemma and (5) hold for each graph with a strictly smaller number of vertices. Fix an arbitrary weighted graph $G = ([r], E)$ with weights $\mathbf{v}$, and a $\rho$. In order to prove the lemma for this $G$ and $\rho$, we distinguish two cases, depending on the minimal valency in $G$.

First, if there is an isolated vertex in $G$, without loss of generality the last vertex, $r$, then $r$ is contained in each basis of $G$. Put $G' = ([r-1], E)$ with weights $\mathbf{v}' = v_1, \ldots, v_{r-1}$; then $N(G') = N$ and $K(G') = K - v_r$, for any $\mathbf{i} = \mathbf{i}', i_r$ we have $f'(\mathbf{i}') = f(\mathbf{i})$ for $f' = f_{G'}$, and by the inductive assumption there is a $c' > 0$, such that

$$S'(n') = S_{G',\mathbf{v}',\rho}(n') = \sum_{\mathbf{i}'} \binom{n'}{\mathbf{i}'} \rho^{f'(\mathbf{i}')} (\mathbf{v}')^{\mathbf{i}'} \sim N(K - v_r)^{n'} \tag{6}$$

and

$$S'(n') - N(K - v_r)^{n'} = o\big((K - v_r - c')^{n'}\big) , \tag{7}$$

for $n' = i_1 + \ldots + i_{r-1}$. Now, by (2), each term in $S(n)$ may be rewritten as a product of an $i_r$-depending and an $i_r$-independent factor: $\binom{n}{\mathbf{i}} \rho^{f(\mathbf{i})} \mathbf{v}^{\mathbf{i}} = \binom{n}{i_r} v_r^{i_r} \cdot \binom{n-i_r}{\mathbf{i}'} \rho^{f'(\mathbf{i}')} (\mathbf{v}')^{\mathbf{i}'}$, which sums up to

$$S(n) = \sum_{i_r} \binom{n}{i_r} v_r^{i_r} S'(n - i_r) ;$$

and by (1), indeed

$$\sum_{i_r} \binom{n}{i_r} v_r^{i_r} N(K - v_r)^{n - i_r} = NK^n .$$

It just remains to give an upper estimate of the growth of $\sum_{i_r}\binom{n}{i_r}v_r^{i_r}(S'(n-i_r)-N(K-v_r)^{n-i_r})$. Fix any $c$ with $0<c<\min(c',K-v_r)$; it is clearly sufficient to prove that for any $\varepsilon>0$, there is an $n_0$, such that for any $n>n_0$ we have

$$\sum_{i_r}\binom{n}{i_r}v_r^{i_r}(S'(n-i_r)-N(K-v_r)^{n-i_r})<\varepsilon(K-c)^n\,. \tag{8}$$

Now, by (7), there is an $n_0'$, such that

$$S'(n')-N(K-v_r)^{n'}<\frac{\varepsilon}{2}(K-v_r-c')^{n'}<\frac{\varepsilon}{2}(K-v_r-c)^{n'}$$

for $n'\geq n_0'$. Hence, certainly, for $n>n_0'$, and putting $\nu=n-i_r$,

$$\begin{aligned}&\sum_{i_r}\binom{n}{i_r}v_r^{i_r}(S'(n-i_r)-N(K-v_r)^{n-i_r})\\ &\quad<\sum_{i_r=0}^{n-n_0'}\binom{n}{i_r}v_r^{i_r}\tfrac{\varepsilon}{2}N(K-c-v_r)^{n-i_r}) \;+\sum_{i_r=n-n_0'+1}^{n}\binom{n}{i_r}v_r^{i_r}S'(n-i_r)\\ &\quad=\sum_{i_r=0}^{n}\binom{n}{i_r}v_r^{i_r}\tfrac{\varepsilon}{2}N(K-c-v_r)^{n-i_r}) \;+\sum_{\nu=0}^{n_0'-1}\binom{n}{\nu}v_r^{-\nu}S'(\nu)\cdot v_r^n\\ &\quad=\tfrac{\varepsilon}{2}(K-c)^n+g(n)v_r^n\,.\end{aligned}$$

for a polynomial $g$ of degree $n_0'-1$. We thus have established "half of (8)"; for the other half, note that $v_r=K-(K-v_r)<K-c$, whence indeed

$$g(n)<\tfrac{\varepsilon}{2}\left(\frac{K-c}{v_r}\right)^n$$

for all $n$ greater than some $n_0''$. Putting $n_0=\max(n_0',n_0'')$ directly yields (8), and we are through with the induction step in this case.

Finally, we should prove the lemma and (5) for $G$, in case $G$ has no isolated vertices; i.e., if $\bigcup E=[r]$. In other words, each $i_j$ appears as a factor in at least one of the terms in $f(\mathbf{i})$. By choosing one such term for each $j$, and noting that no term could be chosen more than twice, for any natural numbers $m$ and $n$ we get

$$f(\mathbf{i})\geq mn \text{ for all } \mathbf{i} \text{ with } \min_j \mathbf{i}\geq 2m \text{ and } \textstyle\sum_j i_j=n\,. \tag{9}$$

Employing that $\rho<1$, we may apply this for some fixed $m$, such that $\rho^m<K\left(\sum_j v_j\right)^{-1}$. With this choice, we shall show that $U(n)$ is bounded by a sum of $2mr+1$ terms, such that each one of the resulting $2mr+1$ sequences satisfies an (8) type of bound. By summing (and since $U(n)$ is non-negative), indeed then (5) follows.

Thus, note that each $\mathbf{i}$ either has $\min\mathbf{i}\geq 2m$, or for at least one $j$ has $i_j\in\{0,\dots,2m-1\}$. Thus,

$$U(n)\leq W(n)+\sum_{j=1}^{r}\sum_{k=0}^{2m-1}W_{jk}(n)$$

where

$$W(n)=\sum_{\min\mathbf{i}\geq 2m}\binom{n}{\mathbf{i}}\rho^{f(\mathbf{i})}\mathbf{v}^{\mathbf{i}}\,,$$

while, for each $j$ and $k$,

$$W_{jk}(n) = \sum_{\substack{i_j=k \\ f(\mathbf{i})>0}} \binom{n}{\mathbf{i}} \rho^{f(\mathbf{i})} \mathbf{v}^{\mathbf{i}} .$$

By the assumption for $m$, indeed $a^m(\sum_j v_j) < K - c$ for some positive $c$, by (1) yielding

$$W(n) < \sum_{\mathbf{i}} \binom{n}{\mathbf{i}} \rho^{mn} \mathbf{v}^{\mathbf{i}} = \big(\rho^m(v_1 + \dots v_r)\big)^n = o\big((K-c)^n\big) .$$

For the remaining series, without loss of generality, it is enough to consider the $W_{rk}$. Put $G' = ([r-1], E')$ and $f' = f_{G'}$, where $E' = E \cap \binom{[r-1]}{2}$. We consider a slightly modified vertex weighting $\mathbf{v}' = (v_1', \dots, v_{r-1}')$, defined as follows:

$$v_j' = \begin{cases} \rho^k v_j & \text{if } \{j,r\} \in E; \\ v_j & \text{else} \end{cases} .$$

Put $K' = K(G, \mathbf{v}')$, $N' = N(G, \mathbf{v}')$, and choose a $c' > 0$, such that indeed $S'(n) - N'(K')^n = o\big((K'-c')^n\big)$, with $S'$ as in (6). Finally, define $T'$ and $U'$ in analogy with $T$ and $U$, such that $S'(n) = T'(n) + U'(n)$.

Treat $k = 0$ separately:

— For $k = 0$, $\mathbf{v}'$ is the restriction of $\mathbf{v}$; we may have $K' = K$; but, even in this case,

$$W_{r0}(n) = U'(n) = o\big((K'-c')^n\big)$$

is small enough.

— For $k > 0$, since any basis of maximal weight in $G$ must contain either $r$ or at least one neighbour $j$ of $r$, and since $v_r > 0$ and $v_j > v_j'$, respectively, we get $K' < K$. Moreover, applying the definitions and (2),

$$W_{rk}(n) = v_r^k \binom{n}{k} S'(n-k) .$$

Thus, choosing a $c > 0$ such that $K - c > K'$, indeed

$$W_{rk}(n) = \mathcal{O}\big(n^r (K')^n\big) = o\big((K-c)^n\big) .$$

Thus (5) is proven in this case too, whence the induction step and thus the entire lemma is proven.

**Remark.** The various factors $v_j^{n_j}$ and $\rho^{n_i n_j}$ may be treated in a more uniform manner, by noting that they may be rewritten $\exp((\ln v_j)n_j)$ and $\exp((\ln \rho)n_i n_j)$, respectively. This motivates the following slightly extended formulation:

Let $h(x_1, \dots, x_r)$ be an (in general inhomogenic) quadratic form, $h(\mathbf{x}) = b + \sum_i d_i x_i + \sum_{i<j} e_{ij} x_i x_j$, with no variable squares terms, and such that all the second degree coefficients $e_{ij}$ are nonpositive. Define a weighted graph $G = ([r], E, \psi)$ by

$$E = \big\{\{i,j\} : e_{ij} < 0\big\} ,$$

and

$$\psi(j) = e^{d_j} .$$

Then, with $N(G)$ and $K(G)$ as before, the sum

$$S(n) = \sum_{\mathbf{i}} \binom{n}{\mathbf{i}} e^{h(\mathbf{i})} \sim e^b N(G) K(G)^n ,$$

and there is a $c > 0$ with $S(n) - e^b N(G) K(G)^n = o\big((K(G)-c)^n\big)$.

Actually, as an extension, this is rather virtual. The conclusions in the lemma do not depend on $\rho$; and (apart from the negligible constant $e^b$) the $S(n)$ in the extension are squeezed between the lemma items with $\rho = \exp(\min e_{ij})$, and those with $\rho = \exp(\max e_{ij} : e_{ij} \neq 0)$.