Symbolic Moment Calculus I.: Foundations and Permutation Pattern Statistics

Doron ZEILBERGER [1]

Abstract: The old workhorse called *linearity of expectation* , by which it is often very easy to compute the expectation (alias first moment) of interesting combinatorial quantities, can also be used to compute higher moments, as well as correlations, but then things get very soon too complicated for mere humans. The computer, used as a symbol-cruncher, can go much further (alas, only by a few orders of magnitude). In this article, a methodology for using Computer Algebra Systems to *automatically* derive higher moments of interesting combinatorial random variables is described, and applied to Pattern Statistics of permutations. This would be hopefully followed by sequels applied to other combinatorial objects like graph-colorings, Boolean functions, and Random Walks. In addition to the intrinsic interest that the actual results might have, it is also hoped that this work will serve as as an *illustrative example* for conducting reliable mathematical experiments that lead to completely rigorous results, by an Overlapping Stages approach.

A. FOUNDATIONS

Introduction

In order to make this article self-contained, I will start with a very elementary, and familiar to most readers, review of one of the most powerful tools of combinatorics, *linearity of expectation.*

Linearity of Expectation

One of the cornerstones of modern combinatorics, and hence of modern mathematics, is the *Problistic Method*, launched by Erdös and beautifully described in Noga Alon and Joel Spencer's magnum opus[AS]. The probablistic method, in its simplest form, may be encapsulated by:

$$E[X] < 1 \quad \Rightarrow \quad P(X = 0) > 0 \quad , \qquad (Paul)$$

and in order to use it, one must, of course, first compute the expectation, $E[X]$, explicitly.

How NOT to compute E[X]

First find a closed-form formula for $P(X = i)$, in terms of i and the integer parameter, n, say, describing the size of the family. Then simplify the sum

$$E[X] = \sum_{i=0}^{\infty} iP(X = i) \quad .$$

[1] Department of Mathematics, Rutgers University (New Brunswick), Hill Center-Busch Campus, 110 Frelinghuysen Rd., Piscataway, NJ 08854-8019, USA. `zeilberg@math.rutgers.edu` , `http://www.math.rutgers.edu/~zeilberg/`.
First version: Jan. 11, 2004. This version: Nov. 8, 2004. Accompanied by the Maple package `SMCper`, available from Zeilberger's homepage. Supported in part by the NSF.

Unless you have a very simple problem, this approach is doomed to failure, since most often even computing $P(X=0)$ is intractable, let alone $P(X=i)$. But even if you managed to find a closed-form expression for $P(X=i)$, performing the summation itself would be most likely intractable.

The Right Way to compute E[X]

In many applications X is a sum of *indicator* random variables whose range is $\{0,1\}$

$$X = \sum_{i=1}^{N} X_i \quad ,$$

and for which $E[X_i]$ are all the same, (or otherwise partitionable into easy families), and trivially computable, since $E[X_i] = P(X_i = 1)$. Then by the *linearity of expectation* we have

$$E[X] = \sum_{i=1}^{N} E[X_i] \quad ,$$

which in the former case, equals $NE[X_1]$.

Some Illustrious Examples

1. The 'sample space' consists of the 2^n sequences of length n of independent coin-tosses ('Bernoulli trials') with $P(H) = p$, and X is the number of Heads. If $X_i(a) = 1$ or 0 according to whether the i^{th} toss is an H, then $E[X_i] = P(X_i = 1) = p$ $(1 \le i \le n)$ and $E[X] = np$. Note that in this extremely trivial and extremely classical case, the 'hard way' is still doable, but takes much longer (at least if you do it 'from scratch').

2. (The 'hat-check 'girl' problem'). The sample space is the set of permutations of $\{1, \ldots, n\}$, and for a permutation π, $X(\pi)$ is the number of its fixed points, i.e. the number of i for which $\pi(i) = i$. If X_i is the indicator random variable for the event $\pi(i) = i$, then obviously $X = \sum_{i=1}^{n} X_i$ and $E[X_i] = (n-1)!/n! = 1/n$, so $E[X] = nE[X_1] = 1$.

3. The sample space is the set of all the $2^{\binom{n}{2}}$ 2-colorings of the edges of the complete graph on n vertices, K_n, and the random variable X is the number of 'monochromatic K_k'. If for any k-subset S of the set of vertices, X_S is the indicator random variable for the event 'all the edges between the vertices of S are of the same color', then $E[X_S] = 2 \cdot (1/2)^{\binom{k}{2}}$, for any S, and since there are $\binom{n}{k}$ such S's we have $E[X] = \binom{n}{k}(1/2)^{\binom{k}{2}-1}$. This old chestnut, due to Erdös [Er], was the *genesis* of the probabilistic method.

4. This example is the starting point for the Pattern statistics of permutations described later in this paper.

The reduction of a sequence of distinct integers $(a_1, \ldots, a_k)$, to be denoted by $Reduce(a_1, \ldots, a_k)$, is the permutation obtained by replacing, for $i = 1, 2, \ldots, k$, the i-th smallest element by i. For example, $Reduce(5, 1, 4, 2, 8) = (4, 1, 3, 2, 5)$.

Let σ be a fixed permutation of size $\{1, \ldots, k\}$. Recall that a permutation π of $\{1, \ldots, n\}$ contains the pattern σ whenever there are $1 \leq i_1 < \ldots < i_k \leq n$ such that

$$Reduce(\pi(i_1), \ldots, \pi(i_k)) \;=\; \sigma \quad .$$

The sample space is the set of permutations of $\{1, \ldots, n\}$, and the random variable $X(\pi)$ is the number of σ-patterns π contains. For example if $\sigma = 132$ then $X(14325) = 3$, corresponding to the places 123, 124, and 134. For any k-subset $S = \{i_1, \ldots, i_k\}$ of $\{1, \ldots, n\}$ (where $1 \leq i_1 < \ldots < i_k \leq n$), X_S is the indicator random variable that is 1 or 0 according to whether or not $(\pi(i_1), \ldots, \pi(i_k))$ reduces to σ. Obviously $X = \sum_S X_S$ and $E[X_S] = 1/k!$ for each such S, hence $E[X] = \binom{n}{k}/k!$. Note that it is the same for all patterns of size k.

Calculating Higher Moments: The General Set-Up

Once again, the way *not* to do it is by using the 'definition'

$$E[X^r] = \sum_{i=0}^{\infty} i^r P(X = i) \quad .$$

Assuming, as above, that X can be written as a sum of indicator random variables

$$X = \sum_{i=1}^{n} X_i \quad ,$$

we have

$$X^r = \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \ldots \sum_{i_r=1}^{n} X_{i_1} X_{i_2} \cdots X_{i_r} \quad ,$$

and hence

$$E[X^r] = \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \ldots \sum_{i_r=1}^{n} E[X_{i_1} X_{i_2} \cdots X_{i_r}] \quad .$$

Alas, now things are much more complicated. It is no longer true that $E[X_{i_1} X_{i_2} \cdots X_{i_r}]$ are all the same, and these quantities depend in an intricate way on how the elementary events engendering $X_{i_1}, \ldots, X_{i_r}$ 'interact' with each other. However, if we invite computers to join in, then there is hope, at least for small r.

It is too much to ask, in all but the most trivial cases, to know *all* the moments $E[X^r]$, i.e. as an explicit expression in n and r, where n is the symbol denoting the size of the combinatorial family. This is because this knowledge is equivalent to knowing, *explicitly*, the full probability distribution, i.e. an explicit expression for $P(X = i)$, in terms of n and i, and this is (usually) intractable.

So what is a poor human to do? $E[X]$ is often almost trivial, and certainly doable, by humans. By enlisting the help of computers, we can hope to find explicit expressions in n, for $r = 2$ (that will

give us variance and correlations), and $r = 3$. This might help us improve lower bounds obtained by the naive inequality ($Paul$) by using the next-in-line in the Bonferroni inequality that may be rephrased as

$$1 - E[X] + E[\binom{X}{2}] - E[\binom{X}{3}] > 0 \quad \Rightarrow \quad P(X = 0) > 0 \quad .$$

Finally the case $r = 4$ will enable us to find the *Kurtosis*, a statistically interesting quantity.

Frustratingly, at least in the case of Pattern statistics described in this first installment, the fourth moment is already beyond the power of current computers, for patterns of length ≥ 4. So we do what we (and our computers) can, and try to find explicit (or at least asymptotic) expressions for low moments.

In order to do what we can, we proceed as follows. The 'compound event' corresponding to when $X_{i_1}X_{i_2}\cdots X_{i_r} = 1$ has a certain *support*, consisting of the 'vertices' that participate in it. (This is very vague, it will all become clear when we will do specific examples). Let's call this support S, a certain subset of the set of vertices. Each such situation is equivalent to a 'canonical configuration' in which the labels of S are as small as possible. It turns out that the total number of reduced configurations (for a fixed r) is finite (albeit usually very big). So in order to compute $E[X^r]$ we naturally get a finite set of reduced configurations, for each of which $E[X_{i_1}X_{i_2}\cdots X_{i_r}]$ is easy to compute. It is also easy to compute, for each reduced configuration, its 'multiplicity', i.e. the number of non-reduced configurations that are equivalent to it, that should be a *symbolic* expression in n (the parameter describing the size of the combinatorial family). If $\mathcal{C}$ is the set of *reduced configurations* then we have

$$E[X^r] = \sum_{c \in \mathcal{C}} multiplicity(c) \cdot E[c] \quad .$$

The problem reduces to *teaching* the computer how to construct the finite set $\mathcal{C}$ and to add up the finitely many expressions.

B. Permutation Pattern Statistics

Pattern Avoidance and Pattern Statistics

The topic of *pattern avoidance* is a very active field in current enumerative combinatorics. Its aim is to enumerate, or talk about, sets of permutations that *avoid* one or more specific patterns. See Miklos Bona's forthcoming book[Bo] for a fascinating and lucid overview. In terms of the present set-up, the goal is to find exactly (or at least bound) the probability that $X = 0$, where X is as in example **4** above, i.e. $X(\pi)$ is the number of occurrences of a given pattern (or patterns). We saw in that example that the first moment is trivially $\binom{n}{k}/k!$ (in the case of a single pattern). What about the higher moments? What about covariances and correlations between random variables enumerating different patterns? This is the purpose of the second, specific, part of this paper.

The General Formula

Our goal is to write a computer program that inputs r arbitrary patterns $\sigma_1, \sigma_2, \dots, \sigma_r$ (not necessarily distinct and not necessarily of the same size), and outputs the expression, in n, (that will turn out to be a polynomial, see below), for

$$E[X_{\sigma_1} \cdots X_{\sigma_r}]$$

over S_n. But this equals

$$\frac{1}{n!} \sum_{\pi \in S_n} X_{\sigma_1}(\pi) \dots X_{\sigma_r}(\pi) \quad .$$

The sum (without the $1/n!$ in front) enumerates the set of $(r+1)$-tuples

$$A(n; \sigma_1, \dots, \sigma_r) := \{ \quad [\pi; T_1, \dots, T_r] \quad \} \quad ,$$

where $\pi \in S_n$, and $T_1, T_2, \dots, T_r$ are subsets of $\{1, 2, \dots, n\}$ such that, for $i = 1, \dots, r$, $Reduce(\pi | T_i) = \sigma_i$, in other words, π contains the pattern σ_i at the places indicated by the set T_i.

Now let's change the order of summation! Let's start with an r-tuple of subsets of $\{1, \dots, n\}$, $[T_1, \dots, T_r]$ and ask how many permutations π are they such that $(\pi; T_1, \dots, T_r)$ is legal. First, let's define

$$T = T_1 \cup T_2 \cup \dots \cup T_r$$

to be the union of the T_i's. Let $t = |T|$. Going back to the full object $(\pi; T_1, \dots, T_r)$, we can form it in four stages.

First pick T, a subset of $\{1, \dots, n\}$ of cardinality t.

Second decide what are the values of $\pi(i)$ for $i \in T$, i.e. the *set*

$$\{\pi(i) \quad | \quad i \in T\} \quad .$$

Thirdly form a 'legal object' on T.

Finally worry about filling-in the entries $\pi(i)$ for $i \notin T$.

There are $\binom{n}{t}$ ways to do the first step, and $\binom{n}{t}$ ways to do the second step, and $(n-t)!$ ways to do the last step. Regarding the Third (and most interesting) step, the restriction of π to T is isomorphic to one in which $T = \{1, \dots, t\}$ and the set of values of π on T, $\{\pi(i) \quad | \quad i \in T\}$, is also $\{1, \dots, t\}$.

This naturally leads us to define the following set

$$\mathcal{A}(t; \sigma_1, \dots, \sigma_r) :=$$

$$\{ \quad [\pi; T_1, \dots T_r] \quad | \quad \pi \in S_t \quad , \quad T_1 \cup \dots \cup T_r \;=\; \{1, \dots, t\} \quad , \quad Reduce(\pi | T_i) = \sigma_i \quad (i = 1 \dots r) \quad \} \quad .$$

Note that for the fixed input parameters: the r patterns $\sigma_1, \ldots, \sigma_r$ and the integer t (which is $\leq |\sigma_1| + \ldots + |\sigma_r|$), this is a finite set, that can be constucted by the computer *explicitly*! Let the cardinality of $\mathcal{A}(t; \sigma_1, \ldots, \sigma_r)$ be denoted by $A(t; \sigma_1, \ldots, \sigma_r)$. Then we have the 'explicit' expression (where $L = |\sigma_1| + \ldots + |\sigma_r|$ is the sum of the sizes of the patterns):

$$E[X_{\sigma_1} \cdots X_{\sigma_r}] = \frac{1}{n!} \sum_{t=0}^{L} \binom{n}{t}^2 (n-t)! A(t; \sigma_1, \ldots, \sigma_r) = \sum_{t=0}^{L} \frac{1}{t!} \binom{n}{t} A(t; \sigma_1, \ldots, \sigma_r) \quad . \tag{MainFormula}$$

In particular, it follows that $E[X_{\sigma_1} \cdots X_{\sigma_r}]$ is a polynomial of degree L in n. This gives rise to the

Sampling Way for Computing $E[X_{\sigma_1} \cdots X_{\sigma_r}]$

By actually computing $E[X_{\sigma_1} \cdots X_{\sigma_r}]$ for $n = 0, \ldots, L$ we can fit the data into a polynomial of degree L. The drawback of this brilliant method is that we need to perform more than $L!$ operations.

If all the patterns are the same, and rather than computing $E[X_\sigma^r]$ one is interested in the more significant quantity $E[(X_\sigma - E[X_\sigma])^r]$ (the r^{th} moment about the mean), then this is a polynomial of degree $r|\sigma| - r + 1$.

Asymptotic Moments

The sets $\mathcal{A}(t; \sigma_1, \ldots, \sigma_r)$ are very complicated and it is unlikely that there is a 'fast' way to compute their cardinalities. Hence we have to resort to the naive way, using the *caveman's formula*

$$|A| = \sum_{a \in A} 1 \quad ,$$

i.e. naive counting, once the computer explicitly constructed the finite set A.

Whenever it is infeasible (or too expensive) to compute the *exact* formula, one tries, at least, to find an *asymptotic* formula for the desired quantity. In the present case, where the expressions are polynomials, this amounts to finding the leading term.

The leading term of $E[X_{\sigma_1} \cdots X_{\sigma_r}]$ is easy (but not interesting). The degree is $L = |\sigma_1| + \ldots + |\sigma_r|$, and $A(L; \sigma_1, \ldots, \sigma_r)$ is simply

$$\left(\frac{L!}{|\sigma_1|! \cdots |\sigma_r|!} \right)^2 \quad ,$$

since we must decide about the *places* and then about the *values*.

However, for the more interesting quantities, moments about the mean, or covariances, we need the asymptotics of expressions of the form

$$E[(X_{\sigma_1} - E[X_{\sigma_1}]) \cdots (X_{\sigma_r} - E[X_{\sigma_r}])],$$

which necessitates computing the first r leading terms of the polynomials describing the straight moments (because of cancellations). For covariances, things are still manageable, and we have (let

$a = |\sigma_1|$, $b = |\sigma_2|$, and $L = a + b$)

$$A(L{-}1; \sigma_1, \sigma_2) = \sum_{s=1}^{a} \sum_{t=1}^{b} \binom{s+t-2}{s-1} \binom{a+b-s-t}{a-s} \binom{\sigma_1(s)+\sigma_2(t)-2}{\sigma_1(s)-1} \binom{a+b-\sigma_1(s)-\sigma_2(t)}{a-\sigma_1(s)} \quad ,$$

obtained by considering the unique place where the support of σ_1 meets the support of σ_2.

A User Manual's for the Maple Package SMCper

All of this (and more!) is implemented in the Maple package `SMCper`. Like all my packages, it can be obtained from

`http://www.math.rutgers.edu/~zeilberg/programs.html` .

Assume that you have Maple, and that you have downloaded `SMCper` to your current directory. To use the package, go into Maple, and type: `read SMCper;`, then follow the on-line instructions.

The main procedures are

`Mishkal(ListPatterns, n)`: Given a list of patterns $ListPatterns = [\sigma_1, \dots, \sigma_r]$, say, and a symbol n, implements $(MainFormula)$ above.

`ACorr(Pat1,Pat2)`: For the asymptotic correlation of the patterns `Pat1` and `Pat2`.

`ACovari(Pat1,Pat2)`: For the asymptotic covariance (divided by n to the power `nops(Pat1)+nops(Pat2)-1`) of two patterns `Pat1` and `Pat2`.

`AllACorr(d)`: All the asymptotic Correlations for all pairs of patterns of size `d`, where only one pair from each equivalence class is displayed.

`AllACorrDiffSize(d1,d2)`: As above but with all pairs of patterns one of size `d1` and the other of size `d2`.

`AllAThirdMom(d)`: All the asymptotic third moments (about the mean), divided by n^{2d-1}, for patterns of length `d`.

`AThirdMom(Pat)`: The asymptotic third moment (about the mean), divided by n^{2d-1}, for the pattern `Pat`.

`AllAVariance(d)`: All the asymptotic variances for patterns of size `d` (divided by n^{2d-1}).

`EListES(Pats,n)`: Given a list of patterns, `Pats =[Pat1, ..., Patr]`,say, and a symbol `n`, computes the formula for the expectation, over S_n, of the random variable $(X_1 - E[X_1]) \cdots (X_r - E[X_r])$, where $X_i(\pi)$ is the number of occurrences of the pattern $Pati$ $(i = 1, \dots, r)$. It does that by using the empirical (but rigorous!) sampling method described above.

`kthMoment(Pat,k,n)`: The `k`-th moment (about the mean) of the random variable `#Pat` defined on S_n.

Some Computer-Generated Theorems

Explicit Expressions for Variances

Since the theorems below involve a *symbol* n, hence are true for infinitely many cases, they can be dignified by the appellation 'Theorem' rather than mere 'fact'. Except for Theorems **1-3**, I doubt that any humans can find them with their naked brains.

Theorem 1: The exact expression for the second moment (about the mean) (alias variance) of the random variable $X_{12}(\pi) :=$ number of occurrences of the pattern 12, over S_n, is

$$\frac{n(2n+5)(n-1)}{72} \quad .$$

Theorem 2: The exact expression for the second moment (about the mean) (alias variance) of the random variable $X_{123}(\pi) :=$ number of occurrences of the pattern 123, over S_n, is

$$\frac{n(n-1)(n-2)(39n^2+102n-157)}{21600} \quad .$$

Theorem 3: The exact expression for the second moment (about the mean) (alias variance) of the random variable $X_{213}(\pi) :=$ number of occurrences of the pattern 213, over S_n, is

$$\frac{n(n-1)(n-2)(21n^2+78n+77)}{21600} \quad .$$

Theorem 4: The exact expression for the second moment (about the mean) (alias variance) of the random variable $X_{1234}(\pi) :=$ number of occurrences of the pattern 1234, over S_n, is

$$\frac{n(n-1)(n-2)(n-3)(982n^3+3645n^2-8623n-2151)}{50803200} \quad .$$

Theorem 5: The exact expression for the second moment (about the mean) (alias variance) of the random variable $X_{1243}(\pi) :=$ number of occurrences of the pattern 1243, over S_n, is

$$\frac{n(n-1)(n-2)(n-3)(718n^3+1137n^2+1829n+13065)}{50803200} \quad .$$

Theorem 6: The exact expression for the second moment (about the mean) (alias variance) of the random variable $X_{1324}(\pi) :=$ number of occurrences of the pattern 1324, over S_n, is

$$\frac{n(n-1)(n-2)(n-3)(694n^3+125n^2+5345n+16729)}{50803200} \quad .$$

Theorem 7: The exact expression for the second moment (about the mean) (alias variance) of the random variable $X_{1342}(\pi) :=$ number of occurrences of the pattern 1342, over S_n, is

$$\frac{n(n-1)(n-2)(n-3)(478n^3+1601n^2+5237n+7369)}{50803200} \quad .$$

Theorem 8: The exact expression for the second moment (about the mean) (alias variance) of the random variable $X_{1423}(\pi) :=$ number of occurrences of the pattern 1423, over S_n, is

$$\frac{n(n-1)(n-2)(n-3)(478n^3+1601n^2+5237n+7369)}{50803200} \quad .$$

Theorem 9: The exact expression for the second moment (about the mean) (alias variance) of the random variable $X_{1432}(\pi) :=$ number of occurrences of the pattern 1432, over S_n, is

$$\frac{n(n-1)(n-2)(n-3)(526n^3+3821n^2-3559n+3961)}{50803200} \quad .$$

Theorem 10: The exact expression for the second moment (about the mean) (alias variance) of the random variable $X_{3142}(\pi) :=$ number of occurrences of the pattern 3142, over S_n, is

$$\frac{n(n-1)(n-2)(n-3)(26n^3+387n^2+895n+639)}{7257600} \quad .$$

Theorem 11: The exact expression for the second moment (about the mean) (alias variance) of the random variable $X_{2143}(\pi) :=$ number of occurrences of the pattern 2143, over S_n, is

$$\frac{n(n-1)(n-2)(n-3)(470n^3+2701n^2+2881n-295)}{50803200} \quad .$$

Explicit Expressions for Third Moments

Theorem 12: The exact expression for the third moment (about the mean) of the random variable $X_{123}(\pi) :=$ number of occurrences of the pattern 123, over S_n, is

$$\frac{n(n-1)(n-2)(1437n^4+5592n^3-11277n^2-33990n+34082)}{6350400} \quad .$$

Theorem 13: The exact expression for the third moment (about the mean) of the random variable $X_{132}(\pi) :=$ number of occurrences of the pattern 132, over S_n, is

$$\frac{n(n-1)(n-2)(129n^4+3705n^3+5355n^2+8655n+11356)}{12700800} \quad .$$

The Winners and Losers in the Variance Game

Theorem 14: The smallest and largest asymptotic variances (respectively) for patterns of length $3 \leq d \leq 6$ are:

$d=3$: 132 with $(7/7200)n^5+O(n^4)$; 123 with $(13/7200)n^5+O(n^4)$;

$d=4$: 2413 with $(13/3628800)n^6 + O(n^5)$; 1234 with $(491/25401600)n^6 + O(n^5)$.

$d=5$: 25314 with $(421/32920473600)n^7 + O(n^6)$; 12345 with $(1171/21946982400)n^7 + O(n^6)$.

$d=6$: 254163 with $(1493/88519495680000)n^8 + O(n^7)$; 123456 with $(829/9484231680000)n^8 + O(n^7)$.

The Most Hostile Pairs

Theorem 15: The least correlated pairs of patterns for patterns of length $3 \leq d \leq 6$, followed by the asymptotic correlations are:

$d=2$: 12 and 21, -1.

$d=3$: 123 and 321, $-12/13 = -.9230769231\ldots$.

$d=4$: 1234 and 3412, $-313/115385 \cdot \sqrt{115385} = -.9214455145\ldots$.

$d=5$: 21354 and 45312, $-1235/1321 = -.9348978047\ldots$.

$d=6$: 213465 and 564312, $-22279/24336 = -.9154750164\ldots$.

Sample Input and output

Lots of more output can be found at the webpage of this article

`http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/smcI.html` .

The Methodology of Overlapping Stages

Experimental Mathematics, as it is commonly understood (see the recent fascinating and revolutionary book by Jon Borwein and David Bailey[BB]), engages in *number crunching* that enables one to formulate and test conjectures that later are to be hopefully followed by human-made 'analytical' proof. On the other hand, the kind of experimental math that this paper narrates involves **symbol-crunching**, and as such its outputs are already theorems, complete with proofs.

Computers are infallible, but humans are not! While my beloved computer, Shalosh B. Ekhad, faithfully executed the Maple program, the program itself was written by yours truly, who while competent and reliable as far as humans go, is nevertheless a lowly human. How do we know that we can trust the output?

One way is just read the program and check that its logic is correct. Hence all the dictums of reliable programming, i.e. modularity, clarity, lots of comments, etc. etc. should be followed in order to facilitate this human checking. But in addition, and more importantly, one should test the program by running sample input, against independent, more basic programs.

Hence one must proceed in *stages* that overlap. Suppose that you want to find a formula $A(n,d)$

for symbolic n but specific (numeric) d. First write a completely naive, exponential-time, program, whose logic is very easy to check. Alas, it blows up for $d \geq 4$, say. Then write a more sophisticated version, $B(n,d)$, and make sure that the outputs coincide for $d \leq 4$. Now you can trust $B(n,d)$, but while it is more efficient than $A(n,d)$, it too explodes for say $d \geq 11$. Then write yet a more sophisticated version $C(n,d)$, valid for $d \leq 20$, say, and test it against $B(n,d)$ in the overlapping range of applicability, $d \leq 10$. Etc. etc.

This methodology was pursued in this project. First I wrote a 'naive' program to compute symbolic moments (`EListES` of SMCper). Then I tested it against the combinatorial symbolic procedure, `Mishkal`, that implements $(MainFormula)$. Then I used this, in turn, to test the procedure for the asymptotic variances and covariances.

Acknowledgement: I wish to thank Thotsaporn Thanatipanonda for numerous corrections.

REFERENCES

[AS] N. Alon and J.H. Spencer, **The Probabilistic Method**, Second Edition, Wiley, New York, 2000.

[Bo] M. Bona, **Combinatorics of Permutations**, Chapman-Hall and CRC Press, Boca Raton, to appear.

[BB] J. Borwein and D. Bailey, **Mathematics by Experiment**. A.K. Peters, Natick, 2004.

[Er] P. Erdös, *Some remarks on the theory of graphs*, Bulletin of the Amer. Math. Soc. **53** (1947), 292-294.