Published in final edited form as:

Commun Stat Theory Methods. 2008 January 1; 37(18): 2815–2824. doi:10.1080/03610920802001896.

Fast Computation by Block Permanents of Cumulative Distribution Functions of Order Statistics from Several Populations

D. H. GLUECK¹, A. KARIMPOUR-FARD¹, J. MANDEL¹, L. HUNTER¹, and K. E. MULLER²

¹Department of Preventive Medicine and Biometrics, University of Colorado Denver, Denver, Colorado, USA

²Department of Epidemiology and Health Policy Research, University of Florida, Gainesville, Florida, USA

Abstract

The joint cumulative distribution function for order statistics arising from several different populations is given in terms of the distribution functions of the populations. The computational cost of our formula in the case of two populations is still exponential in the worst case, but it is a dramatic improvement compared to the general formula by Bapat and Beg. In the case when only the joint distribution function of a subset of the order statistics of fixed size is needed, the complexity is polynomial, for the case of two populations.

Keywords

Block matrix; Computational complexity; Multiple comparison

1. Introduction

We provide results that allow the computation of marginal or joint cumulative distribution functions of order statistics. The order statistics arise from independent but not identically distributed random variables. The computation complexity is reduced in the case when the random variables only have a small number of distributions, rather than a different one for each random variable.

An important application is in calculating power for the Benjamini and Hochberg (1995) procedure. While Glueck et al. (2008) suggested calculating power by integration of the probability density functions of the order statistics, an alternative approach is to calculate the joint cumulative distribution functions of the order statistics as sums of marginal probability

Copyright © Taylor & Francis Group, LLC

Address correspondence to D. H. Glueck, Department of Preventive Medicine and Biometrics, University of Colorado at Denver, Denver, CO 80262, USA; Deborah.Glueck@uchsc.edu.

Publisher's Disclaimer: Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

distributions of high order, using a common set probability theorem (Stirzaker, 2006, p. 35). In power analysis for the Benjamini and Hochberg procedure, the p -values often arise by sampling from two different populations, namely experiments where the null or the alternative hypothesis is true, when all null hypotheses and all alternative hypotheses are the same. Another case when p -values from two populations arise is sampling from data about men or women, or from information about two different types of cancer. The p -values may also arise from sampling from more than two populations, as is in the case when several new drugs are compared (David and Nagaraja, 2003).

In the p -outlier model, there are also p -values from two populations. The second population is where the outlier comes from. However, the focus in the p -outlier theory is different than here. Arnold and Balakrishnan (1989) provided a comprehensive treatment of samples containing a single outlier, with results about moments and functions of the order statistics. Bapat (1990) discussed the multiparameter multinomial model, which produces a similar block structure as our assumption. Childs and Balakrishnan (2006) used results on the moments of order statistics for logistic random variables with p -outliers to compute the bias of various linear estimators of the location and scale parameters of the logistic distribution. Balakrishnan (2007) provided results for the moments of order statistics for p -outlier location and scale models.

For independent, but not identically distributed random variables, Cao and West (1997) gave a recurrence relationship for the distribution function of a single-order statistic. A general formula for computing the joint cumulative distribution function of the order statistics for the independent but not identically distributed case was given by Bapat and Beg (1989). If each order statistic has a different distribution, the result is computationally intractable, because it involves an exponential number of permanents. The complexity of the computation of the permanent by the best algorithms grows exponentially (Knuth, 1998, p. 499). Approximate algorithms for computing the permanent (Forbert and Marx, 2003; Jerrum et al., 2004; Valiant, 1979) with lower asymptotic complexity are still not practical.

The distribution of order statistics for independent, identically distributed random variables is well known, and appears in every basic statistics book; for example, Hogg and Craig (1978, Ch. 4, Sec. 6). David and Nagaraja (2003) and Balakrishnan and Rao (1998) provide a thorough review of order statistics. For identically distributed random variables, the cumulative distribution function is concise and fast to compute.

The computational cost of the formula in this article for the case of two populations is still exponential, but is a dramatic improvement compared to the general formula by Bapat and Beg. In the case when only the joint distribution function of a subset of the order statistics of fixed size is needed, the complexity of the computation is polynomial, in the case of two populations.

2. Notation and Preliminaries

For an $m \times m$ matrix A , with entries a_{ij} , the permanent is given by Aitken (1939, p. 30):

$$\text{per}[A] = \sum_{\pi} \prod_{i=1}^m a_{i,\pi(i)}, \quad (1)$$

where π ranges over all permutations of the set $\{1, 2, \dots, m\}$. Hence, the permanent is defined much like the determinant, but with all signs positive. The permanent can be expanded by row or columns exactly like the determinant. The computational cost of evaluating the permanent

by expansion is $\text{const } mm!$ operations. The computational cost using the best known algorithms is exponential (Knuth, 1998, p. 499).

The following notation will be used in all theorems and proofs in this article without further explicit reference. $X_i, i = 1, \dots, m$, are independent, real-valued random variables with cumulative distribution functions $F_i(x)$. The order statistics $Y_1, Y_2, \dots, Y_m$ are random variables defined by sorting the values of X_i . In particular, $Y_1 \leq Y_2 \leq \dots \leq Y_m$. The arguments of the joint cumulative distribution function of order statistics are customarily written omitting redundant arguments; thus, let $1 \leq n_1 < n_2 < \dots < n_k \leq m$ denote the indices of the remaining arguments and $y_1 \leq y_2 \leq \dots \leq y_k$ their values. The joint cumulative distribution function of the set $\{Y_{n_1}, Y_{n_2}, \dots, Y_{n_k}\}$, which is a subset of the complete set of order statistics, is defined as:

$$F_{Y_{n_1}, \dots, Y_{n_k}}(y_1, \dots, y_k) = \Pr \{(Y_{n_1} \leq y_1) \wedge (Y_{n_2} \leq y_2) \wedge \dots \wedge (Y_{n_k} \leq y_k)\}.$$

Define the index vector $\mathbf{i} = (i_0, i_1, \dots, i_{k+1})$ and the summation index set

$$\mathcal{J} = \left\{ \mathbf{i}: \begin{array}{l} 0 = i_0 \leq i_1 \leq \dots \leq i_k \leq i_{k+1} = m, \\ \text{and } i_j \geq n_j \text{ for all } 1 \leq j \leq k \end{array} \right\}$$

and denote the number of elements of the index set $\mathcal{J}$ by:

$$|\mathcal{J}| = v(n_1, n_2, \dots, n_k; m) = \sum_{i_k=n_k}^m \sum_{i_{k-1}=n_{k-1}}^{i_k} \dots \sum_{i_1=n_1}^{i_2} 1.$$

Note that writing summation over the set $\mathcal{J}$ in terms of loops is straightforward. Using the set $\mathcal{J}$ instead of the loops in this article allows an insight into the structure of the method and its complexity, and it does not tie the mathematical formulation to any particular implementation.

For two sequences a_m and b_m , let $a_m \sim b_m$ denote $\lim_{m \rightarrow \infty} a_m/b_m = 1$. Let const be a generic positive constant independent of m ; that is, const can have a different value every time it is used.

3. Joint Cumulative Distribution Function of Order Statistics

The distribution of the order statistics of a random sample where each sample member is taken from a possibly different population with its own distribution was given by Bapat and Beg (1989).

Theorem 3.1

(Bapat and Beg, 1989). The cumulative distribution function of the order statistics satisfies

$$F_{Y_{n_1}, \dots, Y_{n_k}}(y_1, \dots, y_k) = \sum_{\mathbf{i} \in \mathcal{J}} \frac{P_{i_1, \dots, i_k}(y_1, \dots, y_k)}{(i_1 - i_0)! (i_2 - i_1)! \dots (i_{k+1} - i_k)!}, \quad (2)$$

where $P_{i_1, \dots, i_k}(y_1, \dots, y_k)$ are permanents of block matrices

$$P_{i_1, \dots, i_k}(y_1, \dots, y_k) = \text{per} \left[\left[F_i(y_j) - F_i(y_{j-1}) \right]_{(i_j - i_{j-1}) \times 1} \right]_{j=1, i=1}^{j=k, i=m} \quad (3)$$

with the subscripts indicating the dimensions of blocks created by the repetition of the term in the brackets, and $F_i(y_0) = 0, F_i(y_{k+1}) = 1$.

Now consider drawing a random sample from two populations, each with a different cumulative distribution function, say $F(x)$ and $G(x)$. Sample the first n random variables from the first population with the distribution function F , and then $m - n$ from the second population with the distribution function G . This special form of the distributions allows us to evaluate the joint distribution of the order statistic more efficiently.

Theorem 3.2

Suppose that $F_i(x) = F(x)$, for all $1 \leq i \leq n$, and $F_i(x) = G(x)$, for all $n + 1 \leq i \leq m$. Then:

$$F_{y_{n_1}, \dots, y_{n_k}}(y_1, \dots, y_k) \\ = \sum_{\mathbf{i} \in \mathcal{J}} \sum_{\lambda} \prod_{j=1}^{k+1} \frac{n!(m-n)!}{\lambda_j!(i_j - i_{j-1} - \lambda_j)!} \left[F(y_j) - F(y_{j-1}) \right]^{\lambda_j} \left[G(y_j) - G(y_{j-1}) \right]^{i_j - i_{j-1} - \lambda_j},$$

where $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_{k+1})$ ranges over all integer vectors such that

$$\lambda_1 + \lambda_2 + \dots + \lambda_{k+1} = n, \quad 0 \leq \lambda_j \leq i_j - i_{j-1}, \quad (4)$$

and $F(y_0) = G(y_0) = 0, F(y_{k+1}) = G(y_{k+1}) = 1$.

Proof—The permanents (3) simplify in the case of two population to:

$$P_{i_1, \dots, i_k}(y_1, \dots, y_k) \\ = \text{per} \begin{bmatrix} [F(y_1) - F(y_0)]_{(i_1 - i_0) \times n} & [G(y_1) - G(y_0)]_{(i_1 - i_0) \times (m-n)} \\ [F(y_2) - F(y_1)]_{(i_2 - i_1) \times n} & [G(y_2) - G(y_1)]_{(i_2 - i_1) \times (m-n)} \\ \vdots & \vdots \\ [F(y_{k+1}) - F(y_k)]_{(i_{k+1} - i_k) \times n} & [G(y_{k+1}) - G(y_k)]_{(i_{k+1} - i_k) \times (m-n)} \end{bmatrix} \quad (5)$$

Let $S_1 = \{1, 2, \dots, n\}$ and $S_2 = \{n + 1, n + 2, \dots, m\}$. Write a permutation of $\{1, 2, \dots, m\}$ as $\pi = (\pi_1, \pi_2, \dots, \pi_k, \pi_{k+1})$, where each subsequence π_j has exactly $i_j - i_{j-1}$ terms. The subsequence π_j is a list of the subscripts of the random variables that fall in the interval (y_{j-1}, y_j) . Then the term in the definition of the permanent (1) associated with π is:

$$\prod_{i=1}^m a_{i, \pi(i)} = \prod_{j=1}^{k+1} \left[F(y_j) - F(y_{j-1}) \right]^{\lambda_j} \left[G(y_j) - G(y_{j-1}) \right]^{i_j - i_{j-1} - \lambda_j},$$

where λ_j is the number of random variables with subscripts listed in $\{\pi_j\}$ that are in S_1 .

The number of permutations π such that λ_j is the number of the elements from $\{\pi_j\}$ that are in S_1 is:

$$\left(\frac{n!}{\prod_{j=1}^{k+1} \lambda_j!} \right) \left(\frac{(m-n)!}{\prod_{j=1}^{k+1} (i_j - i_{j-1} - \lambda_j)!} \right) \left(\prod_{j=1}^{k+1} (i_j - i_{j-1})! \right),$$

where the first term is the number of ways to distribute the n elements of S_1 so that set j has λ_j elements (the multinomial coefficient), the second term is the number of ways to distribute the $m - n$ elements of S_1 so that set j has $i_j - i_{j-1} - \lambda_j$ elements, and the last term is the number of permutations that do not change the distribution of the elements S_1 and S_2 into those sets. Thus:

$$\begin{aligned} P_{i_1, \dots, i_k}(y_1, \dots, y_k) &= \sum_{\pi} \prod_{i=1}^m a_{i, \pi(i)} \\ &= \sum_{\lambda} \prod_{j=1}^{k+1} \frac{(i_j - i_{j-1})!}{\lambda_j! (i_j - i_{j-1} - \lambda_j)!} \left[F(y_j) - F(y_{j-1}) \right]^{\lambda_j} \\ &\quad \times \left[G(y_j) - G(y_{j-1}) \right]^{i_j - i_{j-1} - \lambda_j}, \end{aligned}$$

with the sum over all λ that satisfy (4), and it remains to use Theorem 3.1.

We now consider the general case of order statistics of a sample selected from an arbitrary number of populations. The proof of Theorem 3.2 easily carries over, so the proof can be omitted.

Theorem 3.3

Suppose that $F_i = G_1$ for the first m_1 indices i , $F_i = G_2$ for the next m_2 indices i , etc., and $F_i = G_N$ for the last m_N indices i , with

$$m_1 + \dots + m_N = m, \quad m_s > 0 \text{ for all } s.$$

Then:

$$F_{y_{n_1}, \dots, y_{n_k}}(y_1, \dots, y_k) = \sum_{\mathbf{i} \in \mathcal{I}[\lambda_{js}]} \sum_{\lambda_{js}} \prod_{j=1}^{k+1} \prod_{s=1}^N \frac{m_s!}{\lambda_{js}!} \left[G_s(y_j) - G_s(y_{j-1}) \right]^{\lambda_{js}},$$

where the summation is over all integer matrices $[\lambda_{js}]$ size $k+1$ by N such that:

$$\begin{aligned} \lambda_{js} &\geq 0 \text{ for all } j \text{ and all } s, \\ \sum_{j=1}^{k+1} \lambda_{js} &= m \text{ for all } s, \\ \sum_{s=1}^N \lambda_{js} &= i_j - i_{j-1} \text{ for all } j, \end{aligned}$$

and $G_s(y_0) = 0$, $G_s(y_{k+1}) = 1$.

In the particular case when all $m_i = 1$, i.e., every distribution is different because it comes from a different population, Theorem 3.3 gives exactly the same result as Theorem 3.1. With two populations, the complexity of Theorem 3.3 reduces to the complexity of Theorem 3.2, which is less than that of Theorem 3.1, as discussed in the next section.

4. Complexity

We will now compare the relative complexity of Bapat and Beg (1989, Theorem 4.2, p. 84) and the formula from Theorem 3.2. We assume that the evaluation of the cumulative distribution function of each of the statistics takes a constant number of operations.

Lemma 4.1

The number $\mathfrak{Q} = v(n_1, n_2, \dots, n_k; m)$ of permanents in Theorems 3.1–3.3 is bounded by:

$$v(n_1, n_2, \dots, n_k; m) \leq v(1, 2, \dots, k; m) \leq v(1, 2, \dots, m; m) = C_m, \quad (6)$$

where

$$v(1, 2, \dots, k; m) = \binom{m+k}{k} \left(1 - \frac{k}{m+1}\right), \quad (7)$$

and

$$C_m = \frac{1}{m+1} \binom{2m}{m} = \frac{(2m)!}{(m+1)!m!}. \quad (8)$$

Proof. The inequalities in (6) are obtained by taking the smallest numbers for $n_1, n_2, \dots, n_k$ and the largest possible value for k , which both give the largest number of terms. It can be shown, by induction on k , that:

$$v(1, 2, \dots, k; m) = \binom{m+k}{k} - \binom{m+k}{k-1}. \quad (9)$$

Equations (7) and (8) follow from (9) by a direct computation:

$$\begin{aligned} \binom{m+k}{k} - \binom{m+k}{k-1} &= \frac{m+k}{1} \frac{m+k-1}{2} \dots \frac{m+2}{k-1} \frac{m+1}{k} - \frac{m+k}{1} \frac{m+k-1}{2} \dots \frac{m+2}{k-1} \\ &= \binom{m+k}{k} \left(1 - \frac{k}{m+1}\right), \end{aligned}$$

and

$$\binom{m+m}{m} - \binom{m+m}{m-1} = \binom{2m}{m} \left(1 - \frac{m}{m+1}\right) = \frac{1}{m+1} \binom{2m}{m},$$

which concludes the proof.

The numbers C_m defined by (8) are known as the Catalan numbers (Stanley, 1999), and the numbers $a_{k,m} = v(1, 2, \dots, k; m)$ are called the Catalan triangle (Shapiro, 1976). From the Stirling approximation $m! \sim \sqrt{2\pi m} m^m / e^m$, the growth of Catalan numbers is exponential:

$$c_m \sim \text{const } m^{-3/2} 4^m > \text{const } \alpha^m,$$

for any $1 < \alpha < 4$ (with a different const for each α).

Theorem 4.1

THE worst case complexity of computing the distribution function of the order statistics from Theorem 3.1 is:

$$\text{const } C_m m P(m) \sim \text{const } m^{-1/2} 4^m P(m),$$

where $P(m)$ is the number of operations for computing permanent of order m .

Proof—The denominator in Theorem 3.1 requires at most $\text{const } m$ operations, and there are at most C_m terms in the sum by Lemma 4.1.

It is known that the complexity of computing the permanent is bounded by:

$$P(m) \leq \text{const } m^a 2^m$$

for some $a > 0$, e.g., from the Ryser's formula (Knuth, 1998). So, the complexity of the computation of the distribution function from Theorem 3.1 is superexponential in m . Therefore, the computation is practical only for small m .

Fortunately, a drastic reduction of complexity is possible in the case when the order statistics come from two populations. In fact, the complexity reduces still farther when we need only a small number k of order statistics.

Theorem 4.2

Let $C(k, m, n)$ be the number of operations in Theorem 3.2 to evaluate the joint distribution function of k order statistics from m random variables from two populations, with $n \leq m$ of the variables from the first population. Then:

$$C(k, m, n) \leq \text{const } k \binom{m+k}{k} \binom{n+k}{k} \left(1 - \frac{k}{m+1}\right). \quad (10)$$

In the worst case over all k and n , the complexity is bounded by:

$$C(k, m, n) \leq \text{const } m \frac{(2m)^2}{(m!)^4} \sim \text{const } 16^m. \quad (11)$$

For any fixed k , the complexity is bounded by:

$$C(k, m, n) \leq \text{const } m^k n^k \leq \text{const } m^{2k}. \quad (12)$$

In particular, for a fixed k , the complexity is polynomial in m .

Proof—The complexity is bounded by $\text{const } CLM$, where $C = \binom{m+k}{k} \left(1 - \frac{k}{m}\right)$ is the number of terms in the sum over $\mathbf{i}$, L is the number of possible index vectors λ satisfying (4), and M is the complexity of evaluating the products in one term of the sum, which is $M \leq \text{const } k$. To bound L , drop the upper bounds in (4). Thus, L is bounded above by the number of all integer vectors λ such that:

$$\lambda_1 + \lambda_2 + \cdots + \lambda_{k+1} = n, \quad \lambda_j \geq 0 \text{ for all } j,$$

which is the same as the number of ways to distribute n indistinguishable objects to $k + 1$

distinguishable bins, which equals to $\binom{n+k}{k}$. This gives (10).

The bound (11) follows by taking a pessimistic value of k in each term (10) – twice $k = m$, then $k = 0$, and pessimistic value $n = m$. The second part of (11) follows from the Stirling formula.

The polynomial bound (12) follows from (10) and the inequality

$$\binom{p+k}{k} = \frac{(p+k)(p+k-1)\cdots(p+1)}{1 \cdot 2 \cdots k} \leq \text{const}(k) p^k$$

applied with $p = m$ and $p = n$.

Although the complexity of evaluating the cumulative distribution function of order statistics from Theorem 3.1 is exponential in the general case, we have shown in Theorem 4.2 that the complexity is bounded by a polynomial of a small degree when there are only two populations, and the number of order statistics considered, k , is fixed and small. The complexity also depends on n , the number of random variables from the first population, S_1 . In general, n is fixed by the state of nature.

To confirm and illustrate the result, we have conducted a timing experiment. We calculated the joint distribution function in the case of two populations. We considered $k = 1$, $k = 2$, and $k = 3$, and fixed $n = 1$. We measured the amount of time it took to compute the joint distribution function using the general Bapat Beg formula with permanents and the new special formula. Both theorems were implemented in Mathematica. The permanents were computed in Mathematica using the code:

```
Permanent[A_List] := With[v = Array[x Length [A]],
```

```
Coefficient[Times@@(A.v), Times@@v]
```

from Weisstein (2006). This function computes the permanent of matrix A by Vardi's formula

as the coefficient of $x_1 \dots x_m$ in the product $\prod_{i=1}^m (a_{i1} + a_{i2}x_2 + \cdots + a_{im}x_m)$, using symbolic manipulation with automatic caching of partial results by the Mathematica kernel. Amazingly, calculating the permanent from (5) in Mathematica results in times that seem to grow polynomially with m , the number of rows in the permanent. Consequently, for two populations, while the complexity of computation directly from Theorem 3.1 is exponential, the actual time observed while calculating the formulas in Mathematica was polynomial: Mathematica automatically did some of the improvement we have achieved when passing from Theorem 3.1 to Theorem 3.2. We attribute this speedup to the reuse of partial results by the Mathematica kernel.

Mathematica calculates the Bapat Beg formula more rapidly than predicted. In the timing experiment, the observed times for the new formula (Theorem 3.2) are much faster than the Bapat Beg formula. The observed improvement was quite dramatic. The observed improvement is of the order m^2 (Table 1). The observed complexity of the new formula for two populations was about $\text{const } m^k$, which confirms the result of Theorem 4.2 for constant

$n = 1$. All calculations were done on a workstation with Opteron 848 processors running Mathematica 5.2, under the SuSE Linux Enterprise Server 10 operating system.

Mathematica code to calculate the cumulative distribution function for arbitrary collections of order statistics of independent random variables which may have different distributions is available free from the authors. Examples demonstrating the use of the software are also available from the authors.

References

- Aitken, AC. Determinants and Matrices. Oliver and Boyd; New York: 1939.
- Arnold, BC.; Balakrishnan, N. Relations, Bounds and Approximations for Order Statistics. Springer; Berlin: 1989. Lecture Notes in Statistics 53.
- Balakrishnan N. Permanents, order statistics, outliers and robustness. *Revista Matematica Complutense* 2007;20:7–107.
- Balakrishnan, N.; Rao, CR. Order Statistics: Theory & Methods. Vol. 16. Handbook of Statistics. North-Holland; Amsterdam: 1998. Order statistics: an introduction; p. 3–24.
- Bapat RB. Permanents in probability and statistics. *Lin. Algebr. Applic* 1990;127:3–25.
- Bapat RB, Beg MI. Order statistics for non identically distributed variables and permanents. *Sankhyā Ser. A* 1989;51:79–93.
- Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Roy. Statist. Soc. Ser. B* 1995;57:289–300.
- Cao G, West M. Computing distributions of order statistics. *Commun. Statist. Theor. Meth* 1997;26:755–764.
- Childs A, Balakrishnan N. Relations for order statistics from non-identical logistic random variables and assessment of the effect of multiple outliers on the bias of linear estimators. *J. Statist. Plann. Infer* 2006;136(7):2227–2253.
- David, HA.; Nagaraja, HN. Order Statistics. Vol. 3rd ed. John Wiley & Sons; Hoboken, NJ: 2003. Wiley Series in Probability and Statistics, Wiley-Interscience.
- Forbert H, Marx D. Calculation of the permanent of a sparse positive matrix. *Comput. Phys. Commun* 2003;150:267–273.
- Glueck DH, Muller KE, Karimpour-Fard A, Hunter L. Expected power for the false discover rate with independence. *Commun. Statist. Theor. Meth* 2008;37:1855–1866.
- Hogg, RV.; Craig, AT. Introduction to Mathematical Statistics. Vol. 4th ed. Macmillan Publishing Co., Inc; New York: 1978.
- Jerrum M, Sinclair A, Vigoda E. A polynomial-time approximation algorithm for the permanent of a matrix with nonnegative entries. *J. ACM* 2004;51:671–697.
- Knuth, DE. The Art of Computer Programming, Vol. 2: Seminumerical Algorithms. Vol. 3rd ed. Addison-Wesley; New York: 1998.
- Shapiro LW. A Catalan triangle. *Discrete Math* 1976;14:83–90.
- Stanley, RP. Enumerative Combinatorics. Vol. 2. Volume 62 of Cambridge Studies in Advanced Mathematics. Cambridge University Press; Cambridge: 1999.
- Stirzaker, D. Elementary Probability: Second Edition. Cambridge University Press; Cambridge: 2006.
- Valiant LG. The complexity of computing the permanent. *Theoret. Comput. Sci* 1979;8:189–201.
- Weisstein, EW. Permanent. From MathWorld – A Wolfram Web Resource. 2006.
<http://mathworld.wolfram.com/Permanent.html>

Table 1

Fit of timing in Mathematica of the evaluation of the joint distribution of the first k statistics of m variables from two populations ($n = 1$ from one population, $m - n$ from the other). For fixed k , regression was used to fit the logarithm of the time with a linear function of $\log m$, and regression was then used again to fit the coefficients by linear functions of k

Bapat and Beg (1989) Theorem 4.2	New formula Theorem 3.2	Improvement
$10^{-2.9-0.36k} m^{2.0+1.1k}$	$10^{-2.6-0.01k} m^{0.06+1.02k}$	$10^{-0.30-0.34k} m^{1.93+0.09k}$