Generalizations of Denert's Statistic

Jennifer R. Galovich
Dennis E. White
28 March, 1994

Abstract. This paper describes a simple procedure for constructing large families of statistics which are equidistributed on the set of rearrangements of an arbitrary word. We describe conditions which guarantee that all members of a family so generated do indeed have the same distribution. As an application, we use the procedure to construct a family of Mahonian statistics which includes a number of well-known statistics as special cases.

Généralisations de la statistique de Denert

Abstrait. Cet article décrit une procédure assez simple pour construire des grandes familles de statistiques équidistribuées sur la classe de réarrangements d'un mot quelconque. Nous décrivons des conditions qui garantissent que tous les membres d'une famille ainsi engendrée auront vraiment la même distribution. Comme application, nous utilisons cette procédure pour construire une famille de statistiques mahoniennes qui inclut quelques statistiques bien connues comme des cas spéciaux.

In recent years a new statistic on permutations was introduced by Marleen Denert in connection with her work on genus zeta functions of local minimal hereditary orders [2]. D. Foata and D. Zeilberger [3] have shown that Denert's statistic ("den") has the Mahonian distribution on S_n. Subsequently, G.-N. Han showed that Denert's statistic is also Mahonian on arbitrary words [4]. There are a number of other statistics which also have the Mahonian distribution on arbitrary words – perhaps the best known are the inversion number, "inv", and the major index, "maj". Thus the statistics "den", "inv" and "maj" are equidistributed. The principal result of this paper is a simple procedure for constructing large families of equidistributed statistics, a procedure which is rooted in the relationship between Denert's statistic and the inversion number.

After providing the necessary definitions and notation, we describe the procedure for some interesting special cases, and then discuss the conditions under which our results can be generalized. We conclude with some remarks about the possibility of even wider application of these results.

If a is any word on the alphabet $\{1,2, \ldots, n\}$ we will write $W(a)$ to denote the collection of all rearrangements of the word a. If $w \in W(a)$ we let $\bar{w}$ denote its non-decreasing rearrangement. The classical statistics which are of particular interest are described in the following definition:

DEFINITION: If J is any set of (non-negative) integers, we write $\#J$ to denote the cardinality of J and $\sum J$ to denote the sum of the elements of J. Let a be given, and let $w \in W(a)$. We set
(i) $\text{inv } w = \#\{i < j : w_i > w_j\}$
(ii) $\text{maj } w = \sum\{i : w_i > w_{i+1}\}$.

Let $a = 1^{p_1}2^{p_2}\cdots k^{p_k}$ with $\sum p_i = n$. As is well-known, both maj and inv have the Mahonian distribution on $W(a)$, that is,

$$\sum_{w \in W(a)} q^{\text{inv } w} = \sum_{w \in W(a)} q^{\text{maj } w} = \begin{bmatrix} n \\ p_1, p_2, \ldots, p_k \end{bmatrix}.$$

In the course of showing that Denert's statistic is Mahonian on S_n, Foata and Zeilberger gave a convenient encoding which extends easily to arbitrary words. We take this encoding as a definition:

DEFINITION: If w is any word and $\bar{w}$ its non-decreasing rearrangement we define $\text{den } w = \sum d_i(w)$ where

$$d_i(w) = \begin{cases} \#\{j < i : w_j \leq \bar{w}_i \text{ or } w_j > w_i\} & \text{if } \bar{w}_i < w_i \\ \#\{j < i : w_i < w_j \leq \bar{w}_i\} & \text{if } \bar{w}_i \geq w_i \end{cases}.$$

EXAMPLE 1: Let $w = 3122314211$. We have

$\bar{w}$:	1	1	1	1	2	2	2	3	3	4
w :	3	1	2	2	3	1	4	2	1	1
$d_i(w)$:	0	0	2	2	3	2	4	2	5	6

so that $\text{den } w = \sum d_i(w) = 26$.

An alternative description of the Foata-Zeilberger encoding can be given in terms of cyclic intervals [4]:

DEFINITION: Let x and y be elements of $\{1, 2, \ldots, k\}$. We define $C(x, y)$, the *cyclic interval* associated to the pair (x, y) by

$$C(x,y) = \begin{cases} \{x+1, x+2, \ldots, y\} & \text{if } x \leq y \\ \{1, 2, \ldots, y\} \cup \{x+1, x+2, \ldots, k\} & \text{if } x > y \end{cases}.$$

Thus $d_i(w) = \#\{j < i : w_j \in C(w_i, \bar{w}_i)\}$.

We are now in a position to describe the key relationship between Denert's statistic and the inversion number. First, recall that if $w \in W(a)$ then $\operatorname{inv} w = \sum b_i(w)$ where

$$b_i(w) = \#\{j < i : w_j > w_i\}.$$

Thus if $w = 3122314211$ then $\operatorname{inv} w = 0 + 1 + 1 + 1 + 0 + 4 + 0 + 3 + 6 + 6 = 22$.

We can use this encoding to obtain Denert's statistic as follows: decompose the word w into blocks according to the value of $\bar{w}_i$:

$$\begin{array}{rllll} \bar{w}: & 1111 & 222 & 33 & 4 \\ w: & 3122 & 314 & 21 & 1. \end{array}$$

Working from right to left, and treating each block in turn, first compute the contribution of the right-most block to the encoding of the inversion number of w, and then act on the remaining blocks with the cyclic permutation $\gamma = 2341$ (one-line notation). Thus at each iteration the contribution from the right-most block is computed on the basis of the entries to the left which resulted from the previous action of γ. It is convenient to represent these calculations using the following scheme, where the contribution at each iteration is underscored:

$$\begin{array}{rllll} w: & 3122 & 314 & 21 & 1 \\ & 4233 & 421 & 32 & \underline{6} \\ & 1344 & 132 & \underline{25} & \\ & 2411 & \underline{324} & & \\ & \underline{0022} & & & \end{array}$$

Note that the contributions of each position are exactly the values of $d_i(w)$ obtained via the Foata-Zeilberger encoding; thus the sum of these contributions is also equal to $\operatorname{den} w$. Indeed, the cyclic interval description of the d_i guarantees that this procedure always yields Denert's statistic.

The procedure described above can be generalized. If w is a word of length n on $\{1, 2, \ldots, k\}$ and $\sigma \in S_k$ then we write σw to represent the action of σ on w, that is, $\sigma w = \sigma(w_1) \cdots \sigma(w_n)$. Let $\mathbf{C} = (C_1, \ldots, C_r)$ be any composition of n, and let $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_r)$ where each $\sigma_j \in S_k$. First, decompose the word $\sigma_r w$ into blocks $B_1, \ldots, B_r$

of sizes $C_1, \ldots, C_r$ respectively. Now compute the contribution of the letters in block B_r to the encoding of the inversion number of $\sigma_r w$ and then act on the remaining blocks with the permutation σ_{r-1}. Continue computing the contributions of each block in this way, working from right to left and using the permutations $\sigma_r, \ldots, \sigma_1$ in turn. More precisely, for each $j = 1, \ldots, r$ we compute the contribution of $(\sigma_j \cdots \sigma_r)(w_{B_j})$ to the encoding of the inversion number of $\sigma_j \cdots \sigma_r(w_{B_1} \cdots w_{B_j})$. The value of the generalized Denert statistic on w is given by the sum of the contributions from each block. We will denote this value by $GINV(\mathrm{C}, \sigma)(w)$.

EXAMPLE 2: With $w = 3122314211$ as before, $\mathrm{C} = (3,5,2)$ and $\sigma = (4312, 1432, 2341)$ we have:

$$\begin{array}{llll} w: & 312 & 23142 & 11 \\ (\text{action of } \sigma_3) & 423 & 34213 & 22 \\ (\text{action of } \sigma_2) & 243 & 32413 & \underline{55} \\ (\text{action of } \sigma_1) & 321 & \underline{13062} & \\ & \underline{012} & & \end{array}$$

So $GINV(\mathrm{C}, \sigma)(w) = 0+1+2+1+3+0+6+2+5+5 = 25$.

Note that if $w \in W(a)$ where $a = 1^{p_1} \cdots k^{p_k}$, then choosing $\mathrm{C} = (p_1, \ldots, p_k)$, $\sigma_j = \gamma = 23 \cdots k1$ for $j = 1, \ldots, k-1$ and $\sigma_k = 12 \cdots k$ yields the Mahonian statistic den. In fact, we have the following theorem:

Theorem 1: *Let $a = 1^{p_1} \cdots k^{p_k}$ with $\sum p_i = n$. For each choice of a composition $\mathrm{C} = (C_1, \ldots, C_r)$ of n and r-tuple of permutations σ, the statistic $GINV(\mathrm{C}, \sigma)$ has the Mahonian distribution on $W(a)$.*

Thanks to a result of G.-N. Han [4] a similar Mahonian statistic manufacturing machine can be constructed with the major index, maj, playing the role of inv.

DEFINITION: Let $a = 1^{p_1} \cdots k^{p_k}$ with $\sum p_i = n$, and let $w \in W(a)$; also, let $x \in \mathbf{Z} \cup \{\infty\}$, and set $w_{n+1} = x$. For each i, $1 \le i \le n$, let

$$h_i^x(w) = \#\{j < i : w_j \in C(w_i, w_{i+1})\}$$

and let $m^x(w) = \sum h_i^x(w)$.

EXAMPLE 3: Let $w = 3122314211$. Then

$$m^\infty(w) = 0+0+0+1+1+4+4+5+0+6 = 21$$

while

$$m^3(w) = 0 + 0 + 0 + 1 + 1 + 4 + 4 + 5 + 0 + 5 = 20.$$

Note that in the previous example, maj $w = 21 = m^\infty(w)$. Han has shown that for any $w \in W(a)$, $m^\infty(w) =$ maj w, so that m^∞ is certainly Mahonian, and in general we have:

Proposition 2: *Let $a = 1^{p_1} \cdots k^{p_k}$ with $\sum p_i = n$. If $x \in \mathbf{Z} \cup \{\infty\}$ then the statistic m^x has the Mahonian distribution on $W(a)$.*

A new family of statistics, $GMAJ$, can now be constructed: As before, let w be a word of length n on the alphabet $\{1, 2, \ldots, k\}$, let $\mathrm{C} = (C_1, \ldots, C_r)$ be a composition of n and let $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_r)$ where each $\sigma_j \in S_k$. Again, decompose $\sigma_r w$ into r blocks of sizes $C_1, \ldots, C_r$. For $j = 1, \ldots, r-1$ let x_j be the first letter in block j of $\sigma_j \cdots \sigma_r(w_{B_1} \cdots w_{B_j})$, and let $x_r = \infty$. Begin by computing the contribution of $\sigma_r w_{B_r}$ to $m^\infty(w) = m^\infty(w_{B_1} \cdots w_{B_r})$, and continue, working right to left, computing at each stage the contribution of $\sigma_j \cdots \sigma_r(w_{B_j})$ to $m^{x_j}(\sigma_j \cdots \sigma_r(w_{B_1} \cdots w_{B_j}))$. Then sum the contributions to obtain $GMAJ(\mathrm{C}, \boldsymbol{\sigma})(w)$.

EXAMPLE 4: Let $w = 41321314221443$, $\mathrm{C} = (5, 3, 4, 2)$ and $\boldsymbol{\sigma} = (2341, 2143, 4231, 1234)$. The scheme below shows the value of each x_j in bold-face, while the contributions of the positions in each block are underscored:

$$\begin{array}{rllll}
w: & 41321 & 314 & 2214 & 43\infty \\
 & 14324 & 341 & 22414 & \underline{94} \\
 & 23413 & 4322 & \underline{0529} & \\
 & 341244 & \underline{440} & & \\
 & \underline{00020} & & &
\end{array}$$

So $GMAJ(\mathrm{C}, \boldsymbol{\sigma})(w) = 0 + 0 + 0 + 2 + 0 + 4 + 4 + 0 + 0 + 5 + 2 + 9 + 9 + 4 = 39.$

Theorem 3: *Let $a = 1^{p_1} \cdots k^{p_k}$ with $\sum p_i = n$. For each choice of a composition $\mathrm{C} = (C_1, \ldots, C_r)$ of n and r-tuple of permutations $\boldsymbol{\sigma}$, the statistic $GMAJ(\mathrm{C}, \boldsymbol{\sigma})$ has the Mahonian distribution on $W(a)$.*

As in the case of $GINV$ the proof rests on certain properties shared by the "generating" statistics, inv and m^x :

(i) Each of these statistics has the same distribution on $W(a)$ as it does on $W(\sigma a)$ for any $\sigma \in S_k$. (We say that the distribution is *permuteable.*)

This is really a property of the distribution, a property which is also enjoyed by other distributions, e.g., the Eulerian distribution.

(ii) Suppose that $w \in W(a)$ is decomposed as $w = uv$; then each of these statistics can be written as a sum of contributions from the subwords u and v. This decomposition (or *splitting*) has the following properties:

a) The contribution of v depends only on the content (or *type*) of u.

b) Let v be fixed and consider the contribution of the various rearangements of u. Then this contribution, considered as a statistic on $W(u)$, has the Mahonian distribution.

For suppose that $w = uv$ and that u is a word of length l. Then

$$\text{inv } w = \text{ inv } u + \sum_{i=l+1}^{n} b_i(w)$$

while

$$m^x(w) = m^{w_{l+1}}(u) + \sum_{i=l+1}^{n} h_i^x(w).$$

The main result of this paper is, in essence, that any statistic which splits and has a permuteable distribution can serve as the progenitor of a family of equidistributed statistics, just as inv and m^x generate the Mahonian families $GINV$ and $GMAJ$. In fact, an even more general result can be obtained by allowing a new choice of progenitor at each iteration of the procedure. In order to make this more precise, we make the following definitions:

DEFINITION: Let W denote the collection of all words of finite length and let $\bar{W}$ be the set of their non-decreasing rearrangments. If T is any collection of statistics then we say that T is *equidistributed with distribution* F if there is a function $F : \bar{W} \to \mathbb{Z}[q]$ such that for every $t \in T$, if t is defined on $W(a)$ then

$$\sum_{w \in W(a)} q^{t(w)} = F(a).$$

DEFINITION: Let T be an equidistributed collection of statistics, and let τ : $W \times \bar{W} \to T$ be a map which satisfies the following properties:

(i) $\tau(w, a)$ is a statistic defined on $W(a)$;

(ii) for each $(w,a) \in W \times \bar{W}$, $t(w,a)$ splits.
Under these circumstances we will say that τ is *admissible* for T. .

The purpose of the map τ is to select a (possibly new) statistic from T at each iteration, yielding a statistic G^{τ}. As an example we describe a family $GMAJ^{\theta}$:

EXAMPLE 5: Let $\theta : W \times \bar{W} \to \mathbf{Z} \cup \{\infty\}$. For any pair $(w,a) \in W \times \bar{W}$ let $\mathrm{MAJ}^{\theta}(w,a) = m^{\theta(w,a)}$. Then MAJ^{θ} is admissible for any family of Mahonian statistics; when $\theta(w,a) = w_1$ we obtain the statistic $GMAJ$ described earlier. More generally, we have

$$GMAJ^{\theta}(\mathrm{C},\sigma) = G^{\mathrm{MAJ}^{\theta}}(\mathrm{C},\sigma).$$

To illustrate the construction, let C and σ be as in Example 4, and again set $w = 41321314221443$. Suppose that MAJ^{θ} selects the statistics m^1, m^{∞}, m^2 and finally m^3. To compute the value of $GMAJ^{\theta}(\mathrm{C},\sigma)$ we first find the contribution of $\sigma_r w_{B_r}$ to $m^1(\sigma_r w)$, then use m^{∞}, m^2 and m^3 for the remaining iterations:

$$\begin{array}{lllll} w: & 41321 & 314 & 2214 & 431 \\ & 14324 & 341 & 2241\infty & \underline{98} \\ & 23413 & 4322 & \underline{0529} & \\ & 341243 & \underline{440} & & \\ & \underline{00023} & & & \end{array}$$

Summing the contributions from each block, we have $GMAJ^{\theta}(\mathrm{C},\sigma)(w) = 46$. It can be shown that this statistic is equidistributed with the statistics in T, that is, $GMAJ^{\theta}(\mathrm{C},\sigma)$ is Mahonian. In general, we have:

Theorem 4: *Let T be an equidistributed collection with a permuteable distribution F, and let τ be an admissible map for T. Let k be any positive integer, and suppose that $a = 1^{p_1} \cdots k^{p_k}$ with $\sum p_i = n$. Then for each choice of a composition C, and r-tuple σ, the statistic $G^{\tau}(\mathrm{C},\sigma)$ is equidistributed with the statistics in T. Equivalently,*

$$\sum_{w \in W(a)} q^{G^{\tau}(\mathrm{C},\sigma)(w)} = F(a).$$

Proof: The proof is by induction on r, the number of blocks.

REMARKS: (i) It is not difficult to show that if $\theta(w,a) \equiv \infty$ then there is a pair (C',σ') with the property that $GMAJ^{\theta}(\mathrm{C},\sigma) = GINV(\mathrm{C}',\sigma')$. Thus the statistic $GMAJ^{\theta}$ actually subsumes the $GINV$ family. By making appropriate choices for θ, we find that $GMAJ^{\theta}$ can also be specialized to Han's statistic "den^a" or to Rawling's interpolating statistic r-maj [8, 9]

(ii) For certain classes of maps θ it is possible to define a Mahonian statistic GZ^θ which is related to $GMAJ^\theta$ in the same way that the Z-statistic of the q-Dyson Theorem is related to the major index [1]; that is,

$$GZ^\theta(C,\sigma)(w) = \sum_{s<t} GMAJ^\theta(C,\sigma)(w_{st}).$$

In addition, one of us (DW) has used some of our ancillary results to construct a bijection on arbitrary words which sends "Z" to "maj", giving yet another proof that the Z-statistic is Mahonian. (See [5] for a similar result.)

(iii) Although the Eulerian distribution is permuteable, it can be shown that no Eulerian statistic splits (except in a few trivial cases). Nevertheless, there appear to be instances of equidistributed families generated by Eulerian statistics using the procedure described here.

(iv) A slight adaptation of our procedure leads to a family of q-Stirling distributed statistics on restricted growth functions, using the statistic "lb" as the generating statistic [7, 11, 12]. It appears that these methods will extend to other combinatorial objects as well, e.g. barred permutations [10]

References

[1] D. Bressoud, Problems on the Z-statistic, *Discrete Math.* 73 (1988/89) 37-48.

[2] M. Denert, The genus zeta function of hereditary orders in central simple algebras over global fields, *Math. Comp.* 54 (1990) 449-465.

[3] D. Foata and D. Zeilberger, Denert's permutation statistic is indeed Euler-Mahonian, *Studies in Appl. Math.* 83 (1990) 31-59.

[4] G.-N. Han, Calcul Denertien, doctoral thesis, Publ. I.R.M.A. Strasbourg, 1991.

[5] G.-N. Han, Une courte démonstration d'un résultat sur la Z-statistique, *Comptes Rendus Acad. Sci. Paris* 314 (1992) 969-971.

[6] P.A. MacMahon, *Combinatory Analysis*, Vol. 1 (1917) and Vol. 2 (1918), reprinted by Chelsea, 1984.

[7] S.C. Milne, Restricted growth functions, rank row matchings of partition lattices, and q-Stirling numbers, *Adv. Math.* 43 (1982) 173-196.

[8] D. Rawlings, The r-major index, *J. Combin. Theory, Series A* 31 (1981) 175-183.

[9] D. Rawlings, The (q, r)-Simon Newcomb problem, *Linear and Multilinear Algebra* **10** (1981) 253-260.

[10] R. Stanley, Weyl groups, the hard Lefschetz theorem, and the Sperner property, *SIAM J. Alg. Disc. Meth.* **1** (1980) 168-184.

[11] M. Wachs, σ- restricted growth functions and p, q-Stirling numbers, preprint, October 16, 1992.

[12] M. Wachs and D. White, p, q-Stirling numbers and set partition statistics, *J. Combin. Theory, Series A* **56** (1991) 27-46.