

On the distribution of the number of optimal elements for various generalized weak orders

By

JOHN J. LONG, JR. and CARL WAGNER *)

1. Introduction. Let R be a binary relation on a set X and let aR , cR and sR denote, respectively, the asymmetric part of R , the complement of R , and the symmetric part of R . We say that R is *negatively transitive* if cR is transitive. R is called a *generalized weak order* (GWO) if caR is transitive. Generalized weak orders were introduced by Peter Fishburn [2] and include as special cases the *asymmetric weak orders* (asymmetric, negatively transitive relations) familiar to economists as models, respectively, of strict preference and preference-or-indifference. Every symmetric relation on X is also a GWO. Indeed, we shall prove that as $|X| \rightarrow \infty$ the symmetric relations actually predominate among GWOs.

An alternative $x \in X$ is *optimal* for a GWO R if $\forall y \in X, yRx \Rightarrow xRy$. Our aim in this paper is to study the distribution of the number of optimal elements for randomly selected GWOs of several types. Specifically, if X is an n -set, let

$$(1.1) \quad \begin{aligned} \mathcal{G}_n &= \{R: R \text{ is a GWO on } X\} \\ \mathcal{T}_n &= \{R \in \mathcal{G}_n: R \text{ is transitive}\} \\ \mathcal{M}_n &= \{R \in \mathcal{G}_n: R \text{ and } cR \text{ are transitive}\} \\ \mathcal{W}_n &= \{R: R \text{ is a complete weak order on } X\}, \end{aligned}$$

and let the random variables G_n , T_n , M_n and W_n record the number of optimal elements for a GWO R selected at random from \mathcal{G}_n , \mathcal{T}_n , \mathcal{M}_n , and \mathcal{W}_n , respectively.

The combinatorial analysis of the above random variables is considerably facilitated by the existence of a bijection between the set of GWOs on X and the set of ordered partitions of X , each of the blocks of which is equipped with an arbitrary symmetric relation [2, pp. 164–65], [4, p. 147]. The essential details of this correspondence are as follows: a generalized weak order R on a set X partitions X by the equivalence relation $scaR$, and $\hat{X} = X/scaR$ is linearly ordered by $>$, where for all $A, B \in \hat{X}$, $A > B$ iff $(x, y) \in aR$ for all $x \in A$ and $y \in B$. The symmetric relation attached to each $A \in \hat{X}$ is

*) This author's work was supported in part by a grant from the University of Tennessee.

simply the restriction of R to A . Conversely, if \tilde{X} is a partition of X linearly ordered by \succ , and each $A \in \tilde{X}$ is equipped with a symmetric relation R_A , then $R = R^{(s)} \cup R^{(a)}$ is a generalized weak order on X , where $R^{(s)} = \bigcup_{A \in \tilde{X}} R_A$ and $R^{(a)} = \bigcup_{\substack{A, B \in \tilde{X} \\ A \succ B}} A \times B$. This general-

ization of the traditional bijection between weak orders and ordered partitions enjoys the additional feature of pairing GWOs in \mathcal{T}_n with partitions whose blocks carry transitive, symmetric (i.e., partial equivalence) relations, and GWOs in \mathcal{M}_n with partitions whose blocks carry the empty or universal relation [2, Theorem 2]. The blocks of ordered partitions corresponding to complete weak orders are all equipped with the universal relation, and the blocks of those corresponding to asymmetric weak orders with the empty relation [2, p. 167].

It is clear that an alternative $x \in X$ is optimal for a GWO R iff x is an element of the initial block of the ordered partition corresponding to R . Hence the random variables G_n , T_n , M_n , and W_n may be thought of, respectively, as recording initial block cardinalities of randomly chosen elements of \mathcal{G}_n , \mathcal{T}_n , \mathcal{M}_n and \mathcal{W}_n , now regarded as classes of ordered partitions with blocks carrying appropriate relations. In the next two sections we analyze the density functions and moments of the above random variables and determine their limits as $n \rightarrow \infty$.

2. The random variable G_n . Let \mathcal{G}_n denote the set of all ordered partitions of an n -set, the blocks of which carry arbitrary symmetric relations, and let $g_n = |\mathcal{G}_n|$. Let G_n denote the cardinality of the initial block of a randomly chosen $R \in \mathcal{G}_n$. Since there are $2^{\binom{k+1}{2}}$ symmetric relations on a k -set, the density function of G_n is

$$(2.1) \quad f_{G_n}(k) = P(G_n = k) = \binom{n}{k} 2^{\binom{k+1}{2}} g_{n-k} / g_n, \quad 1 \leq k \leq n,$$

where the sequence (g_n) satisfies the recurrence relation

$$(2.2) \quad g_n = \sum_{k=1}^n \binom{n}{k} 2^{\binom{k+1}{2}} g_{n-k} \quad (g_0 = 1).$$

The factor $2^{\binom{k+1}{2}}$ in (2.1) would lead one to guess that $f_{G_n}(k)$ is maximized when $k = n$. This intuition is true in the strongest possible terms, as indicated by the following theorem:

Theorem 2.1. *If G_n denotes the initial block cardinality of a randomly selected $R \in \mathcal{G}_n$, and $Y_n = G_n - n$, then the sequence of random variables (Y_n) converges to zero in mean square and with probability one.*

Proof. By [3, Ch. 10, Th. 1A] it suffices to show that

$$(2.3) \quad \sum_{n=1}^{\infty} E(Y_n^2) < \infty,$$

which we prove by showing that

$$(2.4) \quad E(Y_n^2) = O(n^4/2^n).$$

From $Y_n = G_n - n$ and (2.1), we have

$$(2.5) \quad \begin{aligned} E(Y_n^2) &= E(G_n(G_n - 1)) + (1 - 2n)E(G_n) + n^2 \\ &= n(n - 1)q_n + (1 - 2n)np_n + n^2, \end{aligned}$$

where

$$(2.6) \quad p_n = \sum_{j=0}^{n-1} \binom{n-1}{j} 2^{\binom{j+2}{2}} g_{n-1-j}/g_n$$

and

$$(2.7) \quad q_n = \sum_{j=0}^{n-2} \binom{n-2}{j} 2^{\binom{j+3}{2}} g_{n-2-j}/g_n.$$

If we let $a_n = g_n/2^{\binom{n+1}{2}}$, then

$$(2.8) \quad 0 < 1/a_n \leq q_n \leq p_n \leq 1,$$

since $1/a_n$ is the last term of the sum in (2.7) and since p_n is the probability that some fixed object of an n -set belongs to the first block of a randomly chosen $R \in \mathcal{G}_n$ and q_n the probability that that same object, as well as another fixed object, belongs to the first block.

From (2.5) and (2.8),

$$(2.9) \quad \begin{aligned} E(Y_n^2) &\leq n(n - 1)p_n + (1 - 2n)np_n + n^2 \\ &= n^2(1 - p_n) \leq n^2 a_n(1 - p_n) = n^2(a_n - a_n p_n) \leq n^2(a_n - 1). \end{aligned}$$

Hence, if we can prove that

$$(2.10) \quad a_n - 1 = O(n^2/2^n),$$

(2.4) will follow from (2.9).

To show (2.10) divide each side of the recurrence relation (2.2) by $2^{\binom{n+1}{2}}$, and replace $n - k$ by k , yielding

$$(2.11) \quad a_n - 1 = \sum_{k=1}^{n-1} \binom{n}{k} a_k (1/2)^{k(n-k)}.$$

Letting $b_n = \max_{0 \leq k \leq n} \{a_k\}$, it follows from (2.11) that

$$(2.12) \quad \begin{aligned} 0 \leq a_n - 1 &\leq b_{n-1} \left(4n \left(\frac{1}{2}\right)^n + \left(\frac{1}{2}\right)^{2(n-2)} \sum_{k=2}^{n-2} \binom{n}{k} \right) \\ &\leq b_{n-1} \left(4n \left(\frac{1}{2}\right)^n + \left(\frac{1}{2}\right)^{n-4} \right). \end{aligned}$$

For $n \geq 6$, $a_n \leq 1 + b_{n-1}$. Thus $b_n = O(n)$, and so (2.10) follows from (2.12).

In conclusion we remark that

$$(2.13) \quad \lim_{n \rightarrow \infty} 2^{\binom{n+1}{2}}/g_n = \lim_{n \rightarrow \infty} 1/a_n = \lim_{n \rightarrow \infty} a_n = 1.$$

by (2.10). Since $2^{\binom{n+1}{2}}$ is the number of symmetric relations on an n -set, it follows that as $n \rightarrow \infty$ “almost all” generalized weak orders are simply symmetric relations.

3. The random variables T_n , M_n , and W_n . Let X be an n -set. Exploiting the bijection described in the third paragraph of Sect. 1, we now let \mathcal{T}_n denote the set of all ordered partitions of X with blocks carrying symmetric, transitive (i. e., partial equivalence) relations, \mathcal{M}_n the set of such partitions with blocks carrying the empty or universal relation, and \mathcal{W}_n the set of such partitions with blocks all carrying the universal relation. Let T_n , M_n , and W_n denote the initial block size of an ordered partition chosen at random, respectively, from \mathcal{T}_n , \mathcal{M}_n and \mathcal{W}_n . Letting $t_n = |\mathcal{T}_n|$, $m_n = |\mathcal{M}_n|$, and $w_n = |\mathcal{W}_n|$ and using the fact that a k -set may be equipped with a partial equivalence relation in $B(k + 1)$ ways (where $B(k + 1)$ is the $k + 1$ st Bell number [4, p. 148]) and with the empty or universal relation in two ways, we get the density functions

$$(3.1) \quad f_{T_n}(k) = P(T_n = k) = \binom{n}{k} B(k + 1) t_{n-k}/t_n \quad (1 \leq k \leq n),$$

$$(3.2) \quad f_{M_n}(k) = P(M_n = k) = 2 \binom{n}{k} m_{n-k}/m_n \quad (1 \leq k \leq n),$$

and

$$(3.3) \quad f_{W_n}(k) = P(W_n = k) = \binom{n}{k} w_{n-k}/w_n \quad (1 \leq k \leq n),$$

and the recurrence relations

$$(3.4) \quad t_n = \sum_{k=1}^n \binom{n}{k} B(k + 1) t_{n-k} \quad (t_0 = 1),$$

$$(3.5) \quad m_n = \sum_{k=1}^n 2 \binom{n}{k} m_{n-k} \quad (m_0 = 1),$$

and

$$(3.6) \quad w_n = \sum_{k=1}^n \binom{n}{k} w_{n-k} \quad (w_0 = 1).$$

Using (3.4)–(3.6) [4, pp. 149–50], or a direct argument based on the multiplication of Taylor series, along with the classical [1, p. 43] result

$$(3.7) \quad \sum_{k=0}^{\infty} B(k) z^k/k! = \exp(e^z - 1),$$

one may easily derive the following exponential generating functions:

$$(3.8) \quad \Psi_t(z) = \sum_{n=0}^{\infty} t_n z^n/n! = [1 - (\exp(e^z + z) - 1)]^{-1},$$

$$(3.9) \quad \Psi_m(z) = \sum_{n=0}^{\infty} m_n z^n/n! = [1 - 2(e^z - 1)]^{-1},$$

and

$$(3.10) \quad \Psi_w(z) = \sum_{n=0}^{\infty} w_n z^n/n! = [1 - (e^z - 1)]^{-1}.$$

These generating functions can be used to derive asymptotic estimates of t_n , m_n , and w_n , from which the limiting distributions of T_n , M_n , and W_n follow immediately. Note that in each of (3.8)–(3.10), we have a generating function of the form $\Psi(z) = [1 - F(z)]^{-1}$, where F is entire. Indeed, as we show below, a comprehensive theorem, yielding asymptotic estimates for t_n , m_n , and w_n , may be based on the assumption of analyticity of F on a disk of positive radius. (The numbers g_n , for which $F(z) = \sum_{k \geq 1} \binom{k+1}{2} z^k/k!$, converging only at $z = 0$, are, in the light of the following result, understandably a different case.)

Theorem 3.1. *If $(c_k)_{k \geq 1}$ is a positive sequence such that*

$$(3.11) \quad F(z) = \sum_{k=1}^{\infty} c_k z^k/k!$$

is analytic on a disk of positive radius R , and the sequence $(y_n)_{n \geq 0}$ is defined by the recurrence relation

$$(3.12) \quad y_n = \sum_{k=1}^n \binom{n}{k} c_k y_{n-k} \quad (y_0 = 1),$$

then

$$(3.13) \quad y_n \sim n!/F'(p) p^{n+1},$$

where p is the unique real positive solution of $F(z) = 1$.

Proof. Let

$$(3.14) \quad \Psi(z) = \sum_{n=0}^{\infty} y_n z^n/n!.$$

From (3.11), (3.14), and (3.12) it follows that $\Psi(z)F(z) = \Psi(z) - 1$, i.e., that $\Psi(z) = (1 - F(z))^{-1}$. From (3.11) and the positivity of c_k , $k \geq 1$, we obtain

$$(3.15) \quad F(0) = 0 \quad \text{and} \quad \lim_{\substack{x \rightarrow R^- \\ x \in \mathbb{R}}} F(x) = +\infty,$$

$$(3.16) \quad F'(x) > 0 \quad \text{for } x \in [0, R),$$

and

$$(3.17) \quad |F(z)| \leq F(|z|), \quad |z| < R.$$

By (3.15) and (3.16) there exists a unique real positive p such that $F(p) = 1$. Since $F'(p) \neq 0$, we have

$$(3.18) \quad \lim_{z \rightarrow p} (z - p) \Psi(z) = -1/F'(p),$$

and so the pole p of Ψ is simple, with residue $a = -1/F'(p)$.

Furthermore, there exists a real $r > p$ such that Ψ is analytic for $|z| < r$, except for the aforementioned simple pole at $z = p$. For let $r = \min \{|z|: F(z) = 1 \text{ and } z \neq p\}$, setting $r = R$ if this set is empty. By (3.17), we have $r \geq p$. If $r = p$, then equality holds in (3.17) for some z . But then $\arg(z) = \arg(z^2)$ and $z = |z| = p$, which is a contradiction. Thus $r > p$, as asserted.

Now set

$$(3.19) \quad \varphi(z) = \Psi(z) - (a/(z - p)) = \sum_{n=0}^{\infty} (y_n/n! + a/p^{n+1}) z^n,$$

where $a = -1/F'(p)$ is the residue of Ψ at p . Since φ is analytic at p , (3.19) converges there, and so

$$(3.20) \quad \lim_{n \rightarrow \infty} (y_n p^n/n! + a/p) = 0,$$

from which it follows that

$$(3.21) \quad y_n \sim n! / -ap^{n+1} = n!/F'(p)p^{n+1}.$$

We remark that Bender [1] has established results which coincide with (3.21) when $F(z) = e^z - 1$. We avoid his use of a special case of Darboux's theorem by our choice of the function $\varphi(z)$ of (3.19) to "repair" $\Psi(z)$ [cf. Bender's $g_1(z)$, p. 501].

The following theorem yields, as special cases, the limiting distributions of the random variables T_n , M_n , and W_n :

Theorem 3.2. *Under the hypotheses of Theorem 3.1, with p the unique positive real solution of $F(z) = 1$, if one defines random variables $(Y_n)_{n \geq 1}$ and Y by the density functions*

$$(3.22) \quad f_n(k) = P(Y_n = k) = \binom{n}{k} c_k y_{n-k}/y_n, \quad 1 \leq k \leq n,$$

and

$$(3.23) \quad f(k) = P(Y = k) = c_k p^k/k!, \quad 1 \leq k < \infty,$$

then $(Y_n) \rightarrow Y$ in distribution and the moments of Y_n converge to the moments of Y .

Proof. For each $k \geq 1$, it follows immediately from (3.13) applied to y_{n-k} and y_n that

$$(3.24) \quad f_n(k) = \binom{n}{k} c_k y_{n-k}/y_n \rightarrow c_k p^k/k! = f(k), \quad \text{as } n \rightarrow \infty.$$

Since the densities of the Y_n converge pointwise to the density of Y and these random variables are discrete, $(Y_n) \rightarrow Y$ in distribution.

Setting $k = 1$ in (3.24) yields $f_n(1)/c_1 \rightarrow p$ as $n \rightarrow \infty$, and so $\exists M < R$ (the radius of convergence of F) and $\exists N$ such that $f_n(1)/c_1 \leq M$ for all $n \geq N$. Since

$$(3.25) \quad f_n(k) = (c_k/k!) \prod_{i=0}^{k-1} (f_{n-i}(1)/c_1)$$

for $1 \leq k \leq n$ and $f_n(k) = 0$ for $k > n$, it follows that $\exists B$ such that $f_n(k) \leq B c_k M^k/k!$ for all $n, k \geq 1$. Hence for all $j \geq 1$

$$(3.26) \quad k^j f_n(k) \leq k^j B c_k M^k/k!$$

and so

$$(3.27) \quad E(Y_n^j) = \sum_{k=1}^n k^j f_n(k) \leq B \sum_{k=1}^{\infty} c_k k^j M^k/k! < \infty,$$

since $F(z) = \sum_{k=1}^{\infty} c_k z^k/k!$ and all its derivatives are finite at $z = M$. It follows from the Lebesgue dominated convergence theorem that

$$(3.28) \quad \lim_{n \rightarrow \infty} E(Y_n^j) = \sum_{k=1}^{\infty} k^j f(k) = E(Y^j).$$

In particular, since $F(pz) = \sum_{k=1}^{\infty} (c_k p^k/k!) z^k = \sum_{k=1}^{\infty} f(k) z^k$, we have

$$(3.29) \quad \lim_{n \rightarrow \infty} E(Y_n) = E(Y) = p F'(p),$$

and

$$(3.30) \quad \lim_{n \rightarrow \infty} E(Y_n^2) = E(Y^2) = p F'(p) + p^2 F''(p),$$

so that

$$(3.31) \quad \lim_{n \rightarrow \infty} \text{Var}(Y_n) = \text{Var}(Y) = p F'(p) (1 - p F'(p)) + p^2 F''(p).$$

The above theorem yields the following corollary, which describes the asymptotic behavior of the random variables T_n , M_n , and W_n :

Corollary 3.2.

- (a) $\lim_{n \rightarrow \infty} f_{T_n}(k) = B(k+1)\lambda^k/k!$ ($\lambda \cong 0.3183247$)
 $\lim_{n \rightarrow \infty} E(T_n) = 2(e^\lambda + 1) \cong 1.5119294$
 $\lim_{n \rightarrow \infty} \text{Var}(T_n) = 2\lambda(e^\lambda + 1 - \lambda(e^{2\lambda} + e^\lambda + 1))$

- (b) $\lim_{n \rightarrow \infty} f_{M_n}(k) = 2 (\log(3/2))^k / k!$
 $\lim_{n \rightarrow \infty} E(M_n) = 3 \log(3/2) \cong 1.2163953$
 $\lim_{n \rightarrow \infty} \text{Var}(M_n) = 3 \log(3/2) (1 - 2 \log(3/2))$
- (c) $\lim_{n \rightarrow \infty} f_{W_n}(k) = (\log 2)^k / k!$
 $\lim_{n \rightarrow \infty} E(W_n) = 2 \log 2 \cong 1.3862944$
 $\lim_{n \rightarrow \infty} \text{Var}(W_n) = 2 \log 2 (1 - \log 2).$

P r o o f. Apply (3.24), (3.29), and (3.31) to $Y_n = T_n$ (with $p = \lambda \cong 0.3183247$ the unique positive real solution of $\exp(e^z + z - 1) = 2$), $Y_n = M_n$ (with $p = \log(3/2)$), and $Y_n = W_n$ (with $p = \log 2$), respectively.

It is interesting to note that while the initial block size of any *total* order on an n -set is always 1, even $\lim_{n \rightarrow \infty} E(T_n)$ is only about 1.5. On the other hand, it follows immediately from Theorem 2.1 that $E(G_n) \sim n$.

References

- [1] E. BENDER, Asymptotic methods in enumeration. SIAM Review **16**, 485–515 (1974).
 [2] C. BERGE, Principles of Combinatorics. New York 1971.
 [3] P. FISHBURN, Transitivity. Rev. Economic Studies (1) **46**, 163–173 (1979).
 [4] E. PARZEN, Modern Probability Theory and Its Applications. New York 1960.
 [5] C. WAGNER, Enumeration of generalized weak orders. Arch. Math. **39**, 147–152 (1982).

Eingegangen am 7. 1. 1988

Anschrift der Autoren:

John J. Long, Jr. and Carl Wagner
 Mathematics Department
 University of Tennessee
 Knoxville, TN
 USA 37996–1300