The asymptotic behavior of the Hurwitz binomial distribution*

Jim Pitman

Technical Report No. 500

Department of Statistics
University of California
367 Evans Hall # 3860
Berkeley, CA 94720-3860

June 16, 1998

Abstract

Hurwitz's extension of Abel's binomial theorem defines a probability distribution on the set of integers from 0 to n. This is the distribution of the number of non-root vertices of a fringe subtree of a suitably defined random tree with $n+2$ vertices. The asymptotic behaviour of this distribution is described in a limiting regime where the distribution of the delabeled fringe subtree approaches that of a Galton-Watson tree with a mixed Poisson offspring distribution.

1 Introduction and statement of results

Hurwitz [10] discovered the following identity of polynomials in $n+2$ variables x, y and $z_s, s \in [n] := \{1, \dots, n\}$, which reduces to the binomial expansion of $(x+y)^n$ when

*Research supported in part by N.S.F. Grant DMS97-03961

$z_s \equiv 0$:

$$\sum_{A \subseteq [n]} x(x + z_A)^{|A|-1}(y + z_{\bar{A}})^{|\bar{A}|} = (x + y + z_{[n]})^n \tag{1}$$

where the sum is over all 2^n subsets A of $[n]$, with the notations $z_A := \sum_{s \in A} z_s$, and $|A|$ for the number of elements of A, and $\bar{A} := [n] - A$. See [11, 8, 19, 21] for combinatorial explanations of this identity. As observed in [15], this identity amounts to the fact that for each probability distribution p on the interval of integers $[0, n+1] := \{0, 1, \ldots, n, n+1\}$, the formula

$$P(V(n) = A) = p_0(p_0 + p_A)^{|A|-1}(p_{n+1} + p_{\bar{A}})^{|\bar{A}|} \qquad (A \subseteq [n]) \tag{2}$$

defines the probability distribution of a random subset $V(n)$ of $[n]$, which can be constructed as follows. As a consequence of Cayley's multinomial expansion [4, 18, 15], a probability distribution for a random tree $\mathcal{T}(n)$ labeled by $[0, n+1]$ with root $n+1$ is given by the formula

$$P(\mathcal{T}(n) = \mathbf{t}) = p_{n+1}^{|\mathbf{t}_{n+1}|-1} \prod_{i=0}^{n} p_i^{|\mathbf{t}_i|} \tag{3}$$

for all trees $\mathbf{t}$ labeled by $[0, n+1]$ with root $n+1$, where $|\mathbf{t}_i|$ is the number of children of vertex i in the tree $\mathbf{t}$. See [14, 15, 16, 17] for various constructions of $\mathcal{T}(n)$ and related results. For a vertex x of a rooted tree $\mathbf{t}$ let $V_x(\mathbf{t})$ be the set of non-root vertices of the *fringe subtree of* $\mathbf{t}$ *rooted at* x, that is the set of all vertices v of $\mathbf{t}$ in such that there is a path from x to v in $\mathbf{t}$ directed away from the root of $\mathbf{t}$. See Aldous [3] for background and further references to fringe subtrees.

Theorem 1 [15]. *For a random tree $\mathcal{T}(n)$ with distribution* (3) *on trees labeled by $[0, n+1]$ with root $n+1$, the random set $V(n) := V_0(\mathcal{T}(n))$ of non-root vertices of the fringe subtree of $\mathcal{T}(n)$ rooted at 0 has the Hurwitz distribution* (2).

The purpose of this paper is to present the following result regarding the distribution of the size of such a random set $V(n)$, in a limiting regime as $n \to \infty$ and the probability distribution p on $[0, n+1]$ is allowed to vary as a function of n.

Theorem 2 *Let $V(n)$ be a random subset of $[n]$ with the Hurwitz distribution* (2)*, for $(p_i := p_{i,n}, 0 \le i \le n+1)$ a probability distribution on $[0, n+1]$ which varies with n. Then there exist limits*

$$q_a = \lim_{n \to \infty} P(|V(n)| = a) \text{ for all } a = 0, 1, 2, \ldots \tag{4}$$

if and only if there exists a limit

$$\lambda = \lim_{n\to\infty} np_{0,n} \in [0,\infty]$$

and one of the following conditions obtains:
(i) $\lambda = 0$*, in which case* $(q_a) = (1,0,0,\dots)$
(ii) $\lambda = \infty$*, in which case* $(q_a) = (0,0,0,\dots)$
(iii) $\lambda \in (0,\infty)$ *and the sequence of probability measures* F_n *on* $[0,\infty)$ *defined by*

$$F_n[0,t] := n^{-1} \sum_{i=1}^{n} 1(np_{i,n} \le t)$$

has a weak limit F*, a probability distribution on* $[0,\infty)$ *with mean at most 1; then*

$$q_a = \frac{\lambda}{a!} \int_0^\infty (\lambda + t)^{a-1} e^{-(\lambda+t)} F^{a*}(dt) \qquad (a = 0,1,2,\dots) \tag{5}$$

where F^{a*} *is the a-fold convolution of* F *(in particular* F^{0*} *is a unit mass at* 0 *so that* $q_0 = e^{-\lambda}$*), and* $\sum_a q_a = 1$*.*

In the particular case when $p_{i,n} = \theta/n$ for all $i \in [n]$ and some $0 \le \theta \le 1$ the distribution of $|V(n)|$ is the *quasi-binomial* distribution studied by Consul [5, 6]. In this case $F_n = \delta_\theta$, a unit mass at θ, for every n. The limit distribution (q_a) in (5) for $F = \delta_\theta$ is the *generalized Poisson distribution* with parameters λ and θ:

$$q_a = \frac{\lambda}{a!} (\lambda + a\theta)^{a-1} e^{-(\lambda + a\theta)} \qquad (a = 0,1,\dots). \tag{6}$$

So Theorem 2 is a generalization of Consul's description of this generalized Poisson distribution as a limit of quasi-binomial distributions. Let P_λ denote the usual Poisson distribution with parameter λ, which is the case $\theta = 0$ of (6). In the setting of Theorem 1, as $n \to \infty$ and $np_{0,n}$ remains bounded, it can be shown using results of [15] that the delabeled fringe sub-tree of $\mathcal{T}(n)$ rooted at 0 is well approximated by a modified Galton-Watson branching process in which the root individual has offspring distribution P_λ for $\lambda = np_{0,n}$, and all following individuals have offspring distribution that is the mixed Poisson distribution $\int_0^\infty P_\lambda F_n(d\lambda)$. Details of this approximation will not be given here, but for similar approximations of combinatorially defined random trees by Galton-Watson trees, and related constructions, see [9, 2, 3, 1]. The Galton-Watson approximation to the delabeled fringe sub-tree of $\mathcal{T}(n)$ rooted at 0 explains why the limit distribution (q_a) in (5) has the the following interpretation as a *Lagrangian distribution* [13, 20]:

Theorem 3 *For each probability distribution F on $[0,\infty)$, and each $\lambda \geq 0$, the sequence (q_a) defined by (5) is the distribution of the total progeny $\sum_{i=0}^{\infty} Z_i$ in a Galton-Watson branching process $(Z_0, Z_1, \ldots)$ where the initial number of individuals Z_0 has the Poisson(λ) distribution P_λ, and the offspring distribution is the mixture of Poisson distributions $\int_0^\infty P_\lambda F(d\lambda)$.*

Note that

$$\sum_a q_a = P(\sum_i Z_i < \infty) \tag{7}$$

which equals 1 iff F has mean at most 1, by standard theory of branching processes.

2 Proofs

The interpretation of (q_a) given in Theorem 3 is needed for the proofs of some parts of Theorem 2, so Theorem 3 will be proved first.

Proof of Theorem 3. According to a formula of Otter [12] and Dwass [7], which is reviewed in [17], for a Galton-Watson process $(Z_0, Z_1, \ldots)$ with arbitrary offspring distribution, for every $k = 1, 2, \ldots$ and $a = 1, 2, \ldots$

$$P(\sum_i Z_i = a \,|\, Z_0 = k) = \frac{k}{a} P(S_a = a - k)$$

where S_a is the sum of a independent copies of the generic offspring variable. If the offspring distribution is $\int_0^\infty P_\lambda F(d\lambda)$, then S_1 has the same distribution as $N(T_1)$ where $(N(t), t \geq 0)$ is a homogeneous Poisson process independent of T_1 with distribution F. Hence S_a has the same distribution as $N(T_a)$ for T_a the sum of a independent copies of T_1, so $P(T_a \in dt) = F^{a*}(dt)$. So for $1 \leq k \leq a$

$$P(S_a = a - k) = P(N(T_a) = a - k) = \int_0^\infty e^{-t} \frac{t^{a-k}}{(a-k)!} F^{a*}(dt).$$

It follows that for Z_0 with Poisson(λ) distribution and $a = 1, 2, \ldots$

$$P(\sum_i Z_i = a) = \int_0^\infty f(a, \lambda, t) F^{a*}(dt)$$

where

$$f(a, \lambda, t) := \sum_{k=1}^{a} e^{-\lambda} \frac{\lambda^k}{k!} \frac{k}{a} e^{-t} \frac{t^{a-k}}{(a-k)!} = \frac{\lambda}{a!} (\lambda + t)^{a-1} e^{-(\lambda+t)},$$

and the conclusion of Theorem 3 follows. □

Proof of Theorem 2. Let $V(n)$ be a random subset of $[n]$ with the Hurwitz distribution (2) induced by $(p_i = p_{i,n}, 0 \le i \le n+1)$. So for $a \in [0,n]$

$$P(|V(n)| = a) = \sum_{A \subseteq [n]: |A| = a} p_0 (p_0 + p_A)^{a-1} (p_{n+1} + p_{\bar{A}})^{n-a}. \tag{8}$$

In particular

$$P(|V(n)| = 0) = (1 - p_{0,n})^n \to e^{-\lambda}$$

as $n \to \infty$ iff $np_{0,n} \to \lambda$ for some $\lambda \in [0, \infty]$. Cases (i), (ii) and (iii) now arise according to the value of λ. Consider next

$$P(|V(n)| = 1) = \sum_{i=1}^{n} p_{0,n}(1 - p_{0,n} - p_{i,n})^{n-1} = np_{0,n} \int_0^\infty \left(1 - p_{0,n} - \frac{x}{n}\right)^{n-1} F_n(dx).$$

In case (iii) with $np_{0,n} \to \lambda$ it follows easily from this expression that

$$P(|V(n)| = 1) \sim \lambda \int_0^\infty \left(1 - \frac{\lambda + x}{n}\right)^{n-1} F_n(dx) \to \lambda \int_0^\infty e^{-\lambda - x} F(dx)$$

if F_n converges weakly to F. A similar approximation of multiple sums by multiple integrals shows that in the same limit regime as $n \to \infty$

$$P(|V(n)| = a) \to \frac{\lambda}{a!} \int_0^\infty \cdots \int_0^\infty \left(\lambda + \sum_{i=1}^{a} x_i\right)^{a-1} \exp\left(-\lambda - \sum_{i=1}^{a} x_i\right) F(dx_1) \cdots F(dx_a)$$

for every $a = 1, 2, \ldots$, which yields (5). Since F_n has mean at most 1 for every n, so does any weak limit F. It then follows from the sentence after Theorem 3 that $\sum_a q_a = 1$. Conversely, if a limit q_a of $P(|V(n)| = a)$ exists for all a, then the limit λ of $np_{0,n}$ exists by consideration of $a = 0$. In case (iii), the sequence F_n is tight because of the upper bound 1 on the mean. By passing to subsequences, the limit (q_a) must be of the form (5) for some subsequential weak limit F. To see that this F is in fact the weak limit of the whole sequence F_n it is enough to show that (q_a) determines F uniquely, and this follows easily from Theorem 3. □

References

[1] D. Aldous. Tree-valued Markov chains and Poisson-Galton-Watson distributions. In D. Aldous and J. Propp, editors, *Microsurveys in Discrete Probability*, number 41 in DIMACS Ser. Discrete Math. Theoret. Comp. Sci, pages 1–20, 1998.

[2] D.J. Aldous. A random tree model associated with random graphs. *Random Structures Algorithms*, 1:383–402, 1990.

[3] D.J. Aldous. Asymptotic fringe distributions for general families of random trees. *Ann. Appl. Probab.*, 1:228–266, 1991.

[4] A. Cayley. A theorem on trees. *Quarterly Journal of Pure and Applied Mathematics*, 23:376–378, 1889. (Also in *The Collected Mathematical Papers of Arthur Cayley. Vol XIII*, 26-28, Cambridge University Press, 1897).

[5] P.C. Consul. A simple urn model dependent upon a predetermined strategy. *Sankhya Ser. B*, 36:391–399, 1974.

[6] P.C. Consul. *Generalized Poisson Distributions.* Dekker, 1989.

[7] M. Dwass. The total progeny in a branching process. *J. Appl. Probab.*, 6:682–686, 1969.

[8] J. Françon. Preuves combinatoires des identités d'Abel. *Discrete Mathematics*, 8:331–343, 1974.

[9] G. R. Grimmett. Random labelled trees and their branching networks. *J. Austral. Math. Soc. (Ser. A)*, 30:229–237, 1980.

[10] A. Hurwitz. Über Abel's Verallgemeinerung der binomischen Formel. *Acta Math.*, 26:199–203, 1902.

[11] D.E. Knuth. Discussion on Mr. Riordan's paper. In R.C. Bose and T.A. Dowling, editors, *Combinatorial Mathematics and its Applications*, pages 71–91. Univ. of North Carolina Press, Chapel Hill, 1969.

[12] R. Otter. The multiplicative process. *Ann. Math. Statist.*, 20:206–224, 1949.

[13] A. G. Pakes and T. P. Speed. Lagrange distributions and their limit theorems. *SIAM Journal on Applied Mathematics*, 32:745–754, 1977.

[14] J. Pitman. Coalescent random forests. Technical Report 457, Dept. Statistics, U.C. Berkeley, 1996. To appear in *J. Comb. Theory A*. Available via http://www.stat.berkeley.edu/users/pitman.

[15] J. Pitman. Abel-Cayley-Hurwitz multinomial expansions associated with random mappings, forests and subsets. Technical Report 498, Dept. Statistics, U.C. Berkeley, 1997. Available via http://www.stat.berkeley.edu/users/pitman.

[16] J. Pitman. The multinomial distribution on rooted labeled forests. Technical Report 499, Dept. Statistics, U.C. Berkeley, 1997. Available via http://www.stat.berkeley.edu/users/pitman.

[17] J. Pitman. Enumerations of trees and forests related to branching processes and random walks. In D. Aldous and J. Propp, editors, *Microsurveys in Discrete Probability*, number 41 in DIMACS Ser. Discrete Math. Theoret. Comp. Sci, pages 163–180, Providence RI, 1998. Amer. Math. Soc.

[18] A. Rényi. On the enumeration of trees. In R. Guy, H. Hanani, N. Sauer, and J. Schonheim, editors, *Combinatorial Structures and their Applications*, pages 355–360. Gordon and Breach, New York, 1970.

[19] L.W. Shapiro. Voting blocks, reluctant functions, and a formula of Hurwitz. *Discrete Mathematics*, 87:319–322, 1991.

[20] M. Sibuya, N. Miyawaki, and U. Sumita. Aspects of Lagrangian probability distributions. *Studies in Applied Probability. Essays in Honour of Lajos Takács (J. Appl. Probab.)*, 31A:185–197, 1994.

[21] V. Strehl. Identities of the Rothe-Abel-Schläfli-Hurwitz-type. *Discrete Math.*, 99:321–340, 1992.