LOGISTIC ORDER STATISTICS[1]

By Allan Birnbaum and Jack Dudman

New York University and Reed College

0. Summary. Expected values and standard deviations of order statistics from a logistic distribution are given for sample sizes $n = 1, 2, \cdots, 10$ and $n = 15, 20, 50$, and 100, and summarized in graphs to facilitate interpolation to other sample sizes. For $n \leqq 10$ the results are compared with asymptotic approximations. Related distribution theory is discussed briefly, and related investigations and applications are cited.

1. Distribution theory. A random variable Y will be called *logistic*, or will be said to have the logistic distribution, if $\text{Prob}\{Y \leqq y\} = \Psi(\alpha + \beta y)$, where $-\infty < \alpha < \infty, \beta > 0, -\infty < y < \infty$, and $\Psi(t) = e^t/(1 + e^t) = 1/(1 + e^{-t})$. The corresponding density function is $\beta\psi(\alpha + \beta y)$, where $\psi(t) = \Psi'(t) = e^t/(1 + e^t)^2$. Letting $\bar{\Psi}(t) = 1 - \Psi(t)$, $\psi(t) = \Psi(t)\ \bar{\Psi}(t)$. Tables of Ψ and ψ have been given by Berkson [1].

Interest in the logistic distribution stems in part from the fact that it very closely approximates the normal distribution and seems an equally plausible assumption in many contexts of application [2], [3], and from its convenient mathematical and statistical properties. Haley [10] has shown that the "distance" $d(\beta) = \max_y|\Psi(\beta y) - \Phi(y)|$, between the logistic c.d.f. $\Psi(\beta y)$ with zero mean and scale parameter β and the standard normal c.d.f. $\Phi(y)$, is minimized by taking $\beta = \beta' = 0.5875$, which gives $d(\beta') < 0.01$. Almost as good a fit in this sense to the standard normal is given by the *standard logistic* c.d.f. $\Psi(\beta'' y)$ with zero mean and unit variance, in which $\beta'' = 1/1.81380 = 0.5513$ (as indicated below).

Let $Y_1 \leqq Y_2 \leqq \cdots \leqq Y_n$ denote a sample of n independent observations from a logistic distribution with $\alpha = 0, \beta = 1$, ordered by increasing size. Then it is found readily that the moment generating function (m.g.f.) of an order statistic Y_s is

$$M(w, n, s) = E(e^{wY_s}) = \Gamma(s + w)\, \Gamma(n - s + 1 - w)/\Gamma(s)\, \Gamma(n - s + 1),$$

where Γ denotes the usual gamma function.

We note that the m.g.f. of Fisher's z distribution with $2s$ and $2(n - s + 1)$ degrees of freedom, where $z = \log(F)^{\frac{1}{2}}$ and F has the F-distribution, is $[(n - s + 1)/s]^{w/2} \cdot M(w/2, n, s)$. (In particular, z with 2 and 2 degrees of freedom has the m.g.f. $M(w/2, 1, 1)$, and hence has the logistic c.d.f. $\Psi(2z)$.)

Received July 4, 1961; revised January 3, 1963.

[1] Work supported in part by the Office of Naval Research. The authors are indebted to the referee for helpful suggestions, including the usefulness of logarithmic scaling in Fig. 1.

Thus

$$\text{Prob}\{Y_s \leqq v\} = \text{Prob}\{2z \leqq v + \log[(n - s + 1)/s]\}$$
$$= \text{Prob}\{F \leqq [(n - s + 1)/s]e^v\}$$

where the z and F random variables each have $2s$ and $2(n - s + 1)$ degrees of freedom. Thus the extensive available tables and theoretical knowledge of the z and F distributions provide conveniently corresponding extensive knowledge of the distribution of logistic order statistics.

2. Moments. Cumulants or moments of z or Y_s have been studied and used in various applications by Cornish and Fisher [6] and [9], Plackett [12], and one of the present authors [3]. The following tables are a by-product of the last work.

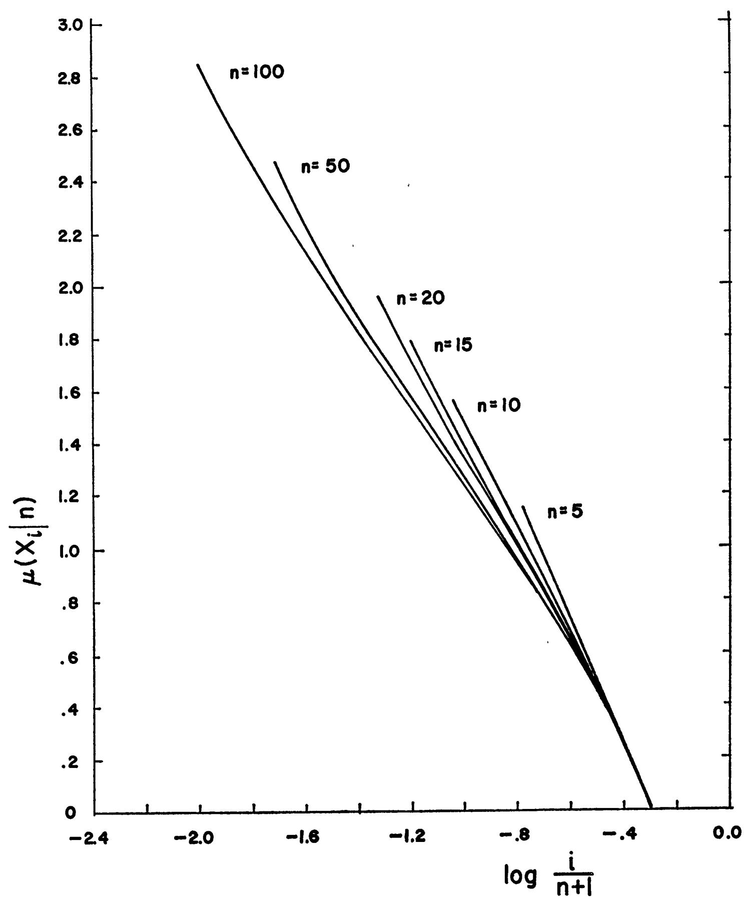

FIG 1.

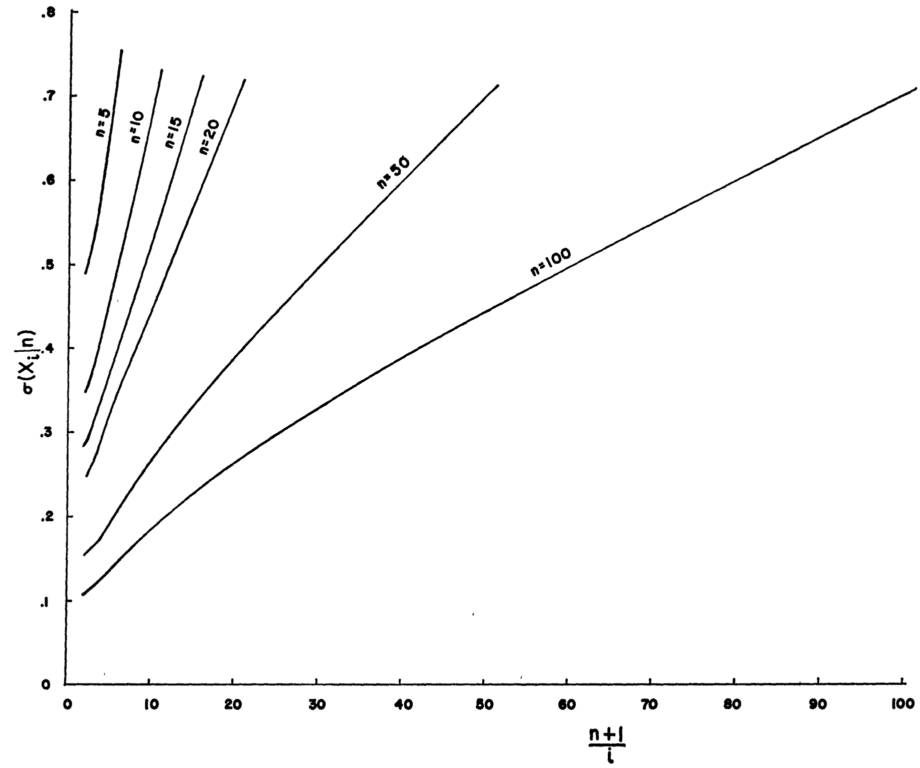

FIG 2.

A comparative study of order statistics from several other distributions, particularly their moments in small samples, was given by Hastings *et al* [11]. The form of their Tables 1 and 2 has been adopted in Table 1 below to facilitate comparisons.

The cumulants of Y_s are given by

$$\kappa_k\,(Y_s) = (\partial^k/\partial w^k) \log M\,(w, n, s)\,|_{w=0} = D_k(s) + (-1)^k\, D_k(n - s + 1),$$

where D_k denotes the polygamma function of order k,

$$D_k(x) = (\partial^k/\partial x^k) \log \Gamma(x),$$

tabulated in [8]. (These formulae for M and the first four cumulants of Y_s were given in [12].) Thus $E(Y_s) = D_1(s) - D_1(n - s + 1)$ and $\mathrm{Var}(Y_s) = D_2(s) + D_2(n - s + 1)$. Setting $n = s = 1$ gives $2D_2(1) = (\beta'')^{-2} = (1.81380)^2$ as the variance of the logistic c.d.f. $\Psi(y)$.

For presentation of tables of moments of logistic order statistics, we adopt the notational scheme of [11] to facilitate comparisons with results therein for other distributions. For each n and s, let $i = n - s + 1$, and let $X_i = \beta'' Y_s$. Then $X_1 \geqq X_2 \geqq \cdots \geqq X_n$ denotes a sample of n independent observations from a *standard logistic distribution* $\Psi(x/\beta'')$ with zero mean and unit variance. Let

$$\mu(X_i \mid n) = E(X_i) = \beta'' E(Y_s) = \beta''[D_1(n - i + 1) - D_1(i)]$$

and

$$\sigma(X_i \mid n) = [\text{Var}(X_i)]^{\frac{1}{2}} = \beta''[D_2(i) + D_2(n - i + 1)]^{\frac{1}{2}}.$$

TABLE 1

Means and standard deviations of order statistics from standard logistic and standard normal distributions. (L = exact logistic, AL = asymptotic logistic, N = exact normal, AN = asymptotic normal. Values for N and AN are from [11].)

		$\mu(X_i \mid n)$				$\sigma(X_i \mid n)$			
n	i	N	AN	L	AL	N	AN	L	AL
1	1	0	0	0	0	1.00000	1.2533	1.0000	1.103
2	1	.56419	.4307	.5513	.382	.82565	.9168	.8343	0.827
3	1	.84628	.6745	.8270	.606	.74798	.7867	.7874	.736
3	2	0	0	0	0	.66983	.7236	.6262	.637
4	1	1.02938	.8416	1.0108	.764	.70122	.7144	.7657	.690
4	2	0.29701	.2533	.2757	.223	.60038	.6340	.5622	.564
5	1	1.16296	.9674	1.1486	.887	.66898	.6670	.7532	.661
5	2	0.49502	.4307	.4594	.382	.55814	.5798	.5313	.524
5	3	0	0	0	0	.53557	.5605	.4900	.494
6	1	1.26721	1.0676	1.2589	.987	.64492	.6331	.7451	.643
6	2	0.64176	.5659	.5973	.503	.52874	.5426	.5131	.498
6	3	0.20155	.1800	.1838	.159	.49620	.5147	.4542	.455
7	1	1.35218	1.1504	1.3508	1.073	.62603	.6702	.7394	.630
7	2	0.75737	.6745	.7075	.606	.50670	.5150	.5012	.482
7	3	0.35271	.3186	.3216	.282	.46875	.4826	.4328	.431
7	4	0	0	0	0	.45874	.4737	.4154	.417
8	1	1.42360	1.2207	1.4295	1.147	.61066	.5867	.7352	.620
8	2	0.85222	.7647	.7994	.691	.48930	.4936	.4927	.469
8	3	0.47282	.4307	.4319	.382	.44807	.4584	.4182	.414
8	4	0.15251	.1397	.1378	.123	.43264	.4447	.3918	.392
9	1	1.48501	1.2816	1.4984	1.212	.59780	.5691	.7319	.613
9	2	0.93230	.8416	.8782	.764	.47508	.4763	.4863	.460
9	3	0.57197	.5244	.5238	.467	.43171	.4393	.4083	.401
9	4	0.27453	.2533	.2481	.223	.41303	.4277	.3760	.376
9	5	0	0	0	0	.40751	.4178	.3668	.367
10	1	1.53875	1.3352	1.5597	1.270	.58681	.5557	.7294	.580
10	2	1.00135	.9085	.9471	.829	.46318	.4619	.4814	.453
10	3	0.65608	.6046	.6025	.541	.41826	.4238	.4006	.392
10	4	0.37572	.3488	.3340	.309	.39756	.4052	.3646	.363
10	5	0.12274	.1142	.1103	.100	.38857	.3973	.3498	.350

TABLE 2

Means and standard deviations of order statistics from the standard logistic distribution

n	i	$\frac{i}{n+1}$	$\mu(X_i \mid n)$	$\sigma(X_i \mid n)$	n	i	$\frac{i}{n+1}$	$\mu(X_i \mid n)$	$\sigma(X_i \mid n)$
15	1	0.0625	1.79268	0.72177		10	0.1961	0.79919	0.19868
	2	0.1250	1.20197	0.46748		11	0.2157	0.73028	0.19137
	3	0.1875	0.88390	0.37993		12	0.2353	0.66602	0.18523
	4	0.2500	0.65418	0.33569		13	0.2549	0.60557	0.18003
	5	0.3125	0.46622	0.31016		14	0.2745	0.54826	0.17559
	6	0.3750	0.30082	0.29510		15	0.2941	0.49356	0.17181
	7	0.4375	0.14768	0.28704		16	0.3137	0.44105	0.16856
	8	0.5000	0	0.28450		17	0.3333	0.39038	0.16580
						18	0.3529	0.34124	0.15345
20	1	0.0476	1.95597	0.71931		19	0.3725	0.29338	0.16148
	2	0.0952	1.37562	0.46094		20	0.3922	0.24658	0.15984
	3	0.1429	1.06933	0.37069		21	0.4118	0.20064	0.15852
	4	0.1905	0.85312	0.32355		22	0.4314	0.15537	0.15748
	5	0.2381	0.68083	0.29475		23	0.4510	0.11062	0.15672
	6	0.2857	0.53381	0.27581		24	0.4706	0.06623	0.15621
	7	0.3333	0.40254	0.26302		25	0.4902	0.02205	0.15596
	8	0.3810	0.28137	0.25450					
	9	0.4286	0.16651	0.24927	100	1	0.0099	2.85444	0.70926
	10	0.4762	0.05513	0.24677		2	0.0198	2.29754	0.44623
						3	0.0297	2.01625	0.35095
50	1	0.0196	2.46952	0.71143		5	0.0495	1.68322	0.26544
	2	0.0392	1.90694	0.44978		10	0.0990	1.24248	0.18795
	3	0.0588	1.61978	0.35559		15	0.1485	0.97815	0.15656
	4	0.0784	1.42428	0.30464		20	0.1980	0.78164	0.13914
	5	0.0980	1.27446	0.27194		25	0.2475	0.62046	0.12818
	6	0.1176	1.15194	0.24889		30	0.2970	0.48031	0.12089
	7	0.1373	1.04752	0.23165		35	0.3465	0.35344	0.11598
	8	0.1569	0.95594	0.21822		40	0.3960	0.23505	0.11278
	9	0.1765	0.87380	0.20747		45	0.4455	0.12178	0.11094
						50	0.4950	0.01103	0.11027

Values of $\mu(X_i \mid n)$ and $\sigma(X_i \mid n)$ for $1 \leqq i \leqq [n + 1/2]$, $1 \leqq n \leqq 10$ are given in Table 1 under heading L (logistic). By symmetry, $\mu(X_i \mid n) = -\mu(X_{n-i+1} \mid n)$ and $\sigma(X_i \mid n) = \sigma(X_{n-i+1} \mid n)$. Table 2 gives such values also for $n = 15$, 20, 50 and 100. Figures 1 and 2 give graphs of $\mu(X_i \mid n)$ and $\sigma(X_i \mid n)$ as functions of $\log(i/n + 1)$ and $n + 1/i$ respectively, for $n = 5, 10, 15, 20, 50$ and 100. These may be useful for interpolation to other values of n.

As in [11], Table 1 also includes approximations given by the asymptotic formulas (following Cramér [7], p. 369) $\mu(X_i \mid n) \doteq \beta'' \log(q_i/p_i)$ and $\sigma(X_i \mid n) \doteq \beta''(np_iq_i)^{-\frac{1}{2}}$, where $p_i = i/n + 1$ and $q_i = n - i + 1/n + 1$. These approximations appear under heading AL (asymptotic logistic).

It would be desirable to augment these results by exact computations of covariances of some logistic order statistics, $\text{Cov}(X_i, X_j)$; evidently numerical integration is required, and this has not been undertaken. The asymptotic approximation to the covariance of two order statistics takes a particularly simple form with the standard logistic distribution:

$$\mathrm{Cov}(X_i, X_j) \doteq \beta''(nq_ip_j)^{-1}, \qquad i \leq j.$$

These relations greatly simplify the formal solutions of various problems of determination of best linear combinations of order statistics for various estimation problems concerning parameters of a logistic distribution, but such formal solutions must be regarded with caution because of the asymptotic approximations used. Exact determination of some covariances of logistic order statistics would be useful to help determine the range of effective accuracy of such approximations and of results derived from them.

Blom [4], [5] has given a class of "nearly best unbiased" estimators, linear in order statistics, which, under general conditions satisfied here, are nearly as efficient as the best unbiased such estimators, and whose determination requires the means but not the covariances of order statistics. Thus such estimators can be determined by use of Tables 1 and 2.

REFERENCES

[1] Berkson, Joseph (1957). Tables for the maximum likelihood estimate of the logistic function. *Biometrics* **13** 28–34.

[2] Berkson, Joseph (1951). Why I prefer logits. *Biometrics* **7** 327–339.

[3] Birnbaum, Allan (1958). Statistical theory of tests of a mental ability (abstract). *Ann. Math. Statist.* **29** 1285.

[4] Blom, Gunnar (1956). On linear estimates with nearly minimum variance. *Ark. Mat.* **3** 365–369.

[5] Blom, Gunnar (1958). *Statistical Estimates and Transformed Beta-Variables.* Wiley, New York.

[6] Cornish, E. A. and Fisher, R. A. (1937). Moments and cumulants in the specification of distributions. *Rev. Inst. Internat. Statist.* **5** 307–322.

[7] Cramér, Harald (1946). *Mathematical Methods of Statistics.* Princeton Univ. Press.

[8] Davis, H. T. (1933). *Tables of the Higher Mathematical Functions* **1, 2.** Principia, Bloomington.

[9] Fisher, R. A. and Cornish, E. A. (1960). The percentile points of distributions having known cumulants. *Technometrics* **2** 209–226.

[10] Haley, D. C. (1952). Estimation of the dosage mortality relationship when the dose is subject to error. Technical Report No. 15, Applied Mathematics and Statistics Laboratory, Stanford Univ.

[11] Hastings, Cecil, Jr., Mosteller, Frederick, Tukey, John W. and Winsor, Charles P. (1947). Low moments for small samples: a comparative study of order statistics. *Ann. Math. Statist.* **18** 413–426.

[12] Plackett, R. L. (1958). Linear estimates from censored data. *Ann. Math. Statist.* **29** 131–142.