PROBABILITY TABLES FOR THE EXTREMAL QUOTIENT[1]

By E. J. Gumbel[2] and James Pickands III

Columbia University and Virginia Polytechnic Institute

0. Summary. Tables are given for the asymptotic distribution function of the extremal quotient, which is valid for a wide class of initial distributions. Selected values of the shape parameter λ are considered. This parameter depends upon the initial distribution, and the size of the sample from which the quotient is drawn. This distribution function is not too different from that of a logarithmic normal distribution. The large sample approach of the distribution function to the logistic one is very slow. A Monte Carlo study shows an unexpectedly good fit to the theory.

1. Distribution of the extremal quotient. Let X be a random variable, symmetrical about zero, that is one such that

$$F(x) = 1 - F(-x). \tag{1}$$

Let X_1 be the smallest, and X_n the largest, observation in a sample of size n. The expression

$$Q = X_n/(-X_1) \tag{2}$$

is called the extremal quotient. Clearly, by symmetry, the probability approaches one that

$$X_1 < 0, \qquad X_n > 0, \tag{3}$$

in which case the extremal quotient is positive. In what follows we assume that n is so large that (3) holds. We define the characteristic largest value u_n, and the extremal intensity α_n, respectively, by

$$F(u_n) = 1 - 1/n, \tag{4}$$

and

$$\alpha_n = nf(u_n). \tag{5}$$

For a large class of initial distributions, the distribution function of the largest value, Prob $\{X_n \leqq x\} = \Phi_n(x)$, converges as n becomes large to

$$\Phi_n(x) = \exp(-e^{-\alpha_n(x-u_n)}). \tag{6}$$

Such initial distributions are said to be of the exponential type.

By symmetry $-X_1$ has the same distribution as X_n. Furthermore, as is well

Received 5 April 1965; revised 5 May 1967.

[1] Prepared with the support of the National Science Foundation, Grant G.K. 695 and the National Institutes of Health, Grant WP 00457.

[2] Deceased.

known, X_1 and X_n are asymptotically independent. So the asymptotic distribution of Q is that of the ratio of two independent positive random variables each having the distribution function (6). In [2] it is shown that the distribution function $H_\lambda(q)$ of the extremal quotient is

$$(7)\quad H_\lambda(q) = [\lambda/(1 - e^{-\lambda})^2]\int_0^1 \exp(-\lambda(z + z^q))\,dz - e^{-\lambda}/(1 - e^{-\lambda})$$

where the parameter $\lambda = \exp(\alpha_n u_n)$, is a function of the initial distribution, and the size n of the sample from which the quotient was drawn. Since u_n has the dimension of X and α_n has the inverse dimension, λ is dimensionless.

For large n and consequently for large values of λ, the expression (7) becomes simply

$$(8)\qquad H_\lambda(q) = \lambda\int_0^1 \exp(-\lambda(z + z^q))\,dz.$$

Since

$$(9)\qquad \log Q = \log X_n - \log(-X_1),$$

and $\log X_n$, and $\log(-X_1)$ are identically distributed and asymptotically independent, $\log Q$ is symmetrical about zero. The most prominent symmetric distribution is the normal one, which is in fact the distribution of the difference of two identically distributed normal variates. Hence, the log-normal distribution is the distribution of the quotient of two independent lognormal variates. It is interesting to note that the two quotients have the same logarithmic symmetry. As shown in [2],

$$(10)\qquad H(1/q) = 1 - H(q).$$

The distribution of the extremal quotient rapidly becomes concentrated with increasing sample size. The concentration is about the median which is unity. In order to compensate for this concentration, the difference $Q - 1$ is multiplied by α_n and u_n. Thus we have the variable

$$(11)\qquad \tau \equiv (Q - 1)\log\lambda = \alpha_n u_n(Q - 1).$$

In [2] it was shown that the distribution of τ approaches the logistic distribution as λ approaches infinity; that is,

$$(12)\qquad \text{for all } x,\ \lim_{\lambda\to\infty} P\{\tau \leqq x\} = 1/(1 + e^{-x}).$$

It should be noted, however, that while a logistic variate has all moments τ has none since Q has none.

2. Tables of the distribution function. Table 1 gives numerical values of the distribution function (7) of the extremal quotient. From equation (10) it is clear that it is sufficient to compute the probabilities for $Q \leqq 1$. To facilitate the computation, a logarithmic transformation was made. First, let $u = \lambda z$, then the integral (8) can be written

$$\int_0^\lambda \exp(-(u + \beta u^q))\,du, \qquad \text{where} \qquad \beta = \lambda^{1-q}.$$

TABLE 1

The distribution function

λ = 2			λ = 5			λ = 9.48		
q	$H_\lambda(q)$	$-\tau$	q	$H_\lambda(q)$	$-\tau$	q	$H_\lambda(q)$	$-\tau$
0.0	0.0000	.69315	0.0	0.0000	1.60944	0.0	0.0000	2.24918
.1	0.0528	.62383	.1	0.0133	1.44894	.1	0.0011	2.02427
.2	0.1120	.55452	.2	0.0426	1.28755	.2	0.0091	1.79935
.3	0.1727	.48520	.3	0.0882	1.12661	.3	0.0337	1.57443
.4	0.2316	.41589	.4	0.1455	0.96566	.4	0.0784	1.34951
.5	0.2870	.34657	.5	0.2086	0.80472	.5	0.1398	1.12459
.6	0.3381	.27726	.6	0.2731	0.64378	.6	0.2114	0.89967
.7	0.3847	.20794	.7	0.3359	0.48283	.7	0.2872	0.67476
.8	0.4270	.13863	.8	0.3952	0.32189	.8	0.3624	0.44984
.9	0.4653	.06931	.9	0.4500	0.16094	.9	0.4340	0.22492
1.0	0.5000	.00000	1.0	0.5000	0.00000	1.0	0.5000	0.00000

λ = 10			λ = 15			λ = 20		
q	$H_\lambda(q)$	$-\tau$	q	$H_\lambda(q)$	$-\tau$	q	$H_\lambda(q)$	$-\tau$
0.0	0.0000	2.30259	0.0	0.0000	2.70805	0.0	0.0000	2.99573
.1	0.0008	2.07233	.1	0.0001	2.43725	.1	0.0000	2.69616
.2	0.0078	1.84207	.2	0.0020	2.16644	.2	0.0000	2.39659
.3	0.0306	1.61181	.3	0.0140	1.89564	.3	0.0077	2.09701
.4	0.0738	1.38155	.4	0.0457	1.62483	.4	0.0317	1.79744
.5	0.1346	1.15129	.5	0.0995	1.35403	.5	0.0792	1.49787
.6	0.2065	0.92103	.6	0.1711	1.08322	.6	0.1487	1.19829
.7	0.2832	0.69078	.7	0.2532	0.81242	.7	0.2330	0.89872
.8	0.3596	0.46052	.8	0.3386	0.54161	.8	0.3240	0.59915
.9	0.4326	0.23026	.9	0.4221	0.27081	.9	0.4146	0.29957
1.0	0.5000	0.00000	1.0	0.5000	0.00000	1.0	0.5000	0.00000

λ = 25			λ = 30			λ = 35		
q	$H_\lambda(q)$	$-\tau$	q	$H_\lambda(q)$	$-\tau$	q	$H_\lambda(q)$	$-\tau$
0.0	0.0000	3.21888	0.0	0.0000	3.40120	0.0	0.0000	3.55535
.1	0.0000	2.89699	.1	0.0000	3.06108	.1	0.0000	3.19981
.2	0.0003	2.57510	.2	0.0001	2.72096	.2	0.0001	2.84428
.3	0.0047	2.25321	.3	0.0031	2.38084	.3	0.0022	2.48874
.4	0.0236	1.93133	.4	0.0184	2.04072	.4	0.0148	2.13321
.5	0.0659	1.60944	.5	0.0564	1.70060	.5	0.0494	1.77767
.6	0.1328	1.28755	.6	0.1207	1.36048	.6	0.1112	1.42214
.7	0.2181	0.96566	.7	0.2063	1.02036	.7	0.1967	1.06660
.8	0.3129	0.64378	.8	0.3040	0.68024	.8	0.2965	0.71107
.9	0.4089	0.32189	.9	0.4042	0.34012	.9	0.4003	0.35553
1.0	0.5000	0.00000	1.0	0.5000	0.00000	1.0	0.5000	0.00000

TABLE 1—*Continued*

$\lambda = 40$			$\lambda = 50$			$\lambda = 60$		
q	$H_\lambda(q)$	$-\tau$	q	$H_\lambda(q)$	$-\tau$	q	$H_\lambda(q)$	$-\tau$
0.0	0.0000	3.68888	0.0	0.0000	3.91202	0.0	0.0000	4.09434
.1	0.0000	3.31999	.1	0.0000	3.52082	.1	0.0000	3.68491
.2	0.0000	2.95110	.2	0.0000	3.12962	.2	0.0000	3.27548
.3	0.0016	2.58222	.3	0.0010	2.73842	.3	0.0006	2.86604
.4	0.0123	2.21333	.4	0.0089	2.34721	.4	0.0069	2.45661
.5	0.0439	1.84444	.5	0.0360	1.95601	.5	0.0304	2.04717
.6	0.1035	1.47555	.6	0.0915	1.56481	.6	0.0826	1.63774
.7	0.1886	1.10666	.7	0.1756	1.17361	.7	0.1655	1.22830
.8	0.2901	0.73778	.8	0.2796	0.78240	.8	0.2711	0.81887
.9	0.3969	0.36889	.9	0.3912	0.39120	.9	0.3866	0.40943
1.0	0.5000	0.00000	1.0	0.5000	0.00000	1.0	0.5000	0.00000

$\lambda = 80$			$\lambda = 100$			$\tau = 144.3$		
q	$H_\lambda(q)$	$-\tau$	q	$H_\lambda(q)$	$-\tau$	q	$H_\lambda(q)$	$-\tau$
0.0	0.0000	4.38203	0.0	0.0000	4.60517	0.0	0.0000	4.97189
.1	0.0000	3.94382	.1	0.0000	4.14465	.1	0.0000	4.47470
.2	0.0000	3.50562	.2	0.0000	3.68414	.2	0.0000	3.97752
.3	0.0003	3.06742	.3	0.0002	3.22362	.3	0.0001	3.48033
.4	0.0045	2.62922	.4	0.0032	2.76310	.4	0.0019	2.98314
.5	0.0233	2.19101	.5	0.0189	2.30259	.5	0.0133	2.48595
.6	0.0700	1.75281	.6	0.0615	1.84207	.6	0.0494	1.98876
.7	0.1504	1.31461	.7	0.1394	1.38155	.7	0.1227	1.49157
.8	0.2581	0.87641	.8	0.2482	0.92103	.8	0.2325	0.99438
.9	0.3794	0.43820	.9	0.3738	0.46052	.9	0.3647	0.49719
1.0	0.5000	0.00000	1.0	0.5000	0.00000	1.0	0.5000	0.50000

$\lambda = 200$			$\lambda = 492.8$			$\lambda = 500$		
q	$H_\lambda(q)$	$-\tau$	q	$H_\lambda(q)$	$-\tau$	q	$H_\lambda(q)$	$-\tau$
0.0	0.0000	5.29832	0.0	0.0000	6.20010	0.0	0.0000	6.21461
.1	0.0000	4.76849	.1	0.0000	5.58009	.1	0.0000	5.59315
.2	0.0000	4.23865	.2	0.0000	4.96008	.2	0.0000	4.97169
.3	0.0000	3.70882	.3	0.0000	4.34007	.3	0.0000	4.35023
.4	0.0011	3.17899	.4	0.0003	3.72006	.4	0.0003	3.72876
.5	0.0097	2.64916	.5	0.0040	3.10005	.5	0.0039	3.10730
.6	0.0405	2.11933	.6	0.0230	2.48004	.6	0.0228	2.48584
.7	0.1092	1.58950	.7	0.0782	1.86003	.7	0.0778	1.86438
.8	0.2191	1.05955	.8	0.1848	1.24002	.8	0.1843	1.24292
.9	0.3567	0.52983	.9	0.3349	0.62001	.9	0.3346	0.62146
1.0	0.5000	0.00000	1.0	0.5000	0.00000	1.0	0.5000	0.00000

TABLE 1—*Continued*

$\lambda = 1000$			$\lambda = 2000$		
q	$H_\lambda(q)$	$-\tau$	q	$H_\lambda(q)$	$-\tau$
0.0	0.0000	6.90776	0.0	0.0000	7.60090
.1	0.0000	6.21698	.1	0.0000	6.84081
.2	0.0000	5.52620	.2	0.0000	6.08072
.3	0.0000	4.83543	.3	0.0000	5.32063
.4	0.0001	4.14465	.4	0.0000	4.56054
.5	0.0020	3.45388	.5	0.0010	3.80045
.6	0.0146	2.76310	.6	0.0093	3.04036
.7	0.0596	2.07233	.7	0.0453	2.28027
.8	0.1607	1.38155	.8	0.1396	1.52018
.9	0.3183	0.69078	.9	0.3024	0.76009
1.0	0.5000	0.00000	1.0	0.5000	0.00000

TABLE 2

The normal parameters u_n, α_n, λ_n

Probability F	Sample Size n		Product $u_n\alpha_n = \lg \lambda_n$	Parameter λ_n	
.90	1.000	1	2.24910	9.479	0
.95	2.000	1	3.39286	2.975	1
.96	2.500	1	3.77158	4.345	1
.98	5.000	1	4.97193	1.443	2
.990	1.000	2	6.20022	4.928	2
.992	1.250	2	6.60030	7.353	2
.995	2.000	2	7.44917	1.718	3
.996	2.500	2	7.85465	2.577	3
.998	5.000	2	9.12405	9.173	3
.9990	1.000	3	10.40509	3.302	4

NOTE: The column following the entries for n and λ_n indicates the power of ten by which the preceding number is to be multiplied.

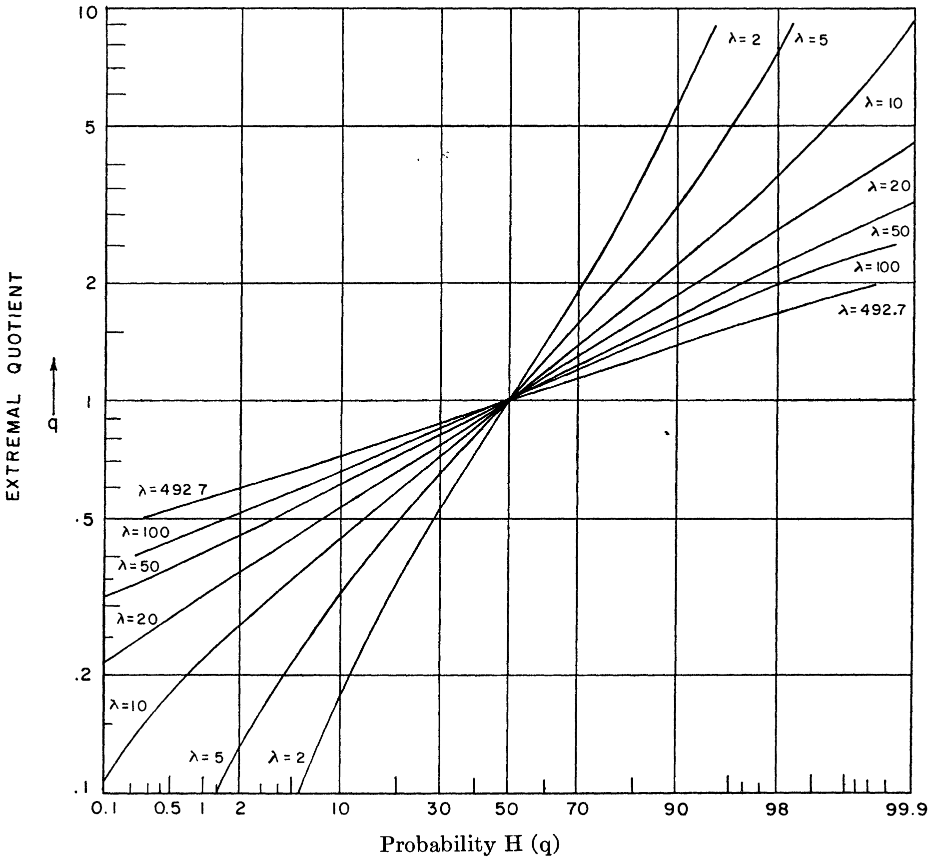
Let $u = -\log \tau$. Then the integral is given by

$$\int_{e^{-\lambda}}^{1} \exp\left(-\beta(-\log \tau)^{q}\right) d\tau.$$

The integral was taken to be the average of the values of the intergrand at the points $1/2m$, $3/2m$, $5/2m$, $\cdots$, $2m - 1/2m$. The integrand is taken to be zero if the argument is less than $e^{-\lambda}$. The value of m was successively doubted until there was agreement to four decimal places. Throughout most of the range of the computation the values of m which gave agreement were 576 and 1152. To have used the original form (8) of the integral would have multiplied the cost of the computations by at least forty.

The distribution functions of the extremal quotient are traced in Graph 1 on logarithmic normal paper for the parameters $\lambda = 2, 5, 10, 20, 50, 100,$ and 492.

GRAPH 1 TABLE 1
Distribution Function of the Extremal Quotient



The curve for $\lambda = 1000$ is indistinguishable from the curve for $\lambda = 492$. It is worth mentioning that the distribution function of the extremal quotient plots nearly as a straight line for large values of λ, although the moments of the extremal quotient do not exist, while all moments of the lognormal distribution do exist.

The values of τ satisfying the relationship (11) are given in Table 1, and are plotted in Graph 2 on logistic probability paper. The approach towards a straight line foreseen by the theory, is very slow and does not yet hold for those values of λ for which the distribution function was computed.

As analytical examples of symmetrical distributions of the exponential type, consider the following three. The first Laplacean distribution function is

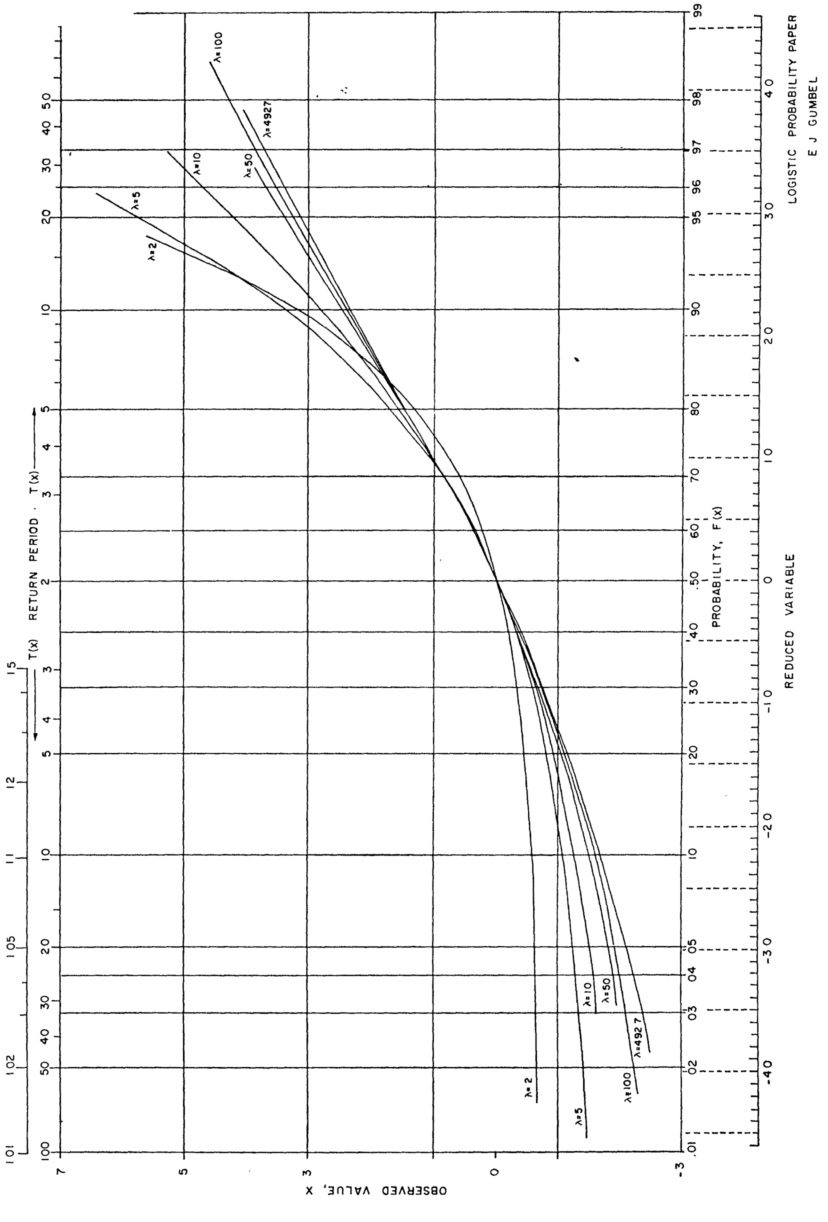
$$F(x) = e^x/2, \qquad \text{if } x \leqq 0,$$
$$= 1 - (e^{-x}/2), \qquad \text{if } x \geqq 0.$$

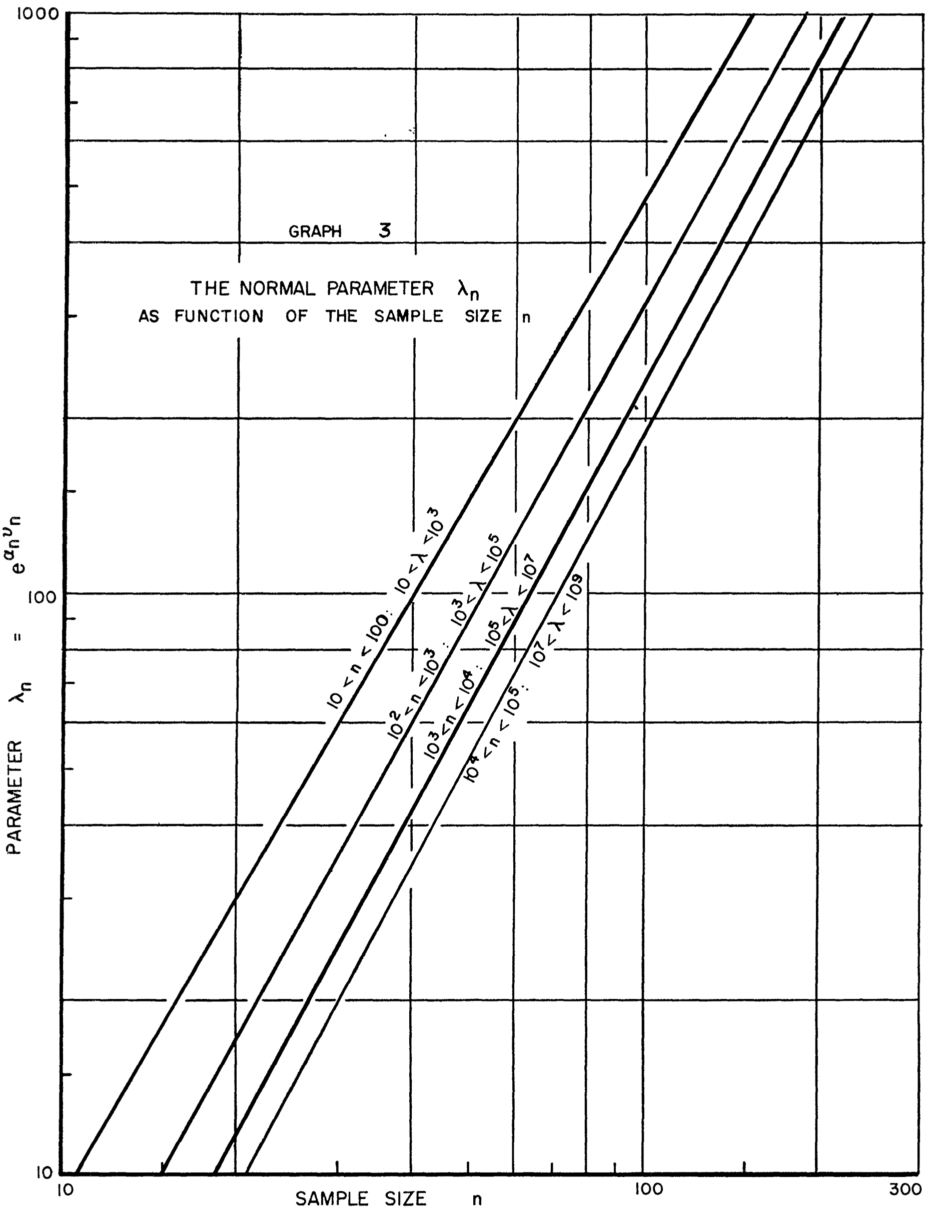
It follows that

$$u_n = \log (n/2), \qquad \alpha_n = 1, \qquad \lambda_n = n/2.$$

GRAPH 2 TABLE 1

Logistic Approximation, T = (Q−1) (LG λ)





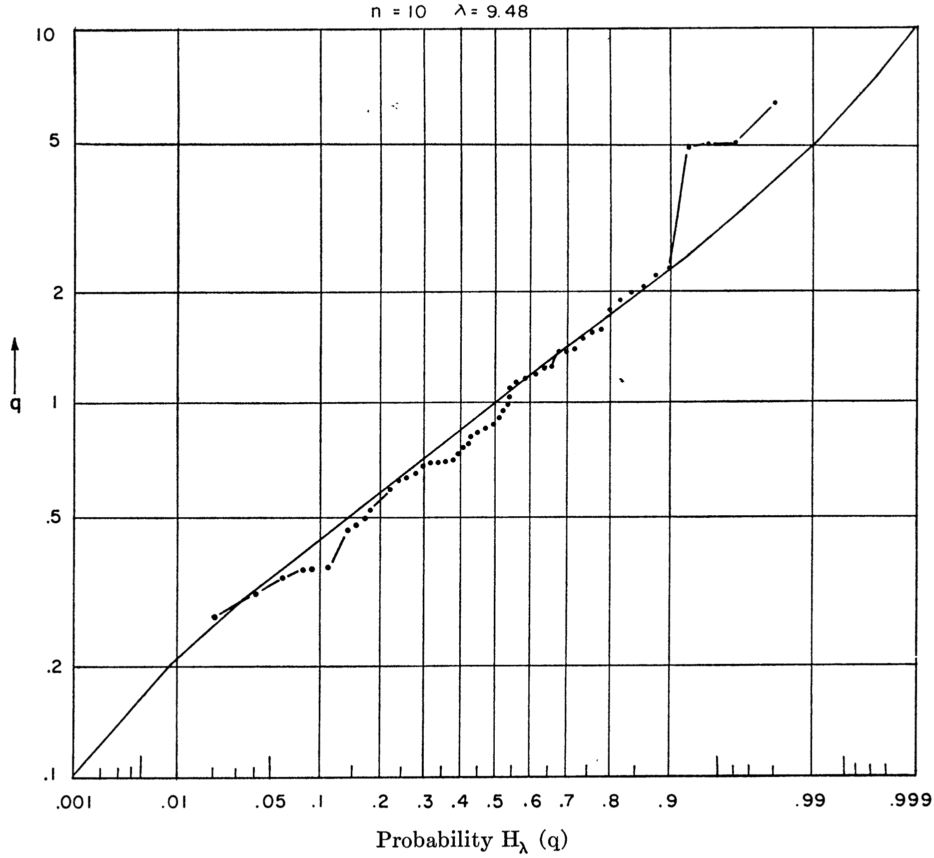
For the logistic distribution function given in equation (12),

$$u_n = \log (n - 1), \qquad \alpha_n = 1 - 1/n.$$

In the normal case, with mean zero, and variance 1, the two parameters u_n and α_n cannot be written in closed form.

GRAPH 4

Extremal Quotient from the Normal Samples of Size



Therefore the numerical values of the parameter λ_n are needed. The definitions (4) and (5) of u_n and α_n lead with the help of the usual normal tables, to the values given in Table 2. The results are traced in Graph 3. The product $\alpha_n u_n$ is practically a linear function of the logarithm of n. As shown in [1], pages 138–9,

$$\log \lambda_n = (2 \log n - \log (2\pi))(1 + o(1)),$$

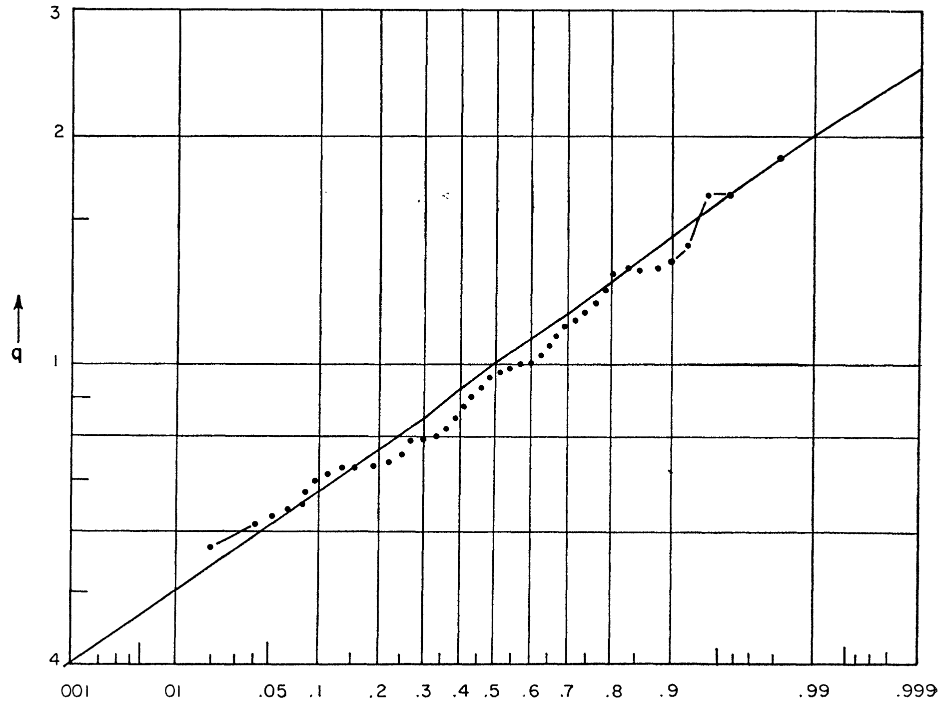
as $n \longrightarrow \infty$.

The graph is read as follows. For $10 < n < 100$, the first slanting line on the left is read to give λ_n. For values on n from 100 to 1000, the second line gives $10^2\lambda_n$. For n in the ranges 10^3 to 10^4, and 10^4 to 10^5 respectively, the third and fourth lines give $10^4\lambda_n$, and $10^6\lambda_n$.

3. Empirical results. In order to test the preceding results, some sampling experiments were made. Although it is known that the distribution of the largest normal value approaches the extremal distribution only slowly, normal extremes were chosen because no other random numbers of an unlimited symmetrical type were available. From the Rand Corporation tables [3], 49 samples were taken, Graph 5

Extremal Quotient from the Normal Samples of Size

n = 50 λ = 144.3

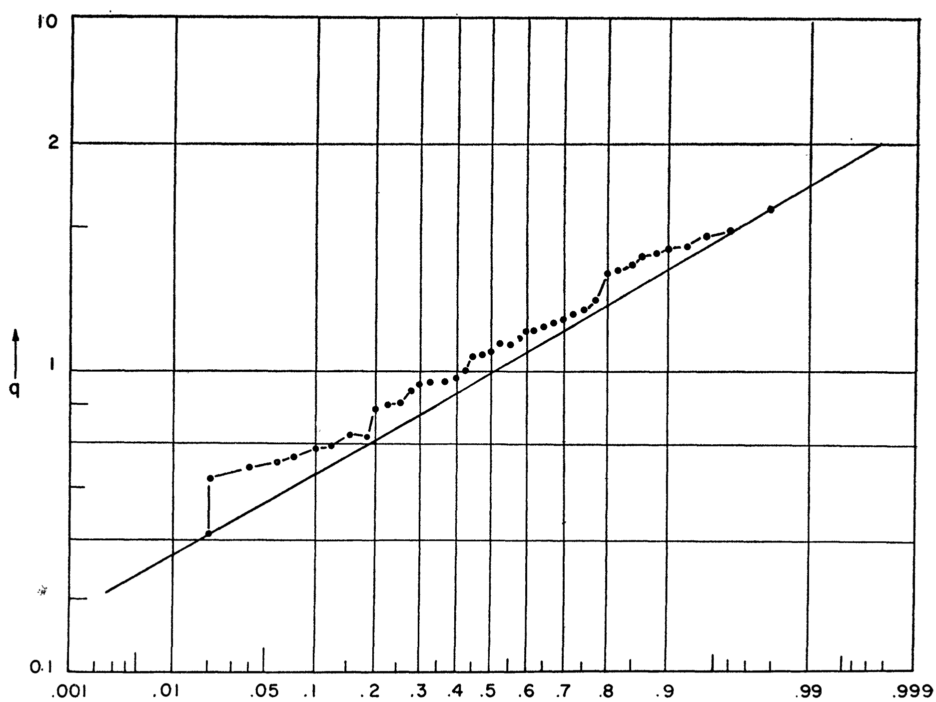


Probability H_λ (q)

Graph 6

Extremal Quotient from the Normal Samples of Size

n = 100 λ = 492.70



Probability H_λ (q)

each of size $n = 10$. For each sample, the extremal quotient was computed. The value of λ corresponding to $n = 10$ is 9.48, as given in Table 2. The 49 observed extremal quotients are plotted on logarithmic normal paper in Graph 4, together with the theoretical curve given in Table 1. Except for a large unexplained jump at the end, the agreement between the observations, and the theoretical values is quite good.

Then 49 samples were taken, each of size fifty, and the same procedure was followed. In this case, $\lambda = 144.3$. The quotients are plotted on Graph 5. Since the distribution contracts considerably with increasing λ, the scale was twice the previous one. The approximation of the sample to the theoretical distribution is excellent.

The same procedure was then followed using 49 samples each of size 100, in which case $\lambda = 492.8$. The results are plotted in Graph 6. The fit to the theory is marred by the fact that all observations, though scattered about a straight line, lie above the theoretical curve.

4. Estimation. In [2], it is proposed to estimate the parameter λ by comparing the expected proportion of the sample for which $\frac{1}{2} < Q < 2$, with the observed proportion. Clearly, however, unless λ is extremely small, all of the sample will lie in this range with very high probability. So this method is not applicable.

The cumbersome nature of the distribution function makes it plain that even on a high speed computer, maximum likelihood estimation would not be feasible. But Q has no moments, and the median is 1, independent of λ. Therefore, neither quantile nor moment methods will suffice either.

The authors have been unable to find a method of estimating this parameter.

REFERENCES

[1] Gumbel, E. J. (1960). *Statistics of Extremes*. Columbia Univ. Press.

[2] Gumbel, E. J. and Keeney, R. D. (1950). The extremal quotient. *Ann. Math. Statist.* **21** 523–528.

[3] The RAND Corporation (1955). A Million Random Digits with 100,000 Normal Deviates.