THE EXTREMAL QUOTIENT

By E. J. Gumbel and R. D. Keeney

New York City and Metropolitan Life Insurance Company

Summary. The extremal quotient is defined as the ratio of the largest to the absolute value of the smallest observation. Its analytical properties for symmetrical, continuous and unlimited distributions are obtained from a study of the auto-quotient defined as the ratio of two non-negative variates with identical distributions. The relation of the two statistics is established by proving that, for sufficiently large samples from an initial distribution with median zero, the largest (or smallest) value may be assumed to be positive (or negative) and that the extremes are independent. It follows that the distribution and the probability of the extremal quotient possess certain symmetries, and that its median is unity. As many moments exist for the extremal quotient as moments and reciprocal moments exist simultaneously for the initial variate. The logarithm of the extremal quotient is symmetrically distributed. These properties hold for all continuous symmetrical unlimited variates which possess a monotonically increasing probability function.

For the exponential type, the asymptotic distribution of the extremal quotient can only be expressed by an integral. In this case, no moments exist. For the Cauchy type, the asymptotic distribution is very simple, and the logarithm of the extremal quotient has the same distribution as the midrange for initial distributions of the exponential type.

It is not necessary to consider asymmetrical distributions since, in this case, for sufficiently large samples, one of the extremes will outweigh the other, unless the distribution is nearly symmetrical or has rapidly varying tails.

1. The auto-quotient and the extremal quotient. Let x and y be two independent non-negative continuous variates, unlimited to the right. Let $f_1(x)$ and $f_2(y)$ be the distributions (probability densities), and let $F_1(x)$ and $F_2(y)$ be the probability functions. Then the joint distribution of the two variates is their product. The quotient

$$Q = x/y \tag{1.1}$$

is also non-negative and unlimited to the right. Since

$$x = yQ; \qquad \frac{dx}{dQ} = y,$$

the joint distribution $\mathrm{w}(y, Q)$ of the quotient Q and the variate y is

$$\mathrm{w}(y, Q) = f_1(yQ)f_2(y)\cdot y, \tag{1.2}$$

and the marginal distribution $h(Q)$ of the variate Q alone becomes

$$h(Q) = \int_0^\infty y f_1(yQ) f_2(y)\, dy. \tag{1.3}$$

The quotient Q possesses a mode if (and only if) $f_1(x)$ possesses a mode.

Assume now that the two variates x and y have the same distribution

$$f_1(x) = f(x); \qquad f_2(y) = f(y) \tag{1.4}$$

with the same parameter values. The quotient of two variates with identical distributions is henceforth called the *auto-quotient* q_a. It may be realized if there are two independent series of observations taken from the same population and ordered in time. Each value from the first series is divided by the corresponding value from the second series. Another realization consists in dividing each value obtained in one series of independent observations by every other value. A third realization is obtained by considering two asymmetrical distributions $f_1(x)$ and $f_2(y)$ where $x \geqq 0$, $y \leqq 0$, and

$$f_2(y) = f_1(-x). \tag{1.4'}$$

The two distributions are called mutually symmetrical, and the auto-quotient is

$$q_a = x/(-y).$$

From the definition of the auto-quotient it follows that the distribution of q_a must be the same as the distribution of its reciprocal $r = 1/q_a$. The proof of this statement is simple. Under the condition (1.4), the distribution $h(q_a)$ becomes, from (1.3)

$$h(q_a) = \int_0^\infty y f(yq_a) f(y)\, dy. \tag{1.5}$$

The distribution $h_1(r)$ of the reciprocal is

$$h_1(r) = \frac{1}{r^2} \int_0^\infty y f(y/r) f(y)\, dy.$$

If y/r is replaced by x, the distribution of r is

$$h_1(r) = h(q_a). \tag{1.6}$$

Thus, the distribution of the auto-quotient of a non-negative unlimited variate is invariant under a reciprocal transformation.

The shape of the distribution $h(q_a)$ and the location of the mode may be obtained from the density of probability $h(1/q_a)$ at the value $1/q_a$ (which differs, of course, from the distribution $h_1(r)$ of $r = 1/q_a$). From (1.5) we obtain

$$h(1/q_a) = \int_0^\infty y f(y/q_a) f(y)\, dy.$$

The transformation

$$y/q_a = z; \qquad dy = q_a\, dz,$$

leads to

$$h(1/q_a) = q_a^2 h(q_a). \tag{1.7}$$

This is a *symmetry relation* for the distribution of the auto-quotient of a non-negative unlimited variate. If q_a is larger than unity,

$$h(1/q_a) > h(q_a). \tag{1.8}$$

If the distribution $h(q_a)$ is continuous for all values of q_a, the derivative of equation (1.7) with respect to q_a leads, for $q_a = 1$, to

$$h'(1) = -h(1). \tag{1.9}$$

If the distribution $h(q_a)$ possesses a unique mode, it must be less than unity.

The moments $\overline{q_a^k}$ are, from (1.5)

$$\begin{aligned}\overline{q_a^k} &= \int_{qa=0}^{q_a=\infty} \int_{y=0}^{y=\infty} q^k y f(qy) f(y)\, dy\, dq \\ &= \int_{y=0}^{y=\infty} \frac{f(y)}{y^k} \int_{q_a=0}^{q_a=\infty} (q_a y)^k f(q_a y)\, d(q_a y)\, dy.\end{aligned}$$

The inner integral is the moment $\overline{y^k}$ of order k of the initial variate y, and the remaining integral is its reciprocal moment $\overline{y^{-k}}$ of order $-k$. Thus

$$\overline{q_a^k} = \overline{y^k} \cdot \overline{y^{-k}} = \overline{q_a^{-k}}. \tag{1.10}$$

The moments of order k and of order $-k$ of q_a exist if the moments and the reciprocal moments of order k for the initial variate exist simultaneously. The second equation in (1.10) also follows immediately from the invariance of q_a under a reciprocal transformation. Even if the initial distribution possesses all moments, the mean $\bar{q}_a$ need not exist, and the same holds, of course, for the mean error and the higher moments. The procedure, usual in economic and meteorological statistics, of calculating the quotients of two series of independent positive variables in order to test whether this ratio is constant may be misleading, especially if the two series happen to be samples taken from the same population. The theoretical mean need not exist, and the calculated mean of the observed quotients need not characterize the relation between the two series.

The probability function $H(Q)$ of the quotient Q obtained from (1.3) is

$$H(Q) = \int_0^Q \int_0^\infty y f_1(zy) f_2(y)\, dy\, dz.$$

Change of the order of integration leads to

$$H(Q) = \int_0^\infty f_2(y) F_1(Qy)\, dy.$$

The probability function $H(q_a)$ of the auto-quotient obtained from (1.4) is

$$H(q_a) = \int_0^1 F(q_a y)\, dF. \tag{1.11}$$

Integration by parts leads to

$$H(q_a) = 1 - q_a \int_0^\infty F(y) f(q_a y)\, dy. \tag{1.12}$$

The boundary condition, $H(0) = 0$; $H(\infty) = 1$ can immediately be verified if the preceding equation is written in the form

$$H(q_a) = 1 - \int_0^\infty F(z/q_a) f(z)\, dz. \tag{1.13}$$

The probability $H(q_a)$ possesses a symmetry relation which is analogous to (1.7). The probability at the value $1/q_a$ is, from (1.11),

$$H(1/q_a) = \int_0^\infty F(y/q_a) f(y)\, dy.$$

If we introduce the variable of integration

$$y = q_a z,$$

we obtain from (1.12)

$$H(q_a) = 1 - H(1/q_a). \tag{1.14}$$

If q_a is any quantile, such that $H(q_a) = P$, its reciprocal $1/q_a$ has the probability $1 - P$. The first quartile (decile) is the reciprocal of the third quartile, (ninth decile) and so on.

For $q_a = 1$, equation (1.14) leads to

$$H(1) = \tfrac{1}{2}. \tag{1.14$'$}$$

The median of the auto-quotient of a positive unlimited variate is unity. From (1.9) it follows that the median surpasses the mode, if a unique mode exists.

Finally, equation (1.14) may be used to construct a symmetrical distribution. If a new variate

$$z = \lg q_a \tag{1.15}$$

with the probability function $H^*(z)$ is introduced, the symmetry relation (1.14) becomes

$$H^*(z) = 1 - H^*(-z). \tag{1.16}$$

The logarithm of the auto-quotient of a positive unlimited variate has a symmetrical distribution about median zero. The geometric mean of q_a exists and is equal to unity.

These results hold if each observed value of a non-negative unlimited variate is divided by each other observed value. They do not hold for the quotients of two specific order statistics because, in general, the fundamental assumption of independence does no longer hold. However, some consequences for the quotients of extreme mth values may be deduced.

Consider a symmetrical unlimited variate. Then the distribution ${}_m\varphi({}_mx)$ of the mth smallest value ${}_mx$, and the distribution $\varphi_m(x_m)$ of the mth largest value x_m are mutually symmetrical in the sense of (1.4′). Therefore the extremal quotient

$$q_m = \frac{x_m}{-{}_mx} \tag{1.17}$$

may be interpreted as an auto-quotient provided that 1) the probability for x_m to be negative, and ${}_mx$ to be positive, may be neglected; 2) the distributions of the mth smallest and the mth largest values are independent. Under these conditions the distribution, the moments, and the probability function of the extremal quotient are obtained from (1.5), (1.10), and (1.11) respectively, if the initial distribution $f(y)$ is replaced by the distribution of the mth largest values $\varphi_m(x_m)$. The symmetry relations (1.7) and (1.14) and their consequence, that the median is equal to unity, hold in particular for $m = 1$, i.e. for the extremal quotient proper.

The validity of the two conditions has now to be established.

a) Consider a symmetrical distribution $f(x)$ with median zero. Then the probability that the largest among n observations, x_n, is equal to or less than a certain x, is $1 - F^n(x)$. The probability P that the largest among n values is positive, i.e. larger than the median, is

$$P = 1 - 2^{-n}. \tag{1.18}$$

If n is sufficiently large, this probability differs from unity by an amount that can be made as small as we please. Even for relatively small samples, say $n = 20$, the probability that the largest value will be positive is of the order $1 - 10^{-6}$. Thus, we expect only one largest value in a million samples of size 20 to be negative. The same argument shows that the smallest value x_1 may be expected to be negative. Thus the postulate

$$x_n \geqq 0; \qquad x_1 \leqq 0, \tag{1.19}$$

is a very weak restriction upon the sample size. If m is sufficiently small, the same result holds for the mth extremes.

b) It is known [7] that the joint distribution $\mathfrak{w}_n(x_1, x_n)$ of the extremes taken from an initial distribution of the exponential type converges, for sufficiently large samples, toward the product of the asymptotic distribution $\varphi(x_n)$ of the largest value, and ${}_1\varphi(x_1)$ of the smallest value. A similar theorem will now be proven for a general class of continuous distributions.

Let ${}_mx$ be the mth smallest observation; let x_l be the lth largest observation where m and l are small compared to n, n being large. Then the joint distribution $\mathfrak{w}_n({}_mx, x_l)$ is

$$\mathfrak{w}_n({}_mx, x_l) = \frac{n!}{(m-1)!(l-1)!(n-m-l)!} \tag{1.20}$$
$$F({}_mx)^{m-1}(F(x_l) - F({}_mx))^{n-m-l}(1 - F(x_l)^{l-1}f({}_mx)f(x_l)).$$

Now the transformation

$$n(1 - F(x_l)) = \xi; \quad nF({}_mx) = \eta; \quad 0 \leqq \xi \leqq n, \quad 0 \leqq \eta \leqq n, \tag{1.21}$$

due to Cramér ([1], p. 371) is used. Then the joint distribution $v_n(\xi, \eta)$ of the new variates ξ and η becomes

$$v_n(\xi, \eta) = \frac{n!}{n^2(m-1)(l-1)!(n-m-l)!}\left(\frac{\xi}{n}\right)^{m-1}\left(1 - \frac{\xi - \eta}{n}\right)^{n-m-l}\left(\frac{\eta}{n}\right)^{l-1},$$

where $m + l$ is small compared to n. As n increases, $v_n(\xi, \eta)$ converges to

$$v(\xi, \eta) = \left(\frac{\xi^{m-1}e^{-\xi}}{(m-1)!}\right)\left(\frac{\eta^{l-1}e^{-\eta}}{(l-1)!}\right),$$

so that in the limit ξ and η are independent. If now the mild restriction is imposed that $F(x)$ be monotonically increasing, (1.21) defines a one to one transformation, and therefore there must exist an inverse function uniquely defining ${}_mx$ as a function of ξ, and x_l as a function of η. From the limiting independence of ξ and η the limiting independence of the extremes ${}_mx$ and x_l follows at once.

Thus the second condition is fulfilled, and the mth extremal quotient shares all properties of the auto-quotient. This holds also for initial symmetrical distributions which do not possess asymptotic distributions of the extremes.

In the following, the two types of initial distributions of an unlimited variate are considered for which asymptotic distributions of the extremes exist, namely, the exponential and the Cauchy type. For simplicity, only the extremal quotient proper, designated by q, is studied. The two asymptotic probabilities of the extremal quotients for these symmetrical distributions are obtained by introducing the asymptotic distributions of the largest value into the probability function (1.11) of the auto-quotient.

2. Application to the exponential type. For symmetrical distributions of the exponential type the asymptotic distribution of the largest value is

$$\varphi(x) = \alpha \exp\left[-\alpha(x - u) - e^{-\alpha(x-u)}\right], \tag{2.1}$$

where u and α are defined in terms of the initial probability $F(x)$ and the initial distribution $f(x)$ by

$$F(u) = 1 - 1/n; \qquad \alpha = nf(u), \tag{2.2}$$

n being the sample size. The distribution (2.1) will now be simplified by introducing a new parameter λ defined by

$$e^{\alpha u} = \lambda > 0. \tag{2.3}$$

To see the meaning of λ, consider Laplace's first distribution, then the so called logistic [6], and the normal distributions, all of which are of the exponential type. In the first two cases we obtain, from (2.2), after some calculations,

$$\alpha = 1, \qquad u = \lg n - \lg 2; \qquad \alpha = 1 - 1/n, \qquad u = \lg (n - 1), \tag{2.4}$$

whereas for the normal distribution, we have asymptotically

$$\alpha = u = \sqrt{2 \lg (n/\sqrt{2\pi})}$$

and

$$\lambda = n^2/(2\pi). \tag{2.4'}$$

For these distributions, and interpreted in this sense, λ is of the order of the sample size or its square.

From (2.3) and (2.1) the distribution $\varphi(x)$ and the probability function $\Phi(x)$ are

$$\varphi(x) = \alpha\lambda \exp [-\alpha x - \lambda e^{-\alpha x}]; \qquad \Phi(x) = \exp [-\lambda e^{-\alpha x}]. \tag{2.5}$$

In order to fulfill the condition (1.19), namely $\Phi(0) = 0$, the distribution $\varphi(x)$ must be truncated at $x = 0$. This leads to the truncated distribution $\varphi_t(x)$ and the truncated probability $\Phi_t(x)$ where

$$\varphi_t(x) = \frac{\alpha\lambda \exp [-\alpha x - \lambda e^{-\alpha x}]}{1 - e^{-\lambda}}; \qquad \Phi_t(x) = \frac{\exp [-\lambda e^{-\alpha x}] - e^{-\lambda}}{1 - e^{-\lambda}}. \tag{2.6}$$

The asymptotic probability function $H_\lambda(q)$ for the extremal quotient of a symmetrical variate of the exponential type is now obtained from (1.11), if y, $f(y)$, and $F(y)$, are replaced by x, $\varphi_t(x)$ and $\Phi_t(x)$, respectively, and the index a is dropped. Consequently, from (2.6),

$$H_\lambda(q) = \frac{1}{(1 - e^{-\lambda})^2} \int_0^\infty \alpha x \exp [-\alpha x - \lambda e^{-\alpha x} - \lambda e^{-\alpha q x}] \, dx - \frac{e^{-\lambda}}{(1 - e^{-\lambda})^2} \int_0^\infty \alpha\lambda \exp [-\alpha x - \lambda e^{-\alpha x}] \, dx.$$

The transformation

$$e^{-\alpha x} = z; \qquad \alpha e^{-\alpha x} \, dx = -dz$$

leads to

$$H_\lambda(q) = \frac{1}{(1 - e^{-\lambda})^2} \int_0^1 \lambda e^{-\lambda(z + z^q)} \, dz - \frac{e^{-\lambda}}{1 - e^{-\lambda}}. \tag{2.7}$$

This probability of the extremal quotient for initial symmetrical distributions of the exponential type is not truely asymptotic since the parameter λ depends upon n. (See Addendum).

Unfortunately, the expression (2.7) cannot be integrated. Therefore the probability function has to be studied in an analytic way. For this purpose we first recall the general properties

$$H(0) = 0; \qquad H(1) = \tfrac{1}{2}; \qquad H(\infty) = 1,$$

valid for any value of λ. Furthermore, for any λ, we have the symmetry relation (1.14). These properties can be verified at once from (2.7).

The numerical values of $H_\lambda(q)$ can easily be calculated for $q = \frac{1}{2}$ and $q = 2$. Consider a value of λ, say of the order 6. Then formula (2.7) may be written

$$H_\lambda(2) = \int_0^1 \lambda e^{-\lambda(z+z^2)}\, dz \tag{2.8}$$
$$= \sqrt{\lambda}\, e^{\lambda/4} \int_0^1 e^{-\lambda(z+\frac{1}{2})^2} \sqrt{\lambda}\, dz.$$

If we introduce

$$\sqrt{\lambda}\,(z + \tfrac{1}{2}) = \frac{t}{\sqrt{2}}; \qquad \sqrt{\lambda}\, dz = \frac{dt}{\sqrt{2}},$$

the probability $H_\lambda(2)$ becomes a difference of two normal probability integrals,

$$H_\lambda(2) = \sqrt{\pi\lambda}\, e^{\lambda/4}\left[1 - F\left(\sqrt{\frac{\lambda}{2}}\right) - \left(1 - F\left(3\sqrt{\frac{\lambda}{2}}\right)\right)\right],$$

where F stands for the normal probability function.

The second expression may be neglected compared to the first one for $\lambda \geqq 4$, whence

$$H_\lambda(2) = \sqrt{\frac{\lambda}{2}}\, e^{\lambda/4} \int_{\sqrt{\lambda/2}}^{\infty} e^{-t^2/2}\, dt. \tag{2.9}$$

The symmetry relation (1.14) leads to the knowledge of $H_\lambda(\frac{1}{2})$. Thus the three probabilities $H_\lambda(\frac{1}{2})$, $H(1)$, and $H_\lambda(2)$ are known.

To see the influence of λ on $H_\lambda(2)$, we use a method due to R. D. Gordon [4]. This author considers a function R_x defined by

$$R_x = e^{x^2/2} \int_x^{\infty} e^{-t^2/2}\, dt, \quad x > 0, \tag{2.10}$$

and proves that

$$\frac{dR}{dx} = xR - 1 < 0; \qquad \frac{d^2R}{dx^2} = x\frac{dR}{dx} + R > 0.$$

It follows that

$$\frac{d}{dx}(xR) > 0.$$

If we substitute $\sqrt{\lambda/2}$ for x, this inequality may be written, from (2.9) and (2.10),

$$\frac{d}{d\sqrt{\frac{\lambda}{2}}}\left(\sqrt{\frac{\lambda}{2}}\, e^{\lambda/4} \int_{\sqrt{\lambda/2}}^{\infty} e^{-t^2/2}\, dt\right) = 2\sqrt{2\lambda}\,\frac{dH_\lambda(2)}{d\lambda} > 0.$$

Consequently $H_\lambda(2)$ increases with λ whereas, from (1.14), the probability $H_\lambda(\frac{1}{2})$ decreases with λ. The following table gives the probabilities $H_\lambda(2)$ and $H_\lambda(\frac{1}{2})$, (2.9) and their differences

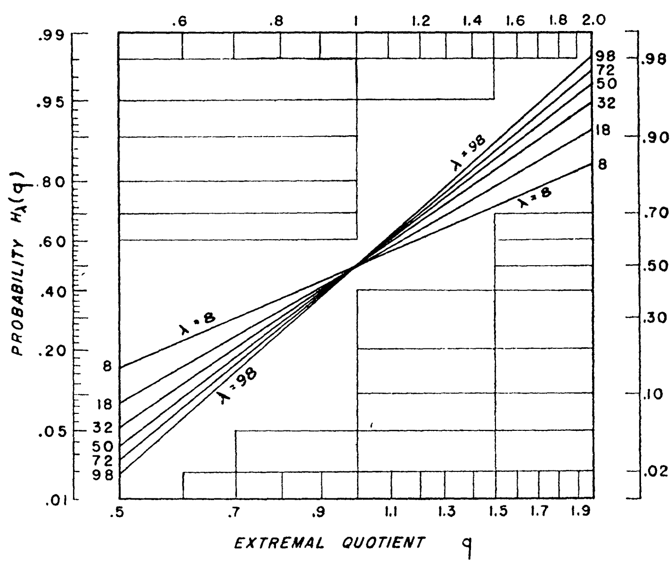
$$P_\lambda(2) = H_\lambda(2) - H_\lambda(\tfrac{1}{2}). \tag{2.11}$$

Asymptotic probabilities of the extremal quotient for symmetrical distributions of the exponential type

Parameter	Probabilities (2.9), (1.14)		Probability (2.11)
λ	$H_\lambda(2)$	$H_\lambda(\frac{1}{2})$	$P_\lambda(2)$
8	.84376	.15624	.68752
18	.91377	.08623	.82754
32	.94661	.05339	.89322
50	.96438	.03562	.92876
72	.97427	.02573	.94854
98	.98087	.01913	.96174

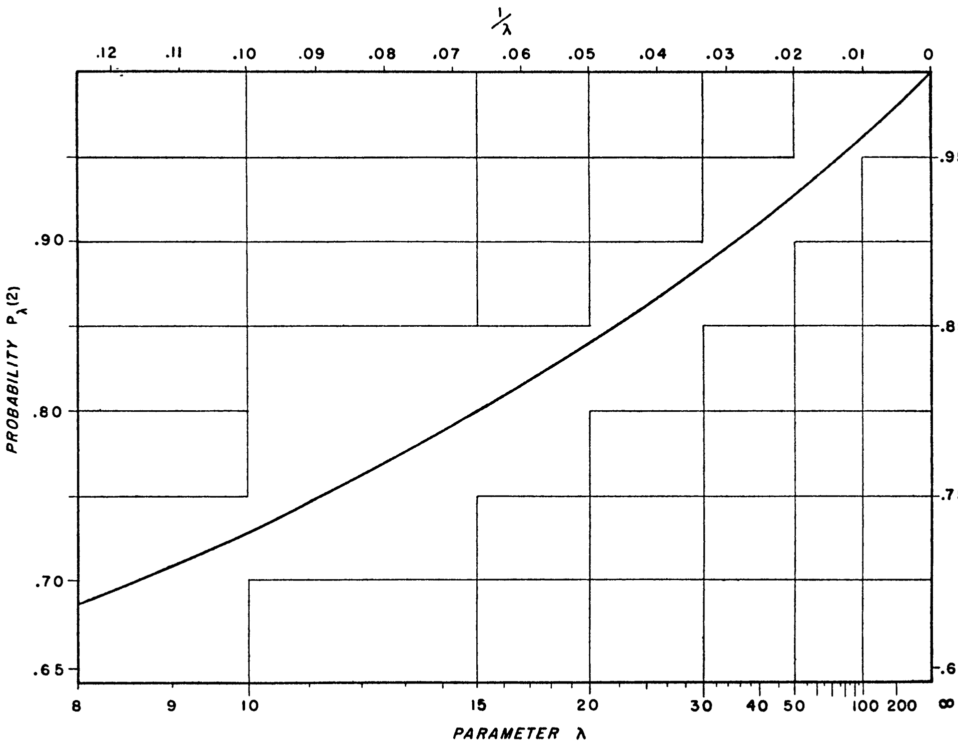
The approximative shape of $H_\lambda(q)$ is traced, for $\lambda = 8, \ldots, 98$, and $\frac{1}{2} < q < 2$ in Graph (1). Since we know from (1.16) that lg q has a symmetrical distribution, we use a logarithmically normal probability paper where q is plotted on the abscissa in a logarithmic scale, and $H_\lambda(q)$ is plotted on the ordinate in a normal probability scale. The probability $P_\lambda(2)$ for any value of q to be contained in the interval $\frac{1}{2} < q < 2$ increases with λ, i.e., with the sample size, and the distribution of the extremal quotient contracts.

1) ASYMPTOTIC PROBABILITY OF THE EXTREMAL QUOTIENT FOR THE EXPONENTIAL TYPE



If the initial distribution is unknown, the parameter λ has to be estimated from the observed extremal quotients. Equation (2.11) may be used for this purpose. We calculate the observed relative frequency $P_\lambda(2)$ of extremal quotients contained between $q = \frac{1}{2}$ and $q = 2$, and substitute it for the probability $P_\lambda(2)$. To facilitate this estimate of λ, we trace $P_\lambda(2)$ against λ in graph (2). The probability $P_\lambda(2)$ is traced on the ordinate in linear scale, and the parameter λ is traced on the abscissa in inverse scale. Thus λ is easily estimated from the observed relative frequency $P_\lambda(2)$.

2) ESTIMATION OF THE PARAMETER λ



The distribution $h_\lambda(q)$ of the extremal quotient obtained by differentiating the probability function (2.7) with respect to q is

$$(2.12) \qquad h_\lambda(q) = \frac{1}{(1 - e^{-\lambda})^2} \int_0^1 \lambda^2 e^{-\lambda(z+z^q)} z^q(-\lg z)\, dz.$$

The symmetry relation (1.7) is easily verified. We now investigate the boundary value $h_\lambda(0)$ and prove that

$$(2.13) \qquad \lim_{q=0} h_\lambda(q) = h_\lambda(0).$$

This is not obvious, since z^q becomes indeterminate if both z and q vanish. For the proof of (2.13), consider the integral

$$(2.14) \qquad \mathrm{I} = \lambda \int_0^1 e^{-\lambda z}(-\lg z)\, dz$$

or

$$\mathrm{I} = (1 - e^{-\lambda}) \lg \lambda - \gamma + e^{-\lambda} \lg \lambda - \epsilon i(-\lambda). \tag{2.15}$$

The last term, the exponential integral, is positive. The value of $h_\lambda(0)$ is thus, from (2.12)

$$h_\lambda(0) = \frac{\lambda e^{-\lambda}(\lg \lambda - \gamma - \epsilon i(-\lambda))}{(1 - e^{-\lambda})^2}, \tag{2.16}$$

The difference

$$\Delta = (1 - e^{-\lambda})^2 (h_\lambda(q) - h_\lambda(0))$$

becomes, from (2.12), (2.15) and (2.16), by the use of the mean value theorem and after expansion

$$\begin{aligned} \Delta &= f(\lambda) \int_0^1 (e^{-\lambda z^q} z^q - e^{-\lambda})\, dz \\ &= f(\lambda) \sum_{\nu=0}^{\infty} \frac{(-1)^\nu \lambda^\nu}{\nu!} \left(\frac{1}{(\nu + 1)q + 1} - 1 \right), \end{aligned}$$

where $f(\lambda)$ is a positive function. Since the series is absolutely convergent, the difference Δ vanishes for $q = 0$, and the density of probability for $q = 0$ is given by (2.16). The condition $h_\lambda(0) \geqq 0$, valid for any distribution, is met provided that

$$\lambda > 1.794 \tag{2.17}$$

By virtue of (2.4) this is a (weak) condition concerning the sample size. From (2.16) it follows that $h_\lambda(0)$ does not vanish although its numerical value is very small.

The existence of at least one mode follows from the fact that the distribution $h_\lambda(q)$ is continuous, very small for $q = 0$, and vanishes for $q = \infty$. Equation (1.9) proves that any mode is inferior to unity. The distribution contracts for increasing values of the parameter. Therefore the mode approaches the median with increasing sample size.

Since the distributions of the exponential type do not possess reciprocal moments it follows from (1.10) that the distribution $h_\lambda(q)$ does not possess moments. The mean extremal quotient $\bar{q}$ diverges. Because the logarithmically normal distribution used in graph (1) as first approximation to the distribution $h_\lambda(q)$ possesses all moments, the distribution $h_\lambda'(q)$ has a much longer tail than the logarithmically normal one.

3. Application to the Cauchy type. For the exponential type, the asymptotic distribution of the extremal quotient can only be expressed in the form of an integral containing a parameter λ which is a function of the sample size. For the Cauchy type, to be defined in the following, the asymptotic distribution will turn out to be very simple.

A distribution of a variate $x \geqq 1$ was said [5] to be of the Pareto type if

$$\lim_{x=\infty} x^k(1 - F(x)) = A; \qquad k > 0; \qquad A > 0. \tag{3.1}$$

We now say that a variate is of the *Cauchy type* if it is unlimited, continuous, subject to (3.1), and symmetrical about zero. Distributions of the Pareto and the Cauchy type do not possess moments of an order equal to or larger than k. However, not all unlimited symmetrical distributions with a finite number of moments are of the Cauchy type.

The simplest example of such a distribution is the Cauchy distribution itself

$$f(x) = \frac{1}{\pi(1 + x^2)}; \qquad F(x) = \tfrac{1}{2} + \frac{1}{\pi} \text{ arc tg } x, \tag{3.2}$$

which possesses no moments. For large absolute values of x, the usual expansion leads to

$$F(x) = 1 - \frac{1}{\pi x} + O(x^{-2}); \qquad F(-x) = \frac{1}{\pi x} - O(x^{-2}).$$

If the factors $O(x^{-2})$ are neglected, the parameters A and k in (3.1) are

$$A = \pi^{-1}; \qquad k = 1. \tag{3.2'}$$

For the Cauchy type, the asymptotic probability $\Pi(x)$ and distribution $\pi(x)$ of the largest value $x = x_n$ established by Fréchet [3], R. A. Fisher [2] and R. von Mises [8] are

$$\Pi(x) = \exp\left[-\left(\frac{u}{x}\right)^k\right]; \qquad \pi(x) = \frac{k}{u}\left(\frac{u}{x}\right)^{k+1} \exp\left[-\left(\frac{u}{x}\right)^k\right], \tag{3.3}$$

where u is defined by (2.2).

The condition (1.19) is fulfilled for any sample size which is so large that the asymptotic distribution of the extremes may be used. The asymptotic probability $H_k(q)$ of the extremal quotient for the Cauchy type is obtained from (1.11), if y, $f(y)$ and $F(y)$ are replaced by x, $\pi(x)$, and $\Pi(x)$, respectively, where the indices n and a are omitted. Consequently, from (3.3),

$$H_k(q) = \int_0^\infty \frac{k}{u}\left(\frac{u}{x}\right)^{k+1} e^{-(u/x)^k-(u/qx)^k}\, dx.$$

From the transformation

$$\left(\frac{u}{x}\right)^k = z; \qquad \frac{k}{u}\left(\frac{u}{x}\right)^{k+1} dx = dz,$$

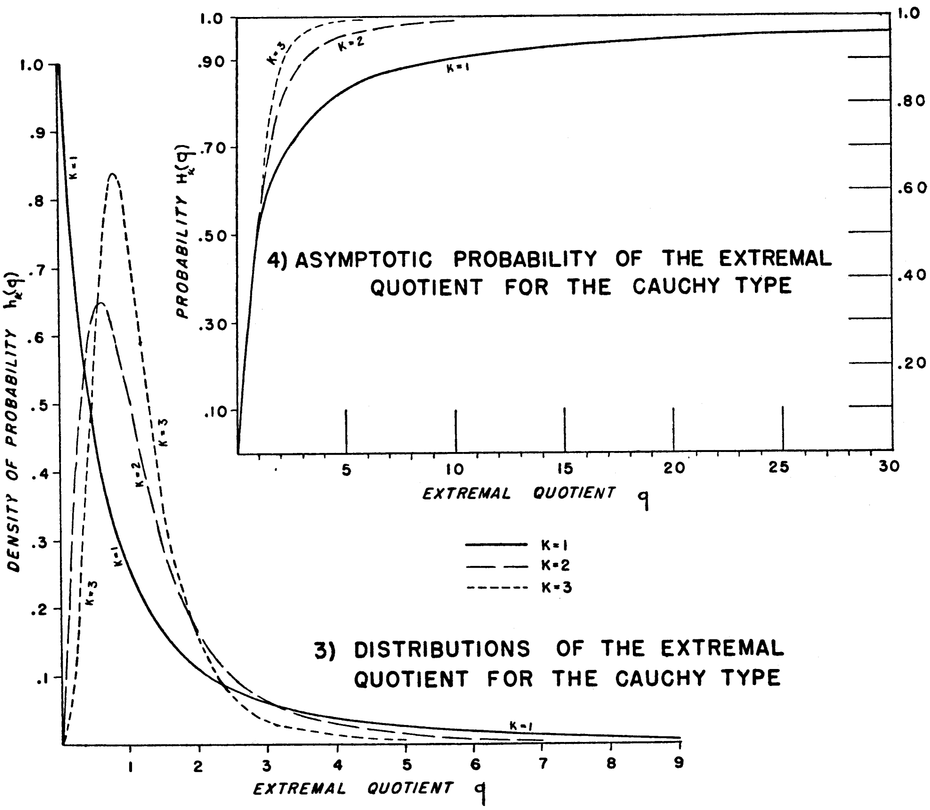
the asymptotic probability $H_k(q)$ and the asymptotic distribution $h_k(q)$ of the extremal quotient become simply

$$H_k(q) = \frac{q^k}{1 + q^k}; \qquad h_k(q) = \frac{kq^{k-1}}{(1 + q^k)^2}, \qquad q \geqq 0. \tag{3.4}$$

Evidently, the symmetry relations (1.7) and (1.14) are fulfilled for any k. The graphs (3) and (4) show the distribution $h_k(q)$ and the probability $H_k(q)$ for the most interesting cases $k = 1, 2, 3$. From

$$\frac{d \lg H_k(q)}{dk} = \lg q(1 - H_k(q))$$

it follows: For k increasing, the probability $H_k(q)$ decreases for $q < 1$, and increases for $q > 1$. *The distribution contracts with increasing values of the parameter* k as shown in the graphs (3) and (4). The more moments that exist in the initial distribution, the more concentrated is the distribution of the extremal quotient.



The density of probability

$$h_k(1) = k/4$$

of the median obtained from (3.4) and (1.14′) increases with k. The mode $\tilde{q}$ of the extremal quotient is obtained from (3.4). For $k > 1$ this leads to

$$\tilde{q}^k = \frac{k-1}{k+1} < 1. \tag{3.5}$$

For $k \leqq 1$ no mode exists, and the distribution diminishes with q. The larger k, the smaller is the distance from the median to the mode, and hence, the smaller the asymmetry. The density of probability of the mode increases with k, and the probability

$$H_k(\tilde{q}) = \tfrac{1}{2}(1 - 1/k) \tag{3.6}$$

approaches $\frac{1}{2}$. The distribution (3.4) belongs to the Pareto type and has no moments of an order equal to or greater than k.

In N samples of sufficiently large size n, the largest quotient q_N^k, defined in the same way as u in equation (2.2), obtained from (3.4)

$$q_N^k = N - 1 \tag{3.7}$$

increases as a root of the number of samples, i.e. very quickly. The higher the order of the highest moments existing, the smaller will the expected largest quotient be.

From (3.4) and the symmetry (1.14) we obtain

$$H_k(q) - H_k(1/q) = 1 - 2/(1 + q^k). \tag{3.8}$$

The larger k, the larger is the percentage of the observations contained in the interval $1/q$ to q.

For a systematic estimate of k, the transformation (1.15) is used. Formula (3.4) leads to the probability $H^*(z)$ and the distribution $h^*(z)$ where

$$H^*(z) = \frac{1}{1 + e^{-kz}}; \qquad h^*(z) = \frac{ke^{-kz}}{(1 + e^{-kz})^2}. \tag{3.9}$$

The logarithm of the extremal quotient for initial distributions of the Cauchy type (where no moments of an order equaling or exceeding k exist) has the logistic distribution, [6], as the midrange $v = x_n + x_1$ for distributions of the exponential type (where all moments exist). The logarithm of the extremal quotient plotted on logistic probability paper should be scattered around a straight line.

The order k of the lowest moment which diverges is obtained from the variance σ_z^2 of the distribution $h^*(z)$ which is [6]

$$\sigma_z^2 = \frac{\pi^2}{3k^2}. \tag{3.10}$$

For the estimate of k from (3.10), σ_z^2 is replaced by the estimate s_z^2 obtained from

$$s_z^2 = \frac{1}{N-1} \sum_{\nu=1}^{N} \lg^2 \frac{x_{n,\nu}}{-x_{1,\nu}}. \tag{3.11}$$

For the Cauchy distribution itself, $k = 1$, and the probability and the distribution of the extremal quotient

$$H_1(q) = q/(1 + q); \qquad h_1(q) = (1 + q)^{-2}$$

are similar to the initial distribution.

The asymptotic distribution of the extremal quotient for initial distributions of the Cauchy type contains one parameter only, the order of the lowest diverging moment in the initial distribution. All other traces of the initial distribution have disappeared.

4. Comparison of the extremal properties for the two types of initial distributions. Assume that the initial distribution is symmetrical, unlimited, and possesses an asymptotic distribution of the extremes. This is not always fulfilled. All moments may exist, and yet the distribution may not belong to the exponential type. No moments may exist, and yet the distribution may not belong to the Cauchy type. If the assumption holds, the initial distribution belongs either to the Cauchy, or to the exponential type.

We take N samples of size n, and estimate the median $\tilde{X}$ of the population from the central value m of the N central values of the samples. Let $X_{1,v}$ and $X_{n,v}$ $(v = 1, 2, \cdots N)$ be the two extremes. If it happens for any v that

$$X_{1,v} > m \text{ or } X_{n,v} < m$$

the sample is too small, and its size has to be increased. The central value q of the observed extremal quotients $q_\nu = (X_{n,\nu} - m)/(m - X_{1,\nu})$ must be near unity.

If the initial distribution is of the exponential type, all moments in the population exist, and the midrange has the logistic distribution. If the initial distribution is of the Cauchy type, no moments of an order greater than k exist, and the logarithm of the extremal quotient has the logistic distribution. The order k can be estimated from the variance (3.11). If all moments in the population diverge, the calculation of the observed moments is futile since they do not characterize the population.

Addendum. The referee of this paper has suggested the following method for obtaining an asymptotic distribution of the extremal quotient for the exponential type. For large values of λ, formula (2.7) becomes, approximately,

$$H_\lambda(q) = \int_0^1 e^{-\lambda(z+z^q)}\, d\lambda z.$$

Let

$$\lambda z = y.$$

Then

$$H_\lambda(q) = \int_0^\lambda \exp\left\{- y\left[1 + \left(\frac{y}{\lambda}\right)^{q-1}\right]\right\} dy.$$

The further transformation

$$e^t = \lambda^{q-1}, q-1 = t/lg\lambda,$$

leads to the probability $H^*(t)$ of the variate t

$$H^*(t) = \int_0^\lambda \exp\{-y[1 + e^{-t}y^{t/\lg\lambda}]\} \, dy,$$

whence asymptotically for $\lambda \to \infty$

$$H^*(t) = \int_0^\infty \exp\{-y(1 + e^{-t})\} \, dy$$
$$= 1/(1 + e^{-t}).$$

Therefore the logistic distribution holds at the same time for both initial types, using the transformation $t = \alpha u(q - 1)$ for the exponential type, and the logarithmic transformation for the Cauchy type.

REFERENCES

[1] H. Cramér, *Mathematical Methods of Statistics*, Princeton University Press, 1946.

[2] R. A. Fisher and L. H. C. Tippett, "Limiting forms of the frequency distribution of the smallest and the largest member of a sample," *Proc. Camb. Philos. Soc.*, Vol. 24 (1928), p. 180.

[3] M. Fréchet, Sur la loi de probabilité de l'écart maximum. *Annales Soc. Polon. Math.*, Vol. 6 (1927).

[4] R. D. Gordon, "Values of Mills ratio of area to boundary ordinate and of the normal probability integral for large values of the argument," *Annals of Math. Stat.*, Vol. 12 (1941), pp. 364–366.

[5] E. J. Gumbel, "The return period of flood flows," *Annals of Math. Stat.*, Vol. 12 (1941), pp. 163–190.

[6] E. J. Gumbel, "Ranges and midranges," *Annals of Math. Stat.*, Vol. 15 (1944), pp. 414–422.

[7] E. J. Gumbel, "On the independence of the extremes in a sample," *Annals of Math. Stat.*, Vol. 17 (1946), pp. 78–81.

[8] R. von Mises, "La distribution de la plus grande de n valeurs," *Revue Math. de l'Union Interbalkanique*, Vol. 1 (1936).