

PARTIAL ATTRACTION OF MAXIMA

BY RICHARD F. GREEN

Technical Report No. 23

Department of Statistics
University of California
Riverside, California 92502

January, 1975

PARTIAL ATTRACTION OF MAXIMA

BY RICHARD F. GREEN

University of California, Riverside

1. Introduction. Certain distributions are stable in the sense that sums of independent random variables having distributions of that type will be of the same type. A distribution such that normed sums of i. i. d. random variables having that distribution converge to a stable law is said to lie in the domain of attraction of that stable law. That is, if there exist sequences $\{a_k\}$, $a_k > 0$ and $\{b_k\}$ such that

$$\frac{\sum_{i=1}^k X_i}{a_k} - b_k \longrightarrow G(x) \text{ weakly}$$

where X_1, X_2, \dots are i. i. d. according to F then F is said to lie in the domain of attraction of G . The normal limit theorem specifies conditions for distributions to lie in the domain of attraction of the normal type. Only stable laws have non-empty domains of attraction.

A weaker condition, namely convergence over some increasing subsequence of sample sizes, defines partial attraction of F to G . That is, if there exist an increasing sequence of natural numbers $\{n_k\}$ and sequences of constants $\{a_k\}$, $a_k > 0$, $\{b_k\}$ such that

$$\frac{\sum_{i=1}^{n_k} X_i}{a_k} - b_k \longrightarrow G(x) \text{ weakly}$$

where X_1, X_2, \dots are i. i. d. according to F then F is said to lie in the domain of partial attraction of G .

A necessary and sufficient condition for a distribution to have a non-empty domain of partial attraction is that it be infinitely divisible. The ideas of attraction and partial attraction are considered in such books as Gnedenko and Kolmogorov (1968) and Feller (1966).

There also exists a theory of stable laws for maxima developed by Fréchet (1927), Fisher and Tippett (1928), von Mises (1936) and Gnedenko (1943) and popularized by Gumbel (1958). A law is stable for maxima if the distribution of the maximum of two independent random variables of that type is of the same type. A distribution F will be said to lie in the domain of attraction for maxima of a law G if there exist sequences a_k , $a_k > 0$ and b_k such that

$$F^k(a_k x + b_k) \xrightarrow{\text{weakly}} G(x)$$

Fisher and Tippett (1928) showed there were three classes of stable laws for maxima and Gnedenko (1943) obtained necessary and sufficient conditions for distributions to lie in the domains of attraction of each of the three types. Only distributions in the three stable classes have non-empty domains of attraction for maxima.

Many distributions frequently used by statisticians lie in the domains of attraction for maxima of the stable laws. Gumbel (1958) has advocated the exploitation of this fact in predicting maximum lifetimes, flood heights and other phenomena involving extremes. Stability for maxima is a mathematical artifact, however. There is

usually no reason for believing that empirical distributions should have tails such that the distribution of normed maxima should converge to some stable type. In fact, maxima can have any distribution and, for the same underlying distribution, the distribution of maxima for a certain sample size may be completely different from that for some other sample size.

It is possible to define partial attraction for maxima analogously to partial attraction for sums. If this is done it will be seen that all distributions have non-empty domains of partial attraction for maxima.

2. Results.

DEFINITION. A distribution F will be said to lie in the domain of partial attraction for maxima of the distribution G if there exists an increasing subsequence of natural numbers $\{n_k\}$ and two sequences of constants $\{a_k\}$, $a_k > 0$, $\{b_k\}$ such that

$$F^{n_k}(a_k x + b_k) \xrightarrow{\text{weakly}} G(x) \text{ as } k \rightarrow \infty.$$

THEOREM 1. Whatever be the distribution G there exists a distribution F that lies in the domain of partial attraction for maxima of G.

PROOF. The proof, by construction, consists of choosing an increasing sequence of sample sizes $\{n_k\}$ and constructing $F(x)$ by putting an ever increasing proportion of $G(x)$ into ever decreasing intervals of $F(x)$.

$$\text{Let } n_k = 10(2^k + 2^{k-1}), \text{ and}$$

$$\begin{aligned}
F(x) &= 99x/50 \text{ for } x \in [0, \frac{1}{2}), \\
&= (G(a_k x + b_k))^{1/n_k}, \text{ for } x \in I_k, \\
&= 1 \text{ for } x \geq 1, \text{ where} \\
I_k &= [1 - 2^{-k}, 1 - 2^{-k-1}), \text{ and} \\
a_k &= (\beta_k - \alpha_k) 2^{k+1}, \\
b_k &= \alpha_k - (\beta_k - \alpha_k) (1 - 2^{-k}) 2^{k+1}, \\
\alpha_k &= G^{-1}((1 - 10^{-2^k})^{n_k}), \\
\beta_k &= G^{-1}((1 - 10^{-2^{k+1}})^{n_k}).
\end{aligned}$$

Here $G^{-1}(y) = \sup(x: G(x) \leq y)$.

For any natural number k the maximum of n_k observations of a variable with distribution $F(x)$ will lie in the interval I_k with probability

$$\begin{aligned}
(1 - 10^{-2^{k+1}})^{n_k} - (1 - 10^{-2^k})^{n_k} &\sim \\
&\sim \exp(-10^{-2^{k+1}}) - \exp(-10^{-2^k})
\end{aligned}$$

which goes to one as k increases. But for sample size n_k values of $\max(X_1, X_2, \dots, X_{n_k})$ in I_k are transformed to values coming from distribution G . \square

Note that the convergence here is at all points, not just at points of continuity.

It is possible to find laws $F(x)$ such that the distribution of the maxima of variables distributed according to these laws converge through different subsequences of sample sizes to different distributions.

This is made precise in the following theorem.

THEOREM 2. Let G_1, G_2, \dots be any countable sequence of probability laws. Then there exists a distribution F that lies simultaneously in the domain of partial attraction for maxima of all the laws G_1, G_2, \dots .

PROOF. The proof is similar to that of Theorem 1 except that here $F(x)$ is constructed by using first one, then another of the laws G_i instead of the same G each time. This is so that each law G_i is included in $F(x)$ denumerably many times.

Let $n_k = 10^{(2^k + 2^{k-1})}$, and

$$\begin{aligned} F(x) &= 99x/50 \text{ for } x \in [0, \frac{1}{2}), \\ &= G_{f(k)}((a_k x + b_k))^{1/n_k} \text{ for } x \in I_k, \\ &= 1 \text{ for } x \geq 1, \text{ where} \end{aligned}$$

$$I_k = [1 - 2^{-k}, 1 - 2^{-k-1}), \text{ and}$$

$f(k) = k - (g_2^{(k)})$, where $g(k)$ is the largest natural number such that $(g_2^{(k)}) < k$. Note that the function $f(k)$ simply produces the familiar sequence of indices 1, 1, 2, 1, 2, 3, 1, 2, 3, 4, 1, 2, etc.

$$a_k = (\beta_k - \alpha_k) 2^{k+1},$$

$$b_k = \alpha_k - (\beta_k - \alpha_k) 2^{k+1} (1 - 2^{-k}),$$

$$\alpha_k = G_{f(k)}^{-1}((1 - 10^{-2^k})^{n_k}),$$

$$\beta_k = G_{f(k)}^{-1}((1 - 10^{-2^{k+1}})^{n_k}). \quad \square$$

Since the family of distribution functions is separable

in the topology of weak convergence we have from Theorem 2 the following theorem.

THEOREM 3. There exists a distribution F that lies simultaneously in the domain of partial attraction for maxima of all probability laws.

PROOF. Let G_1, G_2, \dots be a countable collection of distributions dense in the family of all distributions and apply Theorem 2 to the G_i 's. \square

As Feller remarked concerning some corresponding results for sums, these theorems about partial attraction of maxima possess "curiosity value". They do serve, however, to illustrate an important point, namely, that tails of distributions don't have to be such that the maxima of the random variables they rule will approach some stable limiting distribution. Inferences based on such unwarranted assumptions about the tails of distributions should be regarded with suspicion.

REFERENCES

- FELLER, W. (1966). An Introduction to Probability Theory and Its Applications, Volume II, John Wiley & Sons, New York.
- FISHER, R. A. and TIPPETT, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. Proc. Camb. Phil. Soc. 24 180-190.
- FRÉCHET, M. (1927). Sur la loi de probabilité de l'écart maximum. Ann. de la Soc. Polonaise de Math. (Cracow) 6 93-116.
- GNEDENKO, B. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. Ann. of Math. 44 423-453.
- GNEDENKO, B. V. and KOLMOGOROV, A. N. (1968). Limit Distributions for Sums of Independent Random Variables, Addison-Wesley, Reading, Massachusetts.
- GUMBEL, E. J. (1958). Statistics of Extremes, Columbia University Press, New York.
- MISES, R. VON. (1936). La distribution de la plus grande de n valeurs. Revue math. de l'Union Interbalkanique (Athens) 1 1-20.