5 fr.

Erasmus University Rotterdam
(Netherlands School of Economics,
Econometric Institute)

Report 7612/S

SAMPLE EXTREMES: AN ELEMENTARY INTRODUCTION

by Laurens de Haan

juni 1976.

SAMPLE EXTREMES: AN ELEMENTARY INTRODUCTION

Laurens de Haan

I. Introduction

This note is of a didactic nature. We will derive the limit distributions for maxima of i.i.d. random variables and give sufficient conditions for their domains of attraction. These results can be found in the works of Gnedenko [2] and von Mises [4]. The present exposition contains new proofs which are intended to be sufficiently simple to be used in an elementary course of probability theory.

II. The limit distributions

We will be concerned with the following problem.
Suppose $X_1, X_2, \ldots$ are independent real-valued random variables with common distribution function (df) F. We define for $n = 1, 2, \ldots$

$$Y_n = \max (X_1, X_2, \ldots, X_n).$$

We remark that one can interpret all the results that follow as results for minima by noting that $\min(X_1, \ldots, X_n) = -\max(-X_1, \ldots, -X_n)$.

It follows from the independence of the X_i that

$$P\{Y_n \leq x\} = P\{X_1 \leq x, \ldots, X_n \leq x\} = F^n(x).$$

We ask for conditions which enables one to choose sequences of real constants $a_n > 0$ and $b_n (n = 1, 2, \ldots)$ such that the sequence

$$P\{a_n^{-1} (Y_n - b_n) \leq x\} = F^n (a_n x + b_n)$$

converges weakly to a non-degenerate df G as $n\to\infty$. We first investigate which distribution functions actually can occur as the limit of such a sequence.

Definition 1. A distribution function F is said to belong to the domain of attraction of a non-degenerate distribution function G (notation $F \in \mathcal{D}(G)$) if there exist sequences of real numbers $a_n > 0$ and b_n $(n = 1, 2, \ldots)$ such that

$$(1) \qquad \lim_{n\to\infty} F^n(a_n x + b_n) = G(x)$$

for all continuity points x of G (notation $F^n(a_n x + b_n) \xrightarrow{w} G(x)$).

Our problem can thus be formulated as follows: find all distribution functions with non-empty domains of attraction.

The lemma below (due to Khinchine) implies that a df F cannot be in the domain of attraction of two essentially different df's.

Lemma 1. Let F_n be a sequence of df's and suppose that there exists a non-degenerate df G and sequences of constants $a_n > 0$ and b_n such that

$$(2) \qquad F_n(a_n x + b_n) \xrightarrow{w} G(x).$$

Then for some non-degenerate G_* and sequences of constants $\alpha_n > 0$ and β_n we have

$$(3) \qquad F_n(\alpha_n x + \beta_n) \xrightarrow{w} G_*(x)$$

if and only if for some $a > 0$ and b

$$(4) \qquad \lim_{n\to\infty} a_n^{-1}\alpha_n = a \quad \text{and} \quad \lim_{n\to\infty} a_n^{-1}(\beta_n - b_n) = b.$$

Moreover then for all real x

$$(5) \qquad G_*(x) = G(ax + b).$$

Proof. From geometrical considerations, for instance by using the symmetry of the Lévy metric, it follows that (2) holds if and only if the inverse functions of the left side as functions of x converge weakly to the inverse of the right side of (2), i.e. (2) is equivalent with

$$(6) \qquad \frac{F_n^{-1}(x) - b_n}{a_n} \overset{w}{\to} G^{-1}(x)$$

where the inverse functions are determined in any way consistent with the monotonicity. We can transform (3) similarly.

This immediately implies that (2) and (4) imply (3) and (5).

Now G is non-degenerate i.e. G^{-1} assumes at least two different values, say $G^{-1}(x) > G^{-1}(x_0)$.

The same for G_* with $G^{-1}(x_1^*) > G^{-1}(x_0^*)$. Take $y_1 = \max(x_1, x_1^*)$ and $y_0 = \min(x_0, x_0^*)$. We may assume that y_1 and y_0 are continuity points of G^{-1} and G_*^{-1}.

Applying (6) for x and y_0 we obtain

$$(7) \qquad \frac{F_n^{-1}(x) - F_n^{-1}(y_0)}{a_n} \overset{w}{\to} G^{-1}(x) - G^{-1}(y_0)$$

hence with (6) we have

$$\frac{b_n - F_n^{-1}(y_0)}{a_n} \to G^{-1}(y_0).$$

Applying (7) for $x = y_1$ we get

$$\frac{F_n^{-1}(y_1) - F_n^{-1}(y_0)}{a_n} \to G^{-1}(y_1) - G^{-1}(y_0) > 0.$$

If (3) holds similar relations are true for β_n and α_n with G replaced by G_*. Then (4) follows. As above this implies (5). □

The results of the lemma lead to the following definition.

<u>Definition 2</u>. The df's F_1 and F_2 are of the same <u>type</u> if there exist two constants $a > 0$ and b such that

$$F_2(x) = F_1(ax + b) \quad \text{for all real } x.$$

Clearly this relation between F_1 and F_2 is symmetric, reflexive and transitive. Hence it gives rise to equivalence classes of distribution functions called <u>types</u>. Sometimes we shall indicate a type by one representative of the equivalence class.

By the lemma above the domains of attraction of two df's are identical if these df's are of the same type, and disjoint otherwise so that we can speak of the domain of attraction of a type of df's.

<u>Theorem 1</u>. A non-degenerate df G has a non-empty domain of attraction if and only if there exist real functions $A(s) > 0$ and $B(s)$ defined for $s > 0$ such that

$$G^s(A(s)x + B(s)) = G(x) \tag{8}$$

for all real x and $s > 0$.

<u>Proof</u>. If (3) holds, then by definition $G \in \mathcal{D}(G)$ i.e. $\mathcal{D}(G)$ is non-empty.

Conversely, let $\mathcal{D}(G)$ be non-empty. Then

$$F^n(a_n x + b_n) \xrightarrow{w} G(x)$$

holds for some F, $\{a_n\}$ and $\{b_n\}$. Take $s > 0$, and let $[ns]$ denote the integer part of ns, then

$$\lim_{n\to\infty} F^{[ns]}(a_{[ns]}x + b_{[ns]}) = G(x)$$

and hence

$$\lim_{n\to\infty} F^n(a_{[ns]}x + b_{[ns]}) = \{G(x)\}^{1/s}.$$

Since $G^{1/s}$ is a non-degenerate df we may apply lemma 1 with $\alpha_n = a_{[ns]}$ and $\beta_n = b_{[ns]}$. The conclusion of the theorem follows. □

Remark. It is clear that for G to have a non-empty domain of attraction it is sufficient that (3) holds for s = 2, 3, 4, ...

The theorem suggests the following definition.

Definition 3. A non-degenerate df G is called (max-)stable if there exist real constants $A_n > 0$ and B_n such that

$$G_n\,(A_n x + B_n) = G(x)$$

for all real x and n = 1, 2, ...

In the proof of the main theorem on the class of (max-)stable df's we need the following well-known result.

Lemma 2. Let the function u be non-decreasing. If u satisfies

$$u(t+s) = u(t).u(s)$$

for all real t and s, then either u(t) = 0 for all t or else $u(t) = e^{\rho t}$ for some real constant ρ.

Proof. Suppose $u(t_0) \neq 0$ for some $t_0 \neq 0$. Suppose e.g. $t_0 > 0$. Then for all integers m

$$u(mt_0) = u(t_0)\; u((m-1)t_0) = \ldots = \{u(t_0)\}^m .$$

Hence also for all integers n

$$u(mt_0) = u(n.\frac{mt_0}{n}) = \{u(\frac{mt_0}{n})\}^n ,$$

so

$$u(\frac{m}{n}\, t_0) = \{u(t_0)\}^{m/n} .$$

Now the set $\{\frac{m}{n}.t_0 \mid m,n \in \mathbb{N}\}$ is dense in $\mathbb{R}^+$ and u is monotone, so it follows that u is continuous and $u(tt_0) = \{u(t_0)\}^t$ for t > 0.

Now $u(-t) = u(t)\;\{u(2t)\}^{-1} = \{u(t_0)\}^{-t/t_0}$ and the proof is complete. □

Theorem 2 (Gnedenko). Every stable distribution is of one of the following types:

$$(9)\qquad \Phi_\alpha(x) = \begin{cases} 0 & \text{for } x \leq 0 \\ \exp(-x^{-\alpha}) & \text{for } x > 0 \end{cases}$$

$$(10)\qquad \Psi_\alpha(x) = \begin{cases} \exp -(-x)^{\alpha} & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$$

$$(11)\qquad \Lambda(x) = \exp(-e^{-x}).$$

In (9) and (10) α is a positive constant.

Proof. It is obvious that the df's mentioned above are stable. Conversely suppose G is a stable df i.e. (8) holds for $s > 0$ and real x. Then for $0 < G(x) < 1$

$$-\log -\log G\,(A(s)x + B(s)) - \log s = -\log -\log G(x).$$

Let U be an inverse of $-\log -\log G(x)$. It follows

$$\frac{U\,(x + \log s) - B(s)}{A(s)} = U(x)$$

for all $s > 0$ and all real x. Subtracting the same relation for $x = 0$ we get

$$\frac{U\,(x + \log s) - U\,(\log s)}{A(s)} = U(x) - U(0).$$

Define $A_1(y) = A(e^y)$, $\tilde{U}(x) = U(x) - U(0)$ and $y = \log s$, then for all real y and x

$$(12)\qquad \tilde{U}\,(x + y) - \tilde{U}(y) = \tilde{U}(x).A_1(y).$$

The same equation holds with x and y interchanged, hence for all x and y

(13) $\tilde{U}(x)(1-A_1(y)) = \tilde{U}(y)\ (1 - A_1(x))$.

There are two cases

a) Suppose $A_1(x) = 1$ for all x. Then from (12)

$$\tilde{U}(x + y) = \tilde{U}(x) + \tilde{U}(y).$$

Application of lemma 2 (with a logarithmic transformation) gives $\bar{U}(x) = \rho x$ for some $\rho > 0$, hence G is of type Λ.

b) Suppose there is an x with $A_1(x) \neq 1$. We shall see that this then must be true for all $x \neq 0$. Suppose by contradiction $A_1(y) = 1$ for some $y \neq 0$, then by (13) $\tilde{U}(y) = 0$, hence by (12) $\tilde{U}\ (x + y) - \tilde{U}(x) = 0$ for all x which is impossible. So $A_1(x) \neq 1$ for all $x \neq 0$ and (13) implies for some real $c_1 \neq 0$ and all x

$$\tilde{U}(x) = c_1(1 - A_1(x)).$$

Now (12) gives for $\tilde{\tilde{U}}(x) = c_1^{-1}\ \tilde{U}(x) + 1 = c_1^{-1}\ (U(x) - U(0)) + 1$

$$\tilde{\tilde{U}}\ (x + y) = \tilde{\tilde{U}}(x)\ .\ \tilde{\tilde{U}}(y).$$

By lemma 2 then $\tilde{\tilde{U}}(x) = e^{\rho x}$ with $\rho \neq 0$ because $A_1(x) \neq 1$. For $\rho > 0$ this implies that G is of type $\Phi_{\rho^{-1}}$ and for $\rho < 0$ that G is of type $\Psi_{\rho^{-1}}$. □

III. Domains of attraction

We seek conditions under which maxima from a df F converge to a specified limit distribution. Sufficient conditions will be derived. These are due to von Mises. The conditions are easy to derive, suitable for applications and close to the necessary and sufficient conditions. Our proofs differ from the original ones given by von Mises.

Theorem 3. Suppose F has a positive density F' for all $x \geq x_1$. If for some $\alpha > 0$

$$\lim_{x\to\infty} \frac{x.\ F'(x)}{1 - F(x)} = \alpha,$$

then $F \in \mathcal{D}(\Phi_\alpha)$

Proof. Write $\alpha(x) = \dfrac{x\ F'(x)}{1 - F(x)}$, then for $x \geq x_1$

$$\int_{x_1}^{x} \frac{\alpha(t)}{t}\, dt = -\log(1 - F(x)) + \log(1 - F(x_1)).$$

Hence $$1 - F(x) = \{1 - F(x_1)\} \exp - \int_{x_1}^{x} \frac{\alpha(t)}{t}\, dt.$$

Determine a_n such that $1 - F(a_n) = n^{-1}$,

then for $x > 0$

$$n\{1 - F(a_n x)\} = \exp - \int_{a_n}^{a_n x} \frac{\alpha(t)}{t}\, dt$$

$$= \exp - \int_{1}^{x} \frac{\alpha(a_n s)}{s}\, ds$$

and the latter term converges to $x^{-\alpha}$ as $n\to\infty$.

Hence

$$F^n(a_n x) = \left(1 - \frac{n\{1 - F(a_n x)\}}{n}\right)^n$$

converges to $\Phi_\alpha(x)$ for all x. □

Remark. Two properties of the df's satisfying the conditions of the theorem are also present in the general case: if $F \in \mathcal{D}(\Phi_\alpha)$ then $F(x) < 1$ for all x and (1) holds with $b_n = 0$ for all n.

Theorem 4. Suppose F has a density F' which is positive in some interval (x_1, x_0) and vanishes for $x > x_0$. If for some $\alpha > 0$

$$\lim_{x \uparrow x_0} \frac{(x_0 - x)\, F'(x)}{1 - F(x)} = \alpha,$$

then $F \in \mathcal{D}(\Psi_\alpha)$.

Remark. The condition holds with $\alpha = 1$ if F' has a positive limit for $x \uparrow x_0$.

Proof. As in the previous proof we get for $x_1 < x < x_0$

$$1 - F(x) = \{1 - F(x_1)\}.\ \exp - \int_{x_1}^{x} \frac{\alpha(t)}{x_0 - t}\, dt \text{ with } \alpha(t) = \frac{(x_0 - t)F'(t)}{1 - F(t)}.$$

Determine a_n such that $1 - F(a_n) = n^{-1}$, then for $x < 0$ (with $u = (x_0 - t)/(x_0 - a_n)$)

$$n\{1 - F(x_0 + (x_0 - a_n)x)\} = \exp - \int_{x_0 - (x_0 - a_n)}^{x_0 + (x_0 - a_n)x} \frac{\alpha(t)}{x_0 - t}\, dt$$

$$= \exp \int_{1}^{-x} \frac{\alpha(x_0 + u(x_0 - a_n))}{u}\, du$$

and the latter term converges to $(-x)^\alpha$ as $n \to \infty$ (note that $a_n \uparrow x_0$). Hence $F^n(x_0 + (x_0 - a_n)x) = \left(1 - \frac{n\{1 - F(x_0 + (x_0 - a_n)x)\}}{n}\right)^n$ converges to $\Psi_\alpha(x)$ for all x. □

Remark. One can prove that if $F \in \mathcal{D}(\Psi_\alpha)$ then there always exists a finite x_0 such that $F(x) < 1$ if and only if $x < x_0$. Moreover (1) holds with $b_n = x_0$ for all n.

The domain of attraction of Λ is somewhat more involved. Then both $F(x) < 1$ for all x (as for $D(\Phi_\alpha)$) and $F(x) = 1$ for some x (as for $D(\Psi_\alpha)$) are possible. We need the following lemma.

Lemma 3. Let f be a positive differentiable function on (x_1, x_0) where $x_0 \leq \infty$. Suppose

$$\lim_{t \uparrow x_0} f'(t) = 0 \quad \text{and} \quad \lim_{t \uparrow x_0} f(t) = 0,$$

then

$$\lim_{t \uparrow x_0} \frac{f(t + x f(t))}{f(t)} = 1$$

uniformly in any bounded interval.

Remark. If $x_0 = \infty$, the second condition $f(t) \to 0$ follows from the first one.

Proof. For some $0 \leq \theta(x_1\ t) \leq 1$

$$f(t + x f(t)) = f(t) + x\, f(t) f'(t + x\, \theta(\ x, t)\ f(t)).$$

From $f'(t) \to 0$, $f(t) \to 0$ it follows that $t^{-1}\ f(t) \to 0$ as $t \uparrow x_0$. Hence $1 + x \cdot \theta(x_1 t) \cdot \frac{f(t)}{t}$ remains bounded away from 0 for $t \geq t_0$ and all x in a bounded interval (t_0 does not depend on x). It follows that

$$\lim_{t \uparrow x_0} f'(t + x\, \theta(x, t)\ f(t)) = 0$$

uniformly in any bounded interval. □

Theorem 5 (von Mises). Suppose F has a negative second derivative F" for all x in some interval (x_1, x_0) and let $F'(x)$ vanish for $x \geq x_0$ where x_0 may be finite or infinite. If

$$\lim_{x \uparrow x_0} \frac{F''(x)\,(1 - F(x))}{(F'(x))^2} = -1,$$

then $F \in \mathcal{D}(\Lambda)$.

Proof. Write

$$f(t) = \frac{1 - F(t)}{F'(t)},$$

then f satisfies the assumptions of lemma 3 (If $x_0 < \infty$, $1 - F(t) = \int_t^{x_0} F'(s)\,ds \leq F'(t)\,(x_0 - t)$ because $F'' < 0$; hence $f(t) \to 0$ as $t \uparrow x_0$).

As before we get for $x_1 < x < x_0$

$$1 - F(x) = \{1 - F(x_1)\} \exp - \int_{x_1}^{x} \frac{dt}{f(t)} .$$

Determine b_n such that $1 - F(b_n) = n^{-1}$, then for all x (with $u = (t-b_n)/f(b_n)$)

$$n\,\{1 - F(b_n + x\,f(b_n))\} =$$

$$= \exp - \int_{b_n}^{b_n + x\,f(b_n)} \frac{dt}{f(t)}$$

$$= \exp - \int_0^x \frac{f(b_n)}{f(b_n + u\,f(b_n))}\,du$$

and by the lemma the latter term converges to e^{-x} as $u \to \infty$. Hence

$$F^n(b_n + x\, f(b_n)) = \left(1 - \frac{n\{1 - F(b_n + x\, f(b_n))\}}{n}\right)^n$$

converges to $\Lambda(x)$ for all x. □

Remark. Note that $f(b_n) = \{n.F'(b_n)\}^{-1}$.

Without proof we remark the following about the generality of the given sufficient conditions:

1.[3] If F has a monotone density F' and F is in $\mathcal{D}(\Phi_\alpha)(D(\Psi_\alpha))$ then the conditions of theorem 3 (theorem 4 respectively) are fulfilled. If F has a monotone (increasing) second derivative F" and F is in $\mathcal{D}(\Lambda)$ then the conditions of theorem 5 are fulfilled.

2.[1] If F is in some domain of attraction then there is an F_1 satisfying the conditions of one of the theorems 3, 4 and 5 and in the same domain as F, with

$$\lim_{x \uparrow x_0} \frac{1 - F(x)}{1 - F_1(x)} = 1$$

(hence in particular $F(x) < 1$ if and only if $F_1(x) < 1$).

The following necessary conditions are convenient in deciding whether a given df can possibly belong to the domain of attraction of a (max-)stable df.

1. If $F \in \mathcal{D}(\Phi_\alpha)$ then $F(x) < 1$ for all x and

$$\int_1^\infty t^\rho \, dF(t)$$

is finite for $\rho < \alpha$ and infinite for $\rho > \alpha$.

2. If $F \in D(\Lambda)$ and $F(x) < 1$ for all x then

$$\int_1^\infty t^\rho dF(t)$$

is finite for all ρ.

3. If $F \in \mathcal{D}(\Psi_\alpha)$ then there is an x_0 with the property: $F(x) < 1$ if and only if $x < x_0$. If such an x_0 exists, then $F \in \mathcal{D}(\Psi_\alpha)$ (or $\mathcal{D}(\Lambda)$) if and only if the df $F_1(x) = F(x_0 - 1/x)$ for $x > 0$ is in $\mathcal{D}(\Phi_\alpha)$ (or $\mathcal{D}(\Lambda)$ respectively).

Examples

1) Cauchy's distribution

$$F'(x) = \frac{1}{\Pi} \cdot \frac{1}{1 + x^2} \quad \text{for all } x$$

is in $\mathcal{D}(\Phi_1)$ by the criterion of theorem 3. Here $a_n = n$.

2) Any gamma-distribution

$$F'(x) = (\Gamma(\rho))^{-1} x^{\rho-1} e^{-x}$$

is in $\mathcal{D}(\Lambda)$ by the criterion of theorem 5. Here $b_n = \log n$ and $f(t) = 1$.

3) The normal distribution

$$F'(x) = (2\Pi)^{-\frac{1}{2}} e^{-x^2/2}$$

is in $\mathcal{D}(\Lambda)$ by the criterion of theorem 5. Here $b_n = (2 \log n - \log\log n - \log 4\pi)^{\frac{1}{2}}$ and $f(t) = t^{-1}$.

4) The uniform distribution

$$F'(x) = 1 \quad \text{for all } 0 < x < 1$$

is in $\mathcal{D}(\Psi_1)$ by the criterion of theorem 4. Here $a_n = 1 - n^{-1}$.

5) The distribution

$$F(x) = 1 + e^{1/x} \quad \text{for } x < 0$$

is in $\mathcal{D}(\Lambda)$ by the criterion of theorem 5. Here $b_n = -(\log n)^{-1}$ and $f(t) = t^2$.

6) The tail $1 - F(x)$ of the geometric distribution

$$F(x) = 1 - e^{-[x]}$$

is not asymptotically equivalent to the tail of a continuous df so the geometric df is not in the domain of attraction of any of the stable types.

Acknowledgement. Ideas of A.A. Balkema have been used at several places in this note.

References

1 Balkema, A.A. and L. de Haan (1972). On R. von Mises' condition for the domain of attraction of exp $(-e^{-x})$. Ann. Math. Statist. 43, 1352 - 1354.

2 Gnedenko, B.V. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. Ann. Math. 44, 423 - 453.

3 de Haan, L. (1970). On regular variation and its application to the weak convergence of sample extremes. Mathematisch Centrum, Amsterdam.

4 von Mises, R. (1936). La distribution de la plus grande de n valeurs. In: Selected Papers II (Am. Math. Soc.), 271 - 294.