On accompanying measures and asymptotic expansions in the limit theorem for maximums of random variables

Vladimir I. Piterbarg

Lomonosov Moscow state university

December, 2017, Moscow

Gnedenko Theorem

Let $X_1, ..., X_n, ...,$ be independent identically distributed random variables having distribution function $F(x)$. Denote $M_n := \max(X_i, i = 1, ..., n)$.

Theorem

(R. Fisher, L. H. C. Tippet(1928), B. V. Gnedenko(1943))
Let there exist sequences $a_n > 0, b_n$ and non-trivial distribution function $H(x)$ such, that $P(a_n(M_n - b_n) \le x) \to H(x)$ as $n \to \infty$. Then one can find such $a > 0, b, \gamma$, that $H(ax + b) = H_\gamma(x)$, with
1) (Gumbel) $H_0(x) = \Lambda(x) = \exp(-e^{-x}), x \in \mathbb{R}$;
2) (Frèchet) $H_\gamma(x) = \exp(-x^{-1/\gamma})I(x > 0)$, with $\gamma > 0$;
3) (Weibull) $H_\gamma(x) = \exp(-(-x)^{-1/\gamma})I(x \le 0) + I(x > 0)$, with $\gamma < 0$.

Gnedenko Theorem

Let $X_1, ..., X_n, ...,$ be independent identically distributed random variables having distribution function $F(x)$. Denote $M_n := \max(X_i, i = 1, ..., n)$.

Theorem

(R. Fisher, L. H. C. Tippet(1928), B. V. Gnedenko(1943))
Let there exist sequences $a_n > 0, b_n$ and non-trivial distribution function $H(x)$ such, that $P(a_n(M_n - b_n) \leq x) \to H(x)$ as $n \to \infty$.
Then one can find such $a > 0, b, \gamma$, that $H(ax + b) = H_\gamma(x)$, with
1) (Gumbel) $H_0(x) = \Lambda(x) = \exp(-e^{-x}), x \in \mathbb{R};$
2) (Frèchet) $H_\gamma(x) = \exp(-x^{-1/\gamma})I(x > 0)$, with $\gamma > 0$;
3) (Weibull) $H_\gamma(x) = \exp(-(-x)^{-1/\gamma})I(x \leq 0) + I(x > 0)$, with $\gamma < 0$.

We say that d.f. F belongs to MDA(Λ) if 1) fulfills. (resp. for other two)

Gumbel MDA: von Mises representation

It is known that F belongs to Frèchet MDA iff $1 - F(x)$ is regularly varying at infinity with negative degree. Respectively for Weibull MDA, but with positive degree. It means in particular that the right end point is finite.

Gumbel MDA: von Mises representation

It is known that F belongs to Frèchet MDA iff $1 - F(x)$ is regularly varying at infinity with negative degree. Respectively for Weibull MDA, but with positive degree. It means in particular that the right end point is finite.
We take here $MDA(\Lambda)$ because of it is extremely wide comparing with other two, with quite different behavior of distribution tails, such us Weibull like, log-Weibull like and even slower.

Gumbel MDA: von Mises representation

It is known that F belongs to Frèchet MDA iff $1 - F(x)$ is regularly varying at infinity with negative degree. Respectively for Weibull MDA, but with positive degree. It means in particular that the right end point is finite.
We take here $MDA(\Lambda)$ because of it is extremely wide comparing with other two, with quite different behavior of distribution tails, such us Weibull like, log-Weibull like and even slower.
It is known that $F \in MDA(\Lambda)$ iff for some $x_0 \geq 0$,

$$1 - F(x) = c(x) \exp\left\{-\int_{x_0}^{x} \frac{1}{f(t)} dt\right\}, \quad x \geq x_0, \tag{1}$$

with $f(x)$, positive and absolutely continuous on $[x_0, \infty)$ with density $f'(x)$, $f'(x) \to 0$, and $c(x) \to c > 0$ as $x \to \infty$.
We restrict ourselves with infinite right end point of F.

Gumbel MDA: von Mises representation

Often more flexible form of this assertion is convenient: $F \in MDA(\Lambda)$ if and only if for some $x_0 \geq 0$,

$$1 - F(x) = c(x) \exp\left\{ -\int_{x_0}^{x} \frac{g(t)}{f(t)} dt \right\}, \tag{2}$$

with the same properties of $f(x)$ and $c(x)$, and $g(x) \to 1$ as $x \to \infty$.

Gumbel MDA: von Mises representation

Often more flexible form of this assertion is convenient:
$F \in MDA(\Lambda)$ if and only if for some $x_0 \geq 0$,

$$1 - F(x) = c(x) \exp\left\{ -\int_{x_0}^{x} \frac{g(t)}{f(t)} dt \right\}, \tag{2}$$

with the same properties of $f(x)$ and $c(x)$, and $g(x) \to 1$ as $x \to \infty$.
For both the representations one may take

$$b_n = F^{\leftarrow}(1 - n^{-1}), \quad a_n = f(b_n). \tag{3}$$

There is a wide bibliography on the quality of convergence in Gnedenko limit theorem. We notice here two main directions of studies.

First one is related with restrictions on the tail behavior of F at infinity. First of all it is higher orders regular variation of the tail of F, see

S. Resnick, L. de Haan (1996). Ann. Probab., 24, 1, 97-124.

L. de Haan, Ana Ferreira (2006). Extreme Value Theory. An Introduction. Springer, with many references therein.

There is a wide bibliography on the quality of convergence in Gnedenko limit theorem. We notice here two main directions of studies.

First one is related with restrictions on the tail behavior of F at infinity. First of all it is higher orders regular variation of the tail of F, see

S. Resnick, L. de Haan (1996). Ann. Probab., 24, 1, 97-124.

L. de Haan, Ana Ferreira (2006). Extreme Value Theory. An Introduction. Springer, with many references therein.

The second direction related to concrete expressions of the distributions or family of distributions, such as Gaussian, Gaussian like, Weibull, Weibull like, so on. Here is also a wide bibliography, beginning with

P. Hall(1979) On the Rate of Convergence of Normal Extremes. Journal of Applied Probability 16(2) 433-439

also see the same monograph by de Haan and Ferrira, and

S. Resnick (1987). Extreme values, regular variation, and point processes. Springer-Verlag, New York Berlin Heidelberg,

with many references therein.

Our study belongs rather to this second direction, we use von Mises structure given above, but we suggest another approach: do not investigate immediately
quality of Gumbel approximation
but first
look for better approximations.
Notice that this is very common approach in the study of quality approximation in Central Limit Theorem, with Edgeworth-Cramer approximation and other types of accompanying laws or measures.

Example: Gaussian stationary process

Theorem(2002)

Let $X(t)$, $t \in \mathbb{R}$, be a twice differentiable in square mean Gaussian stationary process with $EX(t) = 0$, $EX^2(t) = 1$, $EX'(t)^2 = 1$. Assume that for its covariance function r and some $a > 0$, $\int_0^T |r(t)|^a dt < \infty$.

Example: Gaussian stationary process

Theorem(2002)

Let $X(t)$, $t \in \mathbb{R}$, be a twice differentiable in square mean Gaussian stationary process with $EX(t) = 0$, $EX^2(t) = 1$, $EX'(t)^2 = 1$. Assume that for its covariance function r and some $a > 0$, $\int_0^T |r(t)|^a dt < \infty$.
Denote $l_T = \sqrt{2 \ln \frac{T}{2\pi}}$ and

$$A_T(x) = \begin{cases} e^{-e^{-x - x^2/2l_T^2}}, & x \geq -l_T^{3/2}, \\ 0, & x < -l_T^{3/2}, \end{cases}$$

$T > 0$.

Example: Gaussian stationary process

Theorem (continued)

Then

1) For some $\gamma > 0$,

$$P\left(\max_{t\in[0,T]} X(t) \le l_T + \frac{x}{l_T}\right) - A_T(x) = O(T^{-\gamma}), \quad T \to \infty$$

uniformly in $x \in \mathbb{R}$.

Example: Gaussian stationary process

Theorem (continued)

Then

1) For some $\gamma > 0$,

$$P\left(\max_{t\in[0,T]} X(t) \leq l_T + \frac{x}{l_T}\right) - A_T(x) = O(T^{-\gamma}), \quad T \to \infty$$

uniformly in $x \in \mathbb{R}$.

2) Also,

$$l_T^2\left(P\left(\max_{t\in[0,T]} X(t) \leq l_T + \frac{x}{l_T}\right) - e^{-e^{-x}}\right) \to \frac{1}{2}e^{-e^{-x}}e^{-x}x^2,$$

as $T \to \infty$, uniformly in $x \in \mathbb{R}$.

Example: Gaussian stationary process

What we see:
1) From the second statement we see that the rate of convergence of the distribution of the maximum to the Gumbel distribution is logarithmic.

Example: Gaussian stationary process

What we see:

1) From the second statement we see that the rate of convergence of the distribution of the maximum to the Gumbel distribution is logarithmic.

It also gives the second term of the asymptotic expansion for the probability.

Example: Gaussian stationary process

What we see:

1) From the second statement we see that the rate of convergence of the distribution of the maximum to the Gumbel distribution is logarithmic.

It also gives the second term of the asymptotic expansion for the probability.

2) The first statement gives the sequence of approximating functions that approaches the Gumbel distribution with the power rate.

Example: Gaussian stationary process

What we see:

1) From the second statement we see that the rate of convergence of the distribution of the maximum to the Gumbel distribution is logarithmic.

It also gives the second term of the asymptotic expansion for the probability.

2) The first statement gives the sequence of approximating functions that approaches the Gumbel distribution with the power rate.

To my best knowledge, there are no similar results in extreme value theory with accompanying laws or charges (signed measures).

Asymptotic expansion and accompanying charges

Taking logarithm in von Mises representation with $g = 1$ and using $b_n = F^{\leftarrow}(1 - n^{-1})$,

$$\int_{x_0}^{b_n} \frac{1}{f(t)} dt = \log(nc(b_n)).$$

Asymptotic expansion and accompanying charges

Taking logarithm in von Mises representation with $g = 1$ and using $b_n = F^{\leftarrow}(1 - n^{-1})$,

$$\int_{x_0}^{b_n} \frac{1}{f(t)} dt = \log(nc(b_n)).$$

Further,

$$F^n(a_n x + b_n) = \left(1 - c(a_n x + b_n) e^{-\int_{x_0}^{a_n x + b_n} \frac{1}{f(t)} dt}\right)^n.$$

Asymptotic expansion and accompanying charges

Taking logarithm in von Mises representation with $g = 1$ and using $b_n = F^{\leftarrow}(1 - n^{-1})$,

$$\int_{x_0}^{b_n} \frac{1}{f(t)} dt = \log(nc(b_n)).$$

Further,

$$F^n(a_n x + b_n) = \left(1 - c(a_n x + b_n) e^{-\int_{x_0}^{a_n x + b_n} \frac{1}{f(t)} dt}\right)^n.$$

From this two expressions, using that $a_n = f(b_n)$ and denoting

$$\gamma_n(x) = \int_{b_n}^{a_n x + b_n} \frac{1}{f(t)} dt - \log \frac{c(a_n x + b_n)}{c(b_n)},$$

after some evaluations we get the following.

Asymptotic expansion and accompanying charges

Proposition

Let $X_1, X_2, ...,$ be i.i.d. random variables with distribution function F, and $M_n := \max(X_1, , ...X_n)$. Assume von Mises representation. Then for a_n, b_n defined above and any x,

$$P(M_n \leq a_n x + b_n) = \exp\left(-e^{-\gamma_n(x)}\right) \exp\left(-\frac{1}{n} \sum_{k=0}^{\infty} \frac{(-1)^k e^{-(k+1)\gamma_n(x)}}{(k+2)n^k}\right).$$

Moreover, for any x, $\gamma_n(x) \to x$ as $n \to \infty$.

Accompanying charge.

Notice that using flexible von Mises representation, with general g, one can also write a similar expansion, with modified $\gamma_n(x)$.

Accompanying charge.

Notice that using flexible von Mises representation, with general g, one can also write a similar expansion, with modified $\gamma_n(x)$.
Other expression for $\gamma_n(x)$;

$$\gamma_n(x) = -\log \frac{1 - F(b_n + a_n x)}{1 - F(b_n)} = \log \frac{1}{n(1 - F(b_n + a_n x)}$$

.

Accompanying charge.

Notice that using flexible von Mises representation, with general g, one can also write a similar expansion, with modified $\gamma_n(x)$.
Other expression for $\gamma_n(x)$;

$$\gamma_n(x) = -\log \frac{1 - F(b_n + a_n x)}{1 - F(b_n)} = \log \frac{1}{n(1 - F(b_n + a_n x)}$$

.
In case F has density one can get von Mises representation with constant $c(x)$, $c(x) \equiv c > 0$, and some other f. In this case,

$$\gamma_n(x) = \int_0^x \left(\frac{f(b_n)}{f(b_n + v)} - 1 \right) dv + x.$$

Accompanying charge.

Notice that using flexible von Mises representation, with general g, one can also write a similar expansion, with modified $\gamma_n(x)$.
Other expression for $\gamma_n(x)$;

$$\gamma_n(x) = -\log \frac{1 - F(b_n + a_n x)}{1 - F(b_n)} = \log \frac{1}{n(1 - F(b_n + a_n x)}$$

.
In case F has density one can get von Mises representation with constant $c(x)$, $c(x) \equiv c > 0$, and some other f. In this case,

$$\gamma_n(x) = \int_0^x \left(\frac{f(b_n)}{f(b_n + v)} - 1 \right) dv + x.$$

Example: For exponential law $f(x) \equiv 1$, hence $\gamma_n(x) = x$ and the rate of convergence is $O(n^{-1})$.

Accompanying charge.

Remark: Looking at the expression

$$\gamma_n(x) = \int_{b_n}^{a_n x + b_n} \frac{1}{f(t)} dt - \log \frac{c(a_n x + b_n)}{c(b_n)},$$

one see that both the summands may play dominating role in the rate of convergence.

Accompanying charge.

Remark: Looking at the expression

$$\gamma_n(x) = \int_{b_n}^{a_n x + b_n} \frac{1}{f(t)} dt - \log \frac{c(a_n x + b_n)}{c(b_n)},$$

one see that both the summands may play dominating role in the rate of convergence.

Example continued: Consider exponential-like distribution $F(x) = 1 - c(x)e^{-x}$, $x \geq 0$, $c(1) = 1$, and $c(x) \to 1/2$ as $x \to \infty$. Taking $c(x) = 1/2 - 1/\log x$, we get the following rate of convergence.

$$\sim \frac{\log(x+1)}{(\log\log n)^2}, \;\; n \to \infty,$$

Second order condition

Consider now the relation between rates of the convergences in terms of γ_n and in terms of the following second order condition.

Second order condition

Consider now the relation between rates of the convergences in terms of γ_n and in terms of the following second order condition. Denote

$$U(t) := \left(\frac{1}{1 - F(t)} \right)^{\leftarrow}; \quad a(t) := f\left(U(t)\right),$$

then since $b_n = F^{\leftarrow}(1 - 1/n)$, $b_n = U(n)$, one can take $a_n = a(n) = f(b_n)$.

Second order condition

Consider now the relation between rates of the convergences in terms of γ_n and in terms of the following second order condition. Denote

$$U(t) := \left(\frac{1}{1-F(t)}\right)^{\leftarrow}; \quad a(t) := f\left(U(t)\right),$$

then since $b_n = F^{\leftarrow}(1-1/n)$, $b_n = U(n)$, one can take $a_n = a(n) = f(b_n)$.

Second order condition:

a) $\exists A(t)$, $A(t) \to 0$, $t \to \infty$, $\mathrm{sign} A(t) = \mathrm{const}$

$$\lim_{t\to\infty} \frac{\frac{U(tx)-U(t)}{a(t)} - \log x}{A(t)} = H(x).$$

$H(x) \neq 0$, $H(x) \neq \infty$ (identically)

Second order condition

Facts (for MDA(Λ):

1. $A(t)$ is regularly varying with non-positive index $\rho \leq 0$.

2.

$$H(x) = \frac{1}{\rho}\left(\frac{x^\rho - 1}{\rho} - \log x\right) \text{ for } \rho < 0,$$

and

$$H(x) = \frac{1}{2}\log^2 x \text{ for } \rho = 0.$$

Second order condition

Facts (for MDA(Λ):
1. $A(t)$ is regularly varying with non-positive index $\rho \leq 0$.
2.
$$H(x) = \frac{1}{\rho}\left(\frac{x^\rho - 1}{\rho} - \log x\right) \text{ for } \rho < 0,$$
and
$$H(x) = \frac{1}{2}\log^2 x \text{ for } \rho = 0.$$
Using S. Resnick and L. de Haan results, from Proposition it follows:
$$P(M_n \leq a_n x + b_n) = e^{-e^{-x} - A(n)H(x)(1+o(1))}(1 + R(x,n)),$$
in particular, $R(x,n) = O(n^{-1})$.

Second order condition

Facts (for MDA(Λ):

1. $A(t)$ is regularly varying with non-positive index $\rho \leq 0$.

2.

$$H(x) = \frac{1}{\rho}\left(\frac{x^\rho - 1}{\rho} - \log x\right) \text{ for } \rho < 0,$$

and

$$H(x) = \frac{1}{2}\log^2 x \text{ for } \rho = 0.$$

Using S. Resnick and L. de Haan results, from Proposition it follows:

$$P(M_n \leq a_n x + b_n) = e^{-e^{-x} - A(n)H(x)(1+o(1))}(1 + R(x,n)),$$

in particular, $R(x,n) = O(n^{-1})$.

Notice different cases: $\rho < -1$; $\rho = -1$; $\rho > -1$.

Example: Weibull tails.

Consider first a distribution with $f(t) = (Cp)^{-1} t^{1-p}$ for $t \geq x_0$, and $c = e^{C x_0^p}$, $p > 0$. That is,

$$1 - F(x) = ce^{-Cx^p}, \; x \geq x_0,$$

Example: Weibull tails.

Consider first a distribution with $f(t) = (Cp)^{-1} t^{1-p}$ for $t \geq x_0$, and $c = e^{Cx_0^p}$, $p > 0$. That is,

$$1 - F(x) = ce^{-Cx^p}, \; x \geq x_0,$$

In this case

$$\gamma_n(x) - x = \log n \left(\left(1 + \frac{x}{p \log n} \right)^p - 1 \right) - x$$

Example: Weibull tails.

Consider first a distribution with $f(t) = (Cp)^{-1} t^{1-p}$ for $t \geq x_0$, and $c = e^{Cx_0^p}$, $p > 0$. That is,

$$1 - F(x) = ce^{-Cx^p}, \; x \geq x_0,$$

In this case

$$\gamma_n(x) - x = \log n \left(\left(1 + \frac{x}{p \log n} \right)^p - 1 \right) - x$$

For $p = 1$ we get power expansion; for $p \neq 1$ we have logarithmic expansion.

Example: Weibull tails.

Consider first a distribution with $f(t) = (Cp)^{-1} t^{1-p}$ for $t \geq x_0$, and $c = e^{Cx_0^p}$, $p > 0$. That is,

$$1 - F(x) = ce^{-Cx^p}, \; x \geq x_0,$$

In this case

$$\gamma_n(x) - x = \log n \left(\left(1 + \frac{x}{p \log n} \right)^p - 1 \right) - x$$

For $p = 1$ we get power expansion; for $p \neq 1$ we have logarithmic expansion.
In particular, for $p = 2$ we have, $\gamma_n(x) - x = \frac{x^2}{4 \log n}$, compare with $x^2/2l_T^2$ in our first example of Gaussian stationary processes.

Example: Log-Weibull tails.

Now take $f(t) = (Cp)^{-1} t \log^{-p+1} t$, $p > 1$, we have,

$$1 - F(x) = ce^{-C \log^p x}, \; x \geq x_0,$$

and get similarly,

$$\gamma_n(x) - x \sim \frac{x^2 \log^{1-1/p} n}{2C^{1/p} p} \exp\left(-C^{-1/p} \log^{1/p} n\right),$$

$n \to \infty$.
That is, the rate of convergence for log-Weibull distribution is generally better that for Weibull one.

A scale

In order to consider the distributions with heavier tails in frames $MDA(\Lambda)$ we may continue a scale, as

$$1 - F(x) = ce^{-C\log x(\log\log x)^a}, \ a > 1, \quad x \geq x_0,$$

$$1 - F(x) = ce^{-C\log x \log\log x(\log\log\log x)^a}, \ a > 1, \quad x \geq x_0,$$

so on.

A scale

In order to consider the distributions with heavier tails in frames $MDA(\Lambda)$ we may continue a scale, as

$$1 - F(x) = ce^{-C \log x (\log\log x)^a}, \; a > 1, \quad x \geq x_0,$$

$$1 - F(x) = ce^{-C \log x \log\log x (\log\log\log x)^a}, \; a > 1, \quad x \geq x_0,$$

so on.
We have for the first, by differentiating,

$$\log x (\log\log x)^a = \int_{x_0}^{x} \frac{1 + \frac{a}{\log\log t}}{t(\log\log x)^{-a}} dt,$$

so that indeed, $F \in MDA(\Lambda)$.
Thus we have some scale for MDA(Gumbel).

A scale

Another definition of the Gumbel index/scale may be as following. Denote

$$\log_k = \overbrace{\log \ldots \log}^{k}$$

and consider the integral

$$G_k = \int_{x_0}^{\infty} \frac{f(t)dt}{t^2 \log t \log_2 t \ldots \log_k t}.$$

The first $k = 0, 1, \ldots$ such that $G_k < \infty$ is called the Gumbel index.

Some bibliography

1. Cheng, S., Jiang, C. (2001). The Edgeworth expansion for distributions of extreme values. *Science in China (SeriesA), 44, 427–437*

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

3. Bingham, N.H., Goldie C. M. and Teugels J. L. (1987). *Regular variation.* Cambridge university press. Cambridge.

4. Davis Richard A.(1982) The rate of convergence in distribution of the maxima. *Statistica Neerlandica 36 , nr. 1, 31-35.*

5. H. Drees, L. de Haan, and D. Li (2003). On large deviations for extremes. *Stat. Prob. Letters 64,51-62.*

6. H. Drees, L. de Haan, and D. Li (2006). Approximations to the tail empirical distribution function with application to testing extreme value conditions. *Statist. Plann.Inference, 136, 10, 3498-3538.*

7. Laurens de Haan, Ana Ferreira (2006). *Extreme Value Theory. An Introduction.* Springer.

8. Fuming Lin , Xinhua Zhang, Zuoxiang Peng & Yingying Jiang (2011). On the Rate of Convergence of STSD Extremes. *Communications in Statistics, Theory and Methods Volume 40, Issue 10, 1795-1806*
9.Hall Peter (1979) On the Rate of Convergence of Normal Extremes. Journal of Applied Probability 16(2) 433-439
10. W. J. Hall, Jon A. Wellner (1979) The rate of convergence in law of the maximum of an exponential sample. *Statistica Neerlandica, 33, 3, 151–154.*
11. Chuandi Liu and Bao Liu Convergence rate of extremes from Maxwell sample. *Journal of Inequalities and Applications 2013, 2013:477*
12. Peng, Z., Nadarajah, S., & Lin, F. (2010). Convergence Rate of Extremes for the General Error Distribution. *Journal of Applied Probability, 47(3), 668-679.*
13. Piterbarg V. I. (2015) *Twenty Lectures About Gaussian Processes.* Atlantic Financial Press London, NewYork, 2015.

14. Resnick Sidney I. (1987). *Extreme values, regular variation, and point processes.* Springer-Verlag, New York Berlin Heidelberg.
15. Sidney Resnick and Laurens de Haan (1996). Second-order regular variation and rates of convergence in extreme-value theory. *Ann. Probab., 24, 1, 97-124.*
16. Xiao Qian Wang, Shi Hong Cheng (2005, 2006), General Regular Variation of the n-th Order and 2nd Order Edgeworth Expansions of the Extreme Value Distribution (I, II) *Acta Mathematica Sinica, English Series Jan., 21(5), 1121–1130, 22 (1), 27–40.*

Many thanks for your attention!