Sums and extremes

8th Conference in Actuarial Science and Finance, Samos, May 29- June 1,2014

Michel Broniatowski

Université Paris 6, France

June 1, 2014

Contents

Direct methods

1. Fixed sample size asymptotics, robustness and loyalty
2. Asymptotics
    1. Light tails
        - all summands are large when the mean is large
        - Erdös-Rényi laws
        - High order statistics and high level aggregates
    2. Heavy tails, conditional independence of the maximum w.r.t. the summands

Integral methods

1. Heavy tails
    - A short proof of a theorem of Gnedenko, Karamata Tauberian Theorem
2. Light tails
    - An Abelian result for the moment generating function
    - Gibbs conditional result under extreme deviation

FIXED SAMPLE SIZE ASYMPTOTICS

X_1, X_2 independent standard Cauchy. Where is (X_1, X_2) when $S_1^2 := X_1 + X_2 > 2t$ for large t?

$$\begin{aligned} & P\left(X_1 > t\epsilon, X_2 > t\epsilon \middle| S_1^2 > 2t\right) \\ & \leq \frac{P(X_1 > t\epsilon)P(X_2 > t\epsilon)}{P\left(\epsilon S_1^2 > 2t\right)} \\ & \propto \frac{t}{(t\epsilon)^2} \to 0 \end{aligned}$$

for all $\epsilon > 0$ as $t \to \infty$.
Only one of the X_i's is large. $(1,0) \in span\,(e_1)\,,\ \ (0,1) \in span\,(e_2)\,.$

X_1, X_2 independent standard $N(0,1)$. Set

$$U := \frac{X_1 + X_2}{2}$$
$$V := \frac{X_1 - X_2}{2}$$

independent gaussian centered.

$$\mathcal{L}\left(\left(\frac{X_1}{t}, \frac{X_2}{t}\right) \middle| S_1^2 > 2t\right) = \mathcal{L}\left(\left(\frac{U}{t}, \frac{U}{t}\right) + \left(\frac{V}{t}, \frac{-V}{t}\right) \middle| S_1^2 > 2t\right)$$

Now $U \perp V$ hence

$$\left(\frac{V}{t}, \frac{-V}{t}\right) \to (0,0)$$

U Gaussian hence $\left(\frac{U}{t}, \frac{U}{t}\right) \to (1,1)$ conditionally on $(U > t)$.

Hence

$$\mathcal{L}\left(\left(\frac{X_1}{t}, \frac{X_2}{t}\right) \middle| S_1^2 > 2t\right) \to \delta_{(1,1)}$$

as $t \to \infty$. $\quad (1,1) \in span\,(e_1, e_2)$

X_1, X_2 independent standard exponential.

$$\mathcal{L}\left((X_1, X_2) | \, S_1^2 = s\right) \text{ uniform on } \left(x_1 + x_2 = s\right), x_{1,2} > 0$$

Hence $P\left(\text{at least one of the } X_i/s \text{ is less than } \epsilon | \, S_1^2 = s\right)$ is twice the normalized distance of the segment from the boundary to $(\epsilon, s - \epsilon)$, hence small for small ϵ (indeed $2\epsilon/s$).
Integrate to get that

$$\mathcal{L}\left(\text{at least one of the } X_i/t \text{ is less than } \epsilon | \, S_1^2 > 2t\right)$$

is small for small ϵ. Although the limit distribution $\mathcal{L}\left((X_1/t, X_2/t) | \, S_1^2 > 2t\right)$ is not a point mass, all terms ars large when the mean is large.
The segment $\left[(1,0), (0,1)\right] \in span\left(e_1, e_2\right)$.

Assumptions

$$T_n := T(X_1, .., X_n) \in \mathbb{R}$$

The upper point of the distribution of T_n is unbounded

There exists $t \to a(t)$ such that $\mathcal{L}\left(X_1^n / a(t) \middle| T_n > t\right) \Longrightarrow \Pi$

and Π not concentrated at 0.

$e_1, .., e_n$ the canonical basis of $\mathbb{R}^n$.

Π is concentrated on $\{x \in \mathbb{R}^n : x_{i_1} = 0, .., x_{i_k} = 0\}$ iff $X_1^n / a(t)$ converges to a r.v Y with $Y_{i_1} = 0, .., Y_{i_k} = 0$, i.e. $\operatorname{supp}\Pi \subset span\left\{e_j, j \notin e_{i_1}, .., e_{i_k}\right\}$.
Consider

$$\inf\left\{n - k : \operatorname{supp}\Pi \subset \bigcup span\left(e_{i_1}, .., e_{i_{n-k}}\right)\right\}$$

where the union is over all $i_1, ..i_{n-k}$ included in $(1, .., n)$.
Then T_n blows when at least $L = n - k$ of the X_i's blow.

Definition

$$BN\left(P_n, T_n\right) := \inf\left\{j : \mathrm{supp}\Pi \subset \bigcup span\left(e_{i_1}, .., e_{i_j}\right)\right\}.$$

The estimator T_n is loyal when $BN\left(P_n, T_n\right)$ s high. (depends on the sampling scheme)

Definition

(Merriam-Webster) LOYAL: Faithful to the lawful government or to the sovereign to whom one is subject: unswerving in allegiance.

An estimator is loyal under a given probability law will take large value only if a vast majority of the sample will force it to.

$$BN\left(Cauchy, S_1^n/n\right) = 1$$
$$BN\left(N(0,1), S_1^n/n\right) = n$$
$$BN\left(Exp(1), S_1^n/n\right) = n$$

Applications

Homogeneous statistics

$$T\left(\lambda X_1, .., \lambda X_n\right) = \lambda^{\alpha} T\left(X_1, .., X_n\right)$$

$\alpha = 1$: $L-$estimators, $\alpha = 2$: variance. More generally U statistics with homogeneous kernel, some M estimators, etc

$$w_\beta(u) = c_\beta \exp\left(-\left|u\right|^{\beta}\right) \quad u \in \mathbb{R}, \beta > 0$$

$$p_\beta\left(u\right) = \left(\beta + 1\right)^{-1} u^{-\beta}, u \in \left(1, \infty\right), \beta > 1$$

Weibull type tails

Theorem

Let T_n be a homogeneous statistics. Assume that

$$(v_1, .., v_n) \to \sum_{i=1}^{n} |v_i|^\beta$$

is minimum at a unique essential point p on the set $\{T_n \geq 1\}$. Then the blowing number of T_n for the i.i.d. sampling is the number of non-zero coordinates of p.

Example

$T_n =$ empirical mean, Blowing number $=n$ when $\beta > 1$,Blowing number $=1$ when $\beta < 1$.

Example

$T_n =$ empirical variance, Blowing number $=n$ when $\beta > 1$, Blowing number $=1$ when $\beta < 1$.

Pareto tails

Theorem

If T_n is a homogeneous with

$$T_n(x_1, .., x_n) \leq \max(x_1, .., x_n)$$
$$T_n(x_1, .., x_n) \leq \theta\,(x + .. + x_n) \text{ for some } \theta > 0$$

then $\mathcal{L}\left(\left(X_1/t, .., X_n/t\right) \middle| T_n > t\right)$ is tight as $t \to \infty$ under i.i.d sampling. Any limiting distribution is concentrated on a union of 1-dimensional canonical subspaces. Hence the blowing number is 1.

See Barbe and B (2004), To be compared with breakdown point; Jureckova and Sen (1996), Kusnier and Mizera (1999), etc.

LIMIT RESULTS UNDER EXCEEDANCES

LIGHT TAILS

$$X_1^n := (X_1, .., X_n)$$

i.i.d.

Assume that for any fixed n it holds

$$\lim_{a \to \infty} P\left(X_1^n \in aB_n \middle| S_1^n \geq na \right) = 0$$

for any Borel set B_n in $\mathbb{R}^n$ such that $(1, .., 1) \notin B_n$.

Fix such a sequence B_n and define a_n such that for any $s > a_n$

$$\sup_{s > a_n} P\left(X_1^n \in sB_n \middle| S_1^n \geq ns \right) \leq 1/n.$$

QUESTION

For which classes of distributions and for which order of magnitude of the conditioning barrier a_n do we have

$$\lim_{n\to\infty} P\left(\cap X_i \in (a_n - \varepsilon_n, a_n + \varepsilon_n) \middle| S_1^n \geq na_n\right) = 1$$

with

$$\lim_{n\to\infty} a_n = \infty$$

and

$$\lim_{n\to\infty} \varepsilon_n / a_n = 0$$

Can we have

$$\lim_{n\to\infty} \varepsilon_n = 0$$

and , if yes, at which rate?

Such questions are also of interest in many branches of physics; see D Sornette (2004) which handles the notion of so-called "democratic localization" of a sample,(fixed n, $a \to \infty$). The definition of the "democratic localization" property is :"the extreme tail behaviour of the sum X of N random variables comes mostly from contributions where the individual variables in the sum are all close to X/N" The precise phrasing amounts to assume that the common density of the independent summands is of the form $\exp(-f(x))$ where f satisfies $f'' > 0$ together with $x^2 f''(x) \to \infty$ as $x \to \infty$, and the density of the sum is of the form $\exp(-f(x/N))$. Applications in fragmentation processes and in turbulence, among others.

NOTATION, HYPOTHESES

Denote

$$C_n := \left(S_1^n / n > a_n \right)$$

and

$$I_n := \cap_{i=1}^n \left(X_i \in \left(a_n - \epsilon_n, a_n + \epsilon_n \right) \right).$$

The real valued random variables $X_1, ..., X_n$ are independent copies of a r.v. X with density f whose support is $\mathbb{R}^+$. We write

$$f(x) := c \exp\left(-d(x)\right)$$

Density functions of the form $f(x) = c \exp(-d(x))$ whose exponent functions d are nearly convex functions.
The i.i.d random variables $X_1, ..., X_n$ have common density f with

$$f(x) = c \exp\Big(-(g(x) + q(x)) \Big)$$

$g(x)$ is a positive convex differentiable function which satisfies

$$\lim_{x\to\infty} g(x)/x = \infty.$$

Let $M(x)$ be some nonnegative continuous function on $\mathbb{R}^+$ for which

$$-M(x) \leq q(x) \leq M(x) \quad \text{for all positive } x$$

together with

$$M(x) = O(\log g(x)) \tag{1}$$

Recall

$$C_n := \left(S_1^n / n > a_n\right)$$

and

$$I_n := \cap_{i=1}^n \left(X_i \in \left(a_n - \epsilon_n, a_n + \epsilon_n\right)\right).$$

Theorem

Let a_n satisfy

$$\liminf_{n\to\infty} \frac{\log g(a_n)}{\log n} > 0$$

then with some control on ϵ_n

$$\lim_{n\to\infty} P(I_n | C_n) = 1.$$

Example

Let $g(x) := x^{\beta}$.

$$\lim_{n\to\infty} \frac{n \log a_n}{a_n^{\beta-2} \epsilon_n^2} = 0.$$

Case 1: $1 < \beta \leq 2$.
Take $a_n = n^{\alpha}$ with $a_n > 1/\alpha$, we need ϵ_n be large enough

$$\epsilon_n >> a_n^{1-\frac{\beta}{2}} \sqrt{n \log a_n}$$

which shows that $\epsilon_n \to \infty$ and $\epsilon_n / a_n \to 0$ possible.
Case 2: $\beta > 2$. $a_n = n^{\alpha}$ with $\alpha > 1/ (\beta - 2)$,arbitrary sequences ϵ_n bounded by below away from 0. The sequence ϵ_n may also tend to 0; indeed $\epsilon_n = O(1/ \log a_n)$ fits.
Let $g(x) := e^x$. Set $a_n := n^{\alpha}$ where $\alpha > 0$ and ϵ_n is rapidly decreasing to 0; indeed we may choose $\epsilon_n = o(\exp(-a_n/4))$ together with $\epsilon_n / \sqrt{n a_n e^{-a_n}} \to \infty$.

Example

Almost Log-concave densities 1: f can be written as

$$f(x) = c(x)\exp\left(-g(x)\right), \qquad 0 < x < \infty$$

with g a twice differentiable convex function with $\lim_{x\to\infty} g(x)/x = \infty$ and for some $x_0 > 0$ and constants $0 < c_1 < c_2 < \infty$,

$$c_1 < c(x) < c_2 \quad \text{for} \quad x_0 < x < \infty.$$

Examples of densities which satisfy the above conditions include the Normal, the hyperbolic density, etc.
Almost Log-concave densities 2: when there exist constants $x_0 > 0$, $\alpha > 0$, and A such that

$$f(x) = Ax^{\alpha-1}l(x)\exp\left(-g(x)\right) \qquad x > x_0$$

where $l(x)$ is slowly varying at infinity.

Erdös-Rényi laws for small increments

Consider a long run $X_1, ..., X_N$ of i.i.d. r.v's and let F be the c.d.f of X_1 which is assumed to satisfy the light tail conditions.

$$J(x) := \sup_t tx - \log E\left(\exp\left(tX_1\right)\right)$$

the Legendre-Fenchel transform of the m.g.f. $t \longrightarrow \log E\left(\exp\left(tX_1\right)\right)$.
Denote $\gamma(u) := J^{\leftarrow}(u) := x$ such that $J(x) = u$ the asymptotic inverse function of J .
We assume that the function γ is asymptotically equivalent to $u \rightarrow \left(-\log\left(1-F\right)\right)^{\leftarrow}(u)$ in the sense that

$$\lim_{x\longrightarrow\infty} \frac{\gamma\left(-\log\left(1-F\right)(x)\right)}{x} = 1.$$

Let $n(N)$, $1 \leq n(N) \leq N$ be an integer sequence and denote

$$M\left(n(N)\right) := \max_{0 \leq j \leq N-n(N)} S_{j+1}^{j+n(N)}$$

the maximum of the sums of the X_i's on blocks of size $n(N)$. Set further (small blocks)

$$c(n(N)) := \frac{\log N}{n(N)} \text{ and } \lim_{N \to \infty} c(n(N)) = \infty.$$

the following result holds

$$\lim_{N \to \infty} \frac{M\left(n(N)\right)}{n(N)\gamma\left(c(n(N))\right)} = 1 \text{ a.s.}$$

Mason (1989). *The function γ behaves like an upper quantile.*

Lemma

The behaviour of the function γ for large values of the argument is closely related to the upper quantile function of X_1

$$\gamma(\log x) = (1 - F)^{\leftarrow}(1/x)\,(1 + o(1))$$

as $x \longrightarrow \infty$.

Assume this presently.

Fix some sequence $n(N)$ define for any positive δ close to 0

$$a_{n(N),\delta} := (1-\delta)\,\gamma\left(c(n(N))\right) =: (1-\delta)\,a_{n(N)}. \tag{2}$$

Hence $a_{n(N)} = (1-F)^{\leftarrow}\left(\exp(-c(n(N))\right)(1+o(1))$ as N tends to infinity.

$$\frac{M\left(n(N)\right)}{n(N)} > a_{n(N),\delta}$$

hold ultimately with probability 1 meaning that that for large N there exists at least one block of consecutive X_i *'s* with length $n(N)$ whose *empirical mean* exceeds $a_{n(N),\delta}$. Therefore all the summands in $X_{j+1},..,X_{j+n(N)}$ satisfy

$$\left|\frac{X_i}{a_{n(N),\delta}} - 1\right| \leq \frac{\epsilon_{n(N),\delta}}{a_{n(N),\delta}} \tag{3}$$

with probability going to 1 as $n \to \infty$. Since δ is arbitrary, the above property holds with $a_{n(N),\delta}$ substituted by $a_{n(N)}$ and $\epsilon_{n(N)}$ defined accordingly.

We may choose the sequences a_n and ϵ_n satisfy

$$\lim_{n\to\infty} \frac{\epsilon_{n(N)}}{a_{n(N)}} = 0.$$

Hence

$$\lim_{n\to\infty} \frac{X_i}{a_{n(N)}} = 1$$

in probability for any i between $j+1$ and $j+n(N)$.

Aggregate forming in the random walk. This phenomenon holds under quite general hypotheses through a discussion on the condition

$$\gamma(\log x) = (1 - F)^{\leftarrow}(1/x)\,(1 + o(1))$$

Globally OK if

$$h(x) := g'(x)$$

is smooth regularly varying.

Height of the aggregate with given length

The case when $-\log(1-F)$ is a regularly varying function
With g a regularly varying function with index $k>1$, As an example consider the case when X_1 has a Weibull distribution on $\mathbb{R}^+$ with scale parameter 1 and shape parameter $k>1$.
Define $n(N)$ and $a_{n(N)}$ through

$$n(N)=(\log N)^{\frac{1}{1+k/\alpha}}$$

and

$$a_{n(N)}:=\gamma\left(\frac{\log N}{n(N)}\right)$$

which entails

$$a_{n(N)}=(\log N)^{\frac{1}{\alpha+k}}(1+o(1))$$

Size of the aggregate with a given height

1-Aggregates with high level: Define for $0 < \gamma \leq 1$

$$a_{n(N)} = (\gamma \log N)^{1/k} .$$

Then $n(N)$ is a constant. The level is of the order of magnitude of the upper quantile of order $N^{-\gamma}$. This result is a limit case since $n(N)$ is bounded; connects with Mason (1989).
2-Aggregates with intermediate level: Define

$$a_{n(N)} = (\log N)^{\frac{1}{\alpha+k}}$$

for positive α. Then

$$n(N) = (\log N)^{\frac{1}{1+k/\alpha}} (1 + o(1)) .$$

3-Aggregates with low level: Define

$$a_{n(N)} = (\gamma \log \log N)^{1/k}.$$

Then

$$n(N) = \frac{\log N}{\gamma \log \log N}(1 + o(1))$$

which are long aggregates.

Law of large numbers for extreme values and properties of aggregates

The above choice for the values of $n(N)$ and of $a_{n(N)}$ is not incidental. Consider the case when $n(N) = 1$; this case is not covered by the the present results; however by Theorem 1 in Mason (1989), Erdös-Rényi law holds, with $a_1 = \gamma\left(c(1)\right)$, asymptotically equivalent to the $1/N$ upper quantile of F. In this case the moving block satisfying (3) shrinks to the maximum $X_{N,N}$ of the X_i 's in the sample $X_1, .., X_N$ for which it is well known that

$$\frac{X_{N,N}}{a_1} \to 1$$

a.s. under the current hypotheses. The extension to the case $n(N) = cst$ holds (Mason (1989) and Barbe and B(2004)) for distributions with Weibull type tails, with a_1 substituted by a_{cst}.

The following Proposition, which extends in a weak sense this classical result of the theory of extreme order statistics to the behaviour of aggregates in the long run holds.

Theorem

Under appropriate regularity conditions then with $a_{n(N)} = \gamma\left(c(n(N))\right)$ and $c(n(N)) \to \infty$

$$\lim_{N\to\infty} \min_{0\le j\le N-n(N)} \left(\max_{j\le i\le j+n(N)-1} \frac{X_i}{a_{n(N)}} \right) = 1 \text{ in probability.} \tag{4}$$

See B and Cao (2014)

LIMIT RESULTS UNDER HEAVY TAILS

Armendariz and Loulakis (2013)

$S_1^n = X_1 + ... + X_n$ Assume the X_i's i.i.d. with common measure μ subexponential

$$F(0) = 0$$

$$\lim_{n\to\infty} \frac{P(X_1 + ... + X_n > x)}{nP(X_1 > x)} = 1.$$

Assume μ be $\Delta-$ subexponential (Asmussen, Foss, Korshunov, 2003),

$$\lim_{n\to\infty} \frac{P(X_1 + ... + X_n \in x + \Delta)}{n\mu(x + \Delta)} = 1$$

and $\Delta := (O, s)$.
Question :
Study

$$\mu_{n,x}^{\Delta}(A) := P\left((X_1, .., X_n) \in A \middle| S_1^n \in x, x + \Delta \right).$$

Define an operator T

$$T(x_1, .., x_n)_k := \begin{array}{l} \max_{1\leq i\leq n} x_i \text{ if } k = n \\ x_n \text{ if } x_k > \max_{1\leq i<k} x_i \text{ and } x_k = \max_{i\geq k} x_i \\ x_k \quad \text{otherwise} \end{array}$$

The operator T exchanges the last and the maximum component of a finite sequence.

Theorem

Suppose G is $\Delta-$ subexponential. Then there exists a sequence q_n such that

$$\lim_{n\to\infty} \sup_{x>q_n} \sup_{A\in\mathcal{B}(\mathbb{R}^{n-1})} \left| \mu_{n,x}^{\Delta} o T^{-1}(A\times\mathbb{R}) - \mu^{n-1}(A) \right| = 0.$$

Conditioning on $(S_1^n \in x+\Delta)$ affects only the maximum in the limit , and the $n-1$ smallest values become asymptotically independent.

Since, with $M_n := \max\left(X_1, .., X_n\right)$

$$\mu_{n,x}^{\Delta}\left(M_n + \sum_{i=1}^{n-1}\left(TX\right)_i \in (x, x+s)\right) = 1$$

when

$$S_{n-1}/\,b_n \text{ converges to a stable law } H$$

then

$$\frac{M_n - x}{b_n} \to -H.$$

A direct consequence, with $\Delta = (0, \infty)$

$$\nu_x\left(A\right) := P\left(X_1 \in A \middle| X_1 > x\right)$$

then

$$\lim_{n\to\infty} \sup_{x>q_n} \left\| \mu_{n,x} o T^{-1} - \left(\mu^{n-1} \times \nu_x\right)\right\|_{t.v.} = 0.$$

An very interesting discussion on the values of s in Δ which yield different limiting behaviours for the conditional measure of the maximum.

INTEGRAL TRANSFORMS, SUMS AND EXTREMES

KARAMATA TAUBERIAN THEOREM, domains of attraction, heavy tails

$X_1, ..X_n$ i.i.d. $F(x) := P(X_1 < x)$, $F(0) = 0$.

$$W_n := \min(X_1, ..X_n)$$

Assume $0 < \alpha < 1$ and F regularly varying at 0 with order α

$$F(x) = x^{\alpha} L(1/x) \text{ as } x \to 0$$

and

$$L \text{ slowly varying at infinity.}$$

Define

$$M(0) = 0$$
$$1 - M(t) = \frac{1}{\Gamma(1-\alpha)} F(1/t), \; t > 0$$

Theorem

(e.g. Feller) $M \in D(G_\alpha)$ *where* G_α *stable law of index* α *on* $\mathbb{R}^+$: *There exists* $(\alpha_n)_n > 0$, $\alpha_n \to 0$

$$M^{n*}(\alpha_n t) \to G_\alpha(t)\ ,\quad t > 0$$

Define

$$1 - G(x) := \int_0^\infty e^{-tx} dM(t)$$

$$\text{Hence} \qquad G(x) = \frac{1}{\Gamma(1-\alpha)} x^\alpha \int_0^\infty e^{-u} u^{-\alpha} L(u/x) du$$

Hence **(Karamata)**

$$G(x) \sim x^\alpha L(1/x) \quad \text{as} \quad x \to 0.$$

Now

$$(1-G)^n\left(\alpha_n^{-1}x\right)=\int_0^\infty e^{-tx}dM^{n*}(\alpha_n t)\quad \text{for all } x>0$$

and

$$\int_0^\infty e^{-tx}dM^{n*}(\alpha_n t)\longrightarrow e^{-x^\alpha}$$

Therefore

$$G\in D\left(\Phi_\alpha\right)\quad \text{for the minimum.}$$

Further

$$G(x)\sim F(x)\text{ as }\quad x\longrightarrow 0.$$

Hence

$$F\in D\left(\Phi_\alpha\right)\quad \text{for the minimum.}$$

For $\alpha\geq 1$, and a part of Gumbel's domain of attraction: the same (change of variables, etc).

All properties for sums translate into properties for extremes

EXAMPLE RATES, rates for stable laws, in the 80'th.

$$D_n := \sup_x \left| P\left(W_n < n^{-1/\alpha} x \right) - \Phi_\alpha(x) \right|$$

$$\text{(H)} \quad \frac{|F(x) - Cx^\alpha|}{x^\alpha} < Kx^\delta \quad \text{for} \quad 0 < x < A$$

$$\text{and} \quad F(x) - Cx^\alpha \text{ monotone on } V(0)$$

Then

α

$$0 < \delta < \alpha \Rightarrow D_n < Kn^{1-1/\alpha}$$

$$0 < \delta < 1 \Rightarrow D_n < K^{1-\delta/\alpha}$$

$$\delta > 1 \Rightarrow D_n < n^{-1-1/\delta}$$

Extensions to maxima, to ≥ 1, to part of $D(\Lambda)$, etc ...

See B (1987)

MOMENT GENERATING FUNCTIONS, Light tails

Gibbs conditioning principles under extreme deviations

An Abelian type Theorem

HYPOTHESES:NEAR LOG-CONCAVITY

Define the density function $p(x)$,

$$p(x) = c \exp\Big(- \left(g(x) - q(x)\right)\Big) \qquad x \in \mathbb{R}_+, \tag{5}$$

where c is some positive constant.
Define $h(x) := g'(x)$, Assume h is smooth regularly varying (nearly Weibull case) or h is smooth rapidly varying ($p(x)$ damps very quickly to 0).
Define the inverse function of h through

$$\psi(u) := h^{\leftarrow}(u) := \inf\{x : h(x) \geq u\}. \tag{6}$$

Juszczak and Nagaev, A. V. (2004), Balkema, Kluppelberg, Resnick (1993).

The i.i.d random variables $X_1, \ldots, X_n$ have common density p with

$$p(x) = c \exp\Big(- (g(x) - q(x)) \Big)$$

$g(x)$ is a positive convex differentiable function which satisfies

$$\lim_{x \to \infty} g(x)/x = \infty.$$

Let $M(x)$ be some nonnegative continuous function on $\mathbb{R}^+$ for which

$$-M(x) \leq q(x) \leq M(x) \quad \text{for all positive } x \text{ with } M(x) = O(\log g(x))$$

Cramer

$$\Phi(t) := E \exp tX < \infty \text{ on } V(0).$$

The r.v. $\mathcal{X}_t$ has *tilted density* defined on $\mathbb{R}$ with parameter t

$$\pi_t(x) := \frac{\exp tx}{\Phi(t)} p(x).$$

The expectation, the three first centered moments of $\mathcal{X}_t$ defined on $\mathcal{N}$

$$m(t) := \frac{d}{dt} \log \Phi(t) \qquad s^2(t) := \frac{d}{dt} m(t) \qquad \mu_j(t) := \frac{d}{dt} s^2(t) \quad , \quad j = 3, 4.$$

Theorem

(Biret, B, Cao, 2014) It holds as $t \to \infty$

$$m(t) \sim \psi(t), \qquad s^2(t) \sim \psi'(t), \qquad \mu_3(t) \sim \psi''(t),$$

$$\mu_j(t) \sim M_j s^j(t) \quad \text{for } j \quad \text{even} > 3$$

$$\mu_j(t) \sim \frac{M_{j+3} - 3jM_{j-1}}{6} \quad \text{for } j \quad \text{odd} > 3$$

where M_j *is the* $j-$*th moment of standard normal density.*

Edgeworth Expansion for some asymptotic array

With t determined by $m(t) = a_n$ define π^{a_n} through

$$\pi^{a_n}(x) = e^{tx} p(x) / \Phi(t).$$

Note

$$m(t) \longrightarrow \text{essup}(\text{support } P) \text{ as } t \longrightarrow \infty.$$

Define $s := s(t)$ and the normalized density of π^{a_n} by

$$\bar{\pi}^{a_n}(x) = s\pi^{a_n}(sx + a_n),$$

and denote by ρ_n the normalized density of n-convolution $\bar{\pi}_n^{a_n}(x)$,

$$\rho_n(x) := \sqrt{n}\bar{\pi}_n^{a_n}(\sqrt{n}x).$$

Theorem

Denote by $\phi(x)$ the standard normal density, uniformly upon x it holds

$$\rho_n(x) = \phi(x)\left(1 + \frac{\mu_3}{6\sqrt{n}s^3}(x^3 - 3x)\right) + o\left(\frac{1}{\sqrt{n}}\right).$$

Local Gibbs conditional result

$$\pi^{a_n}(x) := e^{tx}p(x)/\Phi(t).$$

Theorem

$$p(X_1 = x | S_1^n = na_n) = \pi^{a_n}(x)\left(1 + o(1/\sqrt{n})\right)$$

Extensions: Law of long runs under extreme condition, asymptotic independence, etc. Applications: Rare event simulation, IS

Basic fact: invariance: For all x and all α in the range of X_1

$$p(X_1 = x | S_1^n = na_n) = \pi^{\alpha}(X_1 = x | S_1^n = na_n)$$

where on the LHS, the r.v's X_i 's are sampled i.i.d. under p and on the RHS, sampled i.i.d. under π^{α}.

Bayes, independence

$$p(X_1 = y_1 | S_1^n = na_n - y_1) = \pi^m(X_1 = y_1 | S_1^n = na_n)$$
$$= \pi^a(X_1 = y_1)\frac{\pi^a(S_2^n = na_n - y_1)}{\pi^a(S_1^n = na_n)} = \frac{\sqrt{n}}{\sqrt{n-1}}\pi^a(X_1 = y_1)\frac{\widetilde{\pi_{n-1}}(z_1)}{\widetilde{\pi_n}(0)},$$

where $\widetilde{\pi_{n-1}}$ is the normalized density of S_2^n under i.i.d. sampling under π^a;correspondingly, $\widetilde{\pi_n}$ is the normalized density of S_1^n under the same sampling. A r.v. with density π^a has expectation a_n and variance $s^2(t)$ with $m(t) = a_n$. Here $z_1 := (na_n - y_1) / s(t)\sqrt{n-1}$.

Perform a third-order Edgeworth expansion of $\widetilde{\pi_{n-1}}(z_1)$. It follows

$$\widetilde{\pi_{n-1}}(z_1) = \phi(z_1)\left(1 + \frac{\mu_3}{6s^3\sqrt{n-1}}(z_1^3 - 3z_1)\right) + o\left(\frac{1}{\sqrt{n}}\right),$$

The approximation of $\widetilde{\pi_n}(0)$

$$\widetilde{\pi}(0) = \phi(0)\left(1 + o(\frac{1}{\sqrt{n}})\right).$$

$$\begin{aligned} p(X_1 = y_1 | S_1^n = na_n) &= \left(\frac{\sqrt{n}}{\sqrt{n-1}}\pi^a(X_1 = y_1)\left(1 + o(1/\sqrt{n})\right)\right) \\ &= \left(1 + o(\frac{1}{\sqrt{n}})\right)\pi^{a_n}(X_1 = y_1), \end{aligned}$$

The asymptotic location of x under the conditioned distribution

Let $\mathcal{X}_t$ be a r.v. with density π^{a_n} where $m(t) = a_n$. Recall that $E\mathcal{X}_t = a_n$ and $Var\mathcal{X}_t = s^2$. It holds

$$\log E \exp \lambda \left(\mathcal{X}_t - a_n \right) / s = -\lambda a_n / s + \log \phi \left(t + \frac{\lambda}{s} \right) - \log \phi \left(t \right).$$

$$\log E \exp \lambda \left(\mathcal{X}_t - a_n \right) / s = \frac{\lambda^2}{2} \frac{s^2 \left(t + \frac{\theta \lambda}{s} \right)}{s^2}$$

$t \longrightarrow s(t)$ is self neglecting.
$\left(\mathcal{X}_t - a_n \right) / s$ converge to a standard normal variable $N(0,1)$ in distribution, as $n \rightarrow \infty$. This amounts to say that

$$\mathcal{X}_t = a_n + sN(0,1) + o_{\Pi^{a_n}}(1).$$

Weibull type: $\log p(x) \sim x^\gamma$: $1 < \gamma \leq 2$ the variance of the tilted distribution Π^{a_n} has a non degenerate variance for all a_n (in the standard gaussian case, $\Pi^{a_n} = N\left(a_n, 1\right)$), as does the conditional distribution of X_1 given $\left(S_1^n = na_n\right)$ .If $\gamma > 2$ it concentrates at a_n since $s := s(t) \rightarrow 0$.

Extension to other conditioning events

Let $X_1, .., X_n$ be n i.i.d. r.v's with common density p defined on $\mathbb{R}^d$ and let f denote a measurable function from $\mathbb{R}^d$ onto $\mathbb{R}$ such that $f(X_1)$ has a density p_f . Assume that

$$p_f(x) = \exp - (g(x) - q(x))$$

and we denote accordingly $\phi_f(t)$ its moment generating function. Denote $\Sigma_1^n := f(X_1) + .. + f(X_n)$.

Denote for all a in the range of f

$$\pi_f^a(x) := \frac{\exp tf(x)}{\phi_f(t)} p(x)$$

where t is the unique solution of $m(t) := (d/dt) \log \phi_f(t) = a_n$ and $\Pi_f^{a_n}$ the corresponding probability measure. **Denote P_{f,a_n} the conditional distribution of X_1 given** $(\Sigma_1^n = na_n)$.

Theorem

$$\sup_{C \in \mathcal{B}(\mathbb{R})} \left| P_{f,a_n}(C) - \Pi_f^{a_n}(C) \right| \to 0$$

Differences between Gibbs principle under LDP and under extreme deviation

Consider the application of the above result to r.v's $Y_1, .., Y_n$ with $Y_i := (X_i)^2$ where the X_i's are i.i.d. and are such that Y is light tailed Weibull distribution with parameter larger than 2 . By the Gibbs conditional principle under a point conditioning (Diaconis,Freedman 1981), for *fixed* a, conditionally on $\left(\sum_{i=1}^n Y_i = na\right)$ the generic r.v. Y_1 has a non degenerate limit distribution

$$p_Y^*(y) := \frac{\exp ty}{E \exp tY_1} p_Y(y)$$

and the limit density of X_1 under $\left(\sum_{i=1}^n X_i^2 = na\right)$ is

$$p_X^*(x) := \frac{\exp tx^2}{E \exp tX_1^2} p_X(x)$$

a non degenerate distribution, with $m_Y(t) = a$.
Instead when $a_n \to \infty$ the distribution of X_1 under the condition $\left(\sum_{i=1}^n X_i^2 = na_n\right)$ concentrates sharply at $-\sqrt{a_n}$ and $+\sqrt{a_n}$.

EDP under exceedances

$$p_{A_n}(X_1 = x) := p_{A_n}(X_1 = x \mid S_1^n > a_n)$$

Then for any family of Borel sets B_n such that

$$\liminf_{n \to \infty} P_{A_n}(B_n) > 0$$

it holds

$$P_{A_n}(B_n) = (1 + o(1))\Pi_{a_n}(B_n).$$

as $n \to \infty$. Csiszar (1984) in the LDP domain, etc

Perspectives: Gibbs measures: Erdös-Rényi under dependence for small increments, etc
Long runs under extreme conditions, AR models
Stretched exponential sampling

References

Armendariz,I., Loulakis, M.: Conditional distribution of heavy tailed random variables on large deviations of their sum, St. Proc. Appl. (2012)

Balkema, A.A., Kl ¨uppelberg, C., Resnick, S.I.: Densities with Gaussian tails. Proc. London. Math. Soc.(3) 66(3), 568–588 (1993)

Barbe, Ph., Broniatowski, M Blowing number of a distribution for a statistics and loyal estimators. Statist.Probab. Lett 69(4), 465–475 (2004)

Bingham, N.H., Goldie, C.M., Teugels, J.L.: Regular variation: Cambridge University Press, Cambridge (1989)

Biret, M., Broniatowski, M., Cao, Z.: A sharp Abelian theorem for the Laplace transform. In Limith Theorems and Mathematical Statistics, Festschrift P Deheuvels, to appear (2014)

Broniatowski, M Convergence for the sample extremes via convolutions. Mathematical statistics and probability theory, Vol. A (Bad Tatzmannsdorf, 1986), 43–52, Reidel, Dordrecht, 1987.

Broniatowski, M, Cao, Z.: Light tails; all summands are large when the empirical mean is large. Extremes, 17(2), 305-336 (2014)

Csisz´ar, I.: Sanov property, generalized I -projection and a conditional

Dembo, A., Zeitouni, O.: Refinements of the Gibbs conditioning principle. Probab. Th Rel Fields 104, 1–14 (1996)
Diaconis, P., Freedman, D.A.: Conditional limit theorems for exponential families and finite versions of de Finetti's theorem. J. Th Prob1, 381–410 (1988)
Juszczak, D., Nagaev, A.: Local large deviation theorem for sums of I.I.D. random vectors when the Cram´er condition holds in the whole space. Prob Math. Statist. 24(2), 297–320 (2004)
Mason, D.M.: An extended version of the Erd■os-R´enyi strong law of large numbers. Ann. Probab 17(1),257
Sornette, D.: Critical phenomena in natural sciences, 2d Edition Springer Series in Synergetics (2006)

Thank you