Exponential exact estimation for maximum and minimum
tail of distribution for non-Gaussian random vector.

Formica M.R., Ostrovsky E., and Sirota L.

Università degli Studi di Napoli Parthenope, via Generale Parisi 13, Palazzo Pacanowsky, 80132, Napoli, Italy.

e-mail: mara.formica@uniparthenope.it

Department of Mathematics and Statistics, Bar-Ilan University, 59200, Ramat Gan, Israel.

e-mail: eugostrovsky@list.ru

Department of Mathematics and Statistics, Bar-Ilan University, 59200, Ramat Gan, Israel.

e-mail: sirota3@bezeqint.net

Abstract

We find the exponential exact two-terms non-asymptotic expression for the maximum and minimum distribution of a non-Gaussian, in general case, random vector.

Key words and phrases:

Random vector, distribution of maximum, Lebesgue – Riesz and Grand Lebesgue Spaces and norms, exact exponential tail estimate, Bonferroni’s inequality, Hölder’s inequality, Young – Fenchel transform, moment generation function, subgaussian variables, vectors and correlation; triangle inequality.

1 Statement of problem. Notations. Foreword.

Let \((\Omega = \{\omega\}, \mathcal{B}, P)\) be certain (sufficiently rich) probability space and let \(D = \{1, 2, \ldots, d\}\) be a positive integer numerical set, \(\xi = \xi = \{\xi_1, \xi_2, \ldots, \xi_d\}, d = \ldots\)
2, 3, . . . be a random vector. Define the following important notion of the tail of maximum distribution

\[ Q[S](u) \overset{def}{=} P\left[ \max_{i \in S} \xi_i > u \right], \]  

so that

\[ Q(u) = Q[D](u) \overset{def}{=} P\left[ \max_{i=1,2,...,d} \xi_i > u \right] \]

for the sufficiently great values \( u \), say for \( u \geq 1 \).

In particular,

\[ Q_i(u) := P(\xi_i > u), \quad Q_{i,j}(u) := P(\xi_i > u, \xi_j > u), \quad i, j \in D, \ i \neq j. \]

The case of the minimum distribution

\[ S[\xi](u) \overset{def}{=} P(\min_{i \in D} \xi_i > u), \ u \geq 1. \]

will be considered further.

Our purpose in this report is to deduce the exact exponential tail estimations for these tail probabilities.

We do not suppose the Gaussian distribution of the source vector \( \xi \).

There are a huge numbers of works devoted to this problem, see e.g. [2], [3], [5], [8], [11], [23], [24], [25], [29], [33] etc. The applications of these estimates in the Law of Iterated Logarithm (LIL) and following in statistics are investigated in particular in [16], [33], [34], [35].

Note that as a rule in the mentioned works is considered the case when \( d \rightarrow \infty \).

Preliminary estimations.

We will use the classical Bonferroni’s inequality

\[ \sum Q_i(u) - \sum \sum Q_{i,j}(u) \leq Q[D](u) \leq \sum Q_i(u), \ u \geq 1. \]

where by definition

\[ \sum \overset{def}{=} \sum_{i \in D}, \quad \sum \sum \overset{def}{=} \sum_{i,j \in D; i \neq j}. \]
We present here for beginning some known facts from the theory of one-dimensional random variables with exponential decreasing tails of distributions, see [5], [6], [7], [8], [12], [13], [18], [20], [21], [22], [24], [25], [26].

Let \( \phi = \phi(\lambda), \lambda \in (-\lambda_0, \lambda_0), \) \( \exists \lambda_0 = \text{const} \in (0, \infty) \) be certain even strong convex which takes positive values for positive arguments twice continuous differentiable function, briefly: Young-Orlicz function, such that

\[
\phi(0) = 0, \quad \phi'(0) = 0, \quad \phi''(0) \in (0, \infty).
\]

For instance:

\[
\phi(\lambda) = 0.5 \lambda^2, \lambda \in \mathbb{R}; \quad \text{the so-called subgaussian case.}
\]

We denote the set of all these Young-Orlicz function as \( \Phi : \Phi = \{ \phi(\cdot) \} \). We say by definition that the centered random variable (r.v) \( \xi \) belongs to the space \( B(\phi) \), if there exists certain non-negative constant \( \tau \in [0, \infty) \), such that

\[
\forall \lambda \in (-\lambda_0, \lambda_0) \Rightarrow E \exp(\pm \lambda \xi) \leq \exp(\phi(\lambda \tau)).
\]

(5)

**Definition 2.1.** The minimal non-negative value \( \tau \) satisfying the last relation (5) for all the values \( \lambda \in (-\lambda_0, \lambda_0) \), is named as a \( B(\phi) \) norm of the variable \( \xi \), write \( ||\xi||_{B(\phi)} \equiv \inf \{ \tau, \tau \geq 0, \forall \lambda \in (-\lambda_0, \lambda_0) \Rightarrow E \exp(\pm \lambda \xi) \leq \exp(\phi(\lambda \tau)) \} \). \( (6) \)

For instance: \( \phi(\lambda) = 0.5 \lambda^2, \lambda \in \mathbb{R}; \quad \text{the so-called subgaussian case, write } \xi \in \text{Sub}. \)

We will denote as ordinary the norm in this very popular space by \( || \cdot ||_{\text{Sub}} \):

\[
||\xi||_{\text{Sub}} \equiv ||\xi||_{B(\phi_2)}.
\]

(7)

It is known that if the r.v. \( \xi_i \) are independent and subgaussian, then

\[
|| \sum_{i=1}^{n} \xi_i ||_{\text{Sub}} \leq \sqrt{\sum_{i=1}^{n} ||\xi_i||^2_{\text{Sub}}.}
\]

**Definition 2.2.** The centered r.v. \( \eta \) with finite non-zero variance \( \sigma^2 := \text{Var}(\eta) \in (0, \infty) \) is said to be strictly subgaussian, write: \( \eta \in \text{StSub} \), iff

\[
E \exp(\pm \lambda \eta) \leq \exp(0.5 \sigma^2 \lambda^2), \lambda \in \mathbb{R}.
\]

(7)

For instance, every centered non-zero Gaussian r.v. belongs to the space \( \text{StSub}. \) The Rademacher’s r.v. \( \eta : P(\eta = 1) = P(\eta = -1) = 1/2 \) is also strictly subgaussian. Many other strictly subgaussian r.v. are represented in [5], [24].
It is known that the set \( B(\phi), \phi \in \Phi \) relative the norm (2.3) and ordinary algebraic operations forms a Banach functional rearrangement invariant space, which are equivalent the so - called Grand Lebesgue ones as well as Orlicz exponential spaces. These spaces are very convenient for the investigation of the r.v. having an exponential decreasing tail of distribution, for instance, for investigation of the limit theorem, the exponential bounds of distribution for sums of random variables, non-asymptotical properties, problem of continuous and weak compactness of random fields, study of Central Limit Theorem in the Banach space etc.

Denote by \( \nu(\cdot) \) the Young - Fenchel, or Legendre transform for the function \( \phi : \)

\[
\nu(x) = \nu[\phi](x) \overset{def}{=} \sup_{\lambda:|\lambda| \leq \lambda_0} (\lambda x - \phi(\lambda)) = \phi^*(x).
\] (8)

It is important for us in this preprint to note that if the non - zero r.v. \( \xi \) belongs to the space \( B(\phi) \) then

\[
P(\xi > x) \leq \exp \left( -\nu(x/||\xi||_{B(\phi)}) \right).
\] (9)

The inverse conclusion is also true up to multiplicative constant under suitable conditions.

Further, assume that the centered r.v. \( \xi \) has a finite in some non - trivial neighborhood of origin finite moment generation function

\[
\phi[\xi](\lambda) \overset{def}{=} \max_{\pm} \ln E \exp( \pm \lambda \xi ) < \infty, \; \lambda \in (-\lambda_0, \lambda_0)
\]

for some \( \lambda_0 = \text{const} \in (0, \infty] \). Obviously, the last condition is quite equivalent to the well - known Cramer’s one.

We can and will agree \( \phi[\xi](\lambda) := \infty \) for all the values \( \lambda \) for which

\[
E \exp( |\lambda| \xi ) = \infty.
\]

The introduced in (2.6) function \( \phi[\xi](\lambda) \) is named as natural function for the r.v. \( \xi \); herewith \( \xi \in B(\phi[\xi]) \) and moreover

\[
||\xi||_{B(\phi[\xi])} = 1.
\]

3 Main result. Upper maximum tail estimate.

Let us return to the formulated above problem, \( Q[\xi](u) \) estimation. We suppose that there exists certain function \( \phi \in \Phi \) such that

\[
\forall i \in D \Rightarrow \beta_i := ||\xi_i||_{B(\phi)} \in (0, \infty).
\] (10)
Set also
\[
\beta := \max_{i \in D} \beta_i, \quad r := \text{card}\{ \beta_i : \beta_i = \beta \},
\]
where as ordinary \( \text{card}(M) \) denotes the amount of elements of the set \( M \);
\[
R = R[\beta] \overset{\text{def}}{=} \{ i; \ i \in D, \ \beta_i < \beta \},
\]
\[
\underline{\beta} \overset{\text{def}}{=} \max\{\beta_i, \beta_i < \beta\} = \max\{\beta_i, \ i \in R[\beta]\}.
\]

**Proposition 3.1.** We conclude by means of the Bonferroni’s inequality (4)
\[
Q[\xi](u) \leq r \exp\left( -\nu[\phi](u/\beta) \right) + \sum_{i \in R} \exp\left( -\nu[\phi](u/\beta_i) \right),
\]
and following
\[
Q[\xi](u) \leq r \exp\left( -\nu[\phi](u/\beta) \right) + (d - r) \exp\left( -\nu[\phi](u/\underline{\beta}) \right).
\]

**4 Main result: lower maximum tail estimates.**

We intent to apply here the left-hand side of the relation of Bonferroni (4) and we retain the restrictions (10). Moreover, let us suppose
\[
\exists \delta_i \in (0, \beta_i] \Rightarrow Q_i(u) \geq \exp\left\{ -\nu[\phi](u/\delta_i) \right\}, \ u \geq 1.
\]

Further, we deduce applying the triangle inequality for the \( B(\phi) \) norm
\[
||\xi_i + \xi_j||B\phi \leq \beta_i + \beta_j, \ i \neq j,
\]
that
\[
Q_{i,j}(u) \leq P(\xi_i + \xi_j \geq 2u) \leq \exp\left\{ -\nu[\phi](2u/(\beta_i + \beta_j)) \right\}.
\]

To summarize,

**Proposition 4.1.**
\[
Q[\xi](u) \geq \sum \exp\left\{ -\nu[\phi](u/\delta_i) \right\} - \sum \sum \exp\left\{ -\nu[\phi](2u/(\beta_i + \beta_j)) \right\}, \ u \geq 1.
\]
5 A note about the minimum distribution.

Let us take into account the tail of the minimum distribution for the source random vector (we recall)

$$S[\xi](u) \stackrel{def}{=} \mathbf{P}(\min_{i \in D} \xi_i > u), \ u \geq 1.$$  \hspace{1cm} (18)

We assume that the r.v. $\xi = \vec{\xi}$ is non-negative: $\xi_i \geq 0$ and define the so-called multivariate moment generation function (cf. with a characteristic function)

$$g[\xi](\lambda) = g(\lambda) \stackrel{def}{=} \mathbf{E}\exp(\lambda, \xi),$$  \hspace{1cm} (19)

where $\lambda = \vec{\lambda} \in R^d$, $(\lambda, \xi) = \sum \lambda_i \xi_i$; and as ordinary $(\lambda, \xi) = \sum \lambda_i \xi_i$, $|\lambda| := \sqrt{(\lambda, \lambda)}$.

It will be presumed that this function is finite at last in some neighborhood of origin:

$$\exists \epsilon \in (0, \infty) \ \forall \lambda : |\lambda| < \epsilon \Rightarrow g(\lambda) < \infty.$$  \hspace{1cm} (20)

Of course, if the random vectors $\xi, \eta$ are independent, then $g[\xi + \eta](\lambda) = g[\xi](\lambda) \cdot g[\eta](\lambda)$.

Let now $q_i, i \in D$ be arbitrary numerical vector such that

$$q_i \geq 1, \ \sum_{i} \frac{1}{q_i} = 1.$$  

The Hölder’s inequality give us

$$g[\xi](\lambda) \leq \prod_{i \in D} [\mathbf{E}\exp(\ q_i \ \lambda_i \ \xi_i) ]^{1/q_i} = \prod_{i \in D} \{g[\xi_i](\lambda_i \ q_i) \}^{1/q_i}.$$  

In particular,

$$g[\xi](\lambda) \leq \left\{ \prod_{i \in D} g[\xi_i](\lambda_i \ d) \right\}^{1/d}.$$  

If in addition the r.v. $\{\xi_i\}$ are identical distributed and $\lambda_i = \mu = \text{const} > 0$, then

$$g[\xi](\lambda) \leq g[\xi_1](\mu \ d).$$  

Further, define as usually the so-called Young-Fenchel, or Legendre transform in the multivariate case $x = \vec{x} = \{x_i\} \in R^d$

$$g^*(\vec{x}) = g^*(x) \stackrel{def}{=} \sup_{\lambda \in R^d} [ (\lambda, x) - g(\lambda)].$$  \hspace{1cm} (21)
It is known, see [5], [27] that under formulated restrictions
\[ P(\forall i \in D \Rightarrow \xi_i \geq x_i) = P[ \cap_{i \in D} \{ \xi_i \geq x_i \} ] \leq \]
\[ \exp(-g[\xi]^*(\bar{x})), \ x_i \geq 0. \] (23)

As a consequence:
**Proposition 5.1.**
\[ S[\xi](u) \leq \exp \{ -g^*[\xi](u, u, \ldots, u) \}, \ u \geq 1. \] (24)

### 6 Examples.

**Definition 6.1.** The centered two-dimensional random vector \( \xi = \xi = (\xi_1, \xi_2) \) is said to be subgaussian, if
\[ E \exp(\lambda_1 \xi_1 + \lambda_2 \xi_2) \leq \exp \left\{ 0.5(\sigma_1^2 \lambda_1^2 + 2\rho \sigma_1 \sigma_2 \lambda_1 \lambda_2 + \sigma_2^2 \lambda_2^2) \right\} \] (25)
for arbitrary real numbers \( \lambda_1, \lambda_2 \in R \). Here \( \sigma_1, \sigma_2 = \text{const} > 0, \ \rho = \text{const}, \ |\rho| < 1. \)

For instance, the relation (25) is satisfied for arbitrary non-degenerate, i.e. having strictly positive definite (symmetrical) covariation matrix, centered two-dimensional Gaussian distributed random vector, as well as for the independent subgaussian centered random variables \( (\xi_1, \xi_2) \). The value \( \rho \) may be named as a subgaussian correlation between r.v. \( \xi_1, \xi_2 \).

We conclude by virtue of Proposition 5.1 after simple calculations for the values in particular \( u \geq 1 \)
\[ P(\min(\xi_1, \xi_2) > u) \leq \zeta[\sigma_1, \sigma_2, \rho](u), \] (26)
where
\[ \zeta[\sigma_1, \sigma_2, \rho](u) \defeq \exp \left\{ -\frac{u^2}{2(1-\rho^2)} \frac{\sigma_1^2 + \sigma_2^2 - 2\rho \sigma_1 \sigma_2}{\sigma_1^2 \sigma_2^2} \right\}. \] (27)

Of course,
\[ P(\min(\xi_1, \xi_2) > u) \leq \min\{ \zeta[\sigma_1, \sigma_2, \rho](u), \ \exp \{ -g^*[\xi_1, \xi_2](u, u) \} \}. \] (28)

**Remark 5.1.** The last estimate (28) is exponential asymptotically as \( u \to \infty \) exact for instance for the centered gaussian distributed two-dimensional random vector.
Remark 5.2. This estimate is also asymptotically as \( u \to \infty \) exponential exact for the subgaussian r.v. in both the extremal cases \( \rho = 0 \) and \( \rho = 1 - \rho \). Indeed, let \( \xi_1, \xi_2 \) be independent and standard subgaussian:

\[
\mathbb{E}\xi_{1,2} = 0, \quad ||\xi_1||_{\text{Sub}} = ||\xi_2||_{\text{Sub}} = 1.
\]

The relation (28) give us the following exponential exact estimate

\[
P(\min(\xi_1, \xi_2) > u) \leq \exp\left(-u^2\right), \quad u \geq 1.
\]

Further, let now \( \xi_1 = \xi_2 = \xi \) be a standard subgaussian variable; then

\[
\min(\xi_1, \xi_2) = \xi, \quad \rho = +1.
\]

The relation (28) give us the following exponential exact estimate yet in this limiting case \( \rho = 1 - \rho \)

\[
P(\min(\xi_1, \xi_2) > u) \leq \exp\left(-u^2/2\right), \quad u \geq 1.
\]

Sub - remark 5.2. Notice in addition to the foregoing remark 5.2 that in these conditions, aside from one that \( \rho = +1 \) we assume instead \( \rho = -1 \); i.e. that \( \xi_1 = -\xi_2 \). This case is trivial for us, as long as under these restrictions

\[
P(\min(\xi_1, \xi_2) > u ) = 0, \quad u > 0.
\]

7 Concluding remarks.

It is interest in our opinion to find the exact exponential lower tail estimations for the minimum of random variables.

Acknowledgement. The first author has been partially supported by the Gruppo Nazionale per l’Analisi Matematica, la Probabilità e le loro Applicazioni (GNAMPA) of the Istituto Nazionale di Alta Matematica (INdAM) and by Università degli Studi di Napoli Parthenope through the project “sostegno alla Ricerca individuale”.


References

[1] G. Anatriello and A. Fiorenza. Fully measurable grand Lebesgue spaces. J. Math. Anal. Appl. 422 (2015), no. 2, 783–797.

[2] G. Anatriello and M. R. Formica. Weighted fully measurable grand Lebesgue spaces and the maximal theorem. Ric. Mat. 65 (2016), no. 1, 221–233.

[3] Yu. K. Belyaev and V. I. Piterbarg. Random processes. Sample paths and intersections. Collection of articles, Publishing House ”MIR”, Moscow (1978); 249–257, (in Russian).

[4] M. Sh. Braverman. Bounds on the sums of independent random variables in symmetric spaces. Ukrainian Mathematical Journal, 43 (1991), no. 2, 148–153.

[5] Buldygin V.V., Kozachenko Yu.V. Metric Characterization of Random Variables and Random Processes. 1998, Translations of Mathematics Monograph, AMS, v.188.

[6] V. V. Buldygin, D. I. Mushtary, E. I. Ostrovsky and M. I. Pushalsky. New Trends in Probability Theory and Statistics. Moklas (1992), V.1, 78–92; Amsterdam, Utrecht, New York, Tokyo.

[7] C. Capone, M. R. Formica and R. Giova. Grand Lebesgue spaces with respect to measurable functions. Nonlinear Anal. 85 (2013), no. 2, 125–131.

[8] S. V. Ermakov, and E. I. Ostrovsky. Continuity Conditions, Exponential Estimates, and the Central Limit Theorem for Random Fields. Moscow, VINITY, 1986. (in Russian).

[9] A. Fiorenza. Duality and reflexivity in grand Lebesgue spaces. Collect. Math. 51 (2000), no. 2, 131–148.

[10] A. Fiorenza and G. E. Karadzhov. Grand and small Lebesgue spaces and their analogs, Z. Anal. Anwendungen 23 (2004), no. 4, 657–681.

[11] A. Fiorenza, B. Gupta and P. Jain. The maximal theorem for weighted grand Lebesgue spaces. Studia Math. 188 (2008), no. 2, 123–133.

[12] A. Fiorenza, M. R. Formica and A. Gogatishvili. On grand and small Lebesgue and Sobolev spaces and some applications to PDE’s. Differ. Equ. Appl. 10 (2018), no. 1, 21–46.

[13] A. Fiorenza, M. R. Formica, A. Gogatishvili, T. Kopaliani and J. M. Rakotoson. Characterization of interpolation between grand, small or classical Lebesgue spaces. Nonlinear Analysis, Vol.177, Part B, Dezember 2018, pages 422 - 453.

[14] A. Fiorenza, M. R. Formica and J. M. Rakotoson. Pointwise estimates for $G^{Γ_{c}}$ -functions and applications. Differential Integral Equations, 30, (2017), no. 11 - 12, 809 - 824.

[15] M. R. Formica and R. Giova. Boyd indices in generalized grand Lebesgue spaces and applications. Mediterr. J. Math., 12, (2015), no. 3, 987 - 995.

[16] T. C. Hu. On the law of the iterated logarithm for arrays of random variables. Comm. Statist. Theory Methods 20 (1991), no. 7, 1989–1994.

[17] T. C. Hu and N. C. Weber. On the rate of convergence in the strong law of large numbers for arrays. Bull. Austral. Math. Soc., 45 (1992), no. 3, 479–482.

[18] T. Iwaniec and C. Sbordone. On the integrability of the Jacobian under minimal hypotheses. Arch. Rational Mech. Anal. 119 (1992), no. 2, 129–143.

[19] A. N. Kolmogoroff. Über das Gesetz des iterierten Logarithmus”. (German) Math. Ann., 101 (1929), no. 1, 126–135.

[20] Yu. V. Kozachenko and E. I. Ostrovsky. The Banach Spaces of random variables of sub-Gaussian type. of Probab. and Math. Stat., 32 (1985), (in Russian). Kiev, KSU, 43–57.
[21] Yu.V. Kozachenko, Yu.Yu. Mlavets, and N.V. Yurchenko. Weak convergence of stochastic processes from spaces $F_{\psi}(\Omega)$. STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING, Vol.6, June 2018, pp. 266 - 277.

[22] E. Liflyand, E. Ostrovsky and L. Sirota. Structural properties of bilateral grand Lebesgue spaces. Turkish J. Math. 34 (2010), no. 2, 207–219.

[23] K. W. Ng, Q. H. Tang and H. Yang. Maxima of Sums of Heavy-Tailed Random Variables. ASTIN Bulletin: The Journal of the IAA, 32, (2002), no. 1, pp. 43–55.

[24] E. Ostrovsky. Exponential estimates for random fields and its applications. 1999, OINPE, Moscow - Obninsk.

[25] E. Ostrovsky. Exponential estimate in the Law of Iterated Logarithm in Banach Space. Math. Notes 56 (1994), no. 5-6, 1165–1171.

[26] Ostrovsky E. and Sirota L. Moment Banach spaces: theory and applications. HIAT Journal of Science and Engineering, C, Volume 4, Issues 1 - 2, pp. 233 - 262, (2007).

[27] Ostrovsky E., Sirota L. Prokhorov-Skorokhod continuity of random fields. A natural approach. arXiv:1710.05382v1 [math.PR] 15 Oct 2017

[28] J. Pickands. Maxima of stationary Gaussian processes. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 7 (1967), 190–223.

[29] G. Pisier. Conditions d’entropie assurant la continuité de certains processus et applications à l’analyse harmonique. (French) Seminaire d’analyse fonctionnelle (1980), Exp. No. 13-14, pp. 43 46.

[30] Y. C. Qi. On strong convergence of arrays. Bull. Austral. Math. Soc. 50 (1994), no. 2, 219–223.

[31] S. G. Samko and S. M. Umarkhadzhiev. On Iwaniec-Sbordone spaces on sets which may have infinite measure. Azerb. J. Math. 1 (1) (2011), 67–84.

[32] S. G. Samko and S. M. Umarkhadzhiev. On Iwaniec-Sbordone spaces on sets which may have infinite measure: addendum. Azerb. J. Math 1 (2) (2011), 143–144.

[33] M. S. Sgibnev. On the distribution of the maxima of partial sums. Statist. Probab. Lett. 28 (1996), no. 3, 235–238.

[34] W. F. Stout. Almost sure convergence. Probability and Mathematical Statistics, Vol. 24. Academic Press. New York-London, 1974.

[35] S. H. Sung. An analogue of Kolmogorov’s Law of the Iterated Logarithm for arrays. Bull. Austral. Math. Soc. 54 (1996), no. 2, 177–182.

[36] H. Teicher. Almost certain behavior of row sums of double arrays. Analytical methods in probability theory (Oberwolfach, 1980), pp. 155–165, Lecture Notes in Math., 861, Springer, Berlin-New York, 1981.