Small value probabilities in analysis and mathematical physics

Wenbo V. Li
University of Delaware,
IMA, and
Institute of Applied Mathematics, CAS

Tucson, March, 15, 2012
Arizona School of Analysis and Mathematical Physics

Abstract: Small value probabilities or small deviations study the decay probability that positive random variables behave near zero. In particular, small ball probabilities provide the asymptotic behavior of the probability measure inside a ball as the radius of the ball tends to zero. In this talk, we will provide an overview with connections and recent developments in analysis and mathematical physics, including metric entropy of compact operators, weaker Gaussian correlation inequality, small ball inequalities, symmetrization inequalities in high dimension, and Laplace asymptotics of partition functions.

Small Value Probability

Small value (deviation) probability studies the asymptotic rate of approaching zero for rare events that positive random variables take smaller values. To be more precise, let Y_n be a sequence of *non-negative* random variables and suppose that some or all of the probabilities

$$\mathbb{P}\left(Y_n \leq \varepsilon_n\right), \quad \mathbb{P}\left(Y_n \leq C\right), \quad \mathbb{P}\left(Y_n \leq (1-\delta)\mathbb{E}\, Y_n\right)$$

tend to zero as $n \to \infty$, for $\varepsilon_n \to 0$, some constant $C > 0$ and $0 < \delta \leq 1$. Of course, they are all special cases of $\mathbb{P}\left(Y_n \leq h_n\right) \to 0$ for some function $h_n \geq 0$, but examples and applications given later show the benefits of the separated formulations.

What is often an important and interesting problem is the determination of just how "rare" the event $\{Y_n \leq h_n\}$ is, that is, the study of the *small value (deviation) probabilities* of Y_n associated with the sequence h_n.

If $\varepsilon_n = \varepsilon$ and $Y_n = \|X\|$, the norm of a random element X on a separable Banach space, then we are in the setting of small ball/deviation probabilities.

Deviations: Large vs Small

- Both are estimates of rare events and depend on one's point of view in certain problems.

- Large deviations deal with a class of sets rather than special sets. And results for special sets may not hold in general.

- Similar techniques can be used, such as exponential Chebychev's inequality, change of measure argument, isoperimetric inequalities, concentration of measure, etc.

- Second order behavior of certain large deviation estimates depends on small deviation type estimates.

- General theory for small deviations has been developed for Gaussian processes and measures.

• Some technical difficulties for small deviations: Let X and Y be two positive r.v's (not necessarily ind.). Then

$$\begin{aligned}\mathbb{P}(X+Y>t) &\geq \max(\mathbb{P}(X>t), \mathbb{P}(Y>t)) \\ \mathbb{P}(X+Y>t) &\leq \mathbb{P}(X>\delta t)+\mathbb{P}(Y>(1-\delta)t)\end{aligned}$$

but

$$?? \leq \mathbb{P}(X+Y \leq \varepsilon) \leq \min(\mathbb{P}(X \leq \varepsilon), \mathbb{P}(Y \leq \varepsilon))$$

• Moment estimates $a_n \leq \mathbb{E}\, X^n \leq b_n$ can be used for

$$\mathbb{E}\, e^{\lambda X} = \sum_{n=0} \frac{\lambda^n}{n!} \mathbb{E}\, X^n$$

but $\mathbb{E}$ exp$\{-\lambda X\}$ is harder to estimate.

• Exponential Tauberian theorem: Let V be a positive random variable. Then for $\alpha > 0$

$$\log \mathbb{P}(V \leq \varepsilon) \sim -C_V \varepsilon^{-\alpha} \quad \text{as} \quad \varepsilon \to 0^+$$

if and only if

$$\log \mathbb{E} \exp(-\lambda V) \sim -(1+\alpha)\alpha^{-\alpha/(1+\alpha)} C_V^{1/(1+\alpha)} \lambda^{\alpha/(1+\alpha)}$$

as $\lambda \to \infty$.

Ex: Let X_i, $i \geq 1$, be i.i.d. random variables with $\mathbb{E}\, X_i = 0$ and $\mathbb{E}\, X_i^2 = 1$, $\mathbb{E}\, \exp(t_0|X_1|) < \infty$ for $t_0 > 0$, and $S_n = \sum_{i=1}^n X_i$. Then as $n \to \infty$ and $x_n \to \infty$ with $x_n = o(\sqrt{n})$

$$\log \mathbb{P}\left(\frac{1}{\sqrt{n}} \max_{1 \leq i \leq n} |S_i| \geq x_n\right) \sim -\frac{1}{2} x_n^2$$

and as $n \to \infty$ and $\varepsilon_n \to 0$, $\sqrt{n}\varepsilon_n \to \infty$

$$\log \mathbb{P}\left(\frac{1}{\sqrt{n}} \max_{1 \leq i \leq n} |S_i| \leq \varepsilon_n\right) \sim -\frac{\pi^2}{8} \varepsilon_n^{-2}.$$

•**Open:** Find

$$\log \mathbb{P}\left(\max_{1 \leq i \leq n} |S_i| \leq C\right) \sim -??n.$$

Note that $?? \neq \pi^2/8$.

Ex: Let $L_\mu(n)$ be the length of the longest increasing subsequence (or records) in i.i.d sample $\{(X_i, Y_i)\}_{i=1}^n$ with law μ. Then

$$\lim_{n\to\infty} \frac{L_\mu(n)}{\sqrt{n}} = 2J_\mu.$$

The upper tail is known and for $c > 0$

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}} \log \mathbb{P}\left(L_\mu(n) > (2J_\mu + c)\sqrt{n}\right) = -U_\mu(c).$$

The lower tail is unknown in general, but for $0 < c < 2J_\mu$

$$\log \mathbb{P}\left(L_\mu(n) < (2J_\mu - c)\sqrt{n}\right) \approx -n.$$

See Deuschel and Zeitouni (1999), Aldous and Diaconis (1999), Okounkov (2000), and Li (2004+) for Gaussians.

•**Open:** Find

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(L_\mu(n) < (2J_\mu - c)\sqrt{n}\right).$$

I think there are some recent work on this direction.

Ex: For one-dim Brownian motion $B(t)$ under the sup-norm, we have by scaling

$$\begin{aligned}\log \mathbb{P}\left(\sup_{0\le t\le 1} |B(t)| \le \varepsilon\right) &= \log \mathbb{P}\left(\sup_{0\le t\le T} |B(t)| \le 1\right) = \log \mathbb{P}\left(\tau_2 \ge T\right) \\ &\sim -\frac{\pi^2}{8}\cdot T \sim -\frac{\pi^2}{8}\frac{1}{\varepsilon^2}\end{aligned}$$

as $\varepsilon \to 0$ and $T = \varepsilon^{-2} \to \infty$. Here $\tau_2 = \inf\{s : |B(s)| \ge 1\}$ is the first two-sided exit (or passage) time.

Ex: (One sided exit time)

$$\begin{aligned}\mathbb{P}\left(\sup_{0\le t\le 1} B(t) \le \varepsilon\right) &= \mathbb{P}\left(\sup_{0\le t\le T} B(t) \le 1\right) = \mathbb{P}\left(\tau_1 > T\right) \\ &= \mathbb{P}\left(|B(T)| \le 1\right) \sim (2/\pi)^{1/2} T^{-1/2} = (2/\pi)^{1/2}\varepsilon\end{aligned}$$

where $\tau_1 = \inf\{s : B(s) = 1\}$ is the one-sided exit time.

• For Gaussian process $X(t)$ with $X(0) = 0$, there are very few cases the behavior

$$\mathbb{P}(\sup_{0\le t\le 1} X(t) \le \varepsilon), \quad \varepsilon \to 0^+$$

is known.

Some Formulations for General Processes

Let $X = (X_t)_{t \in T}$ be a real valued stochastic process (not necessary Gaussian) indexed by T.
The large deviation under the sup-norm: $\mathbb{P}\left(\sup_{t \in T}(X_t - X_{t_0}) \geq \lambda\right)$ as $\lambda \to \infty$.

The small ball (deviation) probability: $\log \mathbb{P}(\|X\| \leq \varepsilon)$ as $\varepsilon \to 0$ for any norm $\|\cdot\|$.

The small ball probability under the sup-norm: $\mathbb{P}(\sup_{t \in T} |X_t| \leq \varepsilon)$ as $\varepsilon \to 0$.
Two-sided exit problem: $\mathbb{P}(\sup_{t \in T} |X_t| \leq 1)$ as $|T| \to \infty$.

The lower tail probability: $\mathbb{P}\left(\sup_{t \in T}(X_t - X_{t_0}) \leq \varepsilon\right)$ as $\varepsilon \to 0$ with $t_0 \in T$ fixed.
One-sided exit problem: $\mathbb{P}\left(\sup_{t \in T}(X_t - X_{t_0}) \leq 1\right)$ as $|T| \to \infty$.

•For processes with scaling property, problems equivalent for $\varepsilon \to 0$ and for $|T| \to \infty$.

The Lower Tail Probability

Let $X = (X_t)_{t \in T}$ be a real valued Gaussian process indexed by T. The lower tail probability studies

$$\mathbb{P}\left(\sup_{t\in T}(X_t - X_{t_0}) \le \varepsilon\right) \text{ as } \varepsilon \to 0$$

with $t_0 \in T$ fixed. Some general upper and lower bounds are given in Li and Shao (2004). In particular, for d-dimensional Brownian sheet $W(t)$, $t \in \mathbb{R}^d$,

$$\log \mathbb{P}\left(\sup_{t\in[0,1]^d} W(t) \le \varepsilon\right) \approx -\log^d \frac{1}{\varepsilon}.$$

Many open problems remain and new techniques are needed.

• Known cases: Brownian motion(BM), Brownian bridge, OU process, integrated BM, fractional BM, and a few more.

• The rate for the integrated fractional Brownian motion is related to the singularity of Burger's equation, See Sinai (1992), Molchan (1999, 2001, 2004, 2006), Li and Shao (2005).

• The rate for the m-th integrated Brownian motion is related to the positivity exponent of random polynomials, see Li and Shao (2011+).

Exit Time, Principal Eigenvalue, Heat Equation

Let D be a smooth open (connected) domain in $\mathbb{R}^d$ and τ_D be the first exit time of a diffusion with generator A. For bounded domain D and strong elliptic operator A, by Feynman-Kac formula,

$$\lim_{t\to\infty} t^{-1} \log \mathbb{P}(\tau_D > t) = -\lambda_1(D)$$

where $\lambda_1(D) > 0$ is the principal eigenvalue of $-A$ in D with Dirichlet boundary condition.

Ex: Brownian motion in $\mathbb{R}^d$ with $A = \Delta/2$. Let $v(x,t) = \mathbb{P}_x\{\tau_D \geq t\}$ Then v solves $\begin{cases} \frac{\partial v}{\partial t} = \frac{1}{2}\Delta v & \text{in } D \\ v(x,0) = 1 & x \in D. \end{cases}$ So this type of results can be viewed as long time behavior of $\log v(x,t)$, which satisfies a nonlinear evolution equation.

• Unbounded domain D and/or degenerated differential operator A. Li (2003), van den Berg (2004), Kuelbs and Li (2004), Bañuelos and DeBlassie (2006).

Hamiltonian and Partition Function

One of the basic quantity in various physical models is the associated Hamiltonian (energy function) H which is a nonnegative function. The asymptotic behavior of the partition function (normalizing constant) $\mathbb{E}\, e^{-\lambda H}$ for $\lambda > 0$ is of great interests and it is directly connected with the small value behavior $\mathbb{P}(H \le \epsilon)$ for $\epsilon > 0$ under appropriate scaling.

In the one-dim Edwards model a Brownian path of length t receives a penalty $e^{-\beta H_t}$ where H_t is the self-intersection local time of the path and $\beta \in (0, \infty)$ is a parameter called the strength of self-repellence. In fact

$$H_t = \int_0^t \int_0^t \delta(W_u - W_v) du dv = \int_{-\infty}^{\infty} L^2(t, x) dx$$

It is known, see van der Hofstad, den Hollander and König (2002), that

$$\lim_{t\to\infty}\frac{1}{t}\log\mathbb{E}\,e^{-\beta H_t}=-a^*\beta_{2/3}$$

where $a^*\approx 2.19$ is given in terms of the principal eigenvalues of a one-parameter family of Sturm-Liouville operators. Bounds on a^* appeared in van der Hofstad (1998).

• Chen and Li (2011+): For the one-dim Edwards model, it is not hard to show

$$\lim_{\varepsilon\to 0}\varepsilon^{2/(p+1)}\log\mathbb{P}\{\int_{-\infty}^{\infty}L^p(1,x)dx\le\varepsilon\}=-c_p$$

for some unknown constant $c_p>0$. Bounds on c_p can be given by using Gaussian techniques.

•Chen (2010), Chen and Rosinski (2011): Renormalization and asymptotics for physical models.

•Chen, Li, Rosinski and Shao (2011): Large deviations for local times and intersection local times

SVP for the Martingale Limit of a Galton-Watson Tree

Consider the Galton-Watson branching process $(Z_n)_{n\geq 0}$ with offspring distribution $(p_k)_{k\geq 0}$ starting with $Z_0 = 1$. In any subsequent generation individuals independently produce a random number of offspring according to $\mathbb{P}(N = k) = p_k$. Suppose $\mu = \mathbb{E}\, N > 1$ and $\mathbb{E}\, N \log N < \infty$. Then by Kesten-Stigum theorem, the martingale limit (a.s and in L^1)

$$W = \lim_{n\to\infty} \frac{Z_n}{\mu^n}$$

exists and is nontrivial almost surely with $\mathbb{E}\, W = 1$. WOLG, assume $p_0 = 0$ and $p_k < 1$ for all $k \geq 1$. Then in the case $p_1 > 0$, there exist constants $0 < c < C < \infty$ such that for all $0 < \varepsilon < 1$

$$c\varepsilon^\tau \leq \mathbb{P}(W \leq \varepsilon) \leq C\varepsilon^\tau, \qquad \tau = -\log p_1 / \log \mu$$

and in the case $p_1 = 0$, there exist constants $0 < c < C < \infty$ such that for all $0 < \varepsilon < 1$

$$c\varepsilon^{-\beta/(1-\beta)} \leq -\log \mathbb{P}(W \leq \varepsilon) \leq C\varepsilon^{-\beta/(1-\beta)}.$$

with $\nu = \min\{k \geq 2 : p_k \neq 0\}$ and $\beta = \log \nu / \log \mu < 1$.

•These results are due to Dubuc (1971a,b) in the $p_1 > 0$ case, and up to a Tauberian theorem also in the $p_1 = 0$ case, see Bingham (1988). A probabilistic argument is given in Mörters and Ortgiese (2008).

•Asymptotics for the survival probability in killed branching random walk, Gantert, Hu and Shi (2010).

•Similar results for supercritical branching processes with immigration, Chu, Li and Ren (2011).

Smoothness of the Density via Malliavin Matrix

Consider $F = (F^1, \cdots, F^m) : \Omega \to \mathbb{R}^m$ with $F^i \in D^{1,2}$. Then Malliavin Matrix of F is

$$\gamma_F = (\gamma_F^{ij}), \quad \gamma_F^{ij} = \left\langle DF^i, DF^j \right\rangle$$

Thm:(Bouleau-Hirsch) If $\det(\gamma_F) > 0$, a.s, then the law of F is absolute continuous.

Thm: (Malliavin) If (1) $F^i \in D^\infty$ and (2) $\mathbb{E}|\det \gamma_F|^{-p} < \infty$ for any $p > 0$, then F has a C^∞ density.

•The condition (ii) is called non-degeneracy for F.

•All these have been extended into theory of SDE and SPDE. It is curial to check the non-degeneracy condition which is small value probability.

•Mueller and Nualart (2008): Regularity of the density for the stochastic heat equation.
•Fei, Hu and Nualart (2011+): convergence of densities.
•Nualart (2010, book): Malliavin Calculus and its Applications.

Covering Number, Metric Entropy and ε-nets

Let A be a compact subset in a metric space (E, ρ), and let $\varepsilon > 0$ be given. The *metric entropy* of A is denoted by $H(A, \rho, \varepsilon) = \log N(A, \rho, \varepsilon)$ where

$$\begin{aligned} N(A, \varepsilon) &= N(A, \rho, \varepsilon) = N(A, \varepsilon B_\rho) \\ &= \min \Big\{ n \geq 1 : \exists x_1, \cdots, x_n \in A \\ &\qquad \text{such that } A \subset \cup_{j=1}^{n} (x_j + \varepsilon B_\rho) \Big\}, \end{aligned}$$

and $B_\rho(a; r) = \{x : \rho(x, a) < r\}$ is the open ball of radius r centered at a.

We also say a set $F \subset \mathbb{R}^d$ is an ε-net for A with respect to B if $A \subset \cup_{x \in F}(x + \varepsilon B)$. The smallest cardinality of an ε-net is denoted by $N(A, B, \varepsilon) = N(A, \varepsilon B)$.

• The metric entropy is a natural representation of how many bits you need to send in order to identify an element of a set up to precision ε. It is a tool heavily used in approximation theory, probability and statistics, learning theory, compressive sensing and random matrix theory

Ex:

$$A_2 = \{f \in C[0,1] : f(0) = 0, |f(x) - f(y)| \le |x - y|^\alpha, \forall x, y \in [0,1]\}$$

for $0 < \alpha \le 1$ and $\|f\| = \sup_{0 \le x \le 1} |f(x)|$.

Then $H(A_2, \varepsilon) \approx (1/\varepsilon)^{1/\alpha}$ as $\varepsilon \to 0$.

Ex:

$$\begin{aligned} A_3 &= \{f \in C[0,1] : f(0) = 0, |f(x) - f(y)| \le |x - y|^\alpha, \\ &\quad \forall x, y \in [0,1] \text{and} \quad \mathrm{Var}(f, [0,1]) \le 1\ \} \end{aligned}$$

where $0 < \alpha < 1$.

Then $H(A_3, \varepsilon) \approx (1/\varepsilon) \log(1/\varepsilon)$ as $\varepsilon \to 0$ (Clements, 1963).

Ex:

$$K = \left\{ g \in C[0,1] : g(0) = 0, g \,\text{absol. cont.}, \int_0^1 |g'(s)|^2 ds \le 1 \right\}.$$

Note that $K \subset A_3$ when $\alpha = 1/2$ since

$$\begin{aligned} |g(t) - g(s)| &= \left| \int_s^t g'(u) du \right| \\ &\le (t-s)^{1/2} \left(\int_s^t \left| g'(u) \right|^2 du \right)^{1/2} \\ &\le (t-s)^{1/2} \end{aligned}$$

and $\text{Var}(g) \le \int_0^1 |g'(t)| \, dt \le 1$

- Kolmogorov and Tihomirov (1961): $H(\varepsilon, K, \|\cdot\|_2) \approx 1/\varepsilon$.
- Birman and Solomjak (1967): $H(\varepsilon, K, \|\cdot\|_\infty) \approx 1/\varepsilon$.
- Kuelbs and Li (1993): As $\varepsilon \to 0$

$$(2 - \sqrt{3})/4 \le \varepsilon \cdot H(K, \|\cdot\|_2, \varepsilon) \le 1$$
$$(2 - \sqrt{3})\pi/4 \le \varepsilon \cdot H(K, \|\cdot\|_\infty, \varepsilon) \le \pi.$$

Note that K is the unit ball of the reproducing kernel Hilbert space generated by Brownian motion.

Links between Small Ball and Metric Entropy

As it was established in Kuelbs and Li (1993) and completed Li and Linde (1999), the behavior of

$$\log \mathbb{P}\left(\|X\| \le \varepsilon\right)$$

for Gaussian random element X is determined up to a constant by the metric entropy of the unit ball of the reproducing kernel Hilbert space associated with X, and vice versa.

- The Links can be formulated for entropy numbers of compact operator from Hilbert space to Banach space.

- This is a fundamental connection (both asymptotic and non-asymptotic) that has been used to solve important questions on both directions.

Gaussian Process, Operator and RKHS

The following statements are equivalent:
(i). X is a centered Gaussian random vector with law $\mu = \mathcal{L}(X)$ in a separable Banach space E.
(ii). There exist a separable Hilbert space H and an operator $u : H \to E$ such that $\sum_{j=1}^{\infty} \xi_j u(f_j)$ converges a.s. in E for one (each) ONB $(f_j)_{j=1}^{\infty}$ in H and

$$X \stackrel{d}{=} \sum_{j=1}^{\infty} \xi_j u(f_j)$$

where ξ_j are i.i.d. $N(0,1)$.
(iii). There are $x_1, x_2, \ldots$ in E such that $\sum_{j=1}^{\infty} \xi_j x_j$ converges a.s. in E and

$$X \stackrel{d}{=} \sum_{j=1}^{\infty} \xi_j x_j \ .$$

• The series $\sum_{j=1}^{\infty} \xi_j u(f_j)$ converges a.s. implies that u is compact and the RKHS $H_\mu = u(H)$ with compact unit ball K_μ in E.

• The RKHS H_μ can also be described as the completion of the range of the mapping $S : E^* \to E$ defined by the Bochner integral $Sf = \int_E x f(x) d\mu(x), f \in E^*$.

Ex: For the standard Brownian motion $W(t), 0 \le t \le 1$ on $C[0,1]$, the associated compact operate is the integration operator

$$uf(t) = \int_0^t f(s)ds.$$

The unit ball of the RKHS is

$$K = \left\{ g \in C[0,1] : g(0) = 0, g \,\text{absol. cont.}, \int_0^1 |g'(s)|^2 ds \le 1 \right\}.$$

- Kolmogorov and Tihomirov (1961): $H(\varepsilon, K, \|\cdot\|_2) \approx 1/\varepsilon$.
- Birman and Solomjak (1967): $H(\varepsilon, K, \|\cdot\|_\infty) \approx 1/\varepsilon$.
- Kuelbs and Li (1993): As $\varepsilon \to 0$

$$(2-\sqrt{3})/4 \le \varepsilon \cdot H(\varepsilon, K, \|\cdot\|_2) \le 1$$
$$(2-\sqrt{3})\pi/4 \le \varepsilon \cdot H(\varepsilon, K, \|\cdot\|_\infty) \le \pi.$$

Correlation inequalities

The Gaussian correlation conjecture: For any two symmetric convex sets A and B in a separable Banach space E and for any centered Gaussian measure μ on E,

$$\mu(A \cap B) \geq \mu(A)\mu(B).$$

An equivalent formulation: If $(X_1, \ldots, X_n)$ is a centered, Gaussian random vector, then

$$\mathbb{P}\left(\max_{1\leq i\leq n} |X_i| \leq 1\right) \geq \mathbb{P}\left(\max_{1\leq i\leq k} |X_i| \leq 1\right) \mathbb{P}\left(\max_{k+1\leq i\leq n} |X_i| \leq 1\right)$$

for each $1 \leq k < n$.

- Sidak inequality: The above holes for $k = 1$ or any slab B.

The weaker Correlation inequality: For any $0 < \lambda < 1$, any symmetric, convex sets A and B,

$$\mu(A \cap B)\mu(\lambda^2 A + (1-\lambda^2)B) \geq \mu(\lambda A)\mu((1-\lambda^2)^{1/2}B).$$

In particular,

$$\mu(A \cap B) \geq \mu(\lambda A)\mu((1-\lambda^2)^{1/2}B)$$

and

$$\mathbb{P}(X \in A, Y \in B) \geq \mathbb{P}\Big(X \in \lambda A\Big)\mathbb{P}\Big(Y \in (1-\lambda^2)^{1/2}B\Big)$$

for any centered joint Gaussian vectors X and Y.

The varying parameter λ plays a fundamental role in applications, see Li (1999). It allows us to justify

$$\mu(A \cap B) \approx \mu(A) \quad \text{if} \quad \mu(A) \ll \mu(B).$$

Note also that

$$\mu(\cap_{i=1}^m A_i) \geq \prod_{i=1}^m \mu(\lambda_i A_i)$$

for any $\lambda_i \geq 0$ with $\sum_{i=1}^m \lambda_i^2 = 1$.

For the weaker correlation inequality established in Li (1999), here is a very simple proof given in Li and Shao (2001). Let $a = (1 - \lambda^2)^{1/2}/\lambda$, and (X^*, Y^*) be an independent copy of (X, Y). Then $X - aX^*$ and $Y + Y^*/a$ are independent. Thus, by Anderson inequality

$$\begin{aligned}
\mathbb{P}(X \in A, Y \in B) &\geq \mathbb{P}(X - aX^* \in A, Y + Y^*/a \in B) \\
&= \mathbb{P}(X - aX^* \in A)\mathbb{P}(Y + Y^*/a \in B) \\
&= \mathbb{P}\Big(X \in \lambda A\Big)\mathbb{P}\Big(Y \in (1 - \lambda^2)^{1/2} B\Big).
\end{aligned}$$

Consider the sums of two centered Gaussian random vectors X and Y in a separable Banach space E with norm $\|\cdot\|$.

Thm: If X and Y are independent and

$$\begin{aligned}
\lim_{\varepsilon\to 0} \varepsilon^{\gamma} \log \mathbb{P}\left(\|X\| \le \varepsilon\right) &= -C_X, \\
\lim_{\varepsilon\to 0} \varepsilon^{\gamma} \log \mathbb{P}\left(\|Y\| \le \varepsilon\right) &= -C_Y
\end{aligned}$$

with $0 < \gamma < \infty$ and $0 \le C_X, C_Y \le \infty$. Then

$$\begin{aligned}
\limsup_{\varepsilon\to 0} \varepsilon^{\gamma} \log \mathbb{P}\left(\|X+Y\| \le \varepsilon\right) &\le -\max(C_X, C_Y) \\
\liminf_{\varepsilon\to 0} \varepsilon^{\gamma} \log \mathbb{P}\left(\|X+Y\| \le \varepsilon\right) &\ge -\left(C_X^{1/(1+\gamma)} + C_Y^{1/(1+\gamma)}\right)^{1+\gamma}.
\end{aligned}$$

Thm: If two joint Gaussian random vectors X and Y, *not* necessarily independent, satisfy

$$\begin{aligned}
\lim_{\varepsilon\to 0} \varepsilon^{\gamma} \log \mathbb{P}\left(\|X\| \le \varepsilon\right) &= -C_X, \\
\lim_{\varepsilon\to 0} \varepsilon^{\gamma} \log \mathbb{P}\left(\|Y\| \le \varepsilon\right) &= 0
\end{aligned}$$

with $0 < \gamma < \infty$, $0 < C_X < \infty$. Then

$$\lim_{\varepsilon\to 0} \varepsilon^{\gamma} \log \mathbb{P}\left(\|X+Y\| \le \varepsilon\right) = -C_X.$$

A Symmetrization Inequality for Two Norms

Let $K \subset \mathbb{R}^d$ and $L \subset \mathbb{R}^d$ be two origin symmetric convex bodies, $\|\cdot\|_K$ and $\|\cdot\|_L$ be the corresponding gauges on $\mathbb{R}^d$, that is the norms for which K and L are the unit balls.

Let $C_+ = C_+(\|\cdot\|_K, \|\cdot\|_L, d, a, b,)$ be the optimal constant such that, for all $\mathbb{R}^d$-valued i.i.d. random variables X and Y, and $a, b > 0$,

$$\mathbb{P}(\|X+Y\|_L \le b) \le C_+ \cdot \mathbb{P}(\|X-Y\|_K \le a).$$

•For $d = 1$, it is not hard to show $C_+ \le \lceil 2b/a \rceil + 1$.

•Schultze and Weizsäcker (2007): For $d = 1$ and $a = b$, $C_+ = 2$ which answers an open problem for about 10 years.

•Dong, J. Li and Li (2011+):

$$C_+ \le N(B_L(b), B_K(a/2)),$$

and the bound are optimal for $\|\cdot\|_K = \|\cdot\|_L = \|\cdot\|_\infty$ with $C_+ = \lceil 2b/a \rceil^d$.

Probability of all real zeros for random polynomial in the exponential ensemble

Thm (Li (2011)). The probability that a random polynomial of degree n with i.i.d exponentially distributed coefficients has all real zeros is

$$\mathbb{P}(\text{All zeros are real}) = \mathbb{E} \prod_{1 \le j < k \le n} |U_j - U_k| = \left(\prod_{k=1}^{n-1} \binom{2k+1}{k} \right)^{-1}$$

where U_i are i.i.d uniform on the interval $[0, 1]$.

•In particular, we have

$$p_1^e = 1, \quad p_2^e = \frac{1}{3}, \quad p_3^e = \frac{1}{30}, \quad p_4^e = \frac{1}{1050} \quad p_5^e = \frac{1}{132300}.$$

•Asymptotically, $\log \mathbb{P}(N_n = n) \sim -\log 2 \cdot n^2$ as $n \to \infty$.

•The second identity is a form of Selberg integral.

•Our evaluation of the probability starts with a formula of Zaporozhets (2004) which is based on an integral geometry representation developed by Edelman and Kostlan (1995) and tools from differential geometry.

Other Invited Talks (more to be added):

Xia Chen (University of Tennessee): Laplace asymptotics and Brownian functionals
Frank Gao (University of Idaho): Interplays with Convex geometry and/or bracket entropy
Nguyen Hoi (University of Pennsylvania):Singularity and/or Littlewood-Offer type estimates with applications to random matrices
Yaozhong Hu (University of Kansas): Applications to smoothness of density of Gaussian functionals via Malliavin calculus
Thomas Kuehn (Universitat Leipzig): Interplays with approximation and learning theory
James Kuelbs (University of Wisconsin-Madison): Branching related small value probabilities
Michael Lacey (Georgia Institute of Technology): Small ball inequalities and harmonic analysis
Yimin Xiao (Michigan State University): Interplays with fractal geometry for Gaussian fields

Typical Small Value Behavior

To make precise the meaning of typical behaviors that positive random variables take smaller values, consider a family of *non-negative* random variables $\{Y_t, t \in T\}$ with index set T. We are interested in evaluation $\mathbb{E} \inf_{t \in T} Y_t$ or its asymptotic behavior as the size of the index set T goes to infinity.

Ex: The first passage percolation indexed by paths.

Ex: Random assignment type problems indexed by permutations.

Conj: (Li and Shao) For any centered Gaussian r.v's $(X_i)_{i=1}^n$,

$$\mathbb{E} \min_{1 \le i \le n} |X_i| \ge \mathbb{E} \min_{1 \le i \le n} |\widehat{X}_i|$$

where $\widehat{X}_i$ are ind. centered Gaussian with $\mathbb{E}\,\widehat{X}_i^2 = \mathbb{E}\,X_i^2$.

Yes for $n = 2, 3$.

Gordon, Litvak, Schutt and Werner (2006):

$$2\mathbb{E} \min_{1 \le i \le n} |X_i| \ge \mathbb{E} \min_{1 \le i \le n} |\widehat{X}_i|$$

Expected Lengths of Minimum Spanning Tree (MST)

For a simple, finite, and connected graph G with vertex set $V(G)$ and edge set $E(G)$, we assign a non-negative i.i.d random length ξ_e with distribution F to each edge $e \in E(G)$. The total length of the MST is denoted by

$$L^F_{MST}(G) = \min_T \sum_{e \in T} \xi_e = \sum_{e \in MST(G)} \xi_e.$$

In particular, we use the notation $\mathbb{E}[L^u_{MST}(G)]$ for $U(0,1)$ and $\mathbb{E}[L^e_{MST}(G)$ for exp(1).

•Frieze (1985): For complete graph K_n on n vertices,

$$\lim_{n\to\infty} \mathbb{E}[L^e_{MST}(K_n)] = \lim_{n\to\infty} \mathbb{E}[L^u_{MST}(K_n)] = \zeta(3) = \sum_{k=1}^{\infty} k^{-3} = 1.202...$$

See related results in Steele (1987), Frieze and McDiarmid (1989), Janson (1995). Pennose (1998), Beveridge, Frieze McDiarmid (1998), Frieze, Ruszink and Thoma (2000), Fill and Steele (2004), Gamarnik (2005).

Exact Formula

•Steele (2002):

$$\mathbb{E}[L^u_{MST}(G)] = \int_0^1 \frac{(1-t)}{t} \frac{T_x(G;1/t,1/(1-t))}{T(G;1/t,1/(1-t))}dt,$$

where $T(G:x,y)$ is the Tutte polynomial of G and $T_x(G;x,y)$ is the partial derivative of $T(G;x,y)$ with respect to x.

•Li and X. Zhang (2009): For complete graph K_n,

$$0 < \mathbb{E}[L^e_{MST}(K_n)] - \mathbb{E}[L^u_{MST}(K_n)] = \frac{\zeta(3)}{n} + O\left(n^{-2}\log^2 n\right).$$