Progress and Prospects in Small Value/Deviation Probabilities

Wenbo V. Li
University of Delaware
E-mail: wli@math.udel.edu
St. Petersburg, Sep., 2005

In the past two years (after the first conference) there has been a great deal of progress in various directions. All talks in this conference show important progress and future directions. This talk will highlight some of recent developments, in particular connections with other parts of mathematics and works less represented in this meeting.

We believe a theory of small value probabilities should be developed and centered on:
• systematically studies of the existing techniques and applications
• applications of the existing methods to a variety of fields
• new techniques and problems motivated by current interests of advancing knowledge

Small value 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

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

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

What is often an important and interesting problem is the determination of just how "rare" the event $\{Y_n \le h_n\}$ is, that is, the study of the *small value 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.

• 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

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

as $\lambda \to \infty$.

Geometric Functional Analysis

Large deviation estimates are by now a standard tool in the Asymptotic Convex Geometry, contrary to small deviation results. Very recently, novel applications of small deviation estimates to problems related to the diameters of random sections of high dimensional convex bodies are realized. They imply distinction between the lower and the upper inclusions in the celebrated Dvoretzky Theorem, which says that any n-dimensional convex body has a section of dimension $c \log n$ that is approximately a Euclidean ball. Recall that One of the early manifestations of the concentration of measure phenomenon was V. Milman's proof of Dvoretzky Theorem in the 70s.

PP: Small ball probability and Dvoretzky Theorem, negative moments of a norm (Kahane-Khinchine type inequality for negative exponents), Klartag and Vershynin (2004+).

PP: Gaussian inequalities related to symmetric convex sets with applications to small ball probabilities. Cordero-Erausquin, Fradelizi and Maurey (2004), Latała and Oleszkiewicz (2005+)

Precise Links with Metric Entropy

As it was established in Kuelbs and Li (1993) and completed Li and Linde (1999), the behavior of $\log \mathbb{P}(\|X\| \le \varepsilon)$ 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 Banach space to Hilbert space.
• This is a fundamental connection that has been used to solve important questions on both directions.

PP: Small ball or entropy number for tensors and probabilistic understanding for the tensored Gaussian processes. Gao and Li (2005), Blei and Gao (2005+).
PP: Similar connections for other measures such as stable. One direction is given in Li and Linde (2003) which could be used to disprove the **duality conj.** on entropy numbers of a compact operator.

Gaussian Fields via Riesz Product

This is a new powerful technique developed in Gao and Li (2005+) for the upper bound under sup-norm. In the case of Brownian sheet ($d = 2$), it allows us to give a simple and general approach to avoid ingenious combinatoric arguments used by Talagrand (1994). The basic ideas are

• Choosing Basis: Use (multi-dim) series expansion $X(t) = \sum_{n=1}^{\infty} f_n(t)\xi_n$, where ξ_n are i.i.d. standard normal random variables, and $f_n \in C([0,1]^d)$.

• Choosing Partial Sum: By Andersen's inequality, $\mathbb{P}(\|X\| \le \varepsilon) \le \mathbb{P}(\|Y\| \le \varepsilon)$ where $Y(t)$ is any partial sum $X(t) = \sum_{n \in E} f_n(t)\xi_n$.

• Construct Riesz Product:

$$\mathbb{P}(\|Y\| \le \varepsilon) \le \mathbb{P}(\int Y(t)R(t) \le \varepsilon)$$

where the Riesz product $R(t) = \prod_{n \in F}(1 + \varepsilon_n h_n)$ satisfying $R(t) \ge 0$, $\|R\|_1 = \int R(t)dt = 1$.

PP: Brownian sheet for $d \ge 3$ and other interesting Gaussian fields; entropy number for tensored operators.

The Lower Tail Probability

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

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

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 $W(t)$, $t \in \mathbb{R}^d$,

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

Note that we can write

$$\|X\| = \sup_{f \in D} f(X)$$

so the lower tail formulation is more general than the small ball problem.

PP: Sharper estimates for interesting Gaussian processes/fileds with applications; connections with properties of the generating operator. Li and Shao (2004, 2005+).

Zeros of Random Polynomial

Let $a_0, a_1, \ldots, a_n \in \mathbb{R}$ be i.i.d. Define the random polynomial

$$f_n(x) := \sum_{i=0}^{n} a_i x^i \ .$$

Let N_n denote the number of real zeros of $f_n(x)$.

Dembo, Poonen, Shao and Zeitouni (2002): If $a_i \sim N(0,1)$, then For n even,

$$\mathbb{P}(N_n = 0) = \mathbb{P}(f_n(x) > 0, \forall x \in \mathbb{R}) = n^{-b+o(1)}$$

where

$$b = -4 \lim_{t\to\infty} \frac{1}{t} \log \mathbb{P}\Big(\sup_{0\le s\le t} Y(s) \le 0 \Big)$$

and $\{Y(t),\ t \ge 0\}$ is a centered stationary Gaussian process with $\mathbb{E}\, Y(t)Y(s) = \frac{2e^{-(t-s)/2}}{1+e^{-(t-s)}}$.

PP: Exact value of the positive exponent b; Existence of b in the symmetric stable case; Sharp estimates for small deviation $\mathbb{P}\,(N_n \le (1-\delta)\mathbb{E}\, N_n)$ and large deviation $\mathbb{P}\,(N_n \ge (1+\delta)\mathbb{E}\, N_n)$. Li and Shao (2005).

The Wiener-Hopf Equation

The Wiener-Hopf equation

$$H(x) = \int_0^\infty f(x-y)H(y)dy, \quad x \geq 0$$

is still an active area of study, even the existence and uniqueness of a solution.

Spitzer (1956) has obtained a beautiful formula (Spitzer's identity) from which one can (in principle at least) calculate the joint distribution of any pair $(\max_{0 \leq j \leq n} S_j, S_n)$ knowing the individual distributions of the first n partial sums, $S_0 = 0, S_k = X_1 + \cdots + X_k$. He then used it in Spitzer (1957, 1960a,b) to study the Wiener-Hopf equation. Here is a typical result.

Let $f(x)$ be the density of X, i,e, $F(x) = \int_{-\infty}^x f(t)dt$. If X is symmetric with characteristic function $\phi(\lambda)$, then

$$\lim_{n\to\infty} n^{1/2}\mathbb{P}(\max_{0\leq k\leq n} S_k \leq x) = \pi^{-1/2}H(x)$$

where $H(x)$ is the unique solution (in the class of functions that are non-decreasing, continuous on the right, with $H(0) > 0$) of the Wiener-Hopf equation

$$H(x) = \int_0^\infty f(x-y)H(y)dy$$

and $H(0^+) = 1$. In addition, the Laplace transform of $H(x)$ is given for $\lambda > 0$ by

$$\begin{aligned}\int_{0^-}^\infty e^{-\lambda x} dH(x) &= 1 + \int_{0^+}^\infty e^{-\lambda x} dH(x) \\ &= \exp\left\{-\frac{1}{2\pi}\int_{-\infty}^\infty \frac{\lambda}{\lambda^2+t^2}\log(1-\phi(t))\right.\end{aligned}$$

Moreover, if $\mathbb{E}\,X^2 = \sigma^2 < \infty$, then $H(x)$ has the asymptotic behavior

$$\lim_{x\to\infty}\frac{H(x)}{x} = \frac{\sqrt{2}}{\sigma}.$$

If the variance is infinite, then $H(x) = o(x)$ as $x \to \infty$.

PP: Purely probabilistic arguments with bounds on $\mathbb{P}(\max_{0\le k\le n} S_k \le x)$ and $H(x)$ under weaker moment conditions. Li and C. Zhang (2005+).

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 $\mathbb{P}(H \leq \epsilon)$ 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 (2005+): 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.

PP: Many open questions in the area.

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 $\begin{cases} \frac{\partial v}{\partial t} = \frac{1}{2}\Delta v \text{in } D \\ v(x,0) = 1 \quad x \in D \end{cases}$

So this type of results can be viewed as long time behavior of $\log v(x,t)$, which satisfies a nonlinear evolution equation.

PP: Unbounded domain D and/or degenerated differential operator A. Li (2003), Lifshits and Shi (2003), van den Berg (2004), Bañuelos and Carroll (2004+), Bañuelos and K. Bogdan (2004+), Bañuelos and DeBlassie (2005+).

Brownian pursuit problems

Let $\{W_k(t);\ t \geq 0\}(k = 0, 1, 2, \dots)$ denote independent Brownian motions all starting from 0. Define

$\tau_n = \inf\{t > 0 : W_i(t) = 1 + W_0(t) \text{ for some } 1 \leq i \leq n\}.$

It is known for the exit time τ_n of a cone that

$$\mathbb{P}\{\tau_n > t\} \sim ct^{-\gamma_n}, \quad \text{as} \quad t \to \infty,$$

where γ_n is determined by the first eigenvalue of the Dirichlet problem for the Laplace-Beltrami operator on a subset of the unit sphere $\mathbb{S}^n$ in $\mathbb{R}^{n+1}$.

Conj: Bramson and Griffeath (1991), $\mathbb{E}\,\tau_4 < \infty$.

Li and Shao (2001): $\mathbb{E}\,\tau_5 < \infty$ by using Gaussian distribution identities and the Faber-Krahn isoperimetric inequality.

Li and Shao (2002): $\lim_{n\to\infty} \gamma_n / \log n = 1/4$ by developing a normal comparison inequality (a 'reverse' Slepian's inequality). This verified a conjecture of Kesten (1992).

Ratzkin and Treibergs (2005+): $\mathbb{E}\,\tau_4 < \infty$ by purely analytic estimates of eigenvalue.

Let B_{-i}, $0 \le i \le m-1$ and B_j, $1 \le j \le n$ be independent Brownian motions, starting at 0.

Define the first capture time by

$$\tau_{1,m,n} = \inf\{t > 0 : \max_{1\le j\le n} B_j(t) = \min_{0\le i\le m-1} B_{-i}(t) + 1\}$$

and the overall capture time by

$$\tau_{m,m,n} = \inf\{t > 0 : \max_{1\le j\le n} B_j(t) = \max_{0\le i\le m-1} B_{-i}(t) + 1\}.$$

Then we have

$$\begin{aligned} & \mathbb{P}\left(\tau_{1,m,n} > t\right) \\ = \; & \mathbb{P}\left(\max_{1\le j\le n} \sup_{0\le s\le t} \max_{0\le i\le m-1} (B_j(s) - B_{-i}(s)) < 1\right) \end{aligned}$$

and

$$\begin{aligned} & \mathbb{P}\left(\tau_{m,m,n} > t\right) \\ = \; & \mathbb{P}\left(\max_{1\le j\le n} \sup_{0\le s\le t} \min_{0\le i\le m-1} (B_j(s) - B_{-i}(s)) < 1\right). \end{aligned}$$

Conj: Let

$$\mathbb{P}\left(\tau_{n,n,1} > t\right) \sim ct^{-\beta_n} \quad \text{as} \quad t \to \infty.$$

Then $\beta_n \sim n^{-1} \log n$

Many More Areas

• Special Gaussian Chaos. Kuelbs and Li (2005),
• Determinant of random matrix. Tao and Vu (2004+), Costello, Tao and Vu (2005+).
• Littlewood and Offord type problems.
• Existence in random graphs.
• Combinatorial discrepancy.
• Hadamard conjecture.
• Balancing vectors.

Komlos Conj: Let $x_1, \cdots, x_n \in \mathbb{R}^n$ be arbitrary vectors with $\|x_k\|_2 \le 1$. Then there exist signs $\varepsilon_k = \pm 1$, $1 \le k \le n$ such that

$$\|\sum_{k=1}^{n} \varepsilon_k x_k\|_\infty \le C$$

where C is some numerical constant. That is

$$\mathbb{P}\left(\|\sum_{k=1}^{n} \varepsilon_k x_k\|_\infty \le C\right) \ge \frac{1}{2^n}.$$

PP: It is known from Li (2005+) that most Conj. and results on small values hold for ξ_k in combinatorial discrepancy and balancing vectors. Are there any comparison results between ξ_k and ε_k?

Small Value Phenomenon

Two fundamental problems in probability theory are typical behaviors such as expectations, laws of large numbers and central limit theorems, and rare events such as large deviations. Small value phenomenon comes from both typical behaviors and rare events of the type that positive random variables take smaller values.

- **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. Examples discussed in a short course in Beijing this summer include Gaussian comparison inequalities for $\mathbb{E} \min_{1\le i\le n} |X_i|$, random assignment type problems indexed by permutations, and the first passage percolation indexed by paths.