Small Value Phenomenons in Probability and Statistics

Wenbo V. Li
University of Delaware
E-mail: wli@math.udel.edu
East Lansing, Oct. 2009

Two fundamental problems in probability theory and statistical analysis are typical behaviors such as expectations, laws of large numbers, central limit theorems and approximated sampling distributions, and rare events such as large deviations, significant level and power. Small value probability studies the probability of the rare events that positive random variables take smaller values. The associated typical behavior deals with the expectation of the minimum of a family of positive random variables. We will provide an overview on current and emerging opportunities in the area.

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

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

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 separated 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 (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

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

as $\lambda \to \infty$.

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 $?? \neq \pi^2/8$ and see Li and Zinn (2009+) for more details.

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 (2009+) for Gaussians.

•**Open:** Find

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

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

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

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.

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

The large deviation under the sup-norm:

$$\mathbb{P}\left(\sup_{t \in T}(X_t - X_{t_0}) \geq \lambda\right) \text{ as } \lambda \to \infty.$$

The small ball (deviation) probability:

$$\log \mathbb{P}\left(\|X\| \leq \varepsilon\right) \quad \text{as} \quad \varepsilon \to 0$$

for any norm $\|\cdot\|$.

The small ball probability under the sup-norm:

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

The lower tail probability:

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

with $t_0 \in T$ fixed.

• The last two types of probability can also be viewed as the first exit time problems if the process has scaling property.

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 (2009+).

Precise Links with Metric Entropy

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

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

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 Banach space to Hilbert space.

- This is a fundamental connection that has been used to solve important questions on both directions.

Open: Small ball or entropy number for tensors.

Open: Probabilistic understanding for small balls of the tensored Gaussian.

Open: Similar connections for other measures such as stable. One direction is given in Li and Linde (2003).

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,

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

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).

Sequential Analysis

Several stopping times which arise from problems in approximations for sequential point and interval estimation may be written in the form

$$t_c = \inf\{n \geq m \colon S_n < cn^{\alpha} L(n)\},$$

where $S_n = X_1 + \cdots + X_n$, $X_1, X_2, \cdots$ are i.i.d. *positive* r.v's with $\mathbb{E}\, X_1^2 < \infty$, $L(n) = 1 + O(n^{-1})$, $\alpha > 1$, $m \geq 1$ and $c > 0$.

• The probability of stopped early

$$\mathbb{P}(t_c \leq (1-\delta)\mathbb{E}\, t_c) \sim K_{m,\delta} \cdot c^{(m-1)/2\alpha}, \qquad c \to 0,$$

which is strongly influenced by the initial sample size m.

• The uniform integrability of $|t_c^*|^r$ in c is determined by the behavior of

$$\mathbb{P}(X_1 \leq x) \quad \text{or} \quad \mathbb{P}(S_m \leq x), \qquad x \to 0$$

where

$$t_c^* = \frac{t_c - \mathbb{E}\, t_c}{\sqrt{\mathrm{Var}(t_c)}} \Rightarrow N(0, 2\alpha^2).$$

See Robbins (1959), Chow and Robbins (1965), Starr and Woodroofe (1968, 1972), Woodroofe (1977, 1982), Lai and Siegmund (1977), Yu (1981), Takada (1992), etc.

An Edgeworth Curiosum

Let X_1 and X_2 be i.i.d samples with density $f_k(x-\theta)$, where

$$f_k(x) = 2^{-1}(k-1)(1+|x|)^{-k}, \quad k > 1.$$

Then for $\varepsilon > 0$ small,

$$\mathbb{P}(|\frac{X_1 + X_2}{2} - \theta| \leq \varepsilon) \leq \mathbb{P}(|X_1 - \theta| \leq \varepsilon),$$

i.e. the sample mean provides a bigger error than a single observation under the criterion judged by $\mathbb{P}(|\widehat{\theta} - \theta| \leq \varepsilon)$ for a given $\varepsilon > 0$ small.

•For a detailed study, see S. Stigler (1980), An Edgeworth curiosum. *Ann. Stat*, 8, 931–934.

•For any i.i.d samples X_1 and X_2,

$$\mathbb{P}(|X_1 + X_2| \leq x) \leq 2 \cdot \mathbb{P}(|X_1 - X_2| \leq x), \quad x > 0,$$

and the constant 2 is the best possible.

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

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

• Chen and Li (2009+): For the one-dim Edwards model,

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

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

• Klenke and Morters (2005): Let $l_{m,n}(B(0,1))$ be the (projected) intersection local time of m vs n independent Brownian paths in $\mathbb{R}^d$ for $d = 2, 3$ inside the unit ball $B(0,1) \subset \mathbb{R}^d$. Then

$$\lim_{\varepsilon\to 0}\frac{\log\mathbb{P}(l_{m,n}(B(0,1))\le\varepsilon)}{-\log\varepsilon}=-\frac{\xi_d(m,n)}{4-d}$$

where $\xi_d(m,n)$ are called the non-intersection exponents. The values $\xi_2(m,n)$ are found by Lawler, Schramm and Werner based on SLE. Much less is known in $\mathbb{R}^3$.

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 (2009+)

•Similar results for variants of branching process, Chu, Li and Ren (2010+).

Smoothness of the Density via Malliavin Matrix

Lemma: Let $M(\omega) = (m_{ij})_{n\times n}$ be a symmetric nonnegative definite random matrix with moments of all order for m_{ij}. If for any $p \geq 2$

$$\sup_{|v|=1} \mathbb{P}(v^T M v \leq \varepsilon) = O(\varepsilon^p), \quad \text{as} \quad \varepsilon \to 0^+,$$

Then $\det(M^{-1}) = (\det M)^{-1} \in L^p$ for all $p > 0$.

•In many applications of Malliavin calculus to the smoothness of the density of the solutions of SPDEs, one needs to show the inverse of the determinant of the Malliavin matrix has moments of all orders, or equivalently, the determinant of the Malliavin matrix has negative moments of all orders.

•In fact, the negative moments estimates

$$\mathbb{E}\, V^{-p} < \infty \quad \text{for any/all} \quad p > 0$$

is equivalent to the upper small value estimates

$$\mathbb{P}(V \leq \varepsilon) \leq C_p \varepsilon^p \quad \text{for any/all} \quad p > 0, \quad \text{as} \quad \varepsilon \to 0.$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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):

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

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):

$$\mathbb{E}[L_{MST}^{F}(G)] = \int_0^\infty \frac{1-F(t)}{F(t)} \frac{T_x(G;x,y)}{T(G;x,y)} \, \mathrm{d}t,$$

where $x = 1/F(t)$, $y = 1/(1-F(t))$. In particular,

$$\mathbb{E}[L_{MST}^{e}(G)] = \int_0^1 \frac{1}{t} \frac{T_x(G;1/t,1/(1-t))}{T(G;1/t,1/(1-t))} \, \mathrm{d}t,$$

and for any connected graph G,

$$\mathbb{E}[L_{MST}^{u}(G)] < \mathbb{E}[L_{MST}^{e}(G)].$$

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

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

Combinatorial Optimization

The TSP (travelling salesman problem, i.e. find the shortest route through a set of points) is the paradigm problem in this area.

• Let $L_n = \min_\sigma \sum_{i=1}^n |X_{\sigma(i)} - X_{\sigma(i+1)}|$ be the shortest tour of n i.i.d uniform points $\{X_1, \cdots, X_n\} \subset [0,1]^d$. Then $\mathbb{E}\, L_n / n^{(d-1)/d} \to \beta(d)$. Find "good" estimates on $\beta(d)$.

• Does the Central Limit Theorem hold, i.e. does the length of the optimal tour have a Normal distribution as n tends to infinity?

• Can one prove anything about the geometric structure of the optimal tour?

• Two-sample matching: There is $0 < c_0 < c_1 < \infty$ such that

$$c_0 \le \frac{\mathbb{E}\, M_n}{\sqrt{n \log n}} \le c_1, \quad c_0 \le \frac{\mathbb{E}\, M_n^*}{n^{-1/2} (\log n)^{3/4}} \le c_1$$

where

$$M_n = \min_\sigma \sum_{i=1}^n |X_i - Y_{\sigma(i)}|, \quad M_n^* = \min_\sigma \max_{1 \le i \le n} |X_i - Y_{\sigma(i)}|$$

and $\{X_i\}$ and $\{Y_i\}$ are i.i.d uniform samples on $[0,1]^2$. Show the limiting constants exists.

Small Value Theory

We believe a theory of small value phenomenon 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

Applications of small deviation probabilities

- Chung's law of the iterated logarithm
- Lower limits for empirical processes
- Rates of convergence of Strassen's FLIL
- Rates of convergence of Chung type FLIL
- A Wichura type functional LIL
- Fractal Geometry for Gaussian random fields
- Metric entropy estimates
- Capacity in Wiener space
- Natural Rates of escape for infinite dimensional Brownian motions
- Asymptotic evaluation of Laplace transform for large time
- Onsager-Machlup functionals
- Random fractal laws of the iterated logarithm

All are discussed in details in the survey paper of Li and Shao (2001).

Applications not included in the survey:

- Volume of Wiener sausage and fractional Brownian sausage
- Classical and average Kolmogorov widths
- Hypercontractivity and comparison of moments of iterated maxima and minima
- Cascade relations for intersection exponents of planar Brownian motion
- Estimates of principle eigenvalue of (fractional) Laplacian
- Exit time of Brownian motion from unbounded domain, principal eigenvalue, heat equation
- Entropy and quantization of Gaussian measure
- Regularity of density for functionals of Gaussian processes
- Decaying turbulent transport
- Random sum of vectors
- Cube slicing
- Dvoretzky theorem in geometric functional analysis, negative moments of a norm
- Hamiltonian and Partition Function
- The Wiener-Hopf Equation

- Longest increasing subsequences, longest common increasing subsequences
- Determinant of random matrix
- Littlewood and Offord type problems
- Existence in random graphs.
- Combinatorial discrepancy.
- Hadamard conjecture.
- Most visit sites via isomorphism theorems
- Singularity of Burgers equation
- Galton-Watson tree and DNA testing
- Gaussian free fields
- Etc.