Small Value Problems in Mathematics

Wenbo V. Li
University of Delaware
E-mail: wli@math.udel.edu
Oberwolfach, October 12, 2003

We present a collection of problems and techniques in various parts of mathematics from the point of view of small value probabilities. 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.

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 are being developed for Gaussian measures.

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

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

but

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

• Moment estimates $a_n \le \mathbb{E}X^n \le b_n$ 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 $\alpha > 0$

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

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 $\lambda \to \infty$.

Precise Links with Metric Entropy

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

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

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 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) which could be used to disprove the **duality conj.** on entropy numbers of a compact operator.

Sums of Gaussian Random Vectors

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

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

Open: Find the exact constant in terms of C_X, C_Y and covariances of X and Y.

The following is given in Li (1999) based on a weaker correlation inequality.

Thm: If two joint Gaussian random vectors X and Y, *not* necessarily independent, satisfy $0 < C_X < \infty$ and $C_Y = 0$ with $0 < \gamma < \infty$. Then

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

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 (2002+). 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.

Open: Are there any connections with properties of the generating operator?

Positivity Exponent 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

$$\mathbb{E}\, Y(t)Y(s) = \frac{2e^{-(t-s)/2}}{1 + e^{-(t-s)}}$$

Moreover, $0.4 < b \le 2$. Their numerical simulations for degree $n \le 2^{10}$ suggest $b \approx 0.76 \pm 0.03$.

The estimate is also true for a_i with all finite moments based on KMT strong approximations.

Conj: The result is true for any mean zero a_i in DAN. This is offered in Poonen and Stoll (1999) in the study of Density of Hyperelliptic curve of large genus with odd Jacobian over Q.

Conj: Dembo, Poonen, Shao and Zeitouni (2002). For n even and a_i symmetric stable of order α, $0 < \alpha < 2$,

$$\mathbb{P}(N_n = 0) = n^{-b_\alpha + o(1)}$$

for some $b_\alpha > b$

Open: If $\{X_t, t \geq 0\}$ is a differential stationary Gaussian process with positive correlation, what is the limit

$$\lim_{T\to\infty} \frac{1}{T} \log \mathbb{P}\Big(\sup_{0\leq t\leq T} X_t \leq 0\Big) \ ?$$

Open: Find sharp estimates for small deviation

$$\mathbb{P}\left(N_n \leq (1-\delta)\mathbb{E}\, N_n\right)$$

and large deviation

$$\mathbb{P}\left(N_n \geq (1+\delta)\mathbb{E}\, N_n\right)$$

as $n \to \infty$ for $0 < \delta < 1$, where a_i are i.i.d $N(0,1)$ and $\mathbb{E}\, N_n \sim (2/\pi) \log n$ given in Kac (1943).

Let $W(t)$, $t \geq 0$, be the standard Brownian motion starting at 0. Denote by $W_0(t) = W(t)$ and

$$W_m(t) = \int_0^t W_{m-1}(s) ds, \quad t \geq 0, \quad m \geq 1.$$

the m times integrated Brownian motion for positive integer m. Using integration by parts,

$$W_m(t) = \frac{1}{m!} \int_0^t (t-s)^m dW(s), \quad m \geq 0.$$

The $\mathbb{R}^{m+1}$ valued process

$$(W_0(t), W_1(t), \cdots, W_m(t))$$

is Markov with degenerated generator

$$\mathcal{L} = \frac{\partial^2}{\partial x_0^2} + \sum_{k=1}^{m} x_{k-1} \frac{\partial}{\partial x_k}.$$

Li and Shao (2003+): There exist constants $r_m >$ and $r > 0$ such that

$$\begin{aligned}
\mathbb{P}\left(\sup_{0\leq s\leq \log t} Y(s) \leq 0\right) &\approx t^{-r+o(1)},\\
\mathbb{P}\left(\sup_{0\leq s\leq t} X(s) \leq 1\right) &\approx t^{-r+o(1)},\\
\mathbb{P}\left(\sup_{0\leq t\leq 1} X(t) \leq \varepsilon\right) &\approx \varepsilon^{2r+o(1)},\\
\mathbb{P}\left(\sup_{0\leq t\leq 1} W_m(t) \leq \varepsilon\right) &\approx \varepsilon^{r_m+o(1)},\\
\mathbb{P}\left(\sup_{0\leq s\leq t} W_m(s) \leq 1\right) &\approx t^{-r_m(2m+1)/2+o(1)}
\end{aligned}$$

and $r_m(2m+1)/2$ decrease to r as $m \to \infty$, where $X(t)$ is a centered Gaussian process with

$$\mathbb{E}\, X(t)X(s) = \frac{2st}{t+s}.$$

In particular, $b = 4r \leq 1$ since $r_1 = 1/6$ from McKean (1963), Sinai (1992).

Open: Finding $b = 4r$ and r_m, $m \geq 2$.

Brownian pursuit problems

Let $\{W_k(t);\ t \geq 0\}(k = 0, 1, 2, \ldots)$ 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).

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

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

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

The Exit Time from Unbounded Domain

Consider the first exit time τ_D of (d+1)-dimensional Brownian motion from the unbounded domain

$$D = \{(x, y) \in \mathbb{R}^{d+1} : y > f(x), x \in \mathbb{R}^d\}$$

starting at the point $(x_0, f(x_0) + 1) \in \mathbb{R}^{d+1}$ for some $x_0 \in \mathbb{R}^d$, where the function $f(x)$ on $\mathbb{R}^d$ is convex and $f(x) \to \infty$ as the Euclidean norm $|x| \to \infty$. In Li (2002), very general estimates for the asymptotics of $\log \mathbb{P}(\tau_D > t)$ are found by using Gaussian techniques. In particular, for $f(x) = \exp(|x|^p)$, $p > 0$,

$$\lim_{t \to \infty} \frac{(\log t)^{2/p}}{t} \log \mathbb{P}(\tau_D > t) = -j_\nu^2/2$$

where $\nu = (d-2)/2$ and j_ν is the smallest positive zero of the Bessel function J_ν. Sharp estimates are obtained in Lifshits and Shi (2003) for $f = |x|^\gamma$. the Dirichlet heat kernal is studied in van den Berg (2003+):

Conj: The general lower bound in Li (2002) is sharp.

Connections with heat equation: Let

$$v(x, t) = \mathbb{P}_x\{\tau_D \ge t\}, \quad x \in \mathbb{R}^{d+1}.$$

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 our results can be viewed as long time behavior of $\log v(x, t)$. Furthermore, a close related and useful technique in studying certain asymptotic problems is the logarithmic transformation $V = -\log v(x, t)$ which changes the above equation into a nonlinear evolution equation for V. This can then be viewed as a stochastic control problem.

Connections with principal Dirichlet eigenvalue: For bounded smooth open (connected) domain $\widetilde{D}$, by Feynman-Kac formula,

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

where $\lambda_1(\widetilde{D}) > 0$ is the principal eigenvalue of $-\Delta/2$ in $\widetilde{D}$ with Dirichlet boundary condition.

Fundamental Inequalities

- Gaussian Isoperimetric inequalities; Ehrhard's inequality; S-inequality; Logarithmic Sobolev inequality; Bobkov's inequality; etc.

- Concentration and deviation inequalities

- Comparison inequalities; Anderson's inequality; Shift type inequalities; Slepian's inequality; Gordon's min-max inequalities; Reverse Slepian type inequalities; etc.

- Correlation inequalities.

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

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

Sidak inequality: The above holds for 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

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

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

Smooth Analysis of Simplex Method for Linear Programming

Simplex method for linear programming:

$$\max B^T x \quad s.t. \quad Ax \leq y$$

- Worst case analysis: exponential.
- Average (Gaussian for A) case analysis: polynomial.
- Widely used in practice.

Smooth analysis of simplex method:

$$\max B^T x \quad s.t. \quad (A + \sigma G)x \leq y$$

where $G = (g_{ij})$, $1 \leq i, j \leq n$, with i.i.d normal g_{ij}.

Edelman (1988):

$$\mathbb{P}\left(\|G^{-1}\| > t\right) \leq \frac{\sqrt{n}}{t}.$$

with the best constant. Here $\|A\| = \max_{\|x\|=1} \|Ax\|$ denotes the operator norm of A.

Sanker, Spielman and Teng (2002):

$$\mathbb{P}\left(\|(G+A)^{-1}\| > t\right) \leq \frac{1.823\sqrt{n}}{t}$$

Gaussian Perturbation Conj:

$$\mathbb{P}\left(\|(G+A)^{-1}\| > t\right) \leq \mathbb{P}\left(\|G^{-1}\| > t\right)$$

This is a part of understanding how things behavior under perturbation, such as $Ax = b$ for the input (A, b).

Note that

$$\|M^{-1}\| = \frac{1}{d(M, \mathcal{S})}$$

so this is really a small value problem. Here the distance

$$\begin{aligned} d(M, \mathcal{S}) &= \inf_{S \in \mathcal{S}} d(M, S) \\ &= \inf_{\det(s_{ij})=0} \left(\sum_{i,j=1}^d (m_{ij} - s_{ij})^2\right)^{1/2}. \end{aligned}$$

Intersection Local Times

The mathematical notion of various intersection local times was motivated by the models of polymer physics and quantum field theory.

One of the basic quantity in the study is the associated Hamiltonian (energy function) H which is a nonnegative function of the paths. 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 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).

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.

Open: Small deviation for the mixed intersection time.

$$\int_{-\infty}^{\infty} \prod_{i=1}^{m} L_i^{p_i}(1,x)dx$$

where L_i are i.i.d local times and $p_i \ge 1$.

Open: Small deviation for two-dim re-normalized self-intersection time.

Hitting Probability of a Set

Let X_t, $t \ge 0$, be a fractional Brownian motion on $\mathbb{R}^d$ with index $0 < H < 1$. Then

$$\mathbb{P}(\inf_{1\le t\le 2} |X_t| \le \varepsilon) \begin{cases} \approx \varepsilon^{d-1/H} & \text{if } d > 1/H \\ > \delta & \text{if } d < 1/H \\ \approx (\log 1/\varepsilon)^{-??} & \text{if } d = 1/H \end{cases}$$

The motivations for extending results classical for Brownian motion to the fractional Brownian motions are not only the importance of these processes, but also the force to find proofs that relay upon general principles at a more fundamental level by moving away from crucial properties (such as the Markov property) of Brownian motion. Fractional Brownian motion might not be an object of central mathematical importance but abstract principles are.

Random graphs

Let $G(n,p)$ be a random graph and $\omega(G)$ denote the number of vertices in the maximum clique of the graph G.

Thm: For $k = o(\log n)$,

$$\mathbb{P}(\omega(G(n,1/2)) \le k) = \exp(-n^{2+o(1)})$$

Note that a $o(1)$ in the hyper-exponent leaves lots of room! Also, It is not difficult to show that $\omega(G(n,1/2))$ is concentrated at $2\log_2 n$

Let

$$\mathbb{P}(X_{ij} = 0) = p = p_n, \quad \mathbb{P}(X_{ij} = 1) = 1 - p$$

and

$$H_n = \sum_{1\le i<j<k<m\le n} X_{ij}X_{jk}X_{km}X_{mi}$$

Then

$$\mathbb{P}(H_n = 0) \begin{cases} \to 1 & \text{if} p = \Omega(n^{-1}) \\ \le \text{poly. small} & \text{if} p = n^{-2/3} \\ \le \text{exp. small} & \text{if} p = n^{-1/2} \end{cases}$$

Open: What is the correct cut off behavior?

Littlewood and Offord Type Problems

Let $a_1, a_2, \cdots, a_n$ be vectors in $\mathbb{R}^d$ with $|a_i| \ge 1$ for all i. Let η_i be i.i.d r.v's with $\mathbb{P}(\eta_i = 0) = \mathbb{P}(\eta_i = 1) = 1/2$. Then

$$\mathbb{P}\left(\left|\sum_{i=1}^{n} \eta_i a_i\right| \le D\right) \le \frac{c}{\sqrt{n}}.$$

where $D \ge 1$ is a given constant and c depends only on D.

A considerable literature has been devoted to this problem, beginning with Erdos (1945). A variety of tools from extremal set theory and geometry has been used, see Kleitman (1970), Griggs (1980), Frankl and Füredi (1988).

Slicing the Cube

A cube of dimension n and side 1 is cut by a hyperplane of dimension $n-1$ through its center. The usual $n-1$ measure of the intersection is bounded between 1 and $\sqrt{2}$. Hensley (1979) and Ball (1988).

Thm: Let U_j be i.i.d uniform on $[-a, a]$. Then for any vector $v = (v_1, \cdots, v_n) \in \mathbb{R}^n$ with $|v| = \left(\sum_{j=1}^n v_j^2\right)^{1/2}$

$$\frac{1}{(1+a^2|v|^2)^{1/2}} \leq \mathbb{P}(|\sum_{j=1}^n v_j U_j| \leq 1) \leq \frac{\sqrt{2}}{(1+a^2|v|^2)^{1/2}}$$

Open: Sharp bounds for ε_k or general symmetric X_j with $\mathbb{E}\, X_j^2 = 1$

Combinatorial Discrepancy

Let $(V, \mathcal{F})$ be a set system, where $V = \{1, \cdots, n\}$. Such a combinatorial structure is often called a *hypergraph*. The discrepancy of a set system $\mathcal{F} \subset 2^V$ is

$$\text{disc}(\mathcal{F}) = \min_{\chi} \max_{A \in \mathcal{F}} \left| \sum_{a \in A} \chi(a) \right|$$

where χ ranges over "two-colorings" $\chi : V \to \{-1, +1\}$.

Thm: Any set system $(V, \mathcal{F})$ such that $|V| = |\mathcal{F}| = n$ has $O(\sqrt{n})$ discrepancy. Some set systems have a matching lower bound. Equivalently,

$$\mathbb{P}\left(\max_{F \in \mathcal{F}} \left| \sum_{v \in F} \varepsilon_v \right| \leq C\sqrt{n} \right) \geq \frac{1}{2^n}$$

and

$$\mathbb{P}\left(\max_{F \in \mathcal{F}} \left| \sum_{v \in F} \varepsilon_v \right| \leq c\sqrt{n} \right) = 0 < \frac{1}{2^n}$$

for some constants $C > c > 0$.

Beck-Fiala **Conj.** (1981): disc$(\mathcal{F}) \leq Ct^{1/2}$ if $|\mathcal{F}| \leq t$.

Thm: Let $A = (a_{ij})$, where $a_{ij} = 0$ or 1, be a matrix of size $n \times n$. Then for some $C > 0$

$$\mathbb{P}\left(\max_{1 \leq m \leq n} \max_{1 \leq k \leq n} \left| \sum_{i=1}^m \sum_{j=1}^k a_{ij}\varepsilon_{ij} \right| \leq C(\log n)^4 \right) \geq \frac{1}{2^n}.$$

- The Beck-Fiala conjecture implies $C(\log n)^3$ bound.

- There is a lower bound of $\Omega(\log n)$ given in Beck (1981).

Open: Find the correct order of the lower 'cut off' function.

Open: Gaussian version.

Balancing vectors

Consider an arbitrary pair of symmetric convex bodies, U and V in $\mathbb{R}^n$. Define $\beta(U, V) = \beta_n(U, V)$ as the smallest $r > 0$ satisfying the following: for every sequence $u_1, \cdots, u_n$ of vectors in $U \subset \mathbb{R}^n$ there exists a choice of signs $\varepsilon_1, \cdots, \varepsilon_n = \pm 1$ such that $\sum_{i=1}^n \varepsilon_i u_i \in rV$. And similarly, define $\alpha(U, V) = \alpha_n(U, V)$ as the smallest $r > 0$ such that $\sum_{i=1}^m \varepsilon_i u_i \in rV$ for all $1 \leq m \leq n$. Clearly, $\beta_n(U, V) \leq \alpha_n(U, V)$.

Reformulation: $\beta_n(U, V)$ is the smallest $r > 0$ such that for any $u_i \in U$, $1 \leq i \leq n$,

$$\mathbb{P}\left(\left\| \sum_{i=1}^n \varepsilon_i u_i \right\|_V \leq r \right) \geq \frac{1}{2^n}$$

and $\alpha_n(U, V)$ is the smallest $r > 0$ such that for any $u_i \in U$, $1 \leq i \leq n$,

$$\mathbb{P}\left(\max_{1 \leq m \leq n} \left\| \sum_{i=1}^m \varepsilon_i u_i \right\|_V \leq r \right) \geq \frac{1}{2^n}$$

where $\| \cdot \|_V$ is the norm with the unit ball V in $\mathbb{R}^n$.

Let B_p^n denote the unit L_p-ball in $\mathbb{R}^n$.

• $\beta(B_2^n, B_2^n) \le \sqrt{n}$, i.e. for any $u_1, \cdots, u_n \in \mathbb{R}^n$ with $|u_i|_2 \le 1$, there exist $\eta_1, \cdots, \eta_n = \pm 1$ so that

$$|\eta_1 u_1 + \cdots + \eta_n u_n| \le \sqrt{n}.$$

• Komlos Conjecture (197?): $\beta(B_2^n, B_\infty^n) \le C$ for some absolute constant $C > 0$. It is well known that Komlos Conjecture would imply Beck-Fiala Conjecture.
• Beck and Fiala (1981): $\beta(B_1^n, B_\infty^n) \le 2$.
• Spencer (1985, 1986):

$$\begin{aligned} c\sqrt{n} \ \le \ & \beta(B_\infty^n, B_\infty^n) \le \alpha(B_\infty^n, B_\infty^n) \le C\sqrt{n} \\ & \beta(B_2^n, B_\infty^n) \le C \log n \end{aligned}$$

• Spencer Conj: $\alpha(B_p^n, B_p^n) \le C n^{1/2+o(1)}$ for $1 \le p < \infty$.
• Giannopoulos (1997): $\beta(B_2^n, V) \le 6 \log n$ if $\gamma_n(V) \ge 1/2$ where γ_n is the standard n-dimensional Gaussian measure with density $(2\pi)^{-n/2} e^{-\|x\|_2^2/2}$.
• Banaszczyk (1998): $\beta(B_2^n, V) \le C$ if $\gamma_n(V) \ge 1/2$ and in particular, $\beta(B_2^n, B_\infty^n) \le C\sqrt{\log n}$.

Open: All Conj. and results above hold for ξ_k. Are there any comparison results between ξ_k and ε_k?

Three Permutation Conj: Let σ_i, $i = 1, 2, 3$ be three permutations on $[n]$. Then

$$\mathbb{P}\left(\max_{1\le i\le 3} \max_{1\le m\le n} \left|\sum_{j=1}^m \varepsilon_{\sigma_i(j)}\right| \le c\right) \ge \frac{1}{2^n}$$

for some absolute constant c, independent of n.

The k-permutation Conj: disc$= \Omega(\sqrt{k})$.
• disc$\ge c\sqrt{k}$ via Hadamard matrix.
• disc$\le C(k \log n)$ via the Partial coloring lemma.
• disc$\le C(\sqrt{k} \log n)$ via the entropy method.

Gaussian k-permutation Conj: Under the Gaussian correlation conj.

$$\begin{aligned} & \mathbb{P}\left(\max_{1\le m\le k} \max_{1\le j\le n} \left|\sum_{i=1}^j \xi_{\sigma_m(i)}\right| \le h_k\right) \\ \ge & \prod_{m=1}^k \mathbb{P}\left(\max_{1\le j\le n} \left|\sum_{i=1}^j \xi_{\sigma_m(i)}\right| \le h_k\right) \\ \ge & \prod_{m=1}^k \exp(-cn/h_k^2) = \exp(-cnk/h_k^2) \ge 2^{-n} \end{aligned}$$

if $h_k = \Omega(\sqrt{k})$.
• $h_k = \Omega(k)$ via the weaker Gaussian correlation inequality in Li (1999).

Random trigonometric polynomials

Salem and Zygmund (1954): Given complex numbers $z_1, \cdots z_n$, there exists a choice of $+$ and $-$ such that

$$\sup_{0\le t\le 1} \left|\sum_{k=1}^n \pm z_k e^{ikt}\right| \le C\left(\log n \sum_{k=1}^n |z_k|^2\right)^{1/2}$$

where C is an absolute constant.

Kahane (1980): There is a polynomial

$$P(z) = \sum_{k=1}^n e^{i\theta_k} z^k$$

such that

$$\sup_{|z|=1} |P(z)| \le \sqrt{n}(1 + O(n^{-19/34}(\log n)^{1/2}).$$

Open: Better estimate?

Erdős' Conjecture vs Littlewood's Conjecture

For $\varepsilon_k = \pm 1$, $k \ge 1$, define

$$P_n(t, \varepsilon) = \sum_{k=1}^n \varepsilon_k e^{2\pi k\, ti}.$$

Erdős conjectured that there is an absolute constant $c > 0$ such that for all $n \ge 2$ and all sequence ε_k

$$\sup_{0\le t\le 1} |P_n(t, \varepsilon)| \ge (1 + c)\sqrt{n}.$$

Littlewood conjectured the opposite, that is, there are $c_n \to 0$ and a sequence ε such that

$$\sup_{0\le t\le 1} |P_n(t, \varepsilon)| \le (1 + c_n)\sqrt{n}.$$

Small value reformulation: Take i.i.d ε_k with $\mathbb{P}(\varepsilon_k = \pm 1) = 1/2$,

$$\mathbb{P}\left(\sup_{0\le t\le 1} |P_n(t, \varepsilon)| \le (1 + c)\sqrt{n}\right) = 0 < 2^{-n}$$

vs

$$\mathbb{P}\left(\sup_{0\le t\le 1} |P_n(t, \varepsilon)| \le (1 + c_n)\sqrt{n}\right) \ge 2^{-n}$$

Gaussian version: Replace ε_k by ξ_k.

Discrepancy in Arithmetic Progression

It is a classical result of Van der Waerden that any two coloring of the integers contains an arbitrarily long monochromatic arithmetic progression. Below is a complementary result that not all arithmetic progression can be evenly bi-colored.

Roth's $\frac{1}{4}$ Thm (1964): Any two-coloring of the integers $[n] = \{1, \cdots, n\}$ contains an arithmetic progression whose discrepancy is $\Omega(n^{1/4}$.

Roth (1964): The bound of $\Omega(n^{1/2}(\log n)^{1/2}$.

Sarkozy: The bound of $\Omega(n^{1/3+\delta}$, see Erdos and Spencer (1974).

Beck (1981): The bound of $\Omega(n^{1/4}(\log n)^3)$.

Matousek and Spencer (1996): The bound is tight.

Reformulation: For the collection $\mathcal{A}$ of all arithmetic progressions of $[n]$,

$$\mathbb{P}\left(\max_{A\in\mathcal{F}}\left|\sum_{i\in A}\varepsilon_i\right| \le Cn^{1/4}\right) \ge \frac{1}{2^n}$$

and

$$\mathbb{P}\left(\max_{A\in\mathcal{F}}\left|\sum_{i\in A}\varepsilon_i\right| \le cn^{1/4}\right) = 0 < \frac{1}{2^n}$$

for some constants $C > c > 0$.

- Fourier transform method by Roth.
- Eigenvalue technique by Lovasz and Sos.
- Harmonic analysis approach

Conj:

$$\log\mathbb{P}\left(\max_{A\in\mathcal{F}}\left|\sum_{i\in A}\xi_i\right| \le n^{1/4}\right) \approx -n$$

Discrepancy in Geometry

A set of n points in $\mathbb{R}^d$ is well-spread if the ratio of the largest distance and the smallest distance is less than $cn^{1/d}$ for some absolute constant $c > 0$.

Alexander (1990): For any two-coloring of a well-spread set of n points in $\mathbb{R}^d$, there is an open hyperplane H such that $|D(H)| = \Omega(n^{1/2-1/2d})$.

Matousek (1995): Any set of n points in $\mathbb{R}^d$ can be two-coloring in such a way that the maximum discrepancy of any halfspace is at most proportional to $n^{1/2-1/2d}$.

- Similar to small deviation estimates, the finite difference method is a power tool in discrepancy analysis here.

Hadamard Conjecture: *There exists an Hadamard matrix H_n, or n by n matrix with every entry ± 1 such that $HH^T = nI$ for every $n = 4m$, $m \ge 1$.*

To restate the Hadamard Conjecture, let ε_{ij} be i.i.d random variables with $\mathbb{P}(\varepsilon_{ij} = \pm 1) = 1/2$, $1 \le i, j \le n$. Then the equivalent formulation of the Hadamard Conjecture is

$$\mathbb{P}\left(\max_{1\le j\ne k\le n}\left|\sum_{i=1}^{n}\varepsilon_{ij}\varepsilon_{ik}\right| < 1\right) \ge 2^{-n^2}.$$

for $n = 4m$.

Gaussian Hadamard Conjecture: *Let ξ_{ij}, $1 \le i, j \le n$, be i.i.d standard normal random variables. Then*

$$\log\mathbb{P}\left(\max_{1\le j\ne k\le n}\left|\sum_{i=1}^{n}\xi_{ij}\xi_{ik}\right| < 1\right) \approx -n^2.$$

The quest for proofs of these conjectures is likely to stimulate and to challenge probabilists for years to come.