Recent Developments on Small Value Probabilities

Wenbo V. Li

University of Delaware

E-mail: wli@math.udel.edu

HDP Workshop, Banff, Oct. 9-13, 2011

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 on some recent developments, including symmetrization inequalities in high dimension, smooth Gaussian processes, and branching related processes.

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 V_n be a sequence of *non-negative* random variables and suppose that some or all of the probabilities

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

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 $\mathbb{P}\left(V_n \le h_n\right) \to 0$ 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 $\{V_n \le h_n\}$ is, that is, the study of the *small value (deviation) probabilities* of V_n associated with the sequence h_n.

If $\varepsilon_n = \varepsilon$ and $V_n = \|X\|$, 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 $\lambda \to \infty$.

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 $|T| \to \infty$.

Applications: Partition functions, local times and statistics

•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

•Go, Li and Wellner (2010): How many Laplace transforms of probability measures are there? Applications to bracket entropy in empirical processes theory.

•van der Vaart and van Zanten (2008a,b): Statistical applications for Gaussian priors based on Reproducing kernel Hilbert spaces.

•Gine and Nickl (2011+): Non-parametric estimation, using lower bounds for small ball probabilities for m-th integrated BM.

•Nikitin and Pusev (2011): Refined estimates for weighted L_2-norm.

Integrated processes: Simply integrated r.w.

•Aurzada and Dereich (2011+): For X a LP or RW with $\exists\beta > 0$: $\mathbb{E}\, e^{\beta|X_1|} < \infty$ and $\mathbb{E}\, X_1 = 0$

$$\mathbb{P}\left(\sup_{0\le n\le T}\sum_{i=1}^{n} X_i \le 0\right) \approx \mathbb{P}\left(\sup_{0\le t\le T}\int_0^t X_s ds \le 1\right) = T^{-1/4} e^{O(\log\log T)}$$

•Dembo and Gao (2011+): For X a RW with $\exists\beta > 0$: $\mathbb{E}\, e^{\beta X_1^-} < \infty$, $\mathbb{E}\, X_1 = 0$, (+ some unimportant regularity cond. for X_1^-),

$$\mathbb{P}\left(\sup_{0\le n\le T}\sum_{i=1}^{n} X_i \le 0\right) \approx \sqrt{\frac{\mathbb{E}\,|X_T|}{T\,\mathbb{E}\,|X_1|}} \approx \begin{cases} T^{-1/4} & \text{if } \mathbb{E}\,(X_1^+)^2 < \infty \\ T^{-(1-1/\alpha)/2} & \text{if } X_1^+ \text{ in DoA}(\alpha) \end{cases}$$

•Vysotsky (2011+): For a couple of special cases (all require $\exists\beta > 0$: $\mathbb{E}\, e^{\beta X_1^-} < \infty$, $\mathbb{E}\, X_1 = 0$),

$$\mathbb{P}\left(\sup_{0\le n\le T}\sum_{i=1}^{n} X_i \le 0\right) \sim cT^{-(1-1/\alpha)/2},$$

if X_1^+ in DoA(α), $1 < \alpha \le 2$.

Lower tails for fractional BM

A fractional Brownian motion (FBM) B^H is a centered Gaussian process with covariance

$$\mathbb{E}\, B_t^H B_s^H = \frac{1}{2}\left(|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right), \qquad t, s \in \mathbb{R},$$

where $0 < H < 1$ is the Hurst parameter. For $H = 1/2$, this is a Brownian motion.

•Molchan (1999), Aurzada (2011): For fractional Brownian motion we have, for some $c > 0$,

$$T^{-(1-H)}(\log T)^{-c} \le \mathbb{P}\left(\sup_{0\le t\le T} B_t^H \le 1\right) \le T^{-(1-H)}(\log T)^{c},$$

•Aurzada, Li and Shao (2011+): Fractional integrated Brownian motion.

Small ball probability for smooth Gaussian processes

•Aurzada, Gao, Kühn, Li and Shao (2011): Small deviation probability for a family of smooth Gaussian processes with

$$\mathbb{E}\, X_{\alpha,\beta}(t) X_{\alpha,\beta}(s) = \frac{2^{2\beta+1}(ts)^\alpha}{(t+s)^{2\beta+1}}$$

for $\alpha > 0$ and $\beta > -1/2$.
Thm: For $\alpha > \beta > -1/2$,

$$-\log \mathbb{P}\left(\int_0^1 |X_{\alpha,\beta}(t)|^2\, dt \le \varepsilon^2\right) \sim \kappa_{\alpha,\beta} |\log \varepsilon|^3,$$

where the constant is given by $\kappa_{\alpha,\beta} := \frac{1}{3(\alpha-\beta)\pi^2}$.
For $\alpha > \beta + 1/2 > 0$,

$$\tilde{\kappa}_{\alpha,\beta} |\log \varepsilon|^3 \lesssim -\log \mathbb{P}\left(\sup_{t\in[0,1]} |X_{\alpha,\beta}(t)| \le \varepsilon\right) \lesssim \kappa_{\alpha-1/2,\beta} |\log \varepsilon|^3.$$

•Aurzada (2011+): Path regularity and small deviations of smooth Gaussian processes.

•Lifshits and Linde (2011): Gaussian summation processes on trees (non-smooth, discrete types).

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 $C_+ = \lceil 2b/a \rceil^d$.

An Extension of 123 Theorem

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 $b > a > 0$,

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

•Alon and Yuster (1995) and (independently) Kotlov: For $d = 1$, $C_- \leq 2\lceil b/a \rceil - 1$. In particular, for $a = 1$, $b = 2$, we have $C_- = 3$.

•Alon and Yuster (1995): For $\|\cdot\|_K = \|\cdot\|_L = \|\cdot\|_2$, $C_- \leq M$ if there is no set F of $M + 1$ points in a ball of radius b so that the center belongs to F and the distance between any two pints of F exceeds a. In addition, $C_- = M$ in special settings.

•Dong, J. Li and Li (2011):

$$C_- \leq N(B_L(b) \setminus B_K(a), B_K(a/2)) + 1,$$

and the bound is optimal for $d = 1$.

Our approach for both problems (C_+ and C_-) extends techniques developed in Schultze and Weizs$\ddot{a}$cker (2007) which starts with the following fact:

Lemma: The following two statement are equivalent for a given symmetric matrix $A = (a_{ij})_{n\times n}$:

(i) For all probability measure $p \in \mathcal{P} := \{p : \sum p_i = 1, p_i \geq 0\}$,

$$\sum_{i,j} a_{ij} p_i p_j > 0;$$

(ii) For all $p \in \mathcal{P}$,

$$\max_i \sum_j a_{ij} p_i p_j > 0.$$

•The above fact can be reformulated in the infinite dimensional setting for product measure.

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). The proofs are relying on nontrivial complex analysis and are therefore difficult to generalize, for example to processes with immigration and/or dependent offsprings.

•Examples, near-constancy phenomena and various refinements, see Harris (1948), Karlin and McGregor (1968 a,b), Dubuc (1982), Barlow and Perkins (1987), Goldstein (1987) and Kusuoka (1987), Bingham (1988), Biggins and Bingham (1991), Biggins and Bingham (1993), Biggins and Nadarajah (1994), Hambly (1995), Fleischman and Wachtel (2007, 2009).

•A probabilistic argument is given in Mörters and Ortgiese (2008).

SVP for supercritical branching processes with Immigration

Consider the supercritical branching process with immigration, denoted by $(\mathcal{Z}_n, n \geq 0)$. That is

$$\mathcal{Z}_0 = Y_0, \quad \mathcal{Z}_{n+1} = X_1^n + X_2^n + \cdots + X_{\mathcal{Z}_{n-1}}^n + Y_{n+1}, \quad n \geq 0,$$

where $X_1^n, X_2^n, \cdots$ are independent and identically distributed with the same offspring distribution as X, the $Y_0, Y_1, \cdots$ are i.i.d. with the same immigration distribution $\{q_k, k \geq 0\}$ and the $X's$ and $Y's$ are independent. It is classic result, see Seneta (1970), for example, that

$$\lim_{n\to\infty} \mathcal{Z}_n / m^n = \mathcal{W}$$

exists and is finite a.s. if and only if

$$\mathbb{E} \log^+ Y < \infty \quad \text{and} \quad \mathbb{E}\,(X \log X) < \infty.$$

where here and throughout, $\log^+ x = \log \max(x, 1) \geq 0$.

Thm: (Chu, Li and Ren (2011)) Assume the $X \log X$ and $\log Y$ conditions and $p_0 = 0$.

(a) If $0 < q_0 < 1$, then

$$\mathbb{P}(\mathcal{W} \le \varepsilon) \asymp \varepsilon^{|\log q_0|/\log m} \quad \text{as} \quad \varepsilon \to 0^+.$$

(b) If $q_0 = 0$ and $p_1 > 0$, then

$$\log \mathbb{P}(\mathcal{W} \le \varepsilon) \sim -\frac{\kappa |\log p_1|}{2(\log m)^2} \cdot |\log \varepsilon|^2, \quad \text{as} \quad \varepsilon \to 0^+,$$

with $\kappa = \inf\{n : q_n > 0\}$.

(c) If $q_0 = 0$ and $p_1 = 0$, then

$$\log \mathbb{P}(\mathcal{W} \le \varepsilon) \asymp -\varepsilon^{-\beta/(1-\beta)}, \quad \text{as} \quad \varepsilon \to 0^+,$$

with β being defined as in the case without immigration.

(d) If $p_0 > 0$, then

$$\mathbb{P}(\mathcal{W} \le \varepsilon) \asymp \varepsilon^{|\log h(\rho)|/\log m}, \quad \text{as} \quad \varepsilon \to 0^+,$$

where ρ is the solution of $f(s) = s$ between $(0, 1)$, and h is the generating function of immigration.

• The asymptotic $\asymp$ is best possible in the sense that it can not be improved into the more precise asymptotic $\sim$.

• The oscillation occurs with immigration even there is no oscillation without immigration. This is quite unexpected and demonstrates the significant effects of the immigration.

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.

Probability of all real zeros for random polynomial with 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 with simplification.

•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.

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 such as random matrices.

$\diamondsuit$ W.V. Li, Ten lectures on Small Deviation Probabilities: Theory and Applications, NSF/CBMS Regional Research Conference in the Mathematical Sciences, University of Alabama in Huntsville, June 04-08, 2012.