Ten Lectures on Small Value Probabilities and Applications

L2: Basic Estimates and Techniques for One Variable

Wenbo V. Li

University of Delaware

http://www.math.udel.edu/~wli
wli@math.udel.edu

CBMS Lectures at UAH, June 4-8, 2012

We first formulate several equivalent results for small value probability, including negative moments, exponential moments, Laplace transform and Taubirean theorems. The basic techniques involved are various useful inequalities, motivated from large deviation estimates.

Laplace Transform and Tauberian Theorems

In general, Tauberian theorems involve a transformation of a class of objects such as functions, series, sequences. The transformation is certain type of averaging and must have a continuity property such that certain limit behavior of the original class implies related limit behavior of the image of the transformation. A Tauberian theorem is to reverse the operation, to go from a limit property of the transform to a limit property of the original class.

•The aim in this lecture is mainly on the behavior of $\mathbb{P}(V \le t)$ and the Laplace transform $\mathbb{E}\, e^{-\lambda V}$, with some applications.

•The standard proofs in analysis is based on Karamata's method (1931), see Korevaar (2004, p30-32, p192-194) and Bingham, Goldie and Teugels (1987).

•Here we only present probabilistic arguments and connections.

Polynomial Rate

Thm: For constants $\alpha > 0$ and $C > 0$,

$$\mathbb{E}\, e^{-\lambda V} \sim C/\lambda^{\alpha} \quad \text{as} \quad \lambda \to \infty.$$

if and only if

$$\mathbb{P}(V \le t) \sim \frac{C}{\Gamma(1+\alpha)} t^{\alpha} \quad \text{as} \quad t \to 0.$$

•The following more general statement holds: for $\alpha \ge 0$ and slowly varying function L,

$$\mathbb{E}\, e^{-\lambda V} \sim C/(\lambda^{\alpha} L(\lambda)) \quad \text{as} \quad \lambda \to \infty.$$

if and only if

$$\mathbb{P}(V \le t) \sim \frac{C}{\Gamma(1+\alpha)} t^{\alpha}/L(1/t) \quad \text{as} \quad t \to 0.$$

•Here we only proof the direction that $\mathbb{P}$ implies $\mathbb{E}$ since the argument are instructive and also works even we only have a one-sided bound on probability.

Observe that using integration by parts or checking by Fubini's theorem,

$$\mathbb{E}\, e^{-\lambda V} = \int_0^\infty e^{-\lambda t} d\mathbb{P}(V \le t) = \int_0^\infty \lambda e^{-\lambda t} \mathbb{P}(V \le t) dt. \quad (0.1)$$

From the assumption, given any $\varepsilon > 0$ small, there is $\delta_\varepsilon > 0$ such that for all $0 \le t \le \delta_\varepsilon$,

$$(1-\varepsilon)\frac{C}{\Gamma(\alpha+1)} t^\alpha \le \mathbb{P}(V \le t) \le (1+\varepsilon)\frac{C}{\Gamma(\alpha+1)} t^\alpha.$$

For the lower bound, we have

$$\mathbb{E}\, e^{-\lambda V} \ge \int_0^{\delta_\varepsilon} \lambda e^{-\lambda t} \mathbb{P}(V \le t) dt \quad \ge \quad (1-\varepsilon)\frac{C}{\Gamma(\alpha+1)} \int_0^{\delta_\varepsilon} \lambda e^{-\lambda t} t^\alpha dt$$

and

$$\int_0^{\delta_\varepsilon} \lambda e^{-\lambda t} t^\alpha dt = \lambda^{-\alpha} \int_0^{\delta_\varepsilon \lambda} e^{-x} x^\alpha dx.$$

Thus

$$\liminf_{\lambda \to \infty} \lambda^\alpha \mathbb{E}\, e^{-\lambda V} \ge (1-\varepsilon)\frac{C}{\Gamma(\alpha+1)} \int_0^\infty e^{-x} x^\alpha dx = (1-\varepsilon) C.$$

Taking $\varepsilon \to 0$ gives the desired lower bound.

Similarly, for the upper bound, we have

$$\begin{aligned}
\mathbb{E}\, e^{-\lambda V} &\le \int_0^{\delta_\varepsilon} \lambda e^{-\lambda t} \mathbb{P}(V \le t) dt + \int_{\delta_\varepsilon}^{\infty} \lambda e^{-\lambda t} dt \\
&\le (1+\varepsilon) \frac{C}{\Gamma(\alpha+1)} \lambda^{-\alpha} \int_0^{\delta_\varepsilon \lambda} e^{-x} x^{\alpha} dx + e^{-\delta_\varepsilon \lambda}
\end{aligned}$$

and thus

$$\limsup_{\lambda \to \infty} \lambda^{\alpha} \mathbb{E}\, e^{-\lambda V} \le (1+\varepsilon) C.$$

Taking $\varepsilon \to 0$ finishes the proof. $\square$

Similarly, for the upper bound, we have

$$\begin{aligned}
\mathbb{E}\, e^{-\lambda V} &\leq \int_0^{\delta_\varepsilon} \lambda e^{-\lambda t}\mathbb{P}(V \leq t)dt + \int_{\delta_\varepsilon}^{\infty} \lambda e^{-\lambda t} dt \\
&\leq (1+\varepsilon)\frac{C}{\Gamma(\alpha+1)}\lambda^{-\alpha}\int_0^{\delta_\varepsilon \lambda} e^{-x}x^{\alpha}dx + e^{-\delta_\varepsilon \lambda}
\end{aligned}$$

and thus

$$\limsup_{\lambda\to\infty} \lambda^{\alpha}\mathbb{E}\, e^{-\lambda V} \leq (1+\varepsilon)C.$$

Taking $\varepsilon \to 0$ finishes the proof. $\square$

•The argument from Laplace transform to probability in Korevaar (2004, p30-32, p192-194) is based on Feller's continuity theorem for convergence of a family of transforms. Hence we do not have a one-sided probabilistic estimates like the proof above. It would be interesting to find one. The main point is that we may need to use conditioning argument in probability estimates directly and thus asymptotic is less useful.

•Are there a probabilistic flavor arguments for the other direction? A straightforward argument based on exponential Chebyshev inequality does not seem work well. In fact, assuming $\mathbb{E}\, e^{-\lambda V} \le C/\lambda^\alpha$ for $\lambda > 0$ large, we have,

$$\mathbb{P}(V \le t) = \mathbb{P}(e^{-\lambda V} \ge e^{-\lambda t}) \le e^{\lambda t}\mathbb{E}\, e^{-\lambda V} \le C e^{\lambda t}\lambda^{-\alpha}$$

Minimizing the right hand side by taking $\lambda = \alpha/t$ for $t > 0$ small, we obtain under the assumption $\mathbb{E}\, e^{-\lambda V} \le C/\lambda^\alpha$,

$$\mathbb{P}(V \le t) \le C(e/\alpha)^\alpha t^\alpha$$

which provides the correct rate but not the best constant.
•The one-sided relation

$$\mathbb{P}(V \le t) \le C_1 t^\alpha \quad \text{for some constant } C_1 > 0 \text{ and all } t > 0$$

is equivalent to

$$\mathbb{E}\, e^{-\lambda V} \le C_2 \lambda^{-\alpha} \quad \text{for some constant } C_2 > 0 \text{ and all } \lambda > 0.$$

Exponential Rate

Thm: For $\alpha > 0$ and $\beta \in \mathbb{R}$, or $\alpha = 0$ and $\beta > 0$

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

if and only if

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

as $\lambda \to \infty$.

•A slightly more general formulation of the above result is given in Theorem 4.12.9 of Bingham, Goldie and Teugels (1987), and is called de Bruijn's exponential Tauberian theorem, see also Theorem 3 in Kasahara (1978).

•One direction between the two quantities is easy and follows from the simple connection:

$$\mathbb{P}(V \leq t) = \mathbb{P}(\exp(-\lambda V) \geq \exp(-\lambda t)) \leq \exp(\lambda t)\mathbb{E}\exp(-\lambda V), \tag{0.2}$$

which is just exponential Chebyshev's inequality.

We first consider the case $\alpha > 0$ and $\beta \in \mathbb{R}$ and there are four directions to deal with.
(i). If $\log \mathbb{E}\, e^{-\lambda V} \leq -C\lambda^{\alpha/(1+\alpha)}(\log \lambda)^{\beta/(1+\alpha)}$, then from (0.2),

$$\log \mathbb{P}(V \leq t) \leq \lambda t - C\lambda^{\alpha/(1+\alpha)}(\log \lambda)^{\beta/(1+\alpha)}$$

and thus

$$\limsup_{t \to 0} t^{\alpha} |\log t|^{-\beta} \log \mathbb{P}(V \leq t) \leq -C(C\alpha)^{\alpha}(1+\alpha)^{\beta-1-\alpha}$$

by taking $\lambda = (C\alpha)^{1+\alpha}(1+\alpha)^{\beta-1-\alpha} t^{-(1+\alpha)} |\log t|^{\beta}$. Note that we found λ by approximately minimizing the upper bound, i.e, setting the derivative equals zero and then solving the equation with only dominating terms.

(ii). If $\log \mathbb{P}(V \le t) \ge -Ct^{-\alpha}|\log t|^{\beta}$, then from (0.2)

$$\log \mathbb{E}\, e^{-\lambda V} \ge -\lambda t - Ct^{-\alpha}|\log t|^{\beta}$$

and thus

$$\begin{aligned} & \liminf_{\lambda\to\infty} \lambda^{-\alpha/(1+\alpha)}(\log \lambda)^{-\beta/(1+\alpha)} \log \mathbb{E}\, e^{-\lambda V} \\ \ge \quad & -(1+\alpha)^{1-\beta/(1+\alpha)}\alpha^{-\alpha/(1+\alpha)}C^{1/(1+\alpha)} \end{aligned}$$

by taking $t^{1+\alpha} = C\alpha\lambda^{-1}(\log \lambda/(1+\alpha))^{\beta}$.

(ii). If $\log \mathbb{P}(V \le t) \ge -Ct^{-\alpha}|\log t|^{\beta}$, then from (0.2)

$$\log \mathbb{E}\, e^{-\lambda V} \ge -\lambda t - Ct^{-\alpha}|\log t|^{\beta}$$

and thus

$$\begin{aligned} & \liminf_{\lambda\to\infty} \lambda^{-\alpha/(1+\alpha)}(\log\lambda)^{-\beta/(1+\alpha)} \log \mathbb{E}\, e^{-\lambda V} \\ \ge \quad & -(1+\alpha)^{1-\beta/(1+\alpha)}\alpha^{-\alpha/(1+\alpha)}C^{1/(1+\alpha)} \end{aligned}$$

by taking $t^{1+\alpha} = C\alpha\lambda^{-1}(\log\lambda/(1+\alpha))^{\beta}$.
(iii). If $\log \mathbb{P}(V \le t) \le -Ct^{-\alpha}|\log t|^{\beta}$ for $t \le \delta$, then from (0.1)

$$\log \mathbb{E}\, e^{-\lambda V} \le \int_0^{\delta} \lambda e^{-\lambda t} e^{-Ct^{-\alpha}|\log t|^{\beta}}\, dt + e^{-\lambda\delta}$$

The rest follows from asymptotic analysis (Laplace method).

(ii). If $\log \mathbb{P}(V \le t) \ge -Ct^{-\alpha}|\log t|^{\beta}$, then from (0.2)

$$\log \mathbb{E}\, e^{-\lambda V} \ge -\lambda t - Ct^{-\alpha}|\log t|^{\beta}$$

and thus

$$\begin{aligned} & \liminf_{\lambda\to\infty} \lambda^{-\alpha/(1+\alpha)}(\log\lambda)^{-\beta/(1+\alpha)} \log \mathbb{E}\, e^{-\lambda V} \\ \ge \quad & -(1+\alpha)^{1-\beta/(1+\alpha)} \alpha^{-\alpha/(1+\alpha)} C^{1/(1+\alpha)} \end{aligned}$$

by taking $t^{1+\alpha} = C\alpha\lambda^{-1}(\log\lambda/(1+\alpha))^{\beta}$.
(iii). If $\log \mathbb{P}(V \le t) \le -Ct^{-\alpha}|\log t|^{\beta}$ for $t \le \delta$, then from (0.1)

$$\log \mathbb{E}\, e^{-\lambda V} \le \int_0^{\delta} \lambda e^{-\lambda t} e^{-Ct^{-\alpha}|\log t|^{\beta}}\, dt + e^{-\lambda\delta}$$

The rest follows from asymptotic analysis (Laplace method).
(iv). Analytic argument (approximation approach), see Korevaar (2004). Any there any direct (one-sided) probabilistic argument?

(ii). If $\log \mathbb{P}(V \le t) \ge -Ct^{-\alpha}|\log t|^{\beta}$, then from (0.2)

$$\log \mathbb{E}\, e^{-\lambda V} \ge -\lambda t - Ct^{-\alpha}|\log t|^{\beta}$$

and thus

$$\begin{aligned} & \liminf_{\lambda\to\infty} \lambda^{-\alpha/(1+\alpha)}(\log\lambda)^{-\beta/(1+\alpha)} \log \mathbb{E}\, e^{-\lambda V} \\ \ge \quad & -(1+\alpha)^{1-\beta/(1+\alpha)} \alpha^{-\alpha/(1+\alpha)} C^{1/(1+\alpha)} \end{aligned}$$

by taking $t^{1+\alpha} = C\alpha\lambda^{-1}(\log\lambda/(1+\alpha))^{\beta}$.
(iii). If $\log \mathbb{P}(V \le t) \le -Ct^{-\alpha}|\log t|^{\beta}$ for $t \le \delta$, then from (0.1)

$$\log \mathbb{E}\, e^{-\lambda V} \le \int_0^{\delta} \lambda e^{-\lambda t} e^{-Ct^{-\alpha}|\log t|^{\beta}}\, dt + e^{-\lambda\delta}$$

The rest follows from asymptotic analysis (Laplace method).
(iv). Analytic argument (approximation approach), see Korevaar (2004). Any there any direct (one-sided) probabilistic argument?
•Next we present an instructive probabilistic argument for (iv) in the case $\alpha = 0$ and $\beta > 0$.

Next we present an instructive probabilistic argument in the case $\alpha = 0$ and $\beta > 0$. It is based on the simple observation that

$$\begin{aligned} \mathbb{E}\,\exp(-\lambda V) &= \mathbb{E}\,\exp(-\lambda V)\mathbb{I}_{V\le t} + \mathbb{E}\,\exp(-\lambda V)\mathbb{I}_{V>t} \\ &\le \mathbb{P}(V\le t) + \exp(-t\lambda). \end{aligned} \qquad (0.3)$$

(iii)'. If $\log \mathbb{P}(V \le t) \le -C|\log t|^\beta$, then from (0.3)

$$\log \mathbb{E}\, e^{-\lambda V} \le \log\left(e^{-C|\log t|^\beta} + e^{-t\lambda}\right) \sim -C|\log\lambda|^\beta$$

by taking $t = C\lambda^{-1}(\log\lambda)^\beta$ which was picked to make the two terms approximate equal.
(iv)'. If $\log \mathbb{E}\, e^{-\lambda V} \ge -C(\log\lambda)^\beta$, then from (0.3),

$$\mathbb{P}(V\le t) \ge e^{-C(\log\lambda)^\beta} - e^{-t\lambda} = e^{-C|\log t|^\beta} - e^{-2C|\log t|^\beta} \sim e^{-C|\log t|^\beta}$$

by taking $\lambda = 2Ct^{-1}|\log t|^\beta$. Note that we have a wide range of chooses for λ. $\square$

•When the argument in (iii)' and (iv)' is directly applied to the case $\alpha > 0$, we can obtain the correct rate but not the constant.

•In summary, three useful connections between $\mathbb{P}(V \le t)$ and $\mathbb{E}\, e^{-\lambda V}$ are (0.1), (0.2) and (0.3).
Our first application is for sums of independent random variables, and it is an easy consequence of the Tauberian theorem.

Corollary

If V_i, $1 \le i \le m$, are independent nonnegative random variables such that

$$\lim_{t\to 0} t^{\alpha} \log \mathbb{P}\left(V_i \le t\right) = -d_i, \quad 1 \le i \le m,$$

for $0 < \alpha < \infty$, then

$$\lim_{t\to 0} t^{\alpha} \log \mathbb{P}\left(\sum_{i=1}^{m} V_i \le t\right) = -\left(\sum_{i=1}^{m} d_i^{1/(1+\alpha)}\right)^{1+\alpha}.$$

Proof: The result follows from $\mathbb{E}\, e^{-\lambda \sum_{i=1}^{m} V_i} = \prod_{i=1}^{m} \mathbb{E}\, e^{-\lambda V_i}$.
•We will present direct probabilistic arguments in the next lecture.

Similarly, we have the following in the case of polynomial rates.

Corollary

If V_i, $1 \leq i \leq m$, are independent nonnegative random variables such that

$$\mathbb{P}(V_i \leq t) \sim c_i t^{\alpha_i}$$

as $t \to 0$. Then as $t \to 0$

$$\mathbb{P}(\sum_{i=1}^{m} V_i \leq t) \sim \left(\prod_{i=1}^{m} c_i\right) \cdot \frac{\Gamma(1+\alpha_1)\cdots\Gamma(1+\alpha_m)}{\Gamma(1+\alpha_1+\cdots+\alpha_m)} t^{\alpha_1+\cdots+\alpha_m}$$

A typical application: m-th Integrated BM

Let $X_0(t) = W(t)$ and

$$X_m(t) = \int_0^t X_{m-1}(s)ds, \quad t \geq 0, \quad m \geq 1,$$

which is the m'th integrated Brownian motion or the m-fold primitive. Note that using integration by parts we also have the representation

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

The exact Laplace transform $\mathbb{E} \exp\left(-\lambda \int_0^1 X_m^2(t)dt\right)$ is computed in Chen and Li (2003) and one can find from the exact Laplace transform, for each integer $m \geq 0$,

$$\lim_{\lambda \to \infty} \lambda^{-1/(2m+2)} \log \mathbb{E} \exp\left\{-\lambda \int_0^1 X_m^2(t)dt\right\} = -C_m.$$

where $C_m = 2^{-(2m+1)/(2m+2)} \left(\sin \frac{\pi}{2m+2}\right)^{-1}$.

Then by the Tauberian theorem, We have

$$\begin{aligned} & \log \mathbb{P}\left(\int_0^1 X_m^2(t)dt \le \varepsilon^2\right) \\ \sim\ & 2^{-1}(2m+1)\left((2m+2)\sin\frac{\pi}{2m+2}\right)^{-(2m+2)/(2m+1)} \varepsilon^{-2/(2m+1)}. \end{aligned}$$

•We will use this result in lecture 7.

Feynman Path Integrals

The use of Feynman path integrals to calculate Green functions for certain elementary potentials has received a lot of attention by physicists, see Inomata (1988) and its references. Here we briefly indicate the basic ideas in the calculation of three central forces: Harmonic oscillator, inverse square and Coulomb potential (or hydrogen atom).

Consider the Schrödinger equation

$$\begin{cases} \frac{\partial u}{\partial t} = \frac{\sigma^2}{2}\Delta u - V(x)u & 0 \le t \le 1 \\ u(0,x) = f(x) \end{cases}$$

where $\sigma^2 > 0$, $\Delta = \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_1^2}$ is the Laplace operator and $V : \mathbb{R}^n \to \mathcal{C}$ is a potential. According to the Feynman-Kac formula, the solution of (**??**) can be written in a probabilistic form:

$$u(t,x,f) = \mathbb{E}_x(f(B_t^\sigma)\exp\{-\int_0^t V(B_s^\sigma)ds\})$$

where B_t^σ is the BM with variance σ^2 and $\mathbb{E}_x$ is the expectation with respect to the measure associated with B^σ starting at x.

The exact formulas and asymptotics as $t \to \infty$ can be computed explicitly for

$$u(t,0) = \mathbb{E}\,(\exp\{-\int_0^t V(B_s)ds\})$$

in the case of the following potentials.
(i). Harmonic oscillators, $V(x) = \lambda|x|^2$;
(ii). Inverse square, $V(x) = \lambda/|x|^2$;
(iii). Coulomb potential or hydrogen atom, $V(x) = \lambda/|x|$
where α is a constant and $|x|^2 = x_1^2 + \cdots + x_n^2$, $x \in \mathbb{R}^n$.
•See Pitman and Yor (1981, 1982), Inomata (1988), Hu and Meyer (1988), Hu (1989).

As an example, we consider a general case for the harmonic oscillator and follow the computation of Revuz and Yor (1991, p413-414).

Thm: We have

$$\mathbb{E}_x \exp\{-\int_0^a |B_t|^2 d\nu(dt)\} = (\psi(a)^n \cdot \exp(-\frac{1}{2}\psi'(0)x)$$

where $\phi(t)$ is the unique positive and non-increasing solution of the equation

$$\phi''(t) = -\nu(t)\psi(t), \quad 0 \le t \le a, \quad \psi(0) = 1.$$

In particular,

$$\mathbb{E}_x \exp\{-\frac{\lambda}{2}\int_0^a |B_t|^2 dt\} = (\coth(\lambda a))^{-n/2} \exp\{-\frac{\lambda x}{2\coth(\lambda a)}\}$$

and

$$\log \mathbb{P}(\int_0^a |B_t|^2 dt \le \varepsilon^2) \sim -\frac{n}{8a^2}\varepsilon^{-2}$$

Donsker-Varadhan Theory (1975-9)

In one-dimensional case, the large deviation theory of Donsker-Varadhan can also be used for certain family of potentials. Let $V(x) \geq 0$ be continuous on $\mathbb{R}$. If $V(x) \to +\infty$ as $|x| \to \infty$, then the eigenvalue problem

$$\frac{1}{2}\psi''(x) - V(x)\psi(x) = -\lambda\psi(x) \tag{0.4}$$

has a discrete spectrum, and an old result of Kac (1950) is that, for the least eigenvalue λ_1,

$$\lim_{t\to\infty} \frac{1}{t} \log \mathbb{E}_x \exp\{-\int_0^t V(B_s)ds\} = -\lambda_1 \tag{0.5}$$

where $\mathbb{E}_x$ denotes expectation on the standard Brownian motion starting B_s staring at x.

Now from a pure analytic point of view, the least eigenvalue λ_1 of (0.4) has a variation representation formula

$$\lambda_1 = \inf\left\{\int_{-\infty}^{\infty} V(x)\psi^2(y)dy + \frac{1}{2}\int_{-\infty}^{\infty}(\psi'(x))^2 dx\right\}. \qquad (0.6)$$

where the infimum is taken over all $\psi \in L_2$ such that $\int_{-\infty}^{\infty}\psi^2(x) = 1$. Hence it is reasonable to expect that an expression like the right hand of (0.6) should come from a direct asymptotic evaluation of $\mathbb{E}_x \exp\{-\int_0^t V(B_s)ds$. More important, such a direct connection will allow us to deal with more general Brownian functionals, including those for which there is no associated differential equation at all. This was open for a long time and was solved by Donsker-Varadhan in a series of remarkable papers in late 70's. They developed a whole large deviation theory for the local times of Markov processes.

The L_p-norm for BM

Thm: For any $1 \le p \le \infty$

$$\lim_{\varepsilon \to 0} \varepsilon^2 \log \mathbb{P}\left(\|W(t)\|_p \le \varepsilon \right) = -\kappa_p$$

where

$$\kappa_p = 2^{2/p} p \left(\lambda_1(p)/(2+p)\right)^{(2+p)/p}$$

and

$$\lambda_1(p) = \inf\left\{ \int_{-\infty}^{\infty} |x|^p \phi^2(x) dx + \frac{1}{2} \int_{-\infty}^{\infty} \left(\phi'(x)\right)^2 dx \right\} > 0,$$

the infimum is taken over all $\phi \in L_2(-\infty, \infty)$ such that $\int_{-\infty}^{\infty} \phi^2(x) dx = 1$.

•The case $p = 2$ and $p = \infty$ with $\kappa_2 = 1/8$ and $\kappa_\infty = \pi^2/8$ are well known and the exact distributions in terms of infinite series are known.

From asymptotic point of view for the Laplace transform, Kac (1951) showed by Feynman-Kac formula and eigenfunction expansion that

$$\lim_{t\to\infty} \frac{1}{t} \log \mathbb{E} \exp\left\{-\int_0^t |W(s)|^p ds\right\} = -\lambda_1(p)$$

and $\lambda_1(p)$ is the smallest eigenvalue of the operator

$$Af = -\frac{1}{2}f''(x) + |x|^p f(x)$$

on $L_2(-\infty, \infty)$. Thus we obtain $\lambda_1(p)$ from the classical variation expression for eigenvalues.

•The theorem is first formulated explicitly this way as an lemma in Li (2000). It follows from Kac's result which is by Brownian scaling

$$\lim_{\lambda\to\infty} \lambda^{-2/(2+p)} \log \mathbb{E} \exp\left\{-\lambda \int_0^1 |W(s)|^p ds\right\} = -\lambda_1(p),$$

and the exponential Tauberian theorem with $\alpha = 2/p$.

Some Equivalent Conditions

Next we try to formulate some equivalent conditions for small values probabilities. There are two types to consider here. One is the exponential rate $\mathbb{P}(V \leq t) \leq C \exp(-ct^{-\alpha})$ and the other is the polynomial rate $\mathbb{P}(V \leq t) \leq Ct^{-\alpha}$. We only deal with the exponential rate here. The polynomial case is left as an exercise.

Some Equivalent Conditions

Thm:

The following conditions are equivalent:

(1). There exist $C, c > 0$ such that for any $t > 0$, $\mathbb{P}(V \leq t) \leq Ce^{-ct^{-\alpha}}$ (small value estimate);

(2). There exist $B, b > 0$ such that for any $\lambda > 0$, $\mathbb{E}\, e^{-\lambda V} \leq Be^{-b\lambda^{\alpha/(1+\alpha)}}$ (Laplace transform condition).

(2)$'$. For any fixed $\beta > 0$, there exist $B, b > 0$ such that for any $\lambda > 0$, $\mathbb{E}\, e^{-\lambda V^{\beta}} \leq Be^{-b\lambda^{\alpha/(\beta+\alpha)}}$ (Laplace transform condition).

(2)$''$. There exist $B, b > 0$ such that for any $\lambda > 0$, $\mathbb{E}\, e^{-\lambda V^{\alpha}} \leq Be^{-b\lambda^{1/2}}$ (Laplace transform condition).

(3). There exists $a > 0$ such that $\mathbb{E}\, e^{aV^{-\alpha}} < +\infty$ (inverse/negative ψ_{α}-condition).

(4). For any fixed $0 < \gamma < \alpha$, there exists $d > 0$ such that for any $\lambda > 0$, $\mathbb{E}\, e^{\lambda V^{-\gamma}} \leq De^{d\lambda^{\alpha/(\alpha-\beta)}}$ (Exponential moment condition).

(5). There exist $C' > 0$ such that for any $p > 1$, $\|V^{-1}\|_p = (\mathbb{E}\, V^{-p})^{1/p} \leq C'p^{1/\alpha}$, (inverse/negative moment condition).

Proof: (1) $\Rightarrow$ (2)$'$: Note that $\mathbb{E}\, e^{-\lambda V^\beta} \le \mathbb{P}(V^\beta \le t) + e^{-\lambda t} \le e^{-ct^{-\alpha/\beta}} + e^{-\lambda t}$. Taking $\lambda t = ct^{-\alpha/\beta}$ finishes the argument.
(2)$'$ $\Rightarrow$ (1):

$$\begin{aligned}
\mathbb{P}(V \le t) &= \mathbb{P}(e^{-\lambda V^\beta} \ge e^{-\lambda t^\beta}) \\
&\le e^{\lambda t^\beta} \mathbb{E}\, e^{-\lambda V^\beta} \le e^{\lambda t^\beta - b\lambda^{\alpha/(\beta+\alpha)}} \le e^{-c' t^{-\alpha}}
\end{aligned}$$

by taking $\lambda t^\beta = (b/2)\lambda^{\alpha/(\beta+\alpha)}$.
(1) $\Rightarrow$ (3): Take $a < c$,

$$\begin{aligned}
\mathbb{E}\, e^{aV^{-\alpha}} &= \int_0^\infty a\alpha x^{\alpha-1} e^{ax^\alpha} \mathbb{P}(V^{-1} > x) dx \\
&\le \int_0^\infty a\alpha x^{\alpha-1} \cdot Ce^{-(c-a)x^\alpha} dx < +\infty.
\end{aligned}$$

(3) $\Rightarrow$ (2):

$$\mathbb{E}\, e^{-\lambda V} = \mathbb{E}\, e^{-\lambda V - aV^{-\alpha}} \cdot e^{aV^{-\alpha}} \le \sup_{v>0} e^{-\lambda v - av^{-\alpha}} \cdot \mathbb{E}\, e^{aV^{-\alpha}}.$$

(3) $\Rightarrow$ (4):

$$\mathbb{E}\, e^{\lambda V^{-\gamma}} = \mathbb{E}\, e^{\lambda V^{-\gamma} - aV^{-\alpha}} \cdot e^{aV^{-\alpha}} \leq \sup_{u>0} e^{\lambda u^{\gamma} - au^{\alpha}} \cdot \mathbb{E}\, e^{aV^{-\alpha}}.$$

(4) $\Rightarrow$ (5):

$$\begin{aligned} \mathbb{E}\, V^{-p} &= \mathbb{E}\, V^{-p} e^{-\lambda V^{-\gamma}} \cdot e^{\lambda V^{-\gamma}} \\ &\leq \sup_{u>0} u^{p/\gamma} e^{-\lambda u} \cdot \mathbb{E}\, e^{\lambda V^{-\gamma}} \leq (e\gamma^{-1} p \lambda^{-1})^{p/\gamma} \cdot D e^{d\lambda^{\alpha/(\alpha-\gamma)}}. \end{aligned}$$

Taking $\lambda^{\alpha/(\alpha-\gamma)} = p$ finishes the argument.
(5) $\Rightarrow$ (1): Note that

$$\mathbb{P}(V \leq t) = \mathbb{P}(V^{-p} \geq t^{-p}) \leq t^{p} \mathbb{E}\, V^{-p} \leq (C' t p^{1/\alpha})^{p}.$$

And we can take p such that $C' t p^{1/\alpha} = e^{-1}$.

Corollary

If V_i, $1 \leq i \leq m$, are independent nonnegative random variables such that

$$\mathbb{P}\left(V_i \leq t\right) \leq C_i e^{-c_i t^{-\alpha}}, \quad 1 \leq i \leq m,$$

for $0 < \alpha < \infty$, then

$$\mathbb{P}\left(\sum_{i=1}^{m} V_i \leq t\right) \leq -B_m e^{-b_m t^{-\alpha}}$$

- The result follows from $\mathbb{E}\, e^{-\lambda \sum_{i=1}^m V_i} = \prod_{i=1}^m \mathbb{E}\, e^{-\lambda V_i}$.
- We will present direct arguments in the next Lecture.

Moment Inequalities

The problem of providing bounds on the probability that a certain random variable belongs to a given set, given information on some of its moments, has a very rich history and many applications. Here we will focus on small value type.

•The weakest but most general estimate is the straightforward Chebyshev's inequality: For any $0 \leq t < \mathbb{E}\, V = m_1$,

$$\mathbb{P}(V \leq t) = \mathbb{P}(m_1 - V \geq m_1 - t) \leq \mathbb{P}((m_1 - V)^2 \geq (m_1 - t)^2) \leq \frac{\mathrm{Var}(V)}{(m_1 - t)^2}$$

Note the importance of the assumption $t < m_1$ and also that the right hand side can be bigger than 1 which provides a trivial bound.

•Below we present a refinement and show the technique of using a shifted Chebyshev's inequality, by introducing a parameter and picking the best one at the end, is a very useful trick.

Lemma: For any $0 \le t < \mathbb{E}\, V$,

$$\mathbb{P}(V \le t) \le \frac{\mathrm{Var}(V)}{\mathbb{E}\,(V-t)^2} = \frac{\mathrm{Var}(V)}{\mathrm{Var}(V) + (m_1 - t)^2} \tag{0.7}$$

In particular, $\mathbb{P}(V = 0) \le \frac{\mathrm{Var}(V)}{\mathbb{E}\, V^2} = \frac{m_2 - m_1^2}{m_2}$.
•The bound is sharp in the sense that given the first two moments, $m_1 = \mathbb{E}\, V$ and $m_2 = \mathbb{E}\, V^2$, one can construct a random variable with these moments for which the inequality is an equality. T
•By rewrite the lemma, we have

$$\mathbb{P}(V \ge t) \ge \frac{(\mathbb{E}\, V - t)^2}{\mathbb{E}\,(V-t)^2}$$

for $0 \le t < m_1$. This provides a minor improvement of the well-known Paley-Zygmund lower bound: For $0 \le t < \mathbb{E}\, V$,

$$\mathbb{P}(V \ge t) \ge \frac{(\mathbb{E}\, V - t)^2}{\mathbb{E}\, V^2}$$

which is often applied with $t = 0$ or $t = \lambda \mathbb{E}\, V$, $0 < \lambda < 1$. It is also of interests to note that the proof of Paley-Zygmund is base on Cauchy-Schwarz inequality applied to $V \mathbb{I}_{\{V > t\}}$.

Proof: We use the method of Chebyshev's inequality with a shift parameter. For any $\lambda > t$, we have by Chebyshev's inequality

$$\mathbb{P}(V \leq t) = \mathbb{P}(\lambda - V \geq \lambda - t) \leq \frac{\mathbb{E}\,(\lambda - V)^2}{(\lambda - t)^2} = \frac{\lambda^2 - 2\lambda m_1 + m_2}{(\lambda - t)^2}.$$

Hence

$$\mathbb{P}(V \leq t) \leq \inf_{\lambda > t} \frac{\lambda^2 - 2\lambda m_1 + m_2}{(\lambda - t)^2} = \frac{m_2 - m_1^2}{m_2 - m_1^2 + (m_1 - t)^2}.$$

where the infimum is archived at

$$\lambda = \lambda_0 = \frac{m_2 - m_1 t}{m_1 - t} > t \quad \text{for } t < m_1$$

by simple calculus. This finishes the proof. $\square$

Convex Optimization Approach

Given the first k moments, $m_0 = 1$, $m_1, \cdots, m_k$ of a random variable X with domain $D \subset \mathbb{R}$, general bounds for $\mathbb{P}(X \in S)$ based on Chebyshev's inequality can be obtained from

$$\mathbb{P}(X \in S) \leq \mathbb{P}(g(X) \geq 1) \leq \mathbb{E}\, g(X) = \sum_{j=0}^{k} y_j m_j \tag{0.8}$$

where $g(x) = \sum_{j=0}^{k} y_j x^j$, $g(x) \geq 0$ for any $x \in D$ and $g(x) \geq 1$ for any $x \in S$. The best bounds based on this approach is thus

$$\mathbb{P}(X \in S) \leq \min_{y_0, \cdots, y_k} \left(\sum_{j=0}^{k} y_j m_j \right) \tag{0.9}$$

subject to constrains

$$\sum_{j=0}^{k} y_j x^j \geq 0 \quad \forall x \in D, \quad \text{and} \quad \sum_{j=0}^{k} y_j x^j \geq 1 \quad \forall x \in S.$$

In small value estimates, we have $D = \mathbb{R}^+ = [0, \infty)$ and $S = (0, t)$, $0 \le t \le m_1$. It can be checked that (0.7) for $\mathbb{P}(V \le t)$ can be found this way. Moreover, the following bound based on the first three moments is carried out by this approach in Popescu (1999), see also Bertsimas and Popescu (2005) for an in depth discussion and history, including multivariate settings.

Thm: For $0 \le t \le m_1$,

$$
\begin{aligned}
\mathbb{P}(V \le t) &\le 1 - \frac{(m_2 - tm_1)^3}{(m_3 - tm_2)(m_3 - 2tm_2 + t^2 m_1)} \\
&= 1 - \frac{(\mathbb{E}\, V(V-t))^3}{\mathbb{E}\, V^2(V-t) \cdot \mathbb{E}\, V(V-t)^2}
\end{aligned}
$$

and the bound is tight.

•Similar method has been used in Li and Liu (2009) for truncated variance.

The Moments/Probabilistic Method in Combinatorics

In a typical probabilistic proof of the existence of certain combinatorial result, the basic idea is to define a proper probability distribution on a class of discrete objects and then to show that the probability of a certain event is positive. However, many of these proofs actually give more and show that the probability of the event considered is not only positive but is large. In fact, most probabilistic proofs deal with events that hold with high probability, that is, a probability that tends to one as the size of the problem grow. This can be seen from examples mentioned below. On the other hand, many refined results require one to show that a certain event holds with positive, though very small, probability. Small value type problems appear often in this setting.

•N. Alon and J. Spencer (2000). The probabilistic method.
•Bollobas (2001), Random Graphs.
•J. Spencer, Random structure and Erdős magic, 2004

Levy Concentration Function

Let X be a random vector in $\mathbb{R}^n$. The Levy concentration function of X is defined as

$$L(X, \varepsilon) = \sup_{u \in \mathbb{R}^n} \mathbb{P}(\|X - u\|_2 \leq \varepsilon)$$

A simple but rather weak bound on Levy concentration function follows from Paley-Zygmund inequality.

Lemma: Let X be a random variable with unit variance and with finite fourth moment, and put $M_4 := \mathbb{E}\,(X - \mathbb{E}\,X)^4$. Then for every $\varepsilon \in (0, 1)$, there exists $p = p(M_4, \varepsilon) \in (0, 1)$ such that $L(X, \varepsilon) \leq p$.

•There has been significant interests recently in bounding Levy concentration function for sums of independent random variables; see Rudelson and Vershynin,(2008, 2009, 2010), Vershynin (2011+), for discussion for sums, and tensorization to transfer bounds from random variables to random vectors.

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 $\mathbb{P}(|\hat{\theta}-\theta| \leq \varepsilon)$ for a given $\varepsilon > 0$ small.

•This example is based on the poster "Averaging and Edgewood Expansion". For a detailed study, see S. Stigler (1980), An Edgeworth curiosum. *Ann. Stat*, 8, 931–934.

A Conjecture

For any unit vector $u = (u_1, \cdots, u_n) \in \mathbb{S}^{n-1}$, i.e. $\sum_{j=1}^n u_j^2 = 1$,

$$\mathbb{P}(|\sum_{j=1}^{n} \varepsilon_j u_j| \leq 1) \geq \frac{1}{2}.$$

•An equivalent geometric interpretation of the problem is that any slab $S = \{x \in \mathbb{R}^n : |\langle x, u\rangle| \leq 1\}$ with normal vector u contains at least half of the vertices of the discrete cube $\{-1, 1\}^n \subset \mathbb{R}^n$. The best known bound is $3/8$, I think.

•It is not hard to show a upper bound like

$$\mathbb{P}(|\sum_{j=1}^{n} \varepsilon_j u_j| \leq 1/2) \leq \frac{32}{41}.$$

•Find a better upper bound.

Exercise

Prop: For any positive random variable $V \geq 0$ with non-increasing density, there exist a universal constant $c > 0$ such that

$$\mathbb{P}(V \leq \varepsilon) \geq c(\mathbb{E}\, V^2)^{-1/2} \cdot \varepsilon$$

for all $0 \leq \varepsilon \leq \sqrt{\mathbb{E}\, V^2}$.
$\bullet$The condition implies that for any $\delta > 0, 0 < s \leq t$,

$$\mathbb{P}(s \leq V \leq s + \delta) \geq \mathbb{P}(t < V \leq t + \delta).$$

In fact, it is equivalent to $\pm V$ is symmetric and unimodal.

A Symmetrization Inequality

For any i.i.d X and Y,

$$\mathbb{P}(|X+Y| \leq 1) \leq 3\mathbb{P}(|X-Y| \leq 1)$$

where the constant 3 can be replaced by 2, the best possible.