Probability Theory
Problem Sets w/ Solutions

Rex C. Liu, Courant Institute of Mathematical Sciences

October 19, 2023

Disclaimer:

These are the problem sets for the course *Probability Theory I (MATH.GA 2911)*[1], given by professor Paul Bourgade at New York University in Fall 2022 and the course *Probability Theory II (MATH.GA 2912)*[2], given by professor Nina Holden at New York University in Spring 2023.

The solutions are given by Rex Liu with help from Xiang Fang (UCSB PSTAT), my classmates Andrew Zhang, Asher Miao, Michael Sun, Michael Ye, and references from [Bil86], [GS20], [Wal12]. If you see any mistakes or think that the presentation is unclear and could be improved, please send an email to: cl5682@nyu.edu. All comments and suggestions are appreciated.

[1] This course is aimed primarily for PhD students. Topics include laws of large numbers, weak convergence, central limit theorems, conditional expectation, martingales and Markov chains. The reference text is [Var01].

[2] This course is aimed primarily for PhD students. Topics include: Stochastic processes in continuous time. Brownian motion. Poisson process. Processes with independent increments. Stationary processes. Semi-martingales. Markov processes and the associated semi-groups. Connections with PDEs. Stochastic differential equations. Convergence of processes. There will be no official textbook. Some useful books are [Bas11], [Var07], [KS07], [KS98].

Contents

6. (*law of the iterated logarithm*) The goal of this problem is to prove the iterated logarithm law, first for Gaussian random variables. In other words, for $X_1, X_2 \dots$ i.i.d. standard Gaussian random variables, denoting $S_n = X_1 + \dots + X_n$, we have

$$\mathbb{P}\left(\limsup_{n\to\infty} \frac{S_n}{\sqrt{2n\log\log n}} = 1\right) = 1 \tag{1.6}$$

(a) Prove that

$$\mathbb{P}(X_1 > \lambda) \underset{\lambda\to\infty}{\sim} \frac{1}{\lambda\sqrt{2\pi}} e^{-\frac{\lambda^2}{2}}$$

Proof. Let $x = \lambda + t/\lambda$. Following the change of variable, for every positive λ, one has

$$\mathbb{P}(X > \lambda) = \frac{1}{\sqrt{2\pi}} \int_\lambda^{+\infty} \mathrm{e}^{-x^2/2} \mathrm{d}x = \frac{1}{\lambda\sqrt{2\pi}} \mathrm{e}^{-\lambda^2/2} \int_0^{+\infty} \mathrm{e}^{-t} \mathrm{e}^{-t^2/(2\lambda^2)} \mathrm{d}t$$

When $\lambda \to \infty$, $\mathrm{e}^{-t^2/(2\lambda^2)} \to 1$ and hence $\int_0^{+\infty} \mathrm{e}^{-t}\mathrm{e}^{-t^2/(2\lambda^2)}\mathrm{d}t \to 1$ (as the density function is dominated, the limit can be taken out of the integral by DCT), so that

$$\mathbb{P}(X_1 > \lambda) \underset{\lambda\to\infty}{\sim} \frac{1}{\lambda\sqrt{2\pi}} e^{-\frac{\lambda^2}{2}}$$

□

In the following questions we denote $f(n) = \sqrt{2n\log\log n}$, $\lambda > 1$, $c, \alpha > 0$,

$$\begin{aligned}
A_k &= \left\{ S_{\lfloor\lambda^k\rfloor} \geq cf\left(\lambda^k\right)\right\}, \\
C_k &= \left\{ S_{\lfloor\lambda^{k+1}\rfloor} - S_{\lfloor\lambda^k\rfloor} \geq cf\left(\lambda^{k+1} - \lambda^k\right)\right\}, \\
D_k &= \left\{ \sup_{n\in[\![\lambda^k,\lambda^{k+1}]\!]} \frac{S_n - S_{\lfloor\lambda^k\rfloor}}{f(\lambda^k)} \geq \alpha \right\}
\end{aligned}$$

(b) Prove that for any $c > 1$ we have $\sum_{k\geq 1} \mathbb{P}(A_k) < \infty$ and

$$\limsup_{k\to\infty} \frac{S_{\lfloor\lambda^k\rfloor}}{f(\lambda^k)} \leq 1 \text{ a.s.}$$

Proof. Let $c = 1 + \varepsilon$ for some $\varepsilon > 0$. Following part 6a, with the fact that $S_n \sim \mathcal{N}(0, n)$, one has (as $(1+\varepsilon)f(n)/\sqrt{n} \to \infty$)

$$\begin{aligned}
\mathbb{P}(S_n \geq (1+\varepsilon)f(n)) &= \mathbb{P}(S_n/\sqrt{n} \geq (1+\varepsilon)f(n)/\sqrt{n}) \\
&\sim \frac{1}{(1+\varepsilon)\sqrt{2\log\log n}\sqrt{2\pi}} \frac{1}{(\log n)^{(1+\varepsilon)^2}} \leq \frac{1}{(\log n)^{(1+\varepsilon)^2}}
\end{aligned}$$

Replace n with $\lfloor \lambda^k \rfloor$, one has

$$\mathbb{P}(A_k) \leq \frac{1}{(k \log \lambda)^{(1+\varepsilon)^2}}$$

Hence, the series $\sum_{k \geq 1} \mathbb{P}(A_k)$ converges by comparison test to $\sum_n \frac{1}{n^{1+\varepsilon}}$. Then, following Borel–Cantelli, one has

$$\mathbb{P}\left(\limsup_{k\to\infty} \frac{S_{\lfloor \lambda^k \rfloor}}{f\left(\lambda^k\right)} \geq 1+\varepsilon\right) = 0, \quad \forall \varepsilon > 0$$

i.e. $\limsup_{k\to\infty} \frac{S_{\lfloor \lambda^k \rfloor}}{f(\lambda^k)} \leq 1$ a.s.. □

(c) Prove that for any $c < 1$ we have $\sum_{k\geq 1} \mathbb{P}(C_k) = \infty$ and

$$\mathbb{P}(C_k \text{ i.o.}) = 1$$

Proof. Note that C_k is pairwise independent since they are sums of different X_i, where X_i are i.i.d. Assume λ is large enough so that $f\left(\lambda^{k+1} - \lambda^k\right) \in \mathbb{R}$. Now, set $Y_k := S_{\lfloor \lambda^{k+1} \rfloor} - S_{\lfloor \lambda^k \rfloor}$, so that $Y_k \sim \mathcal{N}\left(0, \left\lfloor \lambda^{k+1} \right\rfloor - \left\lfloor \lambda^k \right\rfloor\right)$. For $\varepsilon > 0$, one has

$$\begin{aligned}
\mathbb{P}\left(Y_k \geq c\left(f\left(\lambda^{k+1} - \lambda^k\right)\right)\right) &\sim \mathbb{P}\left(Y_k \geq c\left(f\left(\left\lfloor \lambda^{k+1} \right\rfloor - \left\lfloor \lambda^k \right\rfloor\right)\right)\right) \\
&\sim \frac{\exp\left(-(1-\varepsilon)^2 \log\left(\log\left(\lfloor \lambda^{k+1} \rfloor - \lfloor \lambda^k \rfloor\right)\right)\right)}{2(1-\varepsilon)\sqrt{\pi \log\log\left(\lfloor \lambda^{k+1} \rfloor - \lfloor \lambda^k \rfloor\right)}} \\
&\geq \frac{\alpha(k+1)^{-(1+\varepsilon)^2}}{\sqrt{\log(k+1)}} \\
&\geq \frac{\beta}{(k+1)\log(k+1)}
\end{aligned}$$

So by comparison test to the series $\sum_n \frac{1}{n}$, one has $\sum_{k\geq 1} \mathbb{P}(C_k) = \infty$. Since C_k are mutually independent, by (the second) Borel-Cantelli, $\mathbb{P}(C_k \text{ i.o.}) = 1$. □

(d) Let $\varepsilon > 0$ and choose $c = 1 - \varepsilon/10$. Prove that almost surely the following inequality holds for infinitely many k:

$$\frac{S_{\lfloor \lambda^{k+1} \rfloor}}{f(\lambda^{k+1})} \geq c\frac{f(\lambda^{k+1} - \lambda^k)}{f(\lambda^{k+1})} - (1+\varepsilon)\frac{f(\lambda^k)}{f(\lambda^{k+1})}$$

Proof. As $f(\lambda^{k+1}) > 0$, it suffices to show that

$$S_{\lfloor \lambda^{k+1} \rfloor} \geq cf(\lambda^{k+1} - \lambda^k) - (1+\varepsilon) f(\lambda^k)$$

Following the result of part 6c, one has

$$S_{\lfloor \lambda^{k+1} \rfloor} \geq cf(\lambda^{k+1} - \lambda^k) + S_{\lfloor \lambda^k \rfloor} \quad \text{i.o.}$$

Now it suffices to show that

$$S_{\lfloor \lambda^k \rfloor} \geq -(1+\varepsilon) f(\lambda^k),$$

but this follows from that S_n is symmetric by 0, so that $\liminf_{k\to\infty} \frac{S_{\lfloor \lambda^k \rfloor}}{f(\lambda^k)} \geq -1$ a.s. ☐

(e) By choosing a large enough λ in the previous inequality, prove that almost surely

$$\limsup_{n\to\infty} \frac{S_n}{f(n)} \geq 1$$

Proof. Note that we λ is large enough, one has (since n increases much faster than $\log\log n$)

$$\frac{f(\lambda^{k+1} - \lambda^k)}{f(\lambda^{k+1})} \to \sqrt{\frac{\lambda-1}{\lambda}}, \quad \frac{f(\lambda^k)}{f(\lambda^{k+1})} \to \frac{1}{\sqrt{\lambda}}$$

Thus, when λ is large enough, the above inequality becomes

$$\frac{S_{\lfloor \lambda^{k+1} \rfloor}}{f(\lambda^{k+1})} \geq (1-\gamma)\sqrt{\frac{\lambda-1}{\lambda}} - (1+\varepsilon)\frac{1}{\sqrt{\lambda}}$$

Taking $\gamma \to 0$ and $\lambda \to \infty$, one has

$$\mathbb{P}\left(\limsup_n \frac{S_{\lfloor \lambda^{k+1} \rfloor}}{f(\lambda^{k+1})} \geq (1-\gamma)\sqrt{\frac{\lambda-1}{\lambda}} - (1+\varepsilon)\frac{1}{\sqrt{\lambda}} = 1 \right) = 1$$

☐

(f) Prove that for any $n \in [\![\lambda^k, \lambda^{k+1}]\!]$ and $S_n > 0$ we have

$$\frac{S_n}{f(n)} \leq \frac{S_{\lfloor \lambda^k \rfloor}}{f(\lfloor \lambda^k \rfloor)} + \frac{S_n - S_{\lfloor \lambda^k \rfloor}}{f(\lfloor \lambda^k \rfloor)}$$

Proof. Note that $f(\lfloor \lambda^k \rfloor) \leq f(n)$ since f is monotonely increasing and $n \geq \lfloor \lambda^k \rfloor$, which implies that

$$\begin{aligned} \frac{1}{f(n)} &\leq \frac{1}{f(\lfloor \lambda^k \rfloor)} \\ \implies \frac{S_n}{f(n)} &\leq \frac{S_n}{f(\lfloor \lambda^k \rfloor)} \\ \implies \frac{S_n}{f(n)} &\leq \frac{S_{\lfloor \lambda^k \rfloor}}{f(\lfloor \lambda^k \rfloor)} + \frac{S_n - S_{\lfloor \lambda^k \rfloor}}{f(\lfloor \lambda^k \rfloor)} \end{aligned}$$

□

(g) Prove that

$$\mathbb{P}(D_k) \underset{k\to\infty}{\sim} 2\mathbb{P}\left(X_1 \geq \frac{\alpha f(\lambda^k)}{\sqrt{\lambda^{k+1} - \lambda^k}}\right) \underset{k\to\infty}{\sim} \frac{c}{\sqrt{\log k}} \left(\frac{1}{k}\right)^{\frac{\alpha^2}{\lambda - 1}}$$

Proof. Note that the *reflection principle* of a random walk on $\mathbb{Z}$ gives us

$$\mathbb{P}\left(\max_{1 \leq k \leq n} X_k \geq b\right) = 2\mathbb{P}(X_n \geq b)$$

Plugging in $\mathbb{P}(D_k)$, one has

$$\mathbb{P}\left(\max_{n \in [\![\lambda^k, \lambda^{k+1}]\!]} \frac{S_n - S_{\lfloor \lambda^k \rfloor}}{f(\lambda^k)} \geq \alpha\right) = 2\mathbb{P}\left(\frac{S_{\lfloor \lambda^{k+1} \rfloor} - S_{\lfloor \lambda^k \rfloor}}{f(\lambda^k)} \geq \alpha\right) \tag{1.7}$$

where $S_{\lfloor \lambda^{k+1} \rfloor} - S_{\lfloor \lambda^k \rfloor} \sim \mathcal{N}(0, \lambda^{k+1} - \lambda^k) \sim X_1 \cdot \sqrt{\lambda^{k+1} - \lambda^k}$, so that

$$2\mathbb{P}\left(\frac{S_{\lfloor \lambda^{k+1} \rfloor} - S_{\lfloor \lambda^k \rfloor}}{f(\lambda^k)} \geq \alpha\right) = 2\mathbb{P}\left(X_1 \geq \frac{\alpha f(\lambda^k)}{\sqrt{\lambda^{k+1} - \lambda^k}}\right)$$

Now, following part 6a, one has

$$\begin{aligned} 2\mathbb{P}\left(X_1 \geq \frac{\alpha f(\lambda^k)}{\sqrt{\lambda^{k+1} - \lambda^k}}\right) &= 2\mathbb{P}\left(X_1 \geq \sqrt{\frac{2\lambda^k \log\log \lambda^k}{\lambda^{k+1} - \lambda^k}} \cdot \alpha\right) \\ &\underset{k\to\infty}{\sim} \frac{1}{\sqrt{\frac{\log k + \log\log \lambda}{\lambda - 1}}\alpha\sqrt{\pi}} e^{-\frac{\alpha^2 (\log\log \lambda^k)}{\lambda - 1}} \\ &\underset{k\to\infty}{\sim} c_1 \cdot \frac{1}{\sqrt{\log k}} \cdot \frac{1}{(k \log \lambda)^{\alpha^2/(\lambda-1)}} \\ &\underset{k\to\infty}{\sim} \frac{c}{\sqrt{\log k}} \left(\frac{1}{k}\right)^{\frac{\alpha^2}{\lambda-1}} \end{aligned}$$

for some constants c_1 and c. □

(h) Prove that for $\alpha^2 > \lambda - 1$, almost surely

$$\limsup_{n\to\infty} \frac{S_n}{f(n)} \leq \limsup_{n\to\infty} \frac{S_{\lfloor \lambda^k \rfloor}}{f(\lambda^k)} + \alpha$$

Proof. Following part 6g, if $\alpha^2 > \lambda - 1$, then $\mathbb{P}\left(D_k \text{ i.o.}\right) = 0$ by Borel-Cantelli, so that

$$\mathbb{P}\left\{ \sup_{n\in\left[\lambda^k,\lambda^{k+1}\right]\cap z} \frac{S_n - S_{\lfloor \lambda^k \rfloor}}{f\left(\lambda^k\right)} < \alpha \right\} = 1$$

for k sufficiently large. Now fix n large enough that $n \in \left[\lambda^k, \lambda^{k+1}\right]$ where k is large enough that the above holds. We obtain, provided $S_n > 0$ that the following holds almost surely (for an appropriately fixed n and all sufficiently large k)

$$\frac{S_n}{f(n)} < \frac{S_n}{f\left(\lambda^k\right)} < \alpha + \frac{S_{\lfloor \lambda^k \rfloor}}{f\left(\lambda^k\right)}$$

Let $k \to \infty$. One has

$$\frac{S_n}{f(n)} \leq \alpha + \limsup_k \frac{S_{\lfloor \lambda^k \rfloor}}{f\left(\lambda^k\right)} \quad \text{a.s.}$$

The result in part 6e implies there is a subsequence $\left(n_j\right)_{j\geq 1}$ such that $S_{n_j} > 0$ for each j, and since the above holds for every n_j (and moreover, for any $S_n > 0$), we may let $n \to \infty$, implying that

$$\limsup_n \frac{S_n}{f(n)} \leq \alpha + \limsup_k \frac{S_{\lfloor \lambda^k \rfloor}}{f\left(\lambda^k\right)} \quad \text{a.s.}$$

□

(i) By choosing appropriate λ and α, prove that almost surely

$$\limsup_{n\to\infty} \frac{S_n}{f(n)} \leq 1$$

Proof. Following part 6h and part 6b above, one has

$$\limsup_{n\to\infty} \frac{S_n}{f(n)} \leq \limsup_{n\to\infty} \frac{S_{\lfloor \lambda^k \rfloor}}{f(\lambda^k)} + \alpha \leq 1 + \alpha$$

where $\alpha > \sqrt{\lambda - 1}$, where $\lambda > 1$ can be arbitrarily close to 1. Hence,

$$\limsup_{n\to\infty} \frac{S_n}{f(n)} \leq 1 + \alpha, \quad \forall \alpha > 0$$

leading us to the desired result. □

(j) State a result similar to (1.6) for i.i.d. uniformly bounded random variables. Which steps in the above proof need to be modified to prove this universality result? How?

Solution. Let $\{\phi_n(x)\}$ be a uniformly bounded orthonormal system of realvalued functions on the interval $[0,1]$. Then there exists a subsequence $\{\phi_{n_k}(x)\}$ and a real-valued function $f(x)$, $\int_0^1 f^2(x)dx = 1, 0 \leqq f(x) \leqq B$, where B is the uniform bound of $\{\phi_n(x)\}$, such that for any arbitrary sequence $\{a_k\}$ of real numbers satisfying

$$A_N = \left(a_1^2 + a_2^2 + \cdots + a_N^2\right)^{1/2} \to \infty \text{ as } N \to \infty,$$
$$M_N = o\left(A_N \left(\log\log A_N\right)^{-1/2}\right) \text{ where } M_N = \max_{k \leqq N} |a_k|$$

we have

$$\limsup \frac{S_N(x)}{\left(2A_N^2 \log\log A_N\right)^{1/2}} = f(x) \quad \text{where } S_N(x) = \sum_{k=1}^N a_k \phi_{n_k}(x)$$

□

3.2 Central Limit Theorem

1. Assume $(\Omega, \mathcal{A}, \mathbb{P})$ is such that Ω is countable and $\mathcal{A} = 2^\Omega$. Prove that convergence in probability and convergence almost sure are the same.

Proof. We have proved in class that in general one has convergence a.s. implies convergence in probability (briefly: $0 = \mathbb{P}\{\limsup\{|X_n - X| > \varepsilon\} \geq \limsup \mathbb{P}\{|X_n - X| > \varepsilon\}, \forall \varepsilon > 0)$.

Now we show that convergence in probability implies convergence a.s. when Ω is countable.

Since $\mathcal{A} = 2^\Omega$, the singletons are measureable. Let $\{\omega_n : n \in \mathbb{N}\}$ be the set of elements whose singletons have positive probability. It suffices to show that if $X_n \xrightarrow{\mathbb{P}} X$, then $X_n(\omega_i) \to X(\omega_i)$ for each $i \in \mathbb{N}$.

Fix $i \in \mathbb{N}, \varepsilon > 0$ and assume that $X_n \xrightarrow{\mathbb{P}} X$. Then there is an N s.t. $\mathbb{P}(\{|X_n - X| \geq \varepsilon\}) < \mathbb{P}(\omega_i)$ whenever $n \geq N$. This implies that if $n \geq N$, then $|X_n(\omega_i) - X(\omega_i)| < \varepsilon$. By definition, $X_n \xrightarrow{\text{a.s.}} X$. □

2. Let $(X_i)_{i\geq 1}$ be i.i.d. Gaussian with mean 1 and variance 3. Show that

$$\lim_{n\to\infty} \frac{X_1 + \cdots + X_n}{X_1^2 + \cdots + X_n^2} = \frac{1}{4} \quad \text{a.s.}$$

9. (*Erdős-Kac*) The goal of this exercise is to prove that if $w(m)$ denotes the number of distinct prime factors of m and k is a random variable uniformly distributed on $[\![1, n]\!]$, then the following convergence in distribution holds:

$$\frac{w(k) - \log\log n}{\sqrt{\log\log n}} \xrightarrow[n\to\infty]{} \mathcal{N}(0,1)$$

(a) Prove that if $(X_n)_{n\geq 1}$ converges in distribution to $\mathcal{N}(0,1)$ and $\sup_{n\geq 1}\mathbb{E}[X_n^{2k}] < \infty$ for any $k \in \mathbb{N}$, then

$$\lim_{n\to\infty}\mathbb{E}\left[X_n^k\right] = \mathbb{E}\left[\mathcal{N}(0,1)^k\right]$$

for any $k \in \mathbb{N}$.

Proof. Consider $k = 1$. Since $\sup_n \mathbb{E}[|X_n|^{1+\varepsilon}] < \infty$ where here we have $\varepsilon = 1$, $(X_n)_n$ is uniformly integrable. So that one has

$$\lim_{\alpha\to\infty}\sup_n \int_{|X_n|>\alpha} |X_n| \,\mathrm{d}\mathbb{P} = \lim_{\alpha\to\infty}\sup_n \mathbb{E}[|X_n| 1_{|X_n|>\alpha}] = 0$$

Then, X is integrable and

$$\lim_{n\to\infty}\mathbb{E}[X_n] = \mathbb{E}\left[\lim_{n\to\infty} X_n\right] = \mathbb{E}[\mathcal{N}(0,1)]$$

Comment. *The first equality may not hold since you may not have $\mathbb{E}[\lim_{n\to\infty} X_n]$ at all. They may lie in different probability spaces.*

For $k > 1$, following the continuous mapping theorem, one has $(X_n^k)_{n\geq 1}$ converges in distribution to $\mathcal{N}(0,1)^k$, and then it follows similarly, as $\sup_{n\geq 1}\mathbb{E}[X_n^{2k}] < \infty$, from above. □

(b) Prove that for any $x \in \mathbb{R}$ and $d \geq 1$ we have

$$\left| e^{\mathrm{i}x} - \sum_{\ell=0}^{d} \frac{(\mathrm{i}x)^\ell}{\ell!} \right| \leq \frac{|x|^{d+1}}{(d+1)!}$$

Proof. We first show by induction that

$$e^{\mathrm{i}x} = \sum_{\ell=0}^{d}\left[\frac{(\mathrm{i}x)^\ell}{\ell!}\right] + \frac{(\mathrm{i}x)^{d+1}}{d!}\int_0^1 (1-u)^d e^{\mathrm{i}ux}\,\mathrm{d}u \tag{1.8}$$

By *fundamental theorem of calculus*, one has $e^{\mathrm{i}x} = 1 + (\mathrm{i}x)\int_0^1 e^{\mathrm{i}ux}\,\mathrm{d}u$, so it is true for $d = 0$. Assume inductively that 1.8 is true for $d-1, d \geq 1$. We show that it is true for d. Integrating by parts with

$$\begin{aligned} U &= e^{\mathrm{i}ux}, & V &= -\frac{(1-u)^d}{d}, \\ \mathrm{d}U &= \mathrm{i}xe^{\mathrm{i}ux}\,\mathrm{d}u, & \mathrm{d}V &= (1-u)^{d-1}\,\mathrm{d}u, \end{aligned}$$

gives

$$\begin{aligned} &\frac{(\mathrm{i}x)^d}{(d-1)!}\int_0^1 (1-u)^d e^{\mathrm{i}ux}\,\mathrm{d}u \\ &= \frac{(\mathrm{i}x)^d}{(d-1)!}\left[-e^{\mathrm{i}ux}\frac{(1-u)^d}{d}\bigg|_{u=0}^{u=1} + \frac{(\mathrm{i}x)}{d}\int_0^1 (1-u)^d e^{\mathrm{i}ux}\,\mathrm{d}u\right] \\ &= \frac{(\mathrm{i}x)^d}{d!} + \frac{(\mathrm{i}x)^{d+1}}{d!}\int_0^1 (1-u)^d e^{\mathrm{i}ux}\,\mathrm{d}u \end{aligned}$$

Hence, one has

$$\begin{aligned} e^{\mathrm{i}x} &= \sum_{\ell=0}^{d-1}\left[\frac{(\mathrm{i}x)^\ell}{\ell!}\right] + \frac{(\mathrm{i}x)^d}{(d-1)!}\int_0^1 (1-u)^{d-1} e^{\mathrm{i}ux}\,\mathrm{d}u \\ &= \sum_{\ell=0}^{d-1}\left[\frac{(\mathrm{i}x)^\ell}{\ell!}\right] + \frac{(\mathrm{i}x)^d}{d!} + \frac{(\mathrm{i}x)^{d+1}}{d!}\int_0^1 (1-u)^d e^{\mathrm{i}ux}\,\mathrm{d}u \\ &= \sum_{\ell=0}^{d}\left[\frac{(\mathrm{i}x)^\ell}{\ell!}\right] + \frac{(\mathrm{i}x)^{d+1}}{d!}\int_0^1 (1-u)^d e^{\mathrm{i}ux}\,\mathrm{d}u \end{aligned}$$

which completes the inductive step.

Now, it follows that

$$\begin{aligned} \left|e^{\mathrm{i}x} - \sum_{\ell=0}^{d}\frac{(\mathrm{i}x)^\ell}{\ell!}\right| &= \left|\frac{(\mathrm{i}x)^{d+1}}{d!}\int_0^1 (1-u)^d e^{\mathrm{i}ux}\,\mathrm{d}u\right| \\ &= \left|\frac{(\mathrm{i}x)^{d+1}}{d!}\right| \cdot \left|\int_0^1 (1-u)^d e^{\mathrm{i}ux}\,\mathrm{d}u\right| \\ &\leq \frac{|x|^{d+1}}{d!} \cdot \frac{1}{d+1} = \frac{|x|^{d+1}}{(d+1)!} \end{aligned}$$

as desired. $\square$

(c) Assume that

$$\lim_{n\to\infty} \mathbb{E}\left[X_n^k\right] = \mathbb{E}\left[\mathcal{N}(0,1)^k\right]$$

for any $k \in \mathbb{N}$. Prove that X_n converges in distribution to X.

Proof. Let $\alpha_k := \mathbb{E}[X^k] = \int_{\mathbb{R}} x^k \mu(\mathrm{d}x)$. It suffices to show that the probability measure μ is unique with the moments $\alpha_1, \alpha_2, \ldots$, since then the distribution of the convergence of X_n is uniquely determined to be $X \sim \mathcal{N}(0,1)$.

Note that for a standard normal, its moments are $0, 1!!, 0, 3!!, 0, 5!!, 0, 7!!, \ldots$, so $\alpha_k \leq k!$ is finite of all orders, implying that $\alpha_k s^k/k! \to 0$ for some positive s. Let $\beta_k = \int_{-\infty}^{\infty} |x|^k \mu(dx)$ be the absolute moments. We first show that

$$\frac{\beta_k r^k}{k!} \xrightarrow[k\to\infty]{} 0 \tag{1.9}$$

for some positive r. Choose $0 < r < s$. Since $\alpha_k s^k/k! \to 0$, one has $2kr^{2k-1} < s^{2k}$ for large k. Since $|x|^{2k-1} \leq 1 + |x|^{2k}$,

$$\frac{\beta_{2k-1} r^{2k-1}}{(2k-1)!} \leq \frac{r^{2k-1}}{(2k-1)!} + \frac{\beta_{2k} s^{2k}}{(2k)!}$$

for large k. Hence 1.9 holds as k goes to infinity through odd values, and $\beta_k = \alpha_k$ for k even, so it holds for all k.

From part 9b, one has

$$\left| e^{\mathrm{i}tx} \left(e^{\mathrm{i}hx} - \sum_{k=0}^{n} \frac{(\mathrm{i}hx)^k}{k!} \right) \right| \leq \frac{|hx|^{n+1}}{(n+1)!},$$

and therefore the characteristic function φ of μ satisfies

$$\left| \varphi(t+h) - \sum_{k=0}^{n} \frac{h^k}{k!} \int_{-\infty}^{\infty} (\mathrm{i}x)^k e^{\mathrm{i}tx} \mu(dx) \right| \leq \frac{|h|^{n+1} \beta_{n+1}}{(n+1)!}.$$

Observe that the integral in the above equation is $\varphi^{(k)}(t)$, the k-th derivative of φ. By 1.9,

$$\varphi(t+h) = \sum_{k=0}^{\infty} \frac{\varphi^{(k)}(t)}{k!} h^k, \quad |h| \leq r.$$

If ν is another probability measure with moments α_k and characteristic function $\psi(t)$, the same argument gives

$$\psi(t+h) = \sum_{k=0}^{\infty} \frac{\psi^{(k)}(t)}{k!} h^k, \quad |h| \leq r.$$

Take $t = 0$; since $\varphi^{(k)}(0) = i^k\alpha_k = \psi^{(k)}(0)$, φ and ψ agree in $(-r, r)$ and hence have identical derivatives there. Taking $t = r - \varepsilon$ and $t = -r + \varepsilon$ in the above two equations shows that φ and ψ also agree in $(-2r + \varepsilon, 2r - \varepsilon)$ and hence in $(-2r, 2r)$. But then they must by the same argument agree in $(-3r, 3r)$ as well, and so on. Thus φ and ψ coincide, and by the uniqueness theorem for characteristic functions, so do μ and ν. □

(d) Let $w_y(m)$ be the number of prime factors of m which are smaller than y. Let $(B_p)_{p \text{ prime}}$ be independent random variables such that $\mathbb{P}(B_p = 1) = 1 - \mathbb{P}(B_p = 0) = 1/p$. Denote

$$W_y = \sum_{p\le y} B_p, \quad \mu_y = \sum_{p\le y} \frac{1}{p}, \quad \sigma_y^2 = \sum_{p\le y}\left(\frac{1}{p} - \frac{1}{p^2}\right)$$

Prove that if $y = n^{o(1)}$, then for any $d \in \mathbb{N}$ we have

$$\mathbb{E}\left[\left(\frac{w_y(k) - \mu_y}{\sigma_y}\right)^d\right] - \mathbb{E}\left[\left(\frac{W_y - \mu_y}{\sigma_y}\right)^d\right] \xrightarrow[n\to\infty]{} 0$$

Proof. Let

$$\delta_p(m) = \begin{cases} 1 & p \mid m \\ 0 & p \nmid m \end{cases}$$

Then,

$$w_y(m) = \sum_{p\le y} \delta_p(m)$$

Note that $\mathbb{E}\left[W_y^d\right]$ is the sum

$$\sum_{u=1}^{d} \sum{}' \frac{d!}{d_1! \cdots d_u!} \frac{1}{u!} \sum{}'' \mathbb{E}\left[B_{p_1}^{d_1} \cdots B_{p_u}^{d_u}\right], \tag{1.10}$$

where $\sum'$ extends over the u tuples $(d_1, \dots, d_u)$ of positive integers satisfying $d_1 + \dots + d_u = d$, and $\sum''$ extends over the u tuples $(p_1, \dots, p_u)$ of distinct primes not exceeding y. Since B_p assumes only the values 0 and 1, from the independence of the B_p and the fact that the p_i are distinct, it follows that the summand in 1.10 is

$$\mathbb{E}\left[B_{p_1} \cdots B_{p_u}\right] = \frac{1}{p_1 \cdots p_u} \tag{1.11}$$

By definition, $\mathbb{E}_n\left[w_y^d\right]$[16] is just 1.10 with the summand replaced by $\mathbb{E}_n\left[\delta_{p_1}^{d_1} \cdots \delta_{p_u}^{d_u}\right]$. Since $\delta_p(m)$ assumes only the values 0 and 1, given that the p_i are distinct, it follows that this

[16] $\mathbb{E}_n$ here means

$$\mathbb{E}_n[f] = n^{-1} \sum_{m=1}^{n} f(m),$$

and we assume that $\mathbb{E}\left[\left(\frac{w_y(k)-\mu_y}{\sigma_y}\right)^d\right]$ is calculated in this way.

summand is

$$\mathbb{E}_n\left[\delta_{p_1}\cdots\delta_{p_u}\right] = \frac{1}{n}\left\lfloor\frac{n}{p_1\cdots p_u}\right\rfloor \tag{1.12}$$

But 1.11 and 1.12 differ by at most $1/n$, and hence $\mathbb{E}\left[W_y^d\right]$ and $\mathbb{E}_n\left[w_y^d\right]$ differ by at most the sum 1.10 with the summand replaced by $1/n$. Therefore,

$$\left|\mathbb{E}\left[W_y^d\right] - \mathbb{E}_n\left[w_y^d\right]\right| \le \frac{1}{n}\left(\sum_{p\le y} 1\right)^d \le \frac{y^d}{n} \tag{1.13}$$

Now

$$\mathbb{E}\left[(W_y-\mu_y)^d\right] = \sum_{j=0}^{d}\binom{d}{j}\mathbb{E}\left[W_y^j\right](-\mu_y)^{d-j},$$

and $\mathbb{E}_n\left[(w_y-\mu_y)^d\right]$ has the analogous expansion. Comparing the two expansions term for term and applying 1.13 shows that

$$\left|\mathbb{E}\left[(W_y-\mu_y)^d\right] - \mathbb{E}_n\left[(w_y-\mu_y)^d\right]\right| \le \sum_{j=0}^{d}\binom{d}{j}\frac{y^j}{n}\mu_y^{d-j} = \frac{1}{n}(y+\mu_y)^d$$

Since $\mu_y \le y$, and since $y^d/n \to 0$ by the assumption $y = n^{\mathrm{o}(1)}$, one has

$$\mathbb{E}\left[\left(\frac{w_y(k)-\mu_y}{\sigma_y}\right)^d\right] - \mathbb{E}\left[\left(\frac{W_y-\mu_y}{\sigma_y}\right)^d\right] \xrightarrow[n\to\infty]{} 0$$

as desired. □

(e) Conclude.

Proof. We first want to show that

$$\frac{w(k)-\log\log n}{\sqrt{\log\log n}} \xrightarrow[n\to\infty]{} \mathcal{N}(0,1)$$

is unaffected if the range of p is further restricted with $w_y(m)$.

Proof. By *Mertens' second theorem*, one has the estimate that

$$\mu_y = \sum_{p\le y}\frac{1}{p} = \log\log y + \mathrm{O}(1)$$

This satisfies (for example let $\log y = \log n/\log\log n$) that $y\to\infty$ slowly enough that $\log y/\log n \to 0$ but fast enough that

$$\sum_{y<p\le n}\frac{1}{p} = \mathrm{o}(\log\log n)^{1/2} \tag{1.14}$$

Now, let $\mathbb{P}_n$ be the probability measure that places mass $1/n$ at each of $1, 2, \ldots, n$. Recall that δ_p is defined as 1 or 0 according as the prime p divides m or not. Note that if $p_1, \ldots, p_u$ are distinct primes, then $\forall i, p_i \mid m$ iff $\prod_i p_i \mid m$, so that one has

$$\mathbb{P}_n\left[m : \delta_{p_i}(m) = 1, \forall i \in [\![1, u]\!]\right] = \frac{1}{n}\left\lfloor \frac{n}{\prod_{i=1}^{u} p_i} \right\rfloor$$

In particular when $u = 1$,

$$\mathbb{E}_n\left[\sum_{p>y} \delta_p\right] = \sum_{y<p\leq n} \mathbb{P}_n\left[m : \delta_p(m) = 1\right] \leq \sum_{y<p\leq n} \frac{1}{p}$$

By 1.14 and Markov's inequality,

$$\mathbb{P}_n\left[m : |w(m) - w_y(m)| \geq \varepsilon(\log\log n)^{1/2}\right] \to 0$$

Therefore the desired result is unaffected if $w_y(m)$ is substituted for $w(m)$. □

Now compare $w_y(m)$ with the corresponding sum $W_y = \sum_{p\leq y} B_p$. The mean and variance of S_n are

$$\mu_y = \sum_{p\leq y} \frac{1}{p}, \quad \sigma_y^2 = \sum_{p\leq y} \frac{1}{p}\left(1 - \frac{1}{p}\right),$$

and each is $\log\log n + o(\log\log n)^{1/2}$ by 1.14. Thus, it suffices to show that

$$\mathbb{P}_n\left[m : \frac{w_y(m) - \mu_y}{\sigma_y} \leq x\right] \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2} du.$$

Since the B_p are bounded, following part 9c, it suffices to show the moments converge to $\mathcal{N}(0, 1)$. Note that B_p should be replaced by $B_p - p^{-1}$ to center it. Thus the d-th moment of $(W_y - \mu_y)/\sigma_y$ converges to that of $\mathcal{N}(0, 1)$. However, it is already proved in part 9d that, as $n \to \infty$, one has

$$\mathbb{E}\left[\left(\frac{w_y(k) - \mu_y}{\sigma_y}\right)^d\right] - \mathbb{E}\left[\left(\frac{W_y - \mu_y}{\sigma_y}\right)^d\right] \xrightarrow[n\to\infty]{} 0,$$

so we are done. □

4 Dependent Random Variables

4.1 Conditioning, Radon-Nikodym Theorem

1. Let X and Y be independent Gaussian random variables with null expectation and variance 1. Show that $\frac{X+Y}{\sqrt{2}}$ and $\frac{X-Y}{\sqrt{2}}$ are also independent $\mathcal{N}(0, 1)$.

Then, one has

$$\begin{aligned}&\mathbb{P}(|S_{n+1}| = y+1 \mid |S_n| = y, \dots, |S_1| = y_1)\\ =& \frac{1}{2}\mathbb{P}(S_{n+1} = y+1 \mid S_n = y) + \frac{1}{2}\mathbb{P}(S_{n+1} = -y-1 \mid S_n = -y) = \frac{1}{2}\end{aligned}$$

and similar for $|S_{n+1}| = y-1$ by symmetry.

Transition matrix:

$$\mathbf{P} := (P_{ij}), \quad P_{i,i+1} = \frac{1}{2}, \forall i > 0, \quad P_{0,1} = 1$$

and zero elsewhere.

□

2. Consider a Markov chain X with state space $\{0, 1, \dots, n\}$ and transition matrix

$$\pi(0,k) = \frac{1}{2^{k+1}},\ 0 \le k \le n-1,\ \pi(0,n) = \frac{1}{2^n}$$
$$\pi(k,k-1) = 1,\ 1 \le k \le n-1,\ \pi(n,n) = \pi(n,n-1) = \frac{1}{2}.$$

(a) Prove that the chain has a unique invariant probability measure μ and calculate it.

Proof. Let $I = \{0, 1, \dots, n\}$. Denote $\mu_k := \mu(k), k \in I$. It suffices to solve the system

$$\begin{cases}\sum_{i=0}^n \mu_i \pi_{ij} = \mu_j, \quad \forall j \in I\\ \mu_0 + \mu_1 + \dots + \mu_n = 1\\ \mu_0, \dots, \mu_n \ge 0\end{cases}$$

After calculation, the system has only one solution that

$$\mu = (\mu_0, \dots, \mu_n): \quad \mu_k = \frac{1}{2^{k+1}}, k \in [\![0, n-1]\!], \quad \mu_n = \frac{1}{2^n}$$

□

(b) Prove that for any $0 \le x_0 \le n-1$, $\pi^{(x_0+1)}(x_0, \cdot) = \mu$.

Proof. We prove by induction and use the *Chapman-Kolmogorov Equation*. For $x_0 = 0$, $\pi^{(1)}(x_0, \cdot)$ is by definition

$$\pi(0,k) = \frac{1}{2^{k+1}},\ 0 \le k \le n-1,\ \pi(0,n) = \frac{1}{2^n}$$

which corresponds with μ. Note that

$$\pi^{(2)}_{1j} = \sum_k \pi_{1k}\pi_{kj} = \pi_{0j} = \mu_j, \forall j \in I; \pi^{(2)}_{2j} = \sum_k \pi_{2k}\pi_{kj} = \pi_{1j}$$

$$\ldots \text{ assume } \pi^{(n-1)}_{(n-2),j} = \mu_j, \forall j \in I, \text{ and note that}$$

$$\pi^{(n-1)}_{(n-1),j} = \sum_k \pi^{(n-2)}_{(n-1),k}\pi_{kj} = \cdots = \pi_{1j},$$

$$\implies \pi^{(n)}_{(n-1),j} = \sum_k \pi^{(n-1)}_{(n-1),k}\pi_{kj} = \sum_k \pi_{1k}\pi_{kj} = \pi_{0j} = \mu_j, \forall j \in I$$

Therefore, by induction, one has for any $0 \leq x_0 \leq n-1$, $\pi^{(x_0+1)}(x_0, \cdot) = \mu$. □

(c) Prove that for any $0 \leq x_0 \leq n$, $\pi^{(n)}(x_0, \cdot) = \mu$.

Proof. When $0 \leq x_0 \leq n-1$, $\pi^{(n)}_{x_0,\cdot} = \pi(0, \cdot) = \mu$ since μ is invariant, which implies $\pi^{(k)}(0, \cdot) = \mu$.

When $x_0 = n$, one has

$$\pi^{(n)}_{x_0,0} = \mathbb{P}(n \to n-1 \text{ at first step}) = \frac{1}{2},$$

$$\pi^{(n)}_{x_0,1} = \mathbb{P}(n \to n \to n-1 \to \cdots \to 1) = \mathbb{P}(n \to n \to n-1) = \frac{1}{2^2}, \ldots$$

Inductively, one has $\pi^{(n)}_{x_0,k} = \frac{1}{2^{k+1}}$ for $0 \leq k \leq n-1$, and $\pi^{(n)}_{x_0,n} = \frac{1}{2^n}$. Hence, $\pi^{(n)}(n, \cdot) = \mu$ as well. □

(d) For any $t \geq 1$, calculate

$$d(t) := \frac{1}{2}\sum_{x=0}^{n} \left|\pi^{(t)}(n, x) - \mu(x)\right|,$$

and plot $t \mapsto d(t)$.

Solution. Note when $t \geq n$, one has $\pi^{(t)}_{n,x} = \mu_x$, since we have proved $\pi(n)_{n,x} = \mu_x$, and for any more steps, the distribution stays invariant. Hence, when $t \geq n$, $d(t) = 0$. When $t < n$, since n can at most go to $n-t$, one has

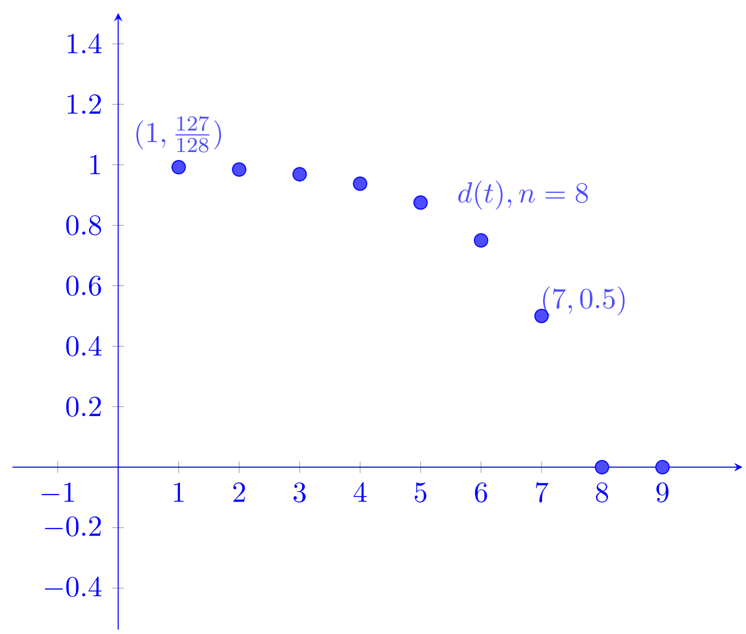
$$\begin{cases} \pi^{(t)}_{n,k} = 0, & k < n-t \\ \pi^{(t)}_{n,k} = 1/2^{k-(n-t)+1}, & n-t \leq k \leq n-1 \\ \pi^{(t)}_{n,n} = 1/2^t \end{cases}$$

Then, one has

$$\begin{aligned} d(1) &= \frac{1}{2}\left(\sum_{0}^{n-2}\mu_x + \left(\frac{1}{2} - \mu_{n-1}\right) + \left(\frac{1}{2} - \mu_n\right)\right) \\ &= 1 - \mu_{n-1} - \mu_n \end{aligned}$$

and, noting that $\pi_{n,x}^{(t)} - \mu_x > 0$ as $k + 1 - (n-t) < k + 1$,

$$\begin{aligned} d(t) &= \frac{1}{2}\left(\sum_{x=0}^{n-t-1}\mu_x + \sum_{x=n-t}^{n}\left|\pi_{n,x}^{(t)} - \mu_x\right|\right) \\ &= \frac{1}{2}\left(1 - \sum_{x=n-t}^{n}\mu_x + 1 - \sum_{x=n-t}^{n}\mu_x\right) \\ &= 1 - \sum_{x=n-t}^{n}\mu_x \end{aligned}$$



□

3. For fixed $p, q \in [0, 1]$, consider a Markov chain X with two states $\{1, 2\}$, with transition matrix

$$\pi = (\pi(i,j))_{1\le i,j\le 2} = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}$$

5. Let X be a real-valued Markov process with transition semigroup $(Q_t)_{t\geq 0}$, let $f:\mathbb{R}\to[0,1]$ be measurable, and let $t_0>0$. Prove that the process $M_t=Q_{t_0-t}f(X_t)$ is a martingale for $t\in[0,t_0]$.

   Proof. We show M_t is a $\mathbb{P}^x$ martingale where $X=(X_t,\mathbb{P}^x)$. By assumption, M_t is integrable and $\mathcal{F}_t$-adapted. It suffices to show $\mathbb{E}[M_t|\mathcal{F}_s]=M_s$ for $0\leq s<t\leq t_0$. Note that, by definition [Gal16, Def. 6.2], $M_s=Q_{t_0-s}f(X_s)=\mathbb{E}[f(X_{t_0})|\mathcal{F}_s]$ while $M_t=\mathbb{E}[f(X_{t_0})|\mathcal{F}_t]$. Since $\mathcal{F}_s\subset\mathcal{F}_t$, by iterated conditioning law, one has

$$\mathbb{E}[M_t|\mathcal{F}_s]=\mathbb{E}\left[\mathbb{E}[f(X_{t_0})|\mathcal{F}_t]|\mathcal{F}_s\right]=\mathbb{E}[f(X_{t_0})|\mathcal{F}_s]=M_s$$

   as desired. □

10 Stochastic Differential Equations

1. In this exercise we will prove weak existence and weak uniqueness of solutions of the stochastic differential equation

$$E(\sigma,b):\quad dX_t=\sigma(X_t)\,dW_t+b(X_t)\,dt \tag{2.5}$$

   where $\sigma,b:\mathbb{R}\to\mathbb{R}$ are bounded and continuous such that $\int_{\mathbb{R}}|b(x)|dx<\infty$ and $\sigma\geq\varepsilon$ for some $\varepsilon>0$. We will also argue pathwise uniqueness if σ is Lipschitz.

   (a) First we study the case $b=0$. Suppose that X solves equation 2.5, and for each $t\geq 0$ define

$$A_t=\int_0^t\sigma(X_s)^2\,ds,\qquad \tau_t=\inf\{s\geq 0:A_s>t\}$$

   Justify the equalities

$$\tau_t=\int_0^t\frac{dr}{\sigma(X_{\tau_r})^2},\qquad A_t=\inf\left\{s\geq 0:\int_0^s\frac{dr}{\sigma(X_{\tau_r})^2}>t\right\}$$

   Proof. When $b=0$, $X_t=\int_0^t\sigma(X_s)\,\mathrm{d}W_s$, is continuous. Then $\sigma(X_t)$ is continuous, so A_t is continuously differentiable, and $\sigma(X_t)>\epsilon$ so A_t is strictly increasing. So τ_t is the inverse of A_t and

$$\frac{d}{dt}\tau_t=\left(\frac{d}{ds}A_s\Big|_{\tau_t}\right)^{-1}=\sigma(X_{\tau_t})^{-2}\quad\Longrightarrow\quad\tau_t=\int_0^t\frac{1}{\sigma(A_{\tau_r})^2}dr,$$

   Then τ_t is increasing, so that

$$A_t:=\inf\left\{s\geq 0:\int_0^t\frac{dr}{\sigma(X_{\tau_r})^2}>t\right\}$$

   as desired. □

(b) In the setting of (a), argue that there is a Brownian motion $(B_t)_{t\geq 0}$ started from x such that, a.s. for every $t \geq 0$, $X_t = B_{\inf\left\{s\geq 0: \int_0^s \sigma(B_r)^{-2} dr > t\right\}}$.

Proof. By part (a), for $b = 0$, X is a continuous local martingale, $\langle X \rangle_t = A_t$ strictly increasing, and $\langle X \rangle_\infty = \infty$, then by [Bas11, Thm. 12.2], $B_t := X_{\tau_t}$ defines a Brownian motion starting at x and

$$X_t = B_{\langle X \rangle_t} = B_{\inf\{s\geq 0: \int_0^s \sigma(X_{\tau_r})^{-2} dr > t\}} = B_{\inf\{s\geq 0: \int_0^s \sigma(B_r)^{-2} dr > t\}}$$

as desired. □

(c) Show that weak existence and weak uniqueness hold for $E(\sigma, 0)$.

Proof. (Weak existence). Let B be a Brownian motion starting at x. Define

$$Y_t := \int_0^t \frac{1}{\sigma(B_s)} dB_s,$$

and let τ_t and A_t be s.t.

$$\tau_t := \langle Y \rangle_t = \int_0^t \frac{1}{\sigma(B_s)^2} ds, \quad A_t := \inf\{s \geq 0 : \tau_s > t\},$$

then since $\langle Y \rangle_\infty = \infty$ a.s., then by [Bas11, Thm. 12.2], $W_t := Y_{A_t}$ defines a Brownian motion starting at x, and

$$\sum_{k=1}^{n-1} \sigma(B_{A_{tk/n}})(Y_{A_{t(k+1)/n}} - Y_{tk/n}) \xrightarrow{\mathbb{P}} \int_0^t \sigma(B_{A_s}) dY_{A_s},$$

so that

$$\int_0^t \sigma(B_{A_s}) dW_s = \int_0^t \sigma(B_{A_s}) dY_{A_s} = \int_0^{A_t} \sigma(B_s) dY_s$$

Also,

$$\int_0^t \sigma(B_s) dY_s = \int_0^t dB_s = B_t,$$

so that we can let $X_t := B_{A_t} = B_{\inf\{s\geq 0: \int_0^s \sigma(B_r)^{-2} dr > t\}}$. Then, $X_0 = x$ and

$$X_t = B_{A_t} = \int_0^t \sigma(B_{A_s}) dW_s = \int_0^t \sigma(X_s) dW_s$$

(Weak uniqueness). We've shown that $X_t = B_{\inf\{s\geq 0: \int_0^s \sigma(B_r)^{-2} dr > t\}}$. Since X is continuous and the finite dimensional distribution is determined by a Brownian motion, the law of X is unique. □

(d) Show that there exists a monotone increasing and twice continuously differentiable function $F : \mathbb{R} \to \mathbb{R}$ such that $F(X_t)$ is a martingale. Give an explicit formula for F in terms of σ and b.

Proof. Suppose $F \in C^2(\mathbb{R})$. By Ito's formula, one has

$$\begin{aligned} F(X_t) &= F(X_0) + \int_0^t F'(X_s)dX_s + \frac{1}{2}\int_0^t F''(X_s)d\langle X\rangle_s \\ &= \int_0^t F'(X_s)\sigma(X_s)dW_s + \int_0^t F'(X_s)b(X_s)dt + \frac{1}{2}\int_0^t F''(X_s)\sigma(X_t)^2dt, \end{aligned}$$

so $F(X_t)$ is a martingale iff the $\mathrm{d}t$ term $F'(X_t)b(X_t) + \frac{1}{2}F''(X_t)\sigma(X_t)^2 = 0$, implying that

$$F'(x) = \exp\left(-\int_0^x \frac{2b(s)}{\sigma(s)^2}ds\right), \quad F(x) = \int_0^x \exp\left(-\int_0^t \frac{2b(s)}{\sigma(s)^2}ds\right)dt,$$

Note that $F' > 0$, so F is monotone increasing as desired. □

(e) Show that $Y_t = F(X_t)$ solves an SDE of the form $dY_t = \sigma'(Y_t)\,dW_t$ and determine the function σ'.

Proof. By part (d), one has $dY_t = dF(X_t) = F'(X_t)\sigma(X_t)dW_t$. Note that $F'(x) \geq \exp\left(-2\epsilon^{-2}\int_{\mathbb{R}}|b(t)|dt\right)$, so $F : \mathbb{R} \to \mathbb{R}$ is bijective, and hence F^{-1} exists. Then, one has

$$E'(\sigma') : \quad dY_t = dF(X_t) = F'(F^{-1}(Y_t))\sigma(F^{-1}(Y_t))dW_t = \sigma'(Y_t)dW_t$$

where $\sigma' = (F' \cdot \sigma) \circ F^{-1}$. □

(f) Using parts (a)-(c), show that weak existence and weak uniqueness hold for $E(\sigma, b)$, along with pathwise uniqueness if σ is Lipschitz.

Proof. First note that by part (c) along with the fact that $\sigma' : \mathbb{R} \mapsto \mathbb{R}$ is continuous and that $\sigma' \geq \epsilon\exp\left(-2\epsilon^{-2}\int_{\mathbb{R}}|b(t)|dt\right)$, weak existence and weak uniqueness hold for $E'(\sigma')$.

Weak existence of $E(\sigma, b)$: fix $x \in \mathbb{R}$. Set $y = F(x)$. There exists a solution Y of $E'_y(\sigma')$. Define $X_t := F^{-1}(Y_t)$. By Itô's formula, we get

$$X_t = x + \int_0^t \frac{dF^{-1}}{dy}(Y_s)\,dY_s + \frac{1}{2}\int_0^t \frac{d^2F^{-1}}{dy^2}(Y_s)\,d\langle Y, Y\rangle_s.$$

By $F^{-1}(F(x)) = x$, we get

$$\frac{dF^{-1}}{dy}(F(x))\frac{dF}{dx}(x) = 1, \quad \frac{d^2F^{-1}}{dy^2}(F(x))\left(\frac{dF}{dx}(x)\right)^2 + \frac{dF^{-1}}{dy}(F(x))\frac{d^2F}{dx^2}(x) = 0$$

Thus,

$$\frac{dF^{-1}}{dy}(Y_s) = \frac{dF^{-1}}{dy}(F(X_s)) = e^{\int_0^X \frac{2b(r)}{\sigma(r)^2}dr}, \quad \frac{d^2F^{-1}}{dy^2}(Y_s) = \frac{2b(X_s)}{\sigma(X_s)^2} e^{2\int_0^{X_s} \frac{2b(r)}{\sigma(r)^2}dr}$$

Since $dY_t = \sigma'(Y_t)\,dW_t = e^{-\int_0^{X_t} \frac{2b(r)}{\sigma(r)^2}dr}\sigma(X_t)\,dW_t$, one has

$$\begin{aligned} X_t &= x + \int_0^t \frac{dF^{-1}}{dy}(Y_s)\,dY_s + \frac{1}{2}\int_0^t \frac{d^2F^{-1}}{dy^2}(Y_s)\,d\langle Y, Y\rangle_s \\ &= x + \int_0^t \sigma(X_s)\,dW_s + \int_0^t b(X_s)\,ds \end{aligned}$$

and so X is a solution of $E_x(\sigma, b)$. Weak uniqueness of X immediately follows from weak uniqueness of Y.

Pathwise uniqueness of $E(\sigma, b)$: given σ is Lipschitz, it suffices to show that σ' is Lipshitz. Indeed, let $c_0 > 0$ be such that $|\sigma(x_1) - \sigma(x_2)| \leq c_0|x_1 - x_2|$, also let

$$c_1 := \sup_{\mathbb{R}} |F''|, \quad c_2 := \sup_{\mathbb{R}} |\dot{F}^{-1}|, \quad c_3 := \sup_{\mathbb{R}} |F'|, \quad c_4 := \sup_{\mathbb{R}} |\sigma|$$

which are all bounded, so that

$$\begin{aligned} |\sigma'(y_1) - \sigma'(y_2)| &= \left|(F'(F^{-1}(y_1))\sigma((F^{-1}(y_1)) - (F'(F^{-1}(y_2))\sigma((F^{-1}(y_2))\right| \\ &\leq (c_1c_4 + c_0c_3) \cdot c_2 \cdot |y_1 - y_2| \end{aligned}$$

is bounded, so σ' is Lipschitz. □

2. Let W be a Brownian motion and let $a > 1/2$ and $z_0 > 0$. This exercise proves that there is a unique positive semimartingale Z such that for every $t \geq 0$,

$$Z_t = z_0 + W_t + \int_0^t \frac{a}{Z_s} ds \tag{2.6}$$

This process is known as a Bessel process.

(a) For $n \in \mathbb{N}$ define $f_n : \mathbb{R} \to \mathbb{R}_+$by $f_n(x) = |x|^{-1} \wedge n$. Justify the existence of a unique semimartingale Z^n that solves

$$Z_t^n = z_0 + W_t + a\int_0^t f_n(Z_s^n)\,ds$$

where either limit goes to 0. Hence,

$$u(x)=\mathbb{E}_x\left[g\left(B_T\right)\right]=\mathbb{E}_x\left[g\left(B_{U_1}\right)\mathbb{1}_{\{U_1<U_0\}}\right]=0, \quad \forall x\in\mathcal{B}_1^\star$$

but then, $u(x)\to 0\neq g(0)=1$ as $x\to 0, x\in\mathcal{B}_1^\star$. Contradiction! □

6. [Gal16, Exercise 7.26] In this exercise, $d\geq 3$. Let K be a compact subset of the open unit ball of $\mathbb{R}^d$, and $T_K:=\inf\{t\geq 0: B_t\in K\}$. We assume that $D:=\mathbb{R}^d\backslash K$ is connected. We also consider a function g defined and continuous on K. The goal of the exercise is to determine all functions $u:\bar{D}\to\mathbb{R}$ that satisfy:

    (P) u is bounded and continuous on $\bar{D}$, harmonic on D, and $u(y)=g(y)$ if $y\in\partial D$.

    (This is the Dirichlet problem in D, but in contrast with [Gal16, Sec. 7.3] above, D is unbounded here.) We fix an increasing sequence $(R_n)_{n\geq 1}$ of reals, with $R_1\geq 1$ and $R_n\uparrow\infty$ as $n\to\infty$. For every $n\geq 1$, we set $T_{(n)}:=\inf\{t\geq 0:|B_t|\geq R_n\}$.

    (a) Suppose that u satisfies (P). Prove that, for every $n\geq 1$ and every $x\in D$ such that $|x|<R_n$

    $$u(x)=\mathbb{E}_x\left[g\left(B_{T_K}\right)\mathbb{1}_{\{T_K\leq T_{(n)}\}}\right]+\mathbb{E}_x\left[u(B_{T_{(n)}})\mathbb{1}_{\{T_{(n)}\leq T_K\}}\right]$$

    Proof. Note that $x\in D=\mathbb{R}^d\backslash K$ but $|x|<R_n$, so the bounded domain is in fact $\mathcal{B}(0,R_n)\backslash K$. Let $T=\inf\{t\geq 0: B_t\notin\mathcal{B}_{R_n}\backslash K\}$. By [Gal16, Prop. 7.7], one has for every $x\in\mathcal{B}_{R_n}\backslash K$, as it either exits first to $>R_n$ or to $\in K$,

    $$u(x)=\mathbb{E}_x[g(B_T)]=\mathbb{E}_x\left[g\left(B_{T_K}\right)\mathbb{1}_{\{T_K\leq T_{(n)}\}}\right]+\mathbb{E}_x\left[u\left(B_{T_{(n)}}\right)\mathbb{1}_{\{T_{(n)}\leq T_K\}}\right]$$

    as desired. □

    (b) Show that, by replacing the sequence $(R_n)_{n\geq 1}$ with a subsequence if necessary, we may assume that there exists a constant $\alpha\in\mathbb{R}$ such that, for every $x\in D$,

    $$\lim_{n\to\infty}\mathbb{E}_x\left[u(B_{T_{(n)}})\right]=\alpha$$

    and that we then have

    $$\lim_{|x|\to\infty}u(x)=\alpha$$

    Proof. We try to apply Liouville's theorem [Eva98, Thm. 2.8] which requires the function defined on all of $\mathbb{R}^n$. Let

    $$f_n(x):=\mathbb{E}_x\left[u\left(B_{T_{(n)}}\right)\right] \quad \forall x\in\mathcal{B}_{R_n}, \quad n\geq 1$$

Note that u is bounded, so is f_n, and by [Gal16, Prop. 7.7.ii], f_n is harmonic (applying strong Markov property). It now suffices to find an increasing (sub)sequence s.t. the limit of f_n converges uniformly on every compact subset $K \subset \mathbb{R}^d$ for every $x \in \mathbb{R}^d$.

We want to show that $\{f_n\}$ is equicontinuous on $\overline{\mathcal{B}(p,r)}$ for every $p \in \mathbb{Q}^d$ and $r \in \mathbb{Q}_+$. Since then, by applying Arzelà-Ascoli theorem, there exists an increasing subsequence n_k such that $f_{n_k}(x)$ converges uniformly on $\overline{\mathcal{B}(p,r)}$, and we're done.

Indeed, let $p \in \mathbb{Q}^d, r \in \mathbb{Q}_+$ and $M := \sup_{z\in\bar{D}} |u(z)|$. Choose $N \geq 1$ such that $\mathcal{B}(p,r) \subset \mathcal{B}_{R_N}$ and $\eta := d(\mathcal{B}(p,r), \partial\mathcal{B}_{R_N}) > 0$. By local estimates for harmonic function, there exists some $c > 0$ such that for $n \geq N$ one has

$$|df_n(x)| \leq \frac{c}{(\eta/2)^{d+1}} \|f_n\|_{L^1(\mathcal{B}(x,\eta/2))} \leq \frac{cM}{\eta/2} \quad \forall x \in \mathcal{B}(p, r+\eta/2)$$

Let $x, y \in \overline{\mathcal{B}(p,r)}$ such that $|x-y| < \frac{\eta}{2cM}\epsilon$ for some $\epsilon > 0$. Then

$$|f_n(x) - f_n(y)| \leq \sup_{z\in B(p,r+\eta/2)} |df_n(z)||x-y| < \epsilon$$

Hence, by letting the subsequence n_k start at N and setting $f(x) := \lim_{k\to\infty} f_{n_k}(x)$, one has $f(x) = \lim_{n\to\infty} \mathbb{E}_x\left[u(B_{T_{(n)}})\right] = \alpha$ a constant by Liouville's theorem.

Now, note that M is finite, and one can always pick some n_j large enough such that $\mathbb{P}_x(T_{(n_j)} > T_K) < \epsilon$ for any $\epsilon > 0$, so that

$$|u(x) - \alpha| \leq c_1 \cdot \epsilon \xrightarrow{|x|\to\infty} 0$$

for some constant c_1, as desired. □

(c) Show that, for every $x \in D$,

$$u(x) = \mathbb{E}_x\left[g(B_{T_K}) \mathbb{1}_{\{T_K<\infty\}}\right] + \alpha\mathbb{P}_x(T_K = \infty)$$

Proof. Note that $\lim_{t\to\infty} |B_t| = \infty$ [Gal16, Thm. 7.17], but we have shown that $u(x) \xrightarrow{|x|\to\infty} \alpha$, so $T_{n_k} < \infty$ a.s. for every $k \geq 1$. Therefore, one has

$$\begin{aligned}
&\mathbb{E}_x\left[u\left(B_{T_{n_k}}\right)\mathbb{1}_{\{T_{n_k}\leq T_K\}}\right] \\
=&\mathbb{E}_x\left[u\left(B_{T_{n_k}}\right)\mathbb{1}_{\{T_{n_k}\leq T_K<\infty\}}\right] + \mathbb{E}_x\left[u\left(B_{T_{n_k}}\right)\mathbb{1}_{\{T_{n_k}<\infty\}\cap\{T_K=\infty\}}\right] \\
&\xrightarrow{k\to\infty} 0 + \alpha\mathbb{P}_x(T_K = \infty)
\end{aligned}$$

By parts (a) and (b), one has

$$\begin{aligned} u(x) &= \lim_{k\to\infty} \mathbb{E}_x\left[g\left(B_{T_K}\right)\mathbb{1}_{\{T_K\le T_{n_k}\}}\right] + \lim_{k\to\infty}\mathbb{E}_x\left[u\left(B_{T_{n_k}}\right)\mathbb{1}_{\{T_{n_k}\le T_K\}}\right] \\ &= \mathbb{E}_x\left[g\left(B_{T_K}\right)\mathbb{1}_{\{T_K<\infty\}}\right] + \alpha\mathbb{P}_x\left(T_K=\infty\right) \end{aligned}$$

as desired. □

(d) Assume that D satisfies the exterior cone condition at every $y \in \partial D$ (this is defined in the same way as when D is bounded). Show that, for any choice of $\alpha \in \mathbb{R}$, the formula of part (c) gives a solution of the problem (P).

Proof. By [Gal16, Prop. 7.7.ii], $u(x)$ is harmonic. Now it suffices to show that

$$\lim_{x\in D\to y} u(x) = g(y)$$

for every $y \in \partial D$. Since then, by [Gal16, Thm. 7.8] we're done. Note that in the theorem D bounded is only required for finite hitting time, so it does not affect its validity here.

Denote $M := \sup_{z\in K}|g(z)|$. Fix $\epsilon > 0$ and $y \in \partial D$. Choose $\delta > 0$ such that

$$|g(z) - g(y)| < \epsilon \quad \forall z \in K \cap B(y,\delta)$$

Choose $\eta > 0$ such that

$$\mathbb{P}_0\left(\sup_{t\le\eta}|B_t| \ge \frac{\delta}{2}\right) < \epsilon$$

Observe that

$$\lim_{x\in D\to y}\mathbb{P}_x\left(T_K > \eta\right) = 0$$

(This proof is the same as the proof of lemma 7.9) and so there exists $\delta' > 0$ such that

$$\mathbb{P}_x\left(T_K > \eta\right) < \epsilon \quad \forall x \in D \cap B\left(y,\delta'\right)$$

Let $x \in D \bigcap B\left(y, \delta' \wedge \frac{\delta}{2}\right)$. Then

$$\mathbb{P}_x\left(\sup_{t\le\eta}|B_t - x| \ge \frac{\delta}{2}\right) = \mathbb{P}_0\left(\sup_{t\le\eta}|B_t| \ge \frac{\delta}{2}\right) < \epsilon$$

and so

$$\begin{aligned}
|u(x)-g(y)| \le\, & \mathbb{E}_x\left[\left|g\left(B_{T_K}\right)-g(y)\right| 1_{\{T_K\le\eta\}}\right]+ \\
& \mathbb{E}_x\left[\left|g\left(B_{T_K}\right)-g(y)\right| 1_{\{\eta<T_K<\infty\}}\right]+(g(y)+\alpha)\mathbb{P}_x\left(T_K=\infty\right) \\
\le\, & \mathbb{E}_x\left[\left|g\left(B_{T_K}\right)-g(y)\right| 1_{\{T_K\le\eta\}} 1_{\left\{\sup_{t\le\eta}|B_t-x|<\frac{\delta}{2}\right\}}\right]+ \\
& 2M\mathbb{P}_x\left(\sup_{t\le\eta}|B_t-x|\ge\frac{\delta}{2}\right)+ \\
& \mathbb{E}_x\left[\left|g\left(B_{T_K}\right)-g(y)\right| 1_{\{\eta<T_K<\infty\}}\right]+(g(y)+\alpha)\mathbb{P}_x\left(T_K=\infty\right) \\
\le\, & \epsilon+2M\epsilon+2M\mathbb{P}_x\left(\eta<T_K<\infty\right)+(g(y)+\alpha)\mathbb{P}\left(T_K=\infty\right) \\
\le\, & \epsilon+2M\epsilon+(3M+\alpha)\mathbb{P}_x\left(T_K>\eta\right)<\epsilon+2M\epsilon+(3M+\alpha)\epsilon
\end{aligned}$$

□

(e) Show that, for every $x\in D$,

$$u(x)=\mathbb{E}_x\left[g\left(B_{T_K}\right)\mathbb{1}_{\{T_K<\infty\}}\right]+\alpha\mathbb{P}_x\left(T_K=\infty\right)$$

Proof. Note that $\lim_{t\to\infty}|B_t|=\infty$ [Gal16, Thm. 7.17], but we have shown in (b) that $u(x)\xrightarrow{|x|\to\infty}\alpha$, so $T_{n_k}<\infty$ a.s. for every $k\ge1$. Then, passing to the limit, one has

$$\begin{aligned}
& \mathbb{E}_x\left[u\left(B_{T_{n_k}}\right)\mathbb{1}_{\{T_{n_k}\le T_K\}}\right] \\
=\, & \mathbb{E}_x\left[u\left(B_{T_{n_k}}\right)\mathbb{1}_{\{T_{n_k}\le T_K<\infty\}}\right]+\mathbb{E}_x\left[u\left(B_{T_{n_k}}\right)\mathbb{1}_{\{T_{n_k}<\infty\}\cap\{T_K=\infty\}}\right] \\
\xrightarrow{k\to\infty}\, & 0+\alpha\mathbb{P}_x\left(T_K=\infty\right)
\end{aligned}$$

By parts (a) and (b), one has

$$\begin{aligned}
u(x) &= \lim_{k\to\infty}\mathbb{E}_x\left[g\left(B_{T_K}\right)\mathbb{1}_{\{T_K\le T_{n_k}\}}\right]+\lim_{k\to\infty}\mathbb{E}_x\left[u\left(B_{T_{n_k}}\right)\mathbb{1}_{\{T_{n_k}\le T_K\}}\right] \\
&= \mathbb{E}_x\left[g\left(B_{T_K}\right)\mathbb{1}_{\{T_K<\infty\}}\right]+\alpha\mathbb{P}_x\left(T_K=\infty\right)
\end{aligned}$$

as desired. □

(f) Assume that D satisfies the exterior cone condition at every $y\in\partial D$ (this is defined in the same way as when D is bounded). Show that, for any choice of $\alpha\in\mathbb{R}$, the formula of part (c) gives a solution of the problem (P).

Proof. By [Gal16, Prop. 7.7.ii], as the entire function inside the expectation is bounded measurable, $u(x)$ is harmonic. Now it suffices to show that

$$\lim_{x\in D\to y}u(x)=g(y)$$

for every $y \in \partial D$. Since then, by [Gal16, Thm. 7.8] we're done. Note that in the theorem D bounded is only required for finite hitting time, so it actually does not affect its validity here.

The remaining follows similarly [Gal16, Thm. 7.8]. Let M be such that $|g(z)| \leq M$ for every $z \in K$. Let $\epsilon > 0$. Choose $\delta > 0$ such that $|g(z) - g(y)| < \epsilon, \forall z \in K \cap \mathcal{B}(y, \delta)$, and choose $\eta > 0$ such that $\mathbb{P}_0 \left(\sup_{t \leq \eta} |B_t| \geq \delta/2\right) < \epsilon$. By [Gal16, Lemma 7.9], one has $\lim_{x \in D \to y} \mathbb{P}_x (T_K > \eta) = 0$, so there exists $\delta_1 > 0$ such that

$$\mathbb{P}_x (T_K > \eta) < \epsilon, \quad \forall x \in D \cap \mathcal{B}(y, \delta_1)$$

Let $x \in D \cap \mathcal{B}(y, \delta_1 \wedge \delta/2)$. Then

$$\mathbb{P}_x \left(\sup_{t \leq \eta} |B_t - x| \geq \frac{\delta}{2}\right) = \mathbb{P}_0 \left(\sup_{t \leq \eta} |B_t| \geq \frac{\delta}{2}\right) < \epsilon,$$

so that

$$\begin{aligned}
&\mathbb{E}_x \left[|g(B_{T_K}) - g(y)| \mathbb{1}_{\{T_K \leq \eta\}}\right] \\
=&\mathbb{E}_x \left[|g(B_{T_K}) - g(y)| \mathbb{1}_{\{T_K \leq \eta\}} \mathbb{1}_{\{\sup_{t \leq \eta} |B_t - x| < \frac{\delta}{2}\}}\right] + \\
&\mathbb{E}_x \left[|g(B_{T_K}) - g(y)| \mathbb{1}_{\{T_K \leq \eta\}} \mathbb{1}_{\{\sup_{t \leq \eta} |B_t - x| \geq \frac{\delta}{2}\}}\right] \\
\leq&\mathbb{E}_x \left[|g(B_{T_K}) - g(y)| \mathbb{1}_{\{T_K \leq \eta\}} \mathbb{1}_{\{\sup_{t \leq \eta} |B_t - x| < \frac{\delta}{2}\}}\right] + \\
&2M\mathbb{P}_x \left(\sup_{t \leq \eta} |B_t - x| \geq \delta/2\right) \\
\leq&\epsilon + 2M\epsilon
\end{aligned}$$

Hence, one has

$$\begin{aligned}
|u(x) - g(y)| \leq &\mathbb{E}_x \left[|g(B_{T_K}) - g(y)| \mathbb{1}_{\{T_K \leq \eta\}}\right] + \\
&\mathbb{E}_x \left[|g(B_{T_K}) - g(y)| \mathbb{1}_{\{\eta < T_K < \infty\}}\right] + (g(y) + \alpha)\mathbb{P}_x (T_K = \infty) \\
&\leq \epsilon + 2M\epsilon + 2M\mathbb{P}_x (\eta < T_K < \infty) + (g(y) + \alpha)\mathbb{P} (T_K = \infty) \\
&\leq \epsilon + 2M\epsilon + (3M + \alpha)\mathbb{P}_x (T_K > \eta) \\
&< \epsilon + 2M\epsilon + (3M + \alpha)\epsilon \xrightarrow{\epsilon \downarrow 0} 0
\end{aligned}$$

as desired. □

12 Convergence of Probability Measures

Remark. For the following, you're allowed to use that *if (E, d) is a separable metric space then the following defines a metric on the set of probability measures on E:*

$$d_E(\mathbb{P}, \mathbb{Q}) = \inf\{\varepsilon > 0 : \mathbb{P}(F) \leq \mathbb{Q}(F_\varepsilon) + \varepsilon \textit{ for all } F \textit{ closed}\}$$

Bibliography

[Bas11] Richard F. Bass. *Stochastic Processes.* Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2011. doi:10.1017/CBO9780511997044.

[Bil86] Patrick Billingsley. *Probability and Measure.* John Wiley and Sons, second edition, 1986.

[Bil99] Patrick Billingsley. *Convergence of Probability Measures.* Wiley-Interscience, 2nd edition, 1999. doi:10.1002/9780470316962.

[BW21] Rabi Bhattacharya and Edward C. Waymire. *Random Walk, Brownian Motion, and Martingales.* Graduate Texts in Mathematics. Springer Cham, 2021. doi:10.1007/978-3-030-78939-8.

[Dur19] Rick Durrett. *Probability: Theory and Examples.* Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 5 edition, 2019. doi:10.1017/9781108591034.

[Eva98] Lawrence C. Evans. *Partial Differential Equations.* Graduate Studies in Mathematics. American Mathematical Society, 2nd edition, 1998. doi:10.1090/gsm/019.

[Gal16] Jean-François Le Gall. *Brownian Motion, Martingales, and Stochastic Calculus.* Graduate Texts in Mathematics. Springer, 2016. doi:10.1007/978-3-319-31089-3.

[GS20] Geoffrey Grimmett and David Stirzaker. *Probability and random processes.* Oxford University Press, fourth edition, 2020.

[Kal21] Olav Kallenberg. *Foundations of Modern Probability.* Probability Theory and Stochastic Modelling. Springer Cham, 3rd edition, 2021. doi:10.1007/978-3-030-61871-1.

[KS98] Ioannis Karatzas and Steven E. Shreve. *Brownian Motion and Stochastic Calculus.* Graduate Texts in Mathematics. Springer New York, NY, 1998. doi:10.1007/978-1-4612-0949-2.

[KS07] Leonid Koralov and Yakov G. Sinai. *Theory of Probability and Random Processes*. Universitext. Springer Berlin, Heidelberg, 2007. doi:10.1007/978-3-540-68829-7.

[Var01] S. R. S. Varadhan. *Probability Theory*. Courant Lecture Notes. American Mathematical Society, 2001. doi:10.1090/cln/007.

[Var07] S. R. S. Varadhan. *Stochastic Processes*. Courant Lecture Notes. American Mathematical Society, 2007. doi:10.1090/cln/016.

[Wal12] John B. Walsh. *Knowing the Odds: An Introduction to Probability*. Graduate Studies in Mathematics. American Mathematical Society, 2012. doi:10.1090/gsm/139.

[Win12] Andreas Winkelbauer. Moments and absolute moments of the normal distribution, 2012. doi:10.48550/ARXIV.1209.4340.

[ZMGR15] Daniel Zwillinger, Victor Moll, I.S. Gradshteyn, and I.M. Ryzhik. Definite integrals of elementary functions. In *Table of Integrals, Series, and Products*, pages 249–519. Academic Press, Boston, eighth edition, 2015. doi:10.1016/B978-0-12-384933-5.00003-5.