Exercises and Solutions for **An Introduction to Stochastic Differential Equations**

Zehao Dou

September 2021

Abstract

Problem 1 Show, using the formal manipulations for Ito's chain rule, that

$$Y(t) := e^{W(t)-\frac{t}{2}}$$

solves the stochastic differential equation

$$\begin{cases} dY &= YdW \\ Y(0) &= 1. \end{cases}$$

(Hint: If $X(t) := W(t) - \frac{t}{2}$, then $dX = -\frac{dt}{2} + dW$.)

Solution of Problem 1 According to the Ito's chain rule: since function $X(t) := W(t) - \frac{t}{2}$ satisfies the equation $X(t) = -\frac{dt}{2} + dW$, for $Y(t) = \exp(X(t))$, it holds that:

$$dY = \left(e^X \cdot \left(-\frac{1}{2}\right) + \frac{1}{2} \cdot e^X\right) dt + e^X dW = 0 + YdW = YdW.$$

Also, it's easy to verify that $Y(0) = 1$. Therefore, the solution

$$Y(t) := e^{W(t)-\frac{t}{2}}$$

actually solves the SDE we want.

Problem 2 Show that,

$$S(t) = s_0 e^{\sigma W(t) + \left(\mu - \frac{\sigma^2}{2}\right)t}$$

solves

$$\begin{cases} dS &= \mu S dt + \sigma S dW \\ S(0) &= s_0. \end{cases}$$

Solution of Problem 2 Again we use the Ito's chain rule: for function $X(t) := W(t) + \frac{\mu - \sigma^2/2}{\sigma} t$, we have:

$$dX = dW + \frac{\mu - \sigma^2/2}{\sigma} dt.$$

Therefore, for function $S(t) = s_0 e^{\sigma \cdot X(t)}$, it holds that:

$$\begin{aligned} dS &= \left(\sigma s_0 e^{\sigma X(t)} \cdot \frac{\mu - \sigma^2/2}{\sigma} + \frac{1}{2} \cdot \sigma^2 s_0 e^{\sigma X(t)}\right) dt + \sigma s_0 e^{\sigma X(t)} dW \\ &= \mu s_0 e^{\sigma X(t)} dt + \sigma s_0 e^{\sigma X(t)} dW = \mu S dt + \sigma S dW. \end{aligned}$$

Also, it's easy to verify that $S(0) = s_0$, which comes to our conclusion.

Problem 3 (1) Let (Ω, μ, P) be a probability space and let $A_1 \subseteq A_2 \subseteq \ldots \subseteq A_n \subseteq \ldots$ be events. Show that

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{m\to\infty} P(A_m).$$

(Hint: Look at the disjoint events $B_n := A_{n+1} - A_n$.)

(2) Likewise, show that if $A_1 \supseteq A_2 \supseteq \ldots \supseteq A_n \supset \ldots$, then

$$P\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{m\to\infty} P(A_m).$$

Solution of Problem 3

(1) Consider the disjoint events $B_n = A_{n+1} - A_n$, then we know that these events are disjoint since for $\forall m < n$, it holds that:

$$B_m = A_{m+1} - A_m \subseteq A_{m+1} \subseteq A_n,\ A_{n+1} - A_n = B_n \cap A_n = \varnothing,$$

so we can conclude that $B_m \cap B_n = \varnothing$, which means $\{B_n\}$ are disjoint events. Also, we have $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=0}^{\infty} B_n$. It is because for $\forall x \in \bigcup_{n=1}^{\infty} A_n$, there exists a smallest k such that $x \in A_k \Rightarrow x \in A_k - A_{k-1} = B_k \subseteq \bigcup_{n=0}^{\infty} B_n$, which leads to

$$\bigcup_{n=1}^{\infty} A_n \subseteq \bigcup_{n=0}^{\infty} B_n.$$

On the other hand, for $\forall x \in \bigcup_{n=0}^{\infty} B_n$, there exists k such that $x \in B_k \subseteq A_{k+1} \subseteq \bigcup_{n=1}^{\infty} A_n$, which leads to

$$\bigcup_{n=1}^{\infty} A_n \supseteq \bigcup_{n=0}^{\infty} B_n.$$

Therefore, we conclude that $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=0}^{\infty} B_n$. Also, notice that $A_m = B_0 \cup B_1 \cup \ldots B_{m-1}$, which leads to

$$P(A_m) = \sum_{n=0}^{m-1} P(B_m).$$

Here, we use the countable additivity of $P(\cdot)$ under disjoint events $\{B_n\}$. Finally, it holds that:

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = P\left(\bigcup_{n=0}^{\infty} B_n\right) = \sum_{n=0}^{\infty} P(B_n) = \lim_{m\to\infty} \sum_{n=0}^{m-1} P(B_n) = \lim_{m\to\infty} P(A_m)$$

In the final step, we use the fact that the sequence $p_m := \sum_{n=0}^{m-1} P(B_n) = P(A_m)$ is non-decreasing and upper bounded by 1, and therefore it converges.

(2) Similarly, we have $A_1^c \subseteq A_2^c \subseteq \ldots \subseteq A_n^c \subseteq \ldots$. By using the conclusion of (1), we have:

$$P\left(\bigcap_{n=1}^{\infty} A_n\right) = 1 - P\left(\bigcup_{n=1}^{\infty} A_n^c\right) = 1 - \lim_{m\to\infty} P(A_m^c) = \lim_{m\to\infty} P(A_m),$$

which comes to our conclusion.

Problem 4 Let Ω be any set and $\mathcal{A}$ any collections of subsets of Ω. Show that there exists a unique smallest σ-algebra μ of subsets of Ω containing $\mathcal{A}$. We call μ the σ-algebra generated by $\mathcal{A}$. (Hint: Take the intersection of all the σ-algebras containing $\mathcal{A}$.)

Solution of Problem 4 Obviously, there exists a σ-algebra containing $\mathcal{A}$ (for example, 2^Ω itself). Now, we take the intersection of all the σ-algebras containing $\mathcal{A}$, denoted by μ. In order to prove that there exists a unique smallest σ-algebra μ of subsets of Ω containing $\mathcal{A}$, we only need to prove that the intersection of all σ-algebras is also a σ-algebra of Ω containing $\mathcal{A}$. On one hand, since all the σ-algebras considered by us contain $\mathcal{A}$, so their intersection μ also contains $\mathcal{A}$. On the other hand,

- For all σ-algebra S that contains $\mathcal{A}$, we have $\varnothing \in S$ by the definition of σ-algebra. Therefore, the intersection holds $\varnothing \in \mu$ obviously.
- For $\forall A \in \mu$, we have $A \in S$ for any σ-algebra S that contains $\mathcal{A}$, so it holds that $A^c \in S$. Therefore, we have $A^c \in \mu$.
- For $A_1, A_2, \ldots, \in \mu$, we know that for any σ-algebra S that contains $\mathcal{A}$, $\bigcup_{i=1}^{+\infty} A_i \in S$. Therefore:
$$\bigcup_{i=1}^{+\infty} A_i \in \mu.$$

After combining all the items above, we know that μ is also a σ-algebra of Ω containing $\mathcal{A}$, which comes to our conclusion.

Problem 5 Show that if $A_1, A_2, \ldots, A_n$ are events, then

$$\begin{aligned} P\left(\bigcup_{i=1}^{n} A_i\right) = &\sum_{i=1}^{n} P(A_i) - \sum_{1\leqslant i<j\leqslant n} P(A_i \cap A_j) \\ &+ \sum_{1\leqslant i<j<k\leqslant n} P(A_i \cap A_j \cap A_k) \\ &- \ldots + (-1)^n P(A_1 \cap A_2 \cap \ldots \cap A_n). \end{aligned}$$

(Hint: Do the case $n = 2$ first and then the general case by induction.)

Solution of Problem 5 We use the method of induction. When $n = 2$, we know that:

$$P(A \cup B) = P(A) + P((A \cup B) - A) = P(A) + P(B - (A \cap B)) = P(A) + P(B) - P(A \cap B).$$

Suppose the condition holds for n, then for $n + 1$, we have:

$$\begin{aligned} &P(A_1 \cup A_2 \cup \ldots \cup A_{n+1}) = P((A_1 \cup A_2 \cup \ldots \cup A_n) \cup A_{n+1}) \\ &= P(A_1 \cup A_2 \cup \ldots \cup A_n) + P(A_{n+1}) - P((A_1 \cup A_2 \cup \ldots \cup A_n) \cap A_{n+1}) \\ &= P(A_{n+1}) + \sum_{i=1}^{n} P(A_i) - \sum_{1\leqslant i<j\leqslant n} P(A_i \cap A_j) + \ldots + (-1)^n P(A_1 \cap A_2 \cap \ldots \cap A_n) \end{aligned}$$

$$
\begin{aligned}
&- P\left((A_1 \cap A_{n+1}) \cup (A_2 \cap A_{n+1}) \cup \ldots \cup (A_n \cap A_{n+1})\right) \\
&= P(A_{n+1}) + \sum_{i=1}^{n} P(A_i) - \sum_{1 \leqslant i < j \leqslant n} P(A_i \cap A_j) + \ldots + (-1)^n P(A_1 \cap A_2 \cap \ldots \cap A_n) \\
&- \sum_{i=1}^{n} P(A_i \cap A_{n+1}) + \sum_{1 \leqslant i < j \leqslant n} P((A_i \cap A_{n+1}) \cap (A_j \cap A_{n+1})) - \ldots \\
&- (-1)^n P((A_1 \cap A_{n+1}) \cap (A_2 \cap A_{n+1}) \cap \ldots \cap (A_n \cap A_{n+1})) \\
&= \sum_{i=1}^{n+1} P(A_i) - \sum_{1 \leqslant i < j \leqslant n+1} P(A_i \cap A_j) + \ldots + (-1)^{n+1} P(A_1 \cap A_2 \cap \ldots \cap A_{n+1}),
\end{aligned}
$$

which completes the induction and comes to our conclusion.

Problem 6 Let $X = \sum_{i=1}^{k} a_i \chi_{A_i}$ be a simple random variable, where the real numbers a_i are distinct, the events A_i are pairwise disjoint, and $\Omega = \bigcup_{i=1}^{k} A_i$. Let $\mu(X)$ be the σ-algebra generated by X.
(1) Describe precisely which sets are in $\mu(X)$.
(2) Suppose the random variable Y is $\mu(X)$-measurable. Show that Y is constant on each set A_i.
(3) Show that therefore Y can be written as a function of X.

Solution of Problem 6
(1) Since the real numbers a_i are distinct, we can assume that

$$a_1 < a_2 < \ldots < a_k$$

without loss of generality. We are going to prove that $\mu(X) = \sigma(\{A_i : \ i \in [k]\})$ which is the σ-algebra generated by the pairwise disjoint events A_i. On one hand, notice that: $\forall i \in [k]$,

$$A_i = \{X = a_i\} \in \mu(X) \ \Rightarrow \ \sigma(\{A_i : \ i \in [k]\} \subseteq \mu(X).$$

On the other hand, for any $B \in \mathcal{B}(\mathbb{R})$, we have:

$$\{X \in B\} = \bigcup_{a_i \in B} A_i \in \sigma(\{A_i : \ i \in [k]\}),$$

which leads to:

$$\mu(X) = \sigma\left(\{X \in B\}\right) \subseteq \sigma(\{A_i : \ i \in [k]\}).$$

To sum up, $\mu(X) = \sigma(\{A_i : \ i \in [k]\})$, which means all the sets in $\mu(X)$ are $\bigcup_{i \in S} A_i$ for some $S \subseteq [k]$.

(2) Random variable Y is $\mu(X)$-measurable. Then, for any $B \in \mathcal{B}(\mathbb{R})$, it holds that event $\{Y \in B\} \in \mu(X)$. From the conclusion of (1), we know that

$$\mu(X) = \left\{ \bigcup_{i \in S} A_i : \ S \subseteq [k] \right\}.$$

Now we suppose that Y is not constant on each set A_i. Without loss of generality, we assume Y is not constant on A_1, then there exists $x, y \in A_1$ such that $Y(x) < Y(y)$. Consider the event $E = \{Y = Y(x)\} \in \mu(X)$, then $x \in E$ but $y \notin E$. Since E must be the union of some A_i-s, from $x \in E$, we know that $A_1 \subseteq E$. From $y \notin E$, we know that $A_1 \cap E = \varnothing$, which contradict with each other. Therefore, Y must be constant on each set A_i.

(3) According to the conclusion of (2), we can assume that Y equals to constant b_i on the set A_i. Then:

$$Y = \sum_{i=1}^{k} b_i \chi_{A_i}.$$

Therefore, Y can be written as $f(X)$ where $f(a_i) = b_i$ for $\forall i \in [k]$, which comes to our conclusion.

Problem 7 Verify:

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}, \quad \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x e^{-\frac{(x-m)^2}{2\sigma^2}} dx = m,$$

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} (x-m)^2 e^{-\frac{(x-m)^2}{2\sigma^2}} dx = \sigma^2.$$

Solution of Problem 7 For the first conclusion, we consider the following equation: denote $S = \int_{-\infty}^{\infty} e^{-x^2} dx$, then:

$$\begin{aligned} S^2 &= \int_{-\infty}^{\infty} e^{-x^2} dx \cdot \int_{-\infty}^{\infty} e^{-y^2} dy = \int_{\mathbb{R}^2} e^{-(x^2+y^2)} dx dy \\ &= \int_{0}^{+\infty} \int_{-\pi}^{\pi} e^{-r^2} r \cdot dr d\theta = \pi \cdot \int_{0}^{+\infty} e^{-r^2} \cdot 2r dr = \pi \cdot \int_{0}^{+\infty} e^{-r^2} dr^2 = \pi. \end{aligned}$$

It is obvious that $S > 0$, therefore $S = \sqrt{\pi}$. For the following two equations, let $x = m + \sqrt{2}\sigma y$, then:

$$\begin{aligned} & \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x e^{-\frac{(x-m)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} (m + \sqrt{2}\sigma y) e^{-y^2} dy \\ = \ & \frac{m}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-y^2} dy + \frac{\sqrt{2}\sigma}{\sqrt{\pi}} \int_{-\infty}^{\infty} y e^{-y^2} dy = m + 0 = m. \end{aligned}$$

Here, we use the conclusion of $\int_{-\infty}^{\infty} e^{-y^2} dy = \sqrt{\pi}$ and the fact that ye^{-y^2} is an odd function, whose integral over $\mathbb{R}$ must be 0.

$$\begin{aligned} & \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} (x-m)^2 e^{-\frac{(x-m)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} 2\sigma^2 y^2 e^{-y^2} dy \\ = \ & \frac{\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{\infty} y \cdot 2y e^{-y^2} dy = \frac{\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{\infty} -y de^{-y^2} = \frac{\sigma^2}{\sqrt{\pi}} \cdot \left(-y e^{-y^2} \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-y^2} dy \right) \\ = \ & \frac{\sigma^2}{\sqrt{\pi}} \cdot \sqrt{\pi} = \sigma^2. \end{aligned}$$

Here, we again use the equation $\int_{-\infty}^{\infty} e^{-y^2} dy = \sqrt{\pi}$.

Problem 8 Suppose A and B are independent events in some probability space. Show that A^c and B are independent. Likewise, show that A^c and B^c are independent.

Solution of Problem 8 Since A and B are independent events, we have:

$$P(A \cap B) = P(A)P(B).$$

Therefore, we have the following two equations:

$$P(A^c \cap B) = P(B) - P(A \cap B) = P(B) - P(A)P(B) = (1 - P(A))P(B) = P(A^c)P(B).$$

$$\begin{aligned} P(A^c \cap B^c) &= 1 - P(A \cup B) = 1 - P(A) - P(B) + P(A \cap B) = 1 - P(A) - P(B) + P(A)P(B) \\ &= (1 - P(A))(1 - P(B)) = P(A^c) \cdot P(B^c). \end{aligned}$$

These two equations conclude that A^c and B are independent. Also, A^c and B^c are independent.

Problem 9 Suppose we have three cards: one is red on both sides, one is red on one side and white on the other side, and one is white on both sides.
(1) Pick a card and then one of its sides at random. What is the probability it is red?
(2) Given that the side of the card is red, what is the probability that the other side is red?

Solution of Problem 9
(1) We denote R as red and W as white, then:

$$P(R) = P(RR) \cdot 1 + P(RB) \cdot \frac{1}{2} + P(BB) \cdot 0 = \frac{1}{3} + \frac{1}{6} = \frac{1}{2}.$$

The probability of red side is $\frac{1}{2}$.

(2) By using Bayes' formula, we have:

$$P(RR|R) = \frac{P(RR) \cdot P(R|RR)}{P(RR) \cdot P(R|RR) + P(RB) \cdot P(R|RB) + P(BB) \cdot P(R|BB)} = \frac{1/3}{1/3 + 1/6 + 0} = \frac{2}{3}.$$

The probability for the other side to be red is $\frac{2}{3}$.

Problem 10 Suppose that $A_1, A_2, \ldots, A_m$ are disjoint events, each of positive probability, such that $\Omega = \bigcup_{j=1}^m A_j$. Prove Bayes' formula:

$$P(A_k|B) = \frac{P(B|A_k)P(A_k)}{\sum_{j=1}^m P(B|A_j)P(A_j)}$$

for $k = 1, 2, \ldots, m$ provided $P(B) > 0$.

Solution of Problem 10 Since $A_1, A_2, \ldots, A_m$ all have positive probabilities, so:

$$P(B|A_j)P(A_j) = P(B \cap A_j).$$

Also, they are disjoint and $\bigcup_{j=1}^m A_j = \Omega$, which leads to $\{B \cap A_j\}_{j \in [m]}$ are also disjoint and $\bigcup_{j=1}^m (B \cap A_j) = B \cap \Omega = B$. Therefore:

$$\frac{P(B|A_k)P(A_k)}{\sum_{j=1}^m P(B|A_j)P(A_j)} = \frac{P(B \cap A_k)}{\sum_{j=1}^m P(B \cap A_j)} = \frac{P(B \cap A_k)}{P(B)} = P(A_k|B),$$

which comes to our conclusion. Here, we used the additivity of probability measures.

Problem 11 During one fall semester 105 women applied to Miskatonic University, of whom 76 were accepted, and 400 men applied, of whom 230 were accepted. During the subsequent spring semester, 300 women applied, of whom 100 were accepted, and 112 men applied, of whom 21 were accepted. Calculate numerically:
(a) the probability of a female applicant being accepted during the fall,
(b) the probability of a male applicant being accepted during the fall,
(c) the probability of a female applicant being accepted during the spring,
(d) the probability of a male applicant being accepted during the spring.
Consider now the total applicant pool for both semesters together and calculate.
(e) the probability of a female applicant being accepted,
(f) the probability of a male applicant being accepted.
Are the university's admission policies biased towards females or towards males?

Solution of Problem 11
(a) $76/105 = 0.7238$.
(b) $230/400 = 0.5750$.
(c) $100/300 = 0.3333$.
(d) $21/112 = 0.1875$.
(e) $(100 + 76)/(300 + 105) = 0.4346$.
(f) $(230 + 21)/(400 + 112) = 0.4902$.

Although in each semester, the probability of acceptance of females is higher than males, the university's admission policies biased towards males.

Problem 12 Let X be a real-valued, $\mathcal{N}(0, 1)$ random variable, and set $Y := X^2$. Calculate the density g of the distribution function for Y.
(Hint: You must find g so that $P(-\infty < Y \leqslant a) = \int_{-\infty}^{a} g dy$ for all a.)

Solution of Problem 12 Notice that $Y = X^2 \geqslant 0$ always holds, which means $P(Y < 0) = 0$. For $\forall t \geqslant 0$, we have:

$$P(-\infty < Y \leqslant t) = P(X^2 \leqslant t) = P(-\sqrt{t} \leqslant X \leqslant \sqrt{t}) = \int_{-\sqrt{t}}^{\sqrt{t}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx = 2\phi(\sqrt{t}).$$

where $\phi(u) = \int_0^u \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx$. Then: $\phi'(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right)$. Now, we can conclude that, the density function for distribution Y is:

$$g(t) = \frac{d}{dt} 2\phi(\sqrt{t}) = \frac{\frac{1}{\sqrt{2\pi}} \cdot \exp(-t/2)}{\sqrt{t}} = \frac{\exp(-t/2)}{\sqrt{2\pi t}} \qquad (t \geqslant 0).$$

To sum up, the density function is:

$$g(t) = \frac{\exp(-t/2)}{\sqrt{2\pi t}} \cdot \mathbb{I}[t > 0].$$

Problem 13 Take $\Omega = [0,1] \times [0,1]$, with $\mathcal{U}$ the Borel sets and P Lebesgue measure. Let $g : [0,1] \to \mathbb{R}$ be a continuous function.
Define the random variables

$$X_1(\omega) := g(x_1), X_2(\omega) := g(x_2) \ \text{ for } \omega = (x_1, x_2) \in \Omega.$$

Show that X_1 and X_2 are independent and identically distributed.

Solution of Problem 13 For $\forall A \in \mathcal{B}(\mathbb{R})$, we have:

$$P(X_1 \in A) = P(g^{-1}(A) \times [0,1]) = P([0,1] \times g^{-1}(A)) = P(X_2 \in A).$$

Here, since g is a continuous function, we know that $g^{-1}(A) \in \mathcal{B}([0,1])$. Therefore, X_1 and X_2 are identically distributed. Next, we are going to prove that they are independently distributed. For $\forall A, B \in \mathcal{B}(\mathbb{R})$, we have:

$$P(X_1 \in A, X_2 \in B) = P(g^{-1}(A) \times g^{-1}(B)) = P'(g^{-1}(A)) \times P'(g^{-1}(B)) = P(X_1 \in A) \cdot P(X_2 \in B),$$

where P' is the Lebesgue measure on $\mathbb{R}$. Therefore, we can conclude that X_1 and X_2 are independently distributed.

Problem 14 Let $f : [0,1] \to \mathbb{R}$ be continuous and define the Bernstein polynomial

$$b_n(x) := \sum_{k=0}^{n} f\left(\frac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k}.$$

Prove that $b_n \to f$ uniformly on $[0,1]$ as $n \to \infty$, by providing the details for the following steps.
(1) Since f is uniformly continuous, for each $\varepsilon > 0$ there exists $\delta(\varepsilon) > 0$ such that $|f(x) - f(y)| \leqslant \varepsilon$ if $|x - y| \leqslant \delta(\varepsilon)$.
(2) Given $x \in [0,1]$, take a sequence of independent random variables X_k such that $P(X_k = 1) = x, P(X_k = 0) = 1 - x$. Write $S_n = X_1 + X_2 + \ldots + X_n$. Then $b_n(x) = \mathbb{E} f\left(\frac{S_n}{n}\right)$.
(3) Therefore

$$\begin{aligned} |b_n(x) - f(x)| &\leqslant \mathbb{E}\left| f\left(\frac{S_n}{n}\right) - f(x) \right| \\ &= \int_A \left| f\left(\frac{S_n}{n}\right) - f(x) \right| dP + \int_{A^c} \left| f\left(\frac{S_n}{n}\right) - f(x) \right| dP, \end{aligned}$$

for $A := \left\{\omega \in \Omega | \, |\frac{S_n}{n} - x| \leqslant \delta(\varepsilon)\right\}$.
(4) Then show that

$$|b_n(x) - f(x)| \leqslant \varepsilon + \frac{2M}{\delta(\varepsilon)^2} V\left(\frac{S_n}{n}\right) = \varepsilon + \frac{2M}{n\delta(\varepsilon)^2} V(X_1),$$

for $M = \max |f|$. Conclude that $b_n \to f$ uniformly.

Solution of Problem 14
(1) For any $\varepsilon > 0, x \in [0,1]$, there exists $\delta_x > 0$ such that $|u - x| < \delta_x \Rightarrow |f(u) - f(x)| \leqslant \varepsilon/2$. Then: $\cup_{x\in[0,1]}(x - \delta_x, x + \delta_x) \supseteq [0,1]$. According to the Heine-Borel Covering Theorem: there exists a finite sub-covering. We denote it by $(a_1, b_1), (a_2, b_2), \ldots, (a_n, b_n)$. Let $\delta = \min_{i,j}\{|a_i - b_j| > 0\}$, then for $\forall x, y \in [0,1]$ such that $|x - y| < \delta$, they belong to the same interval, which leads to $|f(x) - f(y)| \leqslant \varepsilon$, and it comes to our conclusion.

(2) Notice that:

$$\begin{aligned}\mathbb{E}f\left(\frac{S_n}{n}\right) &= f\left(\frac{k}{n}\right) \cdot P(\{\text{There are } k \text{ 1-s and } n-k \text{ 0-s in } X_1, X_2, \ldots, X_n\}) \\ &= f\left(\frac{k}{n}\right) \cdot \binom{n}{k} x^k (1-x)^{n-k} = b_n(x),\end{aligned}$$

which comes to our conclusion.

(3) Notice that for any random variable X, we have:

$$-\mathbb{E}|X| \leqslant \mathbb{E}X \leqslant \mathbb{E}|X| \ \Rightarrow |\mathbb{E}X| \leqslant \mathbb{E}|X|.$$

Therefore:

$$\begin{aligned}|b_n(x) - f(x)| &= |\mathbb{E}[f(S_n/n) - f(x)]| \leqslant \mathbb{E}\left|f\left(\frac{S_n}{n}\right) - f(x)\right| \\ &= \int_A \left|f\left(\frac{S_n}{n}\right) - f(x)\right| dP + \int_{A^c} \left|f\left(\frac{S_n}{n}\right) - f(x)\right| dP,\end{aligned}$$

for $A := \{\omega \in \Omega | \, |\frac{S_n}{n} - x| \leqslant \delta(\varepsilon)\}$.

(4) If $\omega \in A$, we have:$|\frac{S_n}{n} - x| \leqslant \delta(\varepsilon) \Rightarrow \left|f\left(\frac{S_n}{n}\right) - f(x)\right| \leqslant \varepsilon$. Therefore:

$$\int_A \left|f\left(\frac{S_n}{n}\right) - f(x)\right| dP \leqslant \varepsilon \cdot P(A) \leqslant \varepsilon.$$

On the other hand,

$$\begin{aligned}&\int_{A^c} \left|f\left(\frac{S_n}{n}\right) - f(x)\right| dP \leqslant 2M \cdot P(A^c) \leqslant 2M \cdot \mathbb{E}\frac{\left(\frac{S_n}{n} - x\right)^2}{\delta(\varepsilon)^2} \\ &= \frac{2M}{\delta(\varepsilon)^2} V\left(\frac{S_n}{n}\right) = \frac{2M}{n\delta(\varepsilon)^2} V(X_1).\end{aligned}$$

After adding up the two inequalities above, we have:

$$\|b_n - f\|_\infty \leqslant \varepsilon + \frac{2M}{n\delta(\varepsilon)^2} V(X_1).$$

Then we have:

$$\lim_{n\to\infty} \|b_n - f\|_\infty \leqslant \varepsilon.$$

Since $\varepsilon > 0$ can be any positive real number, we have:

$$\lim_{n\to\infty} \|b_n - f\|_\infty = 0,$$

which comes to our conclusion.

Problem 15 Let X and Y be independent random variables, and suppose that f_X and f_Y are the density functions for X, Y. Show that the density function for $X+Y$ is

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(z-y) f_Y(y) dy.$$

(Hint: If $g : \mathbb{R} \to \mathbb{R}$, we have

$$\mathbb{E}[g(X+Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y) g(x+y) dx dy,$$

where $f_{X,Y}$ is the joint density function of X, Y.)

Solution of Problem 15 Notice that,

$$\mathbb{E}[g(X+Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y) g(x+y) dx dy.$$

Let $Z = X+Y$, then:

$$\begin{aligned}
&\int_{-\infty}^{\infty} g(z)\cdot f_Z(z) dz = \mathbb{E}[g(X+Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y) g(x+y) dx dy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_X(x) f_Y(y)\cdot g(x+y) dx dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_Y(y) f_X(z-y) g(z) dz dy \\
&= \int_{-\infty}^{\infty} g(z)\left(\int_{-\infty}^{\infty} f_Y(y) f_X(z-y) dy\right) dz.
\end{aligned}$$

Here, we use the fact that X, Y are independent random variables and we also use the coordinate change of integration. Since g can be any function on Z, we can conclusion that:

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_Y(y) f_X(z-y) dy,$$

which comes to our conclusion.

Problem 16 Let X and Y be two independent positive random variables, each with density

$$f(x) = \begin{cases} e^{-x} & \text{if } x \geqslant 0 \\ 0 & \text{if } x < 0 \end{cases}.$$

Find the density of $X+Y$.

Solution of Problem 16 Notice that for $\forall z \geqslant 0$, we have:

$$\begin{aligned} &P(X+Y \leqslant z) = \mathbb{E}\mathbb{I}[X+Y \leqslant z] = \mathbb{E}_Y[\mathbb{E}\mathbb{I}[X+Y \leqslant z]|Y] = \mathbb{E}_Y P(X \leqslant z-Y) \\ &= \int_0^z e^{-y}dy \cdot \int_0^{z-y} e^{-x}dx = \int_0^z e^{-y}(1-e^{-(z-y)})dy = 1-e^{-z}-ze^{-z} = 1-(z+1)e^{-z}. \end{aligned}$$

After taking derivative over z, we know that the density of $X+Y$ is:

$$g(z) = ze^{-z} \quad (z \geqslant 0).$$

Problem 17 Show that

$$\lim_{n\to\infty} \int_0^1 \int_0^1 \dots \int_0^1 f\left(\frac{x_1+x_2+\ldots+x_n}{n}\right) dx_1 dx_2 \ldots dx_n = f\left(\frac{1}{2}\right)$$

for each continuous function f.

(Hint: Let $X_1, X_2, \ldots, X_n$ be independent random variables, each of which has density function $f_i(x) = 1$ if $0 \leqslant x \leqslant 1$ and $= 0$ otherwise. Then $P\left(\left|\frac{X_1+X_2+\ldots+X_n}{n} - \frac{1}{2}\right| > \varepsilon\right) \leqslant \frac{1}{\varepsilon^2} V\left(\frac{X_1+X_2+\ldots+X_n}{n}\right) = \frac{1}{12\varepsilon^2 n}$.)

Solution of Problem 17 Let $X_1, X_2, \ldots, X_n$ be independent random variables, each of which follows the uniform distribution over $[0,1]$. Then, we only need to prove that:

$$\lim_{n\to\infty} \mathbb{E} f\left(\frac{X_1+X_2+\ldots+X_n}{n}\right) = f(1/2).$$

As we know that:

$$P\left(\left|\frac{X_1+X_2+\ldots+X_n}{n} - \frac{1}{2}\right| > \varepsilon\right) \leqslant \frac{1}{\varepsilon^2} V\left(\frac{X_1+X_2+\ldots+X_n}{n}\right) = \frac{1}{12\varepsilon^2 n}.$$

Since f is continuous function, f is bounded on the closed interval $[0,1]$, which means there exists $L > 0$ such that $|f(x) - f(1/2)| < L$ holds for $\forall x \in [0,1]$. Also, f is continuous at $1/2$. Therefore, for $\forall \varepsilon > 0$, there exists $\delta > 0$ such that $|x - 1/2| < \delta \Rightarrow |f(x) - f(1/2)| < \varepsilon$. Then:

$$P\left(\left|f\left(\frac{X_1+X_2+\ldots+X_n}{n}\right) - f(1/2)\right| > \varepsilon\right) \leqslant P\left(\left|\frac{X_1+X_2+\ldots+X_n}{n} - \frac{1}{2}\right| > \delta\right) \leqslant \frac{1}{12\delta^2 n},$$

which leads to:

$$\left|\mathbb{E} f\left(\frac{X_1+X_2+\ldots+X_n}{n}\right) - f(1/2)\right| \leqslant \varepsilon + \frac{L}{12\delta^2 n}.$$

Let $n \to \infty$:

$$\lim_{n\to\infty} \left| \mathbb{E} f\left(\frac{X_1 + X_2 + \ldots + X_n}{n}\right) - f(1/2) \right| \leqslant \varepsilon.$$

Since ε can be any positive real number, we can finally conclude that:

$$\lim_{n\to\infty} f\left(\frac{X_1 + X_2 + \ldots + X_n}{n}\right) = f(1/2).$$

Problem 18 Prove that:
(1) $\mathbb{E}[\mathbb{E}[X|\mathcal{V}]] = \mathbb{E}[X]$.
(2) $\mathbb{E}[X] = \mathbb{E}[X|\mathcal{W}]$, where $\mathcal{W} = \{\emptyset, \Omega\}$ is the trivial σ-algebra.

Solution of Problem 18 According to the definition of conditional expectation, we know that $\mathbb{E}[X|\mathcal{V}]$ is a random variable on Ω such that $\mathbb{E}[X|\mathcal{V}]$ is $\mathcal{V}$-measurable and $\int_A X dP = \int_A \mathbb{E}[X|\mathcal{V}] dP$ for all $A \in \mathcal{V}$.
(1) $\mathbb{E}[\mathbb{E}[X|\mathcal{V}]] = \int \mathbb{E}[X|\mathcal{V}] d\mu = \int X d\mu = \mathbb{E}[X]$.
(2) Since $\mathcal{W} = \{\emptyset, \Omega\}$ is the trivial σ-algebra, and random variable $\mathbb{E}[X|\mathcal{W}]$ is $\mathcal{W}$-measurable, which means $\mathbb{E}[X|\mathcal{W}]$ is a constant variable. By the conclusion of (1), we have:

$$\mathbb{E}[X|\mathcal{W}] = \mathbb{E}[X],$$

which comes to our conclusion.

Problem 22 Let $W(\cdot)$ be a one-dimensional Brownian motion. Show that

$$\mathbb{E}[W^{2k}(t)] = \frac{(2k)!t^k}{2^k k!} \quad (t > 0).$$

Solution of Problem 22 According to the property of Brownian motions, $W(t) \sim \mathcal{N}(0, t)$. Therefore:

$$\mathbb{E}[W^{2k}(t)] = \mathbb{E}_{z\sim\mathcal{N}(0,1)}(\sqrt{t}z)^{2k} = t^k \cdot \mathbb{E}[z^{2k}] = t^k(2k-1)!! = \frac{(2k)!t^k}{2^k k!},$$

which comes to our conclusion.

Problem 23 Show that if $W(\cdot)$ is an n-dimensional Brownian motion, then so are:
(1) $W(t+s) - W(s)$ for all $s \geqslant 0$.
(2) $cW(t/c^2)$ for all $c > 0$. (Brownian scaling)

Solution of Problem 23
(1) For any $k \in [n]$, $W^k(t)$ is a one-dimensional Brownian motion, then $W^k(t+s) - W^k(s)$ is a Gaussian process and

$$\mathbb{E}[(W^k(u+s) - W^k(s))(W^k(v+s) - W^k(s))] = \min(u+s, v+s) + s - \min(u+s, s) - \min(v+s, s) = \min(u, v),$$

which means $W^k(t+s) - W^k(s)$ is also a one-dimensional Brownian motion. Also $\{W^k(t+s) - W^k(s)\}_{k\in[n]}$ are independent. Therefore, $W(t+s) - W(s)$ is a Brownian motion.

(2) For any $k \in [n]$, $W^k(t)$ is a one-dimensional Brownian motion, then $cW^k(t/c^2)$ is a Gaussian process and

$$\mathbb{E}[cW(u/c^2) \cdot cW(v/c^2)] = c^2 \min(u/c^2, v/c^2) = \min(u, v).$$

So $cW^k(t/c^2)$ is a one-dimensional Brownian motion. Also $\{cW^k(t/c^2)\}_{k\in[n]}$ are independent. Therefore, $cW(t/c^2)$ is a Brownian motion.

Problem 24 Let $W(\cdot)$ be a one-dimensional Brownian motion, and define

$$\overline{W}(t) = \begin{cases} tW\left(\frac{1}{t}\right) & \text{for } t > 0 \\ 0 & \text{for } t = 0 \end{cases}.$$

Show that $\overline{W}(t) - \overline{W}(s)$ is $\mathcal{N}(0, t-s)$ for times $0 \leqslant s \leqslant t$. ($\overline{W}(\cdot)$ also has independent increments and so is a one-dimensional Brownian motion. You do not need to show this.)

Solution of Problem 24 For times $0 \leqslant s \leqslant t$, we know that:

$$\overline{W}(t) - \overline{W}(s) = tW\left(\frac{1}{t}\right) - sW\left(\frac{1}{s}\right) = (t-s)W\left(\frac{1}{t}\right) - s\left(W\left(\frac{1}{s}\right) - W\left(\frac{1}{t}\right)\right).$$

Notice that $0 < \frac{1}{t} \leqslant \frac{1}{s}$, and so $W\left(\frac{1}{t}\right)$ are independent with $W\left(\frac{1}{s}\right) - W\left(\frac{1}{t}\right)$. We know that:

$$W\left(\frac{1}{t}\right) \sim \mathcal{N}(0, 1/t), \quad W\left(\frac{1}{s}\right) - W\left(\frac{1}{t}\right) \sim \mathcal{N}(0, 1/s - 1/t).$$

Therefore, $\overline{W}(t) - \overline{W}(s)$ follows the zero-centered Gaussian distribution with its variance

$$(t-s)^2 \cdot \frac{1}{t} + s^2 \cdot \left(\frac{1}{s} - \frac{1}{t}\right) = t - s.$$

To sum up, we conclude that $\overline{W}(t) - \overline{W}(s) \sim \mathcal{N}(0, t-s)$.

Problem 27 Define $U(t) := e^{-t}W(e^{2t})$, where $W(\cdot)$ is a one-dimensional Brownian motion. Show that

$$\mathbb{E}[U(t)U(s)] = e^{-|t-s|} \quad \text{for all} \quad -\infty < s, t < \infty.$$

Solution of Problem 27 Without loss of generality, we can assume that $s \leqslant t$. Then:

$$\begin{aligned}\mathbb{E}[U(t)U(s)] &= e^{-t-s} \cdot \mathbb{E}[W(e^{2t})W(e^{2s})] = e^{-t-s} \cdot \mathbb{E}\left[W(e^{2s})^2 + W(e^{2s}) \cdot (W(e^{2t}) - W(e^{2s}))\right] \\ &= e^{-t-s} \cdot (e^{2s} + 0) = e^{-t+s} = e^{-|t-s|},\end{aligned}$$

which comes to our conclusion. Here, we use the fact that $W(x) \sim \mathcal{N}(0, x)$ for all $x \geqslant 0$ and $W(x) - W(y) \sim \mathcal{N}(0, x-y)$ for all $0 < y < x$ which is independent of $W(y)$.

Problem 28 Let $W(\cdot)$ be a one-dimensional Brownian motion. Show that

$$\lim_{m\to\infty} \frac{W(m)}{m} = 0 \quad \text{almost surely.}$$

(Hint: Fix $\varepsilon > 0$ and define the event $A_m := \left\{\left|\frac{W(m)}{m}\right| \geqslant \varepsilon\right\}$. Then $A_m = \{|X| \geqslant \sqrt{m}\varepsilon\}$ for the $\mathcal{N}(0,1)$ random variable $X = \frac{W(m)}{\sqrt{m}}$. Apply the Borel-Cantelli Lemma.)

Solution of Problem 28 Notice that, the event $\left\{\lim_{m\to\infty}\frac{W(m)}{m}\neq 0\right\}$ is equivalent to the event that there exists $\delta>0$ such that there are infinitely many m-s satisfy $\left|\frac{W(m)}{m}\right|>\delta$, which means:

$$E:=\left\{\lim_{m\to\infty}\frac{W(m)}{m}\neq 0\right\}=\lim_{k\to\infty}E_k$$

where

$$E_k=\left\{\text{There are infinitely many } m\text{-s satisfy }\left|\frac{W(m)}{m}\right|>\frac{1}{k}\right\}.$$

Notice that

$$E_k=\lim\sup_{m\to\infty}\left\{\left|\frac{W(m)}{m}\right|>\frac{1}{k}\right\}.$$

According to the fact that $W(m)\sim\mathcal{N}(0,m)$, we have:

$$P\left(\left|\frac{W(m)}{m}\right|>\frac{1}{k}\right)=P\left(\left|\frac{W(m)}{\sqrt{m}}\right|>\frac{\sqrt{m}}{k}\right)=\widetilde{\Phi}\left(\frac{\sqrt{m}}{k}\right)\leqslant\exp\left(-\frac{m}{2k^2}\right).$$

Then their sum:

$$\sum_{m=1}^{\infty}P\left(\left|\frac{W(m)}{m}\right|>\frac{1}{k}\right)\leqslant\sum_{m=1}^{\infty}\exp\left(-\frac{m}{2k^2}\right)=\frac{1}{1-\exp(-1/2k^2)},$$

which means the infinite sum actually converges. By using Borel-Cantelli Lemma, we know that:

$$P(E_k)=\left(\lim\sup_{m\to\infty}\left\{\left|\frac{W(m)}{m}\right|>\frac{1}{k}\right\}\right)=0,$$

and then:

$$P(E)=\lim_{k\to\infty}P(E_k)=0\quad\Rightarrow\quad P(\overline{E})=1,$$

which comes to our conclusion.

Problem 29 (1) Let $0<\gamma\leqslant 1$. Show that $f:[0,T]\to\mathbb{R}^n$ is uniformly Holder continuous with exponent γ, it is also uniformly Holder continuous with each exponent $0<\delta<\gamma$.
(2) Show that $f(t)=t^\gamma$ is uniformly Holder continuous with exponent γ on the interval $[0,1]$.

Solution of Problem 29 (1) Since f is uniformly Holder continuous with exponent γ, we have:

$$\|f(x)-f(y)\|\leqslant K|x-y|^\gamma$$

holds for $\forall x,y\in[0,T]$ and a constant K. Then if $|x-y|\leqslant 1$, we have: $|x-y|^\gamma\leqslant|x-y|^\delta$, which directly leads to

$$\|f(x)-f(y)\|\leqslant K|x-y|^\delta.$$

If $|x-y|>1$, we have: $|x-y|^\gamma=|x-y|^{\gamma-\delta}\cdot|x-y|^\delta\leqslant T^{\gamma-\delta}\cdot|x-y|^\delta$. To sum up, for $\forall x,y\in[0,T]$, it holds that:

$$\|f(x)-f(y)\|\leqslant K\max(1,T^{\gamma-\delta})|x-y|^\delta,$$

which comes to our conclusion.

(2) We only need to prove that for $\forall x, y \in [0,1]$, it uniformly holds that:

$$|x^\gamma - y^\gamma| < K|x-y|^\gamma$$

for some constant K. Without loss of generality, we can assume that $0 \leqslant y \leqslant x \leqslant 1$. Let $\delta = x - y$, then we need to prove that:

$$g(y) := (y+\delta)^\gamma - y^\gamma < K\delta^\gamma$$

for some constant K. Since $\gamma \in (0,1]$, we have:

$$g'(y) = \gamma \cdot \left((y+\delta)^{\gamma-1} - y^{\gamma-1}\right) \leqslant 0.$$

It means g is a decreasing function. Therefore:

$$g(y) \leqslant g(0) = \delta^\gamma$$

holds for $\forall y$. We can make $K = 2$ and it comes to our conclusion.