Chapter 2: Basics of Probability Theory

Exercise 2.1 : Describe the sample space in different random experiments.

(1) Toss a coin with two sides denoted head and tail, and stop when tail appeaers for the first time.

(2) In a set of cards with four colors and 13 numbers, randomly pick up five cards (the order does not matter).

(3) In the unit disk, random trajectories that start from the origin and stop when they hit the boundary of the disk.

We assume that we have a fair coin in (1), and the cards are picked up randomly and uniformly in (2), describe the corresponding probability measure.

Exercise 2.2 (Bertrand's Paradox) **:** The French mathematician Joseph Bertrand (1822-1990) asked a question in his book "Probability Computations" (Calculs des probabilités) published in 1889: given a equilateral triangle and its circumscribed circle, what is the probability that a *randomly chosen* chord on the circle is longer than a side of the equilateral triangle? At the same time, he suggested three methods to compute, while obtaining three different results.

(1) The two endpoints of the chord are chosen uniformly at random on the circle.

(2) Choose a point inside the disk uniformly at random then draw the chord going through this point such that the midpoint of the chord is the randomly chosen point.

(3) Choose a point uniformly at random on the circle, draw the radius going through this point, choose a point uniformly at random on the radius and draw the chord which is perpendicular to the radius and going through this point.

In the above three different random experiments, give

(a) the probability space;

(b) the random variable giving the chord length and its density function;

(c) the probability that an chord randomly chosen as such is longer than a side of the equilateral triangle.

Exercise 2.3 : A probability space $(\Omega, \mathcal{A}, \mathbb{P})$ is called atomless (無原子的) if for any $A \in \mathcal{A}$ with $\mathbb{P}(A) > 0$, we can find $B \in \mathcal{A}$ such that $B \subsetneq A$ and $0 < \mathbb{P}(B) < \mathbb{P}(A)$. Otherwise, we say that $(\Omega, \mathcal{A}, \mathbb{P})$ is atomic (有原子的).

(1) Suppose that the measurable space $(\Omega, \mathcal{A})$ is given by $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. Prove that $(\Omega, \mathcal{A}, \mathbb{P})$ is an atomless probability space if and only if $\mathbb{P}$ is an atomless probability measure. (We recall Definition 1.1.12.)

(2) Give respectively an example of an atomless probability space and an atomic probability space.

Given an atomless probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and fix $A \in \mathcal{A}$ with $\mathbb{P}(A) > 0$.

(3) Show that for any $\varepsilon > 0$, we can find $B \in \mathcal{A}$ and $B \subseteq A$ with $0 < \mathbb{P}(B) < \varepsilon$.

(4) Show that if $0 < a < \mathbb{P}(A)$, then there exists $B \in \mathcal{A}$ and $B \subseteq A$ with $\mathbb{P}(B) = a$.

Exercise 2.4 : Fix a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Take $\mathcal{C} \subseteq \mathcal{P}(\Omega)$ satisfying the following conditions,

(a) $\varnothing \in \mathcal{C}$;

(b) if $A \in \mathcal{C}$, then $A^c \in \mathcal{C}$;

(c) if $A, B \in \mathcal{C}$, then $A \cap B \in \mathcal{C}$.

We call such a $\mathcal{C}$ an algebra of sets (集合代數). Additionally, we also assume that $\mathcal{A} := \sigma(\mathcal{C})$.

(1) Prove that for any $B \in \mathcal{A}$, we have $\inf\{\mathbb{P}(A \Delta B) : A \in \mathcal{C}\} = 0$.

(2) Given a bounded random variable X and $\varepsilon > 0$, show that there exists a simple function $Y = \sum_{k=1}^{n} c_k \mathbb{1}_{A_k}$, where $A_k \in \mathcal{C}$, such that $\mathbb{P}(|X - Y| > \varepsilon) < \varepsilon$.

Exercise 2.5 : Let X, Y and Z be real random variables defined on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

(1) Assume that X and Y are equal almost everywhere (with respect to the probability measure $\mathbb{P}$). Prove that X and Y have the same distribution. Does the converse hold?

(2) Assume that X and Y have the same distribution.

  (1) Let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be a Borel function. Prove that $f(X)$ and $f(Y)$ have the same distribution.

  (2) Prove that XZ and YZ does not necessarily have the same distribution.

Exercise 2.6 : Let $X \sim \text{Unif}([0, 1])$ be a uniform random variable on $[0, 1]$.

(1) Let $Y = -\ln(1 - X)$ and find the distribution of Y.

(2) Let $Z = \tan(\pi X - \frac{\pi}{2})$ and find the distribution of Z.

Exercise 2.7 (Question 2.1.21) **:** Construct two bi-dimensional real random variables $X = (X_1, X_2)$ and $X' = (X_1', X_2')$ such that for $j = 1, 2$, X_j and X_j' have the same marginal distribution, where as the distributions of X and X' are not equal.

Exercise 2.8 : We are given a bidimensional real random variable (X, Y) on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Use the method mentioned in Remark 2.1.16 to find the distribution of the following random variables.

(1) Assume that (X, Y) has distribution

$$\lambda\mu e^{-\lambda x - \mu y} \mathbb{1}_{\mathbb{R}_+^2}(x, y)\,\mathrm{d}x\,\mathrm{d}y.$$

Determine the distribution of the random variable $U = \min(X, Y)$.

(2) Assume that (X, Y) has distribution

$$\frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}}\,\mathrm{d}x\,\mathrm{d}y.$$

Determine the distribution of the random variable $\frac{X}{Y}$.

Exercise 2.9 : Let X be a random variable with values in $[0, 1)$ such that its distribution $\mathbb{P}_X$ is the Lebesgue measure on $[0, 1)$. We define the sequence $(X_n)_{n \geqslant 1}$ of random variables as follows,

$$\begin{aligned} X_1 &:= \lfloor 2X \rfloor, \\ X_{n+1} &:= \lfloor 2^{n+1} X - \sum_{k=1}^{n} 2^{n+1-k} X_k \rfloor, \qquad \forall n \geqslant 1. \end{aligned}$$

(1) Check that almost surely all the X_n's take values in the set $\{0, 1\}$. In other words, show that

$$\mathbb{P}\left(\forall n \geqslant 1, X_n \in \{0, 1\}\right) = 1.$$

(Hint: show that the sequence $(X_n)_{n \geqslant 1}$ is the dyadic expansion (二進位展開) of X.)

(2) Describe the σ-algebra generated by X_1, which we denote by $\sigma(X_1)$. (Write down all the elements in this σ-algebra.)

(3) Describe the σ algebra generated by X_1 and X_2, which we denote by $\sigma(X_1, X_2)$. Is this the same as $\sigma(X_2)$, the σ-algebra generated only by X_2? You may want to write down explicitly the elements of these two σ-algebras.

(4) Fix a positive integer $n \geqslant 1$ and find the σ-algebra generated by $X_1, \ldots, X_n$ without writing down all of its elements. How many elements does this σ-algebra contain?

Exercise 2.10 : Compute the expectation and the variance of the probability distributions in Section 2.3.

Exercise 2.11 : Given $p \in (0, 1)$ and if $X \sim \mathrm{Geo}(p)$, then show that for any integers $m, n \geqslant 0$ we have,

$$\mathbb{P}(X \geqslant n) = \mathbb{P}(X \geqslant m + n \,|\, X \geqslant m).$$

We call it the memoryless property (無記憶性質) of the geometric distribution.

Exercise 2.12 : For any $0 \leqslant k \leqslant n$, show that we hvae

$$\lim_{\substack{N,K\to\infty \\ K/N\to p}} f_X(k) = \binom{n}{k} p^k (1-p)^{n-k},$$

where f_X is the mass function of $X \sim \text{Hypergeo}(N, K, n)$. Please give an explanation to this result.

Exercise 2.13 : Consider a sequence $(p_n)_{n\geqslant 1}$ of nonnegative real numbers and a sequence $(X_n)_{n\geqslant 1}$, where X_n is a random variable following the binomial distribution $\text{Bin}(n, p_n)$ for all $n \geqslant 1$. Suppose that we have $np_n \to \lambda$ when n tends to infinity. Then, show that

$$\forall k \in \mathbb{N}_0, \qquad \lim_{n\to\infty} \mathbb{P}(X_n = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Exercise 2.14 : Let X be a random variable with distribution $\mathcal{N}(\mu, \sigma^2)$. Prove the following inequality in two different ways,

$$\forall t > 0, \quad \mathbb{P}(X \geqslant \mu + t) \leqslant \max\Big(\frac{1}{\sqrt{2\pi}} \frac{\sigma}{t}, 1\Big) \exp\big(-\frac{t^2}{2\sigma^2}\big).$$

(1) Use integration by parts.

(2) Use the moment generating function $\mathbb{E}[e^{sX}]$ of X 。

Exercise 2.15 : For any $u > 0$, the density function of the *Cauchy distribution* (柯西分佈) writes,

$$c_u(x) = \frac{1}{\pi} \frac{u}{u^2 + x^2}, \quad x \in \mathbb{R}.$$

(1) Check that for all $u > 0$, c_u defines a probability distribution.

(2) Prove that its expectation does not exist.

(3) Prove that $c_u * c_v = c_{u+v}$.

Exercise 2.16 : Let $Z \sim \mathcal{N}(0, 1)$ and $X := Z^2$. Show that $X \sim \text{Gamma}(\frac{1}{2}, \frac{1}{2})$.

Exercise 2.17 : Let $X \sim \mathcal{N}(0,1)$ be a random variable with standard normal distribution.

(1) Find its moment generating function $M(t) := \mathbb{E}[e^{tX}]$.

(2) Prove by induction that for any nonnegative positive number $k \geqslant 0$, there exists a polynomial P_k satisfying

  (a) $M^{(k)}(t) = P_k(t) e^{t^2/2}$;

  (b) $\deg(P_k) = k$;

  (c) P_k only contains odd degree terms when k is odd; P_k only contains even degree terms when k is even;

  (d) $P_{k+1}(t) = P_k'(t) + tP_k(t)$.

(3) Prove that when k is odd, we have $\mathbb{E}[X^k] = 0$. (Use two different approaches: one using the moment generating function and the other one using integration.)

(4) Compute $\mathbb{E}[X^k]$ for all non-negative integer $k \geqslant 0$.

(5) Use a change of variables and properties of the Γ function to find the same result as in the previous question.

Exercise 2.18 (Estimates on Gaussian integrals) **:** Let f be the density function of the standard normal distribution, that is,

$$\forall x \in \mathbb{R}, \qquad f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

We write F for the distribution function of f, i.e.,

$$\forall x \in \mathbb{R}, \qquad F(x) = \int_{-\infty}^{x} f(y)\,\mathrm{d}y = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{y^2}{2}}\,\mathrm{d}y.$$

(1) Show that we have the following asymptotic behavior when $x \to +\infty$,

$$1 - F(x) \sim \frac{1}{x}\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

(Hint: differentiate the above formula on both the left and the right hand sides.)

(2) Prove the following inequality for $x > 1$,

$$\Big(\frac{1}{x} - \frac{1}{x^3}\Big) f(x) < 1 - F(x) < \frac{f(x)}{x}.$$

(Hint: differentiate again the above formula on both the left and the right hand sides.)

(3) More generally, show that we have the following asymptotic formula for any integer $k \geqslant 0$,

$$\begin{aligned} 1 - F(x) &= f(x)\Big(\sum_{j=0}^{k} (-1)^j \frac{(2j-1)!!}{x^{2j+1}} + o\Big(\frac{1}{x^{2k+1}}\Big)\Big) \\ &= f(x)\Big(\sum_{j=0}^{k} \Big(-\frac{1}{2}\Big)^j \binom{2j}{j} j! + o\Big(\frac{1}{x^{2k+1}}\Big)\Big). \end{aligned}$$

(4) Show that for all $a > 0$, we have the following limit when $x \to +\infty$,

$$\frac{1 - F\left(x + \frac{a}{x}\right)}{1 - F(x)} \to e^{-a}.$$

(5) Show that for all $b > 0$, we have the following limit when $x \to +\infty$,

$$\frac{1 - F(x + b)}{1 - F(x)} \to 0.$$

Exercise 2.19 : On the probability space $(\Omega, \mathcal{A}, \mathbb{P})$, we define the random variable $(X_1, \ldots, X_n)$ with values in $\mathbb{R}^n$ and law

$$\mathbb{1}_{[0,1]^n}(x_1, \ldots, x_n)\,\mathrm{d}x_1 \ldots \mathrm{d}x_n.$$

(1) For any permutation (排列) $\sigma \in \mathcal{S}_n$，let $A_\sigma = \{X_{\sigma_1} < \ldots < X_{\sigma_n}\}$. Construct the random variable $(Y_1, \ldots, Y_n)$ on $(\Omega, \mathcal{A}, \mathbb{P})$ using $(A_\sigma)_{\sigma \in \mathcal{S}_n}$, such that the following statement holds almost surely,

$$Y_1 < \ldots < Y_n \qquad \text{and} \qquad \{Y_1, \ldots, Y_n\} = \{X_1, \ldots, X_n\}.$$

(2) Determine the law of the two vectors of random variables $(Y_1, \ldots, Y_n)$ and $(\frac{Y_1}{Y_2}, \ldots, \frac{Y_{n-1}}{Y_n})$.

Exercise 2.20 (Question 2.4.9) **:** Prove that when X is a d-dimensional real random variable, its covariance matrix K_X is a positive semi-definite (半正定) symmetric matrix. In other words, prove that for all $\lambda = (\lambda_1, \ldots, \lambda_d) \in \mathbb{R}^d$, we have $\lambda K_X \lambda^T \geqslant 0$.

Exercise 2.21 (Question 2.4.10) **:** If A is a matrix of size $n \times d$ and X is a d-dimensional real random variable, define $Y = AX$ and prove that $K_Y = AK_X A^T$.

Exercise 2.22 : Consider the following density function (with respect to the Lebesgue measure),

$$p_0(y) = \frac{1}{\sqrt{2\pi y^2}} \exp\left(-\frac{\ln(y)^2}{2}\right), \quad y > 0.$$

Let Y be a random variable whose density is given by p_0.

(1) Let $X \sim \mathcal{N}(0, 1)$. Find a measurable function $f : \mathbb{R} \to \mathbb{R}$ such that $Y \overset{(d)}{=} f(X)$.

(2) Compute the n-th moment of Y for any non-negative integer $n \geqslant 0$.

(3) Let

$$q(y) = \sin(2\pi \ln(y)) p_0(y).$$

Compute $\int_{\mathbb{R}_{>0}} y^n q(y)\,\mathrm{d}y$ for all non-negative integer $n \geqslant 0$.

(4) For all $a \in \mathbb{R}$, let

$$p_a(y) = p_0(y) + aq(y).$$

Under which condition, p_a is the density function of a probability distribution?

(5) Which conclusion can we deduce from the above questions?

Exercise 2.23 : Given a real sequence $(\mu_k)_{k\geqslant 0}$ satisfying,

$$\limsup_{k\to\infty} \frac{\mu_{2k}^{1/2k}}{2k} = r < \infty.$$

(1) Assume that there exists a distribution function F such that its moments are given by $(\mu_k)_{k\geqslant 0}$. For all $k \geqslant 0$, let $\nu_k = \int |x|^k \,\mathrm{d}F(x)$. Prove that for all $k \geqslant 0$, we have $\nu_{2k+1}^2 \leqslant \mu_{2k}\mu_{2k+2}$ and deduce,

$$\limsup_{k\to\infty} \frac{\nu_k^{1/k}}{k} = r.$$

(2) Let X be a random variable with distribution function F. Write Ψ_X for its characteristic function. Prove that for all $x \in \mathbb{R}$, $t \in \mathbb{R}$ and $n \in \mathbb{N}$, we have,

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

Then, deduce,

$$\forall \xi \in \mathbb{R}, \qquad \left| \Psi_X(\xi + t) - \sum_{k=0}^{n-1} \frac{t^k}{k!} \Psi_X^{(k)}(\xi) \right| \leqslant \frac{|t|^n}{n!}\nu_n.$$

(3) Given $\xi \in \mathbb{R}$, how to choose t so that the following equality holds?

$$\Psi_X(\xi + t) = \sum_{k=0}^{\infty} \frac{t^k}{k!} \Psi_X^{(k)}(\xi).$$

(4) Explain what we have shown here. See Exercise 2.22 and conclude.

Exercise 2.24 (Methods of moments) **:** Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space.

(1) Consider a finite set I and measurable events $A_i \in \mathcal{A}$ for $i \in I$. Show that the following inequality holds,

$$\mathbb{P}\Big(\bigcup_{i\in I} A_i\Big) \leqslant \sum_{i\in I} \mathbb{P}(A_i).$$

(2) Let X be a random variable in $L^2(\Omega, \mathcal{A}, \mathbb{P})$ with expectation $\mu := \mathbb{E}[X] < \infty$. Prove that

$$\mathbb{P}(X \leqslant \mu) > 0 \quad \text{and} \quad \mathbb{P}(X \geqslant \mu) > 0.$$

(3) Let X be a non-constant random variable in $L^2(\Omega, \mathcal{A}, \mathbb{P})$ with non-negative integer values. Show that

$$\mathbb{P}(X = 0) \leqslant \frac{\mathrm{Var}(X)}{\mathbb{E}[X^2]}.$$

Exercise 2.25 : For any subset A of an additive group $(G, +)$, if for all $a, b \in A$, $a + b \notin A$, then we say that A is a sum-free set (無和集合). Given a set of non-zero integers $B = \{b_1, \ldots b_n\}$, we want to prove that there exists a subset $A \subseteq B$ with $|A| > \frac{n}{3}$ such that A is a sum-free set using the methods of moments introduced in Exercise 2.24.

(1) For any positive integer k, prove that $C_k := \{k + 1, \ldots, 2k + 1\}$ is sum-free in the cyclic ring $\mathbb{Z}/(3k + 2)\mathbb{Z}$.

(2) Let $p = 3k + 2$ be a prime number satisfying $p > 2 \max\{|b_i| : 1 \leqslant i \leqslant n\}$. Let X be a uniform random variable on $\{1, \ldots, p - 1\}$. Consider n random variables, $0 \leqslant D_i < p$ such that $D_i \equiv X b_i \pmod p$. Prove that all the D_i's have the same distribution. Determine this distribution and prove that $\mathbb{P}(D_i \in C_k) > \frac{1}{3}$.

(3) Conclude.

Exercise 2.26 (Question 2.4.26) **:** Given a random variable X, prove that the three following statements are equivalent.

(1) X is a sub-Gaussian distribution.

(2) There exist $C, c > 0$ such that $\mathbb{E}[e^{tX}] \leqslant C \exp(ct^2)$.

(3) There exists $C > 0$ such that for all $k \geqslant 1$, we have $\mathbb{E}[|X|^k] \leqslant (Ck)^{k/2}$.

Exercise 2.27 (Question 2.4.27) **:** Let X be a real random variable and $k > 0$. If X is in L^k, prove that when $\lambda \longrightarrow \infty$, we have that

$$\lambda^k \, \mathbb{P}(|X| > \lambda) \longrightarrow 0.$$