Probability Theory

Ikhan Choi

October 23, 2021

Contents

Part I

Probability distributions

Chapter 1

Random variables

1.1 Sample spaces and distributions

sample space of an "experiment" random variables distributions expectation, moments, inequalities

equally likely outcomes coin toss dice roll ball drawing number permutation life time of a light bulb

1.2 Conditional and joint probablity

1.1 (Monty Hall problem)**.** Suppose you're on a game show, and you're given the choice of three doors A, B, and C. Behind one door is a car; behind the others, goats. You pick a door, say A, and the host, who knows what's behind the doors, opens another door, say B, which has a goat. He then says to you, "Do you want to pick door C?" Is it to your advantage to switch your choice?

Proof. Let A, B, and C be the events that a car is behind the doors A, B, and C, respectively. Let X be the event that the challenger picked A, and Y the event that the game host opened B. Note $\{A, B, C\}$ is a partition of the sample space Ω, and X is independent to A, B, and C. Then, $P(A) = P(B) = P(C) = P(X) = 1/3$, and

$$P(Y|X,A) = \frac{1}{2}, \quad P(Y|X,B) = 0, \quad P(Y|X,C) = 1.$$

Therefore,

$$\begin{aligned}
P(C|X,Y) &= \frac{P(X \cap Y \cap C)}{P(X \cap Y)} \\
&= \frac{P(Y|X,C)P(X \cap C)}{P(Y|X,A)P(X \cap A) + P(Y|X,B)P(X \cap B) + P(Y|X,C)P(X \cap C)} \\
&= \frac{1 \cdot \frac{1}{9}}{\frac{1}{2} \cdot \frac{1}{9} + 0 \cdot \frac{1}{9} + 1 \cdot \frac{1}{9}} = \frac{2}{3}.
\end{aligned}$$

Similarly, $P(A|X,Y) = \frac{1}{3}$ and $P(B|X,Y) = 0$. □

1.3 Discrete probability distributions

1.4 Continuous probability distributions

Chapter 2

Measure theory for probability

2.1 Bounded measurable functions

2.1 (Dynkin's π-λ theorem). Let $\mathcal{P}$ be a π-system and $\mathcal{L}$ a λ-system respectively. Denote by $\ell(\mathcal{P})$ the smallest λ-system containing $\mathcal{P}$.

(a) If $A \in \ell(\mathcal{P})$, then $\mathcal{G}_A := \{B : A \cap B \in \ell(\mathcal{P})\}$ is a λ-system.

(b) $\ell(\mathcal{P})$ is a π-system.

(c) If a λ-system is a π-system, then it is a σ-algebra.

(d) If $\mathcal{P} \subset \mathcal{L}$, then $\sigma(\mathcal{P}) \subset \mathcal{L}$.

monotone class

2.2 Polish spaces

2.3 Kolmogorov extension theorem

2.2 (Kolmogorov extension theorem). A *rectangle* is a finite product $\prod_{i=1}^n A_i \subset \mathbb{R}^n$ of measurable $A_i \subset \mathbb{R}$, and *cylinder* is a product $A^* \times \mathbb{R}^{\mathbb{N}}$ where A^* is a rectangle. Let $\mathcal{A}$ be the semi-algebra containing $\varnothing$ and all cylinders in $\mathbb{R}^{\mathbb{N}}$. Let $(\mu_n)_n$ be a sequence of probability measures on $\mathbb{R}^n$ that satisfies *consistency condition*

$$\mu_{n+1}(A^* \times \mathbb{R}) = \mu_n(A^*)$$

for any rectangles $A^* \subset \mathbb{R}^n$, and define a set function $\mu_0 : \mathcal{A} \to [0, \infty]$ by $\mu_0(A) = \mu_n(A^*)$ and $\mu_0(\varnothing) = 0$.

(a) μ_0 is well-defined.

(b) μ_0 is finitely additive.

(c) μ_0 is countably additive if $\mu_0(B_n) \to 0$ for cylinders $B_n \downarrow \varnothing$ as $n \to \infty$.

(d) If $\mu_0(B_n) \geq \delta$, then we can find decreasing $D_n \subset B_n$ such that $\mu_0(D_n) \geq \frac{\delta}{2}$ and $D_n = D_n^* \times \mathbb{R}^{\mathbb{N}}$ for a compact rectangle D_n^*.

(e) If $\mu_0(B_n) \geq \delta$, then $\bigcap_{i=1}^{\infty} B_i$ is non-empty.

Proof. (d) Let $B_n = B_n^* \times \mathbb{R}^{\mathbb{N}}$ for a rectangle $B_n^* \subset \mathbb{R}^{r(n)}$. By the inner regularity of $\mu_{r(n)}$, there is a compact rectangle $C_n^* \subset B_n^*$ such that

$$\mu_0(B_n \setminus C_n) = \mu_{r(n)}(B_n^* \setminus C_n^*) < \frac{\delta}{2^{n+1}}.$$

Let $C_n := C_n^* \times \mathbb{R}^{\mathbb{N}}$ and define $D_n := \bigcap_{i=1}^{n} C_i = D_n^* \times \mathbb{R}^{\mathbb{N}}$. Then,

$$\mu_0(B_n \setminus D_n) \leq \mu_0(\bigcup_{i=1}^{n} B_n \setminus C_i) \leq \mu_0(\bigcup_{i=1}^{n} B_i \setminus C_i) < \frac{\delta}{2},$$

which implies $\mu_0(D_n) \geq \frac{\delta}{2}$.

(e) Take any sequence $(\omega_n)_n$ in $\mathbb{R}^{\mathbb{N}}$ such that $\omega_n \in D_n$. Since each $D_n^* \subset \mathbb{R}^{r(n)}$ is compact and non-empty, by diagonal argument, we have a subsequence $(\omega_k)_k$ such that ω_k is pointwise convergent, and its limit is contained in $\bigcap_{i=1}^{\infty} D_i \subset \bigcap_{i=1}^{\infty} B_n = \varnothing$, which is a contradiction that leads $\mu_0(B_n) \to 0$. $\square$

2.4 Weak convergence

Chapter 3

Independence

3.1 Independent σ-algebras

3.2 Zero-one laws

3.1 (The Kolmogorov zero-one law)**.** Let $X_n : \Omega \to S$ be independent random variables. Let $\mathcal{T}$ be the σ-algebra defined by $\mathcal{T} := \limsup_n \mathcal{F}_n$.

3.2 (The Hewitt-Savage zero-one law)**.** Let $X_n : \Omega \to S$ be i.i.d. random variables.

Part II

Limit theorems

Chapter 4

Laws of large numbers

4.1 Weak laws of large numbers

4.1. Let $X_n : \Omega \to \mathbb{R}$ be uncorrelated random variables.

(a) If $E(X_n) = \mu$ and $E(X_n^2) \lesssim 1$, then $S_n/n \to \mu$ in probability.

(b) If $nP(|X_n| > b_n) \to 0$, $\frac{n}{b_n^2}E(|X|^2\mathbf{1}_{|X|\le b_n}) \to 0$, and $b_n \sim nE(X\mathbf{1}_{|X|\le b_n})$, then $S_n/b_n \to 1$ in probability.

4.2 (Bernstein polynomial)**.** Let $X_n \sim \mathrm{Bern}(x)$ be i.i.d. random variables. Since $S_n \sim \mathrm{Binom}(n,x)$, $E(S_n/n) = x$, $V(S_n/n) = x(1-x)/n$. The L^2 law of large numbers implies $E(|S_n/n - x|^2) \to 0$. Define $f_n(x) := E(f(S_n/n))$. Then, by the uniform continuity $|x-y| < \delta$ implies $|f(x) - f(y)| < \varepsilon$,

$$|f_n(x) - f(x)| \le E(|f(S_n/n) - f(x)|) \le \varepsilon + 2\|f\|P(|S_n/n - x| \ge \delta) \to \varepsilon.$$

4.3 (High-dimensional cube is almost a sphere)**.** Let $X_n \sim \mathrm{Unif}(-1,1)$ be i.i.d. random variables and $Y_n := X_n^2$. Then, $E(Y_n) = \frac{1}{3}$ and $V(Y_n) \le 1$.

large deviation technique: Lp?

4.4 (Coupon collector's problem)**.** $T_n := \inf\{\, t : |\{X_i\}_i| = n\,\}$ Since $X_{n,k} \sim \mathrm{Geo}(1 - \frac{k-1}{n})$, $E(X_{n,k}) = (1 - \frac{k-1}{n})^{-1}$, $V(X_{n,k}) \le (1 - \frac{k-1}{n})^{-2}$. $E(T_n) \sim n\log n$

4.5 (An occupancy problem)**.**

4.6 (The St. Petersburg paradox)**.**

Kolmogorov-Feller

4.2 Strong laws of large numbers

$P(A_n \ i.o.) = 0 \quad iff \quad X_n \to X \ a.s.$ 2.3.14. ? $X_n \to X \ in \ prob \quad iff$ for every subseq, there is further subsequence converging a.s. Thm 2.3.2

infinite monkey

Exercises

Chapter 5

Central limit theorems

Chapter 6

Other limit theorems

large deviation classical summation local limit extreme values

Part III

Stochastic processes

Chapter 7

Martingales

Chapter 8

Markov chains

Chapter 9

Brownian motion

Part IV

Stochastic calculus