What's Probability: A final note for the Week16 lecture

Yunqing LI

July 4, 2020

1 Introduction and Definition

When we talk about probability, it's hard to define what it is. A acceptable way is to mathematically define "expectation" and axioms about expectation, then we can reach more theorems and lemmas based on these axioms.

1.1 Definitions of Ω, H and E

Ω is a set named "sample set", which can be any set.

H is a subset of $\mathbb{R}^{\Omega}$ and should be a linear space.(Formally, It means that $\forall h \in H, c \in \mathbb{R}$ $ch \in H$ and $\forall h_1, h_2 \in H$ $h_1 + h_2 \in H$.) Specially, to ensure one of the of linear exception, H should include 1_{Ω}(a function which maps every element in Ω to 1).

E is a function from H to $\mathbb{R}$.

1.2 Linear Expectation

A linear expectation E should satisfy the following statements:

(1) If $X \in H$ satisfies $X \geq 0$, then $E(X) \geq 0$

(2) If $X \in H$ and $c \in \mathbb{R}$ is a constant, then $E(cX) = cE(X)$

(3) If $X_1, X_2 \in H$, then $E(X_1 + X_2) = E(X_1) + E(X_2)$

(4) $E(1_{\Omega}) = 1$

(5) If sequence of functions X_n is increasing and convergent to X, then $\lim(E(X_n)) = E(X)$

1.3 Sublinear Expectation

A sublinear expectation E should satisfy the following statements:

(1) (monotonicity) If $X \leq Y$, then $E(X) \leq E(Y)$

(2) (constant-preserving) If $c \in \mathbb{R}$ is a constant, then $E(c \cdot 1_\Omega) = c$

(3) (sub-additivity) If $X_1, X_2 \in H$, then $E(X_1 + X_2) \leq E(X_1) + E(X_2)$

(4) (positive homogeneity) If $X \in H$ and $c \geq 0$ is a constant, then $E(cX) = cE(X)$

From the statements it's easy to know that a sublinear expectation is a linear expectation.

1.4 Quantum Expectation

Let $\mathbb{A}$ be a unital *-algebra.

Definition: An unital *-algebra $\mathbb{A}$ is a subset of $\mathbb{C}^\Omega$ if every $f \in \mathbb{A}$ satisfies $\forall m \in \mathbb{N}|f|^m <$ inf and has a unital element 1_Ω for which $1_\Omega * f = f$.

A map $\varphi : \mathbb{A} \to \mathbb{C}$ is a state on $\mathbb{A}$ if it satisfies:

(1) (linear) $\varphi(a + b) = \varphi(a) + \varphi(b)$, $\varphi(\lambda a) = \lambda\varphi(a)$ for $\lambda \in \mathbb{C}$

(2) (positive) $\varphi(a * a) \geq 0$

Note that multiplication in Quantum Expectation can be not commutative.

2 Applications

2.1 Picking balls and Sublinear expectation

Think about a simple game of picking balls among a set Ω, where the probability distribution of picking up each ball is unknown. Use θ to describe this distribution and Θ to describe the set of all distributions. If picking up each ball will make a different cost, then the cost function can be defined as $\varphi : \Omega \to \mathbb{R}$.

Use $E_\theta(\varphi)$ to describe the expectation cost of picking balls under the distribution θ when the cost function is φ. To calculate the expectation of cost equals to calculate $E(\varphi) = \sup_{\theta\in\Theta} E_\theta(\varphi)$, where E is a sublinear expectation.

(1) (monotonicity) If $\varphi_x \leq \varphi_y$, then whatever θ is, the cost of picking any ball under φ_x is lower than under φ_y, so $\forall\theta \in \Theta E_\theta(\varphi_x) \leq E_\theta(\varphi_y)$

(2) (constant-preserving) If $c \in \mathbb{R}$ is a constant, then $\varphi_c = c \cdot 1_\Omega$means that the cost of picking any ball is always c, so $E(\varphi_c) = c$

(3) (sub-additivity) If φ_1, φ_2 are two cost functions, then $E(\varphi_1 + \varphi_2) = \sup_{\theta\in\Theta} E_\theta(\varphi_1 + \varphi_2) \leq \sup_{\theta\in\Theta} E_\theta(\varphi_1) + \sup_{\theta\in\Theta} E_\theta(\varphi_2) = E(\varphi_1) + E(\varphi_2)$

(4) (positive homogeneity) If φ is a cost function and $c \geq 0$ is a constant, then $E(c\varphi) = cE(\varphi)$. Note that if $c \leq 0$ this equality may not hold because $E(\varphi)$ takes the upper bound among $E_\theta(\varphi)$

2.2 From sublinear to linear

Theorem 1. *Let E be a functional defined on a linear space H satisfying sub-additivity and positive homogeneity. Then there exist a family of linear functional $E_\theta, \theta \in \Theta$ such that $E(X) = \max_{\theta\in\Theta} E_\theta(X)$ for $X \in H$*

Theorem 2. *If E is a sublinear expectation then there exist a family of linear expectation $E_\theta, \theta \in \Theta$ such that $E(X) = \max_{\theta\in\Theta} E_\theta(X)$ for $X \in H$*

2.3 Probability in Convex Hull[1][2]

Theorem 3 (Caratheodory's theorem)**.** *Every point x in a $Conv(T)$ can be expressed as $x = \Sigma\lambda_i z_i$ where $\Sigma\lambda_i = 0$ and $\lambda_i \geq 0$, with at most $n+1$ points z_i from T.*

证明. By the definition of convex hull, x can be expressed as $x = \Sigma\lambda_i z_i$ where $\Sigma\lambda_i = 0$ and $\lambda_i \geq 0$ with finite points, suppose at least k points. If $k \leq n+1$ then the theorem is prooved.

If $k \geq n+1$, then $(x_2 - x_1), \cdots, (x_k - x_1)$ has at least $n+1$ elements so they should be linear dependent. So there exist $\mu_j \in R$ such that $\Sigma_{j=2}^{k}\mu_j(x_j - x_1) = 0$ and mu_j is not all zero.

Without loss of generality, take $\mu_k \neq 0$ so $\mu_k x_k = \mu_k x_1 - \Sigma_{j=2}^{k-1}\mu_j(x_j - x_1)$. It means that μ_k can be expressed with $x_1, \cdots, x_{k-1}$, so x can be expressed only with $x_1, \cdots, x_{k-1}$, and k isn't the smallest number of points to express x, leads to a contradiction. □

Theorem 4 (Approximate Carathedory's theroem)**.** *For a set $T \subset \mathbb{R}^n$ and diam$T \leq 1$, then for every $x \in Conv(T)$ and integer k there exists $x_1 \cdots x_k \in T$ such that $|x - \frac{1}{k}\Sigma_{j=1}^{k} x_j|_2 \leq \frac{1}{\sqrt{k}}$*

参考文献

[1] Unknown. Caratheodory theorem. `https://www.tutorialspoint.com/convex_optimization/convex_optimization_caratheodory_theorem.htm`.

[2] Roman Vershynin. Springer Netherlands, University of California, Irvine, June 9, 2020.