PROBABILITY THEORY

Study Note

Author
Sam Ren
Grinnell College
February 5, 2024

Contents

Acknowledgements

I finsihed this study notes by simply collected the concepts and terms from @brightsideofmaths youtube videos, here is the link: Video link. At the same time I would like to sicerely thank to my teachers and friends. Without them this note would not be done.

1 Measure Theory

This section is not necessary for later studying but it helps with understanding those abstract concepts

2 Probability Measure

Probability Measure is the measure with total mass 1, where we call it sample space $\Omega, Area(\Omega) = 1$. In math we write:

$$\mathbb{P} : A \to \mathbb{R}$$

Where A is the collection of subsets, then we have:

$$\mathbb{P}(\Omega) = 1, \mathbb{P}(A) \in [0, 1]$$

It is easy to notice that if two sets(A, B) are disjointed$(A \cap B = \emptyset)$ then we have:

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$$

Theorem 2.1. If we have a sequence of disjoin sets $A_i, i \in \mathbb{N}$ then we have:

$$\mathbb{P}(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$$

Definition 2.2. Let Ω be a set. A collection of subsets $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ ($\mathcal{P}$ is the power set) is called a sigma algebra($\sigma-$Algebra) if:

1. $\Omega, \emptyset \in \mathcal{A}$
2. If $A \in \mathcal{A}$ then $A^c \in \mathcal{A}$
3. If $A_1, A_2, A_3, ... \in \mathcal{A}$ then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$

In Probability theory we call the elements in σ-algebra events. Then we are going to define the probability measure properly.

Definition 2.3. Let $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ is a sigma-algebra. A map $\mathbb{P} : \mathcal{A} \to [0, 1]$ is called probability measure if:

1. $\mathbb{P}(\Omega) = 1$
2. $\mathbb{P}(\emptyset) = 0$
3. $\mathbb{P}(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$ where $i \neq j$

example 2.4. If we have $\Omega = \{1, 2, 3, 4, 5, 6\}$, then $\mathcal{A} = \mathcal{P}(\Omega)$ we have:

$$\mathbb{P} : \mathcal{A} \to [0, 1], \mathbb{P}(A) := \frac{|A|}{|\Omega|}$$

Proof. Prove that $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$ □

2.1 Discrete & Continuous Probability measure

We are going to discuss two situations where the sample space is either discrete or continuous. Where at here the term **Discrete** means the sample space is either countable or finite. For term **Continuous** the sample space is either infinite or uncountable.

Type	discrete	continuous
Sample space	Ω countable or finite	$\Omega \subseteq \mathbb{R}^n$ uncountable, $\Omega \in \mathcal{B}(\mathbb{R}^n)$
$\sigma-$algebra	$\mathcal{A} = \mathcal{P}(\Omega)$	$\mathcal{A} = \mathcal{B}(\Omega)$
Prob Measure	$\mathbb{P} : \mathcal{A} \to [0, 1]$	$\mathbb{P} : \mathcal{A} \to [0, 1]$
Determined by	$\mathbb{P}(\{\omega\}), \forall \omega \in \Omega$	
mass function	$(p_\omega)_{\omega\in\Omega}$ with $p_\omega \geq 0, \sum_{\omega\in\Omega} p_\omega = 1$	$f : \Omega \to \mathbb{R}$ with $f(x) \geq 0, \int_\Omega f(x)dx = 1$
Definition	$\mathbb{P}(A) = \sum_{\omega\in A} p_\omega = 1$	$\mathbb{P}(A) = \int_A f(x)dx$

Table 1: Discrete and continuous sample space explaination table

For continous probability measure we need to ensure that Ω is measuralbe which we are going to discuss later. Here is an example for continuous measure:

example 2.5. For $\Omega = [0, 2]$ we have prob density function $f(x) = \frac{1}{2}, f : \Omega \to \mathbb{R}$. Hence:

$$\mathbb{P}(\Omega) = \int_0^2 f(x)dx = \frac{1}{2}\int_0^2 1dx = 1$$

Therefore we would have conclusion:

$$\mathbb{P}(A) = \int_A f(x)dx = \frac{1}{2}\int_A 1dx = \frac{1}{2}\text{(Lebesgue measure)Length of A}$$

We would discuss the Lebesgue measure later.

2.2 Binomial Distribution

A binomial(two outcomes) distribution is foundamental to the probability theory, and it also require a definition(For now this definition is improper):

Definition 2.6. A binomial distribution can be determined by following features:

1. There is no order
2. With replacements (That is, does not change the sample space)
3. There are only two outcomes

If it is a binomial distribution we denote it as $\mathbb{P} = B(n, p)$. For a binomial distribution we have:

$$\Omega = \{0, 1, 2, 3, ..., n\}, \mathbb{P}(\{k\}) = \binom{n}{k} p^k (1-p)^{n-k}$$

2.3 Product Probability Spaces

In previous sections we already introduced the probability measure($\mathbb{P}$), the σ-algebra($\mathcal{A}$) and sample space(Ω). In this section we are going to introduce the probability space which is important in precisely proof and mathematics content.

Definition 2.7. Probability space is a triple combination$(\Omega_n, \mathcal{A}_n, \mathbb{P}_n), n \in \{1, 2, 3, ...\}$. Then the product space$(\Omega, \mathcal{A}, \mathbb{P})$ is given by:

- $\Omega = \Omega_1 \times \Omega_2 \times \Omega_3 \times ... = \prod_{i \in \mathbb{N}} \Omega_i$
- $\mathcal{A}$=$\sigma(\Omega_1 \times A_2 \times \Omega_3 \times ...)$ The choice of $A_n \in \mathcal{A}$ is not restricted by 2, it can be any number. Hence we just simply replace one Ω_i with A_i to construct.
- $\mathbb{P}(A_1 \times A_2 \times A_3 \times ... \times A_m \times \Omega_{m+1} \times \Omega_{m+2} \times ...) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2) \cdot ... \cdot \mathbb{P}(A_m)$

You can find examples in this Link.

2.4 Hypergeometric Distribution(Multivariant)

For this distribution we consider the sample size n and without replacement and unordered. Let's begin with an example:

example 2.8. Let elements in set S be the color of balls:

$$S = \{0, 1, 2, 3\}$$

If we have two 0 and 1 for others then we have a $(2, 1, 1, 1)$. At the same time there are many functions where mapping set S to $\mathbb{N}$. We denote the set of these functions as $\mathbb{N}^S$. We now can define the sample space:

$$\Omega = \{(k_s)_{s \in S} \in \mathbb{N}^S | \sum_{s \in S} k_s = n\}$$

Let's denote N_s as the number of balls in color s and N as the total number: Then we would have:

$$\mathbb{P}((k_s)_{s \in S}) = \frac{\prod_{s \in S} \binom{N_s}{k_s}}{\binom{N}{k}}$$

Definition 2.9. For hypergeometric distribution, we have (variant)set $S = \{0, 1, 2, ..., n\}$ and sample space $\Omega = \{0, 1, 2, 3, ..., n\}$:

$$\mathbb{P}((k_s)_{s \in S}) = \frac{\prod_{s \in S} \binom{N_s}{k_s}}{\binom{N}{k}}$$

3 Conditional Probability

References