Prob. RC Note

Chia-Cheng, Hao
Yan-Wei, Su

October 13, 2022

Abstract

This is a note for recitation in 2022 Fall undergraduate Probability Theory in the Applied Mathematics department at NYCU. The instructor of the course is Professor Yuan-Chung Sheu (website), and the two TA are Chia-Cheng, Hao and Yan-Wei, Su. The Github site for this note is at https://github.com/18Allen/PT_RC_material/blob/main/Notes/master.pdf

Contents

Chapter 0

RC 0

Recitation 0

0.1 Basic rules

22 Sep. 18:30

- Recitation: Thursday 12:30 to 13:30, 18:30 to 19:30
- My office hour is right after recitation.
- Grade distribution: (quiz correction + attend rc): 1×4; mid correction + attend rc: 2; Attend one rc: 4

0.2 Review of some high school comminatory and probability tricks

Early chapters are about high school counting things over again. We will swiftly go through the concept.

0.2.1 Permutation and Combination

- Permutation
  #ways to form an ordering of m out of n different things.
  $$P(n, m)$$
- Combination
  #ways to form a group of m with n different things.
  $$C(n, m), \quad \binom{n}{m}$$
- Multinomial coefficents [Gra21]
  #ways to divide a set of n elements into r (distinguishable) subsets of $n_1, n_2, \ldots, n_r$ elements.
  $$\frac{n!}{n_1!n_2!\ldots n_r!}$$

0.2.2 Set Operation

- De Morgan's Law

0.3 Axioms of Probability

0.3.1 Probability Space

The **probability space** is a triple $\Omega, \mathcal{F}, P$ that contains

- The **sample space** Ω contains all possible outcome.
- The σ-algebra $\mathcal{F}$ is the **event space**. It is a subset of the power set of Ω we are interested in.
- The **probability measure** P is a function $P : \mathcal{F} \to [0, 1]$ that satisfies the three axioms.
  1. $P(\Omega) = 1$
  2. Non-negative
  3. Countable additivity for disjoint sets in $\mathcal{F}$.

0.3.2 Standard process

A standard process of solving these problems (HW1,2) is to find the size of possible outcome first. Then, finding the size of desired event, and the ratio of the two is the prob.

Example. [Gra21] Example 3.11
You have 10 pairs of socks in the closet. Pick 8 socks at random. For every i, compute the probability that you get i complete pairs of socks.

- # outcome:
- # desirable outcome:
- the probability is :

Example. [Gra21] Problem 3.2 (HW2 problem 2)
Three married couples take seats around a table at random. Compute $P(\text{no wife sits next to her husband})$.
Use Inclusion-Exclusion principle to compute the probability of its complement event.

0.3.3 Why do we need to set σ-algebra: Vitali set

Have you ever wonder: Why would I need $\mathcal{F}$ if I have Ω already? As the textbook said, you won't have any problem with this notion. However, things get messy when we encounter set of infinity size. The idea of "length" will not be clear then. We use **Vitali set** V as an example on $\mathbb{R}$ to show that we can't have a measure on V. This problem is one of the reason that we only put probability measure on $\mathcal{F}$.

For more information: How the Axiom of Choice Gives Sizeless Sets | Infinite Series

0.4 Homework Help

TBD

Chapter 1

RC 1

Recitation 1

1.1 Review

29 Sep. 18:30

1.1.1 Axioms of Probability

A **probability space** is a triple $\Omega, \mathcal{F}, P$ that contains

- The **sample space** Ω contains all possible outcome.
- The σ-algebra $\mathcal{F}$ is the **event space**. It is a subset of the power set of Ω we are interested in.
- The **probability measure** P

1.1.2 Probability (measure)

- The **probability measure** P is a function $P : \mathcal{F} \to [0, 1]$ that satisfies the three axioms.
    1. $P(\Omega) = 1$
    2. Non-negative
    3. Countable additivity for disjoint sets in $\mathcal{F}$.

1.1.3 σ-algebra

$\mathcal{F}$ is call a σ-algebra on a set Ω If

1. $\emptyset \in \mathcal{F}$
2. If $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$
3. If $A_1, A_2, \cdots \in \mathcal{F}$, then $\cup_{i=1} A_i \in \mathcal{F}$

> **Example.** Let $\Omega = \{1, 2, 3, 4, 5, 6\}$, find a minimal* σ-algebra that contains the sets $\{1, 2, 3\}, \{1\}$

Answer: [1]

1.1.4 Conditional Probability

For the general definition, take events A, B, and assume that $P(B) > 0$. The *conditional probability* of the event A given B equals

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

[1] $\mathcal{F} = \{\emptyset, \Omega, \{1\}, \{1, 2, 3\}, \{4, 5, 6\}, \{2, 3, 4, 5, 6\}, \{2, 3\}, \{1, 4, 5, 6\}\}$

1.1.5 Independence

Events $A_1, \dots, A_n$ are independent if

$$P(\cap_{i=1}^n A_i) = \prod_{i=1}^n P(A_i)$$

Please note that it means you can not just check $P(A_i \cap A_j) = P(A_i)P(A_j), i \neq j$. One example is teh following

Example (Pairwise Independence but not independent variables). ([Pan19] Exercise 1.4.2) Consider a regular tetrahedron die painted blue, red and green on three sides and painted in all three colours on the fourth side. If the die is equally likely to land on any side, show that the appearances of these colours on the side it lands on are pairwise-independent but not independent.

1.2 Problems

Exercise. Put r distinguishable balls into n different boxes. What is the probability of all n boxes are occupied?

Probabilistic approach. $A_i = \{\text{Box } i \text{ is occupied}\}$ for all $i = 1, \cdots, n$.

$$\mathbb{P}(\text{No empty boxes}) = \mathbb{P}\Big(\bigcup_{i=1}^n A_i\Big). \tag{1.1}$$

(Inclusion-exclusion principle!)

Combinatorial approach. Define $A(r, n)$ as the number of distributions such that all n boxes are non-empty when you put r balls into them. Then the probability is $A(r, n)/n^r$. Knowing $A(r, n+1)$ can lead us to $A(r, n)$:

$$A(r, n) = \sum_{k=1}^{r-1} \binom{r}{k} A(r-k, n-1). \tag{1.2}$$

Knowing this, we can prove that:

$$A(r, n) = \sum_{\nu=0}^{n} (-1)^\nu \binom{n}{\nu} (n-\nu)^r. \tag{1.3}$$

(Use Induction!)

(Think about it!) What is the probability of exactly m boxes are occupied?

$$\text{number of distributions} \;=\; \binom{n}{m} \times A(r, m). \tag{1.4}$$

$A(r, m) = \mathbb{P}(r \text{ balls into } m \text{ boxes and all of the boxes are occupied}) \times m^r$.

Chapter 2

RC 2

Recitation 1

2.1 Review

13 Oct. 18:30

2.1.1 Random Variables

([Cha] chapter 2) It is a 'function' mapping Ω to some number, i.e

$$X : \Omega \to \mathbb{R}$$

Notation

- $X \in A, \quad A \subseteq \mathbb{R}$: The set containing $\omega \in \Omega$ such that $X(\omega) \in B$. $X \in B = \{\omega \in \Omega | X(\omega) \in B\}$

2.1.2 Discrete Random Variables

The **range** of X is finite or countable.

- **probability mass function**: A measure on x_i given by the measure in Ω

$$f(x_i) = P(X = x_i)$$

- Because $X \in x_i$ for $i = 1, 2, \ldots$ from a disjoint partition (Why?) of Ω, we have

$$\sum_{i=1}^{\infty} f(x_i) = P(\Omega) = 1$$

  TBD(Panchenko assume that Ω is countable)

- Independence of random variables: (We won't go through the detail in class, **strongly recommend** you to check out [Cha] Proposition 3.) It comes from the Independence of events(1.1.5). This is a theorem showing you the equivalent way of defining independence. One using subcollection, one using collection of all.

  Additional note: An infinite sequence of r.v $X_1, X_2, \ldots$ is called Independence if **for any n**, $X_1, X_2, \ldots, X_n$ are independent.

2.1.3 Some Discrete r.v

The most important thing to remember about a r.v (before a test) are 1. PMF; 2. expectation; 3. Varaiance. Examples about the r.v or the relation with other r.v's are helpful, too. You can those info in [Gra21] Chapter 5.

- **Bernoulli r.v** Bernoulli(p)
- **Binomial r.v** binomial(n, p)
- **Poisson r.v** Poisson(λ)
- **Geometric r.v** Geometric(p)

2.1.4 Joint Probability Mass Function

Let $X_1, \ldots, X_n$ be discrete random variables defined on the same sample space. The function

$$f(x_1, \ldots, x_n) = P(X_1 = x_1, \ldots, X_n = x_n)$$

is called the joint probability mass function (joint p.m.f) of the r.v $X_1, \ldots, X_n$.

marginal p.m.f: sum out the rest variables then the one(s) you are interested in, i.e

$$f_i(x) = \sum_{x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n} f(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$$

Theorem. Let $X_1, \ldots, X_n$ be discrete random variables with joint p.m.f f. Suppose that

$$f(x_1, \ldots, x_n) = h_1(x_1) \cdots h_n(x_n)$$

for some p.m.f $h_1, \ldots, h_n$. Then $X_1, \ldots, X_n$ are independent. and $f_i = h_i$, for $i = 1, \ldots, n$.

Example (Joint p.m.f). .

- (In the lecture note) Compute the p.m.f of Geometric r.v from Bernoulli's
- Compute the p.m.f of Binomial r.v from Bernoulli
- (Joint p.m.f when not independent) Consider two independent r.v X, Y follows the p.m.f

$$f(-1) = f(0) = f(1) = \frac{1}{3}$$

Then, we consider $U = XY, V = Y$. Compute the respective p.m.f of U and V. Also, compute the joint p.m.f $f_{U,V}(u, v)$. What do you observe?

Exercise (stability of Poisson random variables). Suppose $X_n \sim \text{Pois}(\lambda)$ are i.i.d. Then Define $S_n = X_1 + \cdots + X_n$. Show that $S_n \sim \text{Pois}(n\lambda)$.

2.2 Preview: Expectation

Let X be be a discrete random variable. The expected value or expectation or mean of X is defined as

$$E(X) = \sum_x x P(X = x)$$

Some simple example is like 'The expected number of head of fair coin toss', or 'average outcome of rolling a six-sided dice' etc.

How do we extend the idea to multiple random variables?

Theorem. (Proposition 6. in [Cha]). Let $X_1, \ldots, X_n$ be discrete random variables and $Y = f(X_1, \ldots, X_n)$ for some function f. Then,

$$E(Y) = \sum_{x_1, \ldots, x_n} f(x_1, \ldots, x_n) P(X_1 = x_1, \ldots, X_n = x_n).$$

With this notion we can use the linearity of expectation.

Exercise (Linearity of expectation). .

- Acquire the expectation of Binomial r.v by direct computation and linearity
- Define two independent r.v X_1 Poisson(λ_1) and X_2 Poisson(λ_2) Calculate the expectation of $Y = X_1 + X_2$ by the linearity of expectation and by the stability of Poisson random variables.

Note The Linearity of expectation does not require the sequence of r.v being independent. (Next time we will show that this is not true for **variance**)

2.3 Extra note

2.3.1 From Bernoulli to Binomial

Exercise (Peak of Binomial distribution). (Feller p.59 q.5) The probability p_k that a given cell contains exactly k balls (with a total r balls and n boxes) is given by the binomial distribution. Show that the most probable number v satisfies

$$\frac{r+n-1}{n} \leq v \leq \frac{r+1}{n}$$

(In other words, it is asserted that $p_0 < p_1 < \cdots < p_{v-1} \leq p_v > p_{v+1} \cdots > p_r$)

(Hint: Don't use derivative to find local extremum).

2.3.2 From Binomial to Poisson

Appendix

Appendix A

Additional Proofs

A.1 Proof of ??

We can now prove **??**.

Proof of ??. See here. ■

Bibliography

[Cha] Sourav Chatterjee. *Lecture Notes For MATH151*. URL: https://souravchatterjee.su.domains//math151notes.pdf.

[Gra21] Janko Gravner. *Lecture Notes for Introductory Probability Introduction to Probability*. sbd, 2021. URL: https://www.math.ucdavis.edu/~gravner/MAT135A/resources/lecturenotes.pdf.

[Pan19] Dmitry Panchenko. *Introduction to Probability Theory*. 2019. URL: https://drive.google.com/file/d/1Rpkr-NCEyqygvypR65RznaZkb36KHB29/view.