
NOTEBOOK FOR MA215 PROBABILITY

Lecturer: Prof.Hong

Author

Hongli Ye

Southern University of Science and Technology

2024 Fall

Contents

1	Lecture 1 Basic of Probability 2024.09.12	3
2	Lecture 2 Probability Space 2024.09.19	4
3	Lecture 3 Conditional Probability and Independence 2024.09.26	6
4	Lecture 4 Discrete Random Variable 2024.10.10	7
5	Lecture 5 Continuous Random Variable 2024.10.12	9
6	Lecture 6 Expectation and Variance of special random variable	12
7	Lecture 7 Function of Random Varibale 2024.10.17	14
8	Lecture 8 Multi-variables 2024.10.24	15
9	Lecture 9 2024.10.29	18
A	Answer for Quizes	20
B	Extension Problem	21
	B.1 Coupon Collector Problem	21

1 Lecture 1 Basic of Probability 2024.09.12

Theorem 1.1. Basic principle of counting

Suppose there are two experiments. Experiment 1 has n results and experiment 2 has m results.

Then together there are $m \times n$ possible outcomes.

This basic theorem could be extended to many finite experiments by induction.

Definition 1.1. Permutation

Permutation means the different ordered arrangement of objects.

Theorem 1.2. Suppose we have n objects. Then there are $n! = \prod_{i=1}^n (i) = 1 \times 2 \times \cdots \times n$ possible permutations.

Theorem 1.3. There are n objects, of which n_1 are alike, n_2 are alike, \dots, n_r are alike.

Then there are $\frac{n!}{n_1! \times n_2! \times \dots \times n_r!}$ possible outcomes.

Definition 1.2. Combination

Combination refers to selecting items from a set where order does not matter.

Theorem 1.4. If we choose r objects from a total of n different objects at a time, then the # possible combinations of $\binom{n}{r}$

Theorem 1.5. Binomial Theorem

For any positive integer $n \geq 1$

$$(x + y)^k = \sum_{k=0}^n \binom{n}{k} x^k y^{n-k}$$

Definition 1.3. Induction

Mathematical Induction is a proof method for natural numbers, consisting of a base case and an inductive step to show a statement holds for all natural numbers.

Mathematical Induction's basic step:

1. Basic step: The case holds when $n = 1$
2. Inductive step: Assume $n = k$ holds for some $k \geq 1$. Then $n = k + 1$ holds.

Quiz 1.1. From 8 women and 6 men, a committee of 3 men and 3 women is to be formed. How many different committees?

1. 2 of the men refuse to serve together?
2. 2 of the women refuse to serve together?
3. 1 man and 1 woman refuse to serve together?

2 Lecture 2 Probability Space 2024.09.19

Probability Space includes Sample Space, Events and Probability Measure.

Probability Space is a special case of measure theory.

Definition 2.1. Sample Space

The sample space S is the set of all possible outcomes of an experiment.

Definition 2.2. Event

An event is a subset of the sample space S , denoted $E \subset S$

Definition 2.3. Set Operation

Let E, F be two events and S is the sample space.

1. **Union:** $E \cup F = \{x | x \in E \text{ or } x \in F\}$
2. **Intersection:** $E \cap F = \{x | x \in E \text{ and } x \in F\}$
3. **Complement:** $E^c = \{x | x \notin E \text{ and } x \in S\}$
4. **Different:** $E - F = \{x | x \in E \text{ or } x \notin F\}$

Definition 2.4. Extension: σ - algebra

Let \mathcal{X} be a non-empty set. \mathcal{F} is said to be a σ -algebra if:

1. $\mathbb{X} \in \mathbb{F}$
2. If $A \in \mathcal{F}$, $A^c \in \mathcal{F}$
3. If $A_1, A_2 \dots \in \mathcal{F}$, then $\bigcup_{i=1}^{\infty} (A_i) \in \mathcal{F}$

Theorem 2.1. De Morgan's Law

For each $n \geq 1$, we have

$$\begin{aligned} \left(\bigcup_{i=1}^n (E_i)\right)^c &= \bigcap_{i=1}^n (E_i^c) \\ \left(\bigcap_{i=1}^n (E_i)\right)^c &= \bigcup_{i=1}^n (E_i^c) \end{aligned}$$

Axiom 2.1. Axiom of Probability Let S be a sample space. For each event E , the probability $P(E)$ satisfies:

1. $0 \leq P(E) \leq 1$
2. $P(S) = 1$
3. For any sequence of mutually exclusive events E_1, E_2, \dots , we have:

$$\sum_{i=1}^{\infty} P(E_i) = 1$$

Theorem 2.2. Basic corollaries:

1. $P(E) = 1 - P(E^c)$
2. If $E \subset F$, then $P(E) \leq P(F)$
3. $P(E \cup F) = P(E) + P(F) - P(E \cap F)$
4. Inclusion-Exclusion Identity: (Extension of the line above)

$$P\left(\bigcup_{i=1}^n\right) = \sum_{i=1}^n P(E_i) - \sum_{i_1 < i_2} P(E_{i_1} \cap E_{i_2}) + \dots + (-1)^{n+1} P\left(\bigcap_{i=1}^n (E_i)\right)$$

Quiz 2.1. There are N cards numbered as $1, 2, \dots, N$. Pick 1 card uniformly at random. Write down the number and return the card. Repeat for n times ($n > N, n = N, n < N$), we get a sequence (x_1, x_2, \dots, x_n) .

1. $P(\text{the sequence is strictly increasing})$
2. $P(\text{the sequence is non-decreasing})$

3 Lecture 3 Conditional Probability and Independence

2024.09.26

Definition 3.1. Conditional Probability

For 2 events E,F such that $P(E) > 0$. The conditional probability F occurs given that E has occurred is denoted by:

$$P(F|E) = \frac{P(F \cap E)}{P(E)}$$

Theorem 3.1. If each outcome of a finite sample space is equally likely, then we may compute the conditional probability of the form $P(F|E)$ by using E as the reduced sample space.

Theorem 3.2. Multiplication Law

For events E, F, we have:

$$P(E \cap F) = P(E) \times P(F|E)$$

More generally:

$$P(E_1 \cap E_2 \cap \dots \cap E_n) = P(E_1) \times P(E_2|E_1) \times P(E_3|(E_1 \cap E_2)) \dots P(E_n | \bigcap_{i=1}^{n-1} (E_i))$$

Definition 3.2. Independence

For two events E, F. We say E and F are independent if:

$$P(E \cap F) = P(E) \times P(F) \text{ or } P(F|E) = P(F)$$

Theorem 3.3. Total Probability Formula

Let A_1, A_2, \dots, A_n be mutually exclusive with $S = \bigcup_{k=1}^n (A_k)$.

Then \forall event B:

$$P(B) = \sum_{i=1}^n P(B|A_i)P(A_i)$$

Theorem 3.4. Bayes's Theorem

Let A_1, A_2, \dots, A_n be mutually exclusive so that $S = \bigcup_{k=1}^n (A_k)$.

Then \forall event B:

$$P(A_j|B) = \frac{P(B|A_j) \times P(A_j)}{\sum_{i=1}^n P(B|A_i)P(A_i)}$$

Quiz 3.1. A gambler has a fair coin and a two-headed coin in his pocket.

1. He selects one of the coins at random; when he flips it, it shows heads. What is the probability that it is the fair coin?
2. Suppose that he flips the same coin a second time and, again, it shows heads. Now what is the probability that it is fair coin?
3. Suppose that he flips the same coin a third time and it shows tails. Now what is the probability that it is the fair coin?

4 Lecture 4 Discrete Random Variable 2024.10.10

Definition 4.1. Discrete Random Variable

A Random Variable $X : S \rightarrow \mathbb{R}$.

If we take on at most a countable number of possible values is called discrete random R.V.

For example: 80 students, for which 70 are male. Choose 1 uniformly at random. Do this for 4 times. Let $X = \#$ of male students chosen.

Then X is a discrete R.V. taking values of $\{0,1,2,3,4\}$

Moreover:

$$\forall k \in \{0, 1, 2, 3, 4\}. P(X = k) = \binom{4}{k} \left(\frac{7}{8}\right)^k \left(\frac{1}{8}\right)^{4-k}$$

This is the probability mass functor of X .

Definition 4.2. Probability Mass Functor

For a discrete random variable X , we can define the probability mass functor(p.m.f), where $p(m)$ of X by

$$p(m) = P(X = m)$$

Definition 4.3. Special Random Variable

1. A random variable is said to be a **Bernoulli** random variable with parameter $p \in [0, 1]$ if:

$$P(X = 0) = 1 - p \text{ and } P(X = 1) = p$$

We say $X \sim \text{Bernoulli}(p)$

2. If we toss a coin independently for n times and let $X = \#$ of heads coming up, then X is said to be a **Binomial** random variable with parameter $p \in [0, 1]$
Denoted by $X \sim \text{Bin}(n, p)$

The possible mass functor of $\text{Bin}(n, p)$ is:

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

For example: Alice is in a class of 80 students, after 100 independent trials. We count X as the # of times where Alice is picked. Then $X \sim \text{Bin}(100, \frac{1}{80})$

Remark: Binomial R.V. equals n times the addition of Bernoulli R.V.

Definition 4.4. Poisson Random variable

Let $X = \text{Bin}(n, \frac{\lambda}{n})$ for some $\lambda > 0$.

Then let $n \rightarrow \infty$, we can get a new p.m.f, which is the p.m.f of Poisson R.V. :

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!} \quad \forall k \geq 0$$

Denoted by $X \sim \text{Poisson}(\lambda)$

Definition 4.5. Geometry Random Variable

There is a coin having probability $p \in (0, 1)$ of coming up heads. Toss the coin until it shows up head. Let $X = \#$ of tosses needed.

Then $X \sim \text{Geometric}(p)$, then p.m.f. of which is :

$$P(X = k) = (1 - p)^{k-1} p \quad \forall k \geq 1$$

Denoted by $X \sim \text{Geometric}(p)$

The definition seems to be different from the Geometry Random Variable in Statistics. But they are actually the same.

Coupon Collector Problem:

Pick one card uniformly at random, record the number and then return the card. Repeat until we collect all the n numbers.

What is the average number of trials needed?

Definition 4.6. Expectation

For a discrete random variable, the expectation of X is defined by:

$$E(X) = \sum_{k=1}^n k P(X = k)$$

Quiz 4.1. Jim is conducting random walk on the real line starting from 0. For each time, independently of anything else, he moves one steps to the right with probability p , and to the left with probability $1 - p$. Let X_n be the position of Jim at time n . Find $P(X_n = k)$ for each $-n \leq k \leq n$

5 Lecture 5 Continuous Random Variable 2024.10.12

Definition 5.1. Probability Density Function

A non-negative function $f : (-\infty, \infty) \rightarrow [0, \infty]$ is called a probability density function (p.d.f.), if:

$$\int_{-\infty}^{\infty} f(x) dx = 1$$

Definition 5.2. Continuous Random Variable

A random variable X is called a continuous random variable if exists a p.d.f. f such that:

$$\forall a, b \in \mathbb{R}, P(a \leq X \leq b) = \int_a^b f(x) dx$$

Remark: Let $b = a$ to get:

$$P(X = a) = P(a \leq X \leq a) = \int_a^a f(x) dx = 0$$

Definition 5.3. Uniform Random Variable

A random variable $X \sim \text{Uniform}(\alpha, \beta)$, if the p.d.f. of X is:

$$f(x) = \frac{1}{\beta - \alpha} 1_{(\alpha, \beta)}(x) = \begin{cases} \frac{1}{\beta - \alpha} & \text{if } \alpha < x < \beta \\ 0 & \text{, otherwise} \end{cases}$$

Indicator function:

$$1_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

For example: Let $X \sim \text{Unif}(1, 5)$, find $P(X > 3.5)$

Solution:

$$P(X > 3.5) = P(X \geq 3.5)$$

$$P(X \geq 3.5) = \lim_{b \rightarrow \infty} P(3.5 \leq X \leq b)$$

$$P(X \geq 3.5) = \int_{3.5}^{\infty} f(x) dx = \frac{3}{8}$$

Definition 5.4. Exponential Random Variable

We say a X is an exponential random variable with parameter $\lambda > 0$ if the p.d.f. is:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Definition 5.5. Memoryless

We say a random variable is memoryless if:

$$P(X > t + s | X > t) = P(X > s) \quad \forall t, s > 0$$

It is easy to prove that all exponential random variables are memoryless.

Theorem 5.1. If X is memoryless, then $X \sim \text{Exp}(\lambda)$ for some $\lambda > 0$

For example: Let X be a continuous random variable with p.d.f.

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x = 0 \end{cases}$$

Calculate $P(50 \leq X \leq 150)$ Solution:

1. Use $\int_{-\infty}^{\infty} f(x) dx = 1$, we can get $\lambda = \frac{1}{100}$
2. $P(50 \leq X \leq 150) = P(X \geq 50) - P(X > 150) = e^{-\frac{50}{100}} - e^{-\frac{150}{100}}$

Definition 5.6. Gamma function:

$$\Gamma(\alpha) = \int_0^{\infty} e^{-y} y^{\alpha-1} dy$$

Moreover:

$$\text{If } \alpha = n \in \mathbb{N}, \text{ then } \Gamma(n) = (n-1)!$$

Definition 5.7. Gamma Random Variable

Let X be a Gamma Random Variable, denoted by $X \sim \text{Gamma}(n, \lambda)$, then its p.d.f. is:

$$f(x) = \frac{x^{n-1} e^{-x/\lambda}}{\lambda^n \Gamma(n)}, \quad x > 0$$

In fact, if X_1, X_2, \dots, X_n are independent $\text{Exp}(\lambda)$, then

$$X_1 + X_2 + \dots + X_n \sim \text{Gamma}(n, \lambda)$$

We can understand the Gamma random variable in both two ways.

Definition 5.8. Normal Random Variable

We say a $X \sim N(\mu, \sigma^2)$ is a normal random variable if the density is:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \forall -\infty < x < \infty$$

Definition 5.9. Expectation

For a continuous random variable X , the expectation of X is defined by

$$EX = E(X) = E[X] = \int_{-\infty}^{\infty} xf(x) dx$$

For any function g , we have:

$$E(g(X)) = \int_{-\infty}^{\infty} g(x)f(x) dx$$

Note: Expectation is actually a integration of a measurement.

Theorem 5.2. Properties of expectation

Let X, Y be two random variables:

1. $\forall c \in \mathbb{R}, E(c) = c$
2. If $X \geq 0$, then $EX \geq 0$
3. If $c \in \mathbb{R}, E(cX) = cEX$
4. $E[X + Y] = E[X] + E[Y]$

By properties 3 and 4, we know that expectation is **linear**.

For example: $X \sim Unif(0, 1)$

$$EX = \int_{-\infty}^{\infty} xf(x) dx = \frac{1}{2}$$

Definition 5.10. Variance

The variance of X is given by:

$$Var(X) = E[(X - EX)^2]$$

Moreover, we could also calculate by:

$$Var(X) = E(X^2) - (EX)^2$$

Quiz 5.1. Prove

$$\mu = \int_{-\infty}^{\infty} xf(x) dx \text{ and } \sigma^2 = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx$$

where:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

6 Lecture 6 Expectation and Variance of special random variable

Theorem 6.1. Expectation and Variance of C.R.V

1. For $X \sim \text{Exp}(\lambda)$, we have:

$$E(X) = \frac{1}{\lambda} \text{ and } \text{Var}(X) = \frac{1}{\lambda^2}$$

2. For $X \sim \text{Gamma}(\alpha, \lambda)$, we have:

$$E(X) = \frac{1}{\lambda} \frac{\Gamma(\alpha + 1)}{\Gamma(\alpha)} \text{ and } \text{Var}(X) = \frac{1}{\lambda^2} \frac{\Gamma(\alpha + 2)}{\Gamma(\alpha)} - \left(\frac{1}{\lambda} \frac{\Gamma(\alpha + 1)}{\Gamma(\alpha)} \right)^2$$

3. For $X \sim N(\mu, \sigma^2)$, we have:

$$E(X) = \mu \text{ and } \text{Var}(X) = \sigma^2$$

Theorem 6.2. Expectation and Variance of D.R.V

For $X \sim \text{Bernoulli}(p)$, we have:

$$E(X) = p \text{ and } \text{Var}(X) = p(1 - p)$$

For $X \sim \text{Bin}(n, p)$, we have:

$$E(X) = np \text{ and } \text{Var}(X) = np(1 - p)$$

For $X \sim \text{Poisson}(\lambda)$, we have:

$$E(X) = \lambda \text{ and } \text{Var}(X) = \lambda$$

For $X \sim \text{Geo}(p)$, we have:

$$E(X) = \frac{1 - p}{p} \text{ and } \text{Var}(X) = \frac{1 - p}{p^2}$$

Definition 6.1. Cumulative Distribution Function

For a random variable X , the cumulative distribution function(c.d.f.) of X is:

$$F_X(b) = P(X \leq b)$$

We notice that:

1. For discrete random variable:

$$F(b) = \sum_{m=-\infty}^{[b]} P(X = m)$$

2. For continuous random variable:

$$F'(b) = f(b)$$

But random variable have forms instead of these two kinds. See the quiz below:

Quiz 6.1. The cumulative distribution function of X is

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{x}{2} & 0 \leq x < 1 \\ \frac{2}{3} & 1 \leq x < 2 \\ \frac{11}{12} & 2 \leq x < 3 \end{cases}$$

- (i) $P(x < 3)$
- (ii) $P(x = 1)$
- (iii) $P(x > \frac{1}{2})$

Theorem 6.3. 1. If $A_n \subset A_{n+1}, \forall n \geq 1$, then:

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n \rightarrow \infty} P(A_n)$$

2. If $B_{n+1} \subset B_n, \forall n \geq 1$, then:

$$P\left(\bigcap_{n=1}^{\infty} B_n\right) = \lim_{n \rightarrow \infty} P(B_n)$$

Theorem 6.4. Properties of Cumulative Distribution Function:

Let F be a cumulative distribution function.

1. F is a non-decreasing function, i.e.:

$$\forall a < b, F(a) \leq F(b)$$

2. $\lim_{b \rightarrow -\infty} F(b) = 0, \lim_{b \rightarrow \infty} F(b) = 1$

3. F is right continuous, i.e.:

$$\forall b \in \mathbb{R}, \forall \lim_{n \rightarrow \infty} b_n = b, \text{ we have } \lim_{n \rightarrow \infty} F(b_n) = F(b)$$

4. F has left limits, i.e.:

$$\forall b \in \mathbb{R}, \forall \lim_{n \rightarrow \infty} (a_n) = a, \text{ we have } \lim_{n \rightarrow \infty} F(a_n) = F(a^-) = F(x < a)$$

Use the **theorem6.3**, we could easily prove.

For example, we take $X \sim \text{Bernoulli}(p)$, then:

$$F_X(b) = \begin{cases} 1 & b \geq 1 \\ 1-p & 0 \leq b < 1 \\ 0, & b < 0 \end{cases}$$

It is a very traditional step function.

7 Lecture 7 Function of Random Varibale 2024.10.17

Theorem 7.1. If $X \sim N(\mu, \sigma^2)$, then:

$$Y = aX + b \sim N(a\mu + b, a^2\sigma^2), a, b \in \mathbb{R}$$

Quiz 7.1. If the pdf of X is:

$$f(x) = \frac{1}{\pi(1+x^2)}, \quad -\infty < x < \infty$$

Show that $Y = \frac{1}{X}$ has the same pdf.

Theorem 7.2. Let X be a continuous random variable with pdf $f_X(x)$. Suppose $g(x)$ is a strictly monotonic(increasingly or decreasing), differentiable function. Then $Y = g(X)$ has a pdf:

$$f_Y(y) = \begin{cases} f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| & \text{if } y = g(x) \text{ for some } x. \\ 0 & \text{if } y \neq g(x), \forall x \end{cases}$$

Proof: $\forall y \in \mathbb{R}, F_Y(y) = P(Y \leq y) = P(g(X) \leq y)$. Assume g is increasing. Then $g(X) \leq y \Leftrightarrow X \leq g^{-1}(y)$. So, $F_Y(y) = P(X \leq g^{-1}(y)) = F_X(g^{-1}(y))$.

Theorem7.2 isn't useful since it has too many restrictions.

Now we do a summary on how to find a probability density function of $Y = g(X)$

1. Find the cdf of $Y = g(X)$, which means do some simple calculation.
2. Differentiate to find the density.
3. Specify in what region the result holds.

Theorem 7.3. Let $F(x)$ be the cdf of any random variable. Define for each $x \in (0, 1)$:

$$F^{-1}(x) = \sup\{y \in \mathbb{R} : F(y) < x\}$$

8 Lecture 8 Multi-variables 2024.10.24

Definition 8.1. Joint cumulative distribution function:

For any random variables X, Y the joint cumulative distribution function of X and Y is defined by:

$$F(a, b) = P(X \leq a, Y \leq b)$$

Obviously, by the axiom of probability, we should have:

$$\lim_{a, b \rightarrow \infty} (F(a, b)) = 1$$

Notice that:

$$P(X \leq a) = \lim_{b \rightarrow \infty} (P(X \leq a, Y \leq b))$$

Denote as $P(X \leq a) = F(a, \infty)$, and this is defined as **Marginal Probability Density Function**.

For discrete multi-random variables, we can define:

Definition 8.2. Joint probability mass function: When X, Y are both discrete random variables with p.m.f is given by p_X, p_Y .

The joint probability mass function:

$$p(X, Y) = P(X = x, Y = y)$$

Similar to the definition above, we have **Marginal Probability Mass Function**:

$$p_X(x) = \sum_y P(X = x, Y = y).$$

Similar to a single variable, we can also define independence in multi-variables.

Definition 8.3. Independent random variables

We say X, Y are independent if $\forall A, B \in \mathbb{R}$,

$$P(X \in A, Y \in B) = p(X \in A)P(Y \in B)$$

From the two definitions above, we could induce that:

Theorem 8.1. Two discrete random variables X, Y are independent if and only if $\forall x, y \in \mathbb{R}$,

$$p(x, y) = P_X(x)P_Y(y)$$

Definition 8.4. Jointly continuous

We say X and Y are jointly continuous if there exist a function $f(x, y)$ such that

$$\forall C \subset \mathbb{R}^2, P((X, Y) \in C) = \int_{(x, y) \in C} f(x, y) dx dy$$

The function $f(x, y)$ is called the **joint probability distribution function** of X and Y .

Definition 8.5. joint cumulative distribution function:

The joint c.d.f. is then given by:

$$f_X(x) = \int_{-\infty}^a \int_{-\infty}^b f(x, y) dx dy$$

The definition of independence is still the same as before.

Definition 8.6. Expectation

For any joint p.m.f $p(x, y)$ or joint p.d.f $f(x, y)$, we have a \forall function $g : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$,

$$E(g(X, Y)) = \sum_m \sum_n (g(m, n) p(m, n))$$

For example:

$$g(X, Y) = 1_{X \in A} 1_{Y \in B}$$

$$E[g(X, Y)] = E[1_{X \in A, Y \in B}] = P(X \in A, Y \in B) = \sum_m \sum_n p(m, n)$$

Or, in continuous situations:

$$E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(X, Y) f(X, Y) dx dy$$

We need to notice that $1_{X \in A}$ and $1_{Y \in B}$ are beneficial **Characteristic functions**.

Another Example: A man and a woman promised to meet at 12 : 30P.M.. Assume the time they arrive are X and Y independently and satisfy:

$$X \sim \text{Unif}(12 : 15, 12 : 45)$$

$$Y \sim \text{Unif}(12 : 00, 1 : 00)$$

1. Calculate $P(\text{the man arrive first})$

$$X \sim \text{Unif}(-0.5, 0.5) Y \sim \text{Unif}(-1, 1)$$

$$\begin{aligned}
P(X < Y) &= \int 1_{(X < Y)} \cdot f(x, y) dx dy \\
&= \frac{1}{2} \int_{-\frac{1}{2}}^{\frac{1}{2}} dx \int_x^1 dy \\
&= \frac{1}{2} \int_{-\frac{1}{2}}^{\frac{1}{2}} (1 - x) dx \\
&= \frac{1}{2}
\end{aligned}$$

2. Find the probability that the first to arrive waits no longer than 5 minutes.

$$\begin{aligned}
P(|X - Y| < \frac{5}{30}) &= \iint 1_{(|X - Y| < \frac{1}{6})} \cdot f(x, y) dx dy \\
&= \frac{1}{2} \int_{-\frac{1}{2}}^{\frac{1}{2}} dx \int_{x - \frac{1}{6}}^{x + \frac{1}{6}} dy \\
&= \frac{1}{6}
\end{aligned}$$

Quiz 8.1. The joint probability distribution function: of X, Y is :

$$f(x, y) = \begin{cases} c & \text{if } x^2 + y^2 \leq R^2 \\ 0 & \text{otherwise} \end{cases}$$

a) Find c

b) Find the marginal probability distribution functions of $f_X(x)$ and $f_Y(y)$.

This is the uniform distribution in circle plates.

Definition 8.7. Bivariate Normal Distribution

The joint probability of bivariate normal distribution is given by:

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)^2}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 + \left(\frac{y-\mu_y}{\sigma_y}\right)^2 - 2\rho\frac{(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y}\right]\right\}$$

We denote it as $(X, Y) \sim N(\mu_x, \mu_y; \sigma_x^2, \sigma_y^2; \rho)$

Definition 8.8. Covariance

The covariance of X, Y is:

$$Cov(X, Y) := E[(X - EX)(Y - EY)]$$

By simple calculation, we know that:

$$Cov(X, X) = E[(X - EX)^2] = Var(X)$$

Now that we have expanded a single variable into bivariable, how can we get higher

dimensions?

We could use **Matrix Form** to gain a beautiful expression of any finite dimension normal distribution.

If we let:

$$\begin{aligned}\vec{x} &= (x, y) \\ \vec{\mu} &= (\mu_x, \mu_y) \\ \Sigma &= \begin{bmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{bmatrix} \\ \Rightarrow \det(\Sigma) &= \sigma_x^2\sigma_y^2(1 - \rho^2) \\ \Sigma^{-1} &= \frac{1}{1 - \rho^2} \cdot \begin{bmatrix} \frac{1}{\sigma_x^2} & -\frac{\rho}{\sigma_x\sigma_y} \\ -\frac{\rho}{\sigma_x\sigma_y} & \frac{1}{\sigma_y^2} \end{bmatrix}\end{aligned}$$

Then we have:

$$f(x, y) = \frac{1}{2\pi\sqrt{\det(\Sigma)}} \cdot \exp\left\{-\frac{1}{2}(\vec{x} - \vec{\mu})\Sigma^{-1}(\vec{x} - \vec{\mu})^T\right\}$$

This is a very beautiful structure with general form.

Theorem 8.2. In Bivariate Normal Distribution, the X, Y independent are equivalent with:

1. $\rho = 0$
2. $Cov(X, Y) = 0$

9 Lecture 9 2024.10.29

To understand the structure below better, we introduce the convolution of two functions.

Definition 9.1. Convolution: Let f and g be two functions

$$(f * g)(t) = \int_{-\infty}^{\infty} f(\tau)g(t - \tau) d\tau \quad (1)$$

Theorem 9.1. Sum of independent random variables

Let X, Y be independent continuous random variables. And $Z = X + Y$, then we have:

$$F_Z(z) = \int_{-\infty}^{\infty} F_X(z-y)f_Y(y)dy$$

Differentiate to obtain:

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} f_X(z-y)f_Y(y)dy \\ &= (f_X * f_Y)(z) \end{aligned}$$

Now we compute an example:

$X, Y \sim \text{Exp}(\lambda)$, and they are independent.

Compute the p.d.f. of $X + Y$:

Solution:

$$\begin{aligned} f_{X+Y}(z) &= \int_{-\infty}^{\infty} f_X(z-y)f_Y(y)dy \\ &= \int_0^z \lambda e^{\lambda(z-y)} \times \lambda e^{-\lambda y} dy \\ &= \lambda^2 e^{-\lambda z} z \end{aligned}$$

By observation, we can easily notice that $X + Y \sim \text{Gamma}(2, \lambda)$

Quiz 9.1. If the joint p.d.f. of (X, Y) is :

$$f(x, y) = \begin{cases} \frac{1}{2}(x+y)e^{-(x+y)}, & 0 < x, y < \infty \\ 0 & \text{otherwise} \end{cases}$$

Find the p.d.f. of $Z = X + Y$.

A Answer for Quizes

1. Quiz 1

- (a) 896
- (b) 1000
- (c) 910

2. Quiz 2

- (a) $\frac{\binom{N}{n}}{N^n}$
- (b) $\frac{\binom{N+n-1}{n}}{N^n}$

3. Quiz 3

- (a) $\frac{1}{3}$
- (b) $\frac{1}{5}$
- (c) 1

4. Quiz 4

$$P(X_n = k) = \begin{cases} \binom{n}{\frac{n+k}{2}} p^{\frac{n+k}{2}} (1-p)^{\frac{n-k}{2}} & \text{if } n+k \text{ odd} \\ 0 & \text{if } n+k \text{ even} \end{cases}$$

5. Quiz 5

6. Quiz 6

- (i) $\frac{11}{12}$
- (ii) $\frac{1}{6}$
- (iii) $\frac{3}{4}$

7. Quiz 7

8. Quiz 8

- (a) $c \frac{1}{\pi R^2}$
- (b) $f_X(x) = \frac{2}{\pi R^2} \sqrt{R^2 - x^2}; f_Y(y) = \frac{2}{\pi R^2} \sqrt{R^2 - y^2}$

9. Quiz 9

10. $f_Z(z) = \frac{1}{2} z^2 e^{-z}$

B Extension Problem

B.1 Coupon Collector Problem

See the original description in Lecture 4.

References

- [1] Wikipedia contributors, “Coupon Collector’s Problem,” Wikipedia, The Free Encyclopedia. [Online]. Available: https://en.wikipedia.org/wiki/Coupon_collector%27s_problem. [Accessed: [2024.10.23]].