

36225  
Probability

ZhenTing Liu

# Contents

## Chapter 1

## Page 3

1.1	Probability and Mass functions	3
1.2	Random	3
1.3	Algorithms	5

# Discrete Random Variables

# Chapter 1

## 1.1 Probability and Mass functions

A **discrete random variable**  $X$  on the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  is defined to be a mapping  $X : \Omega \mapsto \mathbb{R}$  such that

$$\begin{aligned} \text{the image } X(\Omega) &\text{ is a countable subset of } \mathbb{R} \text{ and} \\ \{\omega \in \Omega : X(\omega) = x\} &\in \mathcal{F} \end{aligned}$$

A discrete random variable  $X$  takes value in  $\mathbb{R}$ , but we cannot predict the actual value of  $X$  with certainty since the underlying experiment involves chance. Instead, we would like to measure the probability that  $X$  takes a given value,  $x$  say.

The most interesting things about a discrete random variable are the values that it may take and the probability associated with these values. If  $X$  is a discrete random variable on the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , then its **image**  $\text{im}X$  is the image of  $\Omega$  under  $X$ , that is, the set of values taken by  $X$ .

Henceforth, we abbreviate events of the form  $\{\omega \in \Omega : X(\omega) = x\}$  to the more convenient form  $\{X = x\}$ .

### Definition 1.1.1: The (probability) mass function

the pmf of the discrete random variable  $X$  is the function  $p_X : \mathbb{R} \mapsto [0, 1]$  defined by

$$p_X(x) = \mathbb{P}(X = x)$$

Thus,  $p_X(x)$  is the probability that the mapping  $X$  takes the value  $x$ . Note that  $\text{im}X$  is countable for any discrete random variable  $X$ , and

$$p_X(x) = 0 \quad \text{if } x \notin \text{im } X, \tag{1.1}$$

$$\sum_{x \in \text{im } X} p_X(x) = \mathbb{P}\left(\bigcup_{x \in \text{im } X} \{\omega \in \Omega : X(\omega) = x\}\right) = \mathbb{P}(\Omega) = 1 \tag{1.2}$$

$$0 \leq p_X(x) \leq 1 \forall x \tag{1.3}$$

## 1.2 Random

### Definition 1.2.1: Normed Linear Space and Norm $\|\cdot\|$

Let  $V$  be a vector space over  $\mathbb{R}$  (or  $\mathbb{C}$ ). A norm on  $V$  is function  $\|\cdot\| : V \rightarrow \mathbb{R}_{\geq 0}$  satisfying

- ①  $\|x\| = 0 \iff x = 0 \quad \forall x \in V$
- ②  $\|\lambda x\| = |\lambda| \|x\| \quad \forall \lambda \in \mathbb{R} (\text{or } \mathbb{C}), x \in V$
- ③  $\|x + y\| \leq \|x\| + \|y\| \quad \forall x, y \in V$  (Triangle Inequality/Subadditivity)

And  $V$  is called a normed linear space.

- Same definition works with  $V$  a vector space over  $\mathbb{C}$  (again  $\|\cdot\| \rightarrow \mathbb{R}_{\geq 0}$ ) where ② becomes  $\|\lambda x\| = |\lambda| \|x\|$   $\forall \lambda \in \mathbb{C}, x \in V$ , where for  $\lambda = a + ib$ ,  $|\lambda| = \sqrt{a^2 + b^2}$

**Example 1.2.1 ( $p$ -Norm)**

$V = \mathbb{R}^m$ ,  $p \in \mathbb{R}_{\geq 0}$ . Define for  $x = (x_1, x_2, \dots, x_m) \in \mathbb{R}^m$

$$\|x\|_p = \left( |x_1|^p + |x_2|^p + \dots + |x_m|^p \right)^{\frac{1}{p}}$$

(In school  $p = 2$ )

**Special Case  $p = 1$ :**  $\|x\|_1 = |x_1| + |x_2| + \dots + |x_m|$  is clearly a norm by usual triangle inequality.

**Special Case  $p \rightarrow \infty$  ( $\mathbb{R}^m$  with  $\|\cdot\|_\infty$ ):**  $\|x\|_\infty = \max\{|x_1|, |x_2|, \dots, |x_m|\}$

For  $m = 1$  these  $p$ -norms are nothing but  $|x|$ . Now exercise

**Question 1**

Prove that triangle inequality is true if  $p \geq 1$  for  $p$ -norms. (What goes wrong for  $p < 1$ ?)

**Solution: For Property ③ for norm-2**

**When field is  $\mathbb{R}$ :**

We have to show

$$\begin{aligned} \sum_i (x_i + y_i)^2 &\leq \left( \sqrt{\sum_i x_i^2} + \sqrt{\sum_i y_i^2} \right)^2 \\ &\Rightarrow \sum_i (x_i^2 + 2x_i y_i + y_i^2) \leq \sum_i x_i^2 + 2 \sqrt{\left[ \sum_i x_i^2 \right] \left[ \sum_i y_i^2 \right]} + \sum_i y_i^2 \\ &\Rightarrow \left[ \sum_i x_i y_i \right]^2 \leq \left[ \sum_i x_i^2 \right] \left[ \sum_i y_i^2 \right] \end{aligned}$$

So in other words prove  $\langle x, y \rangle^2 \leq \langle x, x \rangle \langle y, y \rangle$  where

$$\langle x, y \rangle = \sum_i x_i y_i$$

**Note:-**

- $\|x\|^2 = \langle x, x \rangle$
- $\langle x, y \rangle = \langle y, x \rangle$
- $\langle \cdot, \cdot \rangle$  is  $\mathbb{R}$ -linear in each slot i.e.

$$\langle rx + x', y \rangle = r\langle x, y \rangle + \langle x', y \rangle \text{ and similarly for second slot}$$

Here in  $\langle x, y \rangle$   $x$  is in first slot and  $y$  is in second slot.

Now the statement is just the Cauchy-Schwartz Inequality. For proof

$$\langle x, y \rangle^2 \leq \langle x, x \rangle \langle y, y \rangle$$

expand everything of  $\langle x - \lambda y, x - \lambda y \rangle$  which is going to give a quadratic equation in variable  $\lambda$

$$\begin{aligned} \langle x - \lambda y, x - \lambda y \rangle &= \langle x, x - \lambda y \rangle - \lambda \langle y, x - \lambda y \rangle \\ &= \langle x, x \rangle - \lambda \langle x, y \rangle - \lambda \langle y, x \rangle + \lambda^2 \langle y, y \rangle \\ &= \langle x, x \rangle - 2\lambda \langle x, y \rangle + \lambda^2 \langle y, y \rangle \end{aligned}$$

Now unless  $x = \lambda y$  we have  $\langle x - \lambda y, x - \lambda y \rangle > 0$  Hence the quadratic equation has no root therefore the discriminant is greater than zero.

**When field is  $\mathbb{C}$  :**

Modify the definition by

$$\langle x, y \rangle = \sum_i \bar{x}_i y_i$$

Then we still have  $\langle x, x \rangle \geq 0$

## 1.3 Algorithms

---

### Algorithm 1: what

---

**Input:** This is some input

**Output:** This is some output

*/\* This is a comment \*/*

```
1 some code here;
2 x ← 0;
3 y ← 0;
4 if x > 5 then
5   | x is greater than 5 ;                                // This is also a comment
6 else
7   | x is less than or equal to 5;
8 end
9 foreach y in 0..5 do
10  | y ← y + 1;
11 end
12 for y in 0..5 do
13  | y ← y - 1;
14 end
15 while x > 5 do
16  | x ← x - 1;
17 end
18 return Return something here;
```

---