Handout[1] 4

Sept 29, 2022

[Topics]:
Discrete Random Variables

1 Random Variables

Recall that for an experiment, we have different outcomes. For example, if we are talking about the weather tomorrow, the outcomes are in the set {rainny, sunny, cloudy}. Now we want to describe the probability of each outcome, say there is a probability of 0.3 to be rainny, 0.5 to be sunny, and 0.2 to be cloudy. How can we plot a graph to illustrate that? Probably we are unable to do that since it is hard to represent an event on the coordinate. However, if we assign a number to represent an event, then things are easy now. For example, we can assign number "1" to represent the event "tomorrow is rainny", number "2" to represent "tomorrow is sunny", and number "3" to represent "tomorrow is cloudy". Then we can represent each event on the x-axis. "There is a probability of 0.5 that tomorrow is sunny" now is translated into a mathematical language as $\mathbb{P}(2) = 0.5$. On the coordinate, the point (2, 0.5) hence represents such statement.

Definition 1.1 (Random Variable) A random variable is a real-valued outcome. It's a function maps from the sample space S to the real line $\mathbb{R}$.

Example 1.1 Flip a coin, the sample space for the outcome is {Head, Tail}. We can define X to be 1 if the outcome is Head and 0 if Tail.

$$X = \begin{cases} 1 & if \quad Head \\ 0 & if \quad Tail \end{cases}$$

Now X is a random variable, which maps from the sample space {Head, Tail} to two discrete numbers {1,0}. Note that the numbers mean nothing but to represent an event. You can also assign "770" to represent the event "outcome is Head" or number "8189" to represent the evnet "outcome is Tail", or any number you like to represent an event.

Notation : For a random variable X, we use the lower case letter x to denote the possible values. In particular, X is a random variable, x is a real number which represent a realization of the random variable. In the above example, X is random variable, $x = 1$ is a possible value.

2 Discrete Random Variables

Definition 2.1 (Discrete) A set $\mathcal{X}$ is discrete if its number of elements is countable (finite or countably infinite).

Finite is easy to understand, which is simply the total number of elements is a number N$< \infty$.
Countable means there is a one-to-one correspondence with natural number $\mathcal{N} = \{0, 1, 2, 3, \cdots\}$. That is, for each n $\in \mathcal{N}$, there is a correspond X_n. For example, a set $\mathcal{X} = \{1, 10, 100, 1000, \cdots\}$ is counbtable, since we can find a "relation" $X_n = 10^n$. The "relation" may not be a formula, for example, $\mathcal{X} = \{1998, 9, 27, 1997, 8, 11, \cdots\}$ can also be a countable set, where 1998-09-27 is my birthday and 1997-08-11 is Adry's birthday. For each n $\in \mathcal{N}$, there should be a correspond X_n, but we can assign empty set $\emptyset$ to n>N, then the counbtable set becomes finite. So **countable** represents either **finite** or **countably infinite**.

Definition 2.2 (Discrete Random Variable) A random variable is a discrete random variable if it takes finite or countably infinite number of elements.

Definition 2.3 (Support) The **support** of a discrete random variable is the set of possible values the random variable can take.

[1]This handout is made by Hongkai Wang for Econ 329 Economic Statistics. It is adapted from Professor Wiseman's lectures, however, all errors are mine.

The **example 1.1** is a discrete random variable since it takes only two elements "1" and "0". And the support is {0, 1}.

Example 2.1 Let's form the weather example into a random variable. It can be described as:

event	tomorror is rainny	tomorrow is sunny	tomorrow is cloudy
value of random variable	1	2	3
probability	0.3	0.5	0.2

The support is {1, 2, 3}.

3 Probability Mass Function(pmf) for Discrete Random Variables

Definition 3.1 (Probability mass function(pmf)) : For a discrete random variable X, the probability mass function f is defined by:

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

Example 3.1 For the weather example, the pmf's are:

$$\begin{array}{c} f(1) = \mathbb{P}(X = 1) = \mathbb{P}(tomorrow \quad is \quad rainny) = 0.3. \\ f(2) = \mathbb{P}(X = 2) = \mathbb{P}(tomorrow \quad is \quad sunny) = 0.5. \\ f(3) = \mathbb{P}(X = 3) = \mathbb{P}(tomorrow \quad is \quad cloudy) = 0.2. \end{array}$$

Proprties of pmf: Let $f(x_i)(i = 1, 2, 3, \cdots)$ be the pmf of a doscrete random variable, then

$$\begin{array}{ll} (1) & f(x_i) \geq 0 \qquad \forall i, \\ (2) & \sum_i f(x_i) = 1. \end{array}$$

Proof : (1) is easy to see since $f(x_i)$ is a probability and hence is nonnegative. For (2), assume all the possible values of the discrete random variable X form a set E, then

$$\begin{aligned} \sum_i f(x_i) &= \sum_{x_i \in E} f(x_i) \\ &= \sum_{x_i \in E} \mathbb{P}(X = x_i) \\ &= \mathbb{P}(\bigcup_{x_i} \{x_i \in E\}) \\ &= \mathbb{P}(E) \\ &= 1. \end{aligned}$$

4 Cumulative Distribution Function(cdf) for Discrete Random Variables

Definition 4.1 (Cumulative distribution function(cdf)) : The **distribution function** or **cumulative distribution function** F of a random variable is defined by :

$$F(x) = Pr(X \leq x).$$

For discrete random varaible, the cdf $F(x)$ is the sum of the probabilities of all the possible values that less than or euqal to x.

Example 4.1 For the weather example, what are the following cdf's ?

(1). $F(0.5) = Pr(X \leq 0.5) = 0$,
(2). $F(1) = Pr(X \leq 1) = f(1) = 0.3$,
(3). $F(1.314) = Pr(X \leq 1.314) = f(1) = 0.3$,
(4). $F(1.9999) = Pr(X \leq 1.9999) = f(1) = 0.3$,
(5). $F(2) = Pr(X \leq 2) = f(1) + f(2) = 0.3 + 0.5 = 0.8$,
(6). $F(2.5) = Pr(X \leq 2.5) = f(1) + f(2) = 0.3 + 0.5 = 0.8$,
(7). $F(3) = Pr(X \leq 3) = f(1) + f(2) + f(3) = 0.3 + 0.5 + 0.2 = 1$,
(8). $F(3.789) = Pr(X \leq 3.789) = f(1) + f(2) + f(3) = 0.3 + 0.5 + 0.2 = 1$,
(9). $F(100) = Pr(X \leq 100) = f(1) + f(2) + f(3) = 0.3 + 0.5 + 0.2 = 1$.

The graph for the cdf is as follows:

$$F(x) = \begin{cases} 0 \quad if \quad x < 1 \\ 0.3 \quad if \quad 1 \leq x < 2 \\ 0.8 \quad if \quad 2 \leq x < 3 \\ 1 \quad if \quad x \geq 3 \end{cases}$$

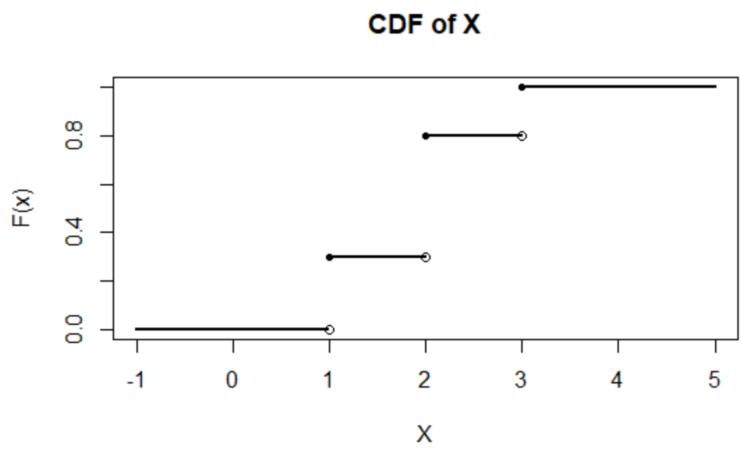

Figure 1: CDF of X, example of weather

CDF $F(x)$ measures the probability that the random variable takes a value less than or equal to x. Some properties then follows:
Properties of cdf : Suppose X is a random variable and $F(x)$ is the cumulative distribution function of X. Then the cdf $F(x)$ possesses the following properties:

$$\begin{aligned} &(1) \quad 0 \leq F(x) \leq 1, \\ &(2) \quad F(x) \text{ is nondecreasing, that is, if } x_1 < x_2, \text{ then } \quad F(x_1) \leq F(x_2), \\ &(3) \quad lim_{x \to -\infty} F(x) = 0, \\ &(4) \quad lim_{x \to +\infty} F(x) = 1. \end{aligned}$$

Proof (1) is trival since $F(x)$ is a probability.
(2) For $x_1 < x_2$, $\{X \leq x_1\} \subset \{X \leq x_2\}$, then $Pr(X \leq x_1) \leq Pr(X \leq x_2)$.
(3) $lim_{x \to -\infty} F(x) = lim_{x_n \to -\infty} Pr(X \leq x_n) = P(\emptyset) = 0$,
(4) $lim_{x \to +\infty} F(x) = lim_{x_n \to +\infty} Pr(X \leq x_n) = P(S) = 1$ (S is the whole sample space).

5 Relation Between pmf and cdf

Recall the definition:

$$pmf : f(x_i) = P(X = x_i)$$
$$cdf : F(x) = Pr(X \leq x)$$

Suppose a discrete random variable X takes nonnegative integers $\{0, 1, 2, 3, \cdots\}$.

1 Given pmf : Suppose we know pmf, i.e, $f(x_i)$ are known, then

$$F(x) = Pr(X \leq x) = \sum_{x_i \leq x} f(x_i).$$

For example, in the weather example, if given:

$$f(1) = 0.3$$
$$f(2) = 0.5$$
$$f(3) = 0.2$$

Then $F(2.2) = f(1) + f(2) = 0.8$.

2 Given cdf : Suppose we know cdf, i.e, $F(x)$ is given, for example, in the weather example, if given

$$F(1.8) = 0.3$$
$$F(2.2) = 0.8$$

Then $f(2) = F(2.2) - F(1.8) = 0.5$.

6 Examples of Discrete Random Variables

6.1 Bernoulli Random Variable

6.1.1 Definition : A Bernoulli random variable takes only two values 0(fail) and 1(success), with probability $1 - p$ and p, where $0 < p < 1$.

6.1.2 pmf : For a Bernoulli random variable X, the pmf is given by:

$$f(0) = P(X = 0) = 1 - p,$$
$$f(1) = P(X = 1) = p.$$

6.2 Binomial Random Variable

6.2.1 Definition : A Binomial random variable X with parameter n and p is a trail of n independent Bernoulli experiments with a success rate(the probability of taking value 1) of p. The value x represents the number of success in the n Bernoulli experiments.

6.2.2 pmf : The probability mass function for Binomial random variable is given by:

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

Where $k = 0, 1, 2, 3, \cdots$,and $0 < p < 1$.
If $n = 1$, then it is reduced to be a Bernoulli random variable.

Notation If X obeys the Binomial distribution with parameters n and p, we denote it as X∼Binomial(n,p).

6.3 Poisson Random Variable

6.3.1 Definition We say that a random variable X obeys the Poisson distribution if all the possible values of X are non-negative integers $\{0, 1, 2, 3, \cdots\}$ and the probability mass function takes the form of

$$f(k) = P(X = k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}.$$

Where $k = 0, 1, 2, 3, \cdots$ and $\lambda > 0$ is a constant. It is also denoted as $X \sim$Poisson(λ).

We can verify that $\sum f(k) = 1$:

$$\begin{aligned}\sum_{k=0}^{\infty} f(k) &= \sum_{k=0}^{\infty} e^{-\lambda} \cdot \frac{\lambda^k}{k!} \\ &= e^{-\lambda} \cdot \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \\ &= e^{-\lambda} \cdot e^{\lambda} \\ &= 1.\end{aligned}$$

Recall the taylor expansion of e^{λ}:

$$\begin{aligned}e^{\lambda} &= \frac{\lambda^0}{0!} + \frac{\lambda^1}{1!} + \frac{\lambda^2}{2!} + \cdots \\ &= \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}.\end{aligned}$$

6.3.2 cdf : If $X \sim$Poisson(λ), then the cdf of X is:

$$\begin{aligned}F(x) &= P(X \leq x) \\ &= \sum_{k \leq x} P(X = k) \\ &= \sum_{k \leq x} e^{-\lambda} \cdot \frac{\lambda^k}{k!}.\end{aligned}$$

For example,

$$\begin{aligned}F(8) &= P(X \leq 8) \\ &= \sum_{k \leq 8} P(X = k) \\ &= \sum_{k=0}^{8} e^{-\lambda} \cdot \frac{\lambda^k}{k!} \\ &= e^{-\lambda}(1 + \lambda + \frac{\lambda^2}{2!} + \cdots + \frac{\lambda^8}{8!}).\end{aligned}$$

And,

$$\begin{aligned}F(5.778899) &= P(X \leq 5.778899) \\ &= \sum_{k \leq 5.778899} P(X = k) \\ &= \sum_{k=0}^{5} e^{-\lambda} \cdot \frac{\lambda^k}{k!} \\ &= e^{-\lambda}(1 + \lambda + \frac{\lambda^2}{2!} + \cdots + \frac{\lambda^5}{5!}).\end{aligned}$$

7 Exercises

Exercise 7.1 Suppose X is a discrete random variable, whose pmf is given by:

$$f(1) = \frac{1}{4}, f(2) = \frac{1}{2}, f(3) = \frac{1}{8}, f(4) = \frac{1}{8}.$$

(1) Find the cdf of X.
(2) Plot the cdf.

Solution :
If $x < 1$, then $F(x) = P(X \leq x) = 0$,
If $1 \leq x < 2$, then $F(x) = P(X \leq x) = P(X = 1) = \frac{1}{4}$,
If $2 \leq x < 3$, then $F(x) = P(X \leq x) = P(X = 1) + P(X = 2) = \frac{1}{4} + \frac{1}{2} = \frac{3}{4}$,
If $3 \leq x < 4$, then $F(x) = P(X \leq x) = P(X = 1) + P(X = 2) + P(X = 3) = \frac{1}{4} + \frac{1}{2} + \frac{1}{8} = \frac{7}{8}$,
If $x \geq 4$, then $F(x) = P(X \leq x) = P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) = \frac{1}{4} + \frac{1}{2} + \frac{1}{8} + \frac{1}{8} = 1$.

The graph of the cdf $F(x)$ is as follows:

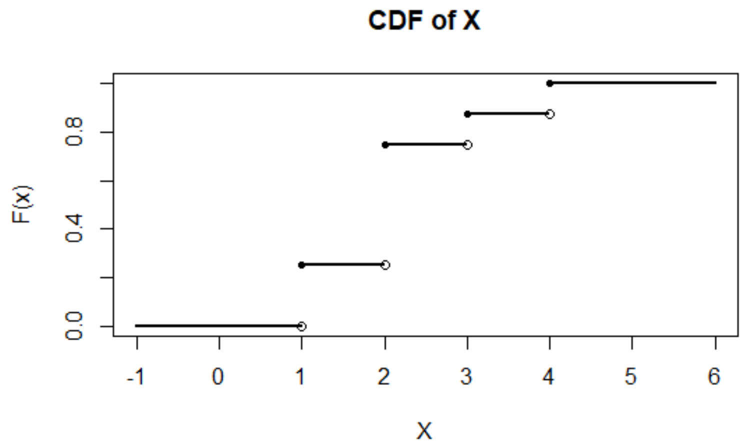

Figure 2: CDF of X, exercise 7.1

Exercise 7.2 The pmf of a random variable is given by $f(x) = c \cdot \frac{\lambda^x}{x!}$ $(x = 0, 1, 2, 3, \cdots)$, where $\lambda > 0$ is a known positive parameter and c is unknown.
(1) Find c to make the pmf valid.
(2) Find $P(X = 0)$.
(3) Find $P(X > 2)$.

Solution :
(1)

$$\begin{aligned} \sum_{x=0}^{\infty} f(x) &= \sum_{x=0}^{\infty} c \cdot \frac{\lambda^x}{x!} \\ &= c \cdot \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} \\ &= c \cdot e^{\lambda} \\ &= 1. \end{aligned}$$

Hence $c = e^{-\lambda}$. And the pmf is then $f(x) = e^{-\lambda} \cdot \frac{\lambda^x}{x!}$.

(2)

$$\begin{aligned} P(X=0) = f(0) &= e^{-\lambda} \cdot \frac{\lambda^0}{0!} \\ &= e^{-\lambda} \cdot \frac{1}{1} \\ &= e^{-\lambda}. \end{aligned}$$

(3)

$$\begin{aligned} P(X>2) &= 1 - P(X \leq 2) \\ &= 1 - (P(X=0) + P(X=1) + P(X=2)) \\ &= 1 - (e^{-\lambda} + e^{-\lambda} \cdot \lambda + e^{-\lambda} \cdot \frac{\lambda^2}{2}) \\ &= 1 - e^{-\lambda}(1 + \lambda + \frac{\lambda^2}{2}). \end{aligned}$$