Some Notes on the random variables:

- ✓ A random variable, named X, can be viewed as a function from the sample space S to a certain subset of R (the set of real values), denote by $\mathbb{R}_X$. As a function, it satisfies all the properties of functions. In particular,
  - ★ Every element in S has one and only one projected value in $\mathbb{R}_X$.
  - ★ However, for every value in $\mathbb{R}_X$, there may exist an arbitrary number of values in S that may be projected to this value.
- ✓ Such a function X defines "equivalent event" between S and $\mathbb{R}_X$. That is, for any subset B of $\mathbb{R}_X$, there is a subset A of S, such that $B = X(A)$. This idea can be written more mathematically:

$$A = \{s \in S | X(s) \in B\}.$$

  Note that since $\mathbb{R}_X$ is the range for X, we have

$$S = \{s \in S | X(s) \in \mathbb{R}_X\}.$$

- ✓ The goal of introducing the equivalent event is to impose probabilities for the elements (subsets) of $\mathbb{R}_X$, which form the "distribution" of X:

$$Pr(B) = Pr(A).$$

Read the notes above together with the example on page 2-6 of the lecture slides. See also pages 2-13 to 2-16 for a full view of this example.

Example 1

- Let $S = \{HH, HT, TH, TT\}$ be a sample space associated with the experiment of tossing two coins.
- Define the random variable (a function)

  X = number of heads obtained.

  $X : S \to \mathbb{R}$, where $\mathbb{R}$ is the set of all real numbers

  such that $X(HH) = 2, X(HT) = 1, X(TH) = 1$ and $X(TT) = 0$.
- In fact the range space, $\boldsymbol{R_X}$, for the random variable X is

  $\{0, 1, 2\}$.

2.2 Discrete Probability Distributions

2.2.1 Discrete Random Variable

Definition 2.3

- Let X be a random variable.
- If the number of possible values of X (i.e., $\boldsymbol{R_X}$, the range space) is **finite or countable infinite**, we call X a **discrete** random variable.
- That is, the possible values of X may be listed as $x_1, x_2, x_3, \cdots$.

✓ Countable or uncountable are both concepts for the number of elements for a set with INFINITELY many elements.

✓ A set is called countable, if we can use a "way" to "count" the set; under such a "way", for an arbitrary element in this set, we can clearly speak out when this element will be counted. For example, the set of positive integers is countable, as we can naturally count $1, 2, 3, 4, \ldots$; the set of integers is also countable, as we can count: $0, 1, -1, 2, -2, 3, -3, \ldots$; the set of the real numbers is uncountable; the set of points between 0 and 1 is uncountable.

Note: the mathematical reasoning of why these sets are uncountable belongs to the content of "measure theory", and is beyond the scope of ST2334.

✓ The elements in a countable set can always be listed as: $x_1, x_2, x_3, \ldots$.

Page 2-22

Probability Function (Continued)

The probability of $X = x_i$ denoted by $f(x_i)$ (i.e. $f(x_i) = \Pr(X = x_i)$, must satisfy the following two conditions.

(1) $f(x_i) \geq 0$ for all x_i.

(2) $\sum_{i=1}^{\infty} f(x_i) = 1$.

These conditions are induced from the properties of probabilities defined on the sample space. In particular, if we define $A_i = \{s \in S | X(s) = x_i\}$, then based on the concepts we discussed in the first page of this document, we have:

★ A_i is the "equivalent set" of $\{x_i\}$, therefore

$$f(x_i) = Pr(X = x_i) = Pr(A_i).$$

★ As $X(\cdot)$ is a "function", therefore, A_i's are disjoint sets.

★ $\mathbb{R}_X$ is the range for X, so $\cup_{i=1}^{\infty} A_i = S$.

Therefore, $1 = Pr(S) = \sum_{i=1}^{\infty} Pr(A_i) = \sum_{i=1}^{\infty} f(x_i)$.

When establishing the probability distribution in some practical problems, make sure that this criteria is satisfied!

Page 2-35

Example 5

- Consider a group of five potential blood donors — A, B, C, D and E — of whom only A and B have type O+ blood.
- Five blood samples, one from each individual, will be typed in random order until an O+ individual is identified.

Prof. Chan used the counting method to solve the question in the lecture video; the lecture slides used the conditional probability method. Make sure that you are able to solve this question by either of these methods.

2.3 Continuous Probability Distributions

2.3.1 Continuous Random Variable

Definition 2.4

- Suppose that $\boldsymbol{R_X}$, the range space of a random variable, X, is an **interval or a collection of intervals.**
- Then we say that X is a **continuous random variable.**

✓ For a discrete random variable, its range must be finite or countable; the random variable has a point mass on each possible value in its range. Here "has a point mass" typically refers to it has positive probability to take the value; for example, "X has a point mass on x_i" means "$P(X = x_i) > 0$".

✓ For a continuous random variable, its range is an interval or a collection of intervals, which means its range is not countable. It has NO point mass on any particular value in its range. This means that for any $x \in \mathbb{R}_X$, we must have $Pr(X = x) = 0$.

This immediately leads to a very good example for the statement: "for a set $B \subset \mathbb{R}_X$, $Pr(X \in B) = 0$ does not imply $B = \emptyset$. " In fact, for any $B \subset \mathbb{R}_X$, if the number of elements in B is countable, we conclude $Pr(X \in B) = 0$.

This also implies if X is a continuous random variable, it is impossible for X to take any particular value. For example, height is a continuous variable and it is impossible to know the exact height of a subject. The value that we get is up to the accuracy provided by an instrument that we use.

✓ Do we have random variables that are in between? Say, can we find a random variable whose range is a collection of intervals, but has point mass on some values in its range? The answer is yes; but is not the focus of this module. Please try to find one such random variable on your own.

Page 2-43 & 44

2.3.2 Probability Density Function

Definition 2.5

- Let X be a continuous random variable.
- The **probability density function (p.d.f.)** $\boldsymbol{f(x)}$, is a function, $f(x)$, satisfying the following conditions:
  1. $f(x) \geq 0$ for all $x \in \boldsymbol{R_X}$,
  2. $\int_{\boldsymbol{R_X}} f(x)\,dx = 1$ or $\int_{-\infty}^{\infty} f(x)\,dx = 1$

     since $f(x) = 0$ for x not in $\boldsymbol{R_X}$.

Probability Density Function (Continued)

Definition 2.5 (Continued)

3. For any c and d such that $c < d$, (i.e. $(c, d) \subset \boldsymbol{R_X}$),

$$\Pr(c \leq X \leq d) = \int_c^d f(x)dx$$

The probability density function (p.d.f.) plays a very similar role as the probability mass function (p.m.f.) for the discrete case. But one should keep in mind that they also have some major difference:

✓ If $f(x)$ is a p.m.f., we must have $f(x) \leq 1$ for any x. In contrast, if $f(x)$ is a p.d.f., this is not necessarily true.

✓ The value of a p.m.f. $f(x)$ has a very obvious probability meaning: it is the probability that the random variable will take the value x. But the value of a p.d.f. $f(x)$ does not have such a meaning (note that it makes nonsense to talk about the probability that the random variable will take a specific value, as this probability is always 0); in stead, the probability that the random variable will be in an interval is evaluated by the areas under the function curve of $f(x)$; the corresponding mathematical statement has been given by item 3 of Definition 2.5 on page 2-44 of the lecture slides.

The p.d.f. reflects how likely the corresponding random variable will fall in a very small neighbourhood of X; namely for a sufficiently small δ, $P(x_0 \leq X \leq x_0 + \delta) \approx f(x_0)\delta$.

For item 2 of Definition 2.5 on page 2-43, we take note that $\int_{\mathbb{R}_X} f(x)dx = 1$ and $\int_{-\infty}^{\infty} f(x)dx = 1$ are exactly the same. This is because for $x \notin \mathbb{R}_X$, $f(x) = 0$, therefore

$$\begin{aligned}\int_{-\infty}^{\infty} f(x)dx &= \int_{x\in\mathbb{R}_X} f(x)dx + \int_{x\notin\mathbb{R}_X} f(x)dx \\ &= \int_{x\in\mathbb{R}_X} f(x)dx + \int_{x\notin\mathbb{R}_X} 0dx = \int_{x\in\mathbb{R}_X} f(x)dx.\end{aligned}$$

Remarks (Continued)

2. For any specified value of X, say x_0, we have

$$\Pr(X = x_0) = \int_{x_0}^{x_0} f(x)dx = 0$$

Hence in the **continuous** case, the probability of X equals to a fixed value is 0 and

$\Pr(c \le X \le d) = \Pr(c \le X < d) = \Pr(c < X \le d) = \Pr(c < X < d)$.

Therefore in the continuous case, $\le$ **and** $<$ can be used **interchangeably** in a probability statement.

Take note of the formulae

$$Pr(c \le X \le d) = Pr(c \le X < d) = Pr(c < X \le d) = Pr(c < X < d),$$

and be aware that they are applicable only when X is a continuous random variable. In general, we have

$$\begin{aligned} Pr(c \le X \le d) &= Pr(c \le X < d) + Pr(X = d) \\ &= Pr(c < X \le d) + Pr(X = c) \\ &= Pr(c < X < d) + Pr(X = c) + Pr(X = d). \end{aligned}$$

Therefore, when X is a discrete random variable, we need to account for whether X has point masses on c and d.

Example 2 (Continued)

- Clearly, $f(x) \geq 0$ and

$$\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{0.5} 0\ dx = \int_{0.5}^{\infty} 0.15\ e^{-0.15(x-0.5)}\ dx$$

$$= 0.15e^{0.075} \int_{0.5}^{\infty} e^{-0.15x}\ dx$$

$$= 0.15e^{0.075} \left[-\frac{1}{0.15} e^{-0.15x} \right]_{0.5}^{\infty}$$

$$= 0.15e^{0.075} \left(0 - \left(-\frac{1}{0.15} \mathrm{e}^{-0.15(0.5)} \right) \right) = 1$$

In line with the lecture video, $\int_{-\infty}^{\infty} f(x)dx = 1$ needs to be checked only when the question asks you to check. For some distributions, checking this may not be an easy task!

2.4 Cumulative Distribution Function

Definition 2.6

- Let X be a random variable, discrete or continuous.

- We define $F(x)$ to be the **cumulative distribution function** of the random variable X (abbreviated as c.d.f.) where

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

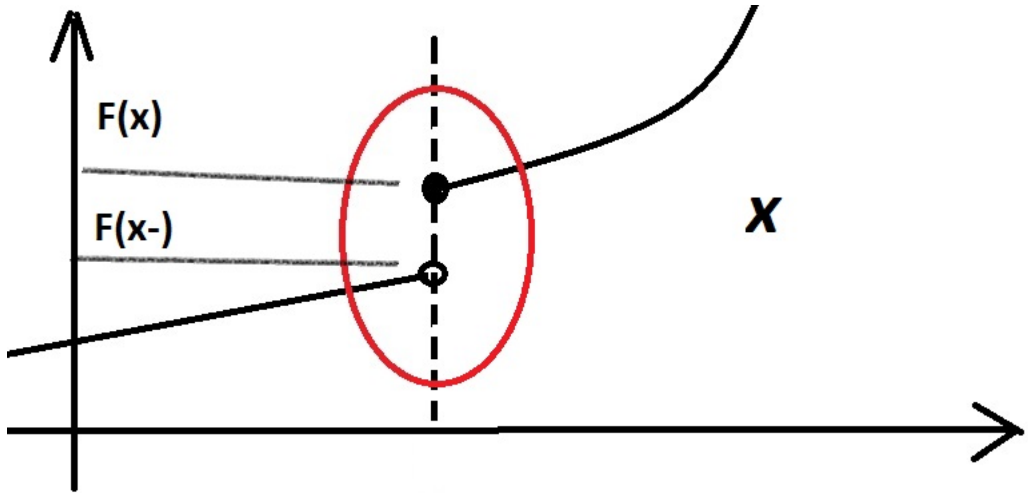
For any random variable, discrete or continuous or "in between", the cumulative distribution function (c.d.f.) is always defined as the one given in this slide.

If $F(x)$ is a c.d.f. for some random variable, then it satisfies

✓ $F(x)$ is a nondecreasing function of x.

✓ We always have $F(x) \to 1$, as $x \to \infty$; $F(x) \to 0$, as $x \to -\infty$.

✓ For every $x \in \mathbb{R}$, $F(x)$ is either continuous, or if it is not continuous, it must be right continuous at x and the left limit exists. In a figure, it must be like this:



Mathematically, it means for any x

$$\lim_{t \to x+} F(t) = F(x) \qquad \lim_{t \to x-} F(t) \text{ exists,}$$

the convergence result of the latter one is denoted as $F(x-)$. See the difference of $F(x)$ and $F(x-)$ in the figure. We have $Pr(X = x) = F(x) - F(x-)$, which is the point mass of random X at x.

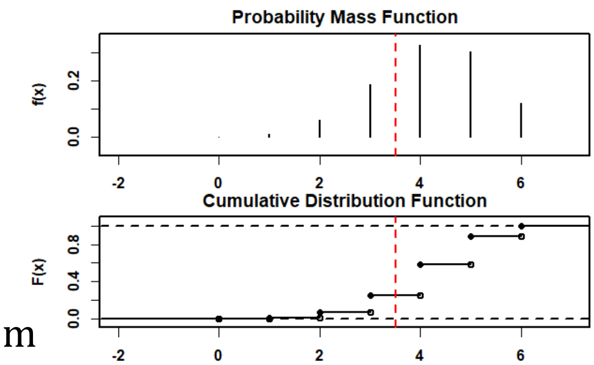
✓ When X is a discrete random variable, its c.d.f. must be a step function in a similar structure as the one given in page 2-60 of the lecture slide. Take note that the mathematical formula of its c.d.f. is also given in this slide.

2.4.1 CDF for Discrete Random Variables

- If X is a **discrete random variable**, then

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

- The c.d.f. of a discrete random variable is a step function.

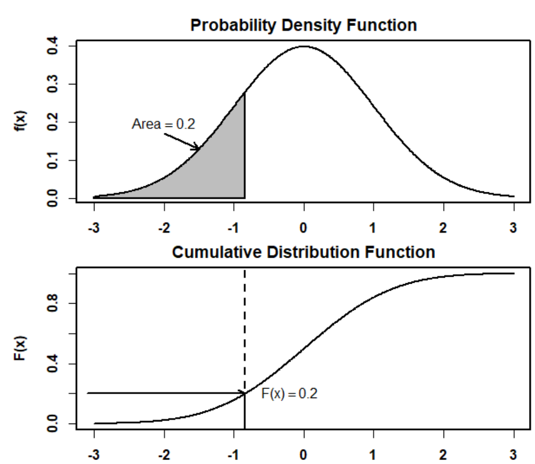


✓ When X is a continuous random variable, its c.d.f. must be a continuous function, with the c.d.f. in a similar structure as the one given in page 2-63 of the lecture slides. This page also provides the mathematical formula of its c.d.f.

2.4.2 CDF for Continuous Random Variables

- If X is a **continuous random variable**, then

$$F(x) = \int_{-\infty}^{x} f(t)dt$$



CDF for Continuous Random Variables (Continued)

- For a **continuous random variable** X,

$$f(x) = \frac{d\,F(x)}{dx}$$

if the derivative exists.

- Also,

$$\begin{aligned}\Pr(a \le X \le b) &= \Pr(a < X \le b) \\ &= F(b) - F(a).\end{aligned}$$

This page gives how we can obtain the p.d.f. when we have the c.d.f. of a continuous random variable in hand.

When X is a discrete random variable, we have $f(x) = F(x) - F(x-)$. So,

✓ when $F(x)$ is continuous at x, $f(x) = 0$;

✓ when $F(x)$ is not continuous at x, then $f(x) > 0$;

therefore, the discontinuous points of $F(x)$ attributes to all the point masses of the corresponding p.d.f.