Continuous Random Variables

- Recall: For *discrete* random variables, only a *finite* or *countably infinite* number of possible values with positive probability (>0).

  ➢ Often, there is interest in random variables that can take (at least theoretically) on an *uncountable* number of possible values, e.g., c.f. | c.f. | defined by prob. ≥0? Ans. NO.

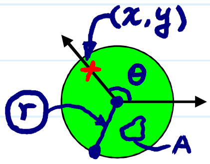
  Note. observed data, always discrete

  $(0, 800)$ ← the weight of a randomly selected person in a population, ← X

  $(0, \infty)$ ← the length of time that a randomly selected light bulb works, ← X

  $(-a, a)$ for some a, or $(-\infty, \infty)$ ← the error in experimentally measuring the speed of light. ← X

  ➢ Example (Uniform Spinner, LNp.3-6, 3-18):



a probability space $(\Omega, \mathcal{F}, P)$

- $\Omega = (-\pi, \pi]$ or $\Omega^* = \{(x,y) \mid x^2+y^2 \le r^2\}$
- For $(a, b] \subset \Omega$, $P((a, b]) = (b-a)/(2\pi)$ or $P(A) = \text{area}(A)/\pi r^2$ for $A \subset \Omega^*$

$P(\{\omega\}) = 0$ for $\forall \omega \in \Omega$

- Consider the random variables:

  $X: \Omega \to \mathbb{R}$, and $X(\omega) = \omega$ for $\omega \in \Omega$,

  $\Omega^* \to \mathbb{R}$, $(x,y) \to \theta$

  Range of X: $(-\pi, \pi]$ ← uncountable set

  Ω, $\{E_A\}$, P → $\mathbb{R}$, $\{A\}$, P_X (X)

$Y: \Omega \to \mathbb{R}$, and $Y(\omega) = \tan(\omega)$ for $\omega \in \Omega$.

Range of Y: $(-\infty, \infty)$ ← uncountable set

$\Omega \to \mathbb{R}$ or $\Omega^* \to \mathbb{R}$

Then, X and Y are random variables that takes on an uncountable number of possible values. ← defined by the "range of r.v."

Recall. LNp.3-18

- Some properties about the distribution of X (or Y), i.e. P_X (or P_Y)
  - $P_X(\{X = x\}) = P(\{x\}) = 0$, for any $x \in \mathbb{R}$. ← $A \subset \mathbb{R}$ ← $E_A \subset \Omega$

    ⇒ Probability for X to take *any single* value is zero

  - But, for $-\pi \le a < b \le \pi$, c.f.

    $P_X(\{X \in (a, b]\}) = P((a, b]) = (b-a)/(2\pi) > 0$. ← $A \subset \mathbb{R}$ ← $E_A \subset \Omega$

    ⇒ Positive probability (>0) is assigned to *any* $(a, b]$

    Why it happen? Note uncountable sum.

- **Q**: Can we still define a probability mass function for X? No. ∵ pmf$(x) = P_X(X=x)$

  Q: If not, what can play a *similar* role like pmf for X?

discrete	continuous
pmf	pdf
countable	uncountable
Σ	$\int$

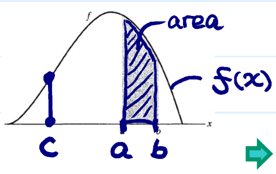
c.f.

Recall. Find area under a curve by integration (*uncountable* sum)

$f(x) \ge 0$

area$((a,b)) = \int_a^b f(x)dx$

$f(c) > 0$, but area$(\{c\}) = \int_c^c f(x)dx = 0$



- **Probability Density Function and Continuous Random Variable**

↳ it plays a similar role like pmf for discrete r.v

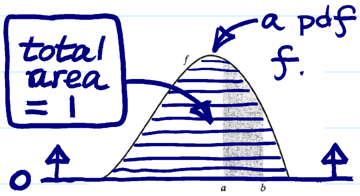
➤ Definition. A function f: $\mathbb{R} \to \mathbb{R}$ is called a probability density function (pdf) if density: 密度, mass = density × volume

LNp.5-6, (i)(ii)(iii) c.f.

1. $f(x) \geq 0$, for all $x \in (-\infty, \infty)$, and
2. $\int_{-\infty}^{\infty} f(x)dx = 1$.

f: not necessary to be a continuous function

total area = 1



f(x): not probability cf area = probability

➤ Definition: A random variable X is called *continuous* if there exists a pdf f such that for any set B of real numbers

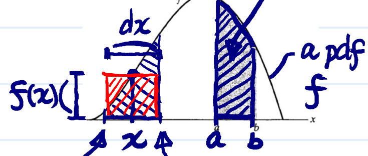
$P(X \in B) = \sum_{x \in B \cap \mathcal{X}} f(x)$

LNp.5-6, (iv) c.f.

area

$$P_X(\{X \in B\}) = \int_B f(x)\, dx.$$

can be regarded as a small prob.



- For example, $P_X(a \leq X \leq b) = \int_a^b f(x)dx$. $\underbrace{}_{B=[a,b]}$

$x - \frac{dx}{2}$ $x + \frac{dx}{2}$

red area = $f(x)\,dx \approx P_X(x - \frac{dx}{2} < X < x + \frac{dx}{2})$

c.f. Theorem in LNp.5-7

➤ Theorem. If f is a pdf, then there must exist a continuous random variable with pdf f.

Sketch of proof. Let $F(x) \stackrel{\text{let}}{=} \int_{-\infty}^{x} f(t)dt$, then show that $F(x)$ is a cdf (exercise, check LNp.5-9~10, (2)(3)(4)). Then, by the theorem in LNp.5-11, there exist a r.v. X s.t. $P_X((a,b]) = F(b) - F(a) = \int_a^b f(x)dx$.

Why? ∵ $F_X(x)$ is defined as $P_X(X \in (-\infty, x])$

➤ Some properties $\because F_X(x) = P(X \in (-\infty, x]) = P(X \in (-\infty, x)) = F_X(x-)$

The cdf of X is a continuous function, i.e. no jumps

- $P_X(\{X = x\}) = \int_x^x f(y)dy = 0$ for any $x \in \mathbb{R}$

- It does not matter whether the intervals are open or close, i.e.,

$\int_a^b f(x)dx$ ‖

$$P(X \in [a, b]) = P(X \in (a, b]) = P(X \in [a, b)) = P(X \in (a, b)).$$

Q: Is the pdf of a continuous r.v. X unique? Hint. Consider $f^*(x) = f(x)$ except at countably many x's

Q: Is the pmf of a discrete r.v. unique? Ans. Yes.

- It is important to remember that the *value* of a pdf $f(x)$ is **NOT** a probability itself

• area = prob.
• but $f(x)$ is not prob.

cf. pmf

- It is quite possible for a pdf to have value greater than 1 cf. pmf ≤ 1

- **Q**: How to interpret the value of a pdf $f(x)$? For small dx,

$$P\left(x - \frac{dx}{2} \leq X \leq x + \frac{dx}{2}\right) = \int_{x-\frac{dx}{2}}^{x+\frac{dx}{2}} f(y)dy \approx f(x) \cdot dx.$$

Ans in LNp.5-5 cf.

$\Rightarrow$ $f(x)dx$ is a measure of how likely it is that X will be near x

i.e. if $f(a) > f(b)$, then X is more likely to take value near a than value near b, e.g. $\frac{f(a)}{f(b)} = 2$ meaning?

➤ We can characterize the distribution of a continuous random variable in terms of its

1. Probability Density Function (pdf)
2. (Cumulative) Distribution Function (cdf)
3. Moment Generating Function (mgf, Chapter 7)

Note cdf is defined for any r.v.'s (LNp.5-8)

- Relation between the pdf and the cdf

cf. relationship between pmf & cdf (LNp.5-11,(6))

➢ Theorem. If F_X and f_X are the cdf and the pdf of a continuous random variable X, respectively, then

- $F_X(x) = P(X \le x) = \int_{-\infty}^{x} f_X(y)dy$ for all $-\infty<x<\infty$

by the definition of cdf

"$P(X\in(-\infty,x])$"

Fundametal Thm of Calculus

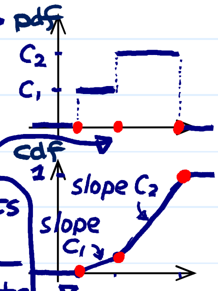
- $f_X(x) = F_X'(x) = \frac{d}{dx}F_X(x)$ at continuity points of f_X

$\because \int_a^b F_X'(x)dx = F_X(b)-F_X(a)$

For continuous r.v. X
when F_X is given $\Rightarrow f_X$ is known
when f_X is given $\Rightarrow F_X$ is known

can define $f_X(x)=0$ at the discontinuity points

not a continuous function



●: discontinuous pts of f_X
●: nondifferential pts of F_X

➢ Some Notes

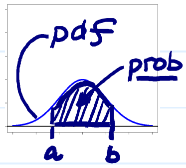
- For $-\infty\le a<b\le \infty$

LNp.5-10,(5)

$P(X\in(a,b])$

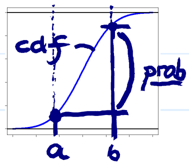
$$P(a < X \le b) = F_X(b) - F_X(a) = \int_a^b f_X(x)dx.$$

$= \int_a^b F_X'(x)dx = F_X(x)\Big|_a^b$



- The cdf for continuous random variables has the *same* interpretation and properties as discussed in the discrete case

LNp.5-9~11



better way to define discrete & continuous r.v.'s

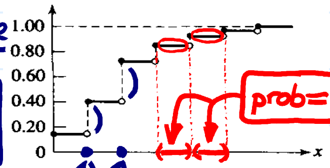
- The only difference is in plotting F_X. In discrete case, there are *jumps* (step function). In continuous case, F_X is a (absolutely) *continuous* non-decreasing function.

cf.

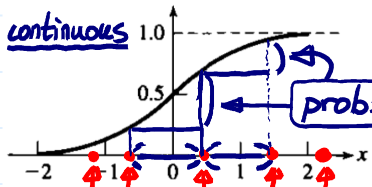
F_X is differentiable almost everywhere.

discrete

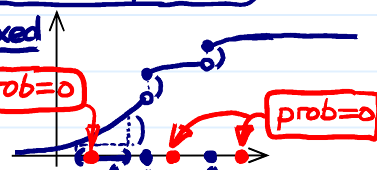
cdf increases ∵ jumps



continuous



mixed



➢ Example (Uniform Distributions)

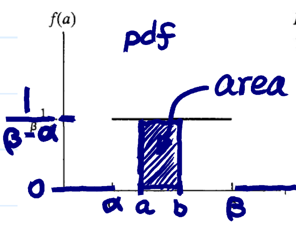
- If $-\infty<\alpha<\beta<\infty$, then

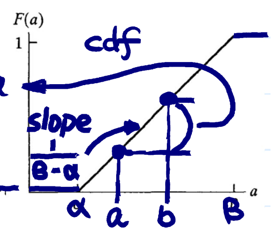
$\alpha<a<b<\beta$
$\frac{f(a)}{f(b)}=1$

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

is a pdf since

1. $f(x) \ge 0$ for all $x\in\mathbb{R}$, and
2. $\int_{-\infty}^{\infty} f(x)dx = \int_\alpha^\beta \frac{1}{\beta-\alpha}dx = \frac{1}{\beta-\alpha}(\beta-\alpha) = 1.$





For $\alpha<a<b<\beta$,
$\int_a^b f(x)dx = \frac{b-a}{\beta-\alpha}$.
① For intervals $\subset(\alpha,\beta)$ with the same length, their prob. are identical.
② prob. $\propto$ length of interval.

- Its corresponding cdf is

$$F(x) = \int_{-\infty}^{x} f(y)dy = \begin{cases} 0, & \text{if } x \le \alpha, \\ \frac{x-\alpha}{\beta-\alpha}, & \text{if } \alpha < x \le \beta, \\ 1, & \text{if } x > \beta. \end{cases} \qquad \frac{1}{\beta-\alpha} y \Big|_{\alpha}^{x}$$

- (exercise) Conversely, it can be easily checked that F is a cdf and $f(x)=F'(x)$ except at $x=\alpha$ and $x=\beta$ (Derivative does not exist when $x=\alpha$ and $x=\beta$, but it does not matter.)

check LNp.5-9~10 (2)(3)(4)

∵ we can assign $f(\alpha)=0, f(\beta)=0$ or any nonegative values. It has no impact on the calculation of prob.

- An example of Uniform distribution is the r.v. X in the Uniform Spinner example (LNp.6-1) where $\alpha=-\pi$ and $\beta=\pi$.

• Transformation

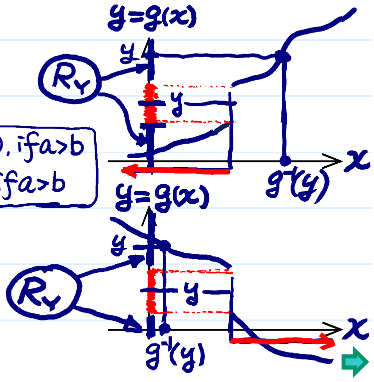
cf. discrete case, LNp 5-12

➢ **Q**: $Y=g(X)$, how to find the distribution of Y?

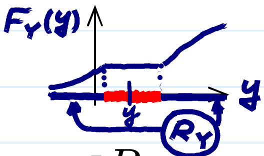
- Suppose that X is a continuous random variable with cdf F_X and pdf f_X.

- Consider $Y=g(X)$, where g is a strictly monotone (increasing or decreasing) function. Let R_Y be the range of g.

i.e., $g(a)>g(b)$, if $a>b$ or $g(a)<g(b)$, if $a>b$



- Note. Any strictly monotone function has an inverse function, i.e., g^{-1} exists on R_Y.



➢ The cdf of Y, denoted by F_Y

$=g(X)$

(1) Suppose that g is a strictly increasing function. For $y \in R_Y$,

$$\begin{aligned} F_Y(y) &= P(Y \le y) \\ &= P(g(X) \le y) = P(X \le g^{-1}(y)) \\ &= F_X(g^{-1}(y)). \end{aligned}$$

same event

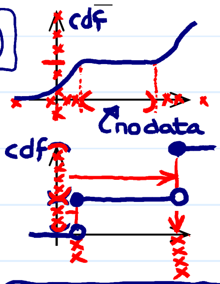
True for any r.v.'s, including $X <$ discrete, continuous and $Y <$ discrete, continuous

Q: what if $y \notin R_Y$?

(2) Suppose that g is a strictly decreasing function. For $y \in R_Y$,

$$\begin{aligned} F_Y(y) &= P(Y \le y) \\ &= P(g(X) \le y) = P(X \ge g^{-1}(y)) \\ &= 1 - P(X < g^{-1}(y)) = 1 - F_X(g^{-1}(y)-) \text{ in general} \\ &\overset{\because X \text{ is a continuous r.v.}}{=} 1 - F_X(g^{-1}(y)). \end{aligned}$$

$P(X \in (-\infty, g^{-1}(y))) = F_X(g^{-1}(y)-)$



F_X is
• nondecreasing
• one-to-one
⊙ strictly increasing
• no jump
⊙ continuous

- Theorem. Let X be a continuous random variable whose cdf F_X possesses a unique inverse F_X^{-1}. Let $Z=F_X(X)$, then Z has a uniform distribution on [0, 1].

$F_X = g \Rightarrow g^{-1} = F_X^{-1}$. $R_Y = [0,1]$

cdf of uniform distribution with $\alpha=0$, $\beta=1$

Proof. For $0 \le z \le 1$, $F_Z(z) = F_X(F_X^{-1}(z)) = \begin{cases} 0, & \text{if } z<0. \\ z. \\ 1, & \text{if } z>1. \end{cases}$

- Theorem. Let U be a uniform random variable on [0, 1] and F is a cdf which possesses a unique inverse F^{-1}. Let $X=F^{-1}(U)$, then the cdf of X is F.

Assume strictly increasing & continuous

$F^{-1}=g \Rightarrow g^{-1}=F$

can be relaxed to any cdfs, including discrete, continuous, mixed

Proof. $F_X(x) = F_U(F(x)) = P(U \le F(x)) = F(x)$.

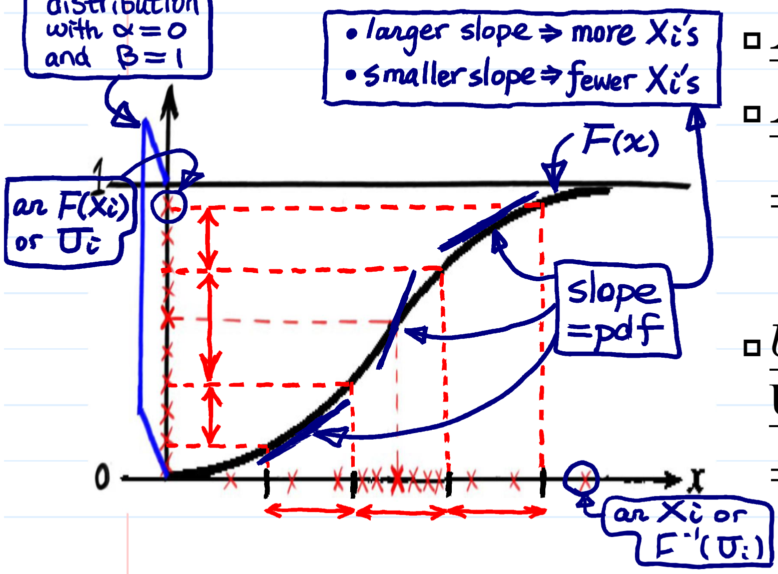
- The 2 theorems are useful for pseudo-random number generation in computer simulation.

亂數

use computer to generate random numbers with a pre-specified distribution.

pdf of uniform distribution with $\alpha=0$ and $\beta=1$

$\Rightarrow$ The key is to generate U(0, 1) random numbers.



- X is r.v. $\Rightarrow$ $F(X)$ is r.v.
- $X_1, \ldots, X_n$: r.v.'s with cdf F (independent)

  $\Rightarrow$ $F(X_1), \ldots, F(X_n)$: r.v.'s with distribution Uniform(0, 1)
- $U_1, \ldots, U_n$: r.v.'s with distribution Uniform(0, 1) (independent)

  $\Rightarrow$ $F^{-1}(U_1), \ldots, F^{-1}(U_n)$: r.v.'s with cdf F