DCP 1206: Probability
Lecture 9 — Probability Density Function and Continuous Random Variables

Ping-Chun Hsieh

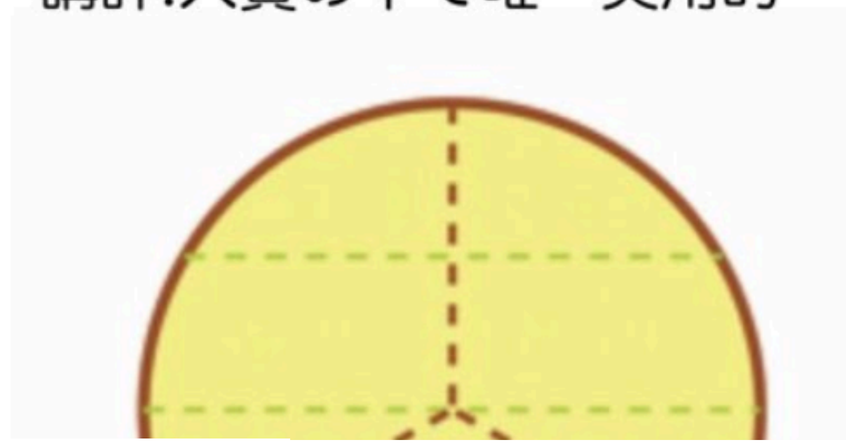
October 16, 2019

Announcements

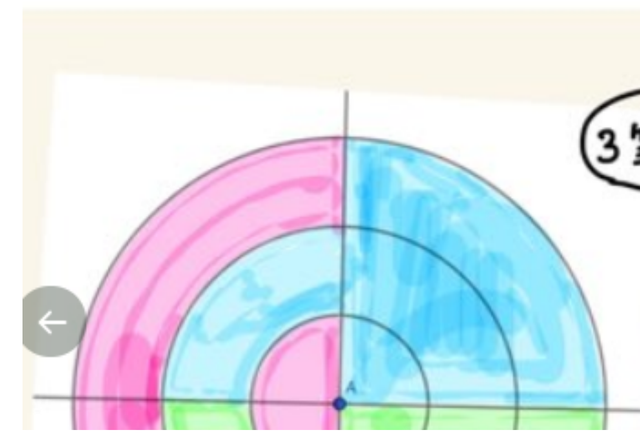
- HW2 is posted on E3 (Due: 10/23 in class)
- Solution to HW1 is also on E3
- Please contact the TAs if you have any questions about HW1
  - 洪偉：hjkl12342002@gmail.com
  - 李昕：sphfs196@gmail.com

How to Cut a Cake into 3 Equal Pieces?

入賞
4等分線をイメージして切る
講評:入賞の中で唯一実用的

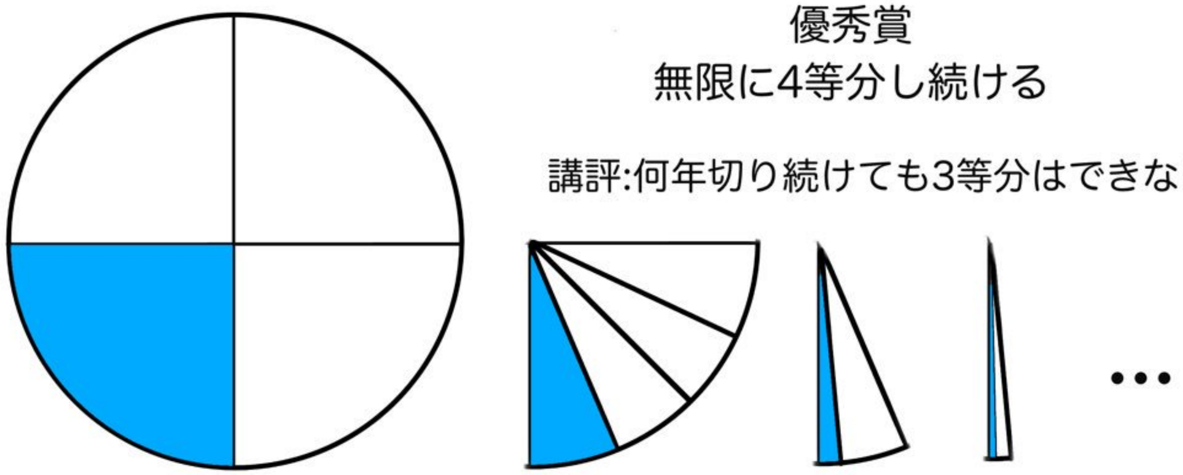


入賞
同心円で切る

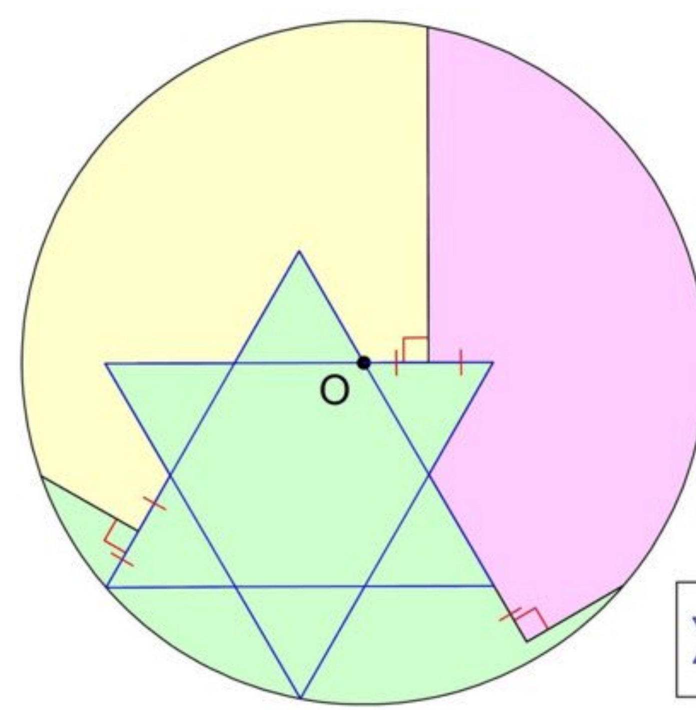


優秀賞
無限に4等分し続ける

講評:何年切り続けても3等分はできない



$$\frac{1}{4}+\frac{1}{4^2}+\frac{1}{4^3}+\frac{1}{4^4}+\cdots=\frac{1}{3}$$



最優秀賞
六芒星で切る

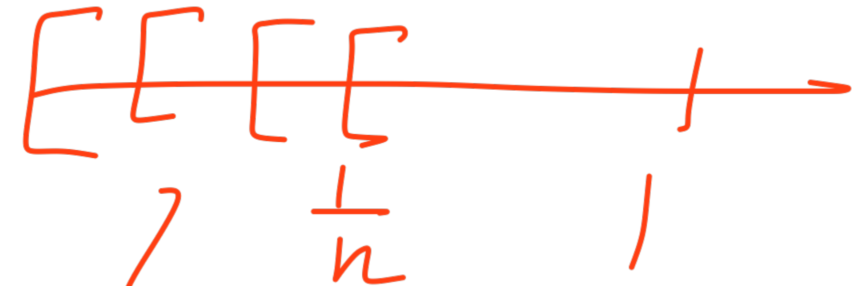
講評:
どうやって思いついた
のかすら謎。
ケーキを切るときに
マウントを取れる。

https://twitter.com/mathlava/status/1163796284760662017?s=21

HW1

- Problem 2: Prove the general form of De Morgan's laws

$$\left(\bigcup_{n=1}^{\infty} S_n\right)^c = \bigcap_{n=1}^{\infty} S_n^c$$



- Finite Union vs Countable Union
- Induction cannot be used to prove "infinite" arguments
- Example: $\bigcup_{n=1}^{k} [\frac{1}{n}, 1] = [\frac{1}{k}, 1]$ is a closed interval

$\bigcup_{n=1}^{\infty} [\frac{1}{n}, 1] = (0, 1]$ is NOT a closed interval

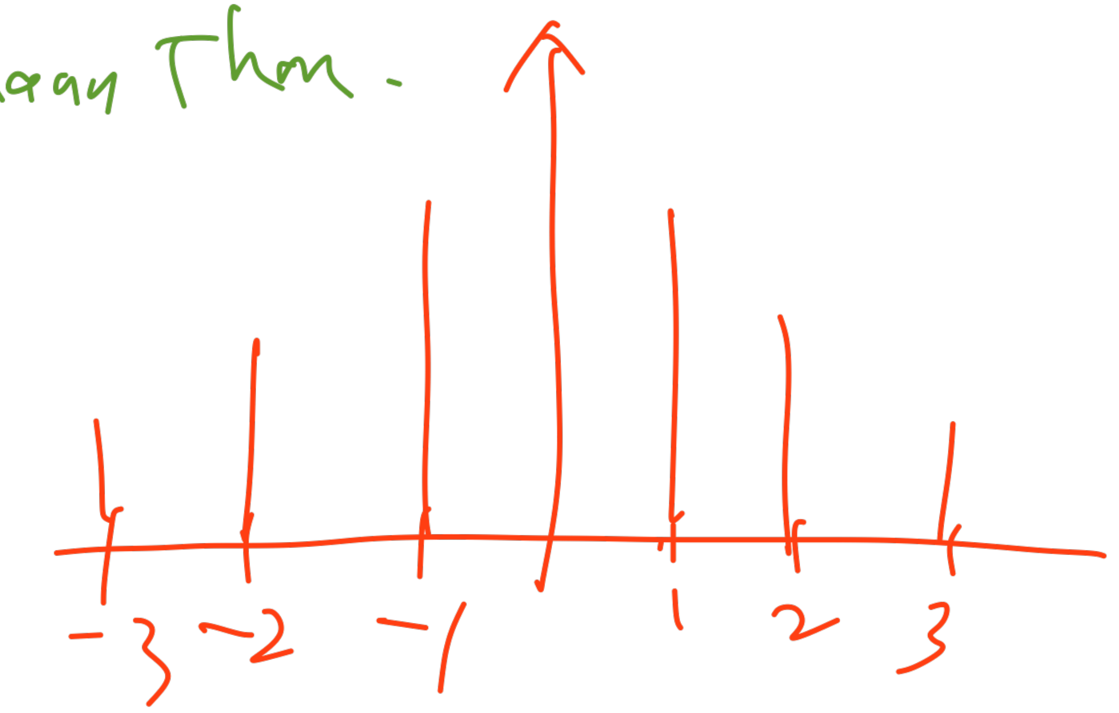
Existence of Moments

- Question: Suppose X is a random variable with PMF $p_X(x)$

$$p_X(k) = \begin{cases} \dfrac{1}{2k(k+1)} & , k = 1,2,3,\cdots \\ \dfrac{1}{2k(k-1)} & , k = -1, -2, -3,\cdots \end{cases}$$

- Does $E[X]$ exist? $\Rightarrow$ Use Riemann Thm.

$$E[X] = \sum_{\text{all } x} x \cdot p_X(x)$$



$$E[|X|] = \infty$$

Rearrangement of Series

- Example: Consider a series $\{a_n\}$: $1, -1, \frac{1}{2}, -\frac{1}{2}, \frac{1}{3}, -\frac{1}{3}, \cdots$ 2k
    - What is $\sum_{n=1}^{\infty} a_n$? $= 0$

$\sum_{n=1}^{\infty} |a_n| = \infty$

$\ln M - 1$ $\ln 2M - 0$

- Example: Rearrange $\{a_n\}$ as $\{b_n\}$:

$\ln M \approx 1, \frac{1}{2}, \frac{1}{3}, \cdots, \frac{1}{M}, -1, \underbrace{\frac{1}{M+1}, \cdots, \frac{1}{2M}}_{M}, -\frac{1}{2}\cdots, \quad - -$

M

- What is $\sum_{n=1}^{\infty} b_n$? $\approx \ln M$

Riemann Rearrangement Theorem

Riemann Rearrangement Theorem:
Let $\{a_n\}$ be a sequence of numbers. If $\{a_n\}$ satisfies that

1. $\sum_{n=1}^{\infty} a_n$ converges
2. $\sum_{n=1}^{\infty} |a_n| = \infty$

1, 100, 10000, ... ∞

0

Then, for any $B \in \mathbb{R} \cup \{\infty\}$, there exists a rearrangement $\{b_n\}$ of $\{a_n\}$ such that $\sum_{n=1}^{\infty} b_n = B$

Existence of Moments (Formally)

Existence of Moments:
Let X be a random variable. Then, the n-th moment of X (i.e. $E[X^n]$) is said to exist if $E[\,|X^n|\,] < \infty$

absolute

This Lecture

1. Expected Value and Variance of Special Discrete Random Variables

2. Probability Density Function (PDF)

3. Expected Value and Variance of Continuous Random Variables

- Reading material: Chapter 6.1-6.3

1. Expected Value and Variance of Special Discrete Random Variables

1. Bernoulli Random Variables

- Example: $X \sim \text{Bernoulli}(p)$
  - How to show that $E[X] = p$?
  - How to show that $\text{Var}[X] = p(1-p)$?

PMF: $P(X=0) = 1-p$

$P(X=1) = p$

$$E[X] = \sum_{x=0,1} x \cdot P_X(x) = p \cdot 1 + (1-p) \cdot 0 = p$$

$$\text{Var}[X] = E[X^2] - (E[X])^2 = p - (p)^2$$

$$= p(1-p)$$

⇓ maximized at $p = \frac{1}{2}$

$$E[X^2] = \sum_{x=0,1} x^2 \cdot p_X(x) = p$$

LOTUS

2. Binomial Random Variables

- Example: $X \sim \text{Binomial}(n, p)$
  - How to show that $E[X] = np$?
  - How to show that $\text{Var}[X] = np(1-p)$?

$$\text{Var}[X] = n\cdot(n-1)p^2 + np - (np)^2 = np(1-p)$$

PMF: $P(X=k) = C_k^n \cdot p^k \cdot (1-p)^{n-k}$

$$E[X] = \sum_{k=1}^{n} k\cdot P(X=k) = \sum_{k=1}^{n} k \cdot C_k^n \cdot p^k \cdot (1-p)^{n-k}$$

$$= \sum_{k=1}^{n} \cancel{k} \frac{n!}{\cancel{k!}\,(n-k)!} p^k \cdot (1-p)^{n-k} \qquad (k-1)!$$

$$= n \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!\,(n-k)!} p^k \cdot (1-p)^{n-k}$$

$$= n\cdot p \left(\sum_{k=1}^{n} C_{k-1}^{n-1} p^{k-1} (1-p)^{n-k} \right)$$

$$\sum_{k=1}^{n} C_{k-1}^{n-1} p^{k-1} (1-p)^{n-k} = 1$$

pmf of Bino$(n-1, p)$

$$\text{Var}[X] = E[X^2] - (E[X])^2$$

$$E[X^2] = \sum_{k=1}^{n} k^2 \cdot P(X=k) = \sum_{k=1}^{n} \cancel{k(k-1)} \cdot \frac{n!}{\cancel{k!}\,(n-k)!} p^k \cdot (1-p)^{n-k} = n\cdot p \ldots$$

$$k^2 = k(k-1) + k$$

$$(k-2)!$$

$$n\cdot(n-1)p^2$$

$$= n\cdot(n-1)p^2 + np$$

Tricks For Deriving $E[X]$ and $\mathrm{Var}[X]$?

1. Reuse $\sum_x p(x) = 1$ and $E[X] = \sum_x xp(x)$ (凑答案)

2. View X as a sum of independent random variables (3~4 weeks)

3. Moment generating functions (4~5 weeks)

3. Poisson Random Variables

- Example: $X \sim \text{Poisson}(\lambda, T)$
  - How to show that $E[X] = \lambda T$?
  - How to show that $\text{Var}[X] = \lambda T$?

PMF:

$$P(X=n) = \frac{e^{-\lambda T} \cdot (\lambda T)^n}{n!}$$

$$E[X] = \sum_{n=1}^{\infty} n \cdot \frac{e^{-\lambda T} \cdot (\lambda T)^n}{n!}$$

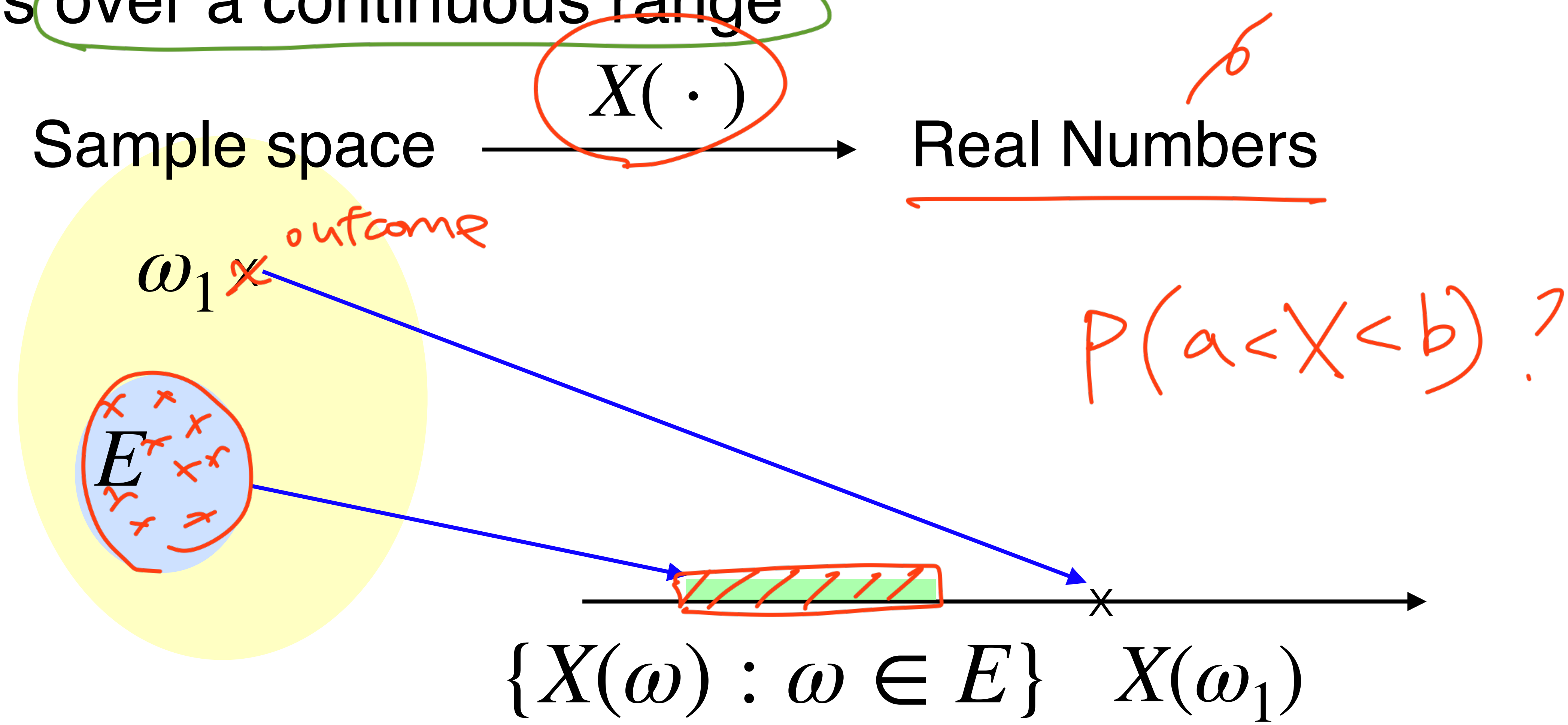
$$= \lambda T \cdot \sum_{n=1}^{\infty} \frac{e^{-\lambda T} \cdot (\lambda T)^{n-1}}{(n-1)!} = \lambda T$$

$\text{Var}[X] =$ BJ4

2. Continuous Random Variables and Probability Density Functions

Continuous Random Variables

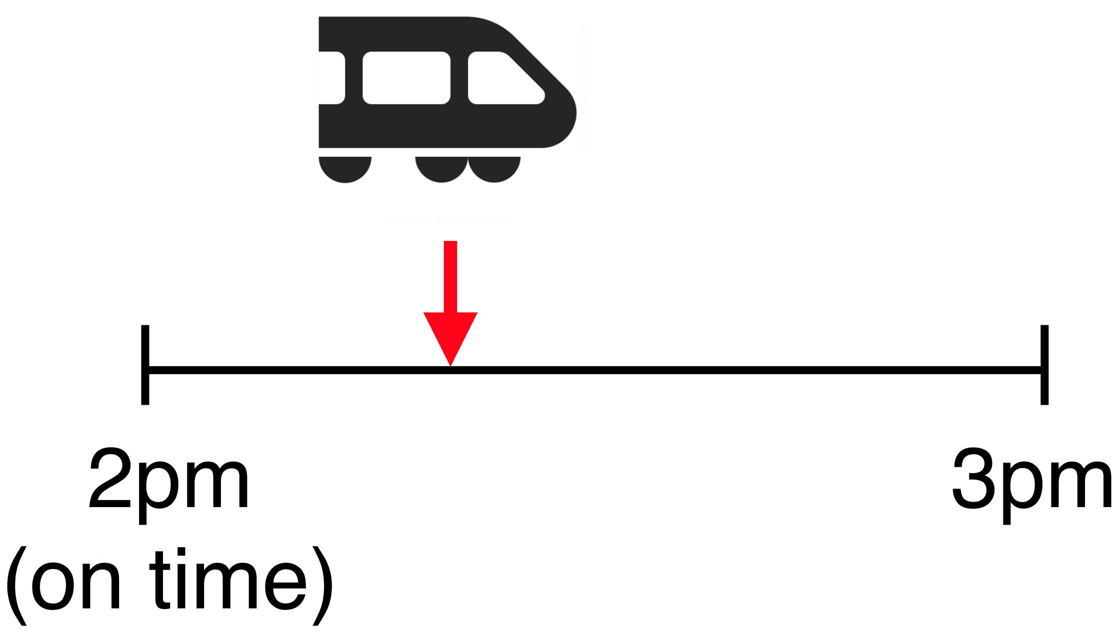
- Continuous random variable: A random variable that takes values over a continuous range



- CDF is still available for a continuous random variable
- How about PMF?

Continuous Random Variables and PMF?

- Example: Train arrival time is between 2pm-3pm (equally likely)

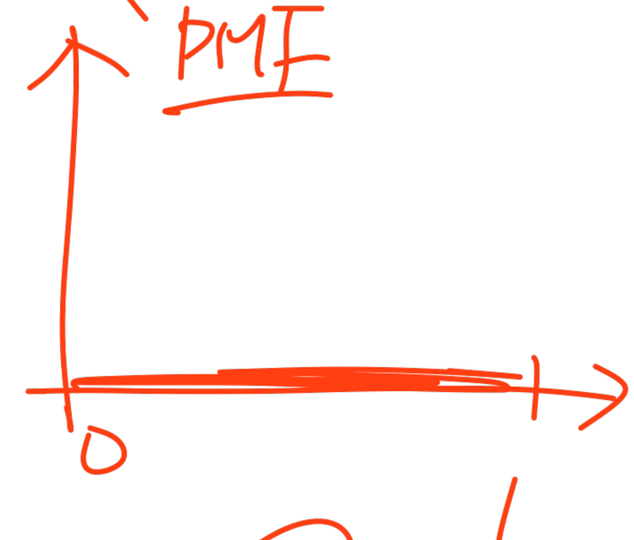


- How to define a random variable?

$$X = \frac{\text{arrival time (in min)} - 2\text{pm}}{60 \text{ min}}$$

?

$X \in [0, 1]$



- P(arrives at exactly 20 min 31.537 sec after 2pm) = ? 0

?

Density / Concentration

B-52 bomber

- Example: B-52 Cocktail

infinitely thin

orange liqueur (40%): 10 ml

milk wine (17%): 10 ml

coffee liqueur (23%): 10 ml

- How much alcohol in total?

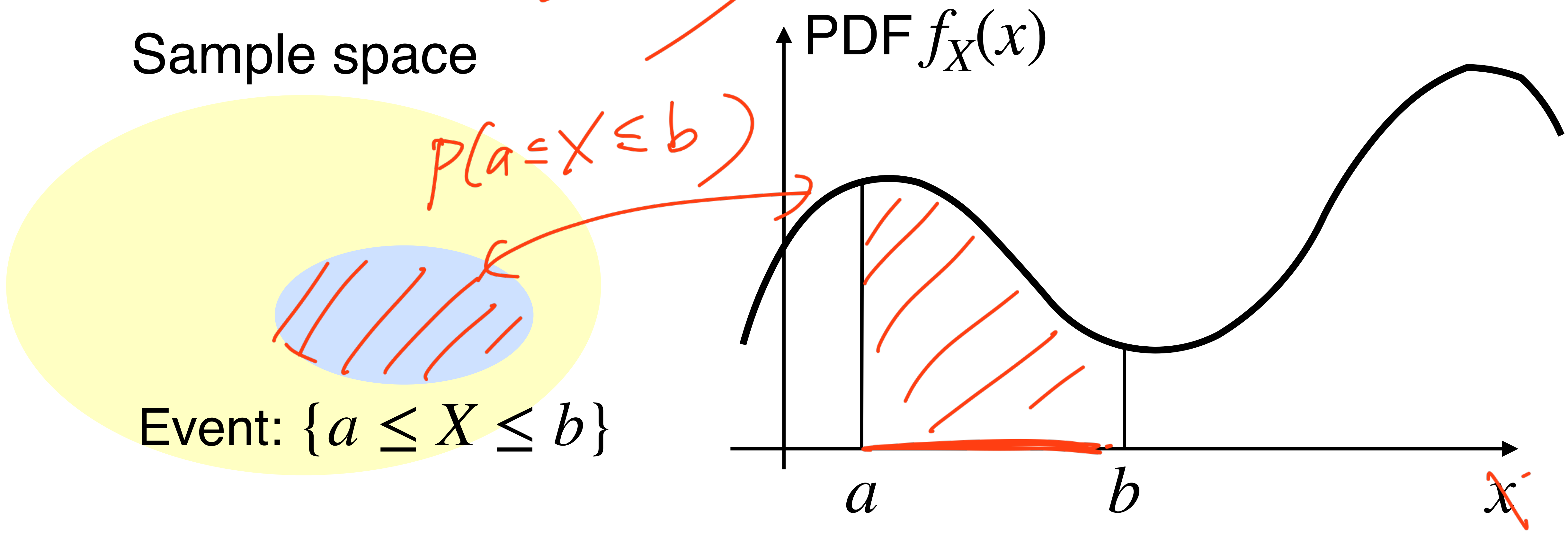
$10ml \times 23\% + 10ml \times 17\% + 10ml \times 40\% = 8ml$

- How much alcohol in the <u>green</u> cross section?

0

Probability Density Function (PDF)

consistency



Probability Density Function (PDF):
Let X be a random variable. Then, $f_X(x)$ is the PDF of X if for every subset B of the real line, we have

$$P(X \in B) = \int_B f_X(x)dx$$

Express Other Quantities Using PDF

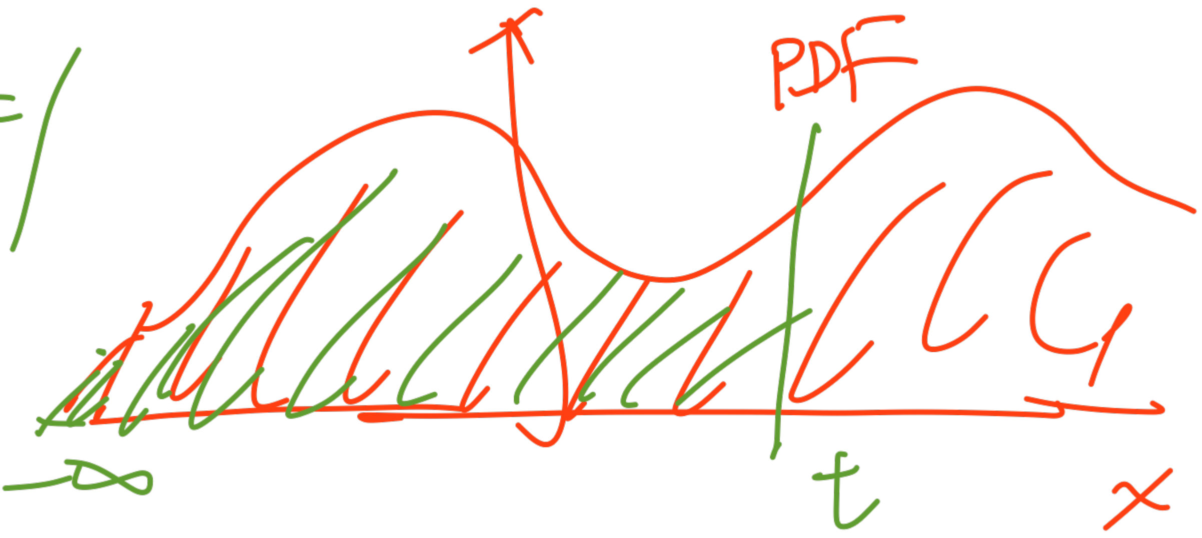
PDF $\rightarrow$ CDF

X, : PDF $= f(x)$

1. $P(X \in \mathbb{R}) = \int_{-\infty}^{+\infty} f(x)\,dx = 1$

$F_X(t)$

2. $P(X \leq t) = \int_{-\infty}^{t} f(x)\,dx$



3. $P(a \leq X \leq b) = \int_{a}^{b} f(x)\,dx$

4. $P(a \leq X < b) = \int_{a}^{b} f(x)\,dx - \underbrace{P(X=b)}_{=0}$

5. $P(a < X < b) = \int_{a}^{b} f(x)\,dx - \underbrace{P(X=a)}_{=0} - \underbrace{P(X=b)}_{=}$

How to Check if a PDF is Valid?

$f(x)$

- Recall: 3 Axioms of Probability

1. $P(X \in \mathbb{R}) = 1$

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

2. $P(X \in A) \geq 0$, for all A

$$\int_A f(x)dx \geq 0$$

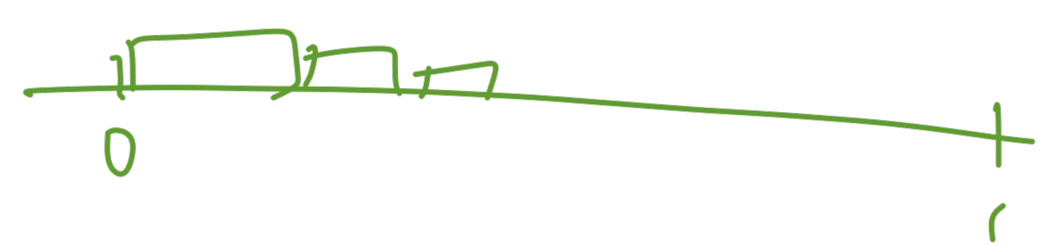
Sufficient condition: $f(x) \geq 0$ for all x

by the def. of integration

3. Let $A_1, A_2, \cdots$ be mutually exclusive sets of real numbers, then

$$P(X \in \bigcup_{i \geq 1} A) = \sum_{i \geq 1} P(X \in A_i)$$

$$A_i = \left(\frac{1}{i+1}, \frac{1}{i}\right)$$



Example: From PDF to CDF (I)

- Example: Consider the following PDF

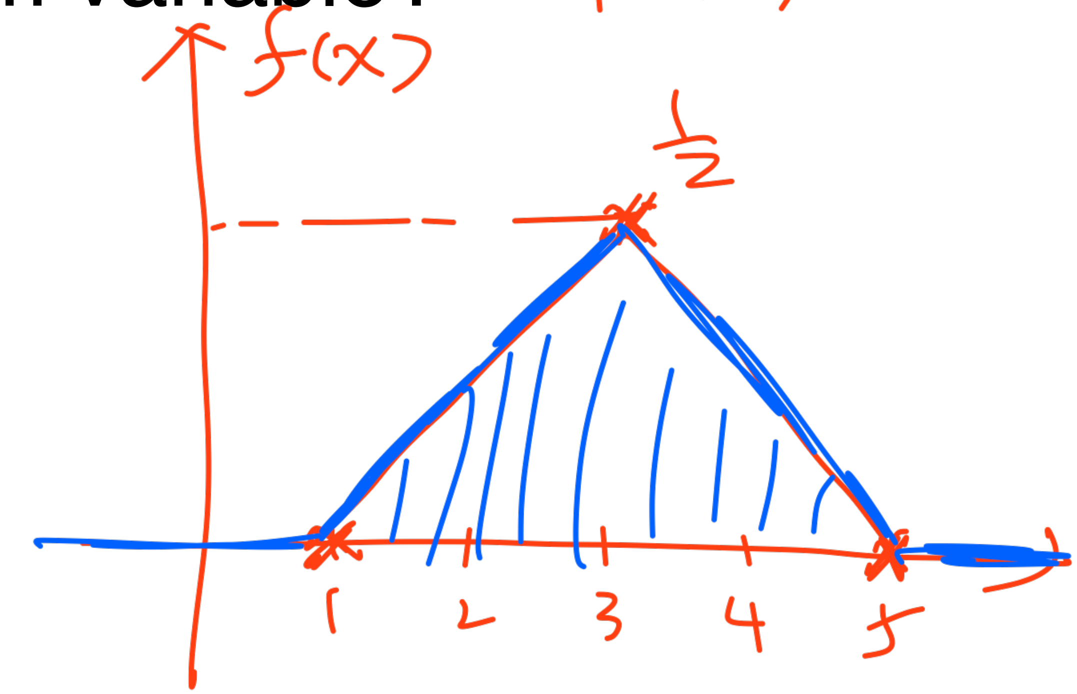
$$f(x) = \begin{cases} \frac{1}{2} - \frac{1}{4}|x - 3| & , 1 \le x \le 5 \\ 0 & , \text{otherwise} \end{cases}$$

- Is $f(x)$ a valid PDF of some random variable?

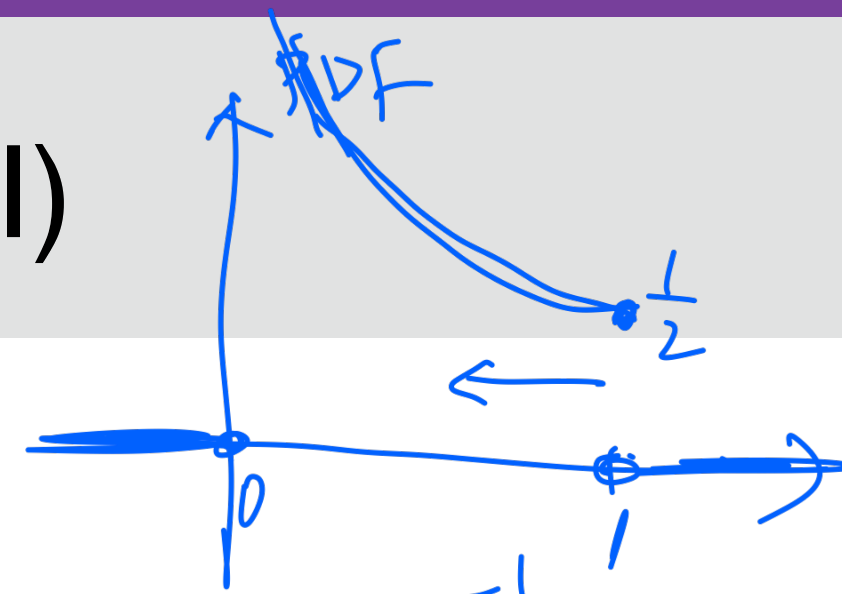
Yes !

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

② $f(x) \ge 0, \forall x$ ✓



Example: From PDF to CDF (II)



- Example: Consider the following PDF

$$f(x) = \begin{cases} \frac{1}{2\sqrt{x}} & ,0 < x \le 1 \\ 0 & \text{, otherwise} \end{cases}$$

$$\frac{1}{2}x^{-\frac{1}{2}}$$

- Is $f(x)$ a valid PDF of some random variable?

① $\int_{-\infty}^{+\infty} f(x)\,dx = 1 \Rightarrow \int_0^1 \frac{1}{2\sqrt{x}}\,dx \overset{?}{=} \frac{\cancel{\frac{1}{2}}\, x^{\frac{1}{2}}}{\cancel{\frac{1}{2}}}\Bigg|_0^1$

$= 1 - 0 = 1$

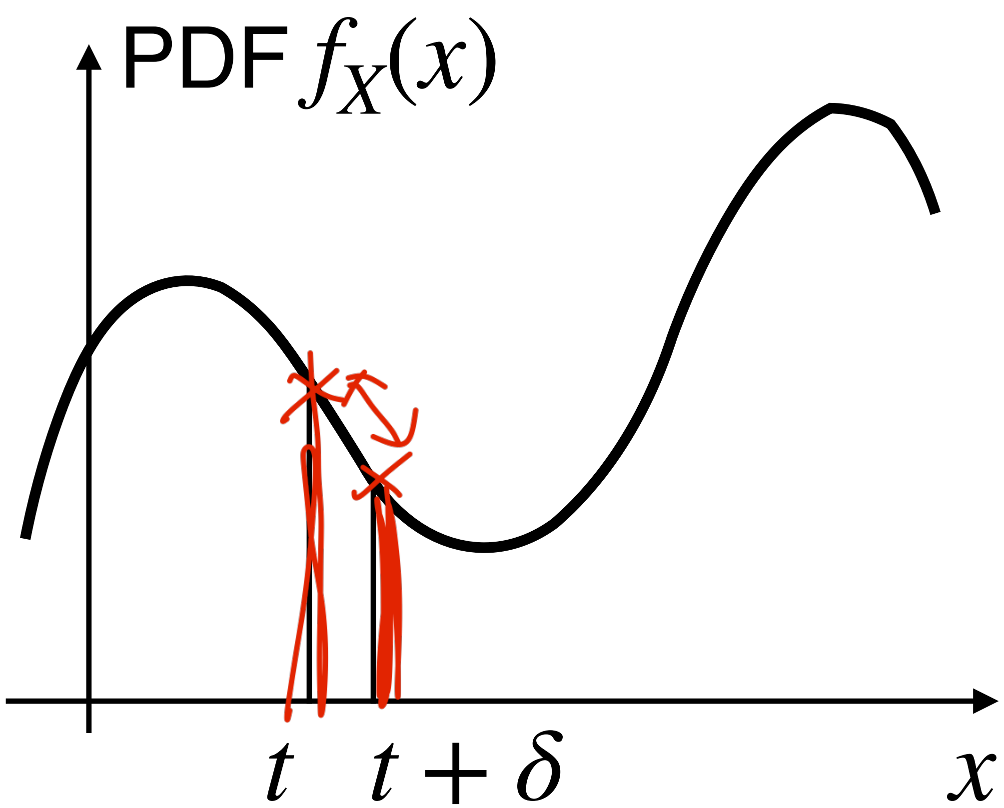
② $f(x) \ge 0, \forall x$ ✓

From CDF to PDF

$f(t) \approx \frac{F_X(t+\delta) - F_X(t)}{\delta}$

$\Rightarrow$ derivative!

- CDF: $F_X(t) = P(X \leq t) = \int_{-\infty}^{t} f_X(x)dx$



- Suppose PDF is continuous
- $F_X(t+\delta) - F_X(t) = ?$

$= P(X \leq t+\delta) - P(X \leq t)$

$= P(t < X \leq t+\delta)$

$= \int_{t}^{t+\delta} f(x)dx \approx f(t) \cdot \delta$

From CDF to PDF (Formally)

Derivative of CDF is PDF:
Let X be a random variable with a CDF $F_X(\cdot)$ and a PDF $f_X(\cdot)$. If $f_X(\cdot)$ is continuous at x_0, then

$$F'_X(x_0) = f_X(x_0)$$

$$\left.\frac{dF_X(x)}{dx}\right|_{x=x_0}$$

- Any similar results in calculus?

Fundamental thm of Calculus

From CDF to PDF: Why Continuity is Required?

- CDF: $F_X(t) = P(X \leq t) = \int_{-\infty}^{t} f_X(x)dx$

integrative

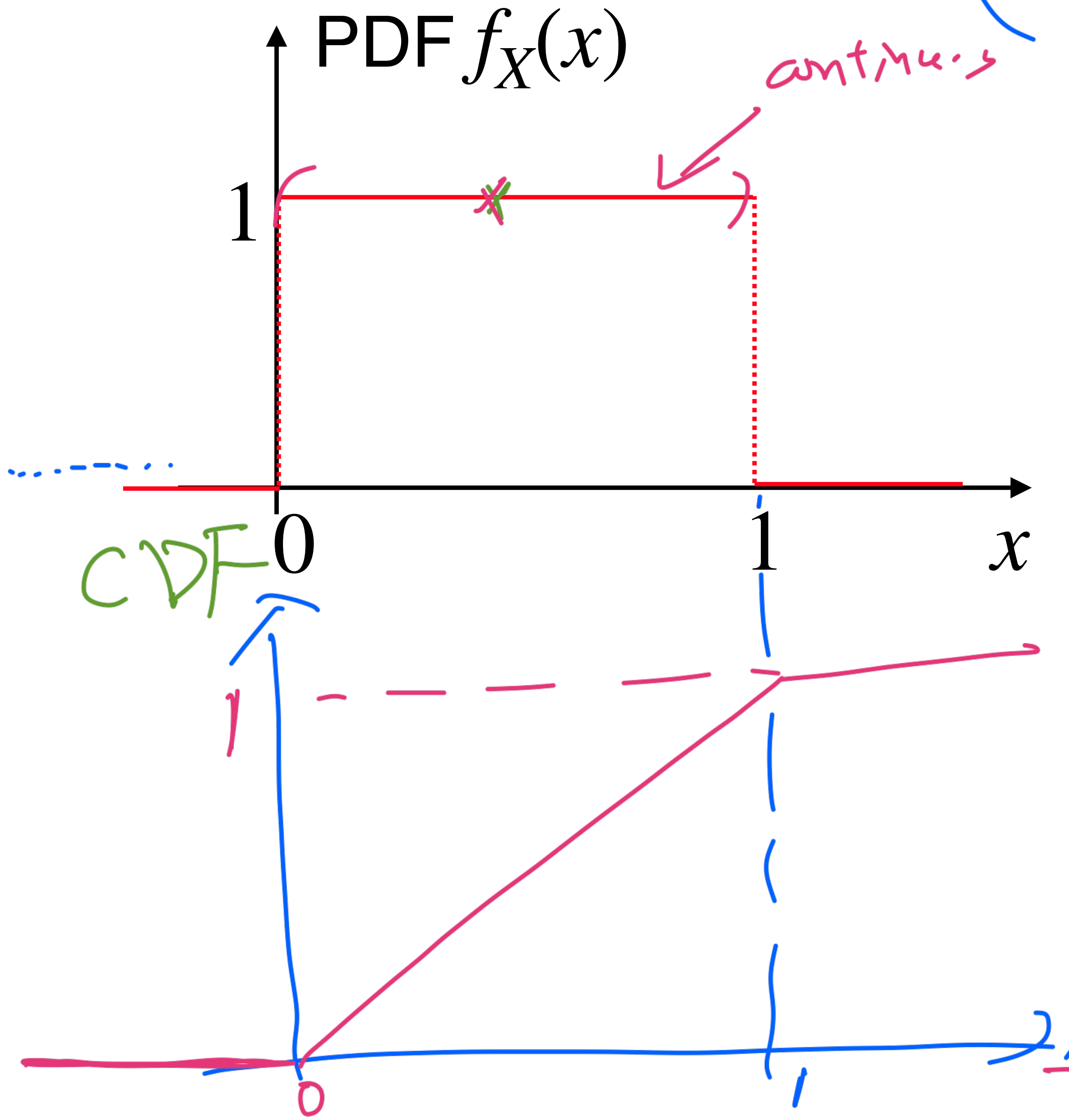
PDF ⟶ CDF

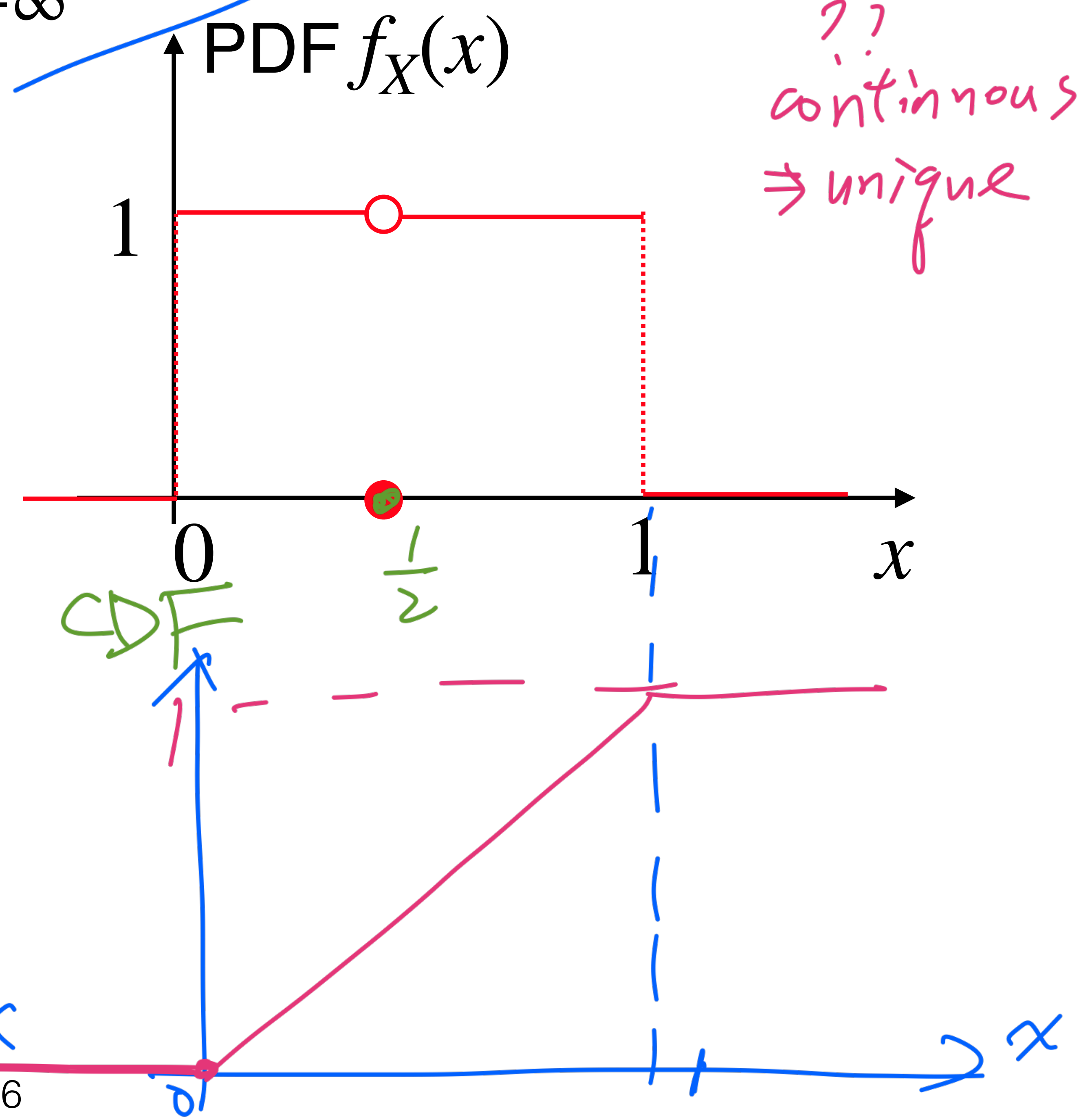
⟵

??

continuous

⇒ unique





3. Expected Value and Variance of Continuous Random Variables

Recall: Expected Value and Variance

Let X be a discrete random variable with a PMF $p_X(x)$ and the set of possible values S. Then, we have

1. $E[X] := \sum_{x \in S} x \cdot p_X(x)$ $\Longleftrightarrow$ $\int x \cdot f(x)\,dx$
2. $E[g(X)] := \sum_{x \in S} g(x) \cdot p_X(x)$ $\Longleftrightarrow$ $\int g(x) f(x)\,dx$ ← LOTUS
3. $\mathrm{Var}[X] := \sum_{x \in S} (X - E[X])^2 \cdot p_X(x)$ $\Longleftrightarrow$ $\int (x - E[x])^2 f(x)\,dx$

Expected Value and Variance of a Continuous R.V.

Let X be a continuous random variable with a PDF $f_X(x)$. Then, we have

1. $E[X] := \int_{-\infty}^{+\infty} x \cdot f_X(x) dx$
2. $E[g(X)] := \int_{-\infty}^{+\infty} g(x) \cdot f_X(x) dx$
3. Var$[(X)] := \int_{-\infty}^{+\infty} (x - E[X])^2 \cdot f_X(x) dx$

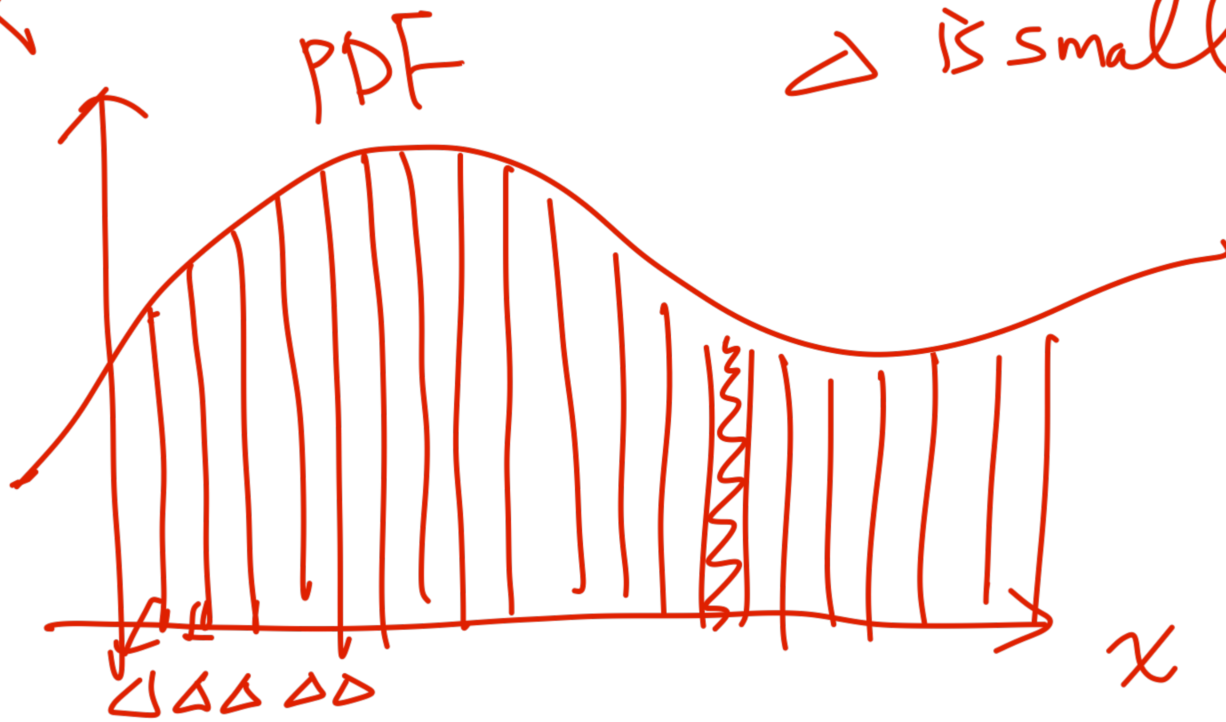
- How to intuitively interpret this?

Integration: Limiting Case of Summation

$$E[X] := \int_{-\infty}^{+\infty} x \cdot f_X(x)dx \approx \sum_{k=-\infty}^{\infty} (k\Delta) f_X(k\Delta)\Delta$$

X

Δ is small



$$E[\tilde{X}] = \sum_{k=-\infty}^{+\infty} (k\Delta)\cdot P(\tilde{X} = k\Delta)$$

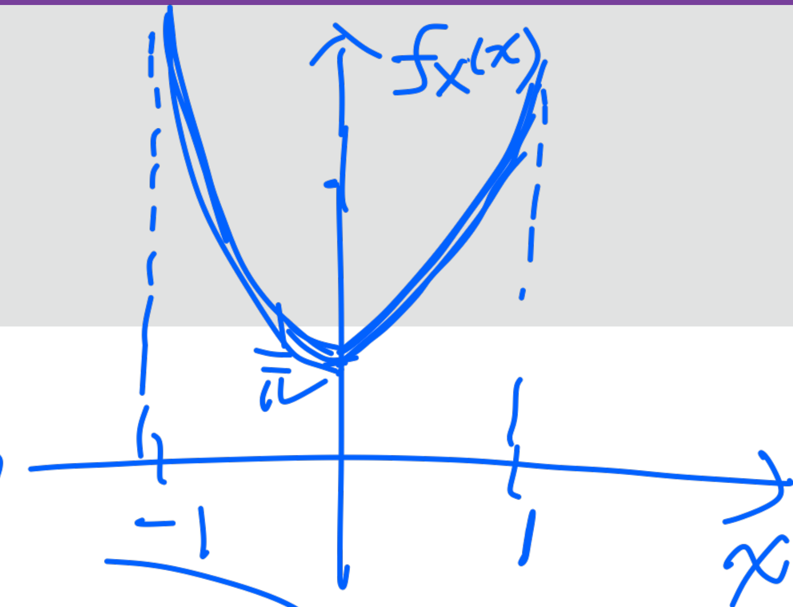
$$= \sum_{k=-\infty}^{+\infty} (k\Delta)\cdot f_X(k\Delta)\cdot \Delta$$

$\tilde{X}$ is a discrete r.v.

$$P(\tilde{X} = k\Delta) = \underbrace{F_X((k+1)\Delta) - F_X(k\Delta)}_{\approx\, f_X(k\Delta)\cdot\Delta}$$

Example

change of variables



- Example: Suppose the PDF of a r.v. X is

$$f_X(x) = \begin{cases} \dfrac{1}{\pi\sqrt{1-x^2}} & , -1 < x < 1 \\ 0 & , \text{otherwise} \end{cases}$$

$\to 0$

$f(0)$?

- What is $E[X]$?

$$E[X] = \int_{-1}^{1} x \cdot \frac{1}{\pi\sqrt{1-x^2}}\,dx = \int_0^1 x \cdot \frac{1}{\pi\sqrt{1-x^2}}\,dx + \int_{-1}^{0} x \cdot \frac{1}{\pi\sqrt{1-x^2}}\,dx$$

$$t = 1 - x^2 \Rightarrow \frac{dt}{dx} = -2x$$

$$\Rightarrow dx = -\frac{1}{2x}\,dt$$

$$= \int_0^1 x \cdot \frac{1}{\pi} \frac{1}{\sqrt{t}} \frac{1}{-2x}\,dt$$

Properties of Discrete R.V. Still Hold for Continuous R.V.?

1. $E[\alpha X + \beta] = \alpha \cdot E[X] + \beta$? ($\alpha, \beta \in \mathbb{R}$)

$$\int (\alpha x + \beta) \cdot f(x) dx \xrightarrow{\text{linearity}} = \int \alpha x \cdot f(x) dx + \int \beta \cdot f(x) dx$$

2. $E[g(X) + h(X)] = E[g(X)] + E[h(X)]$?
(g, h real-valued functions)

linearity of integration

Properties of Discrete R.V. Still Hold for Continuous R.V.?

3. $\text{Var}[X] := E[X^2] - (E[X])^2$?

4. $\text{Var}(X + c) = \text{Var}(X)$?

5. $\text{Var}(aX) = a^2 \cdot \text{Var}(X)$?

Recall: Expected Value of a Discrete Random Variable Using CDF

Expected Value (or Mean / Expectation):

Let X be a non-negative discrete random variable with

- the set of possible values $S = \{x_1, x_2, x_3 \cdots\}$
- CDF of X is $F_X(t)$

Denote $x_0 = 0$. The expected value of X is

$$E[X] = \sum_{i=1}^{\infty} (x_i - x_{i-1}) \cdot \left(1 - F_X(x_i^-)\right)$$

- How about continuous cases?

Expected Value of a Continuous Random Variable Using CDF

Expected Value via CDF:

Let X be a continuous random variable with CDF $F_X(t)$.

The expected value of X is

$$E[X] = \int_0^\infty \underbrace{\overbrace{(1 - F_X(t))}^{P(X>t)}}_{\text{CCDF}} dt - \int_0^\infty \overbrace{F_X(-t)}^{P(X\le -t)} dt$$

- What if X is a non-negative random variable?

$P(X \le -t) = 0$, for all $t \ge 0$

$\Rightarrow E[X] = \int_0^\infty (1 - F_X(t))dt$

- How to prove this?

Proof: Expected Value of a Continuous Random Variable Using CDF

$$E[X] = \int_0^\infty \underbrace{(1 - F_X(t))}_{\int dx} dt - \int_0^\infty F_X(-t)dt$$

Existence of Moments (Formally)

Existence of Moments:
Let X be a random variable. Then, the n-th moment of X (i.e. $E[X^n]$) is said to <u>exist</u> if $E[\,|X^n|\,] < \infty$

- Same definition for both discrete and continuous random variables

Example: Pareto Random Variable

- Example: Suppose $n \geq 1$ and the PDF of a r.v. X is

$$f_X(x) = \begin{cases} \frac{C}{x^{n+1}} & , x \geq C \\ 0 & \text{, otherwise} \end{cases}$$

- What is the value of C?
- For what values of m does $E[X^{m+1}]$ exist?

Next Lecture

1. Special continuous random variables
    - Uniform
    - Normal

1-Minute Summary

1. Expected Value and Variance of Special Discrete Random Variables

- Bernoulli / Binomial / Poisson

2. Probability Density Function (PDF)

- Use PDF to calculate probabilities
- CDF <-> PDF

3. Expected Value and Variance of Continuous Random Variables

- Extension from discrete to continuous cases