1179: Probability
Lecture 11 — Moments and Continuous Random Variables

Ping-Chun Hsieh (謝秉均)

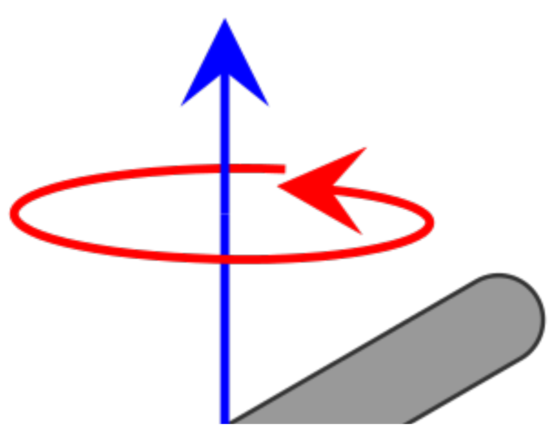
October 20, 2021

Announcements

- HW2 has been posted on E3 (Due: 11/1, 9pm)

Why the Word “Moment”?

- $E[X^n]$: n-th moment
- $E[(X-\mu_X)^n]$: n-th central moment
- An analogy: “Moment of inertia”



$$I = \int_Q \rho x^2 dV$$

ρ : density

H. A. David

A list of over 300
statistics is presen
print. Some of th
in preparing the l

KEY WORDS: F
ogy.

Model, components of variance	Mood, A. M. (1950, p. 342)
——, fixed effects	Scheffé, H. (1956, p. 252)
——, linear	Anderson, R. L., and Bancroft, T. A. (1952, p. 169)
——, mixed	Mood, A. M. (1950, p. 348)
——, random effects	Scheffé, H. (1956, p. 252)
Moment	Pearson, K. (1893, p. 615)
——, factorial	Steffensen, J. F. (1923, title)
—— generating function	Craig, C. C. (1936, p. 55)
Multiple correlation	Pearson, K. (1908, p. 59)
—— correlation coefficient	Pearson, K. (1914, p. 182)
—— comparisons	Duncan, D. B. (1951, p. 178)

Quick Review

- Alternative expression of $E[X]$ for non-negative discrete random variables?

- Law of the Unconscious Statistician?

- Linearity properties of expected values?

- Variance? Any alternative expression?

This Lecture

1. Variance and Moments

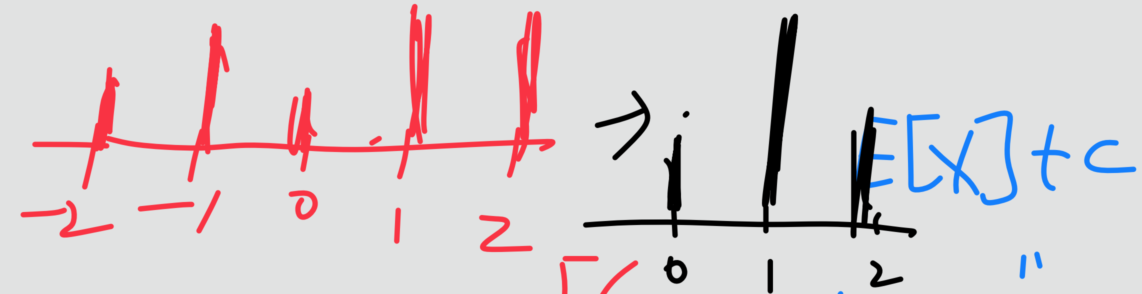
2. Expected Value and Variance of Special Discrete Random Variables

3. Continuous Random Variables

- Reading material: Chapter 4.4-4.5 and 6.1

Variance and Moments

Properties of Variance



1. Var$(X + c) =$ Var(X)?

$$\mathrm{Var}[X+c] = E\left[\left((X+\cancel{c}) - \underbrace{E[X+\cancel{c}]}_{E[X]+c}\right)^2\right] = \mathrm{Var}[X]$$

2. Var$(aX) = a \cdot$ Var(X)?

$$\mathrm{Var}[aX] = E[\underbrace{(aX)^2}_{a^2X^2}] - (\underbrace{E[aX]}_{a\cdot E[X]})^2$$

$$= a^2 \cdot E[X^2] - a^2 \cdot (E[X])^2$$

$$= a^2 \cdot \mathrm{Var}[X]$$

3. Var$(\,|X|\,) =$ Var(X)?

$$X = \begin{cases} 1, & \text{w.p. } \frac{1}{2} \\ -1, & \text{w.p. } \frac{1}{2} \end{cases}$$

$$|X| = 1, \text{ w.p. } 1$$

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

$$= 1 - 0$$

$$\mathrm{Var}[|X|] = 0$$

4. $E(X^2) \geq (E(X))^2$?

$$\mathrm{Var}[X] = \underbrace{E[X^2] - (E[X])^2}_{(*)} \geq 0$$

Variance: An Alternative Explanation

"minimum achievable expected quadratic penalty"

- Example: Suppose we are given a random variable X
  - We need to output a prediction of X (denoted by z)
  - Penalty of prediction is $(X-z)^2$ quadratic
  - What is the minimum expected penalty?

$$g(z) \triangleq \text{Expected penalty} = E[(X-z)^2]$$

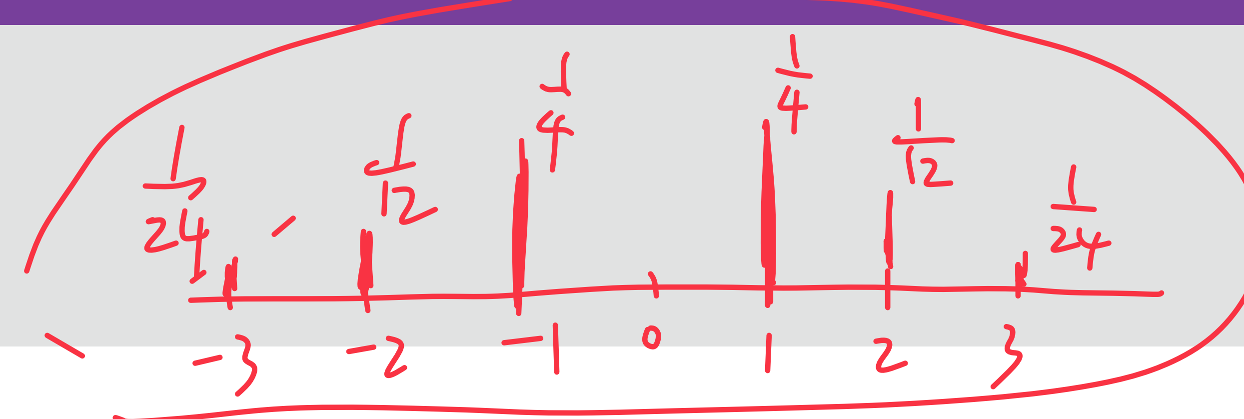
$$= E[X^2 - 2z\cdot X + z^2]$$

$$= \underbrace{E[X^2]}_{} - \underbrace{E[2z\cdot X]}_{= 2z\cdot E[X]} + \underbrace{E[z^2]}_{= z^2}$$

completing the square (配方法)

$$= \underbrace{(z - E[X])^2}_{\geq 0} + \boxed{\left(E[X^2] - (E[X])^2\right)} = \mathrm{Var}[X]$$

Existence of Moments



- Example: 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?

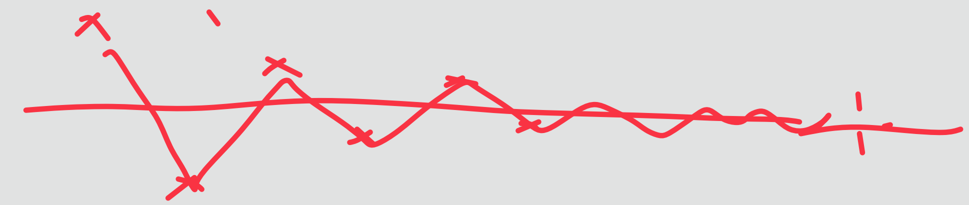
$$E[X] \neq \boxed{\sum_{\text{all } x} x \cdot p_X(x)} \overset{?}{=} 0$$

$E[X]$ does not exist

$$E[|X|] = 2 \cdot \sum_{k=1}^{\infty} k \cdot \frac{1}{2k(k+1)} = \sum_{k=1}^{\infty} \frac{1}{k+1} = \infty$$

Rearrangement of Series

a_n



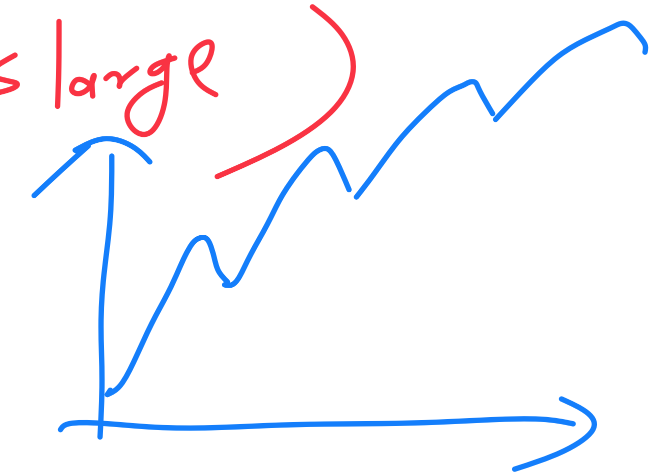
$\sum a_n$

- Example: Consider a series $\{a_n\}$: $1, -1, \frac{1}{2}, -\frac{1}{2}, \frac{1}{3}, -\frac{1}{3}, \cdots$

  $\int \frac{1}{x}dx$

  - What is $\sum_{n=1}^{\infty} a_n$? 0

(Suppose M is large)



$\ln M$

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

$$1, \frac{1}{2}, \frac{1}{3}, \cdots, \frac{1}{M}, -1, \frac{1}{M+1}, \cdots, \frac{1}{2M}, -\frac{1}{2}, \cdots,$$

KM "+" "−"

$K\cdot(M+1)$ terms

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

$\sum_{n=1}^{N} b_n$

Positive: The sum $\approx \ln KM$

Negative: The sum $\approx -\ln K$

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

1.87

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 $E[\,|X^n|\,] < \infty$

- When do we care about the existence of moments?

When are Higher Moments Useful?

Berry-Esseen Theorem:

Let $X_1, X_2, \cdots, X_n$ be i.i.d. random variables with $E[X_1] = 0$, $E[X_1^2] = \sigma^2$ and $E[|X_1|^3] < \infty$. Define $Y = (X_1 + X_2 + \cdots + X_n)/n$. Then, we have

$$|F_Y(t) - \Phi(t)| \leq \frac{C\rho}{\sigma^3\sqrt{n}}$$

- Usually higher moments are used as technical conditions
  - Hence, we usually care about whether $E(|X|^n) < \infty$

Properties of Moments

(Assume discrete random variables)

n : positive integer

- If $E(|X|^{n+1}) < \infty$, then $E(|X|^n) < \infty$? ✓

$$\underbrace{E[|X|^{n+1}]}_{\parallel} < \infty$$

$$\underbrace{\sum_{x:|x|\geq 1} |x|^{n+1} \cdot p_X(x)}_{\text{finite}} + \underbrace{\boxed{\sum_{x:|x|<1} |x|^{n+1} \cdot p_X(x)}}_{<1}$$

$$E[|X|^n] = \boxed{\sum_{x:|x|\geq 1} |x|^n \cdot p_X(x)} + \underbrace{\boxed{\sum_{x:|x|<1} |x|^n \cdot p_X(x)}}_{<1}$$

$$< \underbrace{\sum_{x:|x|\geq 1} |x|^{n+1} \cdot p_X(x)}_{\text{finite}}$$

St. Petersburg Paradox

- Example: We are asked to pay 10000 dollars to play a game.
  - We can keep flipping a fair coin until a head is observed.
  - If the 1st head occurs at n-th toss, then we get a prize of 2^n dollars and the game is over.
  - Shall we play this game?

Conditional Expectation

Conditional Expectation

- Example: Roll a fair 6-sided die once
  - Define $X =$ the number that we observe
  - Given that $X \geq 4$, what is the expected value of X?

(1, 2, 3, 4, 5, 6

Conditional Expectation:

Let X be a discrete random variable with the set of possible values $S = \{x_1, x_2, x_3 \cdots\}$. Let A be an event. The expected value of X conditioned on A

$$E[X|A] := \sum_{x \in S} x \cdot P(X = x|A)$$

Example: Taiwan Receipt Lottery

Suppose

- Example: Suppose we have a receipt at hand
  - Define X = the prize we get
  - What is $E[X]$?
  - Given that the last digit is 7, what is the expected value of X?

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

109年 7-8月 統一發票開獎		
特別獎	13362795	與左欄號碼相同者獎金1000萬元
特獎	27580166	與左欄號碼相同者獎金200萬元
頭獎	53227282 35082085 37175928	頭獎 與頭獎號碼完全相同者獎金20萬元 二獎 與頭獎末7碼相同者各得獎金4萬元 三獎 與頭獎末6碼相同者各得獎金1萬元 四獎 與頭獎末5碼相同者各得獎金4000元 五獎 與頭獎末4碼相同者各得獎金1000元 六獎 與頭獎末3碼相同者各得獎金200元
增開六獎	987 614	末3碼與增開六獎號碼相同者各得獎金200元
正確資訊請以財政部提供為準　中央社祝您幸運中獎		

$\frac{1}{10^8}$

$\frac{1}{10^8}$

$\frac{3}{10^8}$

00000000 ~ 99999999

10^8

$$E[X \mid \text{the last digit is } 7]$$

$$= \sum_{\text{all } x} x \cdot P(X = x \mid \text{the last digit is } 7)$$

$$= 200 \times \frac{1}{100} = 2 \text{ dollars}$$

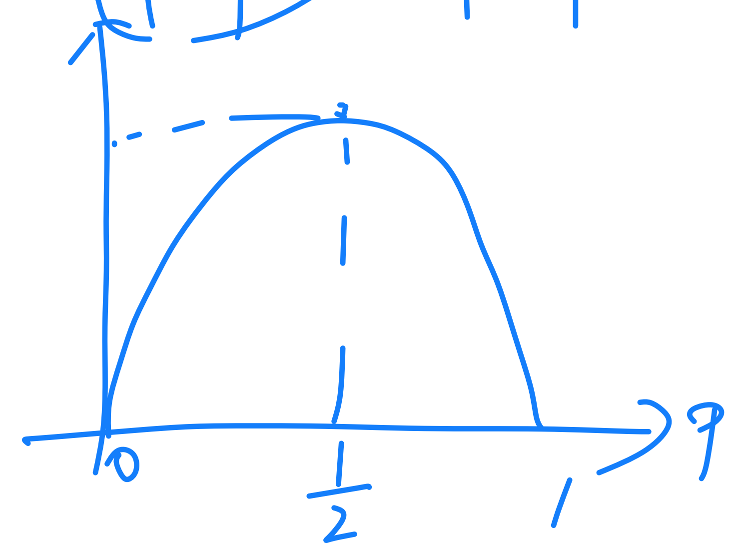
Expected Value and Variance of Special Discrete Random Variables

1. Bernoulli Random Variables

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

$$X = \begin{cases} 1, & \text{w.p. } p \\ 0, & \text{w.p. } 1-p \end{cases}$$

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



$$E[X] = 1 \cdot p + 0 \cdot (1-p) = p$$

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

$$= \underbrace{\left(1^2 \cdot p + 0^2 \cdot (1-p)\right)}_{= p} - p^2 = p \cdot (1-p)$$

If $p = 0$: $\text{Var}[X] = 0$

$p = 1$: $\text{Var}[X] = 0$

2. Binomial Random Variables

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

PMF: $p(X=k) = \begin{cases} C_k^n \cdot p^k \cdot (1-p)^{n-k}, & k=0,1,2,\cdots,n \\ , & \text{otherwise} \end{cases}$

$$E[X] = \sum_{k=0}^{n} k \cdot p(X=k)$$

$$= \sum_{k=0}^{n} k \cdot C_k^n \cdot p^k \cdot (1-p)^{n-k} = \sum_{k=0}^{n} \cancel{k} \cdot \frac{n!}{\cancel{k!}(n-k)!} p^k \cdot (1-p)^{n-k}$$

$n \cdot (n-1)!$

$(k-1)!$

$$n \cdot C_{k-1}^{n-1}$$

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

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

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

2. View X as a sum of independent random variables

3. Moment generating functions

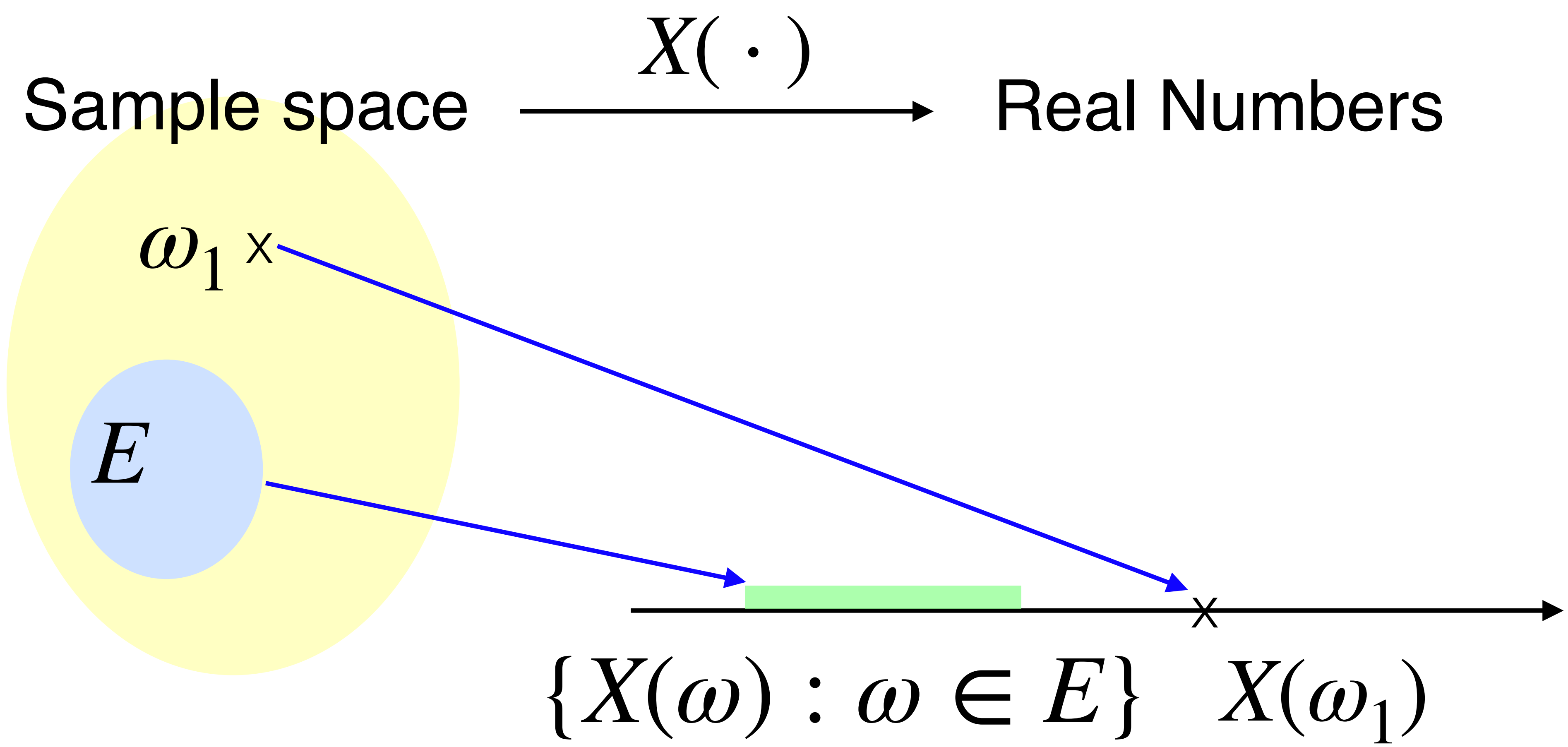
3. Poisson Random Variables

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

3. Continuous Random Variables and Probability Density Functions

Continuous Random Variables

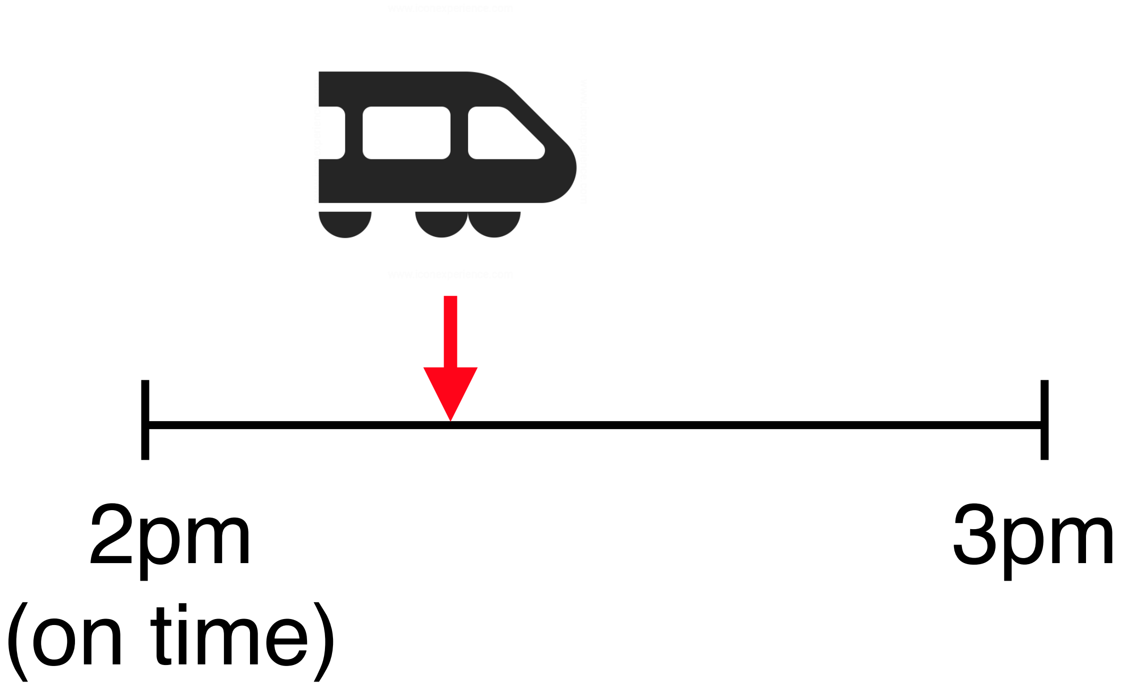
- Continuous random variable: A random variable that takes values over a continuous range



- 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?

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

Density / Concentration

- Example: B-52 Cocktail

orange liqueur (40%): 10 ml

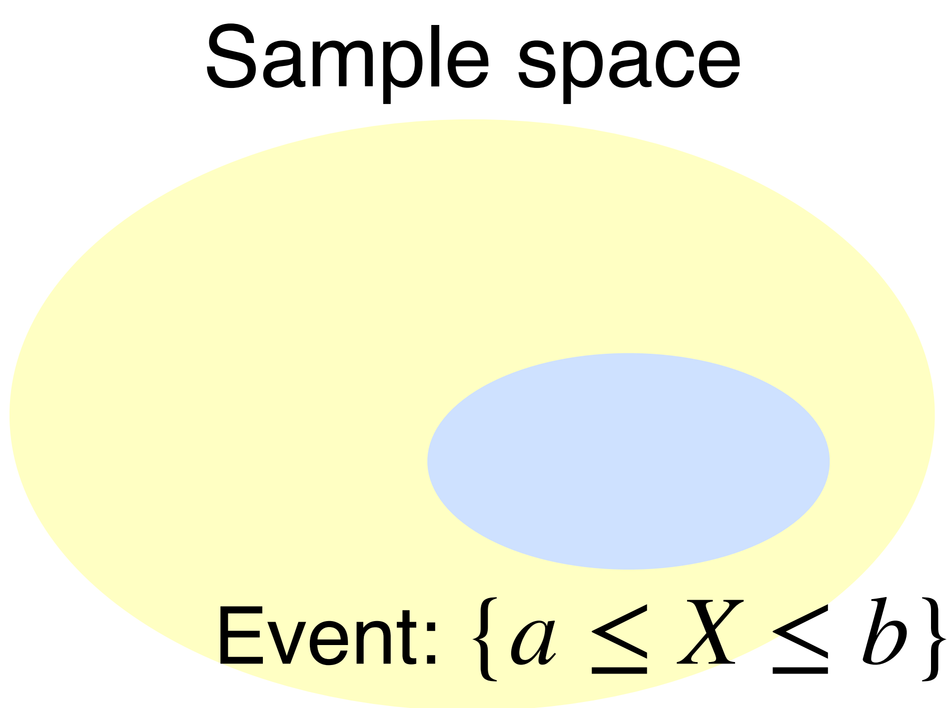
milk wine (17%): 10 ml

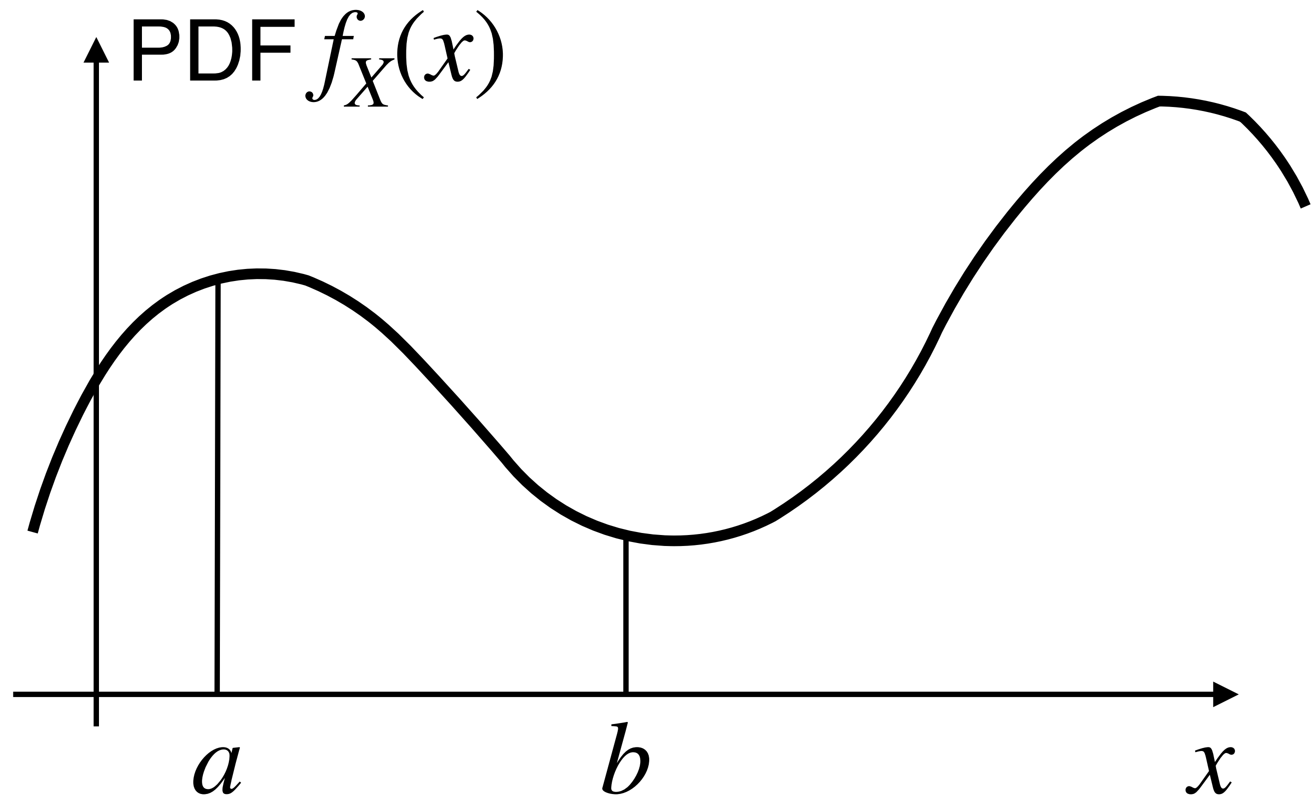
coffee liqueur (23%): 10 ml

- How much alcohol in total (in ml)?

- How much alcohol in the green cross section?

Probability Density Function (PDF)





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

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

2. $P(X \leq t)$ =

3. $P(a \leq X \leq b)$ =

4. $P(a \leq X < b)$ =

5. $P(a < X < b)$ =

How to Check if a PDF is Valid?

- Recall: 3 Axioms of Probability

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

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

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

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

1-Minute Summary

1. Variance and Moments

- Definition / alternative explanation using penalty / properties
- Existence of moments

2. Expected Value and Variance of Special Discrete Random Variables

- Bernoulli / Binomial / Poisson

3. Continuous Random Variables

- Probability density function