1179: Probability
Lecture 9 — Special Discrete Random Variables, Expected Value, and Variance

Ping-Chun Hsieh (謝秉均)

October 13, 2021

Sum of 3 Cubes

- Question: Can we find integers x, y, z such that

$$x^3 + y^3 + z^3 = N, \quad 0 \le N \le 100$$

- $\underline{N = 0 :}$ $\quad (a)^3 + (-a)^3 + 0^3 = 0$
- $\underline{N = 29 :}$ $(3)^3 + (1)^3 + (1)^3 = 29$
- $\underline{N = 9m + 4 \text{ or } 9m + 5 :}$ Not possible
- Unsolved case (since 1954): $N = 42$

Sum of 3 Cubes for $N = 42$

Sum of three cubes for 42 finally solved – using real life planetary computer

Press release issued: 6 September 2019 September 2019

Hot on the heels of the ground-breaking 'Sum-Of-Three-Cubes' solution for the number 33, a team led by the University of Bristol and Massachusetts Institute of Technology (MIT) has solved the final piece of the famous 65-year-old maths puzzle with an answer for the most elusive number of all - 42.

The original problem, set in 1954 at the University of Cambridge, looked for Solutions of the Diophantine Equation x^3+y^3+z^3=k, with k being all the numbers from one to 100.

Beyond the easily found small solutions, the problem soon became intractable as the more interesting answers – if indeed they existed – could not possibly be calculated, so vast were the numbers required.

But slowly, over many years, each value of k was eventually solved for (or proved unsolvable), thanks to sophisticated techniques and modern computers - except the last two, the most difficult of all; 33 and 42.

Professor Andrew Booker
Image credit: University of Bristol

Share this article

Andrew Booker
(University of Bristol)

- Answer:
  - $x = -80538738812075974$
  - $y = 80435758145817515$
  - $z = 12602123297335631$

charityengine

1.3 million hours of computation

- A pure math project or a CS-related project?

Quick Review

- PMF of a Poisson random variable?
- Any nice properties of Poisson random variables?

$$P(X=k) = \begin{cases} \frac{e^{-\lambda T} \cdot (\lambda T)^k}{k!}, & k=0,1,2,\cdots \\ 0, & \text{else} \end{cases}$$

- Binomial
- Sum of independent Poisson is still Poisson

$X_1 \sim \text{Poisson}(\lambda_1, T)$

$X_2 \sim \text{Poisson}(\lambda_2, T)$

$X_1 + X_2 \sim \text{Poisson}(\lambda_1 + \lambda_2, T)$

Recall: An Interview Question

$\frac{e^{-\lambda T}\cdot(\lambda T)^k}{k!}$

$P(X\geq 1)=0.95$

$\boxed{P(X=0)=0.05}$

$= e^{-\lambda T}$

- Example: Suppose we stand at the Fude temple.
  - The probability that we see at least 1 car passing through the temple in 30 minutes is 0.95.
  - What is P(we see at least 1 car in 10 mins) under a Poisson model?

$P(Y\geq 1) = 1-P(Y=0)$

X = # of cars passing through the temple in 30min

Y = # of cars " " in 10min

X, Y are independent

$X \sim \text{Poisson}(\lambda, T = 30\text{min})$

$Y \sim \text{Poisson}(\lambda, T' = 10\text{min})$

$$P(Y=0) = \frac{e^{-\lambda T'}\cdot(\lambda T')^0}{0!} = e^{-\lambda T'} = (e^{-\lambda T})^{\frac{1}{3}} = (0.05)^{\frac{1}{3}}$$

Therefore, we have $P(Y\geq 1) = 1-(0.05)^{\frac{1}{3}}$ □

A, B events

A, B $\ni P(A \cap B) = P(A) \cdot P(B)$
independence

X, Y random variables.

X, Y are independent if $P(A_X \cap A_Y) = P(A_X) \cdot P(A_Y)$

This Lecture

1. Special Discrete Random Variables

2. Expected Value

3. Variance and Moments

- Reading material: Chapter 4.4-4.5 and 5.3

Special Discrete Random Variables

4. Geometric Random Variables

- Example: Play with a claw machine, and each trial is successful with probability 0.7. What is $P(\text{get 1st toy at 10-th trial})$?

- Example: Po-Jung Wang makes a hit with probability 0.28 at each at-bat. What is $P(\text{he makes his 1st hit at 5-th at-bat})$?

- What are the common features?
  - Repetitions of the same Bernoulli experiment
  - Want: how many trials needed until the 1st success?

4. PMF of Geometric Random Variables

- Example: Play with a claw machine, and each trial is successful with probability 0.7. All trials are independent.
  - X = the number of trials until we get the first toy
  - What is the PMF of X?

$P(X=1) = 0.7^1$

$P(X=2) = (0.3)^1 \times (0.7)^1$

$P(X=3) = (0.3)^2 \times (0.7)^1$

$\vdots$

$P(X=k) = (0.3)^{k-1} \times (0.7)^1$

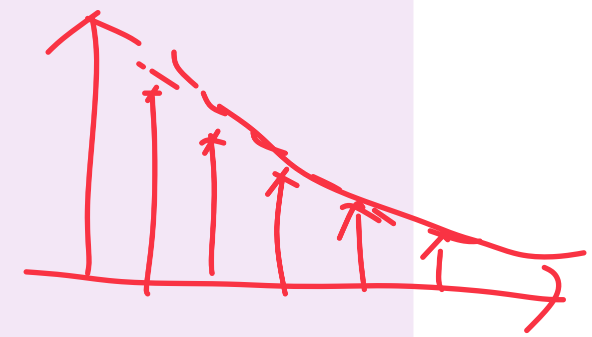
$k \in \mathbb{N}$

geometric sequence

4. Geometric Random Variables (Formally)

Geometric Random Variables: A random variable X is Geometric with parameters p if its PMF is given by

$$P(X=k) = \begin{cases} (1-p)^{k-1}p, & k=1,2,3,\cdots \\ 0, & \text{otherwise} \end{cases}$$



- Do we have $\sum_{k=1}^{\infty} P(X=k) = 1$? $\Longrightarrow \sum_{k=1}^{\infty} (1-p)^{k-1} \cdot p = 1$

$$= \frac{p}{1-(1-p)}$$

首項

公比

CDF of Geometric Random Variables

$$P(X = k) = (1-p)^{k-1}p, \; k = 1,2,3,\cdots \quad t$$

- CDF: $F_X(t) = P(X \le t)$, $t \in \mathbb{R}$

1. $t < 1$: $F_X(t) = 0$

2. $t \in \mathbb{N}$ ($= t \in \{1,2,3,\cdots\}$): $F_X(t) = \sum_{k=1}^{t} (1-p)^{k-1} \cdot p$

3. $t \geq 1$ but $t \notin \mathbb{N}$: $F_X(t) = \sum_{k=1}^{\lfloor t \rfloor} (1-p)^{k-1} \cdot p$

$t = 1.87$

Geometric r.v.: Memoryless Property

$B=\{X>m\}$

$A=\{X=n+m\}$

$P(A\cap B)=P(A)$

- Example: Suppose $X \sim \text{Geometric}(p)$, $p \in (0,1)$
  - $(1-p)^{n-1}\cdot p$
  - What is $P(X=n+m \mid X>m)$? $(n,m \in \mathbb{N})$
  - What is $P(X>n+m \mid X>m)$? $(n,m \in \mathbb{N})$

$$P(X=n+m \mid X>m) = \frac{P(X>m \text{ and } X=n+m)}{P(X>m)} = \frac{P(X=n+m)}{P(X>m)} = \frac{(1-p)^{n+m-1}\cdot p}{(1-p)^m} = (1-p)^{n-1}\cdot p$$

- $P(X=n+m) = (1-p)^{n+m-1}\cdot p$
- $P(X>m) = (1-p)^m$

$$= 1-P(X\le m) = 1-\sum_{k=1}^{m}(1-p)^{m-1}\cdot p = 1-\frac{\cancel{p}\cdot(1-(1-p)^m)}{(1-(1-p))\times \cancel{p}} = (1-p)^m$$

$$P(\underline{X>n+m} \mid \underline{X>m}) = (1-p)^n = P(X>n)$$

$$\frac{P(X>n+m \text{ and } X>m)}{P(X>m)} = \frac{P(X>n+m)}{P(X>m)} = (1-p)^n$$

$P(X>n+m) \to (1-p)^{n+m}$

$P(X>m) \to (1-p)^m$

Memoryless Property:

$$P(X>n+m \mid X>m) = P(X>n)$$

$$P(X=n+m \mid X>m) = P(X=n)$$

5. Discrete Uniform Random Variables

- Example: Roll a 4-sided die, and the numbers 1, 2, 3, 4 are equally likely to occur

- Example: The correct answer to an exam question: A, B, C, D are equally likely

- What are the common features?
  - 1 experiment trial (no repetition) with n equally-likely outcomes
  - Want: Whether a specific outcome occurs

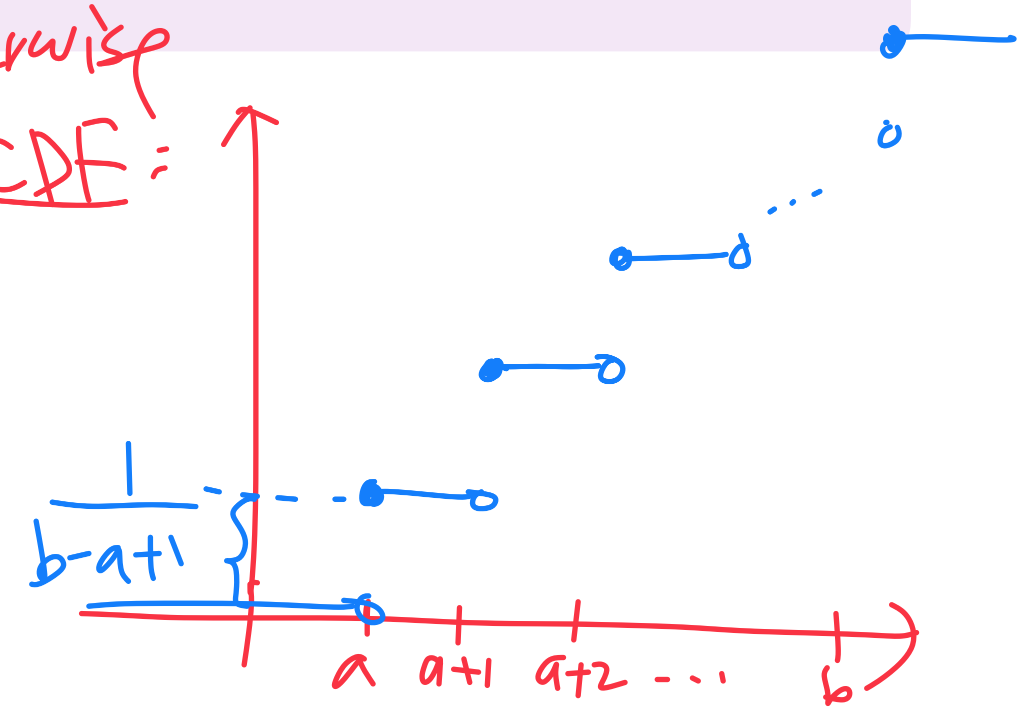
5. Discrete Uniform Random Variables (Formally)

Discrete Uniform Random Variables: A random variable X is discrete uniform with parameters (a, b) ($a, b \in \mathbb{Z}$ with $a \leq b$), if its PMF is given by

$$P(X = k) = \begin{cases} \dfrac{1}{b-a+1}, & k = a, a+1, \cdots, b \\ 0 & \text{otherwise} \end{cases}$$

PMF:

CDF:



Expected Value

Motivation: Guidelines for Decision Making?

- Suppose we are given 2 options:

Option A

$500 \times 10^6 \times \frac{1}{1000}$

$-50 \times 10^3 \times \frac{999}{1000}$

Win 500 millions with prob. 0.1%

Lose 50k dollars, otherwise

Risk-seeking

Option B

Win 50k dollars, with prob. 1

Risk-averse

- Which option will you choose?
- Could you come up with a creative way to get a reward higher than B but with a lower risk than A?

Why Expected Value?

- Example: Imagine you are an investor of Shinemood
  - $X =$ # of waffles sold by Shinemood today
  - Suppose $X \sim$ Poisson($\lambda = 150$, $T =$ 1 day)
- Question 1: How many waffles are expected to be sold today?

Expected Value

PMF of X

- Question 2: In the coming year, how many waffles will be sold per day on average?

Empirical average

$$\frac{X_1 + X_2 + \cdots + X_{365}}{365}$$

- Q1 and Q2 are closely related: Law of Large Numbers

Example: Expected Value and Poisson

- Example: X = # of waffles sold by Shinemood today
  - Suppose $X \sim$ Poisson($\lambda = 150$, $T = 1$ day)
  - PMF of X? How to define the expected value of X?

$$P(X=k) = \begin{cases} \frac{e^{-\lambda T} \cdot (\lambda T)^k}{k!}, & k = 0, 1, 2, \cdots \\ 0, & \text{otherwise} \end{cases}$$

$$E[X] = \sum_{k=0}^{\infty} \underline{k} \cdot P(X=k) = \sum_{k=1}^{\infty} k \cdot P(X=k) = \sum_{k=1}^{\infty} \cancel{k} \cdot \frac{e^{-\lambda T} \cdot (\lambda T)^k}{\cancel{k!}\ (k-1)!}$$

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

$$= \sum_{k=0}^{\infty} \frac{e^{-\lambda T} \cdot (\lambda T)^k}{k!} = 1$$

$$= \lambda T$$

Expected Value of a Discrete R.V. (Formally)

Expected Value (or Mean / Expectation):

Let X be a discrete random variable with

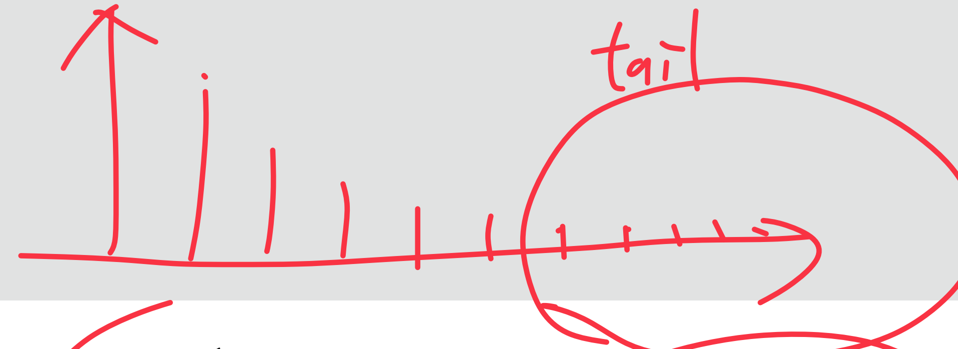
- the set of possible values S
- PMF of X is $p_X(x)$

The expected value of X is defined as

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

- Sometimes we use the notation: $\mu_X \equiv E[X]$

Example: Expected Value



- Example: Suppose X has a PMF: $p_X(n) = \dfrac{1}{n(n+1)}, \; n = 1,2,\cdots$
  - What is $E[X]$?

$$E[X] = \sum_{n=1}^{\infty} n \cdot p_X(n) = \sum_{n=1}^{\infty} \cancel{n} \cdot \frac{1}{\cancel{n}\cdot(n+1)} = \sum_{n=1}^{\infty} \frac{1}{n+1} = \infty$$

① $\frac{1}{2} + \underbrace{\frac{1}{3} + \frac{1}{4}}_{\geq \frac{1}{2}} + \underbrace{\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}}_{\geq \frac{1}{2}} + \cdots$

② $\sum_{n=1}^{\infty} \frac{1}{n+1} \geq \int_1^{\infty} \frac{1}{x}\,dx$

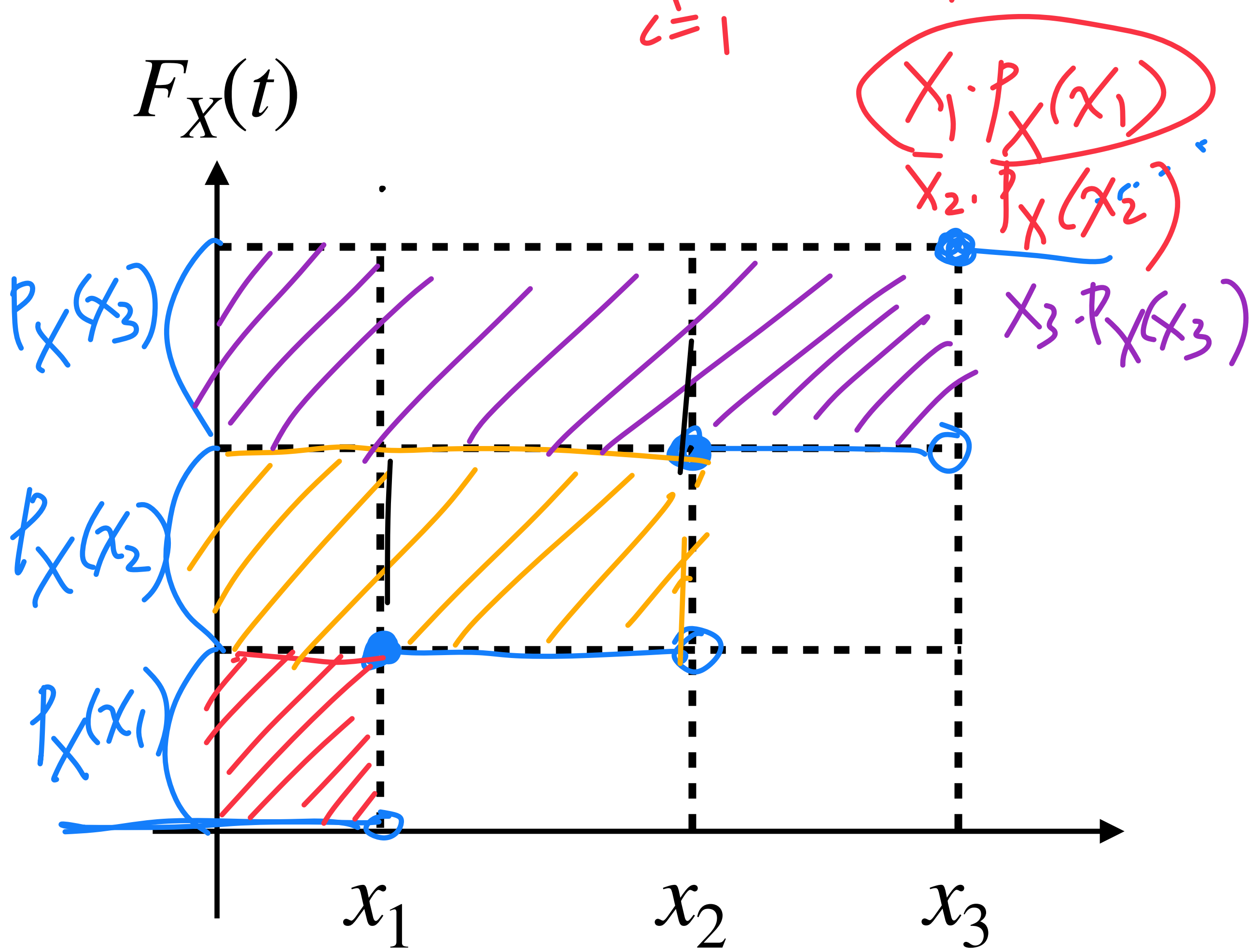
Example: St. Petersburg Paradox

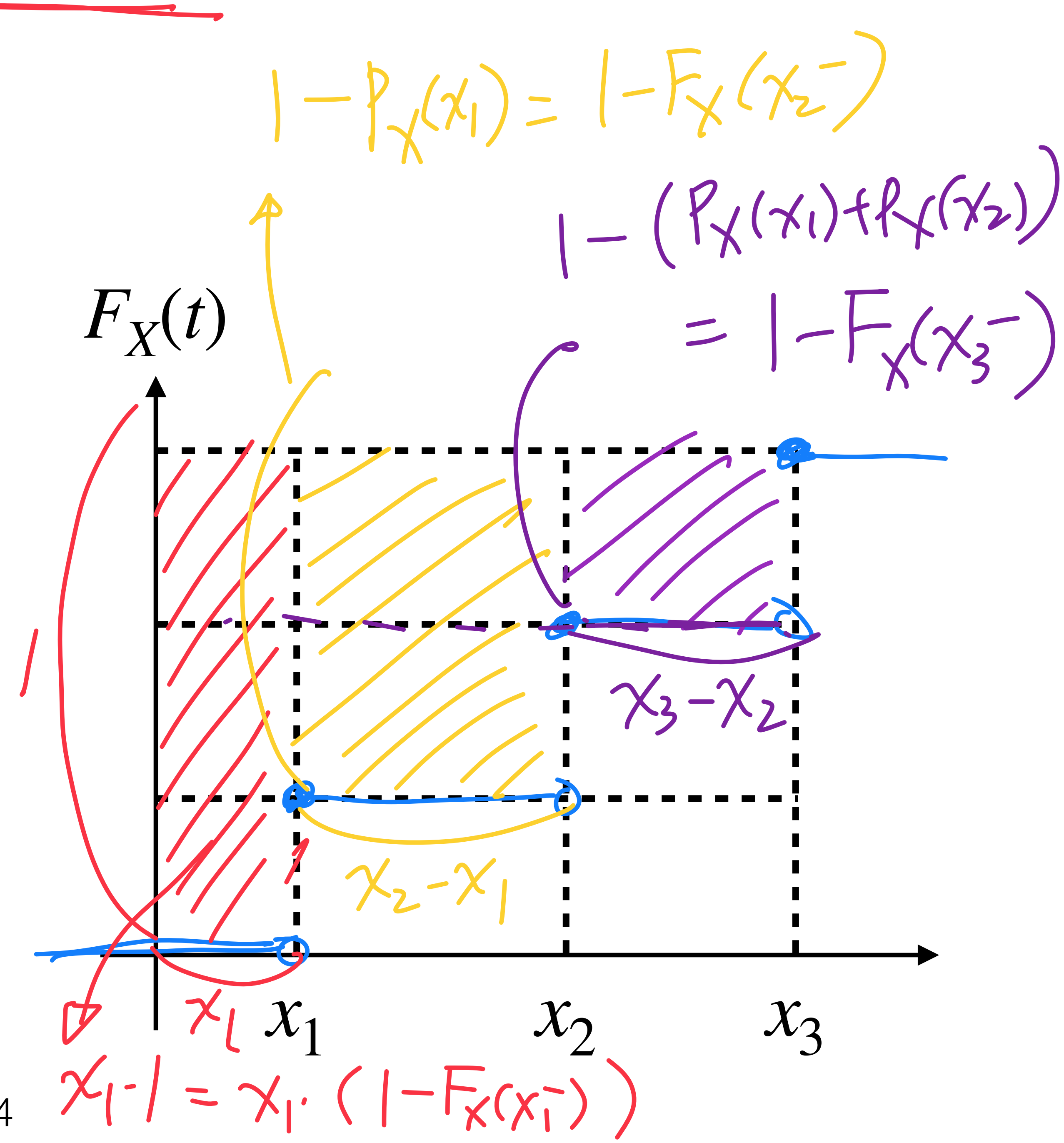
- Example: We are asked to pay 10000 dollars to play a game.
  - We can keep flipping a fair coin until a head is observed.
  - If the 1st head occurs at n-th toss, then we get a prize of 2^n dollars and the game is over.
  - Shall we play this game?

Visualize the Expected Value Using CDF

- Suppose X is a non-negative discrete random variable with
  - The set of possible values $\{x_1, x_2, x_3, \cdots\}$ (assume $x_i < x_{i+1}$)
  - Denote $x_0 = 0$

$$E[X] = \sum_{i=1}^{\infty} x_i \cdot P_X(x_i)$$





Expected Value of a Discrete Random Variable: An Alternative Expression

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

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

- What if $S = \{1,2,3 \cdots\}$?
- How about continuous cases?

$$E[X] = \int_0^{\infty} P(X > t)\,dt$$

Example: Using the Alternative Expression

- Example: Suppose X is a discrete random variable
  - For X, the set of possible values $A = \{2,4,6,8\cdots\}$
  - The CDF of X is $F_X(t) = 1 - \dfrac{1}{t^2}, t \in A$
  - What is $E[X]$?

A Property of Expected Value

Theorem (Expectation of a Function of r.v.):

1. Let X be a discrete random variable with
   - the set of possible values S
   - PMF of X is $p_X(x)$
2. Let $g(\cdot)$ be a real-valued function

The expectation of $g(X)$ is

$$E[g(X)] = \sum_{x \in S} g(x) \cdot p_X(x)$$

- Is this intuitive? Do we need a proof?
- Also called Law of the unconscious statistician

Proof of Law of the Unconscious Statistician

$$E[g(X)] := \sum_{x \in S} g(x) \cdot p_X(x)$$

Linearity of Expected Values (I)

Linearity Property (I):
Let X be a discrete random variable and α, β be real numbers. Then, we have

$$E[\alpha X + \beta] = \alpha \cdot E[X] + \beta$$

- How to show this?

Linearity of Expected Values (II)

Linearity Property (II):

Let X be a discrete random variable and $g(\cdot), h(\cdot)$ be real numbers. Then, we have

$$E[g(X) + h(X)] = E[g(X)] + E[h(X)]$$

- How to show this?

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?

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

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

Example: Taiwan Receipt Lottery

- 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?

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

Variance and Moments

Moments and Others

$$E[g(X)] := \sum_{x \in S} g(x) \cdot p(x)$$

- Example: $g(X) = X^2$
- Example: $g(X) = X^n$
- Example: $g(X) = (X - \mu_X)^2$
- Example: $g(X) = (X - \mu_X)^n$
- Example: $g(X) = e^{tX}$

Variance

Variance (2nd central moment):
Let X be a discrete random variable with the set of possible values S and PMF $p_X(x)$. The variance of X is

$$\text{Var}[X] := E[(X - \mu_X)^2] = \sum_{x \in S} (x - \mu_X)^2 \cdot p_X(x)$$

- Sometimes we use the notation: $\sigma_X^2 \equiv \text{Var}[X]$
- Variance captures the <u>variability</u> of a random variable

Variance: An Alternative Explanation

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

Another Way for Calculating Variance

Theorem:
Let X be a random variable. Then, we have

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

- How to show this?

Properties of Variance and Moments (I)

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

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

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

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

5. Can $\text{Var}(X)$ be infinite?

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

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

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

1-Minute Summary

1. Special Discrete Random Variables

- Geometric / Uniform

2. Expected Value

- Definition / alternative expression

3. Variance and Moments

- Definition / alternative explanation using penalty / properties