香港中文大學（深圳）
The Chinese University of Hong Kong, Shenzhen

DDA2001: Introduction to Data Science

Lecture 5: Random Variable

Zicheng Wang

Recap: 1 - (Discrete) Random Variable and Probability Distributions

Example 1

- Toss a coin three times and count the number of heads.
- The sample space is

$$S = \{(t,t,t),(t,t,h),(t,h,t),(h,t,t),(t,h,h),(h,t,h),(h,h,t),(h,h,h)\}$$

- Lex X count the number of heads. Thus if $s = (t,t,h)$ occurs then $X(s) = 1$.
- X is called a random variable as it takes a numerical value that depends on the outcome of an experiment.

Example 1

- Toss a coin three times and count the number of heads.
- The sample space is

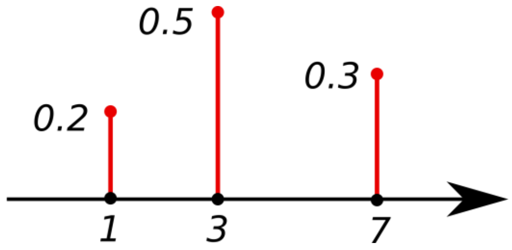
$$S = \{(t,t,t),(t,t,h),(t,h,t),(h,t,t),(t,h,h),(h,t,h),(h,h,t),(h,h,h)\}$$

- Lex X count the number of heads. Thus if $s = (t, t, h)$ occurs then $X(s) = 1$.
- X is called a random variable as it takes a numerical value that depends on the outcome of an experiment.

A map from sample space to real numbers

Probability Distributions

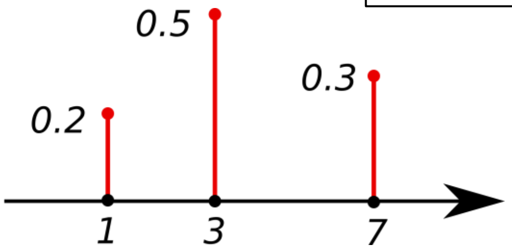
- The probability distribution of a random variable X is a description of the probabilities associated with the possible values of X.
- For discrete random variable, the distribution is just a list of values, e.g., {0.2,0.5,0.3}.



Probability Distributions

- The probability distribution of a random variable X is a description of the probabilities associated with the possible values of X.
- For discrete random variable, the distribution is just a list of values, e.g., {0.2,0.5,0.3}.

Finite or countably infinite



Probability Mass Function

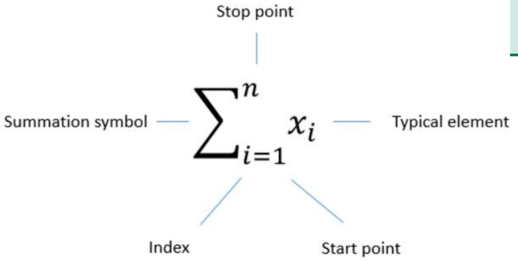
- For a discrete random variable X with possible values $x_1, x_2, \ldots, x_n$. A probability mass function $f(\cdot)$ is a function such that:

✓ $f(x_i) \geq 0$ for all $x_1, x_2, \ldots, x_n$.

✓ $\sum_{i=1}^{n} f(x_i) = 1$

✓ $f(x_i) = \boxed{P(X = x_i)}$ for all $x_1, x_2, \ldots, x_n$.

Probability that the random variable takes value x_i.



Example 1

- Toss a coin three times and count the number of heads.
- The sample space is

$$S = \{(t,t,t),(t,t,h),(t,h,t),(h,t,t),(t,h,h),(h,t,h),(h,h,t),(h,h,h)\}$$

- Lex X count the number of heads. Thus if $s = (t,t,h)$ occurs then $X(s) = 1$.
- What is the pmf for X (f, f_X, P_X)?

Example 1

- Toss a coin three times and count the number of heads.
- Step 1: Figure out the range for X

Example 1

- Toss a coin three times and count the number of heads.
- Step 1: Figure out the range for X
  - $\{0, 1, 2, 3\}$

Example 1

- Toss a coin three times and count the number of heads.
- Step 1: Figure out the range for X
  - $\{0, \ 1, \ 2, \ 3\}$
- Step 2: What are the events $\{X = 0\}, \{X = 1\}, \cdots$

Example 1

- Toss a coin three times and count the number of heads.
- Step 1: Figure out the range for X
    - $\{0, \ 1, \ 2, \ 3\}$
- Step 2: What are the events $\{X = 0\}, \{X = 1\}, \cdots$
    - $\{X = 0\} = \{(t, t, t)\}, \{X = 1\} = \{(t, t, h), (t, h, t), (h, t, t)\} \cdots$

Example 1

- Toss a coin three times and count the number of heads.
- Step 1: Figure out the range for X
  - $\{0,\ 1,\ 2,\ 3\}$
- Step 2: What are the events $\{X = 0\}, \{X = 1\}, \cdots$
  - $\{X = 0\} = \{(t, t, t)\}, \{X = 1\} = \{(t, t, h), (t, h, t), (h, t, t)\} \cdots$
- Step 3: Figure out $\mathrm{P}(X = 0), P(X = 1), \cdots$

Example 1

- Toss a coin three times and count the number of heads.
- Step 1: Figure out the range for X
  - $\{0,\ 1,\ 2,\ 3\}$
- Step 2: What are the events $\{X = 0\}, \{X = 1\}, \cdots$
  - $\{X = 0\} = \{(t, t, t)\}, \{X = 1\} = \{(t, t, h), (t, h, t), (h, t, t)\} \cdots$
- Step 3: Figure out $\mathrm{P}(X = 0), P(X = 1), \cdots$
  - $P(X = 0) = P(\{(t, t, t)\}) = \frac{1}{8}, \cdots$

Example 1

- Toss a coin three times and count the number of heads.
- Step 1: Figure out the range for X
  - $\{0, 1, 2, 3\}$
- Step 2: What are the events $\{X = 0\}, \{X = 1\}, \cdots$
  - $\{X = 0\} = \{(t, t, t)\}, \{X = 1\} = \{(t, t, h), (t, h, t), (h, t, t)\} \cdots$
- Step 3: Figure out $\mathrm{P}(X = 0), P(X = 1), \cdots$
  - $P(X = 0) = P(\{(t, t, t)\}) = \frac{1}{8}, \cdots$
- Step 4: Obtain the pmf for X

Example 1

- Toss a coin three times and count the number of heads.
- Step 1: Figure out the range for X
  - $\{0, \ 1, \ 2, \ 3\}$
- Step 2: What are the events $\{X = 0\}, \{X = 1\}, \cdots$
  - $\{X = 0\} = \{(t,t,t)\}, \{X = 1\} = \{(t,t,h),(t,h,t),(h,t,t)\} \cdots$
- Step 3: Figure out $\mathrm{P}(X = 0), P(X = 1), \cdots$
  - $P(X = 0) = P(\{(t,t,t)\}) = \frac{1}{8}, \cdots$
- Step 4: Obtain the pmf for X
  - $f(0) = P(X = 0) = \frac{1}{8}, \cdots$

Example 1

- Toss a coin three times and count the number of heads.
- The sample space is

$$S = \{(t,t,t),(t,t,h),(t,h,t),(h,t,t),(t,h,h),(h,t,h),(h,h,t),(h,h,h)\}$$

- Lex X count the number of heads. Thus if $s = (t,t,h)$ occurs then $X(s) = 1$.
- What is the cdf (cumulative distribution function) for X?

Example 1

- Toss a coin three times and count the number of heads.
- The sample space is

$$S = \{(t,t,t),(t,t,h),(t,h,t),(h,t,t),(t,h,h),(h,t,h),(h,h,t),(h,h,h)\}$$

- Lex X count the number of heads. Thus if $s = (t,t,h)$ occurs then $X(s) = 1$.
- What is the cdf (cumulative distribution function) for X?
- $F(x) = P(X \leq x)$

Example 1

- $F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)$
- For $x < 0, F(x) = 0$
- For $0 \leq x < 1, F(x) = f(0) = \frac{1}{8}$
- For $1 \leq x < 2, F(x) = f(0) + f(1) = \frac{4}{8}$
- For $2 \leq x < 3, F(x) = f(0) + f(1) + f(2) = \frac{7}{8}$
- For $3 \leq x, F(x) = f(0) + f(1) + f(2) + f(3) = 1$

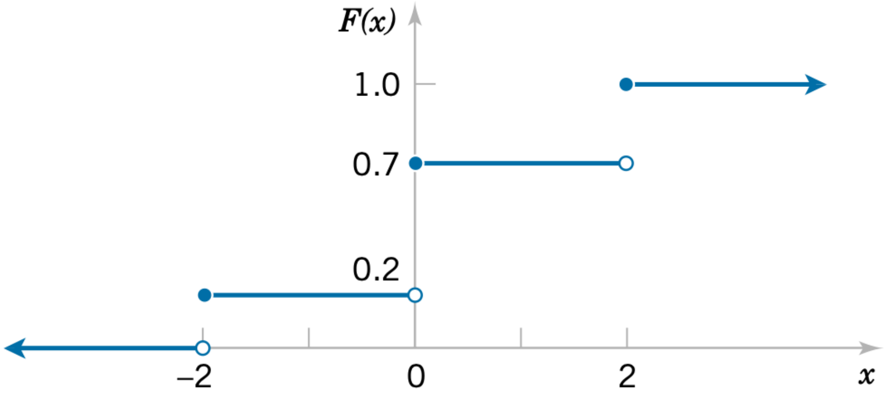
Example 1

- Given the cdf $F(x)$, how to derive the pmf $f(x)$?
- $f(x) = F(x) - \lim_{y \uparrow x} F(y)$

Example 1

- Given the cdf $F(x)$, how to derive the pmf $f(x)$?
- $f(x) = F(x) - \lim_{y \uparrow x} F(y)$
- $f(-1) = F(-1) - \lim_{y \uparrow -1} F(y) = 0 - 0 = 0$
- $f(0) = F(0) - \lim_{y \uparrow 0} F(y) = \frac{1}{8} - 0 = \frac{1}{8}$

$$F(x) = \begin{cases} 0 & x < -2 \\ 0.2 & -2 \le x < 0 \\ 0.7 & 0 \le x < 2 \\ 1 & 2 \le x \end{cases}$$



$f(-2)=0.2 \quad f(0)=0.5 \quad f(2)=0.3$

Mean and Variance

- Mean

$$\mathrm{E}[X] = \Sigma_x x\, P(X = x) = \Sigma_x x\, f(x)$$

- Variance

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

Example 1

- Toss a coin three times and count the number of heads.
- What is $E[X]$?

Example 1

- Toss a coin three times and count the number of heads.
- What is $E[X]$?
- $E[X] = 0 * f(0) + 1 * f(1) + 2 * f(2) + 3 * f(3) = \frac{3}{2}$

Example 1

- Toss a coin three times and count the number of heads.
- What is $E[X]$?
- $E[X] = 0 * f(0) + 1 * f(1) + 2 * f(2) + 3 * f(3) = \frac{3}{2}$
- What is $\mathrm{Var}[X]$?

Example 1

- Toss a coin three times and count the number of heads.
- What is $E[X]$?
- $E[X] = 0 * f(0) + 1 * f(1) + 2 * f(2) + 3 * f(3) = \frac{3}{2}$
- What is $\mathrm{Var}[X]$?
- $\mathrm{Var}[X] = \left(0 - \frac{3}{2}\right)^2 * f(0) + \left(1 - \frac{3}{2}\right)^2 * f(1) + \cdots = \frac{3}{4}$

Example 1

- Toss a coin three times and count the number of heads.
- What is $E[X]$?
- $E[X] = 0 * f(0) + 1 * f(1) + 2 * f(2) + 3 * f(3) = \frac{3}{2}$
- What is $\mathrm{Var}[X]$?
- $\mathrm{Var}[X] = \left(0 - \frac{3}{2}\right)^2 * f(0) + \left(1 - \frac{3}{2}\right)^2 * f(1) + \cdots = \frac{3}{4}$
- Is there any easy way to compute $E[X]$ and $\mathrm{Var}[X]$?

2. Some Useful formulas

Formula 1

Linearity: $\mathrm{E}[\Sigma_i X_i] = \Sigma_i \mathrm{E}[X_i]$

$\boxed{\text{No assumption on } X_i}$

An Example

Toss a coin:

- If Head, you earn 2 dollar
- If Tail, you lose 1 dollar

Suppose you toss twice, how much you will earn on average?

Sample Space	HH	TT	HT	TH
Earnings: $\Sigma_i X_i$	4	-2	1	1
Probability	0.25	0.25	0.25	0.25

Average earning (mean): 4*0.25 - 2*0.25 + 1* 0.25 + 1*0.25 = 1$

An Example

Toss a coin:

- If Head, you earn 2 dollar
- If Tail, you lose 1 dollar

Suppose you toss n times, how much you will earn on average?

Sample Space	HH…H	…	…	TT…T
Earnings: $\Sigma_i X_i$	2n	…	…	-n
Probability	$\left(\frac{1}{2}\right)^n$	…	…	$\left(\frac{1}{2}\right)^n$

Difficult to calculate directly!

An Example

Toss a coin:

- If Head, you earn 2 dollar
- If Tail, you lose 1 dollar

Suppose you toss n times, how much you will earn on average?

Linearity: $E[\Sigma_i X_i] = \Sigma_i E[X_i]$

- For each toss, you win: 2*0.5-1*0.5=0.5
- In total, you win 0.5*n.

Much easier!

Hat Check

n people go to a party and leave their hat with a hat-check person. At the end of the party, she returns hats randomly since she doesn't care about her job. Let X be the number of people who get their original hat back. What is $E[X]$?

Hat Check

n people go to a party and leave their hat with a hat-check person. At the end of the party, she returns hats randomly since she doesn't care about her job. Let X be the number of people who get their original hat back. What is $E[X]$?

Brute Force: $\Omega_X = \{0,1,2,\ldots,n-2,n\}$.

Sample space

Hat Check

n people go to a party and leave their hat with a hat-check person. At the end of the party, she returns hats randomly since she doesn't care about her job. Let X be the number of people who get their original hat back. What is $E[X]$?

Brute Force: $\Omega_X = \{0,1,2, \ldots, n-2, n\}$.

$$p_X(n) = \frac{1}{n!}$$

Hat Check

n people go to a party and leave their hat with a hat-check person. At the end of the party, she returns hats randomly since she doesn't care about her job. Let X be the number of people who get their original hat back. What is $E[X]$?

Brute Force: $\Omega_X = \{0,1,2, \dots, n-2, n\}$.

$$p_X(n) = \frac{1}{n!}$$

$$p_X(0) = ???$$

Too hard → Use linearity!

Hat Check

n people go to a party and leave their hat with a hat-check person. At the end of the party, she returns hats randomly since she doesn't care about her job. Let X be the number of people who get their original hat back. What is $E[X]$?

Quick question: does it matter where you are in line?

Hat Check

n people go to a party and leave their hat with a hat-check person. At the end of the party, she returns hats randomly since she doesn't care about her job. Let X be the number of people who get their original hat back. What is $E[X]$?

Quick question: does it matter where you are in line?

If first in line, $P(\text{get hat back}) = \frac{1}{n}$, because there are n in total.

Hat Check

n people go to a party and leave their hat with a hat-check person. At the end of the party, she returns hats randomly since she doesn't care about her job. Let X be the number of people who get their original hat back. What is $E[X]$?

Quick question: does it matter where you are in line?

If first in line, $P(\text{get hat back}) = \frac{1}{n}$, because there are n in total.
If last in line, $P(\text{get hat back}) = \frac{1}{n}$, because there is 1 left, and the chance it is yours is $\frac{1}{n}$.

Hat Check

n people go to a party and leave their hat with a hat-check person. At the end of the party, she returns hats randomly since she doesn't care about her job. Let X be the number of people who get their original hat back. What is $E[X]$?

For $i = 1, \dots, n$, let $X_i = \begin{cases} 1, & \text{if } i^{th} \text{ person got hat back} \\ 0, & \text{otherwise} \end{cases}$. Then $X = \sum_{i=1}^{n} X_i$.

Hat Check

n people go to a party and leave their hat with a hat-check person. At the end of the party, she returns hats randomly since she doesn't care about her job. Let X be the number of people who get their original hat back. What is $E[X]$?

For $i = 1, \dots, n$, let $X_i = \begin{cases} 1, & \text{if } i^{th} \text{ person got hat back} \\ 0, & \text{otherwise} \end{cases}$. Then $X = \sum_{i=1}^{n} X_i$.

We will use linearity of expectation.

Hat Check

n people go to a party and leave their hat with a hat-check person. At the end of the party, she returns hats randomly since she doesn't care about her job. Let X be the number of people who get their original hat back. What is $E[X]$?

For $i = 1, \dots, n$, let $X_i = \begin{cases} 1, & \text{if } i^{th} \text{ person got hat back} \\ 0, & \text{otherwise} \end{cases}$. Then $X = \sum_{i=1}^{n} X_i$.

We will use linearity of expectation.

$E[X_i] = 1 \cdot P(X_i = 1) + 0 \cdot P(X_i = 0) =$

Hat Check

n people go to a party and leave their hat with a hat-check person. At the end of the party, she returns hats randomly since she doesn't care about her job. Let X be the number of people who get their original hat back. What is $E[X]$?

For $i = 1, \dots, n$, let $X_i = \begin{cases} 1, & \text{if } i^{th} \text{ person got hat back} \\ 0, & \text{otherwise} \end{cases}$. Then $X = \sum_{i=1}^{n} X_i$.

We will use linearity of expectation.

$$E[X_i] = 1 \cdot P(X_i = 1) + 0 \cdot P(X_i = 0) = P(X_i = 1) = P(i^{th} \text{ person got hat back}) = \frac{1}{n}$$

Hat Check

n people go to a party and leave their hat with a hat-check person. At the end of the party, she returns hats randomly since she doesn't care about her job. Let X be the number of people who get their original hat back. What is $E[X]$?

For $i = 1, \ldots, n$, let $X_i = \begin{cases} 1, & \text{if } i^{th} \text{ person got hat back} \\ 0, & \text{otherwise} \end{cases}$. Then $X = \sum_{i=1}^{n} X_i$.

We will use linearity of expectation.

$$E[X_i] = 1 \cdot P(X_i = 1) + 0 \cdot P(X_i = 0) = P(X_i = 1) = P(i^{th} \text{ person got hat back}) = \frac{1}{n}$$

$$E[X] = E\left[\sum_{i=1}^{n} X_i\right] =$$

Hat Check

n people go to a party and leave their hat with a hat-check person. At the end of the party, she returns hats randomly since she doesn't care about her job. Let X be the number of people who get their original hat back. What is $E[X]$?

For $i = 1, \ldots, n$, let $X_i = \begin{cases} 1, & \text{if } i^{th} \text{ person got hat back} \\ 0, & \text{otherwise} \end{cases}$. Then $X = \sum_{i=1}^{n} X_i$.

We will use linearity of expectation.

$$E[X_i] = 1 \cdot P(X_i = 1) + 0 \cdot P(X_i = 0) = P(X_i = 1) = P(i^{th} \text{ person got hat back}) = \frac{1}{n}$$

$$E[X] = E\left[\sum_{i=1}^{n} X_i\right] \overset{\text{Linearity}}{=} \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} \frac{1}{n} = n \cdot \frac{1}{n} = 1$$

Hat Check

n people go to a party and leave their hat with a hat-check person. At the end of the party, she returns hats randomly since she doesn't care about her job. Let X be the number of people who get their original hat back. What is $E[X]$?

For $i = 1, \ldots, n$, let $X_i = \begin{cases} 1, & \text{if } i^{th} \text{ person got hat back} \\ 0, & \text{otherwise} \end{cases}$. Then $X = \sum_{i=1}^{n} X_i$.

NOT independent Random variables (why?)

We will use linearity of expectation.

$$E[X_i] = 1 \cdot P(X_i = 1) + 0 \cdot P(X_i = 0) = P(X_i = 1) = P(i^{th} \text{ person got hat back}) = \frac{1}{n}$$

$$E[X] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} \frac{1}{n} = n \cdot \frac{1}{n} = 1$$

Formula 2

$$\text{E}[g(X)] = \Sigma_x\, g(x)P(X = x) = \Sigma_x\, g(x)f(x)$$

- Note: **$\mathbf{E}[g(X)] \neq g(\mathbf{E}[X])$**

- Toss a coin: head as 1, tail as -1.
- Then $\text{E}[X^2] = 1^2 \times \frac{1}{2} + (-1)^2 \times \frac{1}{2} = 1$ $\boxed{g(x) = x^2}$
- But $(\text{E}[X])^2 = 0$

Variance

$$E[\underbrace{(X - E[X])^2}_{X^2 - 2XE[X] + (E[X])^2}]$$

$$E[X^2] \underbrace{- 2E[X]E[X] + (E[X])^2}_{-\ (E[X])^2}$$

Variance

$$\underbrace{E[(X - E[X])^2]}$$

$$X^2 - 2XE[X] + (E[X])^2$$

$$E[X^2] \underbrace{- 2E[X]E[X] + (E[X])^2}_{-\,(E[X])^2}$$

More useful

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

Variance (Example)

Let X be the outcome of a fair 6-sided die roll. What is $Var(X)$?

Variance (Example)

Let X be the outcome of a fair 6-sided die roll. What is $Var(X)$?

$$Var(X) = E[X^2] - E[X]^2$$

Variance (Example)

Let X be the outcome of a fair 6-sided die roll. What is $Var(X)$?

$$Var(X) = E[X^2] - E[X]^2$$

$$E[X] = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + \cdots + 6 \cdot \frac{1}{6} = 3.5$$

Variance (Example)

Let X be the outcome of a fair 6-sided die roll. What is $Var(X)$?

$$Var(X) = E[X^2] - E[X]^2$$

$$E[X] = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + \cdots + 6 \cdot \frac{1}{6} = 3.5$$

$$E[X^2] = 1^2 \cdot \frac{1}{6} + 2^2 \cdot \frac{1}{6} + 3^2 \cdot \frac{1}{6} + \cdots + 6^2 \cdot \frac{1}{6} = \frac{91}{6}$$

Variance (Example)

Let X be the outcome of a fair 6-sided die roll. What is $Var(X)$?

$$Var(X) = E[X^2] - E[X]^2$$

$$E[X] = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + \cdots + 6 \cdot \frac{1}{6} = 3.5$$

$$E[X^2] = 1^2 \cdot \frac{1}{6} + 2^2 \cdot \frac{1}{6} + 3^2 \cdot \frac{1}{6} + \cdots + 6^2 \cdot \frac{1}{6} = \frac{91}{6}$$

$$Var(X) = E[X^2] - E[X]^2 = \frac{91}{6} - (3.5)^2 = \frac{35}{12}$$

Exercise

- Toss a coin three times and count the number of heads.
- What is $E[X]$?
- $E[X] = 0 * f(0) + 1 * f(1) + 2 * f(2) + 3 * f(3) = \frac{3}{2}$
- What is $\mathrm{Var}[X]$?
- $\mathrm{Var}[X] = \left(0 - \frac{3}{2}\right)^2 * f(0) + \left(1 - \frac{3}{2}\right)^2 * f(1) + \cdots = \frac{3}{4}$
- Is there any easy way to compute $E[X]$ and $\mathrm{Var}[X]$?