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

DDA2001: Introduction to Data Science

Lecture 6: Common Distribution (Discrete)

Zicheng Wang

Recap: 1 - Some Useful formulas

Formula 1

Linearity: $E\left[\sum_i X_i\right] = \sum_i E[X_i]$

$\downarrow$

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

A More General Formula

$$E[\sum_i C_i X_i] = \sum_i C_i E[X_i]$$

C_i is a constant

How to use the linearity of expectation?

Want to calculate $E[X]$, where X is a ‘complicated’ random variable

- Step 1: Figure out a suitable decomposition for X such that $X = \sum_i X_i$
- Step 2: Compute $E[X_i]$
- Step 3: Use the linearity of expectation ($E[X] = \sum_i E[X_i]$)

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 $E[X]$?

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 $E[X]$?

Let's build a probability model

- Hats are labeled {1,2,3,…, n}, where hat m belongs to person m.
- People leave the party one by one in **a random order**.
- Everyone will take away the hat with the smallest label among the untaken hats.
- With same probability, **the order** is any of the n! permutations of {1,2,…, 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 $E[X]$?

- An outcome: people are ordered by (1,2,3,…,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 $E[X]$?

- An outcome: people are ordered by (1,2,3,…,n).
  - Since all people get their original hats: X = 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 $E[X]$?

- An outcome: people are ordered by (1,2,3,…,n).
  - Since all people get their original hats: X = n.
- An outcome: people are ordered by (2,1,3,4,…,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 $E[X]$?

- An outcome: people are ordered by (1,2,3,…,n).
  - Since all people get their original hats: X = n.
- An outcome: people are ordered by (2,1,3,4,…,n).
  - Since only the last n-2 people get their original hats: X = n-2.

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 $E[X]$?

- An outcome: people are ordered by (1,2,3,…,n).
  - Since all people get their original hats: X = n.
- An outcome: people are ordered by (2,1,3,4,…,n).
  - Since only the last n-2 people get their original hats: X = n-2.

- What's the value of P(X=n)?
- How about P(X=n-2)?
- P(X=n-3)….

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 $E[X]$?

- An outcome: people are ordered by (1,2,3,…,n).
  - Since all people get their original hats: X = n.
- An outcome: people are ordered by (2,1,3,4,…,n).
  - Since only the last n-2 people get their original hats: X = n-2.

- What's the value of P(X=n)?
- How about P(X=n-2)?
- P(X=n-3)….

Too hard to calculate f(x)!!!
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 $E[X]$?

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

E[X] = E[Σ_i X_i]= Σ_i E[X_i]

What's the value of E[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 $E[X]$?

$X_i = 1$ if and only if person i is the i-th person who takes away a hat.

How many orders may contribute to the above events?

Fix person i at i-th position of the order, there are still (n-1)! possible orders.

Thus $P(X_i = 1) = \frac{1}{n}$ and $E[X_i]=1/n$, implying $E[X] = 1$.

Formula 2

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

- $g(X)$ is also a random variable
- By definition of expectation, you need to figure out the pmf for $g(X)$ first

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

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

$g(x) = x^2$

- But $(\mathrm{E}[X])^2 = 0$

Variance

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

Variance

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

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

Variance

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

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

- $E[C] = C$ for constant C
- $E[X]$ is a constant
- $E[E[X]] = E[X]$

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

Variance

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

- $E[C] = C$ for constant C
- $E[X]$ is a constant
- $E[E[X]] = E[X]$

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

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 $Var(X)$?

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

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

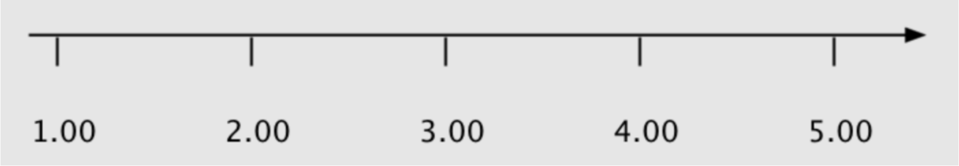
$$E[X^2] = 1^2 \cdot \frac{1}{6} + 2^2 \cdot \frac{1}{6} + 3^2 \cdot \frac{1}{6} + \cdots + 6^2 \cdot \frac{1}{6} = \frac{91}{6}$$

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

2. Common Distribution (Discrete)

Example 1

- You are the owner of a blind box store.
- Suppose that during each minute of the day, there is a probability p that a customer shows up in your store, and a probability $1 - p$ that no one shows up.



- Let X_i be the number of customers arrive in the i-th minute. What is the distribution of X_i?
- Let $Y_N = X_1 + \cdots + X_N$ be the number of customers arrive in the first N minutes of the day. What is the distribution of Y_N?
- Let Z be the minutes taken for the first customer to show up. What is the distribution of Z?

During each minute of the day, there is a probability p that a customer shows up in your store, and a probability $1-p$ that no one shows up.

?	?	?
Whether a customer arrives during a specific minute.	The number of customers arrive in N minutes	At which minute does the first customer arrive

1.Bernoulli distribution

- Take value 1 with probability p and value 0 with probability 1 − p.

1.Bernoulli distribution

- Take value 1 with probability p and value 0 with probability 1 − p.
- $f(1) = p, f(0) = 1 - p$

1.Bernoulli distribution

- Take value 1 with probability p and value 0 with probability 1 − p.
- $f(1) = p, f(0) = 1 - p$

Other Examples

- Toss a coin
- Win or lose in a game

1.Bernoulli distribution

- Take value 1 with probability p and value 0 with probability 1 − p.
- $f(1) = p, f(0) = 1 - p$

Other Examples

- Toss a coin
- Win or lose in a game

Exercise: What is $E[X_i]$ and Var(X_i)?

1.Bernoulli distribution

- Take value 1 with probability p and value 0 with probability 1 – p.
- $f(1) = p, f(0) = 1 - p$

Other Examples

- Toss a coin
- Win or lose in a game

Exercise: What is $E[X_i]$ and Var(X_i)?

- Mean=p
- Variance=p(1-p)

During each minute of the day, there is a probability p that a customer shows up in your store, and a probability $1 - p$ that no one shows up.

Bernoulli Distribution	**?**	**?**
Whether a customer arrives during a specific minute.	The number of customers arrive in N minutes	At which minute does the first customer arrive

2.Binomial distribution

- One Bernoulli trial: take value 1 with probability p and value 0 with probability 1 – p.
- N Bernoulli trials??

2.Binomial distribution

- One Bernoulli trial: take value 1 with probability p and value 0 with probability 1 – p.
- N Bernoulli trials?? -> Binomial distribution

2.Binomial distribution

- One Bernoulli trial: take value 1 with probability p and value 0 with probability 1 – p.
- N Bernoulli trials?? -> Binomial distribution

- With parameters N and p
- N independent experiments
- Each experiment: success (with probability p) or failure (with probability 1 – p).
- X: the number of success.

2.Binomial distribution

- With parameters N and p
- N independent experiments
- Each experiment: success (with probability p) or failure (with probability 1 – p).
- X: the number of success.

$$\Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \qquad \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

2.Binomial distribution

- With parameters N and p
- N independent experiments
- Each experiment: success (with probability p) or failure (with probability 1 − p).
- X: the number of success.

$$\Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \qquad \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

- Mean=np
- Variance=np(1-p)

During each minute of the day, there is a probability p that a customer shows up in your store, and a probability $1 - p$ that no one shows up.

Bernoulli Distribution	Binomial Distribution	?
Whether a customer arrives during a specific minute.	The number of customers arrive in N minutes	At which minute does the first customer arrive

2.Binomial distribution: Application

- A buyer buys N products from a producer.

2.Binomial distribution: Application

- A buyer buys N products from a producer.
- Each product can be flawed (w.p. p) or not (w.p. 1-p)

2.Binomial distribution: Application

- A buyer buys N products from a producer.
- Each product can be flawed (w.p. p) or not (w.p. 1-p)
- If k products are flawed, the producer has to refund R_k

2.Binomial distribution: Application

- A buyer buys N products from a producer.
- Each product can be flawed (w.p. p) or not (w.p. 1-p)
- If k products are flawed, the producer has to refund R_k

- How much the producer will pay to the buyer on average?

2.Binomial distribution: Application

- A buyer buys N products from a producer.
- Each product can be flawed (w.p. p) or not (w.p. 1-p)
- If k products are flawed, the producer has to refund R_k

- How much the producer will pay to the buyer on average?

$$R(N,p) = \sum_{k=1}^{N} \Pr(X = k)\, R_k$$

2.Binomial distribution: Application

- Is increasing production yield rate always beneficial for a factory?

- Production Yield Rate: the percentage of non-defective items of all produced items

2.Binomial distribution: Application

- Suppose the producer can incur cost C(p) to ensure that the product is flawed w.p. p.

2.Binomial distribution: Application

- Suppose the producer can incur cost C(p) to ensure that the product is flawed w.p. p.

- How much to invest?

2.Binomial distribution: Application

- Suppose the producer can incur cost C(p) to ensure that the product is flawed w.p. p.

- How much to invest?

- Choose p such that R(N,p) + C(p) is minimized

3.Geometric distribution

- Continuously draw a Bernoulli R.V.
- The X-th draw is the first success.
- X follows a geometric distribution.

$$\Pr(X = k) = (1 - p)^{k-1} p$$

3.Geometric distribution

- Continuously draw a Bernoulli R.V.
- The X-th draw is the first success.
- X follows a geometric distribution.

$$\Pr(X = k) = (1 - p)^{k-1} p$$

- Mean=1/p
- Variance=(1-p)/p^2

3.Geometric distribution

- Continuously draw a Bernoulli R.V.
- The X-th draw is the first success.
- X follows a geometric distribution.

$$\Pr(X = k) = (1 - p)^{k-1} p$$

Example

- Toss a coin until there is a head.
- Buy a blind box until you get a Harry Potter

During each minute of the day, there is a probability p that a customer shows up in your store, and a probability $1 - p$ that no one shows up.

Bernoulli Distribution	Binomial Distribution	Geometric Distribution
Whether a customer arrives during a specific minute.	The number of customers arrive in N minutes	At which minute does the first customer arrive

3.Geometric distribution: Application

- At each day, a machine breaks down with a probability p.

3.Geometric distribution: Application

- At each day, a machine breaks down with a probability p.

- What’s the expected duration before the machine breaks down?

Connections

- A machine breaks down w.p. p on each date.

Bernoulli distribution	Binomial distribution	Geometric Distribution
Whether the machine breaks down on a spefic date.		

Connections

- A machine breaks down w.p. p on each date.

Bernoulli distribution	Binomial distribution	Geometric Distribution
Whether the machine breaks down on a spefic date.	The number of breadowns during the first N dates.	

Connections

- A machine breaks down w.p. p on each date.

Bernoulli distribution	Binomial distribution	Geometric Distribution
Whether the machine breaks down on a spefic date.	The number of breadowns during the first N dates.	The first date the machine breaks down

Connections

- A machine breaks down w.p. p on each date.

Bernoulli distribution	Binomial distribution	Geometric Distribution
Whether the machine breaks down on a spefic date.	The number of breadowns during the first N dates.	The first date the machine breaks down