



Lecture 3: Random Variable, Part I

Yi, Yung (이용)

EE210: Probability and Introductory Random Processes KAIST EE

May 5, 2021

- (1) Random variable: Idea and formal definition
- (2) Popular discrete random variables
- (3) Summarizing random variables: Expectation and Variance
- (4) (Functions of) multiple random variables
- (5) Conditioning for random variables
- (6) Independence for random variables

May 5, 2021 1 / 45 May 5, 2021 2 / 45

Roadmap

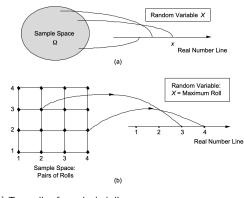
KAIST EE

Random Variable: Idea



- (1) Random variable: Idea and formal definition
- (2) Popular discrete random variables
- (3) Summarizing random variables: Expectation and Variance
- (4) (Functions of) multiple random variables
- (5) Conditioning for random variables
- (6) Independence for random variables

- In reality, many outcomes are numerical, e.g., stock price.
- Even if not, very convenient if we map numerical values to random outcomes, e.g., '0' for male and '1' for female.



(b) Two rolls of tetrahedral dice

Random Variable: More Formally



Example



- Mathematically, a random variable X is a function which maps from Ω to \mathbb{R} .
- Notation. Random variable X, numerical value x.
- Different random variables can be defined on the same sample space.
- For a fixed value x, we can associate an event that a random variable X has the value x, i.e., $\{\omega \in \Omega \mid X(\omega) = x\}$
- Assume that values x are discrete¹ such as $1, 2, 3, \ldots$. For notational convenience,

$$p_X(x) \triangleq \mathbb{P}(X=x) \triangleq \mathbb{P}\Big(\{\omega \in \Omega \mid X(\omega)=x\}\Big)$$

- For a discrete random variable X, we call $p_X(x)$ probability mass function (PMF).
 - ¹Finite or countably infinite.

L3(1)

May 5, 2021 5 / 45

- Rolls a dice, $\Omega = \{1, 2, 3, 4, 5, 6\}$
- Define a random variable X = 1 for even numbers and X = 0 for odd numbers
- Event $A_1 = \{\omega \in \Omega \mid X(\omega) = 1\} = \{2,4,6\} \subset \Omega$, but simply $A_1 = \{X = 1\}$
- Event $A_0 = \{ \omega \in \Omega \mid X(\omega) = 0 \} = \{1, 3, 5\} \subset \Omega$, but simply $A_0 = \{X = 0\}$
- Remember that the random variable X is a function from Ω to \mathbb{R}

L3(1) May 5, 2021 6 / 45

Roadmap



Bernoulli X with parameter $p \in [0, 1]$



- (1) Random variable: Idea and formal definition
- (2) Popular discrete random variables
- (3) Summarizing random variables: Expectation and Variance
- (4) (Functions of) multiple random variables
- (5) Conditioning for random variables
- (6) Independence for random variables

Only binary values

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

In other words, $p_X(0) = 1 - p$ and $p_X(1) = p$ from our PMF notation.

- Models a trial that results in binary results, e.g., success/failure, head/tail
- Very useful for an indicator rv of an event A. Define a rv $\mathbf{1}_A$ as:

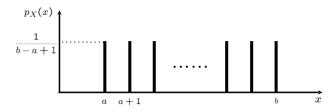
$$\mathbf{1}_A = \begin{cases} 1, & \text{if } A \text{ occurs}, \\ 0, & \text{otherwise} \end{cases}$$

L3(2) May 5, 2021 7 / 45 L3(2) May 5, 2021

¹w.p.: with probability

- integers a, b, where $a \le b$
- Choose a number out of $\Omega = \{a, a+1, \dots, b\}$ uniformly at random.
- $p_X(i) = \frac{1}{b-a+1}, i \in \Omega$

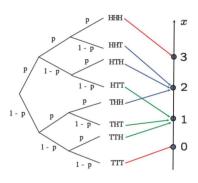
L3(2)



• Models complete ignorance (I don't know anything about X)

- Models the number of successes in a given number of independent trials
- n independent trials, where one trial has the success probability p.

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$$



$$\binom{1}{\binom{n}{k}} = \frac{n!}{\frac{k!(n-k)!}{13(2)}}, \text{ which we read '} n \text{ choose } k'.$$

May 5, 2021 10 / 45

Geometric X with parameter p



May 5, 2021 9 / 45

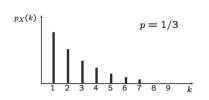
Roadmap



- Infinitely many independent Bernoulli trials, where each trial has success probability p
- Random variable: number of trials until the first success.

$$p_X(k) = (1-p)^{k-1}p$$

 Models waiting times until something happens.



- (1) Random variable: Idea and formal definition
- (2) Popular discrete random variables
- (3) Summarizing random variables: Expectation and Variance
- (4) (Functions of) multiple random variables
- (5) Conditioning for random variables
- (6) Independence for random variables



Average

Definition

$$\mathbb{E}[X] = \sum_{x} x p_X(x)$$

- $p_X(x)$: relative frequency of value x (trials with x/total trials)
- Example. Bernoulli rv with p

$$\mathbb{E}[X] = 1 \times p + 0 \times (1 - p) = p = p_X(1)$$

L3(3)

May 5, 2021 13 / 45

Not very surprising. Easy to prove using the definition.

- If $X \ge 0$, $\mathbb{E}[X] \ge 0$.
- If a < X < b, $a < \mathbb{E}[X] < b$.
- For a constant c, $\mathbb{E}[c] = c$.

L3(3)

May 5, 2021 14 / 45

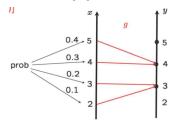
Expectation of a function of a RV

KAIST EE

Variance



- For a rv X, Y = g(X) is also a r.v.
- $\mathbb{E}[Y] = \mathbb{E}[g(X)] = \sum_{x} g(x) p_X(x)$
- Compute $\mathbb{E}[Y]$ for the following:



$$4 \times (0.4 + 0.3) + 3 \times (0.1 + 0.2)$$

= 2.8 + 0.9 = 3.7

= 2.8 + 0.9 = 3.7

- Measures how much the spread of a PMF is.
- What about $\mathbb{E}[X \mu]$, where $\mu = \mathbb{E}[X]$? Zero
- Then, what about $\mathbb{E}[(X \mu)^2]$?

Variance, Standard Deviation

$$var[X] = \mathbb{E}[(X - \mu)^2]$$

$$\sigma_X = \sqrt{\operatorname{var}[X]}$$

Linearity of Expectation

$$\mathbb{E}[aX+b]=a\mathbb{E}[X]+b$$



• $\operatorname{var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$

$$var[X] = \mathbb{E}[X^2 - 2\mu X + \mu^2]$$

= $\mathbb{E}[X^2] - 2\mu \mathbb{E}[X] + \mu^2 = \mathbb{E}[X^2] - \mu^2$

• Y = X + b, var[Y] = var[X]

$$\operatorname{var}[Y] = \mathbb{E}[(X+b)^2] - (\mathbb{E}[X+b])^2$$

• Y = aX, $var[Y] = a^2 var[X]$

$$var[Y] = \mathbb{E}[a^2X^2] - (a\mathbb{E}[X])^2$$

Example: Variance of a Bernoulli rv (p)

$$\mu = \mathbb{E}[X] = 1 \times p + 0 \times (1 - p) = p$$

$$\mathbb{E}[X^2] = 1 \times p + 0 \times (1 - p) = p$$

$$var[X] = \mathbb{E}[X^2] - \mu^2 = p - p^2$$

= $p(1 - p)$

(1) Random variable: Idea and formal definition

(2) Popular discrete random variables

(3) Summarizing random variables: Expectation and Variance

(4) (Functions of) multiple random variables

(5) Conditioning for random variables

(6) Independence for random variables

L3(3)

May 5, 2021 17 / 45

L3(4)

May 5, 2021 18 / 45

Joint PMF

VIDEO PAUSE

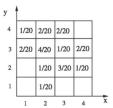
Functions of Multiple RVs



- Joint PMF. For two random variables X, Y, consider two events $\{X = x\}$ and $\{Y = y\}$, and
- $p_{X,Y}(x,y) \triangleq \mathbb{P}(\{X=x\} \cap \{Y=y\})$
- $\sum_{x}\sum_{y}p_{X,Y}(x,y)=1$
- Marginal PMF.

$$p_X(x) = \sum_{y} p_{X,Y}(x,y),$$
$$p_Y(y) = \sum_{y} p_{X,Y}(x,y)$$

Example.



$$p_{X,Y}(1,3) = 2/20$$

$$p_X(4) = 2/20 + 1/20 = 3/20$$

$$\mathbb{P}(X = Y) = 1/20 + 4/20 + 3/20 = 8/20$$

$$p_Z(z) = \mathbb{P}(g(X, Y) = z) = \sum_{(x,y):g(x,y)=z} p_{X,Y}(x,y)$$

• Consider a rv Z = g(X, Y). (Ex) X + Y, $X^2 + Y^2$. Then, PMF of Z is:

Similarly,

$$\mathbb{E}[Z] = \mathbb{E}[g(X,Y)] = \sum_{x} \sum_{y} g(x,y) p_{X,Y}(x,y)$$

- Remember: $\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$
- Similarly,

$$\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$

(easy to prove, using the definition.)

- $\mathbb{E}[X_1 \ldots + X_n] = \mathbb{E}[X_1] + \ldots + \mathbb{E}[X_n]$
- $\mathbb{E}[2X+3Y-Z] = 2\mathbb{E}[X]+3\mathbb{E}[Y]-\mathbb{E}[Z]$

- Example. Mean of a binomial rv Y with (n, p)
- Y: number of successes in n Bernoulli trials with p
- $Y = X_1 + ... X_n$, where X_i is a Bernoulli rv.
- $\mathbb{E}[Y] = n\mathbb{E}[X_i] = n\mathbb{P}(X_i = 1) = np$

Message. When some rv X is written as a linear combination of other rvs, X becomes easy to handle.

- (1) Random variable: Idea and formal definition
- (2) Popular discrete random variables
- (3) Summarizing random variables: Expectation and Variance
- (4) (Functions of) multiple random variables
- (5) Conditioning for random variables
- (6) Independence for random variables

L3(4) May 5, 2021 21 / 45 L3(5)

Conditional PMF: Conditioning on an event

KAIST EE

Example: Conditional PMF



 $A = \{X > 2\}$

May 5, 2021

22 / 45

Remember two probability laws: $\mathbb{P}(\cdot)$ and $\mathbb{P}(\cdot|A)$ for an event A.

•
$$p_X(x) \triangleq \mathbb{P}(X = x)$$

•
$$\mathbb{E}[X] = \sum_{x} x p_X(x)$$

•
$$\mathbb{E}[g(X)] = \sum_{x} g(x) p_X(x)$$

•
$$\operatorname{var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$$

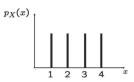
• $p_{X|A}(x) \triangleq \mathbb{P}(X = x|A)$

•
$$\mathbb{E}[X|A] \triangleq \sum_{x} x p_{X|A}(x)$$

•
$$\mathbb{E}[g(X)|A] \triangleq \sum_{x} g(x) p_{X|A}(x)$$

•
$$\operatorname{var}[X|A] \triangleq \mathbb{E}[X^2|A] - (\mathbb{E}[X|A])^2$$

• (Note) $p_{X|A}(x)$, $\mathbb{E}[X|A]$, $\mathbb{E}[g(X)|A]$, and var[X|A] are all just notations!



$$\mathbb{E}[X] = \frac{1}{4}(1+2+3+4) = 2.5$$

$$var[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$$
$$= \frac{1}{4}(1 + 2^2 + 3^2 + 4^2) - 2.5^2$$



$$\mathbb{E}[X|A] = \frac{1}{3}(2+3+4) = 3$$

$$var[X|A] = \mathbb{E}[X^2|A] - (\mathbb{E}[X|A])^2$$
$$= \frac{1}{3}(2^2 + 3^2 + 4^2) - 3^2 = 2/3$$





What do we mean by "conditioning on a rv"? Consider $A = \{Y = y\}$ for a rv Y.

- $p_{X|A}(x) \triangleq \mathbb{P}(X = x|A)$
- $\mathbb{E}[X|A] \triangleq \sum_{x} x p_{X|A}(x)$
- $\mathbb{E}[g(X)|A] \triangleq \sum_{x} g(x) p_{X|A}(x)$
- $\operatorname{var}[X|A] \triangleq \mathbb{E}[X^2|A] (\mathbb{E}[X|A])^2$

- $p_{X|Y}(x|y) \triangleq \mathbb{P}(X=x|Y=y)$
- $\mathbb{E}[X|Y=y] \triangleq \sum_{x} x p_{X|Y}(x|y)$
- $\mathbb{E}[g(X)|Y=y] \triangleq \sum_{x} g(x)p_{X|Y}(x|y)$
- $\operatorname{var}[X|Y = y] \triangleq \mathbb{E}[X^2|Y = y] (\mathbb{E}[X|Y = y])^2$

Conditional PMF

$$p_{X|Y}(x|y) \triangleq \mathbb{P}(X=x|Y=y) = \frac{p_{X,Y}(x,y)}{p_{Y}(y)}$$

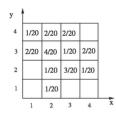
for y such that $p_Y(y) > 0$.

- $\sum_{x} p_{X|Y}(x|y) = 1$
- Multiplication rule

$$p_{X,Y}(x,y) = p_Y(y)p_{X|Y}(x|y)$$
$$= p_X(x)p_{Y|X}(y|x)$$

• $p_{X,Y,Z}(x,y,z) = p_X(x)p_{Y|X}(y|x)p_{Z|X,Y}(z|x,y)$

VIDEO PAUSE



$$p_{X|Y}(2|2) = \frac{1}{1+3+1}$$

$$p_{X|Y}(3|2) = \frac{3}{1+3+1}$$

$$\mathbb{E}[X|Y=3] = 1(2/9) + 2(4/9) + 3(1/9) + 4(2/9)$$

L3(5)

May 5, 2021 25 / 45

L3(5)

May 5, 2021 26 / 45

Remind: Total Probability Theorem (from Lecture 2)



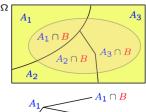
Total Probability Theorem: $B = \{X = x\}$

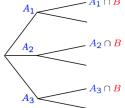


- Partition of Ω into A_1, A_2, A_3
- Known: $\mathbb{P}(A_i)$ and $\mathbb{P}(B|A_i)$
- What is $\mathbb{P}(B)$?

Total Probability Theorem

$$\mathbb{P}(B) = \sum_{i} \mathbb{P}(A_{i}) \mathbb{P}(B|A_{i})$$





• Partition of Ω into A_1, A_2, A_3

Total Probability Theorem

$$p_{X}(x) = \sum_{i} \mathbb{P}(A_{i})\mathbb{P}(X = x|A_{i}) = \sum_{i} \mathbb{P}(A_{i})p_{X|A_{i}}(x)$$





Total Expectation Theorem for $\{A_i\}$

KAIST EE

Total Expectation Theorem for $\{Y = y\}$

KAIST EE

• Partition of Ω into A_1, A_2, A_3

Total Probability Theorem

$$p_X(x) = \sum_i \mathbb{P}(A_i)\mathbb{P}(X = x|A_i) = \sum_i \mathbb{P}(A_i)p_{X|A_i}(x)$$

Total Expectation Theorem

$$\mathbb{E}[X] = \sum_{i} \mathbb{P}(A_i) \mathbb{E}[X|A_i]$$

• Weighted average of expectations from A_i 's perspective

 $\begin{array}{c}
\Omega \\
A_1 \\
A_2 \\
A_2 \\
A_3 \\
A_4 \\
A_1 \\
A_1 \\
A_1 \\
A_2
\end{array}$



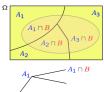
• Partition of Ω into A_1, A_2, A_3

Total Expectation Theorem

$$\mathbb{E}[X] = \sum_{i} \mathbb{P}(A_{i}) \mathbb{E}[X|A_{i}]$$

Total Expectation Theorem

$$\mathbb{E}[X] = \sum_{y} \mathbb{P}(Y = y) \mathbb{E}[X | Y = y] = \sum_{y} p_{Y}(y) \mathbb{E}[X | Y = y]$$





L3(5)

May 5, 2021 29 / 45

L3(5)

May 5, 2021 30 / 45

Example 1: Total Expectation Theorem

• Question. What is $\mathbb{E}(X)$?

(1) Just using the definition of expectation,

$$\mathbb{E}[X] = \frac{1}{9}(0+1+2) + \frac{2}{9}(6+7+8)$$
$$= \frac{3+12+14+16}{9} = 5$$

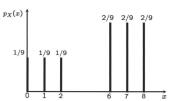
(2) Let's use TET, for which consider

$$A_1 = \{X \in \{0, 1, 2\}\}, \ A_2 = \{X \in \{6, 7, 8\}\}$$

$$\mathbb{E}[X] = \sum_{i=1,2} \mathbb{P}(A_i)\mathbb{E}[X|A_i]$$

$$= 1/3 \cdot 1 + 2/3 \cdot 7 = 5$$





Example 2: Mean of Geometric rv

KAISTEE

- Write softwares over and over, and each time w.p. p of working correctly (independent from previous programs).
- X: number of trials until the program works correctly.
- (Q) $\mathbb{E}(X)$?
- ullet X is a geometric rv
- Direct computation is boring.

$$\mathbb{E}[X] = \sum_{k=1}^{\infty} k(1-p)^{k-1}p = p + 2(1-p)p + 3(1-p)^2p + \cdots$$

Total expectation theorem and a notion of memorylessness helps a lot.







How long do I have to wait? Probability of waiting for more than n mins?

$$\mathbb{P}(X > n + m | X > m)$$



A bus left at time 0





How long do I have to wait? Probability of waiting for more than n mins?





Lin arrives

$$\mathbb{P}(X > n)$$

L3(5)

May 5, 2021 33 / 45

- Some random variable often does not have memory.
- Definition. A random variable X is called memoryless if, for any n, m > 0,

$$\mathbb{P}(X > n + m | X > m) = \mathbb{P}(X > n)$$

• Meaning. Conditioned on X > m, X - m's distribution is the same as the original X.

$$\mathbb{P}(X-m>n|X>m)=\mathbb{P}(X>n)$$

L3(5) May 5, 2021 34 / 45

Background: Memoryless Property of Geometric RVs



Back to Example 2: Mean of Geometric rv



- Theorem. Any geometric random variable is memoryless.
- Remind. Geometric rv X with parameter p

$$\mathbb{P}(X = k) = (1 - \rho)^{k-1} \rho, \quad \mathbb{P}(X > k) = \sum_{i=k+1}^{\infty} (1 - \rho)^{i-1} \rho = (1 - \rho)^k$$

Proof.

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

• Meaning. Conditioned on X > m, X - m is geometric with the same parameter.

• $A_1 = \{X = 1\}$ (first try is success), $A_2 = \{X > 1\}$ (first try is failure).

$$\begin{split} \mathbb{E}[X] &= 1 + \mathbb{E}[X-1] \\ &= 1 + \mathbb{P}(A_1)\mathbb{E}[X-1|X=1] + \mathbb{P}(A_2)\mathbb{E}[X-1|X>1] \\ &= 1 + (1-p)\mathbb{E}[X] \end{split} \tag{from TET)}$$

• Thus,
$$\mathbb{E}[X] = \frac{1}{p}$$

- (1) Random variable: Idea and formal definition
- (2) Popular discrete random variables
- (3) Summarizing random variables: Expectation and Variance
- (4) (Functions of) multiple random variables
- (5) Conditioning for random variables
- (6) Independence for random variables

L3(6) May 5, 2021 37 / 45

y

1/20 2/20 2/20

3 2/20 4/20 1/20 2/20

1/20

1 2 3 4

1/20 3/20 1/20

Two events

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$$

$$\mathbb{P}(A \cap B | C) = \mathbb{P}(A | C) \cdot \mathbb{P}(B | C)$$

• A rv and an event

$$\mathbb{P}(\{X = x\} \cap B) = \mathbb{P}(X = x) \cdot \mathbb{P}(B), \text{ for all } x$$

$$\mathbb{P}(\{X = x\} \cap B | C) = \mathbb{P}(X = x | C) \cdot \mathbb{P}(B | C), \text{ for all } x$$

Two rvs

$$\mathbb{P}(\{X = x\} \cap \{Y = y\}) = \mathbb{P}(X = x) \cdot \mathbb{P}(Y = y), \text{ for all } x, y$$
$$p_{X,Y}(x, y) = p_X(x) \cdot p_Y(y)$$

$$\mathbb{P}(\{X=x\} \cap \{Y=y\} | Z=z) = \mathbb{P}(X=x|Z=z) \cdot \mathbb{P}(Y=y|Z=z), \text{ for all } x, y$$
$$p_{X,Y|Z}(x,y) = p_{X|Z}(x) \cdot p_{Y|Z}(y)$$

L3(6)

May 5, 2021 38 / 45

Example

KAIST EE

Expectation and Variance

 Always true. $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$

• Generally, $\mathbb{E}[g(X,Y)] \neq g(\mathbb{E}[X],\mathbb{E}[Y])$

• However, if $X \perp \!\!\!\perp Y$,

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$$

$$\mathbb{E}[g(X)h(Y)] = \mathbb{E}[g(X)]\mathbb{E}[g(Y)]$$

Proof.

$$\mathbb{E}[g(X)h(Y)] = \sum_{x} \sum_{y} g(x)h(y)p_{X,Y}(x,y)$$
$$= \sum_{x} g(x)p_{X}(x) \sum_{y} h(y)p_{Y}(y)$$

- KAIST EE
- Always true. $var[aX] = a^2 var[X], var[X + a] = var[X]$
- Generally, $var[X + Y] \neq var[X] + var[Y]$ (next slide)
- However, if $X \perp \!\!\!\perp Y$. var[X + Y] = var[X] + var[Y]

$$p_{X,Y}(1,1) = 0$$

$$p_X(1) = 3/20$$

$$p_Y(1)=1/20$$

• $X \perp \!\!\!\perp Y | \{X \leq 2 \text{ and } Y \geq 3\}$?

	VIDEO PAUSE	
Y = 4 (1/3)	1/9	2/9
Y = 3 (2/3)	2/9	4/9
	X = 1 (1/3)	X = 2(2/3)

- Yes.



• Why not generally true?

$$\begin{aligned} \operatorname{var}[X+Y] &= \mathbb{E}[(X+Y)^2] - (\mathbb{E}[X+Y])^2 \\ &= \mathbb{E}[X^2+Y^2+2XY] - \left((\mathbb{E}[X])^2 + (\mathbb{E}[Y])^2 + 2\mathbb{E}[X]\mathbb{E}[Y]\right) \\ &= \operatorname{var}[X] + \operatorname{var}[Y] + 2\left(\mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]\right) \end{aligned}$$

- $\circ \mid X \perp \!\!\! \perp Y \mid$ is a sufficient condition for $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$
- o Also, a necessary condition? we will see later, when we study covariance.

- n people throw their hats in a box and then pick one at random
- X: number of people with their own hat
- $\mathbb{E}[X]$? var[X]?
- All permutations are equally likely as 1/n!. Thus, this equals to picking one hat at a time.
- Key step 1. Define a rv $X_i = 1$ if i selects its own hat and 0 otherwise.

$$X = \sum_{i=1}^{n} X_i.$$

• $\{X_i\}, i = 1, 2, ..., n$: identically distributed (from symmetry)

L3(6)

May 5, 2021 41 / 45

L3(6)

May 5, 2021 42 / 45

Example: The hat problem (2)





- $\mathbb{E}[X] = n\mathbb{E}[X_1] = n\mathbb{P}(X_1 = 1) = n \times \frac{1}{n} = 1.$
- Key step 2. Are X_i s are independent? If yes, easy to get var(X).
- Assume n=2. Then, $X_1=1 \rightarrow X_2=1$, and $X_1=0 \rightarrow X_2=0$. Thus, dependent.

$$\operatorname{var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \mathbb{E}\left[\sum_{i} X_i^2 + \sum_{i,j:i \neq j} X_i X_j\right] - (\mathbb{E}[X])^2$$

$$\mathbb{E}[X_i^2] = \mathbb{E}[X_1^2] = 1 \times \frac{1}{n} + 0 \times \frac{n-1}{n} = \frac{1}{n}$$

$$\mathbb{E}[X_i X_i] = \mathbb{E}[X_1 X_2] = 1 \times \mathbb{P}(X_1 X_2 = 1) = \mathbb{P}(X_1 = 1) \mathbb{P}(X_2 = 1 | X_1 = 1), \quad (i \neq j)$$

- $\mathbb{E}[X^2] = n\mathbb{E}[X_1^2] + n(n-1)\mathbb{E}[X_1X_2] = n\frac{1}{n} + n(n-1)\frac{1}{n(n-1)} = 2$
- var(X) = 2 1 = 1

Questions?

Review Questions



- 1) What is Random Variable? Why is it useful?
- 2) What is PMF (Probability Mass Function)?
- 3) Explain Bernoulli, Binomial, Poisson, Geometric rvs, when they are used and what their PMFs are.
- 4) What are joint and marginal PMFs?
- 5) Describe and explain the total probability/expectation theorem for random variables?
- 6) When is it useful to use total probability/expectation theorem?
- 7) What is conditional independence?

L3(6) May 5, 2021 45 / 45