Last lecture

Probability with equally likely outcomes (Ch 1.4)

- Draw socks from the drawer
- Poker hands

Random Variables (RV) (Ch 2.1)

- Definition

Agenda

Random Variables (RV) (Ch 2.1)

- Probability Mass Function (pmf)
- Mean and Variance (Ch 2.2)

Conditional Probability (Ch 2.3)

- Motivation
- Examples
- Solver

Law of Total Probability (Ch 2.10)

Probability Mass Function (PMF)

Probability Mass Function (PMF)

- $p_X(u) = P\{X = u\}$ for a discrete RV X
- $\sum_i p_X(u_i) = 1$
- Let X be the outcome of a fair die roll
  - $p_X(2) = \frac{1}{6}$
- PMF can determine probabilities of all events determined by X

e.g. fair die

outcome	X
1	1
2	-1
3	1
4	-1
5	1
6	-1

X describes "Even"

$X = -1$

X cannot describe "Multiple of 3"

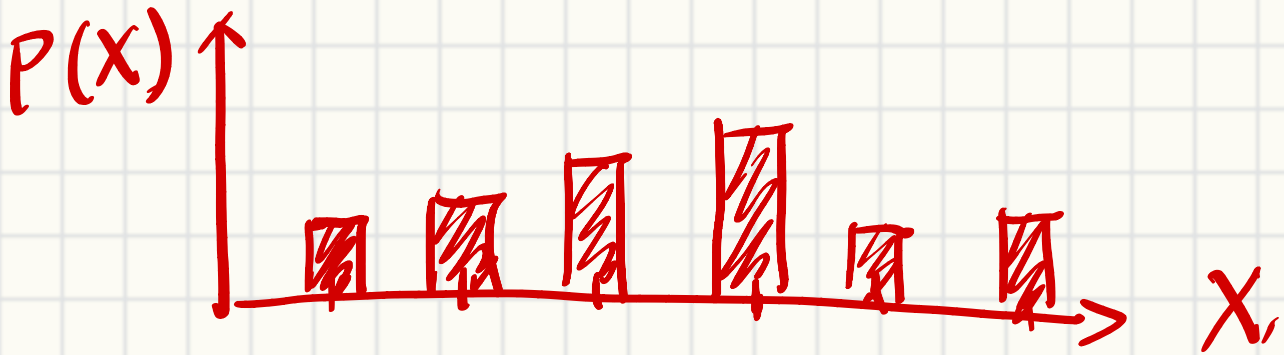
Slido!

Let's create a custom die and plot the PMF

Please vote for your preferred number from 1-6

#2887579

	1	2	3	4	5	6
%	11	15	22	26	11	15



$\Rightarrow$ NOT Equally probable.

P(Roll twice, get a pair)

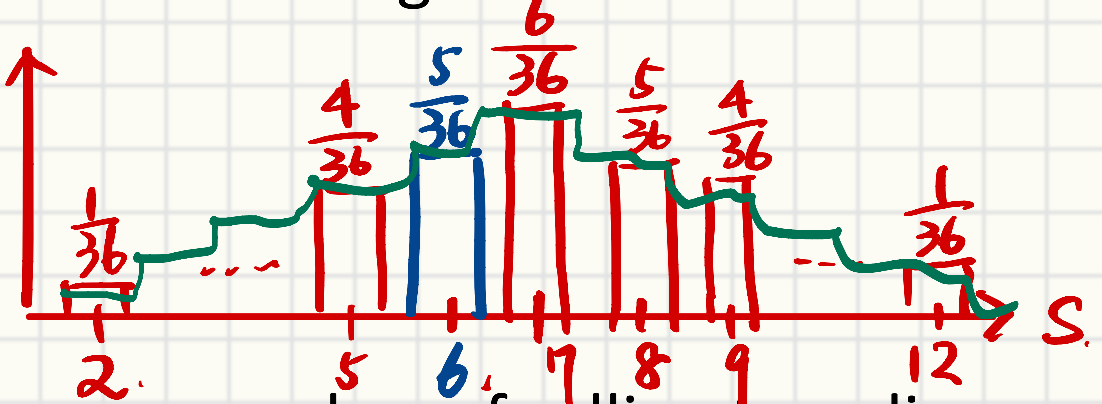
$$= P([1,1]) + P([2,2]) \ldots\ldots P([6,6])$$

$$= (0.11)^2 + (0.15)^2 + (0.22)^2 + (0.26)^2 + (0.11)^2$$

$$+ (0.15)^2$$

Probability Mass Function (PMF)

- Let S be the sum of rolling two dice — S is a RV.
  - $p_S =$

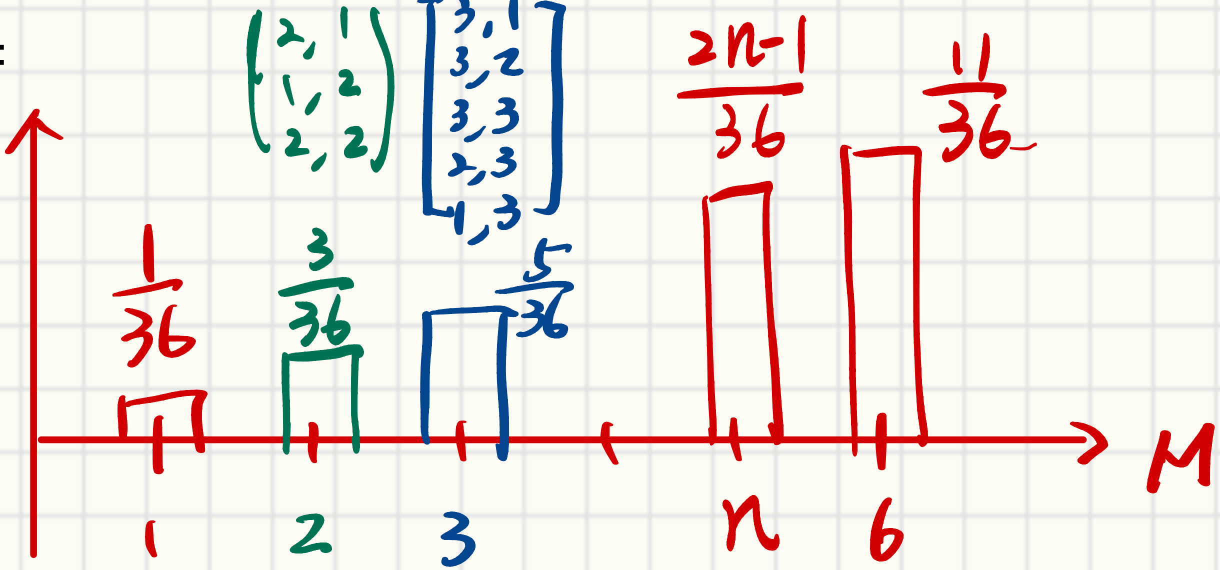


$$P(S=7) = \frac{\|\{(1,6),(2,5)\ \dots\ (6,1)\}\|}{36} = \frac{6}{36} = \frac{1}{6}$$

$$P(S=6) = \frac{\|\{(1,5)\ \dots\ (5,1)\}\|}{36} = \frac{5}{36}$$

$(3,3)$

- Let M be the max number of rolling two dice
  - $p_M =$



Probability Mass Function (PMF)

- Let N be the # of toss until getting first tail

$$P_N(n) = (0.5)^N = \left(\frac{1}{2}\right)^N$$

$P_N(1) = \frac{1}{2}$

$P_N(2) = \frac{1}{2} \times \frac{1}{2}$

$\{H\}$ $\{T\}$

$P_N(3) = \frac{1}{8}$

- Let M be the # of heads observed until getting first tail

$$M \in [0, \infty]$$

$\Rightarrow \{H, HT\}$

$P_M(0) = \frac{1}{2}$

$P_M(1) = \frac{1}{4}$

$$P_M(m) = \left(\frac{1}{2}\right)^{m-1}$$

Q. M & N are dependent or not ?

Mean and Variance (Ch 2.2)

Mean : Expectation of X.

Do we need detailed p_{height}?

- In many cases, we just need mean μ_X instead of p_X
- $\mu_X = E[X] = \sum_i a_i p_X(a_i)$

- X is the number for a die roll

$Y \in \{2, 4, 6, 8, 10, 12\}$

- $Y = 2X$

$E[Y] = \frac{1}{6}\times 2 + \frac{1}{6}\times 4 \cdots\cdots \frac{1}{6}\times 12$

$= 7.$

- $Z = |X - 3|$

$Z \in \{0, 1, 2, 3\}$

$X = 3, \overset{2}{4}, \overset{1}{5}, 6$

$$E[Z] = \frac{1}{6}\times 0 + \frac{2}{6}\times 1 + \frac{2}{6}\times 2 + \frac{1}{6}\times 3$$

$$= \frac{9}{6} = \frac{3}{2}.$$

Function of RV - LOTUS

↳ Law of The Unconciousness

X is RV uniformly sampled from $\{-1, 0, 1, 2, 3\}$

Statistician

- $p_X(x) = \frac{1}{5}.$

$$\begin{array}{cccccc} X & 0 & -1, 1 & 2 & 3 \end{array}$$

$Y = X^2$

① $Y \in \{0, 1, 4, 9\}$

- $\mu_y = E[Y] = \frac{1}{5}\times 0 + \frac{2}{5}\times 1 + \frac{1}{5}\times 4 + \frac{1}{5}\times 9,$

LOTUS

$g(X)=Y$	1	0	1	4	9.
X	-1	0	1	2	3

$$E[Y] = \frac{1}{5}\times 1 + \frac{1}{5}\times 0 + \frac{1}{5}\times 1 + \frac{1}{5}\times 4 + \frac{1}{5}\times 9.$$

Function of RV - LOTUS

X is RV uniformly sampled from $\{-1, 0, 1, 2, 3\}$, $Y = X^2$

But we can also compute $E[Y]$ from p_X!

Mean of RV function $g(X)$ is

$$E[g(X)] = \sum_x g(x)\, p(x)$$

Law of the unconscious statistician (LOTUS)

LOTUS examples

X is rolling a D6, Y is rolling a D8 (8-sided die)

$E[XY] = ?$

$$\sum_{\substack{x\in[1,6]\\ y\in[1,8]}} x y \, p(x)\, p(y)$$

$$= \sum x y \times \frac{1}{6} \times \frac{1}{8}$$

$$= \frac{1}{48}(1+2+3+4+5+6)(1+2+3+4+5+6+7+8) ?$$

$$= \frac{1}{48} \times 21 \times 36 = \frac{63}{4}$$

LOTUS examples

Math magic trick

- Pick a number from 1-10
- Multiply it by 3
- Subtract it by 2
- Divided it by 6 and keep the remainder
- Plus 2
- Is it 3 or 6?

Variance and Standard Deviation

Do you want your salary to be

$p_X(10K) = 1$ or

$p_Y(0) = \cancel{0.09}, p_Y(1000K) = 0.01$

0.99

What if it's bonus?

$p_X(1K) = 1$ or

$p_Y(0) = \cancel{0.09}, p_Y(100K) = 0.01$

0.99

Variance and Standard Deviation

Mean is important… but not complete enough

- Variance $Var(X)$ is how PMF spreads apart from μ_X
- $Var(X) \triangleq E[\underbrace{(X-\mu_X)}_{\text{Deviation}}{}^2] = E[X^2] - (E[X])^2$

(σ_X^2)

$$E[(X-\mu_X)^2]$$

$$= E[X^2 - 2\mu_X X + \mu_X^2]$$

$$= E[X^2] - 2\mu_X \cancel{E[X]}^{\mu_X} + \mu_X^2 = E[X^2] - \mu_X^2$$

Variance and Standard Deviation

Standard deviation $\sigma_X \triangleq \sqrt{Var(X)}$; $Var(X) = \sigma_X^2$

σ_X is of the same unit as X

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

$$= E[X^2] + \cancel{2cE[X]} + \cancel{c^2} - (\mu_X^2 + \cancel{2c\mu_X} + \cancel{c^2})$$

$Var(X + c) =$ $Var(X)$

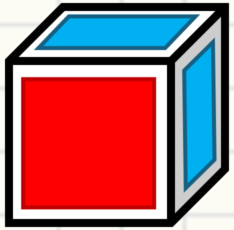
$Var(aX + c) =$ $a^2 \, Var(X)$

Note: $2X \neq X + X$

Law of total probability

There are 3 dice A, B, C in the bag

- $A = [R \times 1; B \times 5]$
- $B = [R \times 2; B \times 4]$
- $C = [R \times 3; B \times 3]$



Draw one die and roll many times

- $P(R_1)$
- $P(R_2|R_1)$