Last lecture

Probability with equally likely outcomes (Ch 1.4)

- Poker hands

Random Variables (RV) (Ch 2.1)

- Definition
- Probability Mass Function (pmf)

Agenda

Random Variables (RV)

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

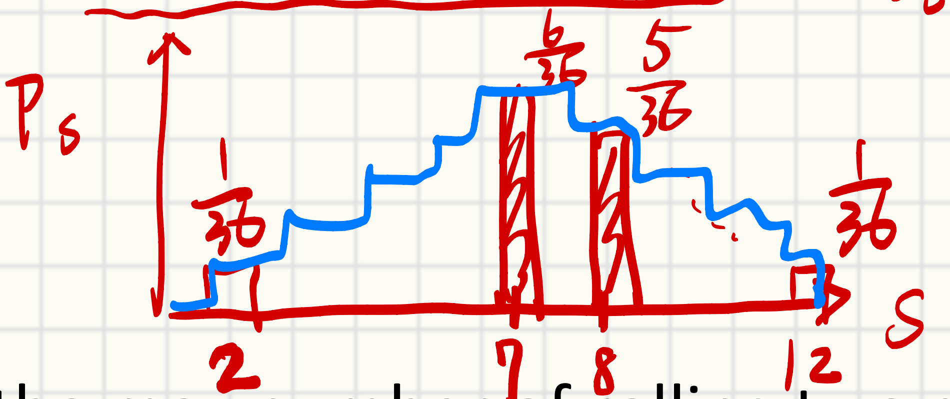
Probability Mass Function (PMF)

6×6

$1+1$ $6+6$

- Let S be the sum of rolling two dice
  - $p_S =$

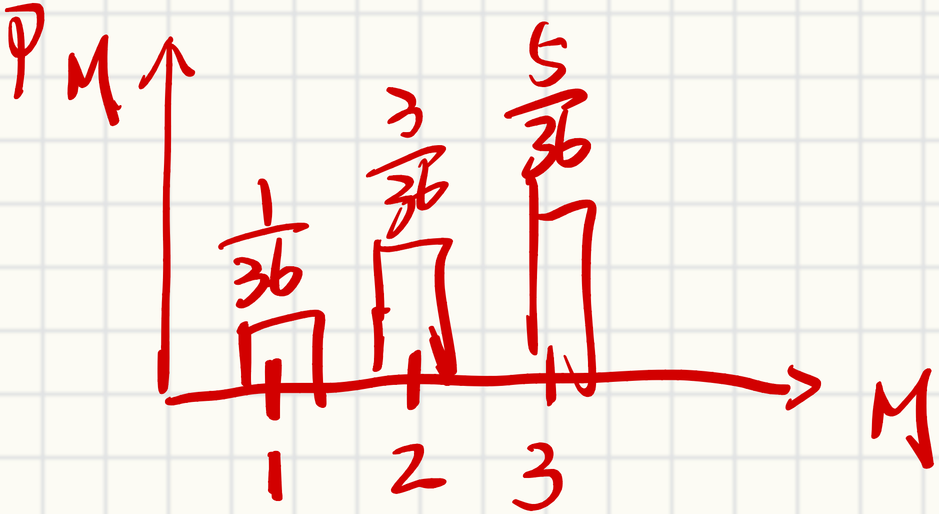
$\Omega_S = \{2 \ldots 12\}$



$1+6$ $2+6$
$2+5$ $\vdots$
$\vdots$ $6+2$
$6+1$

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

$\Omega_M = \{1 \ldots 6\}$



$(2,2)$ $(3,1) \leftrightarrow (1,3)$
$(2,1)$ $(3,2) \leftrightarrow (2,3)$
$(1,2)$ $(3,3)$

Probability Mass Function (PMF)

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

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

$\Omega_N = 1 \ldots \infty$

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

$P_N(2) = \frac{1}{2} \rightarrow H \times \frac{1}{2} \rightarrow T = \frac{1}{4}$

$\sum P_N(k) = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} \ldots = 1$

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

$\Omega_M = 0 \ldots \infty$

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

$P_M(1) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

$N = M+1$

Mean and Variance (Ch 2.2)

Mean

Do we need detailed p_{height}?

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

Expectation ← possible outcome

- X is the number for a die roll $E[X] = \frac{1\times2 + 2\times2 + 3\times2 + 4\times2 + 5\times2 + 6\times2}{6}$ ... $\frac{1+2+3+4+5+6}{6} = 3.5$
- $Y = 2X$ $E[Y] = 7$.
- $Z = |X - 3|$ $\Omega_Z = \{0, 1, 2, 3\}$

$X = 3 \quad 2,4 \quad 1,5 \quad 6$

$E[Z] = 0\times\frac{1}{6} + 1\times\frac{2}{6} + 2\times\frac{2}{6} + 3\times\frac{1}{6} = \frac{9}{6} = 1.5$

Function of RV - LOTUS

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

- $p_X(x) = \begin{cases} \frac{1}{5} & x \in [-1, 3] \\ 0 & \text{else} \end{cases}$

$Y = X^2$

- $\mu_y = E[Y] =$

① $\Omega_{XY} = \{1, 0, 1, 4, 9\}$

$E[Y] = \frac{1+0+1+4+9}{5} = 3$

② $\Omega_Y = \{0, 1, 4, 9\}$

$X = 0, \; -1, 1, \; 2, \; 3$

$0 \times \frac{1}{5} + 1 \times \frac{2}{5} + 4 \times \frac{1}{5} + 9 \times \frac{1}{5}$

P_X

P_Y

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

$$E[g(X)] = \sum_i g(x_i)\, p_X(x_i)$$

Law of the unconscious statistician (LOTUS)

LOTUS examples

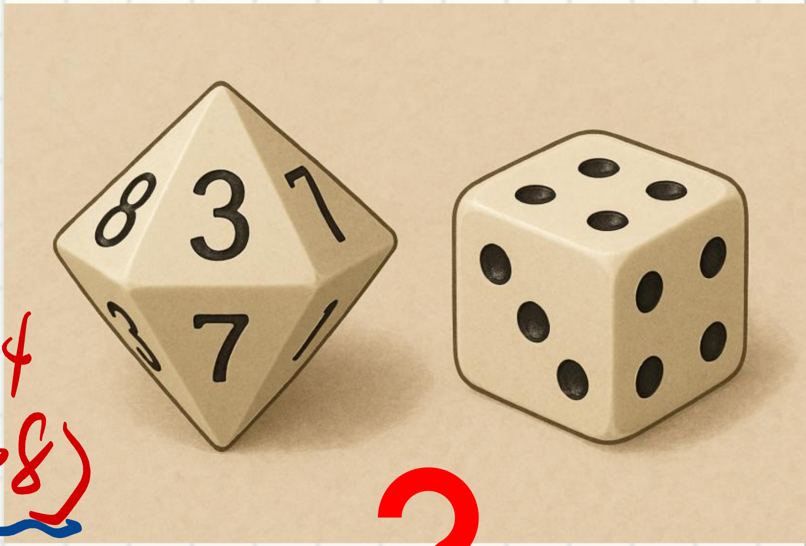
X is rolling a D6, Y is rolling a D8 (8-sided die)

$E[XY] = ?$

$$\sum_i \sum_j \underbrace{x_i}\,\underbrace{y_j}\, \overbrace{\underline{P_{XY}(i,j)}}^{\frac{1}{6}\times\frac{1}{8}}$$

$$= \frac{1}{48}\,\underbrace{\overset{x_i}{(1+2+3+4+5+6)}}_{21}\,\underbrace{\overset{y_j}{(1+2+3+4+5+6+7+8)}}_{36}$$

$$= \frac{21\times \cancel{36}^{3}}{\cancel{48}_{4}} = \frac{63}{4}$$



LOTUS examples

Math magic trick

- Pick a number from 1-10 $\triangleq X$
- Multiply it by 3
- Subtract it by 2
- Divided it by 6 and keep the remainder
- Is it 1 or 4?

$Y = (3X-2)\%6$

$\begin{cases} 1 & X \text{ odd} \\ 4 & X \text{ even.} \end{cases}$

Digit sum

- Pick a two digits number X
- $Y = X -$ sum of digits(X)

Y is $9N$.

$47 - 11 = 36 \% 9 = 0$

$10a+b-(a+b) = 9a$

Variance and Standard Deviation

Do you want your salary to be
$p_X(10K) = 1$ or
$p_Y(0) = 0.99, p_Y(1000K) = 0.01 \leftarrow$ lottery

$\mu_X = \mu_Y = 10K$

Mean is important… but not complete enough

scalar

- Variance $Var(X)$ is how PMF spreads apart from μ_X
- $Var(X) \triangleq E[(X - \mu_x)^2] = E[X^2 - 2\mu_x X + \mu_x^2]$

$= \sum_i (x_i^2 - 2\mu_x x_i + \mu_x^2)\, p_X(x_i)$

$$= \underbrace{\sum_i x_i^2 \, p_X(x_i)}_{} - 2\mu_X \underbrace{\sum_i x_i \, p_X(x_i)}_{E[X]} + \mu_X^2 \cdot 1$$

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

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