Random Variables and Probability Distributions

We can model a sample or process that yields a numeric outcome as a random variable (rv)

Example: Roll a die, get a number between 1-6

We can desribe this discrete random varible using a probability mass function (PMF) that gives a probability for every possible value the rv can have.

X: roll a die

Y: roll a die

x_i	1	2	3	4	5	6
$P(X=x_i)$	1/6	1/6	1/6	1/6	1/6	1/6

Sum 2 dice

$x_i + y_i$	1	2	3	....	10	11	12
$P(X+Y = x_i+y_i)$	0	1/36	2/36	.....	3/36	2/36	1/36

Add up to 1

The Expected Value of X is its theoretical average

consider $\bar{X}$, for a sample of die rolls

$$\bar{X} = \frac{1}{n}(1+6+5+3+\dots+2+1+4+6)$$

$$= \frac{1}{n}(1 \times n_1 + 2 \times n_2 + 3 \times n_3 + \dots + 6 \times n_6)$$

$$= 1 \times \text{proportion of 1s} + 2 \times prop_2 + 3 \times prop_3 + \dots + 6 \times prop_6$$

$$= 1 \times P(X=1) + 2P(X=2) + 3P(X=3) + \dots + 6P(X=6)$$

$$E[X] = \sum_{x_i \in \mathbb{Z}} x_i \times P(X=x_i) = \mu$$

— Greek letter MU μ

$$E[X] = \sum_{x_i \in \mathbb{Z}} x_i P(X = x_i) = \mu \qquad \text{mu} \qquad \mu$$

X is a die roll

$$E[X] = 1 \times P(X=1) + 2P(X=2) + \dots + 6P(X=6)$$
$$= 1 \times \tfrac{1}{6} + 2 \times \tfrac{1}{6} + \dots + 6 \cdot \tfrac{1}{6}$$
$$= 3.5$$

Y is a die roll

$$E[X+Y] = \underset{0}{1 \times \cancel{P(X=1)}} + 2 \times P(X=2) + 3P(X=3) + \dots + 12P(X=12)$$
$$= 2 \times \tfrac{1}{36} + 3 \times \tfrac{2}{36} + \dots + 12 \times \tfrac{1}{36}$$
$$= 7$$

$$E[2 \times X] = 2 \times E[X]$$
$$E[X+Y] = E[X] + E[Y]$$

For a linear combination of random variables

rv

$$E[aX + bY] = aE[X] + bE[Y]$$

constants known

Can't do $X \times Y$ or X/Y or $X^2 \dots$

We can also calculate a theoretical spread: Variance

$$V(X) = \sum_{x_i \in \mathbb{Z}} (x_i - \mu)^2 \times P(X = x_i) = \sigma^2$$

Greek letter sigma σ

The standard dev is the square root of this, σ

$\mu = 3.5$

$$V(X) = (1-3.5)^2 \cdot \tfrac{1}{6} + (2-3.5)^2 \times \tfrac{1}{6} + \dots + (6-3.5)^2 \times \tfrac{1}{6}$$

die roll

$$= 2.91\overline{6}$$

$$V(X+Y) = (2-7)^2 \times \tfrac{1}{36} + (3-7)^2 \tfrac{2}{36} + (4-7)^2 \tfrac{3}{36} + \dots + (12-7)^2 \cdot \tfrac{1}{36}$$

$$V(X+Y) = (2-7)^2 \times 1/36 + (3-7)^2\, 2/36 + (4-7)^2\, 3/36 + \ldots + (12-7)^2\, 1/36$$

$$= 5.8\overline{3}$$

$$V(X-Y) = V(X) + V(Y)$$

$$-Y = (-1 \times Y) \Rightarrow (-1)^2 \cdot Y$$

$$\mathrm{Var}(\underbrace{a}_{\text{known constants}}\overbrace{X}^{\text{r.v.}} + \underbrace{b}_{\text{known constants}}\overbrace{Y}^{\text{r.v.}}) = a^2 V(X) + b^2 V(Y)$$

Only true if X and Y are <u>independent</u>

$$X \perp\!\!\!\perp Y$$