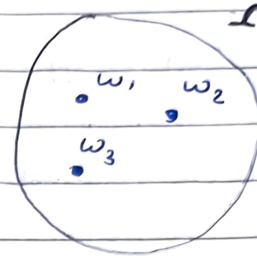
Probability & Random Variables :

Sample Space : Ω



Ω set of outcomes

For each outcome, $\omega \in \Omega$
we assign it a probability

$$P(\omega) \geq 0 \quad \forall \omega \in \Omega$$

$$\sum_{\omega} P(\omega) = 1$$

Event : E any subset of sample space

$$P(E) = \sum_{\omega \in E} P(\omega)$$

Independent events, E_1 & E_2 are independent
if $P(E_1 \cap E_2) = P(E_1) \cdot P(E_2)$

Example : E_1 : Sum of values on two die $= 6$
E_2 : First dice's value is odd

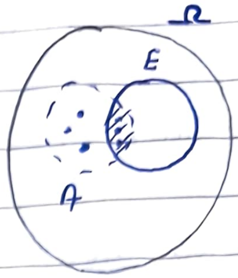
$$P(E_1) = \frac{5}{36} \qquad P(E_2) = \frac{3 \times 6}{36} = \frac{1}{2}$$

$$P(E_1 \cap E_2) = \frac{3}{36}$$

So, $P(E_1 \cap E_2) \neq P(E_1)\, P(E_2)$

E_1 & E_2 are dependent

→ Conditional Probability



Event E has happened

So $P(A|E)$ = Probability of A when E has occured

$$P(A|E) = \sum_{\substack{\omega \in A \\ \omega \notin E}} \frac{P(\omega)}{P(E)} + \sum_{\substack{\omega \in A \\ \omega \in E}} \frac{P(\omega)}{P(E)}$$

$$= 0 + \frac{P(A \cap E)}{P(E)}$$

$$P(A/E) = \frac{P(A \cap E)}{P(E)} = \frac{P(E/A)\,P(A)}{P(E)}$$

if A & E are independent,

$$P\left(\frac{A}{E}\right) = \frac{P(A \cap E)}{P(E)} = \frac{P(A) \cdot P(E)}{P(E)} = P(A)$$

- $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$

- Union Bound :

  $P(A_1 \cup A_2) \leq P(A_1) + P(A_2)$

→ Random Variable

A Random variable is a mapping from Ω to $\mathbb{R}$

set of states to some real number

⇒ **Bayes Rule :**

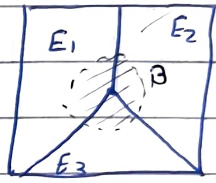
$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

B : Observation
A : Hypothesis

$$= \frac{P(B/A)\ P(A)}{P(B)}$$

A set of events $E_1, E_2 \ldots$ is said to be partition the set of outcomes if

- $E_1 \cap E_2 = \phi$
- $\bigcup E_i = \Omega$



$$Pr[B] = Pr[B \cap E_1] + Pr[B \cap E_2]$$

↪ Law of total probability

⇒ Random variable $X : \Omega \to \mathbb{R}$

Expectation of X $\quad \mathbb{E}X = \sum_i i \cdot Pr[X = i]$

Suppose we conduct the experiment N times & X_i denote the value of X in i^{th} experiment

$$\frac{X_1 + X_2 \cdots X_N}{N} \to \mathbb{E}X \quad \text{as } N \to \infty$$

{ Law of Large Numbers }

Example : Toss N coins

X = number of Heads

$$\mathbb{E}X = \sum_{i=0}^{N} i \Pr[X=i] = \sum_{i=0}^{N} i \binom{n}{i} \left(\frac{1}{2}\right)^{n}$$

$$\mathbb{E}X = \frac{1}{2^{n}} \sum_{i=0}^{N} i \binom{n}{i}$$

$$i \binom{n}{i} = \frac{n!}{i!\,(n-i)!} \times i = \frac{n!}{(i-1)!\,(n-i)!}$$

$$= n \frac{(n-1)!}{(i-1)!\,(n-1-(i-1))!}$$

$$= n \binom{n-1}{i-1}$$

$$\sum_{i=0}^{N} i \binom{n}{i} = \sum_{i=0}^{N} n \binom{n-1}{i-1} = n \sum_{j=0}^{N-1} \binom{n-1}{j}$$

$$= n\, 2^{n-1}$$

$$\mathbb{E}X = \frac{1}{2^{n}} \times n \times 2^{n-1} = \frac{n}{2}$$

We will do it simply later!!

Example: One coin, keep on tossing it till we see a Head

X : # coin tosses

$$\mathbb{E}X = \sum_{i=1}^{\infty} i \times \frac{1}{2^i}$$

$$\mathbb{E}X = \frac{1}{2} + \frac{2}{2^2} + \frac{3}{2^3} \dots$$

$$-\frac{\mathbb{E}X}{2} = -\frac{1}{2^2} - \frac{2}{2^3} \dots$$

$$\frac{\mathbb{E}X}{2} = \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} \dots$$

$$= \left(\frac{1}{1-1/2}\right)\frac{1}{2}$$

$$\frac{\mathbb{E}X}{2} = 1$$

$$\mathbb{E}X = 2$$

Suppose it is a biased coin, $p = \Pr[H]$

then $\mathbb{E}X = \sum i (1-p)^{i-1} p = \frac{1}{p}$

Linearity of Expectation:

$$\mathbb{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbb{E}[X_i]$$

$$X = X_1 + X_2 \dots X_n$$

$$\mathbb{E}X = \mathbb{E}X_1 + \mathbb{E}X_2 \dots \mathbb{E}X_n$$

Examples :

1. n coin tosses

$X =$ no. of heads

Define $X_1, X_2 \dots X_n$

$$X_i = \begin{cases} 1 & \text{if the } i^{th} \text{ coin toss is Heads} \\ 0 & \text{if the } i^{th} \text{ coin toss is Tails} \end{cases}$$

$X = X_1 + X_2 \dots X_n$

$$\mathbb{E} X_i = 1 \, Pr[X_i = 1] + 0 \, Pr[X_i = 0] = \frac{1}{2}$$

$$\mathbb{E}X = \sum_{i=1}^{n} X_i = \frac{1}{2}(n) = \frac{n}{2}$$

2. n envelopes, n letters

X : no. of letters that go in the correct envelope

$X = X_1 + X_2 + X_3 \dots X_n$

where $X_i = \begin{cases} 1 & \text{if } i^{th} \text{ letter is in correct envelope} \\ 0 & \text{otherwise} \end{cases}$

$$\mathbb{E}X_i = 1 * Pr[X_i = 1] = \frac{1}{n}$$

$$\mathbb{E}X = \mathbb{E}X_1 + \mathbb{E}X_2 \dots \mathbb{E}X_n = \frac{1}{n} + \frac{1}{n} + \dots \frac{1}{n} = \frac{n}{n} = 1$$

3. A contains distinct elements (permutation of $\{1, 2 \cdots n\}$)

max = A[0]
for i = 1 to n-1
    if (A[i] > max)
        max = A[i] —— (*)

X = no. of times (*) instruction is executed

$X = X_1 + X_2 \cdots X_n$

where $x_i = \begin{cases} 1 & \text{if } A[i] \text{ is largest from } 1 \text{ to } i, \\ 0, & \text{otherwise} \end{cases}$

$\mathbb{E} X_i = 1 * Pr[X_i = 1] + 0 * [X_i = 0] = 1 * \frac{1}{i}$

$\mathbb{E} X = \mathbb{E} X_1 + \mathbb{E} X_2 \cdots \mathbb{E} X_{n-1}$

$= 1 + \frac{1}{2} + \frac{1}{3} \cdots \frac{1}{n-1}$

$\cong \log_e n$

4. Random walk

Start from 0, toss a coin, if it is head move right, if it is tails move left

X_n : position after n steps

$\mathbb{E} X_n = ?$

$X_n = X_1 + X_2 \dots X_n$

$$X_i = \begin{cases} +1 & \text{if we move right on } i^{th} \text{ step} \\ -1 & \text{if we move left on } i^{th} \text{ step} \end{cases}$$

$\mathbb{E} X_i = +1 \times \frac{1}{2} + -1 \times \frac{1}{2} = 0$

$\therefore \mathbb{E} X_n = 0$

$\mathbb{E}(X^2) = ?$

$X = X_1 + X_2 \dots X_n$

$X^2 = (X_1 + X_2 \dots X_n)^2 = X_1^2 + X_2^2 \dots X_n^2 + 2 \sum_{i \neq j} X_i X_j$

$\mathbb{E} X^2 = \mathbb{E} X_1^2 + \mathbb{E} X_2^2 \dots \mathbb{E} X_n^2 + 2 \sum_{i \neq j} \mathbb{E} X_i X_j$

$= \mathbb{E} X_1^2 + \mathbb{E} X_2^2 \dots \mathbb{E} X_n^2 + 2 \sum_{i \neq j} (\mathbb{E} X_i)(\mathbb{E} X_j)$

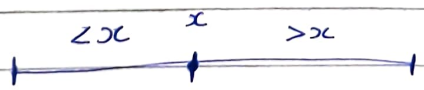
$= 1 + 1 \dots 1 + 0$

$\mathbb{E} X^2 = n$

5) Randomized Quicksort

(Sort an ~~elem~~ array of n elements)

- Pick a random element in the array
- Then divide the array in the two parts



- Recursively sort left & right parts

X = no. of comparisons performed

Let us say that the sorted order is

$a_1 < a_2 < \dots < a_n$

$$X_{ij} = \begin{cases} 1 & \text{if } \cancel{\text{for}}\ a_i\ \&\ a_j \text{ is compared} \\ 0 & \text{otherwise} \end{cases}$$

$$X = \sum_{\text{pairs}(i,j)} X_{ij}$$

$$\mathbb{E}X = \sum_{\text{pairs}(i,j)} \mathbb{E}X_{ij}$$

$$\mathbb{E}X_{ij} = \Pr[X_{ij} = 1] = \Pr[\text{we compare } a_i, a_j]$$

$$= \frac{2}{j-i+1}$$

(we want either of them the pivot element)

$$\mathbb{E}X = 2 \sum_{\text{pairs}(i,j)} \frac{1}{j-i+1} \quad \cancel{\leq 2n}$$

~~let $j-i=d$~~

$$= 2 \sum_{d=1}^{n-1} \sum_{i=1}^{n-d} \frac{1}{d+1}$$

$$= 2 \sum_{d=1}^{n-1} \frac{(n-d)}{(d+1)}$$

$$= 2 \left[\frac{n-1}{2} + \frac{n-2}{3} + \frac{n-3}{4} \dots \frac{1}{n} \right]$$

$$\mathbb{E}X \leq 2n \left(\frac{1}{2} + \frac{1}{3} \dots \frac{1}{n} \right) \Rightarrow \mathbb{E}X \leq 2n \log_e n$$

Markov's Inequality

Let $X \geq 0$ be a non-negative random variable (i.e $X \geq 0$)

$$\Pr(X \geq a) \leq \frac{E[X]}{a}$$ **

Variance $= Var[X] = E(X-\mu)^2$ where $\mu = E[X]$

$$= \sum_i \Pr[X=i] (i-\mu)^2$$

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

- if $E[X] = \mu$, $Var(X) = \sigma^2$
  $E[aX] = a\mu$ $Var(aX) = a^2 Var(X)$

$$E[aX+b] = a\mu + b \quad , \quad Var(aX+b) = a^2 Var(X)$$

Linearity of variance

$$Var(aX+bY) = a^2 Var(X) + b^2 Var(Y) + 2ab\, Cov(X,Y)$$

if $a=b=1$, & (X & Y) independent *

$$Var(X+Y) = Var(X) + Var(Y)$$

$$Cov(X_i, X_j) = E(X_i - \mu_i)(X_j - \mu_j)$$

$= 0$ if X_i, X_j are independent

Date

Let X be a random variable

$$Y = \frac{X_1 + X_2 \cdots X_n}{n}$$ where X_i is an independent copy of X

$$\operatorname{var}(Y) = \frac{1}{n^2}\left[\operatorname{var}(X_1 + X_2 \cdots X_n)\right]$$

$$= \frac{1}{n^2}\left(n \operatorname{Var}(X)\right) = \frac{\operatorname{Var}(X)}{n}$$

→ Chebychev's Inequality

$$\Pr\left[|X-\mu| > t\right] \le \operatorname{Var}(X)/t^2$$

$$\Pr\left[(X-\mu)^2 \geq t^2\right] \le \frac{\operatorname{Var}(X)}{t^2}$$

<u>Example</u>: We toss a coin n times

X : number of heads

$$\mathbb{E}X = \frac{n}{2}$$

$$X = X_1 + X_2 \cdots X_n$$

$$x_i = \begin{cases} 1 & \text{if head} \\ 0 & \text{if tail} \end{cases}$$

$$\operatorname{Var}(X) = \frac{n}{4} \Leftarrow \operatorname{Var}(X_i) = \mathbb{E}(x_i^2) - (\mathbb{E}x_i)^2$$

$$= \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$$

$$\Pr\left[X \geq \frac{3n}{4}\right] \le \frac{n/2}{3n/4} = \frac{2}{3} \quad \left\{\text{by Markov's Inequality}\right\}$$

$$\Pr\left[\left|X - \frac{n}{2}\right| \geq \frac{n}{4}\right] \le \frac{n/4}{(n/4)^2} = \frac{4}{n} \quad \left\{\text{by Chebychev's Inequality}\right\}$$

larger n, lower chances of deviating away from $\mathbb{E}[X]$

Oddy