* Randomized Algorithm

In coin flipping example, the number of Heads in n coin flips is X. Of course, X is an integer between 0 and n, and if the coin is fair.

Question: The homeworks of 20 students are collected in, randomly shuffled and returned to the students. How many students receive their own homework?

To answer this question, we first need to specify the probability space, it should consist of all 20! permutations of the homeworks, each with probability $\frac{1}{20!}$

To make life simple, let's also shrink the class size down to 3 for a while. The following table gives a complete listing of the sample space (of size 3! = 6) together with the corresponding value of X for each sample point.

permutation	Value of X
123	3
132	1
213	1
231	0
312	0
321	1

Scanned by CamScanner

$P(X=3) = \frac{1}{6}$

$P(X=0) = \frac{2}{6} = 1/3$

$P(X=1) = 3/6 = 1/2$

Expectation:

In most applications, however, the complete distribution of a random variable is very hard to calculate. For example, consider the homework example with 20 students. We would have to enumerate $20! = 2.4 \times 10^{18}$ sample points, compute the value of X at each one, and count the number of points at which X takes on each of its possible values!

For these reasons, we seek to compress the distribution into a more compact, convenient form that is also easier to compute. The most widely used such form is the expectation (or mean or average) of the r.v.

The expectation of a discrete random variable is defined as

$$E(X) = \sum a * P(X=a)$$

Where sum is over all possible values taken by r.v.

For homework example (previously dissussed), the expectation is

$$E(X) = \left(0 \times \frac{1}{3} + 1 \times \frac{1}{2} + 3 \times \frac{1}{6}\right)$$

$$= 1$$

Eg. 1 Throw one fair dice. Let X be the number that comes up. Then X takes on values 1, 2, ..., 6 each with probability $\frac{1}{6}$ so.

$$E(X) = \frac{1}{6}(1+2+3+4+5+6) = \frac{21}{6} = \frac{7}{2}$$

② Two dice

x_i	2	3	4	5	6	7	8	9	10	11	12
$p(x_i)$	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{9}$	$\frac{5}{36}$	$\frac{1}{6}$	$\frac{5}{36}$	$\frac{1}{9}$	$\frac{1}{12}$	$\frac{1}{18}$	$1/36$

The expectation is therefore

$$E(X) = \sum_{i=1}^{12} x_i \times p(X = x_i)$$

$$= 7$$

③ A roulette wheel is spun. You bet Rs 1 on black. If a black number comes up, you receive your stake plus Rs 1, otherwise you lose your stake.

Let X be your net winnings

in one game. The X can take values $+1$ and -1, and $Pr[X=1] = \frac{18}{38}$

$Pr[X=-1] = \frac{20}{38}$ [Note: a roulette wheel has 38 slots, the numbers 1, 2, ..., 36 half of which are red & half black, plus 0 and 00, which are green]

$$\therefore E(X) = 1 \times \frac{18}{38} + -1 \times \frac{20}{38}$$

$$= -\frac{1}{19}.$$

$\frac{2}{36}$

So far, we have computed expectations by brute force, ie we have written down the whole distribution & then added up the contributions for all possible values of r.v. The real power of expectations is that in many real life examples, they can be computed much more easily using a simple shortcut.

$$① \; E(X+Y) = E(X) + E(Y)$$

Expectation of a sum of random variables is the sum of their expectations, with no assumptions about the r.v.'s.

Page No
Date

② Also for any constant c, we have

$$E(cX) = c\,E(X)$$

* An Indicator random variable is a variable that indicates whether an event is happening. If A is an event, then indicator random variable I_A is defined as

$$I_A = \begin{cases} 1 & \text{if A occurs} \\ 0 & \text{if A does not occur} \end{cases}$$

eg

Flipping a coin n times.

$X_i \rightarrow$ Indicator Random variable

$$X_i = \begin{cases} 1 & \text{if } i^{th} \text{ coin flip turns in head} \\ 0 & \text{if } i^{th} \text{ coin flip turns in tails} \end{cases}$$

By summing the values of X_i, we can get the total number of heads across the n coin flips.

$$E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E(X_i)$$. linearity of expectation.

Expectation of n coin flips.

sum of expectation of each coin flips.

$$E(X_i) = 1 \cdot p(X_i = 1) + 0 \cdot p(X_i = 0)$$

$$= 1 \cdot \frac{1}{2} + 0 \cdot \frac{1}{2}$$

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

So the expected no. of heads is

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

There are two possibilities: either A occurs, in which case $I_A = 1$ or A does not occur, in which case $I_A = 0$.

$$E(I_A) = 1 \cdot p(A) + 0 \cdot p(\text{not } A)$$

$$= p(A).$$

of

m

→ Two dice rolled.

X = sum of scores of the two dice.

X_i score on dice i.

$$X = X_1 + X_2$$

$$E(X) = E(X_1) + E(X_2)$$

$$= \frac{7}{2} + \frac{7}{2}$$

$$= \underline{\underline{7}}$$

→ Play roulette game 1000 times.

Let X be net winnigs.

$$X = X_1 + X_2 + \cdots + X_{1000}$$

where X_i = net winning in i^{th} game play

$$E(X) = E(X_1) + \cdots + E(X_{1000})$$

$$= -\frac{1}{19} + \qquad + -\frac{1}{19}$$

$$= -\frac{1000}{19} \approx -53$$

So if you play 1000 games, you expect to lose abot Rs. 53.

→ Homework Problem

class of 20 students.

The r.v. X is the number of students who receive their own homework after shuffling.

$$X = X_1 + \cdots + X_{20}$$

$$X_i = \begin{cases} 1 & \text{if student gets own assignment} \\ 0 & \text{if student not get " } \end{cases}$$

Indicator Random variable

$$E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i)$$

=

$$E(X_i) = 1 \cdot P(X_i = 1) + 0 \cdot P(X_i = 0)$$

$$= P(X_i = 1)$$

$$= P(\text{student } i \text{ gets own assg})$$

$$= \frac{1}{20}.$$

$$\therefore E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{20} \frac{1}{20} = 20/20 = 1.$$

So, we see that the expected number of students who gets their own assignments in a class of 20 is 1.

But this is exactly the same answer as we got for a class of size 3.

And indeed, we can easily see from the above calculations that we would get $E(X) = 1$ for any class size n.

- Consider the problem of hiring an office assistant. We interview candidates on a rolling basis and at any given point we want to hire the best candidate we have seen so far. If a better candidate comes along, we immediately fire the old one & hire the new one

HIRE - ASSISTANT(n)

```
best = 0
for i = 1 to n
    interview candidate i
    if candidate i is better than
       candidate best
        best = i
        hire candidate i
```

C_i = cost associated with interviewing people.

C_h = cost of hiring

$C_h >> C_i$ since old is fired & new one is absorbed.

assume $\underline{n} \rightarrow$ total no. of candiates

$m \Rightarrow$ candidate got selected (hired)

$\therefore$ total cost $= (C_i * n + C_h * m)$

What is the worst cost of this algo?

In worst case scenario, the candidates come in order of increasing quality, & we hire every person we interview. Then the hiring cost is $O(C_h n)$

$\therefore$ total cost $O(C_i n + C_h n)$

$= O(n(C_i + C_h))$

Analysis of HireAssistant using I.R.V.

Let X = No. of times we hire a new candidate

$X = X_1 + \dots + X_n$

where $X_i = \begin{cases} 1 & \text{if the candidate is hired} \\ 0 & \text{if not hired.} \end{cases}$

$E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i)$

$E(X_i) = 1 \cdot P(X_i = 1) + 0 \cdot P(X_i = 0)$

$= P(X_i = 1)$

P(Candidate i is hired)

Candidate i is hired when ith candidate is better than all the candidates 1 through i−1.

$$\therefore P(X_i = 1) = \frac{1}{i}$$

$$E(X_i) = \frac{1}{i}$$

$$E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \frac{1}{i}$$

$$= \log n + O(1)$$

$$= O(\log n).$$

The expected no. of candidates hired $= O(\log n)$.

∴ The hiring cost is $O(\log n \cdot C_h)$