

Foundational set of experiments in probability and statistics involve coin toss experiments; a single toss of a coin has two possibilities: (1) head with probability of  $p$ , (2) tail with probability of  $q = 1 - p$ . (Note: Fair coin has  $p = \frac{1}{2} = q$ ; we will not assume this in developing theory below.)

If a coin is tossed a total of  $n$  times (or if  $n$  identical coins are tossed), the probability of seeing a total of  $M$  heads (obviously  $M \leq n$ , and hence  $n - M$  tails) is given by:

$$P(M; n) = \frac{n!}{M!(n-M)!} p^M q^{n-M} \quad (1)$$

If the fraction of heads in  $n$  tosses is denoted by  $M_n = M/n$ , then clearly  $\lim_{n \rightarrow \infty} M_n = p$ . Hence the the probability that  $M_n$  falls in range  $[x, y]$  (with  $0 \leq x \leq y \leq 1$ ) is given by:

$$Prob(x \leq M_n \leq y) = \sum_{\frac{r}{n} \in [x, y]} P(r; n) = \sum_{\substack{r \in \mathbb{Z} \\ nx \leq r \leq ny}} P(r; n) \quad (2)$$

For a fixed number of tosses  $n$ , we can show that (a)  $P(M; n)$  has a single maxima at  $M = pN$ , i.e. at  $M_n = p$ , and (b) around this maxima,  $prob(M_n)$  can be approximated by a Gaussian function. [Hint: Use  $\frac{d}{dM} \ln P(M; n) = 0$  to find the location of extrema in  $\ln P(M; n)$ ; use Sterling Approximation for  $N \gg 1$ ,  $\ln N! = N \ln N - N$ ].

One popular model that can be straightforwardly mapped to coin toss problem is the Random Walk problem: a drunk walker starts at  $x = 0$  and takes a step of  $\delta x = +1$  if the coin toss is head and a step  $\delta x = -1$  if the coin toss is tails. i.e.  $x(t+1) = x(t) + \delta_{C(t), H} - \delta_{C(t), T}$ , where  $C(t)$  denotes the result of coin toss at time  $t$ , and  $\delta$  is the Kronecker delta. Various properties of the trajectory of walker as a function of number of steps is common homework assignments.

An interesting property of any stochastic process (here, random walk problem) is the *first passage* or *first hit* statistics. One famous example is the problem of ‘gambler’s ruin’: the gambler starts with Rs. 10/- with Rs. 1/- per bet (i.e. gain Rs.1 with probability  $p$  or lose Rs 1/- with probability  $q = 1 - p$ ), and stops (‘gambler’s ruin’) when he is left with no money i.e. Rs 0/-. Obviously, he meets the ‘gambler’s ruin’ the first time it happens and is stuck there forever! Note that the first passage time will be different for each realization, and hence is ‘sampled’ from a distribution. You will be numerically calculating this for a random walk problem, and show that the distribution is Poisson Distribution, i.e. probability of hitting time  $t$  is given by  $P(t) \sim e^{-kt}$  where  $k$  is a constant that depends on the threshold (in the example ‘gambler’s ruin’ problem given above, the threshold was Rs. 10/-). Note that this fact about Stochastic Processes (first passage time being a Poisson distributed variable) is known as ‘Cramer’s Theorem’ in Probability and Statistics.