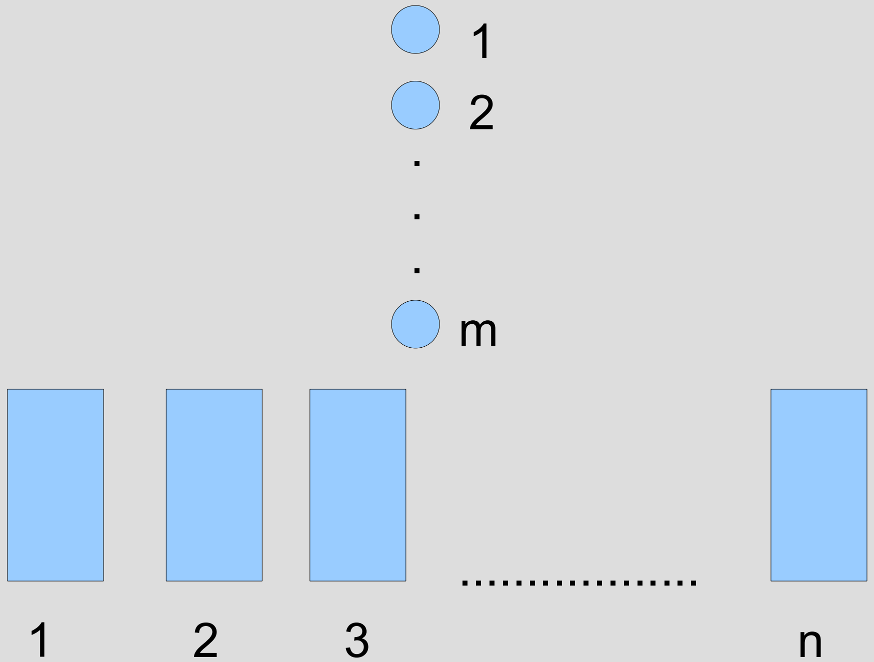
Tail Probability

The probability that a random variable deviates by a given amount from its expectation is referred to as a tail probability.

Occupancy Problem

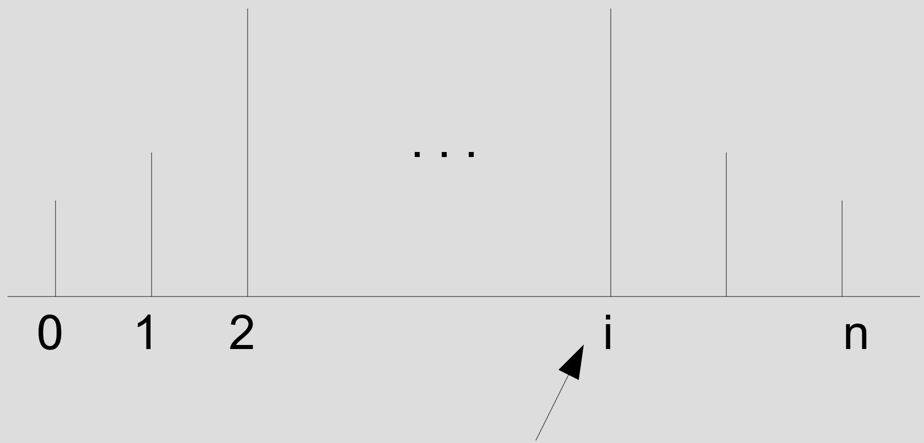
It is a problem of assigning m indistinguishable objects(“balls”) to n distinct classes(“bins”).



One ball is thrown at a time with equal probability of landing into any bin.

Probability that no. of balls in bin 3 be i

It can be found out by finding the probability of bin 3 being selected i times.



The probability distribution is a Binomial Distribution. So,

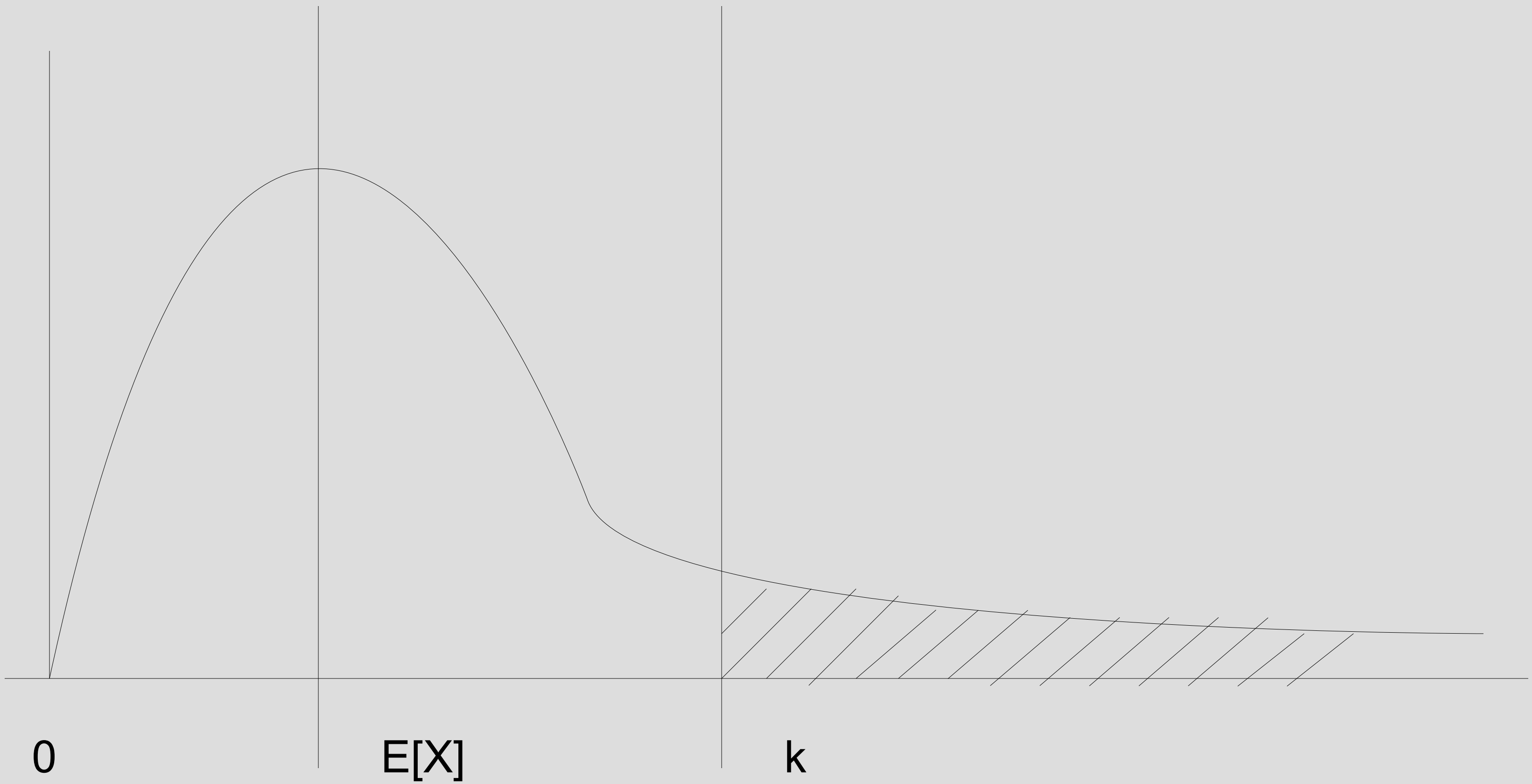
$$\mathbf{P_i = {}^nC_i(1/n)^i(1-1/n)^{n-i}}$$

Expected no. of Balls and SD

Expected Number of balls in each bin E[X] is given by **np**, where n is the number of bins and p=1/n is the probability of any bin getting selected at each iteration.

Standard Deviation is given by sqrt(n.1/n.(n-1)/n) = **sqrt((n-1)/n) < 1**

Probability that a bin has k or more balls



$$\sum_{i=k}^{n} nCi(\frac{1}{n})^i(1-\frac{1}{n})^{n-i}$$

$$\leq \sum_{i=k}^{n} nCi(\frac{ne}{i})^i(\frac{1}{n})^i, Since \ldots nCi \leq (\frac{ne}{i})^i$$

$$\leq \sum_{i=k}^{n} \frac{e^i}{i}$$

$$= (\frac{e}{k})^k + (\frac{e}{k+1})^{k+1} + (\frac{e}{k+2})^{k+2} + \ldots$$

$$\leq (\frac{e}{k})^k + (\frac{e}{k})^{k+1} + (\frac{e}{k})^{k+2} + \ldots$$

$$= (\frac{e}{k})^k[1 + \frac{e}{k} + (\frac{e}{k})^2 + \ldots]$$

Telescopic

$$= (\frac{e}{k})^k \frac{1}{1-\frac{e}{k}}$$

Tail Probability
i>=k

Value of k for which Tail Probability is small

When n=n/2, Tail Probability is

$$(\frac{e}{\frac{n}{2}})^{\frac{n}{2}} \cdot \frac{1}{1-\frac{2e}{n}} \approx \frac{2e}{n}^{\frac{n}{2}}$$

When n=sqrt(n), Tail Probability is

$$(\frac{e}{\sqrt{n}})^{\sqrt{n}} \cdot \frac{1}{1-\frac{e}{\sqrt{n}}} \approx (\frac{e}{\sqrt{n}})^{\sqrt{n}}$$

When k=logn or more Tail Probability is

$$\frac{e}{logn}^{logn} = (\frac{1}{\frac{logn}{2}})^{logn} \leq (\frac{1}{2})^{logn} = \frac{1}{n}$$