Tutorial 6

08 April 2021 16:02

Basic methods: Moments and deviations

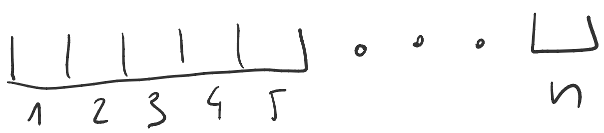
→ occupancy problem

→ drunken sailor problem

→ coupon collector's problem

→ ...

Occupancy problem



n-bins

m balls put into bins at random (uniformly) & (potentionally infinite)

Geometric distribution

→ What is the expected number of balls you need to place before 1 of them lands in bin 1

$X = i$ if ith ball is the first one to land in bin 1

$$\Pr(X=1) = 1/n$$

$$\Pr(X=2) = \left(\frac{n-1}{n}\right) \cdot \frac{1}{n}$$

$$\Pr(X=m) = \left(\frac{n-1}{n}\right)^{m-1} \cdot \frac{1}{n}$$

$E(X) = n$ (Probability of success $p = 1/n$ and $E(X) = \frac{1}{p}$)

$E(X) = n$ (Probability of success $p = \frac{1}{n}$ and $E(X) = \frac{1}{p}$)

$$E(X) = \sum_{i=1}^{\infty} Pr(X=i) \cdot i$$

Q: What is the expected number of empty bins if m balls were placed? (Drunken sailors problem)

$X_i = 1$ if i^{th} bin is empty

$X_i = 0$ otherwise

$$Pr(X_i = 1) = \left(\frac{n-1}{n}\right)^m$$

$$E(X_i) = Pr(X_i = 1) = \left(\frac{n-1}{n}\right)^m$$

$$X = \sum_{i=1}^{n} X_i$$ (X is the number of empty bins in a trial)

$$E(X) = E\left(\sum_i X_i\right) = \sum_i E(X_i) = n \cdot \left(\frac{n-1}{n}\right)^m = \frac{(n-1)^m}{n^{m-1}}$$

Q: How many balls do I expect to drop in order to fill all the bins?

0 bins are occupied

$\Rightarrow$ Prob to fill an empty bin with next ball $= 1$ | exp 1

1 bin is occupied

1 bin is occupied

$\Rightarrow Pr(\text{new bin is filled with next ball}) = \frac{n-1}{n}$ — era 2

2 bin occupied

$\Rightarrow Pr(\text{new}) = \frac{n-2}{n}$ — era 3

⋮

n-1 bins occupied $Pr(\text{new}) = \frac{1}{n}$ — era n

For ith era define $X_i \to$ expected number of balls to drop in order to get to new era

X_i – geometric distribution with success probability $p_i = \frac{n-i+1}{n}$

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

$X = \sum_{i=1}^{n} X_i$ is the total expected number of balls.

$$E(X) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} \frac{n}{n-i+1} = n \cdot \sum_{i=1}^{n} \frac{1}{n-i+1}$$

(reverse order of summands) $$= n \cdot \sum_{i=1}^{n} \frac{1}{i} = n \cdot H_n$$ Harmonic sum of n elements

$$\approx n \cdot \log n$$

Q: Expected number of bins with 2 or more balls when n balls were dropped.

$Pr(j^{th}$ bin has exactly i balls$)$

= binomial distribution = $\binom{n}{i}\left(\frac{1}{n}\right)^i\left(1-\frac{1}{n}\right)^{n-i}$

choose i of n balls

i balls fall into j^{th} bin

$n-i$ balls fall elsewhere

$$\leq \binom{n}{i}\left(\frac{1}{n}\right)^i \qquad \left[\binom{n}{i} \leq \left(\frac{ne}{i}\right)^i\right]$$

basis of natural logarithm

$$\leq \left(\frac{ne}{i}\right)^i\left(\frac{1}{n}\right)^i = \left(\frac{e}{i}\right)^i$$

$\mathcal{E}_j(k)$ = event that j^{th} bin has k or more balls

$$Pr(\mathcal{E}_j(k)) = \sum_{i=k}^{n}\binom{n}{i}\left(\frac{1}{n}\right)^i\left(\frac{n-1}{n}\right)^{n-i}$$

$$\leq \sum_{i=k}^{n}\left(\frac{e}{i}\right)^i = \left(\frac{e}{k}\right)^k + \left(\frac{e}{k+1}\right)^{k+1} + \dots + \left(\frac{e}{n}\right)^n$$

$$\leq \left(\frac{e}{k}\right)^k + \left(\frac{e}{k}\right)^{k+1} + \dots + \left(\frac{e}{k}\right)^n$$

$$= \left(\frac{e}{k}\right)^k \boxed{\sum^{n} \left(\frac{e}{\,}\right)^i}$$

$a<1 \quad \sum^{\infty} a^i = \frac{1}{\,}$

$$= \left(\frac{e}{k}\right)^k \boxed{\sum_{i=0}^{n} \left(\frac{e}{k}\right)^i}$$

$$\boxed{a < 1 \quad \sum_{i=0}^{\infty} a^i = \frac{1}{1-a}}$$

for $n \to \infty$

$\left(\frac{e}{k}\right) < 1$

$(k \geq 3)$ $\quad = \left(\frac{e}{k}\right)^k \cdot \frac{1}{1 - \frac{e}{k}}$

for $\boxed{k} = \left\lceil \frac{e \cdot \ln(n)}{\ln(\ln(n))} \right\rceil$

$$\Pr \varepsilon_j(k) \leq \frac{1}{n^3}$$

$X_j = 1$ if j^{th} bin has k or more balls

$X_j = 0$ otherwise

$X = \sum_j X_j$ (X is the expected number of bins with $\boxed{k}$ or more balls)

$$E(X) = \sum_j E(X_j) = \sum_j \Pr \varepsilon_j(k) \leq \frac{1}{n}$$

Q: What is the probability that at least one bin has k or more balls in it?

$$\Pr\Big[\bigcup_j \mathcal{E}_j(\varepsilon)\Big] \le \sum_i \Pr \mathcal{E}_j(\varepsilon)$$

Mealy's inequality

(equality in case events are mutually exclusive)

$$\le 1/n$$