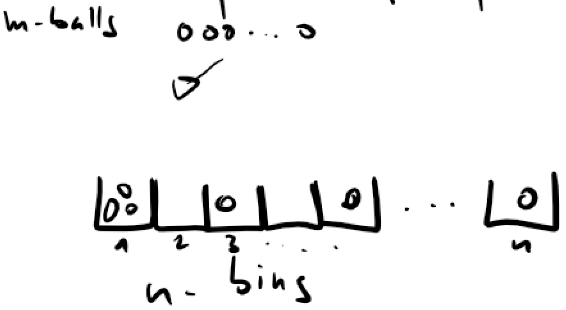
Tutorial VI

26 March 2020 14:07

Basic Methods: Moments and Deviations

→ Occupancy problem (bins and balls)

→ Coupon collector problem



Geometric distribution

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

X Geometric

$=$

1 Probability that 1st ball lands in bin 1 $= \frac{1}{n}$

2 Probability that the 2^{nd} ball is the first one to land in box 1 $\left(\frac{n-1}{n}\right)\cdot\frac{1}{n}$

⋮

m Prob —— // —— $\left(\frac{n-1}{n}\right)^{m-1}\cdot\frac{1}{n}$

$E(X) = \frac{1}{p}$ if the probability of success is p.

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

$$= \sum_{i=1}^{\infty} i \cdot (1-p)^{i-1} \cdot p$$

$$= p \cdot \sum i \cdot (1-p)^{i-1}$$

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

$| \cdot (1-p)$

$$= p \cdot \sum_{i=1}^{\infty} i \cdot (1-p)^{i-1} \qquad 1 \cdot (1-p)$$

$$(1-p)\,E(X) = p \cdot \sum_{i=1}^{\infty} i \cdot (1-p)^{i}$$

$$(1-(1-p))\,E(X) = p \cdot \sum_{i=1}^{\infty} i \cdot (1-p)^{i-1} - p \cdot \sum_{i=1}^{\infty} i\,(1-p)^{i}$$

$$= p \cdot \Big(1 \cdot (1-p)^0 \boxed{- (1-p) + 2 \cdot (1-p)} - 2 \cdot (1-p)^2$$

$$\boxed{+ 3 \cdot (1-p)^2} - 3 \cdot (1-p)^3$$

$$= p \cdot \big[1 + (1-p) + (1-p)^2 + \ldots +$$

$$= p \cdot \sum_{i=0}^{\infty} (1-p)^i$$

$$= \frac{1}{1-(1-p)} \cdot p$$

$$p \cdot E(X) = p \cdot \frac{1}{p}$$

$$E(X) = \frac{1}{p}$$

Q2: What is the expected number of empty bins?

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

$X_i = 0$ otherwise

$$\boxed{Pr(X_i = 1) = \left(\frac{n-1}{n}\right)^m} \qquad \left(E(X_i) = \left(\frac{n-1}{n}\right)^m \cdot 1 + \left(1 - \left(\frac{n-1}{n}\right)^m\right) \cdot 0 \right)$$

$X = \sum_i X_i$ ⇝ number of empty bins

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

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

DRUNKEN SAILOR PROBLEM ↑

Coupon collector problem

I have 0 coupons ⇝ Probability to obtain a new one is = 1

I have 1 coupons ⇝ $\Pr_{(2nd)}(new) = \frac{n-1}{n}$

2 coupons ⇝ $\Pr_{3rd}(new) = \frac{n-2}{n}$

i-1 coupons ⇝ $\boxed{\Pr_{ith}(new)} = \frac{n-i+1}{n}$

these are called "Evas"

X_i ⇝ expected number of tries to get ith coupon

→ X_i is geometric distribution

p ⇝ $E(X_i) = \frac{1}{p_i}$, where $p_i = \frac{n-i+1}{n}$

$$X = \sum_{i=1}^{n} X_i$$

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

Harmonic sum of n elements

$$= n \cdot \sum_{i=1}^{n} \frac{1}{i} = n \cdot H_n$$

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

Q: Expected number of bins with k or more balls

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

n-balls and n-bins

$$\Pr(j^{th} \text{ bin has } \underline{\text{exactly}}\ i \text{ balls}) = \underbrace{\binom{n}{i}}_{\text{choose } i \text{ balls}} \underbrace{\left(\frac{1}{n}\right)^i}_{i \text{ balls fall to } j^{th} \text{ bin}} \underbrace{\left(1-\frac{1}{n}\right)^{n-i}}_{n-i \text{ balls fall somewhere else}}$$

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

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

$$\leq \left(\frac{e}{i}\right)^i$$

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

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

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

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

if $n \to \infty$ $\qquad \left(\frac{e}{k}\right)^k \sum^{\infty} \left(\frac{e}{k}\right)^i$

if $n \to \infty$

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

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

$$\leq \left(\frac{e}{k}\right)^k \cdot \frac{1}{1-\frac{e}{k}}$$

for $k = \left\lceil \frac{e \cdot \ln(n)}{\ln(\ln n)} \right\rceil \leq \frac{1}{n^2}$

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

$X_j := 0$ otherwise

$$X = \sum_i X_i$$

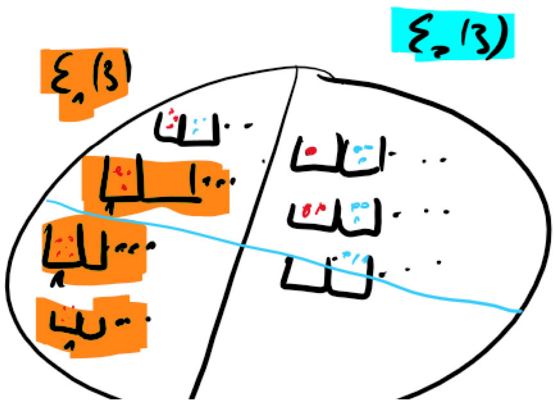
$$E(X) = \sum_i E(X_i) = n\,E(X_i) \leq \frac{1}{n}$$

Q3: What is the probability that at least one box has k or more balls in it.

$$\Pr\left[\bigcup_j \mathcal{E}_j(k)\right] \leq \sum_i \Pr \mathcal{E}_j(k)$$

Meal's inequality

$$\leq \frac{1}{}$$



$\leq \frac{1}{4}$