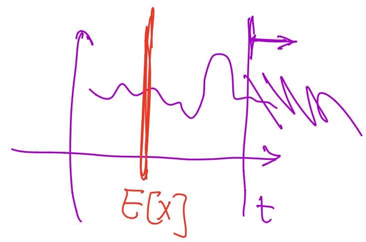
Markov inequality: Consider <u>any</u> Prob. distribution with expected Value $E[X]$

$$P(X \geq t) \leq \frac{E[X]}{t}$$



Proof:

$$f(x) = \begin{cases} 1 & \text{if } x \geq t \text{; i.e. } x/t \geq 1 \\ 0 & \text{otherwise} \end{cases}$$

$$P(X \geq t) = P(f(x) = 1) = E[f(x)]$$

$$\leq E[X/t]$$

$$= \frac{E[X]}{t}$$

$$\boxed{X/t \geq f(t)}$$

Example: R-Q-Sort:

what is the prob. of runtime $\geq \frac{n^2}{c}$

using Markov ineq. ?

Chebyshev inequality

→ Assume we know the expected value μ and variance σ^2

$$\boxed{P(|X-\mu| \geq t\sigma) \leq 1/t^2}$$

Proof:

$$P(|X-\mu| \geq t\sigma) = P((X-\mu)^2 \geq t^2\sigma^2)$$

$$Y = (X-\mu)^2$$

$$\sigma^2 = E[(X-\mu)^2] = E[Y]$$

using Markov. Ineq.

$$P(Y \geq t^2 E[Y]) \leq \frac{E[Y]}{t^2 E[Y]} = 1/t^2 \checkmark$$

Stable Marriage Problem.

given n men, n women each with a pref. list of size n

find a matching b/w men–women

s.t. $\nexists$ X–Y, x–y : x prefers Y over y

Y prefers x over X

Solution:

Rules :- Men Propose
women Accept
- women can switch to better proposal

- Men start from the top of their list, going down, propose to next available woman
- The algorithm stops when there is no more proposals

$$\boxed{\text{worst Case Analysis: } O(n^2)}$$

The Runtime of Stable Marriage in Average?

Observation:

The Algorithm stops when every woman has received proposal from at least one man.

Assumption: The pref. list of men is uniform random
(But for women, it can be arbitrary)

Small change to the Algorithm:

A man will propose to a woman following the rand. uniform distribution.

→ He can propose to the same woman multiple times

⇒ we will find an upperbound on the Avg. runtime of original Alg.

→ following the observation, the algorithm stops when number for all women has been generated

→ This is an application of the Coupon Collector's Problem

→ $O(n\log n)$

Coupon Collector's problem

- Assume there are $\underline{n}$ Coupons $(C_1, C_2, \dots, C_n)$
- Everytime, we generate a random uniform number (independently) between $\underline{1}$ to $\underline{n}$. That is, every time we collect a random Coupon.

e.g. Coupons are $\boxed{1\,|\,2\,|\,3}$

2 2 1 | 2 1 1 2 3 |

— x — | x x x x — |

$\underbrace{\qquad\qquad\qquad}_{\text{epoch}_i}$

epoch$_i$: all trials after collecting $(i-1)$ Coupons, until the i-th Coupon has been Collected

$\Rightarrow$ total $\overbrace{\#\text{ trials}}^{T}$ = Sum of trials in each epoch

Let X_i be the Number of Trials in epoch$_i$

$$\Rightarrow T = \sum_{i=1}^{n} X_i$$

Let P_i be the Prob. that $\underline{a}$ Trial at epoch$_i$ is Successful

$$\Rightarrow P_i = \frac{n-i}{n}$$

X_i follows Geometric Distribution

$$E[X_i] = 1/P_i \qquad \sigma^2[X_i] = \frac{1-P_i}{P_i^2}$$

$$E[T] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} \frac{n}{n-i} = n\sum_{i=1}^{n}\frac{1}{n-i}$$

$$= n\sum_{i=1}^{n} 1/i = n H_n \simeq \boxed{n \ln(n)}$$

* The expected number of Trials to collect all Coupons is in $O(n \log n)$

using Markov inequality,

$$P(T \geq \beta n \ln(n)) \leq \boxed{1/\beta}$$ Markov

$$Var[T] = \sigma^2 = \sum_{i=1}^{n} \sigma_i^2 = \sum \frac{i/n}{\frac{(n-i)^2}{n^2}}$$

$$= \sum \frac{ni}{(n-i)^2} = n \sum_{i=1}^{n} \frac{i}{(n-i)^2}$$

$$= n \sum_{i=1}^{n} \frac{(n-i)}{i^2} = n^2 \sum_{i=1}^{n} 1/i^2 - n \sum_{i=1}^{n} 1/i$$

$$= n^2 \sum 1/i^2 - \cancel{\underbrace{n H_n}_{O(n \log n)}}$$

$$\leq n^2 \sum 1/i^2 = \boxed{C n^2}$$

$$C \approx \pi^2/6$$

Using Chebyshev Ineq., what is the prob.

$$P(T \geq \beta n \ln(n))$$

$$\boxed{\text{Chebyshev}: P(|x-\mu| \geq t\sigma) \leq 1/t^2}$$

$\mu = n \ln(n)$

$\sigma = \sqrt{c}\, n$

$$P(T \geq \beta n \ln(n)) = P(|T-\mu| \geq (\beta-1) n \ln(n))$$

$$t\sigma = t\sqrt{c}\, n = (\beta-1) n \ln(n)$$

$$\Rightarrow t = \frac{(\beta-1)\ln(n)}{\sqrt{c}}$$

$$P(T \geq \beta n \ln(n)) \leq 1/t^2 = \boxed{\frac{c}{(\beta-1)^2 \ln^2(n)}}$$

✓ Chebyshev