CS561, HW2

Ryan Scherbarth, University of New Mexico

August 2024

1. Let X be a random variable that is equal to the number of heads in two flips of a fair coin.

(a) What is $E(X^2)$

$$E[X^2] = (0)^2 P(X = 0) + (1)^2 P(X = 1) + (2)^2 P(X = 2)$$
$$E[X^2] = \frac{3}{2}$$

$$\boxed{E(X^2) = \frac{3}{2}}$$

(b) What is $E(X)^2$

$$X = 0, 1, 2$$
$$P(X = 0, 2) = \frac{1}{4}$$
$$P(X = 1) = \frac{1}{2}$$
$$E[X] = (0)P(X = 0) + (1)P(X = 1) + (2)P(X = 2)$$
$$E[X] = 1$$

$$\boxed{(E[X])^2 = 1}$$

2. 7-3 Alternative quicksort analysis

(a)

Argue that, given an array of size n, the probability that any particular element is chosen as the pivot is $\frac{1}{n}$. Use this probability to define indicator random variables $X_i = I$ (The ith smallest element is chosen as the pivot). What is $E[X_i]$?

We can define the indicator random variable X_i as 1 if x is the ith smallest element, and 0 otherwise.
An array of randomly sorted elements is a uniform distribution. By the definition of a uniform distribution, every element has an equally probably chance of being chosen.
Therefore, we can calculate the expected value via:

$$E[X_i] = (1)(1 - p(X = x)) + (0)(1 - p(X = x))$$
$$E[X_i] = p(X = x)$$

Since $p(X = x) = \frac{1}{n}$,

$$\boxed{E[X_i] = \frac{1}{n}}$$

(b)

Let $T(n)$ be a random variable denoting the running time of quicksort on an array of size n. Argue that
$E[T(n)] = E[\sum_{q=1}^{n} X_q(T(q-1) + T(n-q) + \Theta(n))]$

Let X_q be an indicator random variable that is 1 if the q-th smallest element is the pivot, and 0 otherwise. Given a uniform distribution, the probability of any given element being the pivot is equally likely, therefore $P(X = x) = \frac{1}{n}$

$$E[T(n)] = \sum_{q=1}^{n} \frac{1}{n}(T(q-1) + T(n-q) + \Theta(n))$$

Where $T(q-1)$ represents running time of the sub array of the first $q-1$ elements. $T(n-1)$ represents the subarray of the last n-q elements, and $\Theta(n)$ represents the time taken for the partitioning step itself.
Then, via the linearity of expectation;

$$\boxed{E[T(n)] = E[\sum_{q=1}^{n} X_q(T(q-1) + T(n-q) + \Theta(n))}$$

(c)

Show how to rewrite equation
$E[T(n)] = E[\sum_{q=1}^{n} X_q(T(q-1) + T(n-q) + \Theta(n))$
as
$E[T(n)] = \frac{2}{n}\sum_{q=1}^{n-1} E[T(q)] + \Theta(n)$

$$E[T(n)] = E[\sum_{q=1}^{n} X_q(T(q-1) + T(n-q) + \Theta(n))$$

$$E[T(n)] = \sum_{q=1}^{n} X_q(E[T(q-1)] + E[T(n-q)] + \Theta(n))$$

$$E[T(n)] = X_q\sum_{q=1}^{n} E[T(q-1)] + X_q\sum_{q=1}^{n} E[T(n-q)] + X_q\sum_{q=1}^{n}\Theta(n)$$

$$E[T(n)] = 2X_q\sum_{q=1}^{n-1} E[T(q)] + \Theta(n) \qquad T(q-1) \text{ and } T(n-q) \text{ are sym.}$$

$$E[T(n)] = \frac{2}{n}\sum_{q=1}^{n-1} E[T(q)] + \Theta(n) \qquad X_q = \frac{1}{n}$$

$$\boxed{E[T(n)] = \frac{2}{n}\sum_{q=1}^{n-1} E[T(q)] + \Theta(n)}$$

(d) Show that $\sum_{q=1}^{n-1} qlg(q) \leq \frac{n^2}{2}lg(n) - \frac{n^2}{8}$ for $n \geq 2$

$$\sum_{q=1}^{n-1} q\log(q) = \sum_{q=1}^{\frac{n}{2}-1} q\log(q) + \sum_{q=\frac{n}{2}}^{n-1} q\log(q)$$

$$\sum_{q=1}^{\frac{n}{2}-1} q\log(q) \leq \log(\frac{n}{2})\sum_{q=1}^{\frac{2}{2}-1} q$$

$$\sum_{q=1}^{\frac{n}{2}-1} q = \frac{(\frac{n}{2})(n-1)}{2} \leq \frac{n^2}{8}$$

$$\sum_{q=1}^{\frac{n}{2}-1} q\log(q) = \frac{n^2}{8}(\log(n) - 1)$$

(e)

Using the bound from equation (d), show that the recurrence in equation (c) has the solution $E(T(n)) = O(nlg(n))$.

Show that: $-2n^2 + 2 \leq c * n\log(n)$ for $c \geq 0$ is $O(n\log(n))$
Base case $n = 2$:

$$\begin{aligned} T(2) &\leq c * (2)\log(2) \\ -2(2)^2 + 2 &\leq 2c\log(2) \\ -6 &\leq 2c \\ c &\geq 3 \end{aligned}$$

Inductive Hypothesis:
Let $j > n$:

$$T(j) \leq c * j\log(j)$$

Inductive Step:

$$-2n^2 + 2 \leq c * n\log(n)$$

This is trivially provable since the n^2 factor grows at an asymptotically faster rate that $n\log(n)$, so $f(n)$ will always be less than or equal to $c * n\log(n)$ when $j < n$.
Therefore;

$$\boxed{f(n) = O(n\log(n))}$$

3. You are doing a stress test for the new Iphone. Given a ladder with n rungs, determine the highest run from which you can drop a prototype without it breaking. You want the smallest number of drops.

   (a) You have exactly 2 phones, find highest safe rung in $o(n)$ drops.

$$\begin{aligned} \frac{k(k+1)}{2} &\leq n \\ k^2 + k - 2x &\geq 0 \\ k &= \frac{-1 \pm \sqrt{1+8n}}{2} \\ f(n) &= \Theta\sqrt{n} \end{aligned}$$

$$\boxed{f(n) = \sqrt{n}}$$

   (b) Given k phones, find an algo. for the highest safe run with smallest num drops.

Using $f(n) = \sqrt{n}$ from (a), we can prove by induction that $f(n)$ holds.

Show that $T(k, n) = \sqrt[k]{n} + T(k-1, \sqrt[k]{n}) \leq c * \sqrt[k]{n}$
Base case $n = 1$:

$$\begin{aligned} T(1, n) &= \sqrt[1]{n} \\ &= n \\ T(1, n) &\leq \sqrt[k]{n} \end{aligned}$$

Therefore, the inequality holds when $c = 1$.
Inductive Hypothesis: for all $j < k$, the inequality

$$T(j, n) = \sqrt[k]{n} + T(k-1, \sqrt[k]{n}) \leq c * \sqrt[k]{n}$$

Inductive Step:

$$\begin{aligned} \sqrt[k]{n} + T(k-1, \sqrt[k]{n}) &\leq c * \sqrt[k]{n} \\ &\leq \sqrt[k]{n} + c * \sqrt[k]{n} \\ &\leq (1+c)\sqrt[k]{n} \end{aligned}$$

Therefore,

$$\boxed{T(f, n) = O(\sqrt[k]{n} + T(k-1, \sqrt[k]{n}))}$$

4. The game of Match is played with a special deck of 27 cards. Each card has three attributes: color, shape and number. The possible color values are red, green, and blue. The possible shape values are square, circle, and heart. The possible number values are 1, 2, and 3. Each of the 3 * 3 * 3 = 27 possible combinations is represented by a card in the deck. A match is a set of 3 cards with the property that for every one of the three attributes, either all the cards have the same value for that attribute, or they all have different values for that attribute. For example, the following three cards are a match: (3, red, square), (2 blue square), (1, green, square).

(a) If we shuffle the deck and turn over three cards, what is prob of match?

There are 27 cards, each with 3 values, so $3 * 3 * 3 = 27$. Given that we've already chosen 3 cards freely, the probability that we choose a third card that matches with the previous two is $P(X = x) = \frac{1}{25}$.

(b) Turn over n cards where $n \leq 27$ what is the expected number of matches?

We can calculate the number of possible sets;

$$\binom{n}{3} = \frac{n(n-1)(n-2)}{6}$$

Given that $P(X = x) = \frac{1}{25}$ is the chance that the third card drawn matches the previous two;

$$\binom{n}{3}\frac{1}{25} = \frac{n(n-1)(n-2)}{6}\frac{1}{25} = \frac{n(n-1)(n-2)}{150}$$

We can verify with $n = 27$;

$$\frac{27(26)(25)}{150} = \frac{1}{25} * 2925 = 117$$

So the expected number of matches (including overlap) when 27 cards are turned over is 117.

5. A big square with side length 1 is partitioned into x^2 small squares, each with side length $\frac{1}{x}$, for some positive x. Then n points are distributed independently and uniformly at random in the big square.

(a) Expected number of pairs of points in the same small square.

$$P(2 \text{ points in small square}) = \frac{1}{x^2}$$

We calculate the expected value;

$$\binom{n}{2} * \frac{1}{x^2} = \frac{n(n-1)}{2x^2}$$

For the expected number of pairs of points that fall into the same small square.

$$\begin{aligned}
\frac{n(n-1)}{2x^2} &\geq 1 \\
n(n-1) &\geq 1 \\
x^2 &\leq \frac{n(n-1)}{2} \\
x &\leq \sqrt{\frac{n(n-1)}{2}}
\end{aligned}$$

So when $x \leq \sqrt{\frac{n(n-1)}{2}}$ the expected value is greater than or equal to 1.

(b) Use Markov's inequality for prob. there are at least 2 points in a small square

Markov's inequality says that for a positive random variable X, any $a > 0$:

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

if we let X be the number of pairs of points in the small square;

$$E[X] = \frac{n(n-1)}{2x^2}$$

Let $a = 1$

$$\begin{aligned}
P(X \geq 1) &< \frac{n(n-1)}{2x^2} \\
\frac{n(n-1)}{2x^2} &< 1 \\
n(n-1) &< 2x \\
x &> \sqrt{\frac{n(n-1)}{2}}
\end{aligned}$$

Therefore, for all values greater than $\sqrt{\frac{n(n-1)}{2}}$ our x value will be less than 1.