

5 Tutorial

Thursday, October 10, 2019 3:13 PM

Jensen's Inequality

$f(x)$ is a convex function. (graph below secant line in function domain)

$$f(\mathbb{E}[x]) \leq \mathbb{E}[f(x)]$$

Hockey Stick Identity

- m distinguishable bins and n indistinguishable balls.
there are $\binom{m-1+n}{m-1}$ arrangements

Claim: $\binom{m-1+n}{m} = \sum_{i=0}^{n-1} \binom{i+m-1}{m-1}$

Proof: Fix number of balls in the last bin

$$\text{LHS} = \binom{m-1+n}{m-1} : (n-1) \text{ balls and } (m+1) \text{ balls}$$

RHS: Let $(n-i-1)$ be number of balls in the last bin

Random BST

$$L \sim U(\{1, \dots, n\})$$

X_n : Random variable of BST with n nodes

$$Y_n = 2^{X_n}$$

$Z_i, i \in [n]$ indicator var. for i being the root (first element)

Goal: Find $\mathbb{E}[X_n]$.

Observations:

- $X_n = 1 + \max(X_{i-1}, X_{n-i})$ if i is the root

(i-1) nodes on the left, (n-i) on the right

$$\cdot Y_n = 2 \cdot \max(Y_{i-1}, Y_{n-i}) \quad \text{follows from last point}$$

$$= 2 \sum_{i=1}^n z_i \cdot \max(Y_{i-1}, Y_{n-i})$$

$$\cdot E[z_i] = P(z_i) = \frac{1}{n}$$

Then: $E[Y_n] = \sum_{i=1}^n E[Y_n | z_i = 1] \cdot P(z_i = 1)$ Total probability

$$\begin{aligned}
 (\text{recursive def. of } E[Y_i]) &= \frac{2}{n} \sum_{i=1}^n E[\max(Y_{i-1}, Y_{n-i})] \\
 &\leq \frac{2}{n} \sum_{i=1}^n E[Y_{i-1} + Y_{n-i}] \quad \text{max bounded by sum} \\
 &= \frac{2}{n} \sum_{i=1}^n E[Y_{i-1}] + E[Y_{n-i}] \quad \text{linear} \\
 &= \frac{4}{n} \sum_{i=0}^{n-1} E[Y_i] \quad Y_{i-1} \text{ and } Y_{n-i} \text{ equivalent when summed} \\
 (\dagger) \quad &\leq \frac{1}{4} \binom{n+3}{3} \quad \text{hockey stick, proof by strong induction}
 \end{aligned}$$

By Jensen's inequality, as 2^x is convex:

$$2^{E[X_n]} \leq E[Y_n] \leq \frac{1}{4} \binom{n+3}{3} \leq n^4 \quad (\text{for } n \text{ sufficiently large})$$

$$\text{so } E[X_n] = 4 \log_2(n) \in O(\log n)$$

(expected height)

Problem:

a) Prove (\dagger)

b) Expected number of leaves in a BST. (WTS $\Theta(n)$)

$$\frac{n+1}{3} \quad \text{for } i \in [2, \dots, n-1] \quad P(i \text{ is leaf}) = P(i \text{ after } i-1, i+1 \text{ in perm}) \\ = \frac{1}{3}$$

$$\text{for } i \in \{1, n\}: \quad P(i \text{ is leaf}) = \frac{1}{2} \quad (?)$$

$$(n-1) \frac{1}{3} + 2 \left(\frac{1}{2} \right) = \frac{n+1}{3}$$