CSE 521 HW3

Rohan Mukherjee

October 26, 2023

1. I claim that $h_A = h_B$ iff the minimum value of h among $A \cup B$ occurs in $A \cap B$. The reverse direction is clear, so suppose that the minimum value of h was in $(A \cup B) \setminus (A \cap B)$. Then $h_A < h_B$ or vice versa, a contradiction. Thus,

$$\Pr[h_A = h_B] = \frac{|A \cap B|}{|A \cup B|} = J(A, B)$$

2. We see that

$$\mathbb{E}[Y] = \int_0^1 \Pr[Y \geq x]dx = \int_0^1 \prod_{i=1}^n \Pr[X_i \geq x]dx = \int_0^1 (1-x)^n dx = \int_0^1 x^n dx = \frac{1}{n+1}$$

Also, since $Y^2 = \min\{X_1, \dots, X_n\}^2 = \min\left\{X_1^2, \dots, X_n^2\right\}$, we have that

$$\mathbb{E}\left[Y^2\right] = \int_0^1 \Pr\left[Y^2 \geq x\right]dx = \int_0^1 \prod_{i=1}^n \Pr\left[X_i \geq \sqrt{x}\right]dx = \int_0^1 (1-\sqrt{x})^n dx = \frac{2}{(n+1)(n+2)} \leq \frac{2}{(n+1)^2}$$

Thus $\mathrm{Var}(Y) = \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 \leq \frac{1}{(n+1)^2}$.

3. (a) At the end of the stream Y is the minimum of F_0 independent uniformly distributed r.v.s (since $h(i)$ is a uniform r.v. in $[0,1]$ for each i) in $[0,1]$, and hence by question 2 its expectation is $\frac{1}{F_0+1}$. So $\frac{1}{\mathbb{E}[Y]} - 1 = F_0$.

   (b) We also see that $\mathrm{Var}(Y) \leq \frac{1}{(F_0+1)^2}$. Thus $t(Y) \leq 1$. So Y is an unbiased estimator of $\frac{1}{F_0+1}$ with relative variance at most 1, and hence we can approximate $\frac{1}{F_0+1}$ within a $1 \pm \varepsilon$ multiplicative factor using only $k = O\left(\frac{1}{\varepsilon^2}\log\frac{1}{0.1}\right) = O\left(\frac{1}{\varepsilon^2}\right)$ independent samples of Y with probability $1 - 0.1 = 9/10$. I claim that finding $\frac{1}{F_0+1}$ within a $1 \pm \varepsilon$ multiplicative error is enough to find F_0 within a $1 \pm 4\varepsilon$ multiplicative error. This is true because

$$\frac{1-\varepsilon}{F_0+1} \leq Y \leq \frac{1+\varepsilon}{F_0+1} \iff \frac{F_0-\varepsilon}{1+\varepsilon} \leq \frac{1}{Y} - 1 \leq \frac{F_0+\varepsilon}{1-\varepsilon}$$

   And because

$$\frac{F_0+\varepsilon}{1-\varepsilon} = F_0\frac{1+\varepsilon/F_0}{1-\varepsilon} \leq F_0\frac{1+\varepsilon}{1-\varepsilon} \leq F_0(1+\varepsilon)(1+2\varepsilon) = F_0(1+3\varepsilon+2\varepsilon^2) \leq F_0(1+4\varepsilon)$$

   Where the first inequality holds since $F_0 \geq 1$, the second since $\sum_{k=1}^{\infty} \varepsilon^k \leq \varepsilon/(1-\varepsilon) \leq 2\varepsilon$ for $\varepsilon < 1/2$, and the last when ε is sufficiently small.