Exercise 9.9

Monday, 9 October 2023 14:22

i. For each one of the k answers computed, the probability for each one of them being too large is given by $\Pr[\widetilde{D} > 2D(\sigma)] = \frac{7}{10}$. If the $\frac{k}{2}$-smallest answer is reported, then the probability of the answer being too large is given by the probability of at least $\frac{k}{2}$ answers being too large. This probability is given by $(\Pr[\widetilde{D} > 2D(\sigma)])^{\frac{k}{2}} = \left(\frac{7}{10}\right)^{\frac{k}{2}} > \frac{7}{10}$ for $k \geq 2$.

$2u_n$ over all $i \geq \frac{k}{2}$

ii. We should instead report the answer which is in the 'middle' of the 'correct range'. Here, we have that the smallest $\frac{1}{5}$ answers are likely too small, the next $\frac{1}{10}$ answers are likely in the acceptable range, and the highest $\frac{7}{10}$ answers are likely too high.
Thus, we choose the $\frac{1}{5} + \frac{\frac{1}{10}}{2} = \frac{1}{5} + \frac{1}{20} = \frac{5}{20} = \frac{1}{4}$th lowest answer.

iii. See below.

iv. You would always report the smallest answer; since no answer is too small, the smallest answer is always most likely to fall in the acceptable range.

v. The new algorithm would give a wrong answer only if all of the k answers are wrong. Since the probability of one of the k answers being wrong is equal to $\frac{9}{10}$, the probability of k of them being wrong is $\left(\frac{9}{10}\right)^k$. This means that the success probability is given by $1 - \left(\frac{9}{10}\right)^k$.

iii. Let X_i denote an indicator random variable which is 1 if the i^{th} answer is too large, and let Y_i denote an indicator random variable which is 1 if the i^{th} answer is too small. Let $X = \sum_{i=1}^{k} X_i$ and $Y = \sum_{i=1}^{k} Y_i$. Now, we know that $E[X_i] = \frac{7}{10}$ and $E[Y_i] = \frac{1}{5}$. By linearity of expectation, we then have that $E[X] = \frac{7k}{10}$ and $E[Y] = \frac{k}{5}$.
We then have that the final answer is not a good estimate if either the number of too-large-answers is at least $\frac{3k}{4}$ or if the number of too-small-answers is at least $\frac{k}{4}$. Bounds on these probabilities are given by

$$\Pr\left[X > \frac{3k}{4}\right] = \Pr\left[X > \left(1 + \frac{1}{14}\right) \cdot E[X]\right] < \left(\frac{e^{\frac{1}{14}}}{\left(1 + \frac{1}{14}\right)^{1+\frac{1}{14}}}\right)^{\frac{7k}{10}}$$

$$\Pr\left[Y > \frac{k}{4}\right] = \Pr\left[X > \left(1 + \frac{1}{4}\right) \cdot E[Y]\right] < \left(\frac{e^{\frac{1}{4}}}{\left(1 + \frac{1}{4}\right)^{1+\frac{1}{4}}}\right)^{\frac{k}{5}}$$

Then, it follows that the probability of an answer not being a good estimate is bounded by

$$\Pr[\textit{answer is a good estimate}] > 1 - \left(\frac{e^{\frac{1}{14}}}{\left(1 + \frac{1}{14}\right)^{1+\frac{1}{14}}}\right)^{\frac{7k}{10}} - \left(\frac{e^{\frac{1}{4}}}{\left(1 + \frac{1}{4}\right)^{1+\frac{1}{4}}}\right)^{\frac{k}{5}}$$

$$(1+?)\frac{1}{5} = \frac{1}{4}$$

$$\frac{4}{20} + \frac{?\cdot 4}{20} = \frac{5}{20}$$

$$? \cdot 4 = 1$$

$$? = \frac{1}{4}$$

$$(1+?)\frac{7}{10} = \frac{3}{4}$$

$$\frac{28}{40} + ?\frac{28}{40} = \frac{30}{40}$$

$$?\frac{28}{40} = \frac{2}{40}$$

$$? \cdot 28 = 2$$

$$? = \frac{1}{14}$$

(i) answer is not too large $\Leftrightarrow$ at least half of the answers not too large

We want to bound $\Pr[(\text{number of answers not too large}) \geq \frac{k}{2}]$

define indicator random variables

$$X_i = \begin{cases} 1 & \text{if } i^{th} \text{ answer not too large} \\ 0 & \text{otherwise} \end{cases}$$

$X = \sum_{i=1}^{k} X_i = \#$ answers that are not too large

$\Rightarrow$ compute $E[X]$ and use Chernoff bound

$$E[X] = E\left[\sum_{i=1}^{k} X_i\right] = \sum_{i=1}^{k} E[X_i] = \sum_{i=1}^{k} \Pr[X_i = 1] = k \cdot \frac{3}{10} = \frac{3k}{10}$$

Chernoff bound: $\Pr[X > (1+\delta)E[X]] < \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{E[X]}$

$$\Pr\left[X > \frac{k}{2}\right] = \Pr\left[X > \frac{5}{3}E[X]\right] < \left(\frac{e^{\frac{2}{3}}}{\left(\frac{5}{3}\right)^{\frac{5}{3}}}\right)^{\frac{3k}{10}} < (0.84)^{3k/10}$$

$$\Pr[\text{answer too large}] = 1 - \Pr[\text{answer not too large}] > 1 - (0.84)^{\frac{3k}{10}}$$

the larger k, the greater the probability