CSC473: Advanced Algorithm Design Winter 2020

Worksheet: Variance and Chebyshev's Inequality

Aleksandar Nikolov

In this worksheet you will:

- Recall the concept of *variance* of a random variable;
- Recall some basic facts about variance of sums of random variables;
- Use Chebyshev's Inequality to analyze sampling algorithms.

Variance. The variance of a real-valued random variable X is defined as $\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$. It measures how much X deviates from its expectation on average. The following calculation is often very useful:

$$\begin{aligned}\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] &= \mathbb{E}[X^2 - 2X \cdot \mathbb{E}[X] + \mathbb{E}[X]^2] \\ &= \mathbb{E}[X^2] - 2\mathbb{E}[X]^2 + \mathbb{E}[X]^2 \\ &= \mathbb{E}[X^2] - \mathbb{E}[X]^2.\end{aligned}$$

Chebyshev's Inequality Let X be a random variable. For any $t > 0$,

$$\mathbb{P}(|X - \mathbb{E}[X]| > t) < \frac{\mathrm{Var}(X)}{t^2}.$$

Exercises:

1. Compute the variances of the following random variables.

a. X which is uniform in $\{0, 1\}$.

$\mathrm{Var}(X) =$

b. X which is uniform in $\{0, 1, 2\}$.

$\mathrm{Var}(X) =$

c. X which takes value 0 with probability 1/3, and value 1 with probability 2/3, and Y which takes value 0 with probabiliyt 2/3, and value 1 with probability 1/3

$\mathrm{Var}(X) =$ $\mathrm{Var}(Y) =$

2. Let $X = X_1 + \ldots + X_n$, where $X_1, \ldots, X_n$ are independent and uniform in $\{0, 1\}$. Then

$\mathrm{Var}(X) =$

3. Let S be sampled from $[n] = \{1, \ldots, n\}$ by including each $i \in [n]$ into S independently with probability p.

a.
$\mathbb{E}[|S|] =$

b.
$\mathrm{Var}(|S|) =$

c. Using Chebyshev's inequality, show that, if $\mathbb{E}[|S|] \geq \frac{1}{\varepsilon^2 \delta}$, then

$$\mathbb{P}\Big((1-\varepsilon)\mathbb{E}[|S|] \leq |S| \leq (1+\varepsilon)\mathbb{E}[|S|]\Big) \geq 1 - \delta$$