

Erratum for the book “Algorithms for Big Data”

Moran Feldman

March 7, 2021

In the following, the red text represents the corrections.

- Corollary 3 on Page 36 of Section 2.4 should be

Corollary 3 (Useful form of the Chernoff bound). *Let X_1, X_2, \dots, X_n be n independent numeric random variables taking only values from the range $[0, 1]$. Then, the random variable $X = \sum_{i=1}^n X_i$ obeys*

$$\Pr[X \geq (1 + \delta) \cdot \mathbb{E}[X]] \leq e^{-\frac{\delta^2 \cdot \mathbb{E}[X]}{2 + \delta}} \leq e^{-\frac{\min\{\delta, \delta^2\} \cdot \mathbb{E}[X]}{3}} \quad \text{for every } \delta > 0,$$

and

$$\Pr[X \leq (1 - \delta) \cdot \mathbb{E}[X]] \leq e^{-\frac{\delta^2 \cdot \mathbb{E}[X]}{2}} \quad \text{for every } \delta \in (0, 1).$$

- Footnote 5 on Page 62 of Section 3.2 should be

If k is even, there might be multiple possible median values. We allow Algorithm 3 to return any one of these values.

- The last display math on Page 273 of Section 12.2 should be

$$\begin{aligned} |T_{\leq i}| - |T_{\leq i-1}| &= [n - |CC(G_{\leq i})|] - [n - |CC(G_{\leq i-1})|] \\ &= |CC(G_{\leq i-1})| - |CC(G_{\leq i})|. \end{aligned}$$