Un último uso de la desigualdad de Chebyshev

Lema

$$Pr(\textbf{Length}(S) > 4 \cdot \lfloor n^{\frac{3}{4}} \rfloor) \quad \leq \quad 4 \cdot n^{-\frac{1}{4}}$$

Demostración. Si $\textbf{Length}(S) > 4 \cdot \lfloor n^{\frac{3}{4}} \rfloor$, entonces la menos una de las siguientes condiciones debe ser cierta:

(a) $|\{i \in \{1, \ldots, \textbf{Length}(S)\} \mid S[i] > m\}| \geq 2 \cdot \lfloor n^{\frac{3}{4}} \rfloor$

(b) $|\{i \in \{1, \ldots, \textbf{Length}(S)\} \mid S[i] < m\}| \geq 2 \cdot \lfloor n^{\frac{3}{4}} \rfloor$

Un último uso de la desigualdad de Chebyshev

Vamos a demostrar que la probabilidad de que (a) ocurra es menor o igual a $2 \cdot n^{-\frac{1}{4}}$

De la misma forma se demuestra que la probabilidad que (b) ocurra es menor o igual a $2 \cdot n^{-\frac{1}{4}}$

- De esto se concluye que $\Pr(\textbf{Length}(S) > 4 \cdot \lfloor n^{\frac{3}{4}} \rfloor) \leq 4 \cdot n^{-\frac{1}{4}}$

Un último uso de la desigualdad de Chebyshev

Suponga que (a) es cierto, y sea ℓ la posición de u en la lista L ordenada.

- Tenemos que $\ell \geq \lceil \frac{n}{2} \rceil + 2 \cdot \lfloor n^{\frac{3}{4}} \rfloor$

Dado que $u = R[\lceil \frac{1}{2} \cdot n^{\frac{3}{4}} + n^{\frac{1}{2}} \rceil]$, al menos $\lceil n^{\frac{3}{4}} \rceil - \lceil \frac{1}{2} \cdot n^{\frac{3}{4}} + n^{\frac{1}{2}} \rceil$ elementos de R deben estar en posiciones mayores o iguales a ℓ en la lista L ordenada.

- No podemos asegurar que estos elementos están en posiciones mayores a ℓ puesto que R puede tener elementos repetidos
- Vamos a acotar superiormente la probabilidad de que esto ocurra para obtener un cota superior para la probabilidad de que (a) ocurra

Un último uso de la desigualdad de Chebyshev

Para cada $i \in \{1, \ldots, \lceil n^{\frac{3}{4}} \rceil\}$, definimos una variable aleatoria W_i de la siguiente forma:

$$W_i \quad = \quad \begin{cases} 1 & \text{si la posición de } R[i] \text{ en la lista } L \text{ ordenada} \\ & \text{es mayor o igual a } \lceil \frac{n}{2} \rceil + 2 \cdot \lfloor n^{\frac{3}{4}} \rfloor \\ 0 & \text{en otro caso} \end{cases}$$

Además, definimos la variable aleatoria W como $\displaystyle\sum_{i=1}^{\lceil n^{\frac{3}{4}} \rceil} W_i$

Dado que $\ell \geq \lceil \frac{n}{2} \rceil + 2 \cdot \lfloor n^{\frac{3}{4}} \rfloor$, tenemos entonces que la probabilidad de que (a) ocurra es menor o igual a:

$$\Pr(W \geq \lceil n^{\frac{3}{4}} \rceil - \lceil \frac{1}{2} \cdot n^{\frac{3}{4}} + n^{\frac{1}{2}} \rceil)$$

Un último uso de la desigualdad de Chebyshev

Como L no contiene elementos repetidos obtenemos:

$$
\begin{aligned}
\Pr(W_i = 1) &= \frac{n - (\lceil \frac{n}{2} \rceil + 2 \cdot \lfloor n^{\frac{3}{4}} \rfloor) + 1}{n} \\
&= \frac{\lfloor \frac{n}{2} \rfloor - 2 \cdot \lfloor n^{\frac{3}{4}} \rfloor + 1}{n} \\
&= \frac{\lceil \frac{n}{2} \rceil - 2 \cdot \lfloor n^{\frac{3}{4}} \rfloor}{n}
\end{aligned}
$$

Tenemos entonces que:

$$
\begin{aligned}
\mathrm{E}(W_i) &= \frac{\lceil \frac{n}{2} \rceil - 2 \cdot \lfloor n^{\frac{3}{4}} \rfloor}{n} \\
\mathrm{Var}(W_i) &= \frac{\lceil \frac{n}{2} \rceil - 2 \cdot \lfloor n^{\frac{3}{4}} \rfloor}{n} \cdot \left(1 - \frac{\lceil \frac{n}{2} \rceil - 2 \cdot \lfloor n^{\frac{3}{4}} \rfloor}{n}\right)
\end{aligned}
$$

Un último uso de la desigualdad de Chebyshev

Por lo tanto:

$$\begin{aligned}
\mathsf{E}(W) &= \mathsf{E}\Big(\sum_{i=1}^{\lceil n^{\frac{3}{4}} \rceil} W_i\Big) \\
&= \sum_{i=1}^{\lceil n^{\frac{3}{4}} \rceil} \mathsf{E}(W_i) \\
&= \sum_{i=1}^{\lceil n^{\frac{3}{4}} \rceil} \frac{\lceil \frac{n}{2} \rceil - 2 \cdot \lfloor n^{\frac{3}{4}} \rfloor}{n} \\
&= \lceil n^{\frac{3}{4}} \rceil \cdot \frac{\lceil \frac{n}{2} \rceil - 2 \cdot \lfloor n^{\frac{3}{4}} \rfloor}{n}
\end{aligned}$$

Un último uso de la desigualdad de Chebyshev

Pero tenemos que:

$$
\begin{aligned}
\lceil n^{\frac{3}{4}} \rceil \cdot \frac{\lceil \frac{n}{2} \rceil - 2 \cdot \lfloor n^{\frac{3}{4}} \rfloor}{n} & \leq (n^{\frac{3}{4}} + 1) \cdot \frac{\frac{n}{2} - 2 \cdot n^{\frac{3}{4}} + 3}{n} \\
& = n^{\frac{3}{4}} \cdot \frac{\frac{n}{2} - 2 \cdot n^{\frac{3}{4}} + 3}{n} + \frac{\frac{n}{2} - 2 \cdot n^{\frac{3}{4}} + 3}{n} \\
& = n^{\frac{3}{4}} \cdot \frac{\frac{n}{2} - 2 \cdot n^{\frac{3}{4}}}{n} + 3 \cdot n^{-\frac{1}{4}} + \frac{1}{2} - 2 \cdot n^{-\frac{1}{4}} + \frac{3}{n} \\
& = n^{\frac{3}{4}} \cdot \frac{\frac{n}{2} - 2 \cdot n^{\frac{3}{4}}}{n} + n^{-\frac{1}{4}} + \frac{1}{2} + \frac{3}{n} \\
& = \frac{1}{2} \cdot n^{\frac{3}{4}} - 2 \cdot n^{\frac{1}{2}} + n^{-\frac{1}{4}} + \frac{1}{2} + \frac{3}{n} \\
& \leq \frac{1}{2} \cdot n^{\frac{3}{4}} - 2 \cdot n^{\frac{1}{2}} + 1
\end{aligned}
$$

Concluimos que $\mathrm{E}(W) \leq \frac{1}{2} \cdot n^{\frac{3}{4}} - 2 \cdot n^{\frac{1}{2}} + 1$

Un último uso de la desigualdad de Chebyshev

Para $i, j \in \{1, \ldots, \lceil n^{\frac{3}{4}} \rceil\}$ tal que $i \neq j$ se tiene que W_i es independiente de W_j

Concluimos entonces que:

$$
\begin{aligned}
\mathrm{Var}(W) &= \mathrm{Var}\Big(\sum_{i=1}^{\lceil n^{\frac{3}{4}} \rceil} W_i\Big) \\
&= \sum_{i=1}^{\lceil n^{\frac{3}{4}} \rceil} \mathrm{Var}(W_i) \\
&= \sum_{i=1}^{\lceil n^{\frac{3}{4}} \rceil} \frac{\lceil \frac{n}{2} \rceil - 2 \cdot \lfloor n^{\frac{3}{4}} \rfloor}{n} \cdot \left(1 - \frac{\lceil \frac{n}{2} \rceil - 2 \cdot \lfloor n^{\frac{3}{4}} \rfloor}{n}\right) \\
&= \lceil n^{\frac{3}{4}} \rceil \cdot \frac{\lceil \frac{n}{2} \rceil - 2 \cdot \lfloor n^{\frac{3}{4}} \rfloor}{n} \cdot \left(1 - \frac{\lceil \frac{n}{2} \rceil - 2 \cdot \lfloor n^{\frac{3}{4}} \rfloor}{n}\right)
\end{aligned}
$$

Un último uso de la desigualdad de Chebyshev

Tenemos que:

$$\begin{aligned}
\left(1 - \frac{\lceil \frac{n}{2} \rceil - 2 \cdot \lfloor n^{\frac{3}{4}} \rfloor}{n}\right) \quad &= \quad \frac{n - \lceil \frac{n}{2} \rceil + 2 \cdot \lfloor n^{\frac{3}{4}} \rfloor}{n} \\
&\leq \quad \frac{n - \frac{n}{2} + 2 \cdot n^{\frac{3}{4}}}{n} \\
&= \quad \frac{\frac{n}{2} + 2 \cdot n^{\frac{3}{4}}}{n} \\
&= \quad \frac{1}{2} + 2 \cdot n^{-\frac{1}{4}} \\
&\leq \quad 1
\end{aligned}$$

Un último uso de la desigualdad de Chebyshev

Además tenemos que:

$$\begin{aligned}
\lceil n^{\frac{3}{4}} \rceil \cdot \frac{\lceil \frac{n}{2} \rceil - 2 \cdot \lfloor n^{\frac{3}{4}} \rfloor}{n} \quad &\leq \quad \frac{1}{2} \cdot n^{\frac{3}{4}} - 2 \cdot n^{\frac{1}{2}} + 1 \\
&\leq \quad \frac{1}{2} \cdot n^{\frac{3}{4}} + 1 \\
&\leq \quad n^{\frac{3}{4}}
\end{aligned}$$

Concluimos que $\mathrm{Var}(W) \leq n^{\frac{3}{4}}$

Un último uso de la desigualdad de Chebyshev

Finalmente tenemos que:

$$
\begin{aligned}
\Pr(W \geq \lceil n^{\frac{3}{4}} \rceil - \lceil \frac{1}{2} \cdot n^{\frac{3}{4}} + n^{\frac{1}{2}} \rceil) \quad & \leq \quad \Pr(W \geq n^{\frac{3}{4}} - \frac{1}{2} \cdot n^{\frac{3}{4}} - n^{\frac{1}{2}} - 1) \\
& = \quad \Pr(W \geq \frac{1}{2} \cdot n^{\frac{3}{4}} - n^{\frac{1}{2}} - 1) \\
& = \quad \Pr(W \geq \frac{1}{2} \cdot n^{\frac{3}{4}} - 2 \cdot n^{\frac{1}{2}} + 1 + n^{\frac{1}{2}} - 2) \\
& \leq \quad \Pr(W \geq E(W) + n^{\frac{1}{2}} - 2) \\
& \leq \quad \Pr(W \geq E(W) + n^{\frac{1}{2}} - (1 - \frac{1}{\sqrt{2}}) \cdot n^{\frac{1}{2}}) \\
& \leq \quad \Pr(W \geq E(W) + \sqrt{\frac{n}{2}}) \\
& = \quad \Pr(W - E(W) \geq \sqrt{\frac{n}{2}}) \\
& \leq \quad \Pr(|W - E(W)| \geq \sqrt{\frac{n}{2}})
\end{aligned}
$$

Un último uso de la desigualdad de Chebyshev

Por lo tanto, utilizando la desigualdad de Chebyshev concluimos que:

$$\begin{aligned}
\Pr(W \geq \lceil n^{\frac{3}{4}} \rceil - \lceil \frac{1}{2} \cdot n^{\frac{3}{4}} + n^{\frac{1}{2}} \rceil) &\leq \Pr(|W - E(W)| \geq \sqrt{\frac{n}{2}}) \\
&\leq \frac{\text{Var}(W)}{\frac{n}{2}} \\
&\leq \frac{n^{\frac{3}{4}}}{\frac{n}{2}} \\
&= 2 \cdot n^{-\frac{1}{4}}
\end{aligned}$$

Concluimos que la probabilidad de que (a) ocurra es menor o igual a $2 \cdot n^{-\frac{1}{4}}$. □

El cálculo final

Recuerde que estamos considerando una lista $L[1 \ldots n]$ de números enteros donde n es impar y mayor o igual a 2001.

Para la lista L demostramos lo siguiente:

$$\begin{aligned}
\Pr(Y_1 < \lfloor \frac{1}{2} \cdot n^{\frac{3}{4}} - n^{\frac{1}{2}} \rfloor) &\leq n^{-\frac{1}{4}} \\
\Pr(Y_2 \leq \lceil n^{\frac{3}{4}} \rceil - \lceil \frac{1}{2} \cdot n^{\frac{3}{4}} + n^{\frac{1}{2}} \rceil) &\leq n^{-\frac{1}{4}} \\
\Pr(\textbf{Length}(S) > 4 \cdot \lfloor n^{\frac{3}{4}} \rfloor) &\leq 4 \cdot n^{-\frac{1}{4}}
\end{aligned}$$

El cálculo final

Tenemos entonces que:

$$\Pr(\textbf{CalcularMediana}(L) \text{ retorne } \textit{sin\_resultado}) \quad \leq \quad 6 \cdot n^{-\frac{1}{4}}$$

Dado que $n \geq 2001$ concluimos que

$$\begin{aligned}
\Pr(\textbf{CalcularMediana}(L) \text{ retorne } \textit{sin_resultado}) \quad & \leq \quad 6 \cdot 2001^{-\frac{1}{4}} \\
& < \quad \frac{9}{10}
\end{aligned}$$

El tiempo esperado del algoritmo

Sea p la probabilidad de que **CalcularMediana**(L) retorne resultado

- Tenemos que $p \geq \frac{1}{10}$

¿En promedio cuántas veces se debe llamar a **CalcularMediana**(L) para obtener la mediana de la lista L?

El tiempo esperado del algoritmo

Sea T una variable aleatoria tal que para cada $i \geq 1$:

$$\Pr(T = i) \quad = \quad (1-p)^{i-1} \cdot p$$

Vale decir, T representa el número de llamadas a **CalcularMediana**(L) hasta obtener un resultado

Dado que T tiene distribución geométrica de parámetro p, concluimos que

$$E(T) \quad = \quad \frac{1}{p} \quad \leq \quad 10$$

Concluimos que en promedio se debe llamar 10 veces a **CalcularMediana**(L) para obtener la mediana de la lista L