Algoritmos Aleatorizados

Parte IV

Segundo semestre 2022

IIC2283

Prof. Nicolás Van Sint Jan

Recordatorio: Cálculo de la mediana

Suponga que el procedimiento **Mergesort**(L) ordena una lista L utilizando el algoritmo Mergesort.

El siguiente procedimiento calcula la mediana de una lista de enteros $L[1 \ldots n]$ (suponiendo que n es impar y L no tiene elementos repetidos):

CalcularMediana($L[1 \ldots n]$)
    if $n < 2001$ **then**
        Mergesort(L)
        return $L[\lceil \frac{n}{2} \rceil]$
    else
        sea R una lista de $\left\lceil n^{\frac{3}{4}} \right\rceil$ números enteros escogido con
                distribución uniforme y de manera independiente desde L
        Mergesort(R)

Un algoritmo aleatorizado para el cálculo de la mediana

$d := R\left[\left\lfloor \frac{1}{2} \cdot n^{\frac{3}{4}} - n^{\frac{1}{2}} \right\rfloor\right]$

$u := R\left[\left\lceil \frac{1}{2} \cdot n^{\frac{3}{4}} + n^{\frac{1}{2}} \right\rceil\right]$

$S := \varnothing$

$m_d := 0$

$m_u := 0$

for $i := 1$ **to** n **do**

  if $d \leq L[i]$ **and** $L[i] \leq u$ **then** **Append**$(S, [L[i]])$

  else if $L[i] < d$ **then** $m_d := m_d + 1$

  else $m_u := m_u + 1$

if $m_d \geq \left\lceil \frac{n}{2} \right\rceil$ **or** $m_u \geq \left\lceil \frac{n}{2} \right\rceil$ **or**

  Length$(S) > 4 \cdot \left\lfloor n^{\frac{3}{4}} \right\rfloor$ **then return** *sin_resultado*

else

  Mergesort(S)

  return $S\left[\left\lceil \frac{n}{2} \right\rceil - m_d\right]$

Outline

Calculo de la mediana (cont.)

Desigualdades de Markov y Chebyshev

Acotando la probabilidad

El algoritmo es correcto y eficiente

Ejercicios

Demuestre lo siguiente:

1. Si **CalcularMediana**(L) retorna un número entero m, entonces m es la mediana de L

2. Si se tiene un procedimiento **LanzarMoneda**() que retorna 0 ó 1 con probabilidad $\frac{1}{2}$, entonces existe un algoritmo para construir R que invoca a este procedimiento a los más $c \cdot n^{\frac{3}{4}} \cdot \log_2(n)$ veces, donde c es una constante fija y n es el largo de la lista de entrada
    - Podemos suponer que **LanzarMoneda**() en el peor caso es $\mathcal{O}(1)$

3. **CalcularMediana**(L) en el peor caso es $\mathcal{O}(n)$, suponiendo que n es el largo de L y considerando todas las operaciones realizadas

¿Cuál es la probabilidad de no retornar un resultado?

La llamada **CalcularMediana**(L) puede NO retornar un resultado

- El procedimiento en este caso retorna *sin_resultado*

Para que **CalcularMediana** pueda ser utilizado en la práctica la probabilidad que no entregue un resultado debe ser baja

La probabilidad de no retornar resultado

Sea $L[1 \ldots n]$ una lista de números enteros tal que $n \geq 2001$, n es impar y la mediana de L es m

Defina las siguientes variables aleatorias:

$$\begin{aligned} Y_1 &= \left|\left\{i \in \left\{1, \ldots, \left\lceil n^{\frac{3}{4}} \right\rceil\right\} \mid R[i] \leq m\right\}\right| \\ Y_2 &= \left|\left\{i \in \left\{1, \ldots, \left\lceil n^{\frac{3}{4}} \right\rceil\right\} \mid R[i] \geq m\right\}\right| \end{aligned}$$

Estas son variables aleatorias dado que R es construido escogiendo elementos de L con distribución uniforme (y de manera independiente)

¿ Qué representan Y_1 e Y_2 ?

La probabilidad de no retornar resultado

Lema

CalcularMediana(L) retorna sin_resultado si y sólo si alguna de las siguientes condiciones se cumple:

1. $Y_1 < \left\lfloor \frac{1}{2} \cdot n^{\frac{3}{4}} - n^{\frac{1}{2}} \right\rfloor$
2. $Y_2 \leq \left\lceil n^{\frac{3}{4}} \right\rceil - \left\lceil \frac{1}{2} \cdot n^{\frac{3}{4}} + n^{\frac{1}{2}} \right\rceil$
3. **Length**$(S) > 4 \cdot \left\lfloor n^{\frac{3}{4}} \right\rfloor$

Ejercicio

Demuestre el lema.

La probabilidad de no retornar resultado

Tenemos entonces que la probabilidad de que **CalcularMediana**(L) retorne *sin_resultado* es igual a:

$$\Pr\Big(Y_1 < \left\lfloor \frac{1}{2} \cdot n^{\frac{3}{4}} - n^{\frac{1}{2}} \right\rfloor \ \vee \\ Y_2 \leq \left\lceil n^{\frac{3}{4}} \right\rceil - \left\lceil \frac{1}{2} \cdot n^{\frac{3}{4}} + n^{\frac{1}{2}} \right\rceil \ \vee \ \textbf{Length}(S) > 4 \cdot \left\lfloor n^{\frac{3}{4}} \right\rfloor\Big)$$

Necesitamos entonces acotar superiormente

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

¿ Que herramientas podemos usar para **acotar** estas probabilidades ?

Outline

Calculo de la mediana (cont.)

Desigualdades de Markov y Chebyshev

Acotando la probabilidad

La desigualdad de Markov

Teorema

Sea X una variable aleatoria no negativa. Para cada $a \in \mathbb{R}^+$ se tiene que:

$$\Pr(X \geq a) \quad \leq \quad \frac{\mathrm{E}(X)}{a}$$

Este resultado se conoce como la **desigualdad de Markov**.

Una demostración de la desigualdad de Markov

Demostración

Suponemos que el recorrido de X es un conjunto finito $\Omega \subseteq \mathbb{R}_0^+$:

$$
\begin{aligned}
\mathrm{E}(X) &= \sum_{r \in \Omega} r \cdot \Pr(X = r) \\
&= \Bigg(\sum_{r \in \Omega \,:\, r < a} r \cdot \Pr(X = r) \Bigg) + \Bigg(\sum_{s \in \Omega \,:\, s \geq a} s \cdot \Pr(X = s) \Bigg) \\
&\geq \sum_{s \in \Omega \,:\, s \geq a} s \cdot \Pr(X = s) \\
&\geq \sum_{s \in \Omega \,:\, s \geq a} a \cdot \Pr(X = s) \\
&= a \cdot \Bigg(\sum_{s \in \Omega \,:\, s \geq a} \Pr(X = s) \Bigg) \\
&= a \cdot \Pr(X \geq a)
\end{aligned}
$$

Concluimos que $\Pr(X \geq a) \leq \frac{\mathrm{E}(X)}{a}$. $\square$

La desigualdad de Chebyshev

Teorema

El siguiente resultado se conoce como la **desigualdad de Chebyshev**:

$$\Pr\left(|X - \mathrm{E}(X)| \geq a\right) \quad \leq \quad \frac{\mathrm{Var}(X)}{a^2}$$

Demostración

Utilizando la desigualdad de Markov obtenemos:

$$\begin{aligned}
\Pr(|X - \mathrm{E}(X)| \geq a) &= \Pr((X - \mathrm{E}(X))^2 \geq a^2) \\
&\leq \frac{\mathrm{E}((X - \mathrm{E}(X))^2)}{a^2} \\
&= \frac{\mathrm{Var}(X)}{a^2}
\end{aligned}$$

$\square$

Outline

Calculo de la mediana (cont.)

Desigualdades de Markov y Chebyshev

Acotando la probabilidad

Utilizando la desigualdad de Chebyshev

Lema

$$\Pr\left(Y_1 < \left\lfloor \frac{1}{2} \cdot n^{\frac{3}{4}} - n^{\frac{1}{2}} \right\rfloor\right) \quad \leq \quad n^{-\frac{1}{4}}$$

Demostración

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

$$X_i \quad = \quad \begin{cases} 1 & R[i] \leq m \\ 0 & R[i] > m \end{cases}$$

Tenemos que:

$$Y_1 \quad = \quad \sum_{i=1}^{\lceil n^{3/4} \rceil} X_i$$

Utilizando la desigualdad de Chebyshev

Demostración

Dado que la lista L no contiene elementos repetidos tenemos que:

$$\Pr(X_i = 1) \;=\; \frac{\left\lceil \frac{n}{2} \right\rceil}{n} \;=\; \frac{\frac{n-1}{2} + 1}{n} \;=\; \frac{1}{2} + \frac{1}{2 \cdot n}$$

De esto se deduce que:

$$\begin{aligned}
\mathrm{E}(X_i) &= \frac{1}{2} + \frac{1}{2 \cdot n} \\
\mathrm{Var}(X_i) &= \left(\frac{1}{2} + \frac{1}{2 \cdot n}\right) \cdot \left(1 - \frac{1}{2} - \frac{1}{2 \cdot n}\right) \\
&= \left(\frac{1}{2} + \frac{1}{2 \cdot n}\right) \cdot \left(\frac{1}{2} - \frac{1}{2 \cdot n}\right) \\
&= \frac{1}{4} - \frac{1}{4 \cdot n^2}
\end{aligned}$$

Utilizando la desigualdad de Chebyshev

Demostración

Por lo tanto tenemos que:

$$\begin{aligned}
\mathrm{E}(Y_1) &= \mathrm{E}\left(\sum_{i=1}^{\lceil n^{3/4} \rceil} X_i\right) \\
&= \sum_{i=1}^{\lceil n^{3/4} \rceil} \mathrm{E}(X_i) \\
&= \sum_{i=1}^{\lceil n^{3/4} \rceil} \left(\frac{1}{2} + \frac{1}{2 \cdot n}\right) \\
&= \left\lceil n^{\frac{3}{4}} \right\rceil \cdot \left(\frac{1}{2} + \frac{1}{2 \cdot n}\right)
\end{aligned}$$

Utilizando la desigualdad de Chebyshev

Demostración

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

Concluimos entonces que:

$$
\begin{aligned}
\mathrm{Var}(Y_1) &= \mathrm{Var}\left(\sum_{i=1}^{\lceil n^{3/4} \rceil} X_i\right) \\
&= \sum_{i=1}^{\lceil n^{3/4} \rceil} \mathrm{Var}(X_i) \\
&= \sum_{i=1}^{\lceil n^{3/4} \rceil} \left(\frac{1}{4} - \frac{1}{4 \cdot n^2}\right) \\
&= \left\lceil n^{\frac{3}{4}} \right\rceil \cdot \left(\frac{1}{4} - \frac{1}{4 \cdot n^2}\right) \\
&\leq \frac{1}{4} \cdot \left\lceil n^{\frac{3}{4}} \right\rceil
\end{aligned}
$$

Utilizando la desigualdad de Chebyshev

Demostración

Tenemos que:

$$
\begin{aligned}
\Pr(Y_1 < \left\lfloor \frac{1}{2} \cdot n^{\frac{3}{4}} - n^{\frac{1}{2}} \right\rfloor) &\leq \Pr(Y_1 < \frac{1}{2} \cdot n^{\frac{3}{4}} - n^{\frac{1}{2}}) \\
&\leq \Pr(Y_1 < \frac{1}{2} \cdot \left\lceil n^{\frac{3}{4}} \right\rceil - n^{\frac{1}{2}}) \\
&\leq \Pr(Y_1 < \left\lceil n^{\frac{3}{4}} \right\rceil \cdot (\frac{1}{2} + \frac{1}{2 \cdot n}) - n^{\frac{1}{2}}) \\
&= \Pr(Y_1 < \mathrm{E}(Y_1) - n^{\frac{1}{2}}) \\
&= \Pr(n^{\frac{1}{2}} < \mathrm{E}(Y_1) - Y_1) \\
&\leq \Pr(|Y_1 - \mathrm{E}(Y_1)| > n^{\frac{1}{2}}) \\
&\leq \Pr(|Y_1 - \mathrm{E}(Y_1)| \geq n^{\frac{1}{2}})
\end{aligned}
$$

Utilizando la desigualdad de Chebyshev

Demostración

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

$$\begin{aligned}
\Pr\left(Y_1 < \left\lfloor \frac{1}{2} \cdot n^{\frac{3}{4}} - n^{\frac{1}{2}} \right\rfloor\right) &\leq \Pr(|Y_1 - \mathrm{E}(Y_1)| \geq n^{\frac{1}{2}}) \\
&\leq \frac{\mathrm{Var}(Y_1)}{n} \\
&\leq \frac{1}{4} \cdot \frac{\left\lceil n^{\frac{3}{4}} \right\rceil}{n} \\
&\leq \frac{1}{4} \cdot \frac{n^{\frac{3}{4}} + 1}{n} \\
&\leq \frac{1}{4} \cdot \frac{2 \cdot n^{\frac{3}{4}}}{n} \\
&= \frac{1}{2} \cdot n^{-\frac{1}{4}} \\
&\leq n^{-\frac{1}{4}}
\end{aligned}$$

$\square$

Utilizando nuevamente la desigualdad de Chebyshev

Lema

$$\Pr(Y_2 \leq \left\lceil n^{\frac{3}{4}} \right\rceil - \left\lceil \frac{1}{2} \cdot n^{\frac{3}{4}} + n^{\frac{1}{2}} \right\rceil) \quad \leq \quad n^{-\frac{1}{4}}$$

Ejercicio

Demuestre el lema utilizando las ideas en la demostración del lema anterior.

Un último uso de la desigualdad de Chebyshev

Lema

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

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

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

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

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

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 $\ell \geq \lceil \frac{n}{2} \rceil + 2 \cdot \left\lfloor n^{\frac{3}{4}} \right\rfloor$

Dado que $u = R[\left\lceil \frac{1}{2} \cdot n^{\frac{3}{4}} + n^{\frac{1}{2}} \right\rceil]$, al menos $\left\lceil n^{\frac{3}{4}} \right\rceil - \left\lceil \frac{1}{2} \cdot n^{\frac{3}{4}} + n^{\frac{1}{2}} \right\rceil$ 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 $i \in \{1, \ldots, \left\lceil n^{\frac{3}{4}} \right\rceil\}$, definimos una variable aleatoria W_i de la siguiente forma:

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

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

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

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

Un último uso de la desigualdad de Chebyshev

Como L no contiene elementos repetidos obtenemos:

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

Tenemos entonces que:

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

Un último uso de la desigualdad de Chebyshev

Por lo tanto:

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

Un último uso de la desigualdad de Chebyshev

Pero tenemos que:

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

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

Concluimos entonces que:

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

Un último uso de la desigualdad de Chebyshev

Además tenemos que:

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

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

Un último uso de la desigualdad de Chebyshev

Finalmente tenemos que:

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

Un último uso de la desigualdad de Chebyshev

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

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

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

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:

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

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