MÉTODO DO SEGUNDO MOMENTO

<u>DEF</u>: A VARIÂNCIA DE UMA VARIÁVEL ALEATÓRIA X É DEFINIDA POR

$$Var(X) = \mathbb{E}\Big[\big(\underbrace{X - \mathbb{E}[X]}_{Y}\big)^2\Big]$$

E O DESVIO PADRÃO DE X É O NÚMERO $\sigma(X) = \sqrt{Var(X)}$

PELA DEFINIÇÃO, TEMOS

$$(X - \mathbb{E}[X])^2 = X^2 - 2X\,\mathbb{E}[X] + \mathbb{E}[X]^2 \quad \text{CONSTANTES}$$

$$Var(X) = \mathbb{E}\big[(X - \mathbb{E}[X])^2\big] = \mathbb{E}\big[X^2 - 2X\,\mathbb{E}[X] + \mathbb{E}[X]^2\big]$$

$$= \mathbb{E}[X^2] - \cancel{2}\,\mathbb{E}[X]\,\mathbb{E}[X] + \cancel{\mathbb{E}[X]^2} = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 \geq 0$$

OBS: $\mathrm{Var}(c \cdot X) = \mathbb{E}[c^2 X^2] - \mathbb{E}[cX]^2$

$$= c^2\,\mathbb{E}[X^2] - c^2\,\mathbb{E}[X]^2 = c^2\,\mathrm{Var}(X)$$

Prop. Se $X_1, \dots, X_n$ são variáveis aleatórias independentes, então

$$Var(X_1 + \dots + X_n) = Var(X_1) + \dots + Var(X_n)$$

Prova Lembre-se que, se X_i e X_j são independ., então

$$\mathbb{E}[X_i \cdot X_j] = \mathbb{E}[X_i] \cdot \mathbb{E}[X_j].$$

Logo $Var(X_1 + \dots + X_n) = \mathbb{E}[(X_1 + \dots + X_n)^2] - \mathbb{E}[(X_1 + \dots + X_n)]^2$

$$= \mathbb{E}\Big[\sum_i X_i^2 + 2\sum_{i<j} X_i X_j\Big] - \Big(\sum_i \mathbb{E}[X_i]\Big)^2$$

$$= \sum_i \mathbb{E}[X_i^2] + \cancel{2\sum_{i<j} \mathbb{E}[X_i]\,\mathbb{E}[X_j]} - \Big(\sum_i \mathbb{E}[X_i]^2 + \cancel{2\sum_{i<j} \mathbb{E}[X_i]\,\mathbb{E}[X_j]}\Big)$$

$$= \sum_i \mathbb{E}[X_i^2] - \mathbb{E}[X_i]^2 = \sum_i Var(X_i)$$

$\blacksquare$

PROPOSIÇÃO (DESIGUALDADE DE CHEBYCHEV).

MARKOV: $\mathbb{P}(X \geq \lambda) \leq \frac{\mathbb{E}[X]}{\lambda}$

SE X É UMA VARIÁVEL ALEATÓRIA E $\lambda > 0$, ENTÃO

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq \lambda) \leq \frac{Var(X)}{\lambda^2}$$

PROVAR: APLICAMOS MARKOV A $|X - \mathbb{E}[X]|$:

$$\mathbb{P}(|X-\mathbb{E}[X]| \geq \lambda) = \mathbb{P}((X-\mathbb{E}[X])^2 \geq \lambda^2) \leq \frac{Var(X)}{\lambda^2}$$

$\blacksquare$

SE TOMAMOS $\lambda = t \cdot \underbrace{\sigma}_{\sqrt{Var(X)}}(X)$, OBTEMOS

$$\mathbb{P}(|X-\mathbb{E}[X]| \geq t \cdot \sigma(X)) \leq \frac{Var(X)}{\lambda^2} = \frac{Var(X)}{t^2 \cdot Var(X)} = \frac{1}{t^2}$$

ISSO É: X ESTÁ A UMA DISTÂNCIA DE NO MÁXIMO $t\sigma(X)$ COM PROB. PELO MENOS $1 - t^{-2}$

DADO $A \subseteq \mathbb{N}$, DENOTAMOS POR $\sum(A)$ O CONJUNTO DE NÚMEROS QUE PODEM SER OBTIDOS PELA SOMA DE NÚMEROS DISTINTOS DE A.

$$\sum(A) = \left\{ \sum_{s \in S} s \ : \ S \subseteq A \right\}$$

EX: $A = \{1, 5, 8\}$.

$$\sum(A) = \{0, 1, 5, 8, 6, 9, 13, 14\}$$

NOTE QUE $\left|\sum(A)\right| \leq 2^{|A|}$

EX: $A = \{2, 3, 4, 5\}$

$$\sum(A) = \{0, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14\}$$

$$\left|\sum(A)\right| = 13 < 16 = 2^{|A|}$$

Dizemos que A tem somas distintas se $|\sum(A)| = 2^{|A|}$.

Def: Para cada $n \in \mathbb{N}$, denote por $f(n)$ o tamanho do maior subconjunto de $[n]$ com somas distintas.

Obs: $f(n) \geq 1 + \lfloor \log_2 n \rfloor$ basta considerar $A = \{2^i : 0 \leq i \leq \lfloor \log_2 n \rfloor\}$

$$A = \{1, 2, 4\}$$

$$\sum(A) = \{0, 1, 2, 4, 3, 5, 6, 7\}$$

Teorema: Para todo $n \in \mathbb{N}$, temos

$$f(n) \leq \log_2 n + \log_2 \log_2 n + O(1)$$

existe C absoluta tal que ....

Teorema: Existe $C \in \mathbb{N}$ t.q., para todo $n \in \mathbb{N}$, temos

$$f(n) \leq \log_2 n + \log_2 \log_2 n + C$$

Teorema: Para todo $n \in \mathbb{N}$, temos

$$f(n) \leq \log_2 n + \log_2 \log_2 n + O(1)$$

Prova: Seja $A = \{a_1, \ldots, a_m\} \subseteq [n]$ um conj. com somas distintas contido em $[n]$.

Sejam $X_1, \ldots, X_m$ var. aleat. independentes, com distrib. de Bernoulli com parâmetro $1/2$.

Note que $Var(X_i) = \mathbb{E}[X_i^2] - \mathbb{E}[X_i]^2 = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$

Considere a var. aleat. $X = a_1 X_1 + \cdots + a_m X_m$.

Temos que

$$\mu := \mathbb{E}[X] = \frac{a_1 + \cdots + a_m}{2}$$

NOTE QUE $\mathrm{Var}(X_i) = \mathbb{E}[X_i^2] - \mathbb{E}[X_i]^2 = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$

CONSIDERE A VAR. ALEAT. $X = a_1 X_1 + \cdots + a_m X_m$.

TEMOS QUE

$$\mu := \mathbb{E}[X] = \frac{a_1 + \cdots + a_m}{2}$$

CONCLUIMOS QUE

$$\begin{aligned}\mathrm{Var}(X) &= \mathrm{Var}(a_1 X_1) + \cdots + \mathrm{Var}(a_m X_m) \\ &= a_1^2 \mathrm{Var}(X_1) + \cdots + a_m^2 \mathrm{Var}(X_m) = \frac{a_1^2 + \cdots + a_m^2}{4} \\ &\le \frac{m \cdot n^2}{4}\end{aligned}$$

$\le n$

logo $\sigma(X) = \frac{n\sqrt{m}}{2}$

Seja $t > 1$, pela des. de Chebychev, temos

$$\mathbb{P}\left(|X - \mathbb{E}[X]| \geq \frac{tn\sqrt{m}}{2}\right) \leq \frac{1}{t^2}$$

e

$$\mathbb{P}\left(|X - \mathbb{E}[X]| < \frac{tn\sqrt{m}}{2}\right) \geq 1 - \frac{1}{t^2} \qquad = \mathbb{P}(X \in B)$$

Por outro lado, como A tem somas distintas, para $x \in \mathbb{N}$, temos

$$\mathbb{P}(X = x) = \begin{cases} 0 & \text{se } x \notin \Sigma(A) \\ 1/2^m & \text{se } x \in \Sigma(A) \end{cases}$$

Defina $B = \left\{ x \in \mathbb{N} : |x - \mathbb{E}[X]| < \frac{tn\sqrt{m}}{2} \right\} = \left[\mathbb{E}[X] - \frac{tn\sqrt{m}}{2}, \mathbb{E}[X] + \frac{tn\sqrt{m}}{2} \right]$

$$\mathbb{P}(X \in B) \leq \frac{tn\sqrt{m} + 1}{2^m}$$

$$\mathbb{P}(X \in B) \geq 1 - \frac{1}{t^2}$$

$$\mathbb{P}(X \in B) \leq \frac{t n \sqrt{m} + 1}{2^m}$$

$$\frac{t n \sqrt{m} + 1}{2^m} \geq 1 - \frac{1}{t^2} = \frac{t^2 - 1}{t^2}$$

SABEMOS QUE

$m \leq n$

$$2^m \leq \underbrace{\frac{t^2}{t^2-1} t}_{c} n \sqrt{m} + 1 = c n \sqrt{m} + 1 \leq c' n \sqrt{m}$$

$$m \leq c'' + \log_2 n + \frac{1}{2} \log_2 m$$

$$\leq c'' + \log_2 n + \frac{1}{2} \log_2 \left(c'' + \log_2 n + \frac{1}{2} \log_2 m \right)$$

$m \leq n$

$$\leq c'' + \log_2 n + \frac{1}{2} \log_2 \left(c'' + \log_2 n + \frac{1}{2} \log_2 n \right)$$

$$m \leq C'' + \log_2 n + \frac{1}{2}\log_2 m$$
$$\leq C'' + \log_2 n + \frac{1}{2}\log_2\left(C'' + \log_2 n + \frac{1}{2}\log_2 m\right)$$

$m \leq n \Rightarrow$
$$\leq C'' + \log_2 n + \frac{1}{2}\log_2\left(C'' + \log_2 n + \frac{1}{2}\log_2 n\right)$$

$$\leq C''' + \log_2 n + \frac{1}{2}\log_2\log_2 n \qquad \blacksquare$$

$$C'' + \frac{3}{2}\log_2 n \leq 2\log_2 n$$
$$\log_2\left(C'' + \frac{3}{2}\log_2 n\right) \leq \log_2\left(2\log_2 n\right) = 1 + \log_2\log_2 n$$