3 - wariancja, Markow, Czebyszew i kolekcjoner kuponów

Wednesday, 31 January 2024 20:50

Wariancja

$$Var(X) = \mathbb{E}\left[(X-\mathbb{E}[X])^2\right] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

$\sigma[X] = \sqrt{Var(X)}$ - odchylenie standardowe

Kowariancja

$$Cov(X,Y) = \mathbb{E}\left[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])\right] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

$$Var(X+Y) = Var(X) + Var(Y) + 2\,Cov(X,Y)$$

Jeśli X i Y są niezależne:

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$$

Jeśli X i Y są niezależne:

$Cov(X,Y) = 0$ oraz $Var(X+Y) = Var(X) + Var(Y)$

Nierówność Markowa

Jeśli X przyjmuje wartości nieujemne, to dla każdego $a > 0$

$$\mathbb{P}[X \geq a] \leq \frac{\mathbb{E}[X]}{a}$$

I - indykator $\frac{X}{a} \geq 1 \Rightarrow I \leq \frac{X}{a}$

$$\mathbb{P}[X \geq a] = \mathbb{E}[I] \leq \mathbb{E}\left[\frac{X}{a}\right] = \frac{\mathbb{E}[X]}{a}$$

Nierówność Czebyszewa

Dla każdego $a > 0$

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

$$\mathbb{P}\left[|X-\mathbb{E}[X]| \geq a\right] = \mathbb{P}\left[|X-\mathbb{E}[X]|^2 \geq a^2\right]$$

$$\leq \frac{\mathbb{E}\left[(X-\mathbb{E}[X])^2\right]}{a^2} = \frac{Var(X)}{a^2}$$

Kolekcjoner kuponów

X - czas zbierania

$$\mathbb{E}[X] = n \ln n + O(n)$$

Z Markowa

$$\mathbb{P}[X \geq 2nH_n] \leq \frac{1}{2}$$

$X_i \sim Geo\left[\frac{n-i+1}{n}\right]$

$Var(Geo(p)) = \frac{1-p}{p^2} \leq \frac{1}{p^2}$

$$Var(X) = \sum_{i=1}^{n} Var(X_i) \leq \sum_{i=1}^{n} \left(\frac{n}{n-i+1}\right)^2$$

$$= n^2 \sum_{i=1}^{n} \frac{1}{i^2} = n^2 \frac{\pi^2}{6}$$

Z Czebyszewa

$$\mathbb{P}\left[|X - nH_n| \geq nH_n\right] \leq \frac{n^2 \frac{\pi^2}{6}}{n^2 H_n^2} = \Theta\left(\frac{1}{\ln^2 n}\right)$$

Z Union-bounda

Szansa, że po $n \ln n + cn$ krokach nie wyszkaliśmy i-tego kuponu: $\left(1-\frac{1}{n}\right)^{n(\ln n + c)} < e^{-\ln n + c} = \frac{e^{-c}}{n}$

Zatem $\mathbb{P}[X \geq n\ln n + cn] < e^{-c}$