28.05.10 32° μάθημα

Ανισότητες - Νόμοι Μεγάλων Αριθμών - Κεντρικό Οριακό Θεώρημα

① Ανισότητα Markov

Θεώρημα: Έστω $\underline{X \geq 0}$ τ.μ. Τότε $P(X \geq a) \leq \frac{E[X]}{a}$ $\forall a > 0$

Απόδειξη: Για X συνεχή με σ.π.π. $f_X(x)$:

$$E[X] = \int_0^\infty x f_X(x)\,dx \geq \int_a^\infty x f_X(x)\,dx \geq \int_a^\infty a f_X(x)\,dx =$$

$$= a\int_a^\infty f_X(x)\,dx = a P(X \geq a) \Rightarrow P(X \geq a) \leq \frac{E[X]}{a}$$

Για X διακριτή όμοια (με αθροίσματα)

Εναλλακτικά, για X συνεχή ή διακριτή

$$I = I(x) = \begin{cases} 1, & x \geq a \\ 0, & x < a \end{cases} \quad \text{τότε} \quad a \cdot I \leq X \Rightarrow E[aI] \leq E[X]$$

$$\Rightarrow a E[I] \leq E[X] \Rightarrow a P(X \geq a) \leq E[X] \Rightarrow P(X \geq a) \leq \frac{E[X]}{a}$$

② Ανισότητα Chebyshev

Θεώρημα: Έστω X τ.μ. με $E[X] = \mu$, $Var[X] = \sigma^2$. Τότε:

$$P(|X-\mu| \geq a) \leq \frac{\sigma^2}{a^2}, \quad \forall a > 0$$

Απόδειξη: $P(|X-\mu| \geq a) = P(|X-\mu|^2 \geq a^2) \overset{\text{Markov}}{\leq} \frac{E[(X-\mu)^2]}{a^2} = \frac{\sigma^2}{a^2}$

③ Τι σημαίνει $Var[X] = 0$;

Θεώρημα: $Var[X] = 0 \Leftrightarrow X = E[X]$ με π.θ. 1

Απόδειξη: $(\Rightarrow)$ $\forall a>0 : P(|X-\mu|\geq a) \leq \frac{\sigma^2}{a^2} = 0$

Άρα $P(|X-\mu|\geq a) = 0$, $\forall a>0$

αύξουσα ακολουθία ενδεχομένων

$P(X\neq\mu) = P(|X-\mu|>0) = P\left(\bigcup_{n=1}^{\infty}\{|X-\mu|\geq \frac{1}{n}\}\right) =$

$= \lim_{n\to\infty} P(|X-\mu|\geq\frac{1}{n}) = 0 \Rightarrow P(X=\mu) = 1-0 = 1.$

$(\Leftarrow)$ $X = E[X]$ με πιθ 1 $\Rightarrow X - E[X] = 0$ με πιθ 1

$\Rightarrow E[(X-E[X])^2] = 0$

$\Rightarrow Var[X] = 0$

④ Νόμος Μεγάλων Αριθμών (ΝΜΑ) - ασθενής

Αν $X_1, X_2 \dots$ ανεξάρτητες και ισόνομες τ.μ. με $E[X_i]=\mu$, τότε για $S_n = X_1+X_2+\dots+X_n$ και $\bar{X}_n = \frac{S_n}{n}$: Δειγματικός μέσος, ισχύει $P(|\bar{X}_n-\mu|\geq\varepsilon)\to 0$ καθώς $n\to\infty$, $\forall\varepsilon>0$

Απόδειξη: (μόνο για την περίπτωση που $Var[X_i]=\sigma^2<\infty$)

$$S_n = \sum_{i=1}^{n} X_i \Rightarrow E[S_n] = n\mu \text{ και } Var[S_n] = n\sigma^2$$

$$\Rightarrow E[\bar{X}_n] = \mu \text{ και } Var[\bar{X}_n] = \frac{\sigma^2}{n}$$

$$P(|\bar{X}_n-\mu|\geq\varepsilon) \overset{\text{Chebyshev}}{\leq} \frac{Var(\bar{X}_n)}{\varepsilon^2} = \frac{\sigma^2}{\varepsilon^2 n} \to 0 \text{ καθώς } n\to\infty$$

$$\Rightarrow \lim_{n\to\infty} P(|\bar{X}_n-\mu|\geq\varepsilon) = 0.$$

(5) Κεντρικό Οριακό Θεώρημα (ΚΟΘ)

Έστω $X_1, X_2 \ldots$ ανεξάρτητες και ισόνομες τ.μ. με $E[X_i]=\mu$ και $Var[X_i]=\sigma^2<\infty$ τότε

$$\lim_{n\to\infty} P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \le x\right) = \underbrace{\Phi(x)}_{P(Z\le x)} = \int_{-\infty}^{x} \underbrace{\frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}}_{\text{σ.π.π. κατ. } N(0,1)} du$$

$$\overset{\shortparallel}{\lim_{n\to\infty} P\left(\frac{S_n - E[S_n]}{\sqrt{Var[S_n]}} \le x\right)} \qquad Z\sim N(0,1)$$

Διαισθητικά

$X_1, X_2, \ldots$ μετρήσεις επαναλαμβ. πειρ

$S_n \approx N(n\mu, n\sigma^2)$ για μεγάλα n

και

$\bar{X}_n \approx N(\mu, \frac{\sigma^2}{n})$ γιατί για $\bar{X}_n = \dfrac{X_1+X_2+\ldots+X_n}{n}$

ισχύει $$\lim_{n\to\infty} P\left(\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \le x\right) = \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$$

$$\overset{\shortparallel}{\lim_{n\to\infty} P\left(\frac{\bar{X}_n - E[\bar{X}_n]}{\sqrt{Var[\bar{X}_n]}} \le x\right)}$$

(6) Παράδειγμα

T_i χρόνος ζωής λαμπτήρα

$T_i \sim Exp$ με $E[T_i]=200$ ώρες

Τυχαίο Δείγμα από 49 λαμπτήρες

ανεξ και ισον

Να υπολογισθούν προσεγγιστικά

P_1: P (ο συνολικός χρόνος ζωής τους να υπερβεί τις 10000 ώρες) = ;

P_2: P (το πολύ 14 από αυτούς να ζήσουν λιγότερο από 140 ώρες) = ;

Λύση: Έστω T_k = χρόνος ζωής k λαμπτήρα

$$S_{49} = \sum_{k=1}^{49} T_k \quad \text{συνολ. χρόνος ζωής}$$

$$P_1 = P(S_{49} > 10.000) = P\left(\frac{S_{49} - E[S_{49}]}{\sqrt{Var[S_{49}]}} \geq \frac{10000 - E[S_{49}]}{\sqrt{Var[S_{49}]}} \right)$$

$$T_k \sim Exp\left(\tfrac{1}{200}\right)$$

$$E[T_k] = 200 \quad \Rightarrow \quad E[S_{49}] = 49 \cdot 200 = 9.800$$

$$Var[T_k] = 40.000 \quad \Rightarrow \quad \sqrt{Var[S_{49}]} = \sqrt{49 \cdot 40.000} = 1400$$

$$P_1 \overset{\text{κοθ}}{\simeq} P\left(Z \geq \frac{10000 - 9800}{1400}\right), \quad Z \sim \mathcal{N}(0,1)$$

$$= P(Z \geq 1/7) = 1 - P(Z \leq 1/7) = 1 - \Phi(1/7)$$

$$I_k = \begin{cases} 1, & \text{ο χρόνος ζωής του } k \text{ λαμπτήρα} < 140 \text{ ώρες} \\ 0, & \text{διαφορετικά} \end{cases}$$

$$S_{49} = \sum_{k=1}^{49} I_k \sim Bin(49, P(T_k < 140)) = \#\text{λαμπτήρων που ζουν} < 140\ h$$

$$P_2 = P(S_{49} \leq 14) = P\left(\frac{S_{49} - E[S_{49}]}{\sqrt{Var[S_{49}]}} \leq \frac{14 - E[S_{49}]}{\sqrt{Var[S_{49}]}} \right) = *$$

$$S_{49} \sim Bin(49, p)$$

$$p = P(T_k < 140) = 1 - e^{-\frac{140}{200}} = 0.5$$

$$E[S_{49}] = n \cdot p = 49 \cdot 0.5 = 24.5$$

$$Var[S_{49}] = n\,p\,(1-p) = 49 \cdot 0.5 \cdot 0.5 = 12.25$$

$$* = P\left(\frac{S_{49} - 24.5}{\sqrt{12.25}} \leq \frac{14 - 24.5}{\sqrt{12.25}} \right) \overset{\text{κοθ}}{\simeq} P(Z \leq -3), \quad Z \sim \mathcal{N}(0,1)$$

$$= \Phi(-3) = 1 - \Phi(3) = 0.001$$

⑦ Παράδειγμα

4000 κάτοικοι

10 άτομα κατά μέσο όρο χρειάζεται να εισαχθούν σε νοσοκομείο τη μέρα

Προσεγγιστικός υπολογισμός μικρότερου αριθμού κλινών σε ιατρ. κέντρο ώστε η πόλη να εξυπηρετείται με π.θ. 95% τη μέρα

Λύση: $I_i = \begin{cases} 1, & \text{αν ο κάτοικος } i \text{ χρειαστεί νοσοκομείο} \\ 0, & \text{διαφορετικά} \end{cases}$

S_{4000} = # ατόμων που χρ. το νοσ.

Ψάχνω ελάχιστο k ώστε:

$$P(S_{4000} \le \underset{\#\text{κλινών}}{k}) \ge 0.95$$

$$P(I_i = 1) = \frac{10}{4000} = \frac{1}{400}$$

$$E[S_{4000}] = 4000 \cdot \frac{1}{400} = 10$$

$$Var[S_{4000}] = 4000 \cdot \frac{1}{400} \cdot \frac{399}{400} = \frac{3990}{400} \simeq 9{,}975$$

$$\Rightarrow P\left(\frac{S_{4000} - 10}{\sqrt{9{,}975}} \le \frac{k-10}{\sqrt{9{,}975}}\right) \ge 0.95$$

$$\Rightarrow \Phi\left(\frac{k-10}{\sqrt{9{,}975}}\right) \ge 0.95 = \Phi(1{,}645)$$

$$\Rightarrow \frac{k-10}{\sqrt{9.975}} \ge 1{,}645 \Rightarrow k \ge 10 + 1.645 \cdot \sqrt{9.975} \Rightarrow k \ge 16.$$