32ο Μάθημα Πιθανοτήτων — 28/5/2010

Ανισότητες - Νόμος Μεγάλων Αριθμών - Κεντρικό Οριακό Θεώρημα

1) Ανισότητα Markov

ΘΕΩΡΗΜΑ (ισχύει για όλες τις τ.μ.)

$X \geq 0$ τ.μ. Τότε $P(X \geq \alpha) \leq \frac{E[X]}{\alpha}$ $\forall \alpha > 0$

Απόδειξη:

• Για X συνεχή με σ.π.π. $f_X(x)$

$$E[X] = \int^{+\infty} x f_X(x) dx \geq \int_\alpha^\infty x f_X(x) dx \geq \int_\alpha^\infty \alpha f_X(x) dx = \alpha \int_\alpha^\infty f_X(x) dx =$$

$$= \alpha P(X \geq \alpha) \Rightarrow P(X \geq \alpha) \leq \frac{E[X]}{\alpha}$$

• Για X διακριτή ομοίως. Εναλλακτικά για X συνεχή ή διακριτή

$$I = I(X) = \begin{cases} 1, & X \geq \alpha \\ 0, & X < \alpha \end{cases} \quad \text{Τότε } \alpha \cdot I \leq X$$

$$\Rightarrow E[\alpha I] \leq E[X] \Rightarrow \alpha E[I] \leq E[X]$$

$$\Rightarrow \alpha P(X \geq \alpha) \leq E[X].$$

② Ανισότητα Chebyshev

Θεώρημα:

X τ.μ. με $E[X]=\mu$, $Var[X]=\sigma^2$. Τότε: $P(|X-\mu|\geq a) \leq \dfrac{\sigma^2}{a^2}$

Απόδειξη:

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

(Για $a<0$ δεν έχει νόημα, αφού η απόλυτη τιμή είναι πάντα μεγαλύτερη από αρνητικούς αριθμούς και άρα $P(|X-\mu|\geq a)=1$).

③ Τι σημαίνει $Var[X]=0$;

Θεώρημα:

$Var[X]=0 \Leftrightarrow X=E[X]$ με πιθαν. 1.

Απόδειξη:

$(\Leftarrow)$ $X=E[X]$ με πιθ. 1 $\Rightarrow X-E[X]=0$ με πιθ. 1 $\Rightarrow E[(X-E[X])^2]=0 \Rightarrow Var[X]=0$

$(\Rightarrow)$ Για κάθε $a>0$: $P(|X-\mu|\geq a) \leq \dfrac{\sigma^2}{a^2}=0$.

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

$$P(X\neq\mu) = P(|X-\mu|>0) = P\Big(\bigcup_{n=1}^{\infty} \overbrace{\{|X-\mu|\geq \tfrac{1}{n}\}}^{\text{αύξουσα ακολουθία ενδ.}}\Big) = \lim_{n\to\infty} P(|X-\mu|\geq \tfrac{1}{n}) = 0.$$

Άρα: $P(X=\mu)=1-0=1$.

④ Νόμος Μεγάλων Αριθμών (ΝΜΑ) - Ασθενής

Αν $X_1, X_2, \ldots$ ανεξ. και ισόνομες τ.μ. με $E[X_i]=\mu$ τότε

$S_n = X_1+X_2+\ldots+X_n$

$\bar{X}_n = \dfrac{S_n}{n}$: Δειγματικός Μέσος

Ισχύει: $P(|\bar{X}_n-\mu|\geq\varepsilon) \to 0$ καθώς $n\to\infty \quad \forall\varepsilon>0$.

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

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

$\Rightarrow E[\bar{X}_n]=\mu$, $Var[\bar{X}_n]=\dfrac{\sigma^2}{n}$

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

⑤ Κεντρικό Οριακό Θεώρημα (ΚΟΘ)

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

<u>Τότε:</u>

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

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

<u>Διαισθητικά:</u> $X_1, X_2, \ldots$ μετρήσεις επαναλαμβ. πειράματος

$S_n \approx N(n\mu, n\sigma^2)$ για μεγάλα n, $\bar{X}_n \approx N(\mu, \frac{\sigma^2}{n})$

Επίσης για $\bar{X}_n = \dfrac{X_1+X_2+\ldots+X_n}{n}$ ισχύει:

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

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

⑥ Παράδειγμα

T_k χρόνος ζωής λαμπτήρα, $T_k \sim Exp$ με $E[T_k]=200$ ώρες

Τυχαίο δείγμα (= ανεξ + ισόνομες) από 49 λαμπτήρες.

Να υπολογιστούν προσεγγιστικά:

P(ο συνολικός χρόνος ζωής τους να υπερβεί τις 10.000 ώρες) = ; $= P_1$

P(το πολύ 14 λαμπτήρες να ζήσουν λιγότερο από 140 ώρες) = ; $= P_2$

<u>Λύση:</u>

T_k = χρόνος ζωής κ λαμπτήρων

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

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

$T_k \sim Exp\left(\frac{1}{200}\right)$, $E[T_k]=200$, $Var[T_k]=40.000$

$\Rightarrow \sqrt{Var[S_{49}]} = \sqrt{49 \cdot 40000} = 1400$

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

$P_1 \overset{ΚΟΘ}{\simeq} P\left(Z \geq \frac{10.000 - 9.800}{1400}\right), \quad Z \sim N(0,1)$

$= P(Z \geq \frac{1}{7}) = 1 - P(Z < \frac{1}{7}) = 1 - \Phi(\frac{1}{7})$

Για την P_2 έχουμε:

T_k: χρόνος ζωής του κ λαμπτήρα

$I_k: \begin{cases} 1, & \text{ο χρόνος ζωής του κ λαμπ.} < 140 \text{ ώρες} \\ 0, & \text{διαφορ.} \end{cases}$

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

Άρα: $P_2 = P(S_{49} \leq 14)$

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

$p = P(T_k < 140) = 1 - e^{-\frac{140}{200}} \approx 0,5$

$E[S_{49}] = np = 49 \cdot 0,5 = 24,5$

$Var[S_{49}] = np(1-p) = 49 \cdot 0,5 \cdot 0,5 = 12,25$

Επομένως: $P_2 = P(S_{49} \leq 14) = P\left(\frac{S_{49} - 24,}{\sqrt{12,25}} \leq \frac{14 - 24,5}{\sqrt{12,25}}\right)$

$\overset{ΚΟΘ}{\simeq} P(Z \leq -3), \quad Z \sim N(0,1)$

$= \Phi(-3) = 1 - \Phi(3) = 0,001.$

(7) Παράδειγμα

4.000 κάτοικοι

10 άτομα κατά μέσο όρο χρειάζονται να εισαχθούν σε νοσοκομείο τη μέρα. Προσεγγ. υπολ. μικρότερου αριθμού κλινών σε ιατρικό κέντρο ώστε η πόλη να εξυπηρ. με πιθ. 95% τη μέρα

$$I = \begin{cases} 1, & \text{αν ο κάτοικος χρειάζεται το νοσοκ.} \\ 0, & \text{διαφορ.} \end{cases}$$

S_{4000} = # ατόμων που χρειάζ. το νοσοκ.

Ψάχνω ελάχιστο k ώστε: $P(S_{4000} \le \underset{\#\text{ κλινών}}{k}) \geq 0{,}95$

$$P(I_i = 1) = \frac{10}{4000} = \frac{1}{400}, \quad E[S_{4000}] = 4000\,\frac{1}{400} = 10$$

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

$$P(S_{4000} \le k) \geq 0{,}95 \Rightarrow P\left(\frac{S_{4000}-10}{\sqrt{9{,}975}} \le \frac{k-10}{\sqrt{9{,}975}}\right) \geq 0{,}95$$

$$\Rightarrow \Phi\left(\frac{k-10}{\sqrt{9{,}975}}\right) \geq 0{,}95 = \Phi(1{,}645) \Rightarrow \frac{k-10}{\sqrt{9{,}975}} \geq 1{,}645 \Rightarrow k \geq 16.$$