24/5/13

Νόμος των μεγάλων αριθμών (ΝΜΑ)

1. Νόμος Μεγάλων Αριθμών (Ασθενής)

$X_1, X_2, \ldots$: ανεξάρτητες, ισόνομες με $E[X_i] = \mu < \infty$

$\hookrightarrow \sum_x x f_X(x)$ ή $\int_{-\infty}^{\infty} x f_X(x)\,dx$

$\Downarrow$

$\bar{X}_n = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{S_n}{n}$ → Δειγματικός Μέσος

$\Downarrow$

$\forall \varepsilon > 0 \quad \lim_{n\to\infty} (P(\bar{X}_n - \mu) > \varepsilon) = 0$

Απόδειξη (απλοποιημένη όταν $\exists \; Var[X] = \sigma^2 < \infty$)
(Πλήρης απόδειξη → Khintchin)

$P(|\bar{X}_n - \mu| > \varepsilon)$
$= P(|\bar{X}_n - E[\bar{X}_n]|) > \varepsilon \leq \frac{Var[\bar{X}_n]}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} \xrightarrow{n\to\infty} 0$

$E[\bar{X}_n] = E[\frac{1}{n}\sum_{i=1}^{n} X_i] = \frac{1}{n}\sum_{i=1}^{n} E[X_i] = \mu$

$Var[\bar{X}_n] = Var[\frac{1}{n}\sum X_i] = \frac{1}{n^2}\sum_{i=1}^{n} Var[X_i] = \frac{\sigma^2}{n}$

2. Κεντρικό Οριακό Θεώρημα

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

$S_n = \sum_{i=1}^{n} X_i \quad , \quad \bar{X}_n = \frac{S_n}{n}$

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

$$\lim_{n\to\infty} P\left(\frac{\bar{X}_n - E[\bar{X}_n]}{\sqrt{Var[\bar{X}_n]}}\right) = P(Z\in A) \quad , \quad Z\sim N(0,1)$$

Άλλες Μορφές

$$\lim_{n\to\infty} P\left(\frac{S_n - n\cdot\mu}{\sigma\sqrt{n}} \in A\right) = P(Z\in A) \quad , \quad Z\sim N(0,1)$$

$$\lim_{n\to\infty} P\left(\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \in A\right) = P(Z\in A) \quad , \quad Z\sim N(0,1)$$

Συνήθως $A = (-\infty, x]$

$$\lim_{n\to\infty} P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x\right) = \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

↳ συνάρτηση κατανομής της $N(0,1)$

$$\lim_{n\to\infty} P\left(\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \leq x\right) = \Phi(x)$$

3. Διαισθητική έννοια του Κ.Ο.Θ.

$X_1, X_2, \ldots$: ανεξάρτητες, ισόνομες (ότι κατανομή και να έχουν) με $E[X_i] = \mu$, $Var[X_i] = \sigma^2$ τότε για μεγάλα n

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

↑ προσεγγίζεται

4. Ιδέα Απόδειξης

Θα αποδείξουμε ότι $\lim_{n\to\infty} P\left(\frac{S_n - n\cdot\mu}{\sigma\sqrt{n}} \leq x\right) = \Phi(x)$

Λήμμα (Θεώρημα συνέχειας)

$Z, Z_1, Z_2, \ldots$ τ.μ. με σ.κ. $F_Z(x), F_{Z_1}(x), F_{Z_2}(x), \ldots$

και ροπογεννήτριες $M_Z(t), M_{Z_1}(t), M_{Z_2}(t), \ldots$

Τότε $\lim_{n\to\infty} M_{Z_n}(t) = M_Z(t) \Rightarrow \lim_{n\to\infty} F_{Z_n}(x) = F_Z(x)$ για κάθε σημείο συνέχειας της F_Z

Εφαρμόζω το Λήμμα για $Z_n = \frac{S_n - n\cdot\mu}{\sigma\sqrt{n}}$, $Z \sim N(0,1)$

Αρχικά υποθέτω ότι $\mu = 0$, $\sigma^2 = 1$

Τότε αρκεί να δείξω ότι: $\lim_{n\to\infty} \underbrace{P(\underbrace{\frac{S_n}{\sqrt{n}}}_{Z_n} \leq x)}_{F_{Z_n}(x)} = \underbrace{P(Z \leq x)}_{F_Z(x)}$

Αρκεί $\lim_{n\to\infty} M_{Z_n}(t) = M_Z(t) = e^{t^2/2}$

$M_{Z_n}(t) = E[e^{tZ_n}] = E[e^{\frac{tS_n}{\sqrt{n}}}] = M_{S_n}\left(\frac{t}{\sqrt{n}}\right) = M_X\left(\frac{t}{\sqrt{n}}\right)^n$

Αρκεί να δείξουμε ότι $\lim_{n\to\infty} M_X\left(\frac{t}{\sqrt{n}}\right)^n = e^{t^2/2}$

$\Leftrightarrow \lim_{n\to\infty} n\cdot L\left(\frac{t}{\sqrt{n}}\right) = \frac{t^2}{2}$, $L = \log M_X$

Εχουμε $\lim_{n\to\infty} n \cdot L\left(\frac{t}{\sqrt{n}}\right) = \lim_{n\to\infty} \frac{L\left(\frac{t}{\sqrt{n}}\right)}{n^{-1}} \underset{D.H.}{\overset{\left(\frac{0}{0}\right)}{=}} \lim_{n\to\infty} \frac{L'\left(\frac{t}{\sqrt{n}}\right)\left(-\frac{1}{2}\right)n^{-3/2}\, t}{-1 n^{-2}}$

$$= \lim_{n\to\infty} \frac{L'\left(\frac{t}{\sqrt{n}}\right)}{2 n^{-1/2}} \underset{D.H.}{\overset{\left(\frac{0}{0}\right)}{=}} \lim_{n\to\infty} \frac{L''\left(\frac{t}{\sqrt{n}}\right) t^2 \left(-\frac{1}{2}\right) n^{-3/2}}{2\left(-\frac{1}{2}\right) n^{-3/2}}$$

- $L(t) = \log M_x(t)$

$$= \frac{t^2}{2} L''(0) = 1$$

- $L'(t) = \frac{M_x'(t)}{M_x(t)}$
- $L''(t) = \frac{M_x''(t) \cdot M_x(t) - M_x'^2(t)}{M_x^2(t)}$

$\Rightarrow L''(0) = \frac{1 \cdot 1 - 0^2}{1} = 1$

Στη γενικη περίπτωση αν $E[X_i] = \mu$, $Var[X_i] = \sigma^2$ τότε $X_i' = \frac{X_i - \mu}{\sigma}$ τότε το Κ.Ο.Θ. ισχύει

Οπότε $\lim_{n\to\infty} P\left(\frac{S_n'}{\sqrt{n}} \le x\right) = \Phi(x)$

Όμως $S_n' = \frac{S_n - n \cdot \mu}{\sigma}$

5. Άσκηση

Δείγμα από $n = 49$ λαμπτήρες

$T_k \sim \exp$, $E[T_k] = 200$ ώρες, $k = 1, 2, \dots, 49$

Να υπολογίσετε προσεγγιστικά:

$P_1 = P($συνολικός χρόνος ζωής ≥ 10.000 ώρες$) = ;$

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

Λύση

$P_1 = P(S_{49} \geq 10.000)$, όπου $S_{49} = \sum_{k=1}^{49} T_k$

$P_1 = P(S_{49} \geq 10.000) = P\left(\frac{S_{49} - E[S_{49}]}{\sqrt{Var[S_{49}]}} \geq \frac{10.000 - E[S_{49}]}{\sqrt{Var[S_{49}]}}\right)$ $Var[T_k] = \frac{1}{\lambda^2} = 40.000$

$\overset{\text{K.O.Θ.}}{\simeq} P\left(Z \geq \frac{10.000 - 49 \cdot 200}{\sqrt{49 \cdot 40.000}}\right)$, $Z \sim N(0,1)$

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

$P_2 = P($το πολύ 14 απ' τους 49 να ζήσουν λιγότερο από 140 ώρες$)$

$I_k = \begin{cases} 1, & \text{ο λαμπτήρας κ ζει λιγότερο από 140 ώρες} \\ 0, & \text{διαφορετικά} \end{cases}$

$P_2 = P\left(\underbrace{\sum_{k=1}^{49} I_k}_{=S'_{49}} \leq 14\right) = P\left(\frac{S'_{49} - E[S'_{49}]}{\sqrt{Var[S'_{49}]}} \leq \frac{14 - E[S'_{49}]}{\sqrt{Var[S'_{49}]}}\right)$

$\overset{\text{K.O.Θ.}}{\simeq} P\left(Z \leq \frac{14 - 49 \cdot 0{,}5}{\sqrt{49 \cdot 0{,}25}}\right)$, $Z \sim N(0,1) = P\left(Z \leq -\frac{10{,}5}{3{,}5}\right) = P(Z \leq -3)$

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

- $E[I_k] = P(I_k = 1) = P(T_k \leq 140) = \int_0^{140} \lambda e^{-\lambda t} dt = 1 - e^{-\frac{140}{200}} \simeq 0{,}5$

  $\lambda = 1/200$

- $Var[I_k] = p(1-p) = 0{,}5 - 0{,}5^2 = 0{,}25$

34ο Μάθημα (τελευταίο) 27/05/13

Κεντρικό Οριακό Θεώρημα

Ασκήσεις - Εφαρμογές

① Κ.Ο.Θ.

$X_1, X_2, \ldots$: ανεξάρτητες και ισόνομες $E[X_i] = \mu < \infty$

$Var[X_i] = \sigma^2 < \infty$

$$S_n = \sum_{i=1}^{n} X_i, \quad \bar{X}_n = \frac{S_n}{n}$$

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

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

$$Z \sim N(0,1)$$

② Τρόπος Εργασίας

$$P(a \le S_n \le b) = ;$$

$$\| $$

$$P\left(\frac{a-n\mu}{\sigma\sqrt{n}} \le \frac{S_n-n\mu}{\sigma\sqrt{n}} \le \frac{b-n\mu}{\sigma\sqrt{n}}\right)$$

$$\approx$$

$$\Phi\left(\frac{b-n\mu}{\sigma\sqrt{n}}\right) - \Phi\left(\frac{a-n\mu}{\sigma\sqrt{n}}\right)$$

Αν οι X_i παίρνουν ακέραιες τιμές τότε

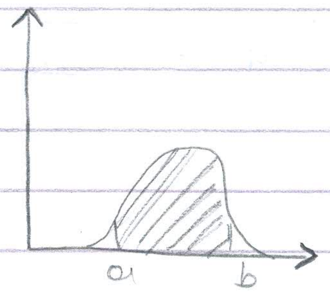
$$P(a \le S_n \le b) \quad , a,b \in \mathbb{Z}$$

$$\|$$

$$P\left(a-\frac{1}{2} \le S_n \le b+\frac{1}{2}\right)$$

$$\|$$

$$P\left(a-\frac{3}{4} \le S_n \le b+\frac{4}{5}\right)$$



κ.λ.π.

⊛ Ποιο να διαλέξω για να εφαρμόσω το ΚΟΘ;

Τεχνική διόρθωσης συνέχειας

Όταν S_n ακέραια, $a,b \in \mathbb{Z}$

$$P(a \le S_n \le b)$$

$$\|$$

$$P\left(a-\frac{1}{2} \le S_n \le b+\frac{1}{2}\right)$$

και εφαρμόζω εδώ Κ.Ο.Θ. (διόρθωση συνέχειας)

③ Παράδειγμα

Παίκτης ρίχνει ζάρι.

Ζάρια	1	2	3	4	5	6
Κέρδος	-1	-2	-3	3	2	1

Προσεγγιστικός υπολογισμός πιθανότητας σε 42 ρίψεις να έχει συνολικό κέρδος τουλάχιστον 7.

X_i = κέρδος στην i ρίψη

$S_{42} = \sum_{i=1}^{42} X_i$: κέρδος συνολικό

$$P(S_{42} \geq 7) = P(S_{42} \geq 6.5) = P\left(\frac{S_{42} - E[S_{42}]}{\sqrt{Var[S_{42}]}} \geq \frac{6.5 - E[S_{42}]}{\sqrt{Var[S_{42}]}}\right)$$

Διόρθωση συνέχειας

$$E[X_1] = \frac{1}{6}(-1) + \frac{1}{6}(-2) + \frac{1}{6}(-3) + \frac{1}{6}\cdot 3 + \frac{1}{6}\cdot 2 + \frac{1}{6}\cdot 1 = 0$$

$$Var[X_i] = E[X_i^2] - \cancel{E[X_i]}^{\,0}{}^2$$
$$= \frac{1}{6}(-1)^2 + \frac{1}{6}(-2)^2 + \frac{1}{6}(-3)^2 + \cdots = \frac{28}{6}$$

$$E[S_{42}] = 42\,E[X_i] = 42 \cdot 0 = 0$$

$$Var[S_{42}] = 42\,Var[X_i] = 42 \cdot \frac{28}{6} = 14^2$$

$$P(S_{42} \geq 7) = P\left(\underbrace{\frac{S_{42} - E[S_{42}]}{\sqrt{Var[S_{42}]}}}_{\approx Z \sim N(0,1)} \geq \frac{6.5 - 0}{\sqrt{14^2}}\right) \overset{ΚΟΘ}{\simeq} P\left(Z \geq \frac{6.5}{14}\right)_{Z \sim N(0,1)} =$$

$$= 1 - \Phi\left(\frac{6.5}{14}\right)$$

④ Άσκηση

Πόλη 4000 κατοίκων

10 άτομα/ημέρα χρειάζονται νοσηλεία κατά μ.ο.

Προσεγγιστικός υπολογισμός ελάχιστου αριθμού κλινών ώστε η πόλη να εξυπηρετείται χωρίς διακομιδές ασθενών σε άλλες πόλεις με πιθανότητα 95%

—— ◦ ——

$I_k = \begin{cases} 1, & \text{αν ο κάτοικος } k \text{ χρειάζεται νοσηλεία μια συγκεκριμένη μέρα} \\ 0, & \text{διαφορετικά} \end{cases}$

$S_{4000} = \sum_{k=1}^{4000} I_k = \#$ κατοίκων που χρειάζονται νοσηλεία τη συγκεκριμένη μέρα

Θέλουμε το ελάχιστο x ώστε $P(S_{4000} \le x) \ge 0{,}95$

$$P(S_{4000} \le x) \ge 0{,}95 \overset{\text{ΔΙΟΡΘΩΣΗ ΣΥΝΕΧΕΙΑΣ}}{\Longleftrightarrow} P(S_{4000} \le x + 1/2) \ge 0{,}95 \Leftrightarrow$$

$$\Leftrightarrow P\left(\frac{S_{4000} - E[S_{4000}]}{\sqrt{Var[S_{4000}]}} \le \frac{x + \frac{1}{2} - E[S_{4000}]}{\sqrt{Var[S_{4000}]}} \right) \ge 0{,}95$$

$$\overset{\text{Κ.Ο.Θ.}}{\Longleftrightarrow} P\left(Z \le \frac{x + 0{,}5 - 10}{\sqrt{9{,}9}} \right) \ge 0{,}95 \Leftrightarrow \Phi\left(\frac{x - 9{,}5}{\sqrt{9{,}9}} \right) \ge 0{,}95 = \Phi(1{,}96)$$

$Z \sim N(0,1)$

$$\left(\begin{array}{l} E[I_k] = 1/400 \\ Var[I_k] = E[I_k^2] - E[I_k]^2 = 1/400 - 1/(400)^2 = 399/400^2 \\ E[S_{4000}] = 4000\, E[I_k] = 10 \\ Var[S_{4000}] = 4000 \cdot \frac{399}{(400)^2} = \frac{3990}{400} \simeq 9{,}9 \end{array} \right)$$

$$\Leftrightarrow \frac{x - 9{,}5}{\sqrt{9{,}9}} \ge 1{,}96 \Leftrightarrow x \ge 9{,}5 + 1{,}96\sqrt{9{,}9}$$

5 Άσκηση

Ημερήσιο εισόδημα χαρτοπαίκτη $\sim$ Uniform$([-5,5])$ σε χιλιάδες €.

Προσεγγιστικός υπολογισμός

i) Πιθανότητα σε 48 μέρες να κερδίσει τουλάχιστον 30.

ii) Το ποσό s ώστε με πιθανότητα τουλάχιστον 95% το εισόδημα σε 48 ημέρες $\leq s$.

iii) Το πλήθος των ημερών που πρέπει να παίζει ώστε με πιθανότητα τουλάχιστον 95% το εισόδημα κατά απόλυτη τιμή < 5.0

$P(S_n \underset{>}{\overset{\leq}{=}} x) \geq y$

μέρες ποσό επίπεδο βεβαιότητας

X_i = εισόδημα τη μέρα i

$$E[X_i] = \int_{-5}^{5} x \cdot \frac{1}{b-a}\,dx = 0 = \frac{\overset{-5}{a}+\overset{5}{b}}{2}$$

$$Var[X_i] = E[X_i^2] - \cancel{E[X_i]}^{2} = \int_{-5}^{5} x^2 \frac{1}{b-a}\,dx$$

$$= \frac{1}{10}\left[\frac{x^3}{3}\right]_{x=-5}^{5} = \frac{25}{3} = \frac{(b-a)^2}{12}$$

$E[S_n] = 0 \cdot n = 0$

$Var[S_n] = n \cdot \frac{25}{3} = \frac{25n}{3}$

(i) $P(S_{48} \geq 30) = P\left(\frac{S_{48} - E[S_{48}]}{\sqrt{Var[S_{48}]}} \geq \frac{30-0}{\sqrt{\frac{25 \cdot 48}{3}}}\right)$

$$\overset{KOΘ}{\simeq} P\left(Z \geq \frac{30}{20}\right) = 1 - \Phi(1,5)$$

(ii) $s = ;$

$$P(|S_{48}| \leq s) \geq 0{,}95 \Leftrightarrow P(-s \leq S_{48} \leq s) \geq 0{,}95 \Leftrightarrow$$

$$\Leftrightarrow P\left(\frac{-s - E[S_{48}]}{\sqrt{Var[S_{48}]}} \leq \frac{S_{48} - E[S_{48}]}{\sqrt{Var[S_{48}]}} \leq \frac{s - E[S_{48}]}{\sqrt{Var[S_{48}]}}\right) \geq 0{,}95 \Leftrightarrow$$

$E[S_{48}] = 0$, $Var[S_{48}] = 48 \cdot \frac{25}{3} = 400$

$$\overset{KOΘ}{\Leftrightarrow} P\left(\frac{-s}{20} \leq Z \leq \frac{s}{20}\right) \geq 0{,}95 \quad , Z \sim N(0,1)$$

$$\Leftrightarrow \Phi\left(\frac{s}{20}\right) - \Phi\left(-\frac{s}{20}\right) \geq 0{,}95 \xLeftrightarrow{\Phi(-x) = 1 - \Phi(x)}$$

$$\Leftrightarrow 2\Phi\left(\frac{s}{20}\right) - 1 \geq 0{,}95 \Leftrightarrow \Phi\left(\frac{s}{20}\right) \geq 0{,}975 = \Phi(1{,}96)$$

$$\Leftrightarrow \frac{s}{20} \geq 1{,}96 \Leftrightarrow s \geq 20 \cdot 1{,}96$$

(iii) $n = ;$ $\quad P(|S_n| < 50) \geq 0{,}95 \Leftrightarrow P(-50 < S_n < 50) \geq 0{,}95 \Leftrightarrow$

$$\Leftrightarrow P\left(\frac{-50}{\sqrt{\frac{25n}{3}}} < \frac{S_n - E[S_n]}{\sqrt{Var[S_n]}} < \frac{50}{\sqrt{\frac{25n}{3}}}\right) \geq 0{,}95 \Leftrightarrow$$

$$\overset{KOΘ}{\Leftrightarrow} P\left(\frac{-50}{\sqrt{\frac{25n}{3}}} < Z < \frac{50}{\sqrt{\frac{25n}{3}}}\right) \geq 0{,}95 \quad , Z \sim N(0,1) \Leftrightarrow$$

$$\Leftrightarrow \Phi\left(\frac{50}{\sqrt{\frac{25n}{3}}}\right) - \Phi\left(\frac{-50}{\sqrt{\frac{25n}{3}}}\right) \geq 0{,}95 \Leftrightarrow 2\Phi\left(\frac{50}{\sqrt{\frac{25n}{3}}}\right) - 1 \geq 0{,}95$$

$$\Leftrightarrow \Phi\left(\frac{50}{\sqrt{\frac{25n}{3}}}\right) \geq 0{,}975 = \Phi(1{,}96) \Leftrightarrow \frac{50}{\sqrt{\frac{25n}{3}}} \geq 1{,}96 \Leftrightarrow$$

$$\Leftrightarrow \frac{10\sqrt{3}}{\sqrt{n}} \geq 1{,}96 \Leftrightarrow n \leq \left(\frac{10\sqrt{3}}{1{,}96}\right)^2$$

⑥ Σύνοψη του μαθήματος - Τυπικές ασκήσεις

1| Έννοιες Δειγματου Χώρου Πιθανότητας → Μοντελοποίηση

2| Κλασσική Πιθανότητα → Συνδυαστική
#ευνοϊκών, #δυνατών

3| Πειράματα Τύχης σε Στάδια ⊛ → Υπολογισμοί Πιθ. με Πολλαπλ. Νόμο
Θεώρημα Ολ.Πιθ.
τύπος Bayes

4| Διακριτές κ' συνεχείς τ.μ. ⊛ → Υπολογισμοί Πιθανοτ. E, Var, κλ.π. όταν ή σ.π. η σ.π.π. είναι γνωστή

5| Ιδιότητες E, Var, Cov, ρ ⊛ → Υπολογισμοί με Θεωρ. Διπλής Μέσης Τιμής κτλ.

6| Ειδικές κατανομές → Υπολογισμοί

7| Πιθανογεννήτριες
Ροπογεννήτριες → Υπολογισμοί (π.χ. ροπογεν. αθρ.)
Θεωρητικό

8| Ανισότητες
Νόμος Μεγάλων Αριθμών
ΚΟΘ ⊛ → Υπολογισμοί Προσεγγιστικοί Πιθαν. με βάση το Κ.Ο.Θ.