22/12/2017

Κεντρικό Οριακό Θεώρημα
Εφαρμογές - Ασκήσεις

1. ΚΟΘ

$X_1, X_2, \dots$ ανεξ ισόνομες τμ, $E[X_i] = \mu$, $Var[X] = \sigma^2$

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

$$\lim_{n\to\infty} P\left(\frac{S_n - n\mu}{\sqrt{n}\,\sigma} \le x \right) = \Phi(x) = P(\overset{N(0,1)}{\tilde{Z}} \le x)$$

και

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

2. Χρήση του ΚΟΘ

Μας ενδιαφέρει ο προσεγγιστικός υπολογισμός $P(a \le S_n \le b)$ για μεγάλα n $(n \ge 30)$

$$P(a \le S_n \le b) = P\left(\frac{a - n\mu}{\sqrt{n}\,\sigma} < \frac{S_n - n\mu}{\sqrt{n}\,\sigma} < \frac{b - n\mu}{\sqrt{n}\,\sigma} \right) \overset{\text{ΚΟΘ}}{\simeq} P\left(\frac{a - n\mu}{\sqrt{n}\,\sigma} < \overset{N(0,1)}{\tilde{Z}} < \frac{b - n\mu}{\sqrt{n}\,\sigma} \right)$$

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

3. Διόρθωση συνέχειας

Αν $X_1, X_2, \dots$ ακέραιες τ.μ. και $a, b \in \mathbb{Z}$

τότε,

$$P(a \le S_n \le b) = P(a - x \le S_n \le b + y) \ , \ x, y \in [0,1]$$

Έχω ελευθερία να προσεγγίσω την πιθανότητα μέσω του ΚΟΘ με πολλούς τρόπους:

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

Στην περίπτωση αυτή, η βέλτιστη προσέγγιση είναι για $x = y = 1/2$, δηλ.

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 τώρα εφαρμόζω ΚΟΘ.

4. Άσκηση

Παίκτης ρίχνει ζάρι Y ζαριά: χάνει Y, $Y = 1, 2, 3$ / κερδίζει $7 - Y$, $Y = 4, 5, 6$

Προσεγγ. υπολογισμός σε 42 ρίψεις να έχει συνολικό κέρδος τουλ. 7.

Έστω X_i = κέρδος από την i ζαριά

y	1	2	3	4	5	6
x	-1	-2	-3	3	2	1
$P(X=x)$	1/6	1/6	1/6	1/6	1/6	1/6

$P(S_{42} \geq 7)$, οπου $S_{42} = \sum_{i=1}^{42} X_i$, X_i ανεξ ισόνομες

$42 \geq 30$ αρα το ΚΟΘ είναι εφαρμόσιμο

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

$Var[X_i] = E[X_i^2] - (\cancel{E[X]})^2 = (-1)^2\frac{1}{6} + (-2)^2\frac{1}{6} + (-3)^2\frac{1}{6} + 3^2\frac{1}{6} + 2^2\frac{1}{6} + 1^2\frac{1}{6} = \frac{28}{6} = \frac{14}{3}$

Αρα,

$P(S_{42} \geq 7) = P(S_{42} \geq 6,5) = P\left(\frac{S_{42} - 42\cdot 0}{\sqrt{42\cdot 14/3}} \geq \frac{6.5 - 42\cdot 0}{\sqrt{42\cdot 14/3}}\right)$

Διορθωση συνεχειας

$\overset{ΚΟΘ}{\simeq} P\left(\underset{\sim N(0,1)}{Z} \geq \frac{6.5}{14}\right) = 1 - \Phi\left(\frac{6.5}{14}\right)$

5. Άσκηση (προσοχη!)

Δειγμα 49 λαμπτηρων

T_i = Ο χρόνος ζωής ενός λαμπτηρα $\sim Exp$, $E[T_i] = 200$ ώρες

$P_1 = P$(Συνολ. χρόνος ζωης ≥ 10.000 ωρες) = ;

$P_2 = P$(Το πολύ 14 λαμπτηρες από τους 49 να ζήσουν λιγότερο από 140 ωρες) = ;

$P_1 = P(S_{49} \geq 10.000)$, $S_{49} = \sum_{i=1}^{49} T_i$

Η T_i έχει σ.π.π. $f_{T_i}(t) = \begin{cases} \frac{1}{200} e^{-t/200} & , t > 0 \\ 0 & , διαφορ \end{cases}$

(Εχω χρησιμοποιησει ότι: $T \sim Exp(\lambda)$, $E[T] = \frac{1}{\lambda}$, $Var[T] = \frac{1}{\lambda^2}$
$f_T(t) = \begin{cases} \lambda e^{-\lambda t} & , t > 0 \\ 0 & , διαφ \end{cases}$)

$E[T_i] = 200$

$Var[T_i] = 200^2 = 40.000$

δεν μπορώ να κανω διόρθωση συνεχειας γιατι είναι συνεχής

$P_1 = P(S_{49} \geq 10.000) = P\left(\frac{S_{49} - E[S_{49}]}{\sqrt{Var[S_{49}]}} \geq \frac{10.000 - E[S_{49}]}{\sqrt{Var[S_{49}]}}\right) \underset{49 \geq 30}{\overset{ΚΟΘ}{\sim}}$

$P\left(Z \geq \frac{10.000 - 49\cdot 200}{\sqrt{49\cdot 40.000}}\right) = P\left(Z \geq \frac{10.000 - 9.800}{7\cdot 200}\right) = P\left(Z \geq \frac{1}{7}\right)$

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

X = # λαμπτηρων που ζουν λιγότερα από 140 ώρες

$P_2 = P(X \leq 14)$

$X \sim Bin(49, p)$

$p = P$(χρ. ζωής λαμπτήρα ≤ 140 ωρες) (*)

$$X = \underbrace{\sum_{i=1}^{49} I_i}_{S_{49}} \quad , \quad I_i = \begin{cases} 1, & \text{ο } i \text{ λαμπτήρας ζησει} \leq 140 \text{ ωρες} \\ 0, & \text{διαφ.} \end{cases}$$

Άρα:

$$P_2 = \sum_{i=1}^{14} \binom{49}{k} p^k (1-p)^{49-k}$$

ακριβής τύπος (κακός υπολογιστικά)

Προσεγγιστικά, έχω:

$$P_2 = P(X \leq 14) = P(S_{49} \leq 14) \overset{\text{διόρθωση συνέχειας}}{=} P(S_{49} \leq 14.5) = P\left(\frac{S_{49} - E[S_{49}]}{\sqrt{Var[S_{49}]}} \leq \frac{14.5 - E[S_{49}]}{\sqrt{Var[S_{49}]}} \right)$$

$$\underset{49 \geq 30}{\overset{\text{κοθ}}{\sim}} P\left(Z \leq \frac{14.5 - E[S_{49}]}{\sqrt{Var[S_{49}]}} \right) = \Phi\left(\quad \right)$$

Έχω: $E[S_{49}] = np$

$Var[S_{49}] = np(1-p)$

(*) $= F_{T_i}(140) = 1 - e^{-\frac{140}{200}} \simeq 0.5 \quad \left(T_i \sim Exp(\lambda) \Rightarrow F_{T_i}(t) = 1 - e^{-\lambda t}, \ t > 0 \right)$

άρα, $E[S_{49}] = 49 \cdot 0,5 = 49/2$

$Var[S_{49}] = 49/4$

και

$$P_2 = \Phi\left(\frac{14.5 - \frac{49}{2}}{\sqrt{49/4}} \right) = \Phi\left(-\frac{20}{7} \right) = 1 - \Phi\left(\frac{20}{7} \right) = \ldots$$

6. Άσκηση

Πόλη 4.000 κατοίκων

10 άτομα/ημέρα χρειάζονται νοσηλεία κατα μέσο όρο

Προσεγγ. υπολογισμός ελάχιστου # κλινών ωστε να εξυπηρ. η πόλη με πιθ. 95%.

P(η πόλη εξυπηρετείται) ≥ 0.95

$$P(\underbrace{\text{ο αριθμός των ασθ. που χρειάζονται νοσηλεία σε μια συγκ μέρα}}_{X \sim Bin(4000,\, p)} \leq c) \geq 0.95$$

c: ο ζητούμενος αριθμός κλινών

όπου $p = P$(ένας κάτοικος της πόλης χρειάζεται νοσηλεία μια συγκ μέρα)

Όμως, $E[X] = 10 = 4000 \cdot p \rightarrow p = \frac{1}{400}$

Άρα $X = S_{4000} = \sum_{i=1}^{4000} I_i$

όπου,

$$I_i = \begin{cases} 1, & \text{αν ο } i \text{ κατοικος χρειαζ. νοσηλεια} \\ 0, & \text{διαφορ.} \end{cases}$$

$E[I_i] = \frac{1}{400}$

$Var[I_i] = \frac{1}{400} \cdot \left(1 - \frac{1}{400}\right) = \frac{399}{400^2}$

Έχω:

$P(S_{4000} \le c) \geq 0.95$

$P\left(\frac{S_{4000} - E[S_{4000}]}{\sqrt{Var[S_{4000}]}} \le \frac{c - E[S_{4000}]}{\sqrt{Var[S_{4000}]}}\right) \geq 0.95$

$\overset{\text{ΚΟΘ}}{\Longleftrightarrow}$

$P\left(Z \le \frac{c - 4000 \cdot \frac{1}{400}}{\sqrt{4000 \; 399/400^2}}\right) \geq 0.95$

$P\left(Z \le \frac{c - 10}{\sqrt{9.9}}\right) \geq 0.95$

$\Leftrightarrow \Phi\left(\frac{c-10}{\sqrt{9.9}}\right) \geq 0.95$

$\Leftrightarrow \Phi\left(\frac{c-10}{\sqrt{9.9}}\right) \geq \Phi(1.96)$

$\overset{\Phi \uparrow}{\Longleftrightarrow} \frac{c-10}{\sqrt{9.9}} \geq 1,96$

$\Leftrightarrow c \geq 1.96\sqrt{9.9} + 10$

$\Leftrightarrow c_{min} = 16$

7. Σύνοψη εφαρμογών του ΚΟΘ

2 περιπτώσεις:

→ Άμεση Εφαρμογή $P(a \le S_n \le b) \xrightarrow{\text{κανονικοποιώ}} : \Phi\left(\frac{b-n\mu}{\sqrt{n}\sigma}\right) - \Phi\left(\frac{a-n\mu}{\sqrt{n}\sigma}\right)$

+ Διόρθωση συνέχειας για τις διακοπές

→ Έμμεση Εφαρμογή

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

$X \sim Bin(n,p)$

$X = \sum_{i=1}^{n} I_i$

καπ

3 τυποι προβλημάτων - ερωτημάτων:

(1) n, x γνωστά και ψάχνω $P(S_n \leq x) =$;

a, b γνωστά » $P(a \leq S_n \leq b) =$;

(2) n, p γνωστά » $P(S_n \leq x) \geq p$, $x =$;

Η διαδικασία είναι: $P(S_n \leq x) \geq p \quad \ldots \Leftrightarrow \Phi\left(\frac{x-b}{c}\right) \geq p \Leftrightarrow \Phi\left(\frac{x-b}{c}\right) \geq \Phi(a)$

$\Leftrightarrow \frac{x-b}{c} \geq a \Leftrightarrow x \geq \ldots$

(3) x, p γνωστά

ωστε $P(S_n \leq x) \geq p$ » $n =$;