MEM-262 Παραμετρική Στατιστική

Τμήμα Μαθηματικών και Εφ. Μαθηματικών, Πανεπιστήμιο Κρήτη

Διδάσκων : Κώστας Σμαραγδάκης

Διάλεξη 4 : 13-10-2020

Ισόνομες και ανεξάρτητες τυχαίες μεταβλητές

Συμβολισμός : $i.i.d$ = independent and identically distributed

$$X_1, X_2, \ldots, X_n \quad i.i.d$$

δηλαδή,

$$F_{X_1}(x) = F_{X_2}(x) = \cdots = F_{X_n}(x)$$

$$F_{X_1 X_2 \ldots X_n}(x_1, x_2, \ldots, x_n) = \prod_{j=1}^{n} F_{X_j}(x_j)$$

- Εάν $X_1, X_2, \ldots, X_n$ είναι $i.i.d$ τότε $g(X_1), g(X_2), \ldots, g(X_n)$ είναι επίσης $i.i.d$

Ορισμός : $X_1, X_2, \ldots \quad i.i.d$

Λέμε ότι μια ακολουθία $X_1, X_2, \ldots$ είναι $i.i.d$ ανν για κάθε $n \in \{2, 3, \ldots\}$ και διαφορετικούς μεταξύ τους δείκτες $k_1, k_2, \ldots, k_n \in \mathbb{N}$ οι τυχαίες μεταβλητές $X_{k_1}, X_{k_2}, \ldots, X_{k_n}$ είναι $i.i.d$.

Εστώ $\{X_n\}_{n=1}^{\infty}$ ακολουθία τυχαίων μεταβλητών στο χώρο πιθανότητας $(\Omega, \mathcal{G}, P)$.

Ορισμός : Συγκλίση σχεδόν βεβαίως

$\{X_n\}_{n=1}^{\infty}$ συγκλίνει σε μια τυχαία μεταβλητή X στον ίδιο χώρο πιθανότητας **σχεδόν βεβαίως** και γράφουμε $X_n \xrightarrow{a.s.} X$ αν

$P(\omega \in \Omega : \lim X_n(\omega) \neq X(\omega)) = 0$

$$P(\lim_{n\to\infty} X_n = X) = P(\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)) = 1$$

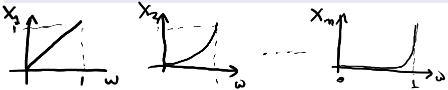
Παράδειγμα : $X_n(\omega) = \omega^n, \quad \omega \in \Omega = [0,1]$

$\mathcal{G} = \mathcal{B}([0,1])$ $\quad P((\alpha,b)) = b-\alpha$

$P([0,1/2]) = 1/2 - 0 = 1/2$

$[0,2]$

$P((\alpha,b)) = \frac{b-\alpha}{2}$



$$\lim_{n\to\infty} X_n(\omega) = \begin{cases} 0 & \text{αν } 0 \le \omega < 1 \\ 1 & \text{αν } \omega = 1 \end{cases} \qquad X(\omega) = 0$$

$$P(\omega \in [0,1] : \lim_{n\to\infty} X_n(\omega) = 0) = P([0,1)) = 1 - 0 = 1 \qquad X_n \xrightarrow{a.s.} X$$

$P(\{\omega : X_n(\omega) = 0\}) = 1$

$\forall \varepsilon > 0 \quad \exists m \quad$ T.ω $\quad P(\{\omega : |X_n(\omega) - X(\omega)| < \varepsilon \quad \forall n \ge m\}) = 1$

$P(\{\omega : |X_n(\omega) - X(\omega)| \ge \varepsilon \quad \forall n \ge m\}) = 0$

Εστώ $\{X_n\}_{n=1}^{\infty}$ ακολουθία τυχαίων μεταβλητών στο χώρο πιθανότητας $(\Omega, \mathcal{G}, P)$.

Ορισμός : Συγκλίση κατά πιθανότητα

$\{X_n\}_{n=1}^{\infty}$ συγκλίνει σε μια τυχαία μεταβλητή στον ίδιο χώρο πιθανότητας **κατά πιθανότητα** και γράφουμε $X_n \xrightarrow{p} X$ αν

$P(\omega : |X_n(\omega) - X(\omega)| \geq \varepsilon) \to 0$

$$\text{Για κάθε } \epsilon > 0 \text{ ισχύει } \lim_{n\to\infty} P(|X_n - X| \geq \epsilon) = 0$$

Νόμος των μεγάλων αριθμών

$|\mu| < \infty$

Έστω $X_1, X_2, \ldots$ *i.i.d* με πεπερασμένη μέση τιμή μ και διασπορά σ^2, τότε:

$\sigma^2 = Var(X_1) = Var(X_2) = \ldots = Var(X_n)$

$$\bar{X}_n \xrightarrow{p} \mu,$$

$\mu = \mathbb{E}(X_1) = \mathbb{E}(X_2) = \ldots = \mathbb{E}(X_n)$

όπου $\bar{X}_n = \dfrac{\sum_{i=1}^{n} X_i}{n}$. $= \dfrac{X_1 + X_2 + \ldots + X_n}{n} \xrightarrow{p} \mathbb{E}(X_1)$

Νόμος των μεγάλων αριθμών

$$\bar{X}_n = \frac{\sum_{i=1}^n X_i}{n}, \quad \bar{X}_n \xrightarrow{p} \mu \qquad \underline{\sigma^2 < \infty}$$

$$\mathbb{E}(\bar{X}_n) = \mathbb{E}\left\{\frac{1}{n}\sum_{i=1}^n X_i\right\} = \frac{1}{n}\mathbb{E}\left\{\sum_{i=1}^n X_i\right\} = \frac{1}{n}\sum_{i=1}^n \mathbb{E}\{X_i\} = \frac{1}{n} n\, \mathbb{E}(X_1) = \mu$$

$$\mathrm{Var}(\bar{X}_n) = \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^n X_i\right) = \frac{1}{n^2}\sum_{i=1}^n \mathrm{Var}(X_i) = \frac{1}{n^2} n\, \sigma^2 = \frac{\sigma^2}{n}$$

Έστω $\varepsilon > 0$ Chebyshev

$$\mathbb{P}(|\bar{X}_n - \mu| \geq \varepsilon) \leq \frac{\mathrm{Var}(\bar{X}_n)}{\varepsilon^2} = \frac{1}{n}\frac{\sigma^2}{\varepsilon^2} \xrightarrow{n\to\infty} 0$$

$$P(|\bar{X}_n - \mu| \geq \varepsilon) \longrightarrow 0 \quad \text{άρα} \quad \bar{X}_n \xrightarrow{P} \mu$$

Εστώ $\{X_n\}_{n=1}^{\infty}$ ακολουθία τυχαίων μεταβλητών στο χώρο πιθανότητας $(\Omega, \mathcal{G}, P)$.

Ορισμός : Συγκλίση κατά κατανομή

$\{X_n\}_{n=1}^{\infty}$ συγκλίνει σε μια τυχαία μεταβλητή στον ίδιο χώρο πιθανότητας **κατά κατανομή** και γράφουμε $X_n \xrightarrow{d} X$ αν

$$\text{Για κάθε σημείο συνέχειας της } F_X \text{ ισχύει } \lim_{n\to\infty} F_{X_n}(x) = F_X(x)$$

Παράδειγμα $X_n \in (0, n]$ $\quad F_{X_n}(x) = 1 - \left(1 - \frac{x}{n}\right)^n \quad , 0 < x \leq n$

ΝΔΟ $\quad X_n \xrightarrow{d} X$

$\lim_{x\to\infty} F_X(x) = 1$

$F_{X_n}(x) \xrightarrow{n\to\infty} 1 - e^{-x} = F_X(x)$ $\quad X \sim Exp(1)$

$f_X(x) = F_X'(x) = e^{-x}$

$f_X(x) = 0 \quad \forall x > 1$

Άσκηση : $X_n \in [0,1]$ $\quad F_{X_n}(x) = x - \frac{\sin(2n\pi x)}{2n\pi} \quad , 0 \leq x \leq 1$

$F_{X_n}(x) = F_{X_n}(1), x > 1$

Σύγκλιση τυχαίων μεταβλητών

Πρόταση : $X_n \xrightarrow{a.s.} X \Rightarrow X_n \xrightarrow{p} X$

$\varepsilon > 0$

$A_n = \{\omega : |X_n(\omega) - X(\omega)| < \varepsilon\} = \{|X_n - X| < \varepsilon\}$

$B_n = \{\omega : |X_k(\omega) - X(\omega)| < \varepsilon \quad \forall k \geq n\}$ $\qquad B_{n+1} = \{\omega : \overbrace{|X_k(\omega) - X(\omega)| < \varepsilon}\quad \forall k \geq n+1\}$

$B_n \subset A_n \qquad B_n \subset B_{n+1} \qquad B_1 \subset B_2 \subset B_3 \subset \cdots \qquad B = \bigcup_{n=1}^{\infty} B_n$

$\lim_{n\to\infty} P(B_n) = P(B)$

$X_n \xrightarrow{a.s.} X \qquad \exists\, m > 0 \quad \tau\omega \quad P(\overbrace{\{|X_n(\omega) - X(\omega)| < \varepsilon \quad \forall n \geq m\}}^{C}) = 1$

$C \subseteq B \qquad P(C) = 1 \Rightarrow P(B) = 1$

$1 = P(C) \leq P(B) \leq P(\Omega) = 1$

$B_n \subset A_n \subset \Omega \qquad \lim_{n\to\infty} P(A_n) = 1 \qquad \lim_{n\to\infty} P(A_n^c) = 0 \qquad X_n \xrightarrow{p} X$

$A = (A \cap B) \cup (A \cap B^c)$

Πρόταση : $X_n \xrightarrow{p} X \Rightarrow X_n \xrightarrow{d} X$

$\lim_{n\to\infty} P(|X_n - X| > \varepsilon) = 0$

$P(A) = P(A \cap B) + P(A \cap B^c)$

$X_n \xrightarrow{P} X \Leftrightarrow \forall \varepsilon > 0, \delta > 0 \quad \exists n_0 : P(|X_n - X| > \varepsilon) < \delta \quad \forall n \geq n_0$

$P(A \cap B) \leq P(B)$

$\{X_n - X < -\varepsilon\} \cup \{X_n - X > \varepsilon\}$

$F_{X_n}(x) = P(X_n \leq x) = P(X_n - X \leq x - X) =$

$= P(\{X_n - X \leq x - X\} \cap \{X_n - X \geq -\varepsilon\}) + P(\{X_n - X \leq x - X\} \cap \{X_n - X < -\varepsilon\}) \leq$

$\leq P(-\varepsilon \leq x - X) + P(X_n - X < -\varepsilon) \leq P(X \leq x + \varepsilon) + P(|X_n - X| > \varepsilon) =$

$\leq F_X(x + \varepsilon) + \delta$

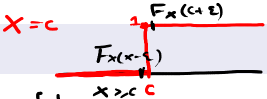
όμοια : $F_X(x - \varepsilon) \leq F_{X_n}(x) + \delta$

$$F_X(x - \varepsilon) - \delta \leq F_{X_n}(x) \leq F_X(x + \varepsilon) + \delta$$

$\varepsilon, \delta \to 0 \qquad F_{X_n}(x) = F_X(x) \qquad X_n \xrightarrow{d} X$

Σύγκλιση τυχαίων μεταβλητών

Πρόταση : $X_n \xrightarrow{d} c \in \mathbb{R} \Rightarrow X_n \xrightarrow{p} c$



$$X_n \xrightarrow{d} c \qquad F_{X_n}(x) \longrightarrow F(x) = \begin{cases} 1, & x \geq c \\ 0, & x < c \end{cases}$$

$$P(|X_n - c| > \varepsilon) = P(\{X_n - c > \varepsilon\} \cup \{X_n - c < -\varepsilon\}) =$$

$$= P(X_n > c+\varepsilon) + P(X_n < c-\varepsilon) =$$

$$1 - P(X_n \leq c+\varepsilon)$$

$$= 1 - F_{X_n}(c+\varepsilon) + F_{X_n}(c-\varepsilon) \xrightarrow{n\to\infty} 1 - 1 + 0 = 0$$

$$F_{X_n}(c+\varepsilon) \to F_X(c+\varepsilon) \qquad F_{X_n}(c-\varepsilon) \to F_X(c-\varepsilon)$$

$$X_n \xrightarrow{P} c$$

Θεώρημα Slutsky

Έστω $X_n \xrightarrow{d} X, \;\; Y_n \xrightarrow{p} c \in \mathbb{R}$ και g συνεχής συνάρτηση τότε

$$g(X_n, Y_n) \xrightarrow{d} g(X, c)$$

Ειδικές περιπτώσεις

Εάν $X_n \xrightarrow{d} X, \;\; Y_n \xrightarrow{p} c \in \mathbb{R}$ και

- $g(x, y) = x + y$ τότε $X_n + Y_n \xrightarrow{d} X + c$
- $g(x, y) = xy$ τότε $X_n Y_n \xrightarrow{d} cX$
- $g(x, y) = x/y$ και $c \neq 0$ τότε $X_n/Y_n \xrightarrow{d} X/c$
- $g(x, y) = (x, y)^T$ τότε $(X_n, Y_n)^T \xrightarrow{d} (X, c)^T$

Έστω $X_n \xrightarrow{d} X$, $Y_n \xrightarrow{p} c \in \mathbb{R}$ και g συνεχής συνάρτηση τότε

$c-\varepsilon \le Y_n \le c+\varepsilon$

$$X_n + Y_n \xrightarrow{d} X + c$$

$$F_{X_n+Y_n}(z) = P(X_n + Y_n \le z) =$$

$$= P(\{X_n+Y_n \le z\} \cap \{|Y_n - c| \le \varepsilon\}) + P(\{X_n + Y_n \le z\} \cap \{|Y_n - c| > \varepsilon\}) \le$$

$X_n + c - \varepsilon \le$

$$\le P(\{X_n \le z - c + \varepsilon\}) + P(|Y_n - c| > \varepsilon)$$

ομοια :

$$F_{X_n}(z-c-\varepsilon) \le F_{X_n+Y_n}(z) + P(|Y_n - c| > \varepsilon)$$

$$F_{X_n}(z-c-\varepsilon) - P(|Y_n - c| > \varepsilon) \le F_{X_n+Y_n}(z) \le F_{X_n}(z-c+\varepsilon) + P(|Y_n-c| > \varepsilon)$$

$$F_{X_n}(z-c-\varepsilon) \le \liminf_{n\to\infty} F_{X_n+Y_n}(z) \le \limsup_{n\to\infty} F_{X_n+Y_n}(z) \le F_{X_n}(z-c+\varepsilon)$$

$\varepsilon \to 0$

$$F_{X_n+Y_n}(z) = F_{X_n}(z-c) = F_{X_n+c}(z) \qquad X_n + Y_n \xrightarrow{d} X + c$$

Θεώρημα Mann-Wald

Έστω ακολουθία τυχαίων μεταβλητών $X_1, X_2, \ldots$ στον ίδιο χώρο πιθανότητας και g συνεχής συνάρτηση τότε:

$$X_n \xrightarrow{a.s.} X \Rightarrow g(X_n) \xrightarrow{a.s.} g(X)$$

$$X_n \xrightarrow{p} X \Rightarrow g(X_n) \xrightarrow{p} g(X)$$

$$X_n \xrightarrow{d} X \Rightarrow g(X_n) \xrightarrow{d} g(X)$$

Σύγκλιση της δειγματικής διασποράς

Θα δείξουμε ότι

$$\frac{\sum_{k=1}^{n}(X_k - \bar{X}_n)^2}{n-1} \xrightarrow{p} \sigma^2$$