Τετάρτη 24/02/21

Θεώρημα Glivenko - Contelli :

Η εμπειρική α.σ.κ συγκλίνει σχεδόν βέβαια στην $F(x)$, δηλαδή,

$$\sup_{x \in \mathbb{R}} |\hat{F}(x) - F(x)| \xrightarrow{a.s} 0$$

Η σχεδόν βέβαιη σύγκλιση για την $\hat{F}(x)$ σημαίνει ότι

η $P\left(\lim_{n \to \infty} \sup_{x \in \mathbb{R}} |\hat{F}(x) - F(x)| = 0\right) = 1$

Από τον ισχυρό νόμο των μεγάλων αριθμών έχουμε ότι για οποιοδήποτε (αλλά σταθερό) $x \in \mathbb{R}$ .. $P(\lim_{n\to\infty} \hat{F}(x) = F(x)) = 1$

Στην ουσία το supremum της διαφοράς μεταξύ της εμπειρικής α.σ.κ και της $F(x)$ στην εκφώνηση του θεωρήματος είναι το μέγιστο των διαφορών $\hat{F}(x_i) - F(x_i)$ για πεπ/νο αριθμό $i = 1, \dots m$.

Εφόσον η F είναι συνεχής μπορούμε να βρούμε m σημεία τ/ω $-\infty < x_0 < x_1 < \dots < x_m < +\infty$ και $F(x_j) - F(x_{j-1}) \leq \varepsilon$, $j \in (1, .. m)$ για δεδομένο $\varepsilon > 0$.

Έστω οποιοδήποτε σημείο $x \in \mathbb{R}$. Υπάρχει $j \in \{1, .. m\}$ τ/ω $x_{j-1} \leq x \leq x_j$

Εφόσον η Α.σ.κ είναι μη φθίνουσα έχουμε ότι

$$\hat{F}(x) - F(x) \leq \hat{F}(x_j) - F(x_j) \overset{(*)}{=} (\hat{F}(x_j) - F(x_j)) + F(x_j) - F(x_{j-1}))$$
$$\leq (\hat{F}(x_j) - F(x_j)) + \varepsilon \leq \max_{j \in \{0, \dots m\}} |\hat{F}(x_j) - F(x_j)| + \varepsilon.$$

Ομοίως έχουμε

$$\hat{F}(x) - F(x) \geq \hat{F}(x_{j-1}) - F(x_j) = (\hat{F}(x_{j-1}) - F(x_{j-1})) + (F(x_{j-1}) - F(x_j))$$
$$< (\hat{F}(x_{j-1}) - F(x_{j-1})) - \varepsilon \geq -\max_{j \in \{0 .. m\}} |\hat{F}(x_{j-1}) - F(x_{j-1})| - \varepsilon.$$

Συνδυάζοντας τις δύο παραπάνω ανισότητες

$$\sup |\hat{F}(x) - F(x)| \leq \varepsilon + \max_{j \in \{0, m\}} |\hat{F}(x_j) - F(x_j)|. \quad (2)$$

(*) $\hat{F}(x_j) - F(x_j) = \hat{F}(x_j) - F(x_j) + F(x_{j-1}) - F(x_{j-1})$

$x_{j-1} < x < x_j \Rightarrow F(x_{j-1}) < F(x) < F(x_j)$

Άρα $\hat{F}(x_j) - F(x_j) \leq \hat{F}(x_j) - F(x_j) + F(x) - F(x_{j-1})$.

Η συνέχεια της απόδειξης βασίζεται στην εφαρμογή του ισχυρού νόμου των μεγάλων αριθμών.

Έστω $A_j = \{\lim_{n\to\infty} |\hat{F}(x_j) - F(x_j)| \neq 0\}$

Ξέρουμε ήδη ότι $P(A_j) = 0$. Έστω

$A = \lim_{n\to\infty} \max_{j\in\{0..m\}} |\hat{F}(x_j) - F(x_j)| \neq 0\}$.

Παρατηρούμε ότι $\bigcup_{j=0}^{m} A_j = A$

$$P(A) = P\left(\bigcup_{j=0}^{m} A_j\right) \leq \sum_{j=0}^{m} P(A_j) = 0$$

Άρα $\max_{j\in\{0,..m\}} |\hat{F}(x_j) - F(x_j)| \xrightarrow{a.s.} 0$ (3)

Αντικαθιστώντας την (3) στην (2) παίρνουμε ότι

$$\limsup_{n\to\infty} \sup_{x\in\mathbb{R}} |\hat{F}(x) - F(x)| \leq \varepsilon$$

αν $\varepsilon \to 0$ προκύπτει το ζητούμενο

Εκτίμηση με ιστόγραμμα και a.σ.κ. στην $\mathbb{R}$.
όσο το δείγμα μεγαλώνει τόσο πιο καλή εκτίμηση έχουμε.