Διάλεξη 1

$$\begin{bmatrix} X \\ \text{σ.κ.}\ F \\ P \end{bmatrix} \xrightarrow{\text{τ.δ.}} X_1, \dots, X_\nu \longrightarrow \begin{matrix} \hat{X}_\nu \\ \hat{F}_\nu \\ \hat{P}_\nu \end{matrix}$$

(θεωρ. + άγν) (εκτίμηση)

[εισαγωγικά σχόλια]

- Έστω $X_1, \dots, X_\nu$ τ.δ. από άγνωστη **κατανομή** με σ.κ. F.

Ελλείψει άλλης πληροφορίας, θα χρησιμοποιούσαμε μόνο το δείγμα για να βγάλουμε κάποιο συμπέρασμα για την F. Χωρίς να κάνουμε παραμετρικές υποθέσεις, θέλουμε $\hat{F}_\nu$.

[$\hat{\theta}_\nu$ πεπερασμένης διάστασης, $\hat{F}_\nu$ είναι άπειρης διάστασης]

(μη παραμετρικά) → (προβλήματα που συναντάμε για 1η φορά χωρίς προηγ. εμπειρία).

[1η κατασκευή είναι αρκετά απλή:

Στη βάση του δείγματος που έχουμε, να φτιάξουμε μία τ.μ. που να προσεγγίζει την άγνωστη X:

$$\hat{X}_\nu = \begin{cases} X_1, & \text{με πιθ } \frac{1}{\nu} \\ \vdots & \\ X_\nu, & \text{με πιθ. } \frac{1}{\nu} \end{cases} \Rightarrow \hat{X}_\nu \sim \text{discr. Unif}(\{X_1, \dots, X_\nu\}).$$

[για να προσεγγίσουμε την X, βάζουμε σε μία κληρωτίδα τις τιμές που πήραμε, και εφόσον δεν υπάρχει λόγος να ευνοήσουμε κάποια από αυτές, επιλέγουμε τυχαία μέσα από αυτήν].

- Λέμε <u>εμπειρική κατανομή</u> την παραπάνω κατανομή, και <u>εμπειρική σ.κ.</u>, τη σ.κ. της εμπειρικής κατανομής

$$\forall x \in \mathbb{R}, \quad \mathbb{F}_\nu(x) = \frac{1}{\nu} \sum_{i=1}^{\nu} \mathbb{1}_{(-\infty, x]}(X_i) = \frac{\#\{i : X_i \le x\}}{\nu} \overset{*}{=}$$

(τυχαία συνάρτηση του x)

και γράφουμε $F_\nu(x) = \mathbb{F}_\nu(x)[\omega]$ για μία πραγματοποίηση.

($\overset{*}{=}$ ποσοστό παρατηρήσεων $\le x$)

<u>Παράδ</u>: $F \to \mathcal{N}(0,1)$: τ.δ. $X_1, \dots, X_\nu$ $\mathcal{N}(0,1)$

γράφημα (για $\nu = 10$) (page 1) [άλμα $\frac{1}{10}$ σε κάθε παρατήρηση, όρια 0 και 1 αντίστοιχα, μήκος διακεκομμ. $\frac{1}{10}$, στο επό. τμήμα όχι άλλη παρατήρηση.

Ως σ.κ. διακριτής τ.μ., είναι δεξιά συνεχής + ↗ (⇒ ∃ όρια από τα αριστερά).

$$P(X = x) = F(x) - F(x-).$$

— Προσθέτουμε τη θεωρητ. σ.κ. $\Phi(x)$ (page 2).

(2)

- <u>Εξετάζουμε την οριακή συμπεριφορά καθώς $\nu \to \infty$</u>

[Καθώς το ν-μεγαλώνει, θα θέλαμε η εμπειρική κατανομή που φτιάξαμε, να προσεγγίζει όλο καλύτερα την άγνωστη κατανομή και να μπορούμε να το οπτικοποιήσουμε μέσα απο τις σ.κ.]

(page 3 : γράφημα)

Σε ένα Π.Τ. μπορούμε να δούμε την εξέλιξη του γραφήματος της ε.σ.κ. για $\nu = 10$, $\nu = 100$, $\nu = 1000$, $\nu = 10.000$.

Έχουμε ένδειξη για τη σύγκλιση : $F_\nu(x) \longrightarrow F(x)$, $\forall x \in \mathbb{R}$ (για το συγκεκριμένο ω).

- <u>Εξετάζουμε και άλλα ω, μέσω διαφορ. Π.Τ.</u>

Βλέπουμε 4 διαφορετικά γραφήματα που αντιστοιχούν σε 4 διαφορετικές πραγματοποιήσεις (4 διαφ. Π.Τ.). Παρατηρούμε ότι η τάξη μεγέθους των αποκλίσεων είναι παρόμοια. (εκτός απο σημειακή σύγκλιση έχαμε ένδειξη για ομοιόμορφη σύγκλιση)

Για $\nu = 10$ και για $\nu = 10.000$ (τα γραφήματα σχεδόν συμπίπτουν)

(page 4)

$$\forall x \in \mathbb{R}, \quad \mathbb{F}_\nu(x) \xrightarrow{a.s.} F(x) \qquad \left(\begin{matrix}\text{ακολουθία} \\ \text{τ.μ} \xrightarrow{a.s.} \text{τ.μ.}\end{matrix}\right).$$

<u>Απόδειξη</u>

$$\mathbb{F}_\nu(x) = \frac{1}{\nu} \sum_{i=1}^{\nu} \underbrace{1_{\{X_i \le x\}}}_{g(X_i)} = \overline{g(X)_\nu} \underset{(I.N.M.A)}{\xrightarrow{a.s}} E[g(X)] = E\left[1_{\{X \le x\}}\right]$$

$$\overset{\|}{P(X \le x)} = F(x), \forall x \in \mathbb{R}$$

($\{g(X_i)\}_{i \ge 1}$ είναι ακολουθία α.ι.τ.μ., με $E(g(X)) = F(x)$)

Είναι απλό και χρήσιμο αποτέλεσμα, για να μπορέσουμε όμως να ελέγξαμε την καταλληλότητα μιας θεωρητικής κατανομής, χρειαζόμαστε να μελετήσαμε συνολικά τη σύγκλιση για όλα τα x [καθολικό μέτρο] να ισχυροποιήσουμε τη σύγκλιση για...

<u>Οι αποστάσεις στο γράφημα μειώνονται ομοιόμορφα</u>

[ένδειξη: $\sup_x |F_\nu(x) - F(x)| \to 0$]

Μπορούμε ν.δ.ο.

$$\sup_x |\mathbb{F}_\nu(x) - F(x)| =: \|\mathbb{F}_\nu - F\|_\infty \xrightarrow{a.s.} 0$$

[Θεώρημα Glivenko - Cantelli : ομοιόμορφη σύγκλιση της $\mathbb{F}_\nu$ στην F με πιθ. 1.

[Η ακολουθία των εμπειρικών σ.κ. συγκλίνει στη θεωρητική σ.κ ομοιόμορφα, με πιθ. 1] π.χ. αν θέσουμε (Σχόλια). ③

1) $$D_\nu = \| \mathbb{F}_\nu - F \|_\infty$$ τότε αυτό

εκφράζει την ομοιόμορφη απόσταση της ε.σ.κ. απο τη θεωρητική σ.κ.), είναι ένα ρυθμιστής της απόστασης απο την υποτιθ. θεωρητ. κατανομή και λέγεται στατιστικό <u>Kolmogorov – Smirnov</u> και χρησιμοποιείται για τον έλεγχο καλής προσαρμογής μιας υποψήφιας θεωρητ. κατανομής στα δεδομένα [page 5: γράφημα KS για απόσταση]

2) $$|\mathbb{F}_\nu(x) - F(x)| \xrightarrow{a.s.} 0 \quad (\forall x, \text{ απο I.N.M.A.}).$$

$$\sup_x |\mathbb{F}_\nu(x) - F(x)| \xrightarrow{a.s.} 0$$ (Το πιο βασικό παράδειγμα ομοιόμορφου I.N.M.A.)

$\boxed{\text{Θεώρημα G-C}}$ Αν $(X_\nu)_{\nu \geq 1}$ ακολουθία α.λ.τ.μ. με σ.κ. F, τότε

$$\| \mathbb{F}_\nu - F \|_\infty \xrightarrow{a.s.} 0$$

<u>Απόδειξη</u>

Έστω $\varepsilon > 0$. Η $F \uparrow$ με τιμές στο $[0,1]$ $\Longrightarrow$

$\exists$ διαμέριση $-\infty = x_0 < x_1 < \dots < x_{k-1} < x_k = +\infty$:

$$\boxed{F(x_i -) - F(x_{i-1}) \leq \varepsilon \;, \; \forall 1 \leq i \leq k} \quad (1)$$, για κάποιο $k \geq 1$

[π.χ. $t_0 = -\infty$, $t_i = \sup\{x > t_{i-1} : F(x) \leq i\varepsilon\}$, $i = 1, \dots, [\frac{1}{\varepsilon}]$, $t_{[\frac{1}{\varepsilon}]+1} = +\infty$.
Έστω $\{x_0, \dots, x_k\}$ τα διακεκριμένα t_i, $0 \leq i \leq [\frac{1}{\varepsilon}]+1 \Rightarrow F(x_i) \geq i\varepsilon$ και $F(x_i-) - F(x_{i-1}) \leq \varepsilon$]
$F(x_i-) \leq i\varepsilon$

Έστω $x \in \mathbb{R}$, τότε $\exists\, 1 \leq i \leq k$: $x \in [x_{i-1}, x_i)$ και

$$\mathbb{F}_\nu(x) - F(x) \leq \mathbb{F}_\nu(x_i -) - F(x_{i-1}) \overset{(1)}{\leq} \mathbb{F}_\nu(x_i -) - F(x_i -) + \varepsilon \leq \max_i (\,) + \varepsilon$$

$$F(x) - \mathbb{F}_\nu(x) \leq F(x_i -) - \mathbb{F}_\nu(x_{i-1}) \leq F(x_{i-1}) - \mathbb{F}_\nu(x_{i-1}) + \varepsilon \leq \max_i + \varepsilon$$

$$\Rightarrow |\mathbb{F}_\nu(x) - F(x)| \leq \left(\max_{1 \leq i \leq k} \{\mathbb{F}_\nu(x_i -) - F(x_i -)\} \vee \max_{1 \leq i \leq k} \{F(x_{i-1}) - \mathbb{F}_\nu(x_{i-1})\} \right) + \varepsilon \quad \forall x \in \mathbb{R}.$$

$$\Rightarrow \| \mathbb{F}_\nu - F \|_\infty = \sup_x |\mathbb{F}_\nu(x) - F(x)| \leq \left\{ \max_{1 \leq i \leq k} \{\mathbb{F}_\nu(x_i -) - F(x_i -)\} \vee \max_{1 \leq i \leq k} \{F(x_{i-1}) - \mathbb{F}_\nu(x_{i-1})\} \right\} + \varepsilon$$

$$\Rightarrow \limsup_\nu \| \mathbb{F}_\nu - F \|_\infty \overset{a.s.}{\leq} 0 + \varepsilon$$ $\left(\mathbb{F}_\nu(x_i -) - F(x_i -) \xrightarrow{a.s.} 0 \;\&\; \downarrow a.s.\ 0 \right)$.

$$\Rightarrow \limsup_\nu \| \mathbb{F}_\nu - F \|_\infty \overset{a.s.}{=} 0 \Rightarrow \lim_\nu \| \mathbb{F}_\nu - F \|_\infty \overset{a.s.}{=} 0 .$$

Παρατήρηση

- $F(x) = P\left(X \leq x\right) = E\left[\mathbb{1}_{(-\infty, x]}(X)\right] = E\left[f_x(X)\right]$, όπου

$$f_x(u) = \mathbb{1}_{(-\infty,x]}(u) = \begin{cases} 1, & u \leq x \\ 0, & u > x \end{cases}.$$

[μέση τιμή μιας οικογένειας συναρτήσεων].

Ορίζουμε $\mu f := E\left[f(X)\right]$, όταν $X \sim \mu$ (μ, μέτρο πιθανότητας)

Άρα $Pf = E\left[f(X)\right]$, όταν $X \sim P$

$\mathbb{P}_\nu f = E\left[f(X)\right]$, όταν $X \sim \mathbb{P}_\nu$ (εδώ $X = \hat{X}_\nu$).

Μάλιστα $\mathbb{P}_\nu f = \frac{1}{\nu} \sum_{i=1}^{\nu} f(X_i)$

$f \to f_x$, τότε $\underline{Pf_x = F(x)}$ και $\underline{\mathbb{P}_\nu f_x = \mathbb{F}_\nu(x)}$

Τελικά

$$\|\mathbb{F}_\nu - F\|_\infty = \sup_x \left|\mathbb{F}_\nu(x) - F(x)\right| = \sup_x \left|\mathbb{P}_\nu(f_x) - Pf_x\right|$$

$$= \sup_{f \in \mathcal{F}} \left|\mathbb{P}_\nu(f) - Pf\right| = \|\mathbb{P}_\nu - P\|_{\mathcal{F}}, \to$$ απόσταση μέτρων πιθανότητας.

όπου $$\mathcal{F} = \left\{ \mathbb{1}_{(-\infty, x]} : x \in \mathbb{R} \right\}$$

<u>Ιδέα</u>: να αναζητήσουμε γενικότερες κλάσεις συναρτήσεων, όπου η σύγκλιση αυτή $\mathbb{P}_\nu f \xrightarrow{a.s} Pf$, να είναι ομοιόμορφη στην κλάση $\mathcal{F}$.

Εξ'άλλου από $\underline{\text{I.N.M.A.}}$ $\mathbb{P}_\nu f = \frac{1}{\nu} \sum_{i=1}^{\nu} f(X_i) = \overline{f(X)}_\nu \xrightarrow{a.s.} E f(X) = Pf$,

όταν $E|f(X)| = P|f| < \infty$.

[Τα αφηρημένα θεωρήματα G-C κάνουν αυτή τη σύγκλιση ομοιόμορφη ως f, όταν $f \in \mathcal{F}$]

<u>Ορσ</u>: Μια κλάση $\mathcal{F}$- μετρησίμων συνάρτησεων $f : \mathcal{X} \to \mathbb{R}$ λέγεται P- Glivenko-Cantelli (ή P-GC) αν

$$\|\mathbb{P}_\nu - P\|_{\mathcal{F}} = \sup_{f \in \mathcal{F}} \left|\mathbb{P}_\nu f - Pf\right| \xrightarrow{a.s} 0.$$

(συγκλίνει ομοιόμορφα σ.β. στο 0).

Παρατήρηση

Για $r \geq 1$, $f \in \mathcal{L}^r(P) \iff \int |f|^r \, dP < \infty$ ή $P|f|^r < \infty$.

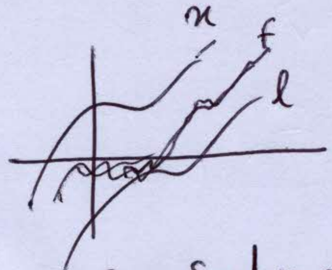
$X \sim P$, $E|f(X)|^r$

Ο ορισμός P-GC προϋποθέτει $P|f| < \infty$ ή $f \in \mathcal{L}^1(P)$.

Αναζητούμε ικανές συνθήκες

Έστω l, u 2 συναρτήσεις.

- $[l,u] \overset{\text{ορσ}}{=} \left\{ f : \mathcal{X} \to \mathbb{R} \;\middle|\; \underbrace{l(x)}_{\text{ομοιόμορφο κάτω φράγμα}} \leq f(x) \leq \underbrace{u(x)}_{\text{ομοιόμορφο άνω φράγμα}}, \forall x \in \mathcal{X} \right\}$.



Το λέμε **bracket** (υποστήριγμα) $[l,u]$. [βραχίονας? αγκύλη?]

- **ε-bracket** στον $\mathcal{L}^r(P)$ είναι ένα bracket $[l,u]$, όπου

$$P(u-l)^r \leq \varepsilon^r \quad \left(\iff \|u-l\|_{P,r} = \left(P(u-l)^r\right)^{1/r} \leq \varepsilon \right)$$

Για $r=1$, $P(u-l) \leq \varepsilon$ ή ισοδ. $\|u-l\|_{P,1} \leq \varepsilon$

- αριθμό bracketing (bracketing number) $N^r_{[\,]}(\varepsilon, \mathcal{F}, P)$, ονομάζουμε τον ελάχιστο αριθμό από ε-brackets που χρειαζόμαστε για να καλύψουμε την οικογένεια $\mathcal{F}$

Θεώρημα Glivenko-Cantelli

Κάθε οικογένεια $\mathcal{F}$-μετρησίμων συναρτήσεων με $N^1_{[\,]}(\varepsilon, \mathcal{F}, P) < \infty$, $\forall \varepsilon > 0$ είναι P-GC.

Απόδειξη

Έστω $\varepsilon > 0$. Από υπόθεση $N^1_{[\,]}(\varepsilon, \mathcal{F}, P) < \infty \Rightarrow \exists k \geq 1$, $l_1, \dots, l_k, u_1, \dots, u_k$: αν $f \in \mathcal{F}$, τότε $f \in [l_i, u_i]$ για κάποιο $1 \leq i \leq k$.

$$\begin{aligned}
\mathbb{P}_v f - Pf &= \mathbb{P}_v(l_i) + \mathbb{P}_v(f - l_i) + P(u_i - f) - Pu_i \\
&\leq \mathbb{P}_v(l_i) + \mathbb{P}_v(u_i - l_i) + \underbrace{P(u_i - l_i)}_{\leq \varepsilon} - Pu_i \leq \mathbb{P}_v u_i - Pu_i + \varepsilon \\
&\leq \max_{1 \leq i \leq k} \{ (\mathbb{P}_v - P) u_i \} + \varepsilon
\end{aligned}$$

$$\begin{aligned}
Pf - \mathbb{P}_v f &= Pl_i + P(f - l_i) + \mathbb{P}_v(u_i - f) - \mathbb{P}_v u_i \leq Pl_i + \underbrace{P(u_i - l_i)}_{\leq \varepsilon} + \mathbb{P}_v(u_i - l_i) - \mathbb{P}_v u_i \\
&\leq Pl_i - \mathbb{P}_v l_i + \varepsilon \leq \max_{1 \leq i \leq k} (P - \mathbb{P}_v) l_i + \varepsilon
\end{aligned}$$

$$\Rightarrow |\mathbb{P}_v f - Pf| \leq \left(\max_{1 \leq i \leq k} \{ (\mathbb{P}_v - P) u_i \} \vee \max_{1 \leq i \leq k} \{ (P - \mathbb{P}_v) l_i \} \right) + \varepsilon, \quad \forall f \in \mathcal{F}$$

$$\Rightarrow \sup_{f \in \mathcal{F}} |\mathbb{P}_v f - Pf| \leq A_v + \varepsilon \Rightarrow \limsup_v \|\mathbb{P}_v - P\|_{\mathcal{F}} \overset{a.s.}{\leq} \varepsilon \Rightarrow \|\mathbb{P}_v - P\|_{\mathcal{F}} \xrightarrow{a.s.} 0.$$