Εκτίμηση αθροιστικής συνάρτησης κατανομής

Μη Παραμετρική Στατιστική

Παναγιώτης Παπασταμούλης
Επίκουρος Καθηγητής
Τμήμα Στατιστικής ΟΠΑ

papastamoulis@aueb.gr

Πέμπτη, 26/3/2020

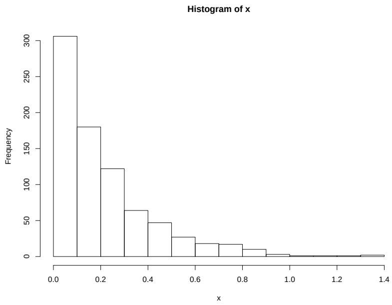
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ & ΤΕΧΝΟΛΟΓΙΑΣ ΤΗΣ ΠΛΗΡΟΦΟΡΙΑΣ
SCHOOL OF INFORMATION SCIENCES & TECHNOLOGY

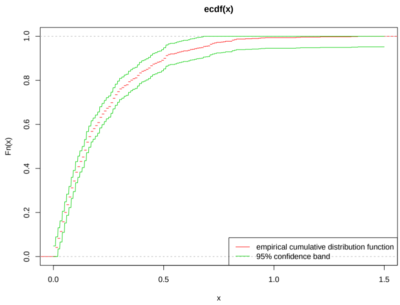
ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ
DEPARTMENT OF STATISTICS

Nerve data

- Cox and Lewis (1966)
- $n = 799$ χρόνοι αναμονής μεταξύ διαδοχικών συσπάσεων νευρικής ίνας



Ιστόγραμμα δεδομένων



Εμπειρική συνάρτηση κατανομής + 95% ζώνη εμπιστοσύνης

Εκτίμηση αθροιστικής συνάρτησης κατανομής

Έστω $X_1, \ldots, X_n$ τυχαίο δείγμα από κάποια (άγνωστη) συνάρτηση κατανομής F (διακριτή ή συνεχής).

Ορισμός (Εμπειρική συνάρτηση κατανομής)

Η εμπειρική συνάρτηση κατανομής ορίζεται ως

$$\widehat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathrm{I}\{X_i \leqslant x\}. \tag{1}$$

Στον παραπάνω ορισμό:

$$\mathrm{I}\{X_i \leqslant x\} = \begin{cases} 1, & \text{αν } X_i \leqslant x \\ 0, & \text{διαφορετικά.} \end{cases}$$

Παράδειγμα ($n = 10$)

- Δεδομένα:

$$(-0.90, 0.18, 1.59, -1.13, -0.08, 0.13, 0.71, -0.24, 1.98, -0.14)$$

- Διατεταγμένο δείγμα

$$(-1.13, -0.90, -0.24, -0.14, -0.08, 0.13, 0.18, 0.71, 1.59, 1.98)$$

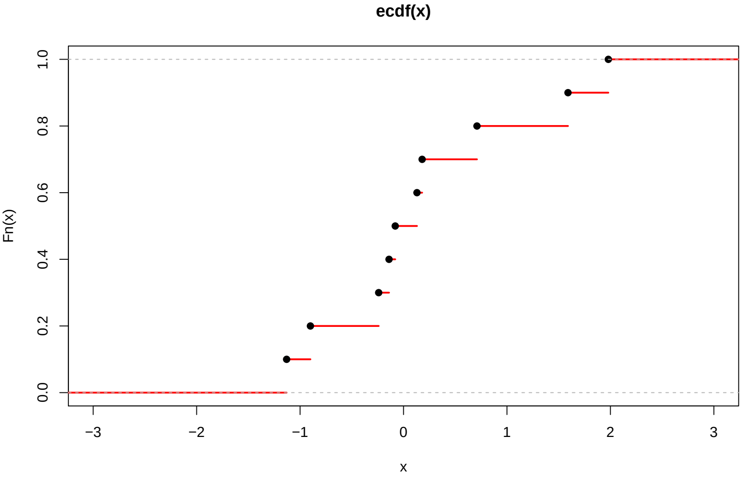
- Υπολογισμός $\widehat{F}_n(x)$

x	$\widehat{F}_n(x)$
$(-\infty, -1.13)$	0
$[-1.13, -0.90)$	$1/10$
$[-0.90, -0.24)$	$2/10$
$[-0.24, -0.14)$	$3/10$
$[-0.14, -0.08)$	$4/10$
$[-0.08, 0.13)$	$5/10$
$[0.13, 0.18)$	$6/10$
$[0.18, 0.71)$	$7/10$
$[0.71, 1.59)$	$8/10$
$[1.59, 1.98)$	$9/10$
$[1.98, \infty)$	1

- Παράδειγμα: $\widehat{F}_n(1.64) = 0.9$ (πραγματική τιμή: $F(1.64) \approx 0.95$)

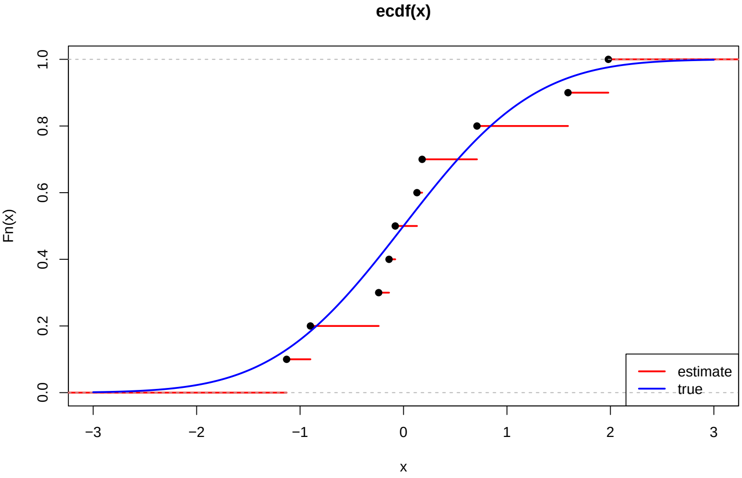
Παράδειγμα ($n = 10$)

Δεδομένα: $(-0.90, 0.18, 1.59, -1.13, -0.08, 0.13, 0.71, -0.24, 1.98, -0.14)$

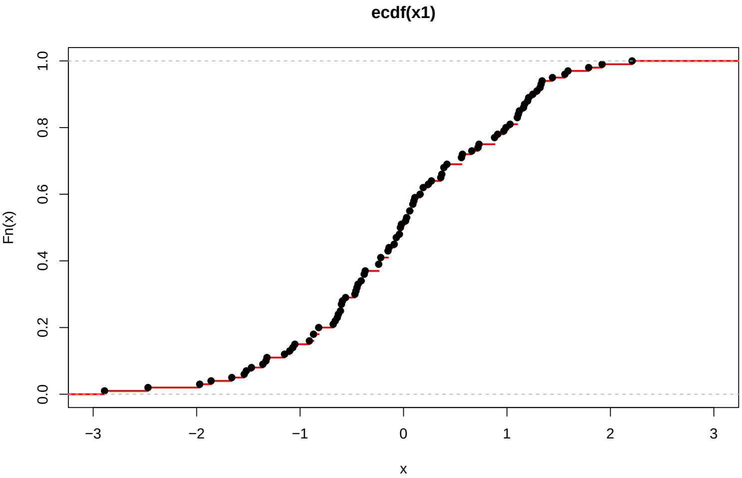


Παράδειγμα ($n = 10$)

$$\text{Δεδομένα}: \ (-0.90, 0.18, 1.59, -1.13, -0.08, 0.13, 0.71, -0.24, 1.98, -0.14)$$

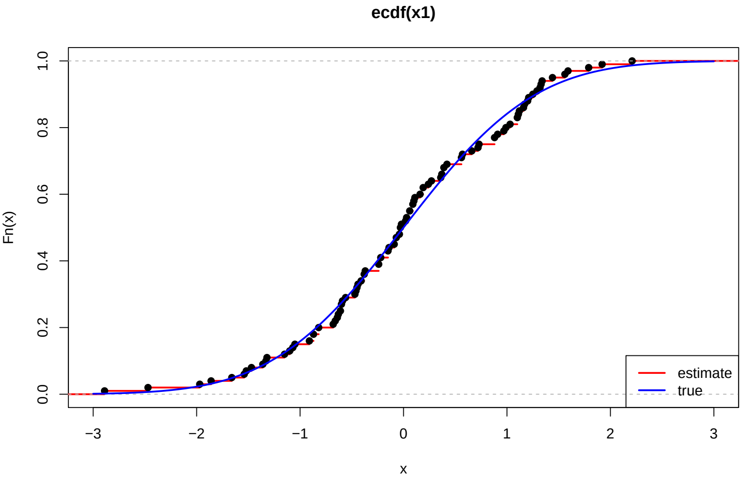


Παράδειγμα ($n = 100$)



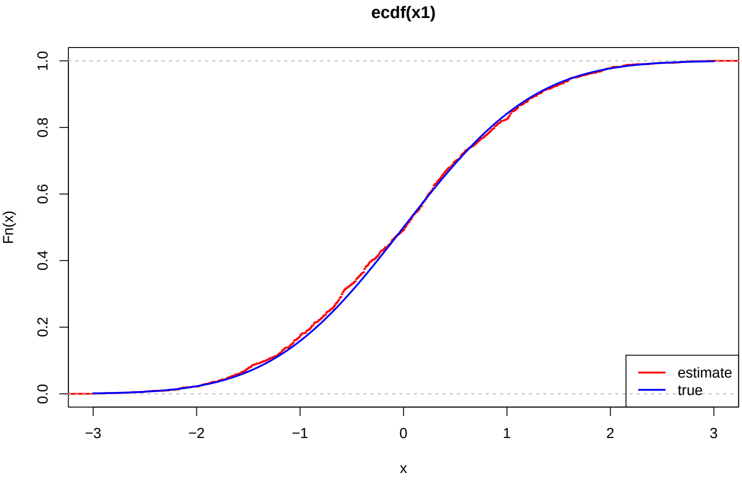
Παράδειγμα: $\widehat{F}_n(1.64) = 0.97$ (πραγματική τιμή: $F(1.64) \approx 0.95$)

Παράδειγμα ($n = 100$)



Παράδειγμα: $\widehat{F}_n(1.64) = 0.97$ (πραγματική τιμή: $F(1.64) \approx 0.95$)

Παράδειγμα ($n = 1000$)



Παράδειγμα: $\widehat{F}_n(1.64) = 0.95$ (πραγματική τιμή: $F(1.64) \approx 0.95$)

Ιδιότητες Εμπειρικής Συνάρτησης Κατανομής

1. **Μέση τιμή και διασπορά**

$$\mathrm{E}\widehat{F}_n(x) = F(x) \quad \text{και} \quad \mathrm{Var}\widehat{F}_n(x) = \frac{F(x)(1 - F(x))}{n}$$

για κάθε $x \in \mathbb{R}$.

2. **Θεμελιώδες θεώρημα Glivenko-Cantelli.** Για κάθε συνάρτηση κατανομής F ισχύει ότι:

$$\sup_{x\in\mathbb{R}} \left| \widehat{F}_n(x) - F(x) \right| \overset{a.s.}{\to} 0$$

3. **Ανισότητα DKW (Dvoretzky-Kiefer-Wolfowitz).** Για κάθε $\varepsilon > 0$:

$$\mathrm{P}\left(\sup_{x\in\mathbb{R}} \left| \widehat{F}_n(x) - F(x) \right| > \varepsilon \right) \leqslant 2e^{-2n\varepsilon^2}$$

(για όλα τα n).

Απόδειξη για το (1)

- Για δοθέν $x \in \mathbb{R}$, έστω η τυχαία μεταβλητή

$$Y = n\widehat{F}_n(x) = \sum_{i=1}^{n} \mathrm{I}\{X_i \leqslant x\} = \text{«πόσα } X \text{ από τα } n \text{ είναι } \leqslant x\text{»}.$$

Απόδειξη για το (1)

- Για δοθέν $x \in \mathbb{R}$, έστω η τυχαία μεταβλητή
$$Y = n\widehat{F}_n(x) = \sum_{i=1}^{n} \mathrm{I}\{X_i \leqslant x\} = \text{«πόσα } X \text{ από τα } n \text{ είναι } \leqslant x\text{»}.$$
- Ξεκάθαρα: $Y \sim \mathcal{B}(n, F(x))$.

Απόδειξη για το (1)

- Για δοθέν $x \in \mathbb{R}$, έστω η τυχαία μεταβλητή

$$Y = n\widehat{F}_n(x) = \sum_{i=1}^{n} \mathrm{I}\{X_i \leqslant x\} = \text{«πόσα } X \text{ από τα } n \text{ είναι} \leqslant x\text{»}.$$

- Ξεκάθαρα: $Y \sim \mathcal{B}(n, F(x))$.
- Συνεπώς για τη μέση τιμή

$$\begin{aligned} \mathrm{E}(Y) &= nF(x), \quad \forall x \in \mathbb{R} \Leftrightarrow \\ \mathrm{E}(n\widehat{F}_n(x)) &= nF(x), \quad \forall x \in \mathbb{R} \Leftrightarrow \\ \mathrm{E}(\widehat{F}_n(x)) &= F(x), \quad \forall x \in \mathbb{R} \end{aligned}$$

Απόδειξη για το (1)

- Για δοθέν $x \in \mathbb{R}$, έστω η τυχαία μεταβλητή

$$Y = n\widehat{F}_n(x) = \sum_{i=1}^{n} \mathrm{I}\{X_i \leqslant x\} = \text{«πόσα } X \text{ από τα } n \text{ είναι} \leqslant x\text{»}.$$

- Ξεκάθαρα: $Y \sim \mathcal{B}(n, F(x))$.
- Συνεπώς για τη μέση τιμή

$$\begin{aligned} \mathrm{E}(Y) &= nF(x), \quad \forall x \in \mathbb{R} \Leftrightarrow \\ \mathrm{E}(n\widehat{F}_n(x)) &= nF(x), \quad \forall x \in \mathbb{R} \Leftrightarrow \\ \mathrm{E}(\widehat{F}_n(x)) &= F(x), \quad \forall x \in \mathbb{R} \end{aligned}$$

- Διασπορά

$$\begin{aligned} \mathrm{Var}(Y) &= nF(x)(1 - F(x)), \quad \forall x \in \mathbb{R} \Leftrightarrow \\ n^2\mathrm{Var}(\widehat{F}_n(x)) &= nF(x)(1 - F(x)), \quad \forall x \in \mathbb{R} \Leftrightarrow \\ \mathrm{Var}(\widehat{F}_n(x)) &= \frac{F(x)(1 - F(x))}{n}, \quad \forall x \in \mathbb{R}. \end{aligned}$$

Σκιαγράφηση απόδειξης του (2)

Λήμμα

Η κατανομή της τυχαίας μεταβλητής $\sup_{x\in\mathbb{R}} \left|\widehat{F}_n(x) - F(x)\right|$ *δεν εξαρτάται από την* F *(αν* F *συνεχής).*

- Αν $X_i \sim F$, τότε $U_i := F(X_i) \sim \mathcal{U}(0,1)$
- Πράγματι, για $0 \leqslant t \leqslant 1$:

$$\begin{aligned} F_U(t) &:= \mathrm{P}(U_i \leqslant t) \\ &= \mathrm{P}(F(X_i) \leqslant t) \\ &= \mathrm{P}(X_i \leqslant F^{-1}(t)) \\ &= F\left(F^{-1}(t)\right) \\ &= t. \end{aligned}$$

Σκιαγράφηση απόδειξης του (2)

Συνεπώς

$$
\begin{aligned}
\sup_{x\in\mathbb{R}} \left|\widehat{F}_n(x) - F(x)\right| &= \sup_{t\in[0,1]} \left|\widehat{F}_n(F^{-1}(t)) - F(F^{-1}(t))\right| \\
&= \sup_{t\in[0,1]} \left|\frac{1}{n}\sum_{i=1}^{n} \mathrm{I}\{X_i \leqslant F^{-1}(t)\} - t)\right| \\
&= \sup_{t\in[0,1]} \left|\frac{1}{n}\sum_{i=1}^{n} \mathrm{I}\{F(X_i) \leqslant t\} - t)\right| \\
&= \sup_{t\in[0,1]} \left|\frac{1}{n}\sum_{i=1}^{n} \mathrm{I}\{U_i \leqslant t\} - t)\right| \\
&= \sup_{t\in[0,1]} \left|\widehat{G}_n(t) - t)\right|,
\end{aligned}
$$

όπου $\widehat{G}_n(\cdot)$ η εμπειρική συνάρτηση κατανομής των $U_i = F(X_i)$, $i = 1, \ldots, n$.

Συνέπεια του (3): Μία ζώνη εμπιστοσύνης για την F

Μέσω της ανισότητας DKW μπορούμε να κατασκευάσουμε μία **ζώνη εμπιστοσύνης** για τη συνάρτηση κατανομής (F).

- Έστω $\alpha \in (0, 1)$.
- Ορίζουμε

$$L_n(x) = \max\left\{\widehat{F}_n(x) - \sqrt{\frac{1}{2n}\log\frac{2}{\alpha}}, 0\right\}$$

$$U_n(x) = \min\left\{\widehat{F}_n(x) + \sqrt{\frac{1}{2n}\log\frac{2}{\alpha}}, 1\right\}$$

Λήμμα ($100(1-\alpha)\%$ DKW ζώνη εμπιστοσύνης για την F)

Γιά κάθε F και για κάθε n

$$\mathrm{P}(L_n(x) \leqslant F(x) \leqslant U_n(x) \text{ για κάθε } x) \geqslant 1 - \alpha.$$

Απόδειξη

Από την ανισότητα DKW για $\varepsilon_n^2 = \frac{1}{2n}\log\frac{2}{\alpha}$ προκύπτει ότι

$$\begin{aligned}
\mathrm{P}\left(\sup_{x\in\mathbb{R}}\left|\widehat{F}_n(x) - F(x)\right| > \varepsilon_n\right) &\leqslant 2e^{-2n\varepsilon_n^2} \Leftrightarrow \\
\mathrm{P}\left(\left|\widehat{F}_n(x) - F(x)\right| > \varepsilon_n\right) &\leqslant 2e^{-2n\frac{1}{2n}\log\frac{2}{\alpha}}, \quad \forall x \in \mathbb{R} \Leftrightarrow \\
\mathrm{P}\left(\left|\widehat{F}_n(x) - F(x)\right| > \varepsilon_n\right) &\leqslant \alpha, \quad \forall x \in \mathbb{R} \Leftrightarrow \\
\mathrm{P}\left(\left|\widehat{F}_n(x) - F(x)\right| \leqslant \varepsilon_n\right) &\geqslant 1-\alpha, \quad \forall x \in \mathbb{R} \Leftrightarrow \\
\mathrm{P}\left(-\varepsilon_n \leqslant \widehat{F}_n(x) - F(x) \leqslant \varepsilon_n\right) &\geqslant 1-\alpha, \quad \forall x \in \mathbb{R} \Leftrightarrow \\
\mathrm{P}\left(-\widehat{F}_n(x) - \varepsilon_n \leqslant -F(x) \leqslant -\widehat{F}_n(x) + \varepsilon_n\right) &\geqslant 1-\alpha, \quad \forall x \in \mathbb{R} \Leftrightarrow \\
\mathrm{P}\left(\widehat{F}_n(x) - \varepsilon_n \leqslant F(x) \leqslant \widehat{F}_n(x) + \varepsilon_n\right) &\geqslant 1-\alpha, \quad \forall x \in \mathbb{R}
\end{aligned}$$

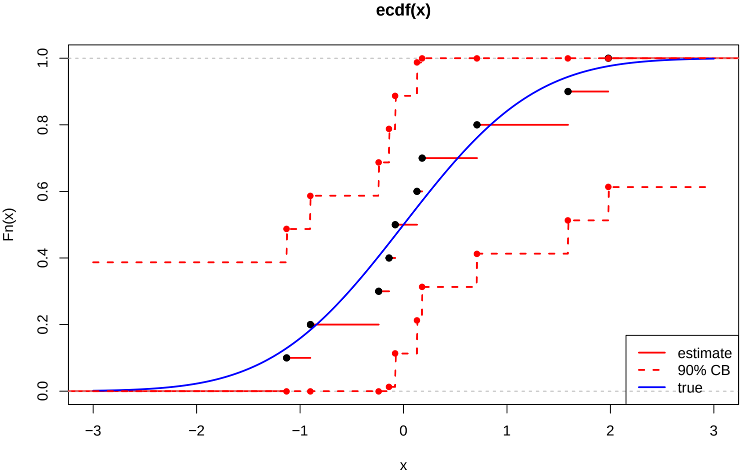
η απόδειξη ολοκληρώνεται θέτοντας $L(x) = \max\{\widehat{F}_n(x) - \varepsilon_n, 0\}$ και $U(x) = \min\{\widehat{F}_n(x) + \varepsilon_n, 1\}$.

Παράδειγμα: 90% ζώνη εμπιστοσύνης $(n = 10)$

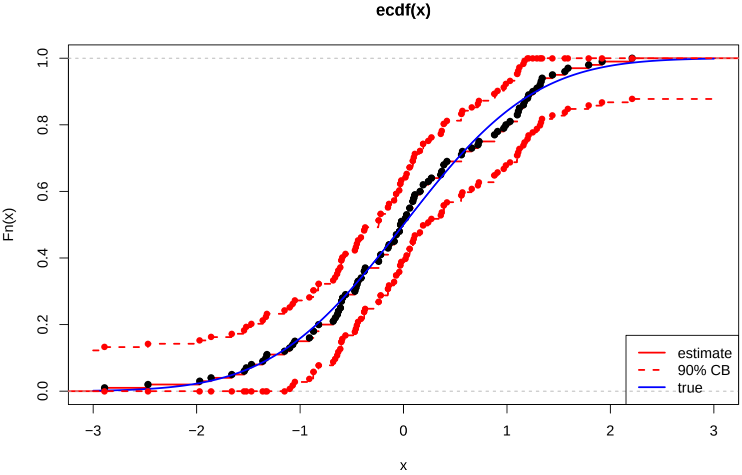
- $n = 10$
- $\alpha = 0.1$
- $\varepsilon_n = \sqrt{\frac{1}{2n} \log \frac{2}{\alpha}} \approx 0.387$
- $L_n(x) = \max\{\widehat{F}_n(x) - 0.387, 0\}$
- $U_n(x) = \min\{\widehat{F}_n(x) + 0.387, 1\}$

x	$\widehat{F}_n(x)$	$L_n(x)$	$U_n(x)$
$(-\infty, -1.13)$	0	0	0.387
$[-1.13, -0.90)$	0.1	0	0.487
$[-0.90, -0.24)$	0.2	0	0.587
$[-0.24, -0.14)$	0.3	0	0.687
$[-0.14, -0.08)$	0.4	0.013	0.787
$[-0.08, 0.13)$	0.5	0.113	0.887
$[0.13, 0.18)$	0.6	0.213	0.987
$[0.18, 0.71)$	0.7	0.313	1
$[0.71, 1.59)$	0.8	0.413	1
$[1.59, 1.98)$	0.9	0.513	1
$[1.98, \infty)$	1	0.613	1

Παράδειγμα: 90% ζώνη εμπιστοσύνης $(n = 10)$



Παράδειγμα: 90% ζώνη εμπιστοσύνης ($n = 100$)



Διαστήματα εμπιστοσύνης για την $F(x)$ σε δοθέν x

- Θέλουμε να καθορίσουμε ℓ_n, u_n τέτοια ώστε

$$\mathrm{P}(\ell_n \leqslant F(x) \leqslant u_n) \geqslant 1 - \alpha$$

- Όμως

$$Y = n\widehat{F}_n(x) \sim \mathcal{B}(n, F(x))$$

- Οπότε το πρόβλημα μπορεί να αναδιατυπωθεί ως εξής:
  - Παρατηρούμε $Y \sim \mathcal{B}(n, p)$
  - $p = F(x) \in [0, 1]$ (άγνωστη παράμετρος)
  - και θέλουμε να βρούμε διάστημα εμπιστοσύνης για το p.
- Μερικές τεχνικές
  1. Clopper-Pearson: Ακριβές Δ.Ε. μέσω της διωνυμικής κατανομής
  2. Wald: ασυμπτωτικό Δ.Ε
  3. Wilson: ασυμπτωτικό Δ.Ε.
  4. μετασχηματισμός $\arcsin$: ασυμπτωτικό Δ.Ε. με σταθεροποιήση (ασυμπτωτικής) διασποράς

Ασυμπτωτικό ΔΕ (Wald)

- Έστω $\widehat{p}_n = Y/n$. Από Κεντρικό Οριακό Θεώρημα

$$\sqrt{n}\frac{\widehat{p}_n - p}{\sqrt{p(1-p)}} \xrightarrow{d} \mathcal{N}(0,1)$$

Ασυμπτωτικό ΔΕ (Wald)

- Έστω $\widehat{p}_n = Y/n$. Από Κεντρικό Οριακό Θεώρημα

$$\sqrt{n}\frac{\widehat{p}_n - p}{\sqrt{p(1-p)}} \xrightarrow{d} \mathcal{N}(0,1)$$

- (Ασθενής) Νόμος Μεγάλων Αριθμών

$$p_n \xrightarrow{P} p$$

Ασυμπτωτικό ΔΕ (Wald)

- Έστω $\widehat{p}_n = Y/n$. Από Κεντρικό Οριακό Θεώρημα

$$\sqrt{n}\frac{\widehat{p}_n - p}{\sqrt{p(1-p)}} \xrightarrow{d} \mathcal{N}(0,1)$$

- (Ασθενής) Νόμος Μεγάλων Αριθμών

$$p_n \xrightarrow{P} p$$

- Θεώρημα Slutsky

$$\sqrt{n}\frac{\widehat{p}_n - p}{\sqrt{p(1-p)}} = \sqrt{n}\frac{\frac{\widehat{p}_n - p}{\sqrt{\widehat{p}_n(1-\widehat{p}_n)}}}{\sqrt{\frac{p(1-p)}{\widehat{p}_n(1-\widehat{p}_n)}}} \xrightarrow{d} \mathcal{N}(0,1)$$

Ασυμπτωτικό ΔΕ (Wald)

- Έστω $\widehat{p}_n = Y/n$. Από Κεντρικό Οριακό Θεώρημα

$$\sqrt{n}\frac{\widehat{p}_n - p}{\sqrt{p(1-p)}} \xrightarrow{d} \mathcal{N}(0,1)$$

- (Ασθενής) Νόμος Μεγάλων Αριθμών

$$p_n \xrightarrow{P} p$$

- Θεώρημα Slutsky

$$\sqrt{n}\frac{\widehat{p}_n - p}{\sqrt{p(1-p)}} = \sqrt{n}\frac{\frac{\widehat{p}_n - p}{\sqrt{\widehat{p}_n(1-\widehat{p}_n)}}}{\sqrt{\frac{p(1-p)}{\widehat{p}_n(1-\widehat{p}_n)}}} \xrightarrow{d} \mathcal{N}(0,1)$$

- Λύνοντας ως προς p προκύπτει ότι το τυχαίο διάστημα

$$\left[\widehat{p}_n - z_{\alpha/2}\sqrt{\frac{\widehat{p}_n(1-\widehat{p}_n)}{n}}, \widehat{p}_n + z_{\alpha/2}\sqrt{\frac{\widehat{p}_n(1-\widehat{p}_n)}{n}}\right]$$

είναι το $100(1-\alpha)\%$ Α.Δ.Ε ίσων ουρών για το $p = F(x)$.

Παράδειγμα ($n = 10$)

- Δεδομένα:

$$(-0.90, 0.18, 1.59, -1.13, -0.08, 0.13, 0.71, -0.24, 1.98, -0.14)$$

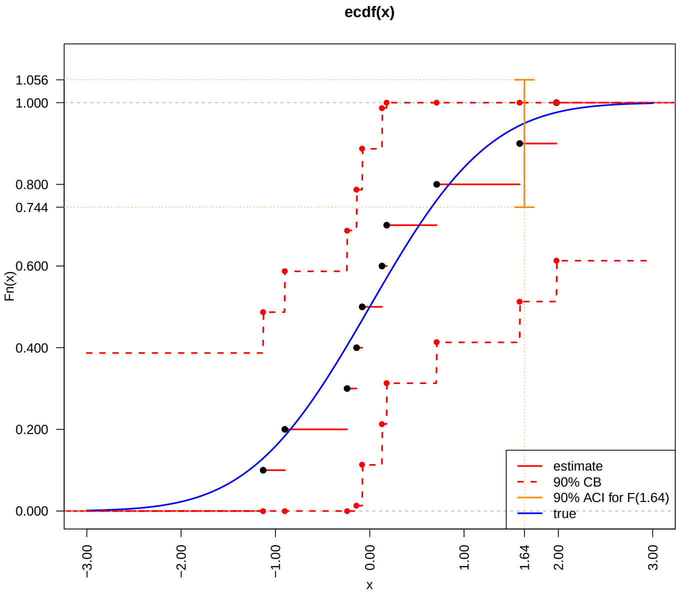
- Να υπολογιστεί το 90% ΑΔΕ ίσων ουρών του Wald για την $F(1.64)$.

- Υπολογισμός $\widehat{F}_n(x)$

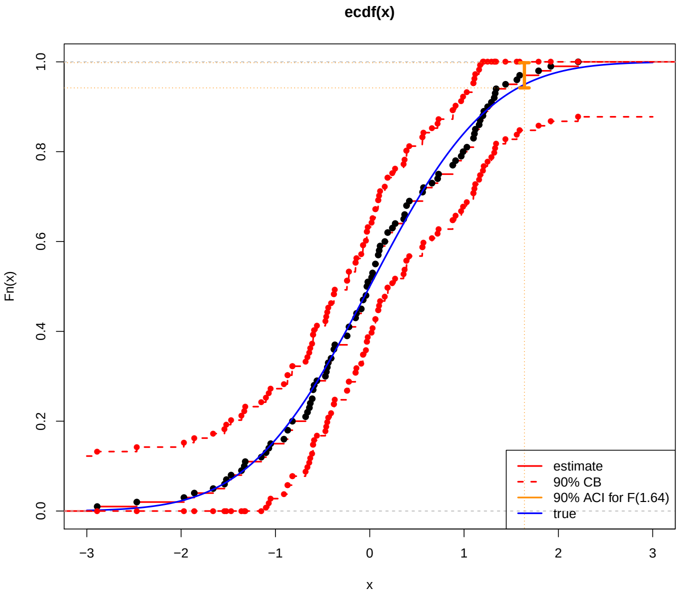
x	$\widehat{F}_n(x)$
$(-\infty, -1.13)$	0
$[-1.13, -0.90)$	1/10
$[-0.90, -0.24)$	2/10
$[-0.24, -0.14)$	3/10
$[-0.14, -0.08)$	4/10
$[-0.08, 0.13)$	5/10
$[0.13, 0.18)$	6/10
$[0.18, 0.71)$	7/10
$[0.71, 1.59)$	8/10
$[1.59, 1.98)$	9/10
$[1.98, \infty)$	1

- Είναι $\widehat{p}_n = \widehat{F}_n(1.64) = 0.9$, $\alpha = 0.1$, $z_{0.05} \approx 1.64$ και $n = 10$
- 90% ΑΔΕ για $F(1.64)$: $[0.744, 1.056]$
- Παρατηρήστε ότι τα άκρα του ΑΔΕ μπορεί να υπερβαίνουν τα όρια του παραμετρικού χώρου.

Παράδειγμα ($n = 10$)



Παράδειγμα ($n = 100$): $[0.942, 0.998]$



Έλεγχος καλής προσαρμογής Kolmogorov-Smirnov

Ενδιαφέρει ο έλεγχος της

$$H_0 : F(x) = F_0(x), \forall x \in \mathbb{R}, \quad \text{vs} \quad H_1 : \exists x : F(x) \neq F_0(x).$$

- Ορίζουμε ως στατιστική συνάρτηση Kolmogorov-Smirnov την

$$D_n = \sup_{x \in \mathbb{R}} \left| \widehat{F}_n(x) - F(x) \right|.$$

- «Μεγάλες» τιμές της $D_n \Rightarrow$ απόρριψη H_0
- Ασυμπτωτικά,

$$\mathrm{P}(\sqrt{n} D_n > k) \xrightarrow{H_0} 2 \sum_{j=1}^{\infty} (-1)^{j+1} e^{-2j^2k^2}$$

- (Ασυμπτωτικό) p-value:

$$2 \sum_{j=1}^{\infty} (-1)^{j+1} e^{-2j^2k^2},$$

όπου k η πραγματοποίηση της $\sqrt{n} D_n$ στο δείγμα.

Παράδειγμα ($n = 10$)

- $(-0.90, 0.18, 1.59, -1.13, -0.08, 0.13, 0.71, -0.24, 1.98, -0.14)$
- Να ελεγχθεί η υπόθεση ($\alpha = 5\%$)

$$H_0 : F(x) = F_0(x), \forall x \in \mathbb{R}, \quad \text{VS} \quad H_1 : \exists x : F(x) \neq F_0(x),$$

όπου $F_0(x)$ η συνάρτηση κατανομής της $\mathcal{N}(0,1)$.

- Υπολογισμός $\widehat{F}_n(x)$

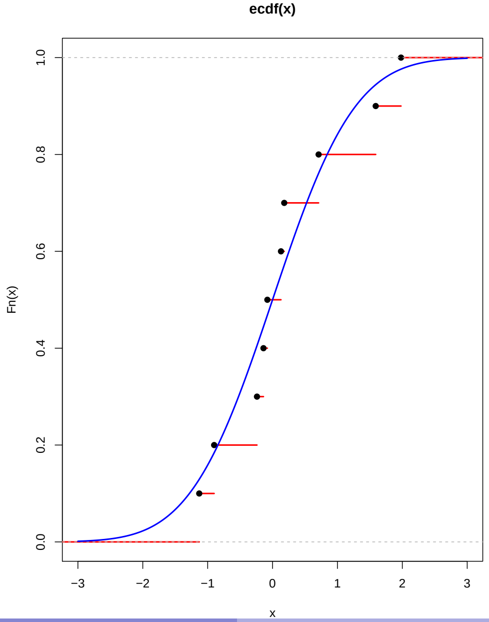
x	$\widehat{F}_n(x)$
$(-\infty, -1.13)$	0
$[-1.13, -0.90)$	1/10
$[-0.90, -0.24)$	2/10
$[-0.24, -0.14)$	3/10
$[-0.14, -0.08)$	4/10
$[-0.08, 0.13)$	5/10
$[0.13, 0.18)$	6/10
$[0.18, 0.71)$	7/10
$[0.71, 1.59)$	8/10
$[1.59, 1.98)$	9/10
$[1.98, \infty)$	1

- Συνάρτηση κατανομής της $\mathcal{N}(0,1)$: $F_0(x) = \int\limits_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{t^2}{2}\right\} \mathrm{d}t.$

Παράδειγμα (συνέχεια)

σελ 24 σημειώσεων Ιωαννίδη :

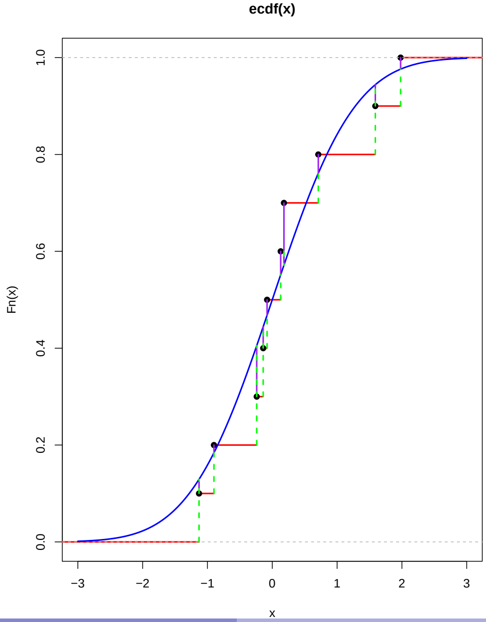
Αρκεί το στατιστικό να υπολογιστεί ως max επί των $x = X_i$ και των $x = X_i - \varepsilon$ για μικρό ε.



Παράδειγμα (συνέχεια)

σελ 24 σημειώσεων Ιωαννίδη :

> *Αρκεί το στατιστικό να υπολογιστεί ως max επί των* $x = X_i$ *και των* $x = X_i - \varepsilon$ *για μικρό* ε.



Παράδειγμα (συνέχεια)

- Έστω
  - $\widehat{F}_n(x_0^-) = \lim_{x \to x_0^-} \widehat{F}_n(x)$
  - $D^-(x) = |\widehat{F}_n(x^-) - F_0(x)|$
  - $D^+(x) = |\widehat{F}_n(x) - F_0(x)|$
- Για κάθε x στο παρατηρηθέν δείγμα υπολογίζουμε $D^-(x)$ και $D^+(x)$
- Η στατιστική συνάρτηση Kolmogorov-Smirnov ισούται με

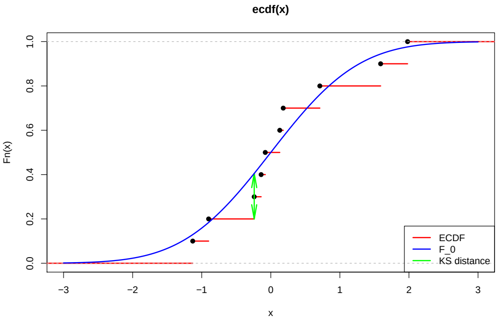
$$D_n = \max_x \{D^-(x), D^+(x)\}.$$

Παράδειγμα (συνέχεια)

	x	$\widehat{F}_n(x^-)$	$\widehat{F}_n(x)$	$F_0(x)$	D^+	D^-
1	-1.130	0.000	0.100	0.129	0.029	0.129
2	-0.900	0.100	0.200	0.184	0.016	0.084
3	-0.240	0.200	0.300	0.405	0.105	0.205
4	-0.140	0.300	0.400	0.444	0.044	0.144
5	-0.080	0.400	0.500	0.468	0.032	0.068
6	0.130	0.500	0.600	0.552	0.048	0.052
7	0.180	0.600	0.700	0.571	0.129	0.029
8	0.710	0.700	0.800	0.761	0.039	0.061
9	1.590	0.800	0.900	0.944	0.044	0.144
10	1.980	0.900	1.000	0.976	0.024	0.076

- Συνεπώς $D_n = 0.205$.
- Η στήλη $F_0(x)$ υπολογίστηκε μέσω της `pnorm()` στην `R`.

Παράδειγμα (συνέχεια)



- Είναι $k = \sqrt{n} D_n = \sqrt{10} \times 0.205 \approx 0.648$
- (Ασυμπτωτικό) p-value: $2 \sum_{j=1}^{\infty} (-1)^{j+1} e^{-2j^2k^2} \approx 0.794 > 0.05$
- Αποτυγχάνουμε να απορρίψουμε την H_0 σε επίπεδο σημαντικότητας 5%.