Διάλεξη 1

Μαθηματική Στατιστική

$\downarrow$

Με τι ασχολείται;

Ο κλάδος της Στατιστικής που ασχολείται με την ανάλυση **δεδομένων** με μαθηματικές τεχνικές.

Δεδομένα

$$\underbrace{x_1, x_2, \ldots, x_n}_{\text{παρατηρήσεις που προέρχονται από τη μελέτη κάποιου φαινομένου}}$$

- αποτελέσματα επαναλαμβανόμενων ρίψεων ενός κέρματος
- χρόνοι ζωής λαμπτήρων μιας συγκεκριμένης μάρκας
  χρόνοι ζωής ασθενών μετά από μια σοβαρή επέμβαση
- μισθοί κατοίκων μιας συγκεκριμένης χώρας
- μια χρονοσειρά της εξέλιξης της τιμής μιας μετοχής
- χρόνοι άφιξης πελατών σε μια τράπεζα ή των μηνυμάτων μας στο gmail
- παρουσία ή όχι μιας συγκεκριμένης ασθένειας σε μια ομάδα ατόμων

Τύποι Προβλημάτων

- **απόφαση**: αντικατάσταση ενός φαρμάκου από κάποιο άλλο,
  καθορισμός ασφαλίστρων

- **πρόβλεψη**: δείκτης χρηματιστηρίου,
  βαθμός Στατιστική

- **εξήγηση**: συμπεριφορά καταναλωτών (κατανομή της ζήτησης προϊόντων),
  συγκέντρωση όζοντος ως συνάρτηση άλλων παραγόντων
  (π.χ. θερμοκρασία, νέφος)

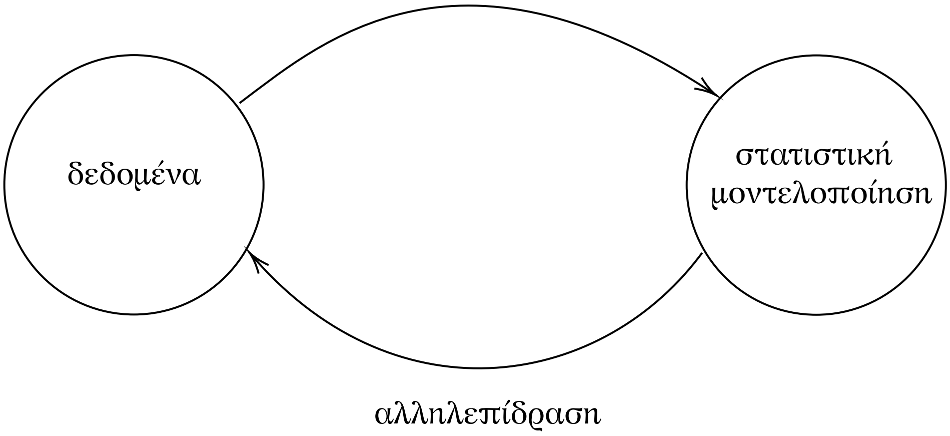
- **εντοπισμός**: θέση μέσω GPS,
  κρίσιμη περιοχή στο DNA που συνδέεται με μετάλλαξη γονιδίων

Τύποι Δεδομένων

- ποσοτικές (συνεχείς, διακριτές)
- ποιοτικές / κατηγορικές (φύλο ή επάγγελμα)
- διατακτικές (σειρά προτίμησης)
- εικόνες
- χρονοσειρές
- γραφήματα
- βίντεο

Προεπεξεργασία Δεδομένων

- καθαρισμός δεδομένων
- συμπλήρωση χαμένων δεδομένων
- διαγραφή μη πλήρων δεδομένων
- μετασχηματισμοί δεδομένων (εξαρτάται από την εφαρμογή)



Το στατιστικό μοντέλο είναι το εργαλείο εκείνο που θα μας επιτρέψει να κατανοήσουμε όσο γίνεται καλύτερα τον άγνωστο μηχανισμό παραγωγής των δεδομένων.

Κεντρική Ιδέα

$$\begin{array}{ccccl} x_1 & x_2 & \cdots & x_n & \to \ \text{δεδομένα} \\ \uparrow & \uparrow & & \uparrow & \\ X_1 & X_2 & \cdots & X_n & \to \ \text{τυχαίες μεταβλητές} \end{array}$$

Τα δεδομένα θεωρούμε ότι αντιστοιχούν σε πραγματοποιήσεις κάποιων τυχαίων μεταβλητών. Έτσι λοιπόν αυτό που έχουμε παρατηρήσει υποθέτουμε ότι έχει προκύψει ως αποτέλεσμα κάποιου τυχαίου πειράματος.

Στις Πιθανότητες οι κατανομές θεωρούνται γνωστές, ενώ στη Στατιστική, οι κατανομές, αλλά και η σύνδεσή τους θα μπορούσε να είναι εντελώς ή μερικώς άγνωστη.

Ένα **στατιστικό μοντέλο** είναι μία υπόθεση που κάνουμε για τη φύση της άγνωστης κατανομής του τυχαίου διανύσματος $(X_1, X_2, \ldots X_n)$.

Σε αυτό το μάθημα, συνήθως θα υποθέτουμε ότι οι

$$X_1, X_2, \ldots, X_n \underbrace{\text{ είναι ανεξάρτητες και ισόνομες τ.μ.}}_{\text{τυχαίο δείγμα}}$$

με κοινή **άγνωστη** συνάρτηση κατανομής F

ή

συνάρτηση πιθανότητας f (διακριτή)

συνάρτηση πυκνότητας πιθανότητας f ((απόλυτα) συνεχών)

μοντελοποίηση

$\to$ υποθέσεις για την υποψήφια κατανομή

$F \in \mathcal{F} \to$ οικογένεια κατανομών (αποτελούν τις υποψήφιες F)

Διαδικασία Επίλυσης ενός Στατιστικού προβλήματος

① Πως θα μοντελοποιήσουμε την παραγωγή των δεδομένων; Ποιά είναι η $\mathcal{F}$;

② Μέθοδοι / Κριτήρια επιλογής $\widehat{F}$ (της εκτιμήτριας)

(**κλασική** / μπεϋζιανή; / μέγιστη πιθανοφάνεια, εκτίμηση ροπών, ελάχιστα τετράγωνα)

③ Καθορισμός ακρίβειας (π.χ. διάστημα εμπιστοσύνης)

④ Αλληλεπίδραση με τα δεδομένα, επιστροφή στο αρχικό πρόβλημα, σύγκριση μοντέλων, ...

Προσοχή!

Σε δεδομένο πρόβλημα $\exists$ συνήθως πολλές λύσεις (διαφορετική μεθοδολογία ή κριτήρια επιλογής) και όχι κατ' ανάγκη βέλτιστη λύση.

Υποθέσεις του μαθήματος Μαθηματικής Στατιστικής

Οι τυχαίες μεταβλητές $X_1, X_2, \ldots, X_n$:

(1) είναι ανεξάρτητες και ισόνομες τυχαίες μεταβλητές (τυχαίο δείγμα)

(2) είναι μονοδιάστατες

(3) έχουν κοινή άγνωστη κατανομή που προσδιορίζεται πλήρως από μία παράμετρο πεπερασμένης διάστασης (παραμετρική Στατιστική)

(4) η παράμετρος αυτή, έστω θ, είναι μια άγνωστη σταθερά (υπό εκτίμηση) και οτιδήποτε συμπεράνουμε γι' αυτήν πρέπει να στηριχθεί μόνο στις $X_1, \ldots, X_n$ χωρίς τη δυνατότητα ενσωμάτωσης κάποιας **εκ των προτέρων** γνώσης ή **πίστης** για τη συμπεριφορά αυτής της παραμέτρου

(5) δε μπορούν να εξηγηθούν καλύτερα από άλλες βοηθητικές μεταβλητές καθώς υποθέτουμε ότι απουσιάζει τέτοιου είδους πληροφορία

(6) έχουν κατανομή με "εύκολα" διαχειρίσιμη μαθηματική μορφή που μας επιτρέπει να υπολογίζουμε γενικά ποσότητες ενδιαφέροντος σε κλειστή μορφή

Πιθανότητες Ι

$\downarrow$

Μαθηματική Στατιστική

Γραμμικά Μοντέλα (5)

Μπεϋζιανή Στατιστική (4)

Υπολογιστική Στατιστική (6)

Πολυμεταβλητή Ανάλυση Δεδομένων (2)

Μη Παραμετρική Στατιστική (3)

[+ Θεωρία Αξιοπιστίας $\rightarrow$ γίνεται και ένα κομμάτι Στατιστικής]

Παράδειγμα

Μας δίνονται τα εξής 10 δεδομένα από επαναλαμβανόμενες ρίψεις ενός κέρματος

$$\kappa \;\; \kappa \;\; \kappa \;\; \gamma \;\; \gamma \;\; \gamma \;\; \gamma \;\; \kappa \;\; \gamma \;\; \gamma$$

Στόχος: Να "εξηγήσετε" την παραγωγή των δεδομένων

Στατιστικό μοντέλο: Υποθέσεις

(Α) οι ρίψεις είναι ανεξάρτητες

(Β) οι ρίψεις είναι ισόνομες

(Γ) $X_i \sim \mathrm{Be}(\theta),\ 1 \leq i \leq n,\ \theta \in (0,1)$ άγνωστο

Παραμετρικό μοντέλο είναι η οικογένεια $\mathrm{Be}(\theta),\ \theta \in (0,1) \equiv \Theta \rightarrow$ ο παραμετρικός χώρος

Τροποποιημένος στόχος $\rightarrow$ να εκτιμήσουμε το θ
(δηλαδή να μαντέψουμε το άγνωστο θ)

Μετασχηματισμός δεδομένων

$$\begin{array}{cccccccccc} 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 1 \\ \shortparallel & \shortparallel & & & & & & & & \shortparallel \\ x_1 & x_2 & & & & & & & & x_{10} \end{array} \longrightarrow \text{μέγεθος δείγματος}$$

Γενικά **μέγεθος δείγματος** λέμε το πλήθος των παρατηρήσεων / δεδομένων
που διαθέτουμε

Πληθυσμό λέμε την άγνωστη κατανομή.

[Ένα τυχαίο δείγμα προέρχεται από κάποιον πληθυσμό $X_1, \ldots, X_n \sim F($ ή $P)$]

Προβλήματα που θα μας απασχολήσουν

(i) τρόποι εύρεσης $\widehat{\theta}$ εκτιμητριών [Σημειακή Εκτίμηση]

(ii) Καθορισμος ακρίβειας εκτίμησης [Διαστήματα Εμπιστοσύνης]

(iii) **Έλεγχοι Υποθέσεων**

π.χ. πιστεύουμε ότι $p = \frac{1}{2}$ και δεν εγκαταλείπουμε αυτή την "πίστη" εκτός εάν από τα δεδομένα έχουμε ισχυρές ενδείξεις **εναντίον** αυτής της υπόθεσης.

Διάλεξη 2

Εστιάζουμε στην **παραμετρική** Στατιστική.

Η άγνωστη συνάρτηση κατανομής $F \in \mathcal{F} = \{F_\theta : \theta \in \Theta\},\ \Theta \subset \mathbb{R}^d$

ή $P \in \mathcal{P} = \{P_\theta : \theta \in \Theta\}$ (θ ευκλείδια παράμετρος)

ή αντίστοιχα **συνάρτηση πιθανότητας** (διακριτή)

ή **συνάρτηση πυκνότητας πιθανότητας** (συνεχής) f_θ

Ξεκινάμε με **εκτιμητική** και ειδικά **σημειακή εκτίμηση**. Εκμεταλλευόμαστε το τυχαίο

δείγμα $X_1, X_2, \ldots, X_n$ για να "μάθουμε" για το θ. (εκτιμάμε το θ)

$$0\ \ 0\ \ 0\ \ 1\ \ 1\ \ 1\ \ 1\ \ 0\ \ 1\ \ 1 \qquad (0 \leftrightarrow \kappa\ ,\ 1 \leftrightarrow \gamma)$$

$X_i \sim \mathrm{Be}(\theta),\ 0 < \theta < 1$. Πώς εκτιμάμε το θ;

• Λύση (πιθανή)

$*$ Εκφράζει το θ κάποιο θεωρητικό χαρακτηριστικό της κατανομής;

π.χ. εδώ $\theta = P_\theta(X_i = 1) \longrightarrow$ είναι το βάρος που δίνει η κατανομή στο σημείο 1.

Μπορούμε με το δείγμα να εκφράσουμε την αντίστοιχη ιδιότητα;

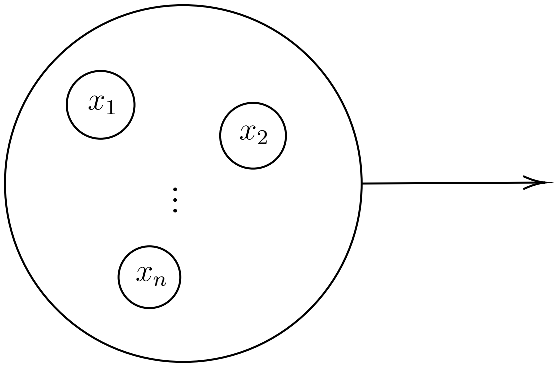
$$\begin{array}{cccc} X_1, & X_2, & \ldots, & X_n \\ \uparrow & \uparrow & & \uparrow \\ x_1 & x_2 & \ldots & x_n \end{array} \qquad \left(\begin{array}{l} \text{αντίγραφα κάποιας υποθετικής} \\ \text{τυχαίας μεταβλητής} \sim F \end{array} \right)$$

Ξέρουμε κάποιες τιμές της θεωρητικής. Μπορούμε να φτιάξουμε ένα τυχαίο πείραμα που να μοντελοποιεί τη γνώση που έχουμε για τη θεωρητική κατανομή;

Πώς με τα δεδομένα $x_1, x_2, \ldots, x_n$ μπορώ να πάρω μία τιμή;

Δε μπορούμε να ευνοήσουμε κάποια από αυτές!

τις βάζουμε σε μία κληρωτίδα



επιλέγουμε μία στην τύχη

Ο μηχανισμός αυτός θα θέλαμε να προσεγγίζει το θεωρητικό.

$$\left[\begin{array}{l}\text{όσο το } n \text{ μεγαλώνει, τόσο καλύτερη γίνεται η προσέγγιση}\\ \text{κομμάτι της μη-Παραμετρικής Στατιστικής}\end{array}\right]$$

Άρα υποθέτουμε ότι έχουμε κάποια

$$X^* = \begin{cases} x_1 & \text{, με πιθανότητα } \dfrac{1}{n} \\ x_2 & \text{, με πιθανότητα } \dfrac{1}{n} \\ \vdots & \qquad \vdots \\ x_n & \text{, με πιθανότητα } \dfrac{1}{n} \end{cases}$$

Αυτή η $X^* \sim \text{discr.Unif}\left(\{x_1, x_2, \dots, x_n\}\right)$

Στο συγκεκριμένο πρόβλημα επειδή $x_i \in \{0, 1\}$ θα γράφεται

$$X^* = \begin{cases} 1 & \text{, με πιθανότητα } \dfrac{\#x_i = 1}{n} \\ 0 & \text{, με πιθανότητα } \dfrac{\#x_i = 0}{n} \end{cases}$$

Τελικά $\widehat{\theta} = P_\theta\widehat{(x_i = 1)} = P(X^* = 1) = \dfrac{\sum_{i=1}^{n} x_i}{n} \quad \left(\widehat{\theta}(x_1, x_2, \dots, x_n)\right)$

στο συγκεκριμένο δείγμα πήραμε

$$\widehat{\theta} = \frac{x_1 + x_2 + \cdots + x_n}{n} \qquad \text{π.χ. εδώ } \widehat{\theta} = \frac{6}{10}$$

μετατρέπεται σε μία στατιστική διαδικασία

$$\begin{array}{cccc} x_1, & x_2, & \dots, & x_n \\ \downarrow & \downarrow & & \downarrow \\ X_1 & X_2 & & X_n \end{array}$$

$$\widehat{\theta}(X_1, X_2, \dots, X_n) = \frac{X_1 + X_2 + \cdots + X_n}{n} \quad \text{(ο δειγματικός μέσος)}$$

$\longrightarrow$ συνάρτηση του τυχαίου δείγματος.

Παρατήρηση 0.0.1. Η $\widehat{\theta}$ που φτιάξαμε λέγεται **εμπειρική** εκτιμήτρια του θ, διότι στηρίζεται στην εμπειρική κατανομή του τυχαίου δείγματος $X_1, X_2, \ldots, X_n$ (αυτή που βάζει πιθανότητα $\frac{1}{n}$ σε κάθε μία από αυτές)

π.χ. $\underset{\text{(θεωρητικό)}}{P_\theta(X=1)} \to P(X^*=1) =$ ποσοστό των παρατηρήσεων του δείγματος που είναι 1

μία πιθανότητα $\xrightarrow{\text{εκτιμάται}}$ ποσοστό

Παρατήρηση 0.0.2. Μία εκτιμήτρια είναι συνάρτηση του τυχαίου δείγματος και άρα είναι τυχαία μεταβλητή με κατανομή που εξαρτάται γενικά από το θ.

π.χ. εδώ $\widehat{\theta} = \dfrac{\sum_{i=1}^{n} X_i}{n} \sim \dfrac{1}{n}\underset{\substack{\downarrow \\ \text{κατανομή που εξαρτάται από το } \theta}}{\text{Bin}(n,\theta)} \quad \left(= \dfrac{Y}{n},\ Y \sim \text{Bin}(n,\theta)\right)$

Παρατήρηση 0.0.3. Έδω ερμηνεύεται και ως **εμπειρικός μέσος** αφού

$$\underbrace{\mathbb{E}_\theta(X_i)}_{\text{θεωρητική μέση τιμή}} = \theta \quad \xrightarrow{\text{εμπειρική}} \quad \widehat{\theta} = \mathbb{E}(X^*) = \underset{\text{δειγματικός μέσος}}{\dfrac{\sum_{i=1}^{n} X_i}{n}}$$

Παρατήρηση 0.0.4.

$$\widehat{\theta} = \frac{\sum_{i=1}^{n} X_i}{n} \in [0,1], \quad \text{όμως } \theta \in (0,1).$$

Γενικά θέλουμε μία εκτιμήτρια να παίρνει τιμές εντός του παραμετρικού χώρου

$\Downarrow$

επεκτείνουμε τον παραμετρικό χώρο $\Theta = [0,1]$

δηλαδή επεκτείνουμε με $\underset{\substack{\downarrow \\ \text{Be}(0) \equiv 0}}{\theta = 0}$ και $\underset{\substack{\downarrow \\ \text{Be}(1) \equiv 1}}{\theta = 1}$

Θα γράφουμε $X = (X_1, X_2, \ldots, X_n)$ για το τυχαίο δείγμα, δηλαδή στη μορφή τυχαίου διανύσματος.

Ορισμός 0.0.1. (Στατιστική Συνάρτηση)

Λέμε **στατιστική συνάρτηση** (ή στατιστικό, ή δειγματοσυνάρτηση) κάθε τυχαία μεταβλητή

$T = T(X)$, δηλαδή κάθε συνάρτηση του τυχαίου δείγματος.

[το πεδίο ορισμού της $T(\cdot)$ πρέπει να περιέχει το σύνολο τιμών της X]

Παρατήρηση 0.0.5. Κάθε στατιστική συνάρτηση είναι συνάρτηση **μόνο** του τυχαίου δείγματος και δεν εμπλέκει άγνωστες παραμέτρους, αλλά ούτε και άλλες τυχαίες μεταβλητές εκτός των $X_1, X_2, \ldots, X_n$.

π.χ. αν θ είναι παράμετρος τότε $\overline{X}_n - \theta$ δεν είναι στατιστική συνάρτηση

ή π.χ. $X_1^2 + X_2^2 + \cdots + X_n^2 - U$, όπου η U είναι τυχαία μεταβλητή που δεν είναι συνάρτηση

των $X_1, X_2, \ldots, X_n$, τότε δεν είναι στατιστική συνάρτηση

Όμως $\overline{X}_n = \dfrac{X_1 + X_2 + \cdots + X_n}{n}$

$$\overline{X_n^2} = \frac{X_1^2 + X_2^2 + \cdots + X_n^2}{n}$$

$\left(\overline{X}_n\right)^2, \quad \dfrac{X_1 + X_n}{n}, \quad \max\{X_1, X_2, \ldots, X_n\}$ είναι στατιστικές συναρτήσεις

Ορισμός 0.0.2. (Εκτιμήτρια)

Αν $g : \Theta \to \mathbb{R}$ ή $\mathbb{R}^d$, τότε κάθε στατιστική συνάρτηση $T = T(X)$ με

πεδίο τιμών το $g(\Theta)$ λέγεται εκτιμήτρια της παραμετρικής συνάρτησης $g(\theta)$.

(κάθε συνάρτηση της παραμέτρου, είναι παράμετρος)

Παρατήρηση 0.0.6. Ο ορισμός αυτός είναι πολύ γενικός “για να έχει ενδιαφέρον”. Απέχει πολύ από την έννοια μίας λογικής ή (περισσότερο) καλής εκτιμήτριας.

Κατ' αρχήν θα μας φανούν χρήσιμες οι εμπειρικές εκτιμήτριες (είναι λογικές).

$$\text{θεωρητικό χαρακτηριστικό} \quad \to \quad \text{δειγματικό χαρακτηριστικό}$$

$$\theta = \theta(F) \quad \to \quad \underbrace{\widehat{\theta} = \theta(F^*)}_{\text{εμπειρική}}$$

Ιδιαιτέρως, η μέθοδος αυτή έχει πολύ ενδιαφέρον στις ροπές.

	<u>θεωρητικό</u>	<u>δειγματικό</u>
μέση τιμή	$\mathbb{E}(X_1)$	$\overline{X}_n = \dfrac{\sum_{i=1}^{n} X_i}{n}$
ροπή k-τάξης	$\mathbb{E}\left(X_1^k\right)$	$\overline{X_n^k} = \dfrac{\sum_{i=1}^{n} X_i^k}{n}$
κεντρική ροπή k-τάξης $\downarrow$ ειδικά	$\mathbb{E}\left[(X_1 - \mathbb{E}(X_1))^k\right]$	$\dfrac{\sum_{i=1}^{n} \left(X_i - \overline{X}_n\right)^k}{n}$
διασπορά	$\mathbb{V}(X_1)$	$\underbrace{\dfrac{\sum_{i=1}^{n} \left(X_i - \overline{X}_n\right)^2}{n}}_{\text{δειγματική διασπορά}}$
....		

όλες αυτές είναι της μορφής $\mathbb{E}\left[g(X_1)\right] \to \mathbb{E}\left[g(X^*)\right]$

Επεκτείνεται και σε άλλα χαρακτηριστικά όπως συνάρτηση κατανομής, πιθανότητες, ποσοστημόρια, διάμεσος, ακραίες τιμές (για φραγμένες τυχαίες μεταβλητές).

Αυτός ο τρόπος εξαγωγής εκτιμητριών είναι στην ουσία μη-παραμετρικός.

Ασκήσεις

① Έστω $x_1, x_2, \ldots, x_n$ τυχαίο δείγμα από $\mathcal{N}(\mu, \sigma^2)$ όπου $\mu,\ \sigma^2$ είναι άγνωστα, $\mu \in \mathbb{R},\ \sigma^2 > 0$.

(α) Να βρεθεί ο παραμετρικός χώρος του μοντέλου.

(β) Να προταθούν εκτιμήτριες των άγνωστων παραμέτρων.

Λύση

(α) $\theta = (\mu, \sigma^2)\,,\ \mu \in \mathbb{R},\ \sigma^2 > 0$. Άρα $\Theta = \mathbb{R} \times (0, +\infty)$

(β) $\mu = \mathbb{E}_\theta(X_1) \Rightarrow$ εμπειρική εκτιμήτρια, δηλαδή δειγματικός μέσος

$$\widehat{\mu} = \mathbb{E}(X^*) = \overline{X}_n$$

$$\sigma^2 = \mathbb{V}_\theta(X_1) = \mathbb{E}_\theta\left(X_1 - \mathbb{E}_\theta\left(X_1\right)\right)^2$$

$\Big\downarrow$ εμπειρική

$$\widehat{\sigma^2} = \mathbb{E}\left(X^* - \mathbb{E}(X^*)\right)^2 = \mathbb{V}(X^*)$$

$$= \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2}{n} \to \text{δειγματική διασπορά}$$

Τελικά $\widehat{\theta} = \left(\widehat{\mu}, \widehat{\sigma^2}\right)$

Διάλεξη 3

Εκτιμήτριες Ροπών

Για τις εμπειρικές εκτιμήτριες είδαμε ότι αν $\theta = \theta(F)$, τότε $\widehat{\theta} = \theta(F^*)$.

Στην παραμετρική Στατιστική έχουμε ότι κάθε χαρακτηριστικό ενδιαφέροντος γράφεται $g(\theta)$ αφού $F = F(\theta)$.

Η ιδέα για να εκτιμήσουμε το $\theta = (\theta_1, \theta_2, \ldots, \theta_s)$ είναι να φτιάξουμε ένα σύστημα εξισώσεων που εμπλέκει ροπές.

Σύστημα Εξισώσεων

$$\begin{array}{lll} (= g_1(\theta)) & \mathbb{E}(X_1) = \overline{X} & (= \mathbb{E}(X^*)) \\ (= g_2(\theta)) & \mathbb{E}(X_1^2) = \overline{X^2} & (= \mathbb{E}\left[(X^*)^2\right]) \\ \vdots & \vdots & \vdots \\ (= g_s(\theta)) & \mathbb{E}(X_1^s) = \overline{X^s} & (= \mathbb{E}\left[(X^*)^s\right]) \end{array}$$

Επεκτείνουμε $> s$ αν χρειάζεται...

Η λύση του συστήματος των εξισώσεων (αν υπάρχει) θα μας δώσει μία εκτιμήτρια **ροπών** (ε.ρ.), ας πούμε $\overline{\theta} = \left(\overline{\theta}_1, \overline{\theta}_2, \ldots, \overline{\theta}_s\right)$

Παράδειγμα 0.0.1. Έστω $X_1, X_2, \ldots, X_n$ τυχαίο δείγμα από $\mathcal{N}(\mu, \sigma^2)$, με μ, σ^2 άγνωστα. Να βρεθεί η εκτιμήτρια ροπών $\overline{\theta} = \left(\overline{\mu}, \overline{\sigma^2}\right)$ του $\theta = (\mu, \sigma^2)$.

Λύση

$$\left.\begin{array}{l}\mathbb{E}_\theta(X_1)=\overline{X}\\[2em] \underbrace{\mathbb{E}_\theta\left(X_1^2\right)}_{\mathbb{V}_\theta(X_1)+\mathbb{E}_\theta^2(X_1)}=\overline{X^2}\end{array}\right| \Leftrightarrow \left.\begin{array}{l}\mu=\overline{X}\\[2em] \sigma^2+\mu^2=\overline{X^2}\end{array}\right| \Leftrightarrow \begin{array}{l}\mu=\overline{X}\\[2em] \sigma^2=\overline{X^2}-\left(\overline{X}\right)^2=\dfrac{\sum\limits_{i=1}^{n}\left(X_i-\overline{X}\right)^2}{n}=M_2\end{array}$$

- Έτσι $\overline{\theta}=\left(\overline{X},M_2\right)$

Σχόλια

① εκτιμήτρια ροπών (ε.ρ.) = εμπειρική εκτιμήτρια (ε.ε.)

② **κεντρική** ροπή k-τάξης κάποιας τυχαίας μεταβλητής X

$$\mu_k\equiv\mathbb{E}(X-\underset{\substack{\downarrow\\ \mathbb{E}(X)}}{\mu})^k,\quad k\geq 2$$

π.χ. $\mu_2=\mathbb{E}(X-\mu)^2=\sigma^2$

$\mu_3=\mathbb{E}(X-\mu)^3$

Η διαδικασία που χρησιμοποιήσαμε είναι η **κεντροποίηση** αφαιρώντας τη μέση τιμή. $X\to X-\mu$

Παρατήρηση 0.0.7. το σ^2 είναι μία χρήσιμη παράμετρος, πιο χρήσιμη από τη $\mathbb{E}(X^2)$

③ κεντροποίηση → τυποποίηση

Αν έχει μέση τιμή + τυπική απόκλιση (ή διασπορά) τότε

$$X\xrightarrow{\text{τυποποίηση}}\boxed{\frac{X-\mu}{\sigma}}\qquad\left(\text{κεντροποίηση } \mathbb{E}(X-\mu)=0\right)$$

$$\mathbb{E}\left(\frac{X-\mu}{\sigma}\right)=0,$$

$$\mathbb{V}\left(\frac{X-\mu}{\sigma}\right)=\mathbb{E}\left(\frac{X-\mu}{\sigma}\right)^2=\frac{\mathbb{E}(X-\mu)^2}{\sigma^2}=\frac{\sigma^2}{\sigma^2}=1$$

$$\Rightarrow\mathbb{V}\left(\frac{X-\mu}{\sigma}\right)=1.$$

παίρνουμε $\mathbb{E}\left(\frac{X-\mu}{\sigma}\right)^3 \equiv \lambda \rightarrow$ **συντελεστής λοξότητας** ή **λοξότητα**

$\mathbb{E}\left(\frac{X-\mu}{\sigma}\right)^4 \equiv \kappa \rightarrow$ **συντελεστής κύρτωσης** ή **κύρτωση**

Δύο παράμετροι ενδιαφέροντος είναι το κ και το λ τα οποία μπορούμε να τα εκτιμάμε είτε με τα αντίστοιχα δειγματικά (ποιές είναι;), δηλαδή τις εμπειρικές εκτιμήτριες, ή ως $\lambda(\theta)$ ή $\kappa(\theta)$ τα χρησιμοποιούμε στη μέθοδο των ροπών ή με τη βοήθεια μιας $\widehat{\theta}$ κάνουμε **plug-in** και εκτιμάμε $\widehat{\lambda(\theta)} = \lambda\left(\widehat{\theta}\right) \rightarrow$ εκτιμήτρια ή $\widehat{\kappa(\theta)} = \kappa\left(\widehat{\theta}\right)$. Όμοια $\sigma^2 = \sigma^2(\theta) \Rightarrow \widehat{\sigma^2} = \sigma^2\left(\widehat{\theta}\right)$

④ **plug-in εκτιμήτριες**

Αν $\eta = g(\theta)$, τότε παίρνουμε εκτιμήτρια του η, μέσω της διαδικασίας

$$\widehat{\eta} = \widehat{g(\theta)} = g\left(\widehat{\theta}\right)$$

αρκεί να έχουμε εκτιμήτρια για το θ

⑤ Αντί να χρησιμοποιούμε απ' ευθείας τις ροπές, στη μέθοδο των ροπών μπορούμε να βάζουμε **κεντρικές** ροπές. (προκύπτουν ισοδύναμα συστήματα εξισώσεων)

π.χ.

$$\mathbb{E}_\theta(X_1) = \overline{X}$$
$$\mathbb{V}_\theta(X_1) = \mathbb{E}_\theta(X_1-\mu)^2 = M_2$$

Παράδειγμα 0.0.2. $X_1, X_2, \dots, X_n$ τυχαίο δείγμα $\mathcal{N}(\mu, \sigma^2)$. Θέλουμε $\overline{\theta}$.

$$\left.\begin{aligned} \mathbb{E}_\theta(X_1) &= \overline{X} \\ \mathbb{V}_\theta(X_1) &= M_2 \end{aligned}\right| \Leftrightarrow \left.\begin{aligned} \mu &= \overline{X} \\ \sigma^2 &= M_2 \end{aligned}\right| \Leftrightarrow \overline{\theta} = \left(\overline{X}, M_2\right)$$

Υπενθύμιση

Βασικές Διακριτές					
Κατανομή / Παράμετρος	$\mathrm{Be}(p)$	$\mathrm{Bin}(N,p)$	$\overset{(\geq 0)}{\mathrm{Geo}}(p)$	$\mathcal{P}(\lambda)$	$\overset{(\geq 0)}{\mathrm{NegBin}}(N,p)$
μέση τιμή	p	Np	$\frac{1-p}{p}$	λ	$\frac{N(1-p)}{p}$
διασπορά	$p(1-p)$	$Np(1-p)$	$\frac{(1-p)}{p^2}$	λ	$\frac{N(1-p)}{p^2}$

Βασικές Συνεχείς					
Κατανομή / Παράμετρος	$\mathcal{U}(\alpha,\beta)$	$\mathcal{N}(\mu,\sigma^2)$	$\mathrm{Exp}(\theta)$	$\mathrm{G}(\alpha,\theta)$	$\mathrm{B}(\alpha,\beta)$
μέση τιμή	$\frac{\alpha+\beta}{2}$	μ	$\frac{1}{\theta}$	$\frac{\alpha}{\theta}$	$\frac{\alpha}{\alpha+\beta}$
διασπορά	$\frac{(\beta-\alpha)^2}{12}$	σ^2	$\frac{1}{\theta^2}$	$\frac{\alpha}{\theta^2}$	$\frac{\alpha\cdot\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$

Κατανομή Βήτα

$X \sim \mathrm{B}(\alpha, \beta)$, με $\alpha > 0,\ \beta > 0$, αν $f(x) \propto x^{\alpha-1}(1-x)^{\beta-1}, \quad 0 < x < 1$

$$\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx \equiv \mathrm{B}(\alpha, \beta)$$

ολοκλήρωμα του Euler α′ είδους

σταθερά κανονικοποίησης

$$\frac{x^{\alpha-1}(1-x)^{\beta-1}}{c} \to \text{πυκνότητα}$$

$$\Downarrow$$

$$f(x) = \frac{1}{\mathrm{B}(\alpha, \beta)} x^{\alpha-1}(1-x)^{\beta-1}, \quad 0 < x < 1$$

Παράδειγμα 0.0.3. Έστω $X_1, X_2, \ldots, X_n$ τυχαίο δείγμα $\mathcal{U}[0, \theta], \quad \theta > 0$

$$\left(\mathcal{U}(0, \theta) \to f_\theta(x) = \frac{1}{\theta}, \quad 0 < x < \theta\right)$$

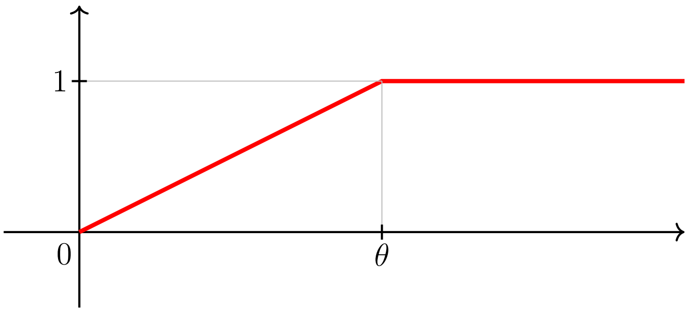
$$\left(\mathcal{U}[0, \theta] \to f_\theta(x) = \frac{1}{\theta}, \quad 0 \le x \le \theta\right)$$

Να βρεθεί η εκτιμήτρια ροπών του θ και να αναζητηθεί μία άλλη εμπειρική εκτιμήτρια.

Λύση

$\mathbb{E}_\theta(X_1) = \overline{X}$ άρα $\dfrac{\theta}{2} = \overline{X}$ άρα $\boxed{\overline{\theta} = 2\overline{X}}$

Παρατηρούμε ότι $\theta = \sup\{x \in \mathbb{R} : F_\theta(x) < 1\}$



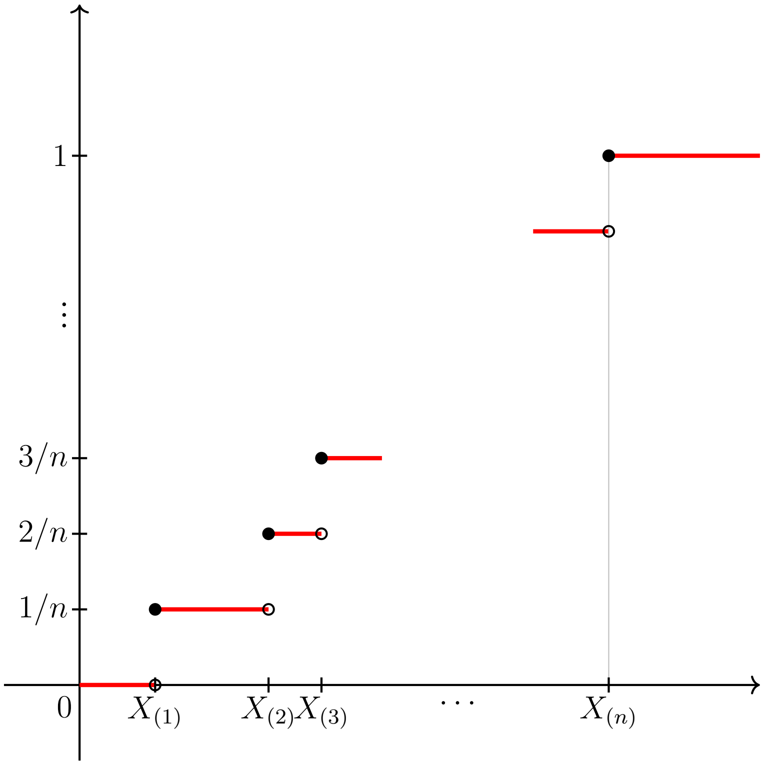
$\widehat{\theta} = \theta\,(F^*)$

πεπερασμένο σύνολο τιμών

$X_{(k)}$: k-διατεταγμένη τιμή των $X_1, X_2, \ldots, X_n$

$\sup \to \max \quad X_{(1)} = \min\{X_1, X_2, \ldots, X_n\}$

$X_{(n)} = \max\{X_1, X_2, \ldots, X_n\}$



$\sup\{x \in \mathbb{R} : F_\theta(X) < 1\}$
$\qquad\quad\downarrow$
$\sup\{x \in \mathbb{R} : F^*(X) < 1\} = \max\{X_1, X_2, \dots, X_n\} = \boxed{X_{(n)}}$

Άρα $\widehat{\theta} = X_{(n)}$. $\quad \left(\overline{\theta} = 2\overline{X}\right)$

εκτιμήτρια ροπών $2\overline{X}$ **vs** εμπειρική εκτιμήτρια $X_{(n)}$

Ποιά είναι η καλύτερη;

Πώς συγκρίνουμε δύο εκτιμήτριες;

Κριτήρια Σύγκρισης

Για να μπορέσουμε να κρίνουμε την ποιότητα μιας εκτιμήτριας κατ' αρχήν χρειαζόμαστε την έννοια του σφάλματος.

Σφάλματα Εκτιμητριών

Έστω $\widehat{\theta}$ μία εκτιμήτρια του $\theta \in \Theta$.

- **Σφάλμα**: $\widehat{\theta} - \theta, \ \forall \theta \in \Theta$: τυχαία συνάρτηση του θ
  $\downarrow \quad \downarrow$ (και τυχαία μεταβλητή και άγνωστη παράμετρο)
  τ.μ. άγνωστο

- **Απόλυτο Σφάλμα**: $\left|\widehat{\theta} - \theta\right|, \ \forall \theta \in \Theta$: τυχαία συνάστηση του θ

- **Τετραγωνικό Σφάλμα**: $\left(\widehat{\theta}-\theta\right)^2, \ \forall\theta\in\Theta$

Δεν είναι χρήσιμες από μόνες τους διότι είναι τυχαίες συναρτήσεις
(εξαρτώνται από την πραγματοποίηση)

$\Rightarrow$ θα ήταν καλύτερο να βλέπουμε τη μέση συμπεριφορά τους.

- **Μέσο Σφάλμα** $\rightarrow$ (κριτήριο μεροληψίας)

$b_{\widehat{\theta}}(\theta) = \mathbb{E}_\theta\left(\widehat{\theta}-\theta\right), \ \forall\theta\in\Theta = \mathbb{E}_\theta\left(\widehat{\theta}\right)-\theta$

λέγεται και **συνάρτηση μεροληψίας** της $\widehat{\theta}$ για το θ (ή **μεροληψία**)

Αν $b_{\widehat{\theta}}(\theta)=0, \ \forall\theta\in\Theta$, τότε η εκτιμήτρια λέγεται **αμερόληπτη** (α.ε.) για το θ.

Διαφορετικά λέγεται **μεροληπτική**

- **Μέσο Απόλυτο Σφάλμα** (Mean Absolute Error)

ΜΑΣ ή MAE : $\mathbb{E}_\theta\left|\widehat{\theta}-\theta\right|, \ \forall\theta\in\Theta.$

- **Μέσο Τετραγωνικό Σφάλμα** (Mean Square Error) [πιο συχνά]

ΜΤΣ ή MSE : $\mathbb{E}_\theta\left(\widehat{\theta}-\theta\right)^2$

Πρόταση 0.0.1.

$$\mathrm{MSE}_{\widehat{\theta}}(\theta) = b^2_{\widehat{\theta}}(\theta) + \mathbb{V}_\theta\left(\widehat{\theta}\right)$$

Απόδειξη

$$\mathrm{MSE} = \mathbb{E}_\theta\left[\underset{\substack{\big\downarrow \\ \text{τ.μ.}}}{\left(\widehat{\theta}-\theta\right)^2}\right] = \mathbb{E}^2_\theta\left(\widehat{\theta}-\theta\right) + \mathbb{V}_\theta\left(\underset{\substack{\big\downarrow \\ \text{τ.μ.}}}{\widehat{\theta}} - \underset{\substack{\big\downarrow \\ \text{σταθερά}}}{\theta}\right) = b^2_{\widehat{\theta}}(\theta) + \mathbb{V}_\theta\left(\widehat{\theta}\right)$$

$$\left[\begin{array}{l} \widehat{\theta}-\theta = X \\ \mathbb{E}(X^2) = \mathbb{E}^2(X) + \mathbb{V}(X) \\ \mathbb{V}(X-c) = \mathbb{V}(X) \end{array}\right]$$

Το $b^2_{\widehat{\theta}}(\theta)$ είναι συνάρτηση του θ

Σχόλιο Το ΜΤΣ είναι πιο εύκολο να το υπολογίζουμε ως $b^2+\mathbb{V}$

Διάλεξη 4

Είδαμε $\overline{X}_n,\ M_2,\ X_{(n)},\ X_{(1)}$. Αυτά είναι παραδείγματα σημαντικών εκτιμητριών (ε.ρ. + ε.ε.).

κριτήρια σύγκρισης $\to$ μεροληψία, MSE (και διασπορά)

Πρόταση 0.0.2. Αν $X_1, \ldots, X_n$ από κατανομή με $\mathbb{E}(X_1) = \mu$ και $\mathbb{V}(X_1) = \sigma^2$, τότε

(i) $\mathbb{E}\left(\overline{X}_n\right) = \mu,\ \forall \mu$ $\big($ο $\overline{X}_n$ είναι α.ε. του $\mu\big)$ και $\mathbb{V}\left(\overline{X}_n\right) = \dfrac{\sigma^2}{n}$.

(ii) $\mathbb{E}(M_2) = \dfrac{n-1}{n} \cdot \sigma^2$ (η M_2 μας οδηγεί σε μεροληπτική εκτιμήτρια του σ^2).

Ορίζουμε $S^2 = \dfrac{1}{n-1} \cdot \sum\limits_{i=1}^{n} \left(X_i - \overline{X}\right)^2$ είναι α.ε του σ^2 (διόρθωση μεροληψίας).

Τότε

$$\mathbb{E}(S^2) = \sigma^2, \quad \forall \sigma^2 > 0 \text{ και } \mathbb{V}(S^2) = \frac{\mu_4}{n} - \frac{n-3}{n(n-1)}\sigma^4,$$

με $n \geq 4,\ \mu_4 = \mathbb{E}(X-\mu)^4$ και υποθέτουμε ότι $\mathbb{E}(X_1^4) < +\infty$.

Απόδειξη

(i) $\mathbb{E}\left(\overline{X}\right) = \mathbb{E}\left(\dfrac{1}{n}\sum\limits_{i=1}^{n} X_i\right) = \dfrac{1}{n}\sum\limits_{i=1}^{n} \underbrace{\mathbb{E}(X_i)}_{=\mu} = \dfrac{1}{n} \cdot n\mu = \mu$ και

$$\mathbb{V}\left(\overline{X}\right) = \mathbb{V}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} \underbrace{\mathbb{V}(X_i)}_{=\sigma^2} = \frac{n}{n^2} \cdot \sigma^2 = \frac{\sigma^2}{n}.$$

(ii) $M_2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^2 = \overline{X^2} - \left(\overline{X}\right)^2$

$$\Rightarrow \mathbb{E}(M_2) = \mathbb{E}\left(\overline{X^2}\right) - \mathbb{E}\left[\left(\overline{X}\right)^2\right] \overset{\mathbb{E}(Y^2)=\mathbb{V}(Y)+\mathbb{E}^2(Y)}{=\!=\!=\!=\!=\!=\!=\!=} \mathbb{E}(X^2) - \left[\mathbb{V}\left(\overline{X}\right) + \mathbb{E}^2\left(\overline{X}\right)\right]$$

$$= \mathbb{V}(X) + \mathbb{E}^2(X) - \mathbb{V}\left(\overline{X}\right) - \mathbb{E}^2(X)$$

$$= \mathbb{V}(X) - \mathbb{V}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)$$

$$= \mathbb{V}(X) - \frac{1}{n^2}\sum_{i=1}^{n}\underbrace{\mathbb{V}(X_i)}_{=\sigma^2}$$

$$= \sigma^2 - \frac{n}{n^2}\cdot\sigma^2$$

$$= \frac{n-1}{n}\sigma^2.$$

$\mu_2 = \mathbb{E}(X-\mu)^2 = \sigma^2 \quad (\mu_2 \equiv \sigma^2)$ και $M_2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2, \quad \mathbb{E}(M_2) = \underbrace{\frac{n-1}{n}}_{\equiv c}\sigma^2$

$\Rightarrow S^2 \overset{\text{διορθ. μερολ}}{=\!=\!=\!=\!=} \frac{1}{c}M_2 = \frac{n}{n-1}\cdot\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 \ \mathbb{E}(S^2) = \sigma^2, \quad \forall \sigma^2 > 0$ (α.ε. του σ^2).

Είχαμε μια $\widetilde{\theta} \Rightarrow \mathbb{E}\left(\widetilde{\theta}\right) = c\cdot\theta, \quad \forall\theta$ και c σταθερά.

Άσκηση

Έστω τ.δ. $X_1, \ldots, X_n \sim \mathcal{N}(\mu, \sigma^2)$. Να συγκριθούν η μεροληψία και το MSE των M_2, S^2, $\frac{1}{n+1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)$. Ποια πετυχαίνει το καλύτερο MSE;

Παρατήρηση 0.0.8. Όταν για εκτιμήτρια U του $g(\theta)$ ισχύει ότι $\mathbb{E}_\theta(U) = g(\theta), \forall\theta \in \Theta$, τότε η U λέμε ότι είναι α.ε. του $g(\theta)$ και αν για κάποιο V έχουμε $\mathbb{E}_\theta(V) = c\cdot g(\theta)$, τότε η $U = \frac{1}{c}V$ είναι α.ε. του $g(\theta)$. Αυτό λέγεται διόρθωση μεροληψίας.

Λήμμα 0.0.1.
$$\sum_{i\neq j}(X_i - X_j)^2 = 2n\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2.$$
(Η $\overline{X}$ εξαρτάται από όλες τις X_i.)

Απόδειξη

$$\begin{aligned}\sum_{i\neq j}(X_i - X_j)^2 &= \sum_{i,j}(X_i - X_j)^2 = \sum_{i,j}\left(X_i - \overline{X} + \overline{X} - X_j\right)^2 \\ &= \sum_{i,j}\left(X_i - \overline{X}\right)^2 + \sum_{i,j}\left(X_j - \overline{X}\right)^2 + 2\sum_{i,j}\underbrace{\left(X_i - \overline{X}\right)\left(\overline{X} - X_j\right)}_{=0} \\ &= 2\sum_{i,j}\left(X_i - \overline{X}\right)^2 = 2\sum_i\sum_j\left(X_i - \overline{X}\right)^2 \\ &= 2n\sum_i\left(X_i - \overline{X}\right)^2.\end{aligned}$$

Από Λήμμα $\Rightarrow \sum_i \left(X_i - \overline{X}\right)^2 = \frac{1}{2n}\sum_{i\neq j}(X_i - X_j)^2.$

Για τη μέση τιμή έχουμε,

$$M_2 = \frac{1}{n}\sum_i\left(X_i - \overline{X}\right)^2 \Rightarrow M_2 = \frac{1}{2n^2}\sum_{i\neq j}(X_i - X_j)^2 \Rightarrow$$
$$\mathbb{E}(M_2) = \frac{1}{2n^2}\cdot\mathbb{E}\left(\sum_{i\neq j}(X_i - X_j)^2\right) = \frac{1}{2n^2}\cdot n(n-1)\mathbb{E}(X_1 - X_2)^2. \qquad (\star)$$

Όμως,

$$\mathbb{E}(X_1 - X_2)^2 = \mathbb{V}(X_1 - X_2) \overset{X_1 \perp X_2}{=} \mathbb{V}(X_1) + \mathbb{V}(X_2) \overset{\mathbb{V}(X_i)=\sigma^2}{=} 2\sigma^2,$$

γιατί η $X_1 - X_2$ έχει μέση τιμή $\underbrace{0}_{=\mu-\mu}$.

$(\star) \Rightarrow \mathbb{E}(M_2) = \frac{2(n-1)}{2n}\sigma^2 = \frac{n-1}{n}\sigma^2$. Άρα $S^2 = \frac{1}{n-1}\cdot\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$ είναι α.ε. του σ^2.

Επόμενο ερώτημα $\mathbb{V}(S^2) = ?$

Παρατήρηση 0.0.9.

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\Big[\underbrace{X_i-\mu}_{Y_i}-\underbrace{\overline{X-\mu}}_{\overline{Y}}\Big]^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2,$$

όπου $\mathbb{E}(Y_i)=0 \to$ κεντροποιημένες. $\Rightarrow$ χ.β.γ. υποθέτουμε $\mathbb{E}(X_i)=0$.

$$\begin{aligned}\mathbb{V}(S^2) &= \frac{1}{(n-1)^2}\mathbb{V}\left(\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2\right) = \frac{1}{(n-1)^2}\mathbb{V}\left(\frac{1}{2n}\sum_{i\neq j}(X_i-X_j)^2\right)\\ &= \frac{1}{4n^2(n-1)^2}\cdot\mathbb{V}\left[\sum_{i\neq j}(X_i-X_j)^2\right]. \qquad (1)\end{aligned}$$

Υπενθύμιση

$$\underbrace{\mathbb{V}\left(\sum_i X_i\right)}_{=\mathbb{C}\left(\sum_i X_i,\sum_j X_j\right)} = \sum_i \mathbb{V}(X_i)+\sum_{i\neq j}\mathbb{C}(X_i,X_j)$$

$$\begin{aligned}\mathbb{V}\left[\sum_{i\neq j}(X_i-X_j)^2\right] &= \sum_{i\neq j}\mathbb{V}(X_i-X_j)^2+\sum_{\substack{i\neq j\\ k\neq l\\ (i,j)\neq(k,l)}}\mathbb{C}\left[(X_i-X_j)^2,(X_k-X_l)^2\right]\\ &= n(n-1)\mathbb{V}(X_1-X_2)^2+n(n-1)\sum_{\substack{k\neq l\\ (k,l)\neq(1,2)}}\mathbb{C}\left[(X_1-X_2)^2,(X_k-X_l)^2\right]. \qquad (2)\end{aligned}$$

τρόπων που το $(1,2)$ έχει 1 μόνο κοινό σημείο με το (k,l) είναι $4(n-2)$.

τρόπων που το $(1,2)$ έχει 2 μόνο κοινά σημεία (χωρίς το $(1,2)=(k,l)$) με το (k,l) είναι 1.

Αν το $(1,2)$ δεν έχει κοινά σημεία με το (k,l) τότε το $\mathbb{C}(\cdot,\cdot)=0$.

Άρα,

$$\begin{aligned}\mathbb{V}\left[\sum_{i\neq j}(X_i-X_j)^2\right] &= n(n-1)\mathbb{V}(X_1-X_2)^2+n(n-1)\mathbb{C}\left[(X_1-X_2)^2,(X_2-X_1)^2\right]\\ &\quad+4n(n-1)(n-2)\mathbb{C}\left[(X_1-X_2)^2,(X_1-X_3)^2\right]\\ &= 2n(n-1)\mathbb{V}(X_1-X_2)^2+4n(n-1)(n-2)\mathbb{C}\left[(X_1-X_2)^2,(X_1-X_3)^2\right]. \qquad (3)\end{aligned}$$

Επίσης,

$$\mathbb{V}\left[(X_1 - X_2)^2\right] = 2(\mu_4 + \sigma^4), \tag{4}$$

$$\mathbb{C}\left[(X_1 - X_2)^2, (X_1 - X_3)^2\right] = \mu_4 - \sigma^4 \tag{5}$$

όπου με αντικατάσταση έχουμε $\mathbb{V}(S^2) = \dfrac{\mu_4}{n} - \dfrac{n-3}{n(n-1)} \cdot \sigma^4$.

Επιστρέφουμε στη σύγκριση εκτιμητριών. Στο $\mathcal{U}[0,\theta]$ έχουμε $\widetilde{\theta} = 2\overline{X}$ και $\widehat{\theta} = X_{(n)}$. Πώς συγκρίνονται;

Κατ' αρχήν,

$$\mathbb{E}\left(\widetilde{\theta}\right) = \mathbb{E}\left(2\overline{X}\right) = 2\underbrace{\mathbb{E}\left(\overline{X}\right)}_{\mu=\frac{\theta}{2}} = \theta \ \forall\theta > 0 \Rightarrow \widetilde{\theta} = 2\overline{X} \text{ είναι α.ε. του } \vartheta.$$

$\mathbb{E}(X_{(n)}) = ?$. Χρειαζόμαστε την κατανομή της $X_{(n)}$.

Γενικά για δειγματικά μέγιστα, αν X_i είναι συνεχείς, έχουν σ.π.π. ας πούμε f και σ.κ. F. $F_{X_{(n)}}(x) = F_{(n)}(x) = ?$

$$F_{(n)}(x) = \mathbb{P}(X_{(n)} \leq x), \quad \left[\max_{1\leq i\leq n} X_i \leq x \Leftrightarrow X_i \leq x \ \forall 1 \leq i \leq n\right].$$

$$\{X_{(n)} \leq x\} = \{\max\{X_1, \ldots, X_n\} \leq x\} = \bigcap_{i=1}^{n} \{X_i \leq x\} = \{X_1 \leq x, X_2 \leq x, \ldots, X_n \leq x\}$$

Άρα,

$$F_{(n)}(x) = \mathbb{P}\left(X_1 \leq x, X_2 \leq x, \ldots, X_n \leq x\right) \overset{\text{λόγω ανεξ.}}{=} \prod_{i=1}^{n} \underbrace{\mathbb{P}(X_i \leq x)}_{F(x)} \overset{\text{ισόνομες}}{=} F^n(x), \quad \forall x \in \mathbb{R}.$$

Τελικά,

$$f_{(n)}(x) = (F^n(x))' = nF^{n-1}(x)F'(x) = nF^{n-1}(x)f(x), \quad x \in \mathbb{R}$$

$X_i \sim \mathcal{U}[0,\theta] \ldots X_{(n)} \to f_{(n)}$

Παρατήρηση 0.0.10. Αν $X \sim \mathcal{U}[0,\theta]$, τότε εφαρμόζοντας μετασχηματισμό κλίμακας έχουμε $X = \theta \cdot U,\ U \sim \mathcal{U}[0,1]$.
Γενικά, αν $Y = \alpha X + \beta,\ Y$ συνεχής, τότε $f_Y(y) = \dfrac{1}{\alpha} \cdot f_X\left(\dfrac{Y-\beta}{\alpha}\right)$

$$X_{(n)} = \max\{X_1, \ldots, X_n\} = \max\{\theta \cdot U_1, \ldots, \theta \cdot U_n\} = \theta \cdot \max\{U_1, \ldots, U_n\} = \theta \cdot U_{(n)},$$

όπου $X_i = \theta \cdot U_i, \{U_1, \ldots, U_n\}_{1 \le i \le n}$ ανεξ. και $\mathcal{U}[0,1]$.

Άρα,

$$X_{(n)} = \theta \cdot U_{(n)} \Rightarrow \mathbb{E}(X_{(n)}) = \theta \cdot \mathbb{E}(U_{(n)}).$$

Επομένως,

$$\begin{aligned} f_{U_{(n)}}(u) &= n \cdot F^{n-1}_{\mathcal{U}[0,1]}(u) \cdot f_{\mathcal{U}[0,1]}(u), \forall u \in [0,1], \ (0, \text{ αλλιώς}) \\ &= n \cdot u^{n-1} \cdot 1, \forall u \in [0,1], \quad \big(F_{\mathcal{U}[0,1]}(u) = u, 0 \le u \le 1\big) \\ &= n \cdot u^{n-1} \\ &= n \cdot u^{n-1}(1-u)^{1-1}, (\text{μου προέκυψε Βήτα } \alpha x^{\alpha-1}(1-x)^{\beta-1}) \end{aligned}$$

$\Rightarrow U_{(n)} \sim \mathrm{B}(n, 1)$.

$$\mathbb{E}(U_{(n)}) = \frac{\alpha}{\alpha+\beta} = \frac{n}{n+1} \Rightarrow \mathbb{E}(X_{(n)}) = \frac{n}{n+1}\theta, \ \forall \theta > 0.$$

Είναι μεροληπτική εκτιμήτρια του θ και μάλιστα το υποεκτιμά γιατί $\dfrac{n}{n+1}\theta < \theta$.

Διάλεξη 5

Σταματήσαμε σε εκτίμηση από $\mathcal{U}[0,\theta]$ με $\overline{\theta} = 2\overline{X}$, που είναι α.ε. και $\widehat{\theta} = X_{(n)}$ με $\mathbb{E}(\widehat{\theta}) = \frac{n}{n+1} \cdot \theta, \ \forall \theta > 0$, δηλαδή υποεκτιμά το θ.

Με διόρθωση μεροληψίας έχουμε $\widetilde{\theta} = \frac{n+1}{n} \cdot \widehat{\theta} \to$ α.ε. του θ.
Επίσης, $b_{\widehat{\theta}}(\theta) = \mathbb{E}_\theta(\widehat{\theta}) - \theta = -\frac{1}{n+1} \cdot \theta$.

Σχόλιο Θα είχαμε την τάση να πούμε ότι η $\overline{\theta}$ είναι "καλύτερη" από την $\widehat{\theta}$. Όπως θα δούμε είναι χειρότερη.

Πρέπει να προσέχουμε το δίπολο μεροληψίας-διασποράς.

$$\text{ΜΤΣ}(\overline{\theta}) \overset{\text{α.ε.}}{=} \mathbb{V}(\overline{\theta}) = \mathbb{V}(2\overline{X}) = 4\mathbb{V}(\overline{X}) = 4 \cdot \frac{\theta^2}{12n} = \frac{\theta^2}{3n}$$

$$\text{ΜΤΣ}(\widetilde{\theta}) \overset{\text{α.ε.}}{=} \mathbb{V}(\widetilde{\theta}) = \mathbb{V}\left(\frac{n+1}{n} \cdot X_{(n)}\right) = \frac{(n+1)^2}{n^2}\mathbb{V}(X_{(n)}) \tag{6}$$

$$X_{(n)} = \theta \cdot U_{(n)} \text{ και } U_{(n)} \sim \mathcal{B}(\underbrace{n}_{=\alpha}, \underbrace{1}_{=\beta}), \text{ όπου } \mathbb{V}(\text{B}(\alpha,\beta)) = \frac{\alpha \cdot \beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.$$

Άρα, θα είναι

$$\mathbb{V}(X_{(n)}) = \theta^2 \cdot \mathbb{V}(U_{(n)}) = \theta^2 \cdot \frac{n}{(n+1)^2(n+2)} \tag{7}$$

Από (6), (7) έχουμε

$$\text{ΜΤΣ}(\widetilde{\theta}) = \frac{(n+1)^2}{n^2} \cdot \frac{n}{(n+1)^2(n+2)} \cdot \theta^2 = \frac{\theta^2}{n(n+2)}$$

$$\text{ΜΤΣ}(\widehat{\theta}) = \text{ΜΤΣ}(X_{(n)}) \overset{\text{όχι α.ε.}}{=} b^2_{X_{(n)}}(\theta) + \mathbb{V}(X_{(n)})$$

$$\begin{aligned}
\Rightarrow \text{ΜΤΣ}(X_{(n)}) &= \frac{\theta^2}{(n+1)^2} + \frac{n}{(n+1)^2(n+2)} \cdot \theta^2 \\
&= \left[\frac{1}{(n+1)^2} + \frac{n}{(n-1)^2(n+2)}\right] \cdot \theta^2 \\
&= \frac{2}{(n+1)(n+2)} \cdot \theta^2
\end{aligned}$$

Η τάξη σύγκλισης στο 0 είναι πιο σημαντική από μια μικρή διαφορά μεροληψίας.

Τελικά, έχουμε

Εκτιμήτρια	Μεροληψία	ΜΤΣ
$\overline{\theta}$	0	$\dfrac{\theta^2}{3n}$
$\widetilde{\theta}$	0	$\dfrac{\theta^2}{n(n+2)}$
$\widehat{\theta}$	$-\dfrac{\theta}{n+1}$	$\dfrac{2\theta^2}{(n+1)(n+2)}$

Ορισμός 0.0.3. Μια ακολουθία (α_n) θετικών πραγματικών αριθμών είναι $O(\beta_n)$,

αν $\dfrac{\alpha_n}{\beta_n} \leq M$, για κάποιο $M > 0$, από κάποιο n και πέρα.

Η (β_n) προτιμάται για την τάξη σύγκλισης της (α_n).

Εφαρμογή

$$\text{ΜΤΣ}(\overline{\theta}_n) = O\left(\frac{1}{n}\right) \quad \text{ενώ} \quad \text{ΜΤΣ}(\widehat{\theta}_n) = \text{ΜΤΣ}(\widetilde{\theta}_n) = O\left(\frac{1}{n^2}\right).$$

Ερωτήματα

1. Υπάρχουν πάντα α.ε. του $g(\theta)$; (υπαρξιακό θέμα)
2. Αν υπάρχουν, πώς τις βρίσκουμε; (μέθοδος εύρεσης-κατασκευής)
3. Όταν υπάρχουν πολλές α.ε. πώς τις συγκρίνουμε; (π.χ. κριτήριο διασποράς)
4. Σε μια συγκεκριμένη κλάση εκτιμητριών έχουν οι α.ε. το μικρότερο ΜΤΣ; (δίπολο μεροληψίας-διασποράς)$\to$ όχι γενικά

Απαντήσεις

1. Όχι
   π.χ. αν $X \sim \text{Be}(p)$, $0 < p < 1$ και έστω $g(p) = p^2$. Τότε ν.δ.ο. δεν υπάρχει α.ε. του p^2 ($\forall p \in (0,1)$, και άρα θέστε $U = U(X)$ κάποια υποψήφια α.ε. και φτάστε σε άτοπο).

Κατανομή χ^2_n

$\to$ τη λέμε κατανομή χ^2 με n βαθμούς ελευθερίας.

Ορισμός 0.0.4. Έστω X μια τ.μ. Θα λέμε ότι ακολουθεί την κατανομή χ^2 με n βαθμούς ελευθερίας, αν

$$X = \sum_{i=1}^{n} {Z_i}^2,$$

όπου $Z_i \sim \mathcal{N}(0,1)$ (τυπική κανονική) και $Z_1, \ldots, Z_n$ ανεξάρτητες.

[ασκ. στην R: να δείτε τη σ.π.π. της χ_n^2, για διάφορες τιμές του n.]
εφαρμογές→ στις εκτιμήτριες, στα διαστήματα εμπιστοσύνης και στους ελέγχους υποθέσεων.

Ερώτηση Έστω

$$Y_n = \frac{\sum\limits_{i=1}^{n} \overbrace{Z_i^2}^{=g(Z_i)}}{n},$$

όπου $Z_i \sim \mathcal{N}(0,1)$ και είναι και ανεξ. τ.μ. και γνωρίζουμε ότι

$$(Z_n)_{n>1} \text{ α.ι.τ.μ. } \Rightarrow g(Z_n)_{n\geq 1} \text{ α.ι.τ.μ. }.$$

Διάφοροι τρόποι σύγκλισης ακολουθίας τ.μ.

Έστω $W_i = Z_i^2,\ 1 \leq i \leq n$. Τότε,

$$Y_n = \frac{\sum\limits_{i=1}^{n} W_i}{n} = \overline{W}_n,$$

όπου $\overline{W}_n$ ο δειγματικός μέσος μιας ακολουθίας α.ι.τ.μ. και

$$\overline{W}_n \overset{a.s.}{\to} \mathbb{E}(W_1) = 1,$$

από τον Ι.Ν.Μ.Α. (ισχυρός νόμος μεγάλων αριθμών).

Έχουμε ότι

$$\mathbb{E}(W_1) = \mathbb{E}(Z_1^2) \underset{\mathbb{E}(Z_1)=0}{=} \mathbb{V}(Z_1) = 1.$$

a.s.: almost surely→ σχεδόν βεβαίως ≡ με πιθανότητα 1.

$$X_n \overset{a.s.}{\to} X \Leftrightarrow \mathbb{P}\Big(\{\omega \in \Omega : \lim \underbrace{X_n(\omega)}_{\text{ντετερμ. ακολ.}} = X(\omega)\}\Big) = 1,\ [X_n, X : \Omega \to \mathbb{R}]$$

$\big[$τυπικά οι ακολ. $X_n(\omega)$ συγκλίνουν στο $X(\omega)\big]$.

Γραφικά, όταν κάνουμε μια πραγματοποίηση, και βλέπουμε τη συμπεριφορά της ακολουθίας τότε αυτή θα συγκλίνει [όταν $X = 1$ ή c, τότε $\to c$].

Διάλεξη 6

Διερεύνηση της κατανομής χ_n^2

$$X = \sum_{i=1}^{n} Z_i^2,\ \{Z_i\}_{1\leq i\leq n} \text{ είναι ανεξ. } \mathcal{N}(0,1) \text{ και έστω } X = Z^2,\ n=1.$$

Κατ' αρχήν παρατηρούμε ότι η $X \geq 0$. Έστω $x > 0$. Τότε,

$$\begin{aligned}
F_X(x) &= \mathbb{P}(X \leq x) = \mathbb{P}(Z^2 \leq x) = \mathbb{P}(|Z| \leq \sqrt{x}) \\
&= \mathbb{P}(-\sqrt{x} \leq Z \leq \sqrt{x}) \overset{Z \text{ συνεχής}}{=} F_Z(\sqrt{x}) - F_Z(-\sqrt{x}) \\
&= \Phi(\sqrt{x}) - \Phi(-\sqrt{x}) \quad (\text{υπενθύμιση: } \Phi(-x) = 1 - \Phi(x))
\end{aligned}$$

$$\Rightarrow f_X(x) = F_X'(x) = \frac{1}{2\sqrt{x}}\left[\phi(\sqrt{x}) + \underbrace{\phi(-\sqrt{x})}_{=\phi(\sqrt{x})}\right] = \frac{1}{\sqrt{x}}\phi(\sqrt{x})$$

$$= \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{x}} e^{-\frac{1}{2}\cdot(\sqrt{x})^2} = \frac{1}{\sqrt{2\pi}} \cdot x^{-\frac{1}{2}} \cdot e^{-\frac{1}{2}\cdot x},$$

$$\left[\text{με } \phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}},\ x \in \mathbb{R}\right]$$

$$\propto x^{\alpha-1}e^{-\theta x}, \text{ όπου } \alpha = \frac{1}{2},\ \theta = \frac{1}{2} \text{ που αντιστοιχεί σε } \mathrm{G}\left(\frac{1}{2}, \frac{1}{2}\right)$$

Καταλήξαμε ότι $Z^2 \sim X_1^2 \equiv \mathrm{G}\left(\frac{1}{2}, \frac{1}{2}\right) \Rightarrow$ αν $X \sim \chi_n^2$, τότε X αντιστοιχεί σε άθροισμα $\mathrm{G}\left(\frac{1}{2}, \frac{1}{2}\right)$ n-το πλήθος.

$$\mathrm{G}(\alpha_1, \theta) + \mathrm{G}(\alpha_2, \theta) + \ldots + \mathrm{G}(\alpha_n, \theta) \xrightarrow{+\text{ανεξ.}} \mathrm{G}\left(\sum_{i=1}^{n} \alpha_i, \theta\right)$$

$\Rightarrow \chi_n^2 \equiv \mathrm{G}\left(\frac{n}{2}, \frac{1}{2}\right).$

Υπενθύμιση

- Αν $Y = g(X)$, τότε αν η g δεν είναι "1-1", τότε δουλεύουμε με τη σ.κ. F_Y και τη βρίσκουμε ως συνάρτηση της F_X.

- Αν g είναι "1-1", g, g^{-1} είναι παραγωγίσιμες τότε για X συνεχή τ.μ.

$$f_Y(y) = f_X\left(g^{-1}(y)\right) \cdot \left|\frac{dg^{-1}(y)}{dy}\right|, \quad y \in S_Y,$$

όπου $S_Y = \{y \in \mathbb{R} : f_Y(y) > 0\}$
$\downarrow$
κατανομή που εξαρτάται από το θ

Εφαρμογές

① Έστω $X_1, X_2, \ldots, X_n$ τ.δ. από $\mathcal{N}(\mu, \sigma^2)$, όπου μ γνωστό και σ^2 άγνωστο. Τότε $\widehat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$.
(αφού το μ είναι γνωστό ⇒ παραμένει εκτιμήτρια. Άρα δεν έχει νόημα να το εκτιμήσουμε γιατί έτσι θα μεγαλώναμε τη μεταβλητότητα της εκτιμήτριας)

$$\frac{n\widehat{\sigma^2}}{\sigma^2} = \sum_{i=1}^{n}\left(\frac{X_i - \mu}{\sigma}\right)^2 = \sum_{i=1}^{n} Z_i^2 \sim \chi_n^2, \quad Z_i \sim \mathcal{N}(0,1) \text{ και ανεξ.}$$

$$\boxed{\frac{n\widehat{\sigma^2}}{\sigma^2} \sim \chi_n^2} \quad \text{ή} \quad \boxed{\widehat{\sigma^2} \sim \frac{\sigma^2}{n} \cdot \chi_n^2}$$

② Έστω $X_1, X_2, \ldots, X_n$ τ.δ. από $\mathcal{N}(\mu, \sigma^2)$, όπου μ, σ^2 άγνωστα.

$$(n-1)S^2 = \sum_{i=1}^{n}(X_i - \overline{X})^2 (\text{ Το } \overline{X} \text{ είναι απαραίτητο, αφού } \mu \text{ άγνωστο})$$

$$\Rightarrow \frac{(n-1)S^2}{\sigma^2} = \sum_{i=1}^{n}\left(\frac{X_i - \overline{X}}{\sigma}\right)^2 \sim \chi_{n-1}^2.$$

Προσοχή $\frac{X_i - \overline{X}}{\sigma} \nsim \mathcal{N}(0,1)$ και δεν έχουμε και ανεξαρτησία. Πρακτικά θα λέγαμε ότι επειδή εκτιμάμε το μ, χάνεται ένας βαθμός ελευθερίας, διότι δημιουργείται μία σχέση εξάρτησης στις $Y_i = X_i - \overline{X}$ (αθροίζουν στο 0).

③ Για τ.δ. από $\mathcal{N}(\mu, \sigma^2)$

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1} \equiv \mathbf{G}\left(\frac{n-1}{2}, \frac{1}{2}\right)$$

$$\Rightarrow S^2 \sim \frac{\sigma^2}{n-1} \cdot \mathbf{G}\left(\underbrace{\frac{n-1}{2}}_{=\alpha}, \underbrace{\frac{1}{2}}_{=\beta}\right)$$

Τελικά για τ.δ. $\mathcal{N}(\mu, \sigma^2)$

$$\mathbb{E}(S^2) = \frac{\sigma^2}{n-1} \cdot \frac{n-1}{2} \cdot 2 = \sigma^2 \quad \forall \sigma^2 > 0$$

$\Rightarrow S^2$ είναι α.ε. του σ^2.

$$\text{ΜΤΣ}(S^2) = \mathbb{V}(S^2) = \left(\frac{\sigma^2}{n-1}\right)^2 \cdot \frac{n-1}{2} \cdot 4 = \frac{2}{n-1}\sigma^4 \quad \forall \sigma^2 > 0$$

αφού $\mathbb{E}\left(\mathbf{G}(\alpha, \theta)\right) = \dfrac{\alpha}{\theta}$ και $\mathbb{V}(\mathbf{G}(\alpha, \theta)) = \dfrac{\alpha}{\theta^2}$.

Σχόλιο Θυμηθείτε ότι

$$\mathbb{V}(S^2) = \frac{\mu_4}{n} - \frac{n-3}{n(n-1)}\sigma^4,$$

όπου $\mu_4 = \mathbb{E}\left[(X-\mu)^4\right] = 3\sigma^4$ για $\mathcal{N}(\mu, \sigma^2)$.

Πριν δούμε περιπτώσεις εύρεσης α.ε. και α.ε.ε.δ (α.ε. ελάχιστης διασποράς), θα απαιτούσαμε τουλάχιστον αν δεν είναι α.ε. τότε η μεροληψία, ΜΤΣ $\underset{n\to\infty}{\longrightarrow} 0$, καλές ασυμπτωματικές ιδιότητες μιας εκτιμήτριας.

Ορισμός 0.0.5. Μια εκτιμήτρια $\widehat{\theta}_n$ λέγεται **ασυμπτωματικά αμερόληπτη εκτιμήτρια** (α.α.ε.) του θ, αν $\mathbb{E}_\theta(\widehat{\theta}_n) \underset{n\to\infty}{\longrightarrow} \theta, \quad \forall \theta \in \Theta$ ή ισοδύναμα $b_{\widehat{\theta}_n}(\theta) \underset{n\to\infty}{\longrightarrow} 0, \quad \forall \theta \in \Theta$.

Σχόλιο Γενικότερα επεκτείνεται και για $\lambda = g(\theta), \quad \mathbb{E}_\theta(\widehat{\lambda}_n) \to \lambda = g(\theta), \quad \forall \theta \in \Theta$.

Ορισμός 0.0.6. Μια εκτιμήτρια $\widehat{\theta}_n$ του θ, λέγεται L^2-**συνεπής**, αν

$$\mathbb{E}_\theta(\widehat{\theta}_n - \theta)^2 \equiv MSE_{\widehat{\theta}_n}(\theta) \underset{n\to\infty}{\longrightarrow} 0, \quad \forall\theta \in \Theta.$$

Σχόλιο Έστω X_n, X τ.ω. $\mathbb{E}(X_n^2), \mathbb{E}X^2 < +\infty$. Τότε $X_n \xrightarrow{L^2} X$, αν $\mathbb{E}|X_n - X|^2 \to 0$, για $X_n = \widehat{\theta}_n$, $X = \theta$ με θ σταθερά.

Πρόταση 0.0.3. Αν $MSE_{\widehat{\theta}_n}(\theta) \to 0$, $\forall\theta \in \Theta$, τότε η $\widehat{\theta}_n$ είναι α.α.ε. του θ.

Απόδειξη

$$MSE_{\widehat{\theta}_n}(\theta) = b^2_{\widehat{\theta}_n}(\theta) + \mathbb{V}(\widehat{\theta}_n) \Rightarrow 0 \le b^2_{\widehat{\theta}_n}(\theta) \le MSE_{\widehat{\theta}_n}(\theta) \xrightarrow[n\to\infty]{\text{υπόθ}} 0$$

$$\Rightarrow b_{\widehat{\theta}_n}(\theta) \to 0,\ \forall\theta \Rightarrow \text{ η } \widehat{\theta}_n \text{ είναι α.α.ε. του } \theta.$$

Παράδειγμα 0.0.4. Σε τ.δ. από $\mathcal{N}(\mu, \sigma^2)$, μ, σ^2 άγνωστα η $M_{2,n} = \dfrac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$ είναι μεροληπτική εκτιμήτρια του σ^2. Όμως είναι α.α.ε., διότι

$$\mathbb{E}(M_{2,n}) = \frac{n-1}{n} \cdot \sigma^2 \underset{n\to\infty}{\longrightarrow} \sigma^2.$$

Προφανώς α.ε. $\Rightarrow$ α.α.ε., άρα η S_n^2 είναι α.α.ε..

Παράδειγμα 0.0.5. Σε τ.δ. από $\mathcal{U}[0, \theta]$, η $\widehat{\theta}_n = X_{(n)}$ είναι μεροληπτική, διότι

$$\mathbb{E}(\widehat{\theta}_n) = \frac{n}{n+1} \cdot \theta \underset{n\to\infty}{\longrightarrow} \theta,\ \forall\theta > 0$$

άρα η $\widehat{\theta}_n$ είναι α.α.ε. του θ.

Ορισμός 0.0.7. Αν $\widehat{\theta}_n \xrightarrow{\text{a.s.}} \theta,\ \forall\theta \in \Theta$, τότε η $\widehat{\theta}_n$ λέγεται **ισχυρά συνεπής** εκτιμήτρια του θ.

- Μια ακολ. τ.μ. (X_n) λέμε ότι συγκλίνει **κατά πιθανότητα** ή **στοχαστικά** σε μία τ.μ. X και γράφουμε $X_n \xrightarrow{p} X$ (in-probability), αν

$$\forall\epsilon > 0, \quad \mathbb{P}(|X_n - X| > \epsilon) \xrightarrow[n\to\infty]{} 0$$

ή ισοδύναμα

$$\forall\epsilon > 0, \quad \mathbb{P}(|X_n - X| \leq \epsilon) \xrightarrow[n\to\infty]{} 1$$

- Αν $\widehat{\theta}_n \xrightarrow{p} \theta,\ \theta \in \Theta$, τότε η $\widehat{\theta}_n$ λέγεται (ασθενώς) συνεπής εκτιμήτρια του θ

$$\left[\forall\epsilon > 0, \quad \mathbb{P}(|\theta_n - \theta| > \epsilon) \xrightarrow[n\to\infty]{} 0\right]$$

Πρόταση 0.0.4. ισχυρά συνεπής $\Rightarrow$ συνεπής

Απόδειξη

ισχυρά συνεπής $\Rightarrow \widehat{\theta}_n \xrightarrow{\text{a.s.}} \theta, \quad \forall\theta \in \Theta$. Όμως $\xrightarrow{\text{a.s.}} \quad \Rightarrow \quad \xrightarrow{\text{p}}$. Άρα $\widehat{\theta}_n \xrightarrow{\text{p}} \theta, \quad \theta \in \Theta$, δηλ. η $\widehat{\theta}_n$ είναι συνεπής.

Παράδειγμα 0.0.6. Ας υποθέσουμε ότι έχουμε ένα τ.δ. $\mathcal{N}(\mu, \sigma^2),\ \mu, \sigma^2$ άγνωστα. Ο $\overline{X}_n$ είναι μια καλή εκτιμήτρια του μ και από τον Ι.Ν.Μ.Α. έχουμε ότι

$$\overline{X}_n \xrightarrow{\text{a.s.}} \mathbb{E}_\theta(X_1) = \mu,\ \forall\mu \in \mathbb{R}.$$

Ισχυρός Νόμος των Μεγάλων Αριθμών

Έστω (X_n) ακολουθία α.ι.τ.μ. με $\mathbb{E}\,|X_1| < +\infty$ (δηλαδή υπάρχει η μέση τιμή) και $\mu = \mathbb{E}(X_1)$. Τότε $\overline{X}_n \xrightarrow{\text{a.s.}} \mu$.

Έτσι λοιπόν με την προηγούμενη εφαρμογή έχουμε ότι ο δ.μ. $\overline{X}_n$ είναι ισχυρά συνεπής εκτιμήτρια του $\mu \Rightarrow \overline{X}_n$ είναι συνεπής εκτιμήτρια του μ.
[Ασθενής Νόμος των Μεγάλων Αριθμών: $\overline{X}_n \xrightarrow{\text{p}} \mu$]

Πρόταση 0.0.5. L^2-συνεπής $\Rightarrow$ συνεπής

Απόδειξη

Έστω $\epsilon > 0$.

$$\mathbb{P}\left(\left|\widehat{\theta}_n - \theta\right| > \epsilon\right) = \mathbb{P}\left[\underbrace{(\widehat{\theta}_n - \theta)^2}_{X} > \underbrace{\epsilon^2}_{\alpha}\right]$$

$$\overset{Markov}{\leq} \frac{\mathbb{E}_\theta(\widehat{\theta}_n - \theta)^2}{\epsilon^2} \xrightarrow[n\to\infty]{} 0 \text{ (από υπόθεση) } \forall \theta \in \Theta.$$

$$\boxed{\text{ανισότητα Markov: } \mathbb{P}(X > \alpha) \leq \frac{\mathbb{E}(X)}{\alpha},\ \alpha > 0}$$

Τελικά $\widehat{\theta}_n \xrightarrow{p} \theta,\ \forall \theta \in \Theta$ δηλαδή η $\widehat{\theta}_n$ είναι συνεπής εκτιμήτρια του θ.

Πόρισμα 0.0.2. Αν η $\widehat{\theta}_n$ είναι L^2**-συνεπής**, τότε η $\widehat{\theta}_n$ είναι **α.α.ε. + συνεπής**.
$\not\Leftarrow$ δεν ισχύει (σκεφτείτε αντιπαράδειγμα).

Παράδειγμα 0.0.7. Ας υποθέσουμε ότι $X_1, \ldots, X_n$ είναι τ.δ. από $\mathcal{U}[0, \theta]$. Είδαμε ότι η $\overline{\theta}_n = 2\overline{X}_n$ είναι α.ε. του θ. Επιπλέον η $\widehat{\theta}_n = \dfrac{n+1}{n} X_{(n)}$ είναι επίσης α.ε. του θ και η $X_{(n)}$ είναι α.α.ε. του θ, ενώ δεν είναι α.ε.

Είδαμε ότι MSE(εκτιμητριών) $\xrightarrow[n\to\infty]{} 0,\ \forall \theta \in \Theta.$ $\Rightarrow$ η $\overline{\theta}_n,\ \widetilde{\theta}_n,\ \widehat{\theta}_n$ είναι L^2-συνεπείς $\Rightarrow$ $\overline{\theta}_n,\ \widetilde{\theta}_n,\ \widehat{\theta}_n$ είναι και συνεπείς (μαζί με α.α.ε.). Είναι ισχυρά συνεπείς;

Θεώρημα Συνεχούς Απεικόνισης ($\xrightarrow{\text{a.s.}}$, $\xrightarrow{\text{p}}$)

Αν $X_n \xrightarrow{a.s./p} X$ και g είναι συνεχής, τουλάχιστον στο S_X (στο στήριγμα της X), τότε $g(X_n) \xrightarrow{a.s./p} g(X)$.

Συνέχεια Παραδείγματος

$$\overline{\theta}_n = 2\overline{X}_n,\ \overline{X}_n \xrightarrow{a.s.} \mathbb{E}(X_1) = \frac{\theta}{2},\ \forall \theta > 0$$

$$\overset{\text{Θ.Σ.Α.}}{\Rightarrow} \overline{\theta}_n \xrightarrow{a.s.} 2 \cdot \frac{\theta}{2} = \theta,\ \forall \theta > 0 \Rightarrow \overline{\theta}_n \text{ είναι ισχυρά συνεπής.}$$

Διαφορετικά,

$$\overline{\theta}_n = \overline{2X}_n = \overline{Y}_n \xrightarrow[\text{I.N.M.A.}]{a.s.} \mathbb{E}_\theta(Y_1) = 2\mathbb{E}_\theta(X_1) = 2 \cdot \frac{\theta}{2} = \theta, \ \forall \theta > 0.$$

[Συμληρώστε με $X_{(n)}$]

Αν ζητούσαμε μόνο συνέπεια της εκτιμήτριας

$$\overline{\theta}_n = 2\overline{X}_n = \underbrace{\overline{2X}_n}_{\text{α.ι.τ.μ.}} \xrightarrow[\text{A.N.M.A.}]{p} \mathbb{E}_\theta(2X_1) = 2 \cdot \frac{\theta}{2} = \theta, \ \forall \theta > 0$$

ή με Θ.Σ.Α. ($\xrightarrow{p}$).

π.χ. αν θέλουμε να εκτιμήσουμε το θ^2.

$g(\theta) = \theta^2$. Θέτουμε $\lambda = g(\theta) = \theta^2$. Έχουμε π.χ. $\widetilde{\theta}_n$ είναι ισχυρά συνεπής.

Άρα,

$\widetilde{\theta}_n \xrightarrow{a.s.} \theta, \ \forall \theta > 0$

$\Rightarrow$ Θέτουμε $\widetilde{\lambda}_n = g(\widetilde{\theta}_n) = (\widetilde{\theta}_n)^2 \xrightarrow{a.s.} \theta^2$ (λόγω Θ.Σ.Α. για τη x^2 που είναι συνεχής)

$\Rightarrow$ η $\widetilde{\lambda}_n$ είναι ισχυρά συνεπής εκτιμήτρια του $\lambda = g(\theta)$.

ΜΑΘΗΜΑΤΙΚΗ ΣΤΑΤΙΣΤΙΚΗ - QUIZ 1, 23 Οκτώβρη 2020

ΕΡΩΤΗΣΕΙΣ:

1. Για 2 κεντροποιημένες (ως προς μέση τιμή) και ανεξάρτητες τ.μ. X, Y, όταν $\mathbb{V}(X), \mathbb{V}(Y) < \infty$

   ◯ $\mathbb{V}(XY) = \mathbb{V}(X)V(Y)$ ◯ $\mathbb{V}(X - Y) = 0$ ◯ $\mathbb{V}(X - Y) = \mathbb{V}(X) - \mathbb{V}(Y)$

2. Όταν αναζητούμε αρχικά μία εκτιμήτρια ροπών, τότε αν υπάρχει λύση είναι πάντα

   ◯ εκτιμήτρια ◯ στατιστική συνάρτηση ◯ μοναδική

3. Αν $X^2 \sim Geo(p)$ στο $\{1, 2, \ldots\}$, τότε μία εκτιμήτρια ροπών του p είναι

   ◯ $\overline{X^2}$ ◯ $\frac{1}{\overline{X}^2}$ ◯ $\frac{1}{\overline{X^2}}$

4. Αν $\widehat{\theta}$ είναι αμερόληπτη εκτιμήτρια του θ, και g συνεχής τότε :

   ◯ $\mathbb{E}\left(g(\widehat{\theta})\right) = g(\theta)$ ◯ $g\left(\mathbb{E}(\widehat{\theta})\right) = g(\theta)$ ◯ $\mathbb{E}\left(g(\widehat{\theta})\right) \geq g(\theta)$

5. Αν έχουμε ένα n-τυχαίο δείγμα από $Unif[\theta, 0]$, $\theta < 0$, τότε α.ε. του θ είναι η :

   ◯ $\frac{n+1}{n}X_{(1)}$ ◯ $X_{(1)}$ ◯ $-\frac{n+1}{n}X_{(1)}$

6. Αν $X = \mu + Z$, όπου $\sim$ $Cauchy(0, 1)$ και μ άγνωστο, τότε εμπειρικά το εκτιμάμε με

   ◯ τη δειγματικό μέσο ◯ το δειγματικό εύρος ◯ τη δειγματική διάμεσο

7. Το ΜΤΣ μίας εκτιμήτριας U του $g(\theta)$ είναι

   ◯ τυχαία μεταβλητή ◯ τυχαία συνάρτηση του θ ◯ συνάρτηση του θ

8. Αν μία εκτιμήτρια $\widehat{\theta}_n$ του θ είναι L^2-συνεπής, συμπεραίνουμε ότι

   ◯ είναι ισχυρά συνεπής ◯ $\mathbb{E}(\widehat{\theta}_n) \xrightarrow[n\to\infty]{} \theta$ $(\forall\theta)$ ◯ $\mathbb{E}[(\widehat{\theta}_n)^3] \xrightarrow{n\to\infty} \theta^3$ $(\forall\theta)$

9. Αν μία εκτιμήτρια $\widehat{\theta}_n$ του θ είναι ασυμπτωτικά αμερόληπτη, συμπεραίνουμε ότι

   ◯ $\mathbb{E}(\widehat{\theta}_n) \xrightarrow[n\to\infty]{} \theta$ $(\forall\theta)$ ◯ έχει μηδενική μεροληψία ◯ είναι L^2-συνεπής

10. Ο Ασθενής Νόμος των Μεγάλων Αριθμών εμφανίζεται στις εκτιμήτριες ως

    ◯ ισχυρή συνέπεια ◯ L^2-συνέπεια ◯ συνέπεια

ΚΑΛΗ ΕΠΙΤΥΧΙΑ!

Απαντήσεις

1	Α
2	Β
3	Γ
4	Β
5	Α
6	Γ
7	Γ
8	Β
9	Α
10	Γ

Σχολιασμός του Quiz 1

① Για δύο κεντροποιημένες (ως προς μέση τιμή) και ανεξάρτητες τ.μ. X, Y όταν $\mathbb{V}(X), \mathbb{V}(Y) < +\infty$ ισοδύναμα $\mathbb{E}(X^2), \mathbb{E}(Y^2) < +\infty$, ισχύει:

$$\boxed{\mathbb{V}(XY) = \mathbb{V}(X)\,\mathbb{V}(Y)}$$

Κεντροποιημένη: $X \mapsto X - \mathbb{E}(X) =: \widetilde{X}$, $\mathbb{E}\left(\widetilde{X}\right) = 0$.

Τυποποιημένη: $X \mapsto \dfrac{X - \mathbb{E}(X)}{\sqrt{\mathbb{V}(X)}} =: \widetilde{X}$, $\mathbb{E}\left(\widetilde{X}\right) = 0$, $\mathbb{V}\left(\widetilde{X}\right) = 1$.

Απόδειξη

Για κεντροποιημένες τ.μ. είναι $\mathbb{V}(X) = \mathbb{E}(X^2)$ και $\mathbb{V}(Y) = \mathbb{E}(Y^2)$. Άρα:

$$\begin{aligned}
\mathbb{V}(XY) &= \mathbb{E}\left(X^2Y^2\right) - \mathbb{E}^2(XY) \\
&= \mathbb{E}\left(X^2\right)\mathbb{E}\left(Y^2\right) - \mathbb{E}^2(X)\,\mathbb{E}^2(Y) \\
&= \mathbb{V}(X)\,\mathbb{V}(Y)
\end{aligned}$$

Σχόλια

- Η κεντροποίηση είναι απαραίτητη συνθήκη για να ισχύει η ανωτέρω ισότητα.
- Αν X, Y, ανεξάρτητες τ.μ. τότε για κάθε $g, h : \mathbb{R} \to \mathbb{R}$ συναρτήσεις ισχύει ότι οι τ.μ. $g(X), h(Y)$ είναι επίσης ανεξάρτητητες.
- Ισχύει προφανώς $\mathbb{E}(XY) = \mathbb{E}(X)\,\mathbb{E}(Y)$ και $\mathbb{E}(g(X)\,h(Y)) = \mathbb{E}(g(X))\,\mathbb{E}(h(Y))$, για κάθε $g, h : \mathbb{R} \to \mathbb{R}$.
- Υπενθυμίζεται ότι αν X, Y, ανεξάρτητες τ.μ. τότε $\mathbb{V}(X - Y) = \mathbb{V}(X) + \mathbb{V}(Y)$.

② Όταν αναζητούμε αρχικά μία εκτιμήτρια ροπών, τότε αν υπάρχει λύση **είναι στατιστική συνάρτηση**.

Πράγματι, το αποτέλεσμα είναι συνάρτηση του τυχαίου δείγματος. Η εκτιμήτρια απαιτεί η σ.σ. να παίρνει τιμές εντός του Θ ($T_n \in \Theta$). Δεν υπάρχει εξασφάλιση αρχικά ότι θα δίνει τιμές στο Θ. Π.χ. θυμηθείτε ότι κάνουμε ελέγχους περιορισμών και διορθώσεις π.χ. $N \to$ ακέραια παράμετρος [ένα μειονέκτημα της μεθόδου των ροπών]. Επίσης, η λύση δεν είναι κατ' ανάγκη μοναδική.

③ Αν $X^2 \sim \mathrm{Geo}\,(p)$ στο $\{1,2,3,\ldots\}$ τότε $\overline{p} = \dfrac{1}{\overline{X^2}}$.

Απόδειξη

Αν $Y_i = X_i^2$ είναι $\overline{Y} = \overline{X^2}$. Επομένως:

$$\mathbb{E}\,(\mathrm{Geo}\,(p)) = \frac{1}{p} \Rightarrow \overline{p} = \frac{1}{\overline{Y}} = \frac{1}{\overline{X^2}}$$

④ Η σωστή απάντηση είναι $g\left(\mathbb{E}\left(\widehat{\theta}\right)\right) = g\,(\theta)$, για κάθε θ και για οποιαδήποτε g όχι κατ' ανάγκη συνεχή.

Σχόλια

- Αν ήταν $\mathbb{E}\left(g\left(\widehat{\theta}\right)\right) = g\,(\theta)$ για κάθε θ, τότε η $g\left(\widehat{\theta}\right)$ είναι α.ε. του $g\,(\theta)$. Μία συνεχής συνάρτηση g δεν μεταφέρει τις αμεροληψίες της $\widehat{\theta}$ στην $g\left(\widehat{\theta}\right)$ για την $g\,(\theta)$ [plug-in]. Αν όμως η g είναι γραμμική τότε αυτό ισχύει.
- Αν η g είναι κυρτή και υπάρχουν οι μέσες τιμές τότε $\mathbb{E}\left(g\left(\widehat{\theta}\right)\right) \geq g\left(\mathbb{E}\left(\widehat{\theta}\right)\right) = g\,(\theta)$ (θετική αμεροληψία). Αντίστοιχα, αν η g είναι κοίλη και υπάρχουν οι μέσες τιμές τότε $\mathbb{E}\left(g\left(\widehat{\theta}\right)\right) \leq g\left(\mathbb{E}\left(\widehat{\theta}\right)\right) = g\,(\theta)$ (αρνητική αμεροληψία). Οι ανίσοτητες αυτές οφείλονται στον Jensen (βλέπε Πιθανότητες).
  Για να είναι γνήσιες οι παραπάνω ανισότητες δεν πρέπει:
  i. Η X να είναι εκφυλισμένη (σταθερή με πιθανότητα 1).
  ii. Η $g : \Delta \to \mathbb{R}$ να είναι αφφινική, δηλαδή αν για κάθε $x, y \in \Delta$ και για κάθε $\lambda \in [0,1]$ ισχύει:

  $$g\,(\lambda x + (1-\lambda)\,y) = \lambda g\,(x) + (1-\lambda)\,g\,(y)$$

  Για παράδειγμα, μια τέτοια g είναι η γραμμική.

⑤ Η σωστή απάντηση είναι το $\dfrac{n+1}{n} X_{(1)}$.

Απόδειξη

Έστω ανεξάρτητες $X_1, \ldots, X_n \sim \mathcal{U}\,[\theta, 0]$ με $\theta < 0$. Είναι $X_1, \ldots, X_n < 0 \Rightarrow X_{(1)} < 0$ και $\theta < X_{(1)} \Rightarrow \theta < \mathbb{E}\left(X_{(1)}\right)$. Επιπλέον:

$$-X_1, \ldots, -X_n \sim \mathcal{U}\,[0, \lambda]$$

όπου $\lambda := -\theta > 0$. Από θεωρία α.ε. του λ είναι η $\dfrac{n+1}{n}\,(-X)_{(n)}$ δηλαδή η $-\dfrac{n+1}{n} X_{(1)}$. Επομένως, η α.ε. του θ είναι η $\dfrac{n+1}{n} X_{(1)}$.

[$\max\{-X_1, \ldots, -X_n\} = -\min\{X_1, \ldots, X_n\}$]

⑥ Η σωστή απάντηση είναι ότι η δειγματική διάμεσος είναι μία κατάλληλη εκτιμήτρια και αυτό διότι η κατανομή $\mathrm{Cauchy}\,(0,1)$ δεν έχει μέση τιμή.

7 Η σωστή απάντηση είναι: συνάρτηση του θ, αφού ΜΤΣ $= \mathbb{E}_\theta \left(\widehat{\theta}_n - \theta\right)^2 = g\left(\theta\right)$.

Παίρνουμε μέση τιμή $\Rightarrow$ φεύγει η τυχαιότητα.

$\overline{X}_n$ είναι τυχαία μεταβλητή

$\overline{X}_n - \mu$ είναι τυχαία συνάρτηση του μ (βλέπω και τ.μ. + άγνωστη παράμετρο)

$\omega \to \overline{X}_n(\omega) - \mu = \overline{x}_n - \mu,\ \overline{x}_n$ είναι πραγματικός αριθμός

άρα έχουμε μία συνάρτηση του μ

8 Αν μια εκτιμήτρια $\widehat{\theta}_n$ του θ είναι L^2-συνεπής, τότε $\mathbb{E}\left(\widehat{\theta}_n\right) \overset{n\to\infty}{\longrightarrow} \theta$, για κάθε θ.

Για να ήταν ισχυρά συνεπής θα έπρεπε επιπλέον $\mathbb{V}_\theta\left(\widehat{\theta}_n\right) \overset{n\to\infty}{\longrightarrow} 0$.

Επίσης, δεν μπορούμε να συμπεράνουμε ότι $\mathbb{E}\left(\left(\widehat{\theta}_n\right)^3\right) \overset{n\to\infty}{\longrightarrow} \theta^3$ για κάθε θ, διότι δεν γνωρίζουμε αν έχει ροπή τρίτης τάξης.

9 Αν μία εκτιμήτρια είναι ασυμπτωτικά αμερόληπτη, συμπεραίνουμε ότι $\mathbb{E}\left(\widehat{\theta}_n\right) \overset{n\to\infty}{\longrightarrow} \theta$, για κάθε θ.

Για να έχει μηδενική μεροληψία (α.ε. του θ) θα έπρεπε $b_{\widehat{\theta}_n} = (\theta) = \mathbb{E}_\theta\left(\widehat{\theta}_n\right) - \theta = 0$. Άρα $\Leftrightarrow \mathbb{E}_\theta\left(\widehat{\theta}_n\right) = \theta$ για κάθε θ.

10 Ο Α.Ν.Μ.Α. εμφανίζεται στις εκτιμήτριες ως συνέπεια. Αντίστοιχα, ο Ι.Ν.Μ.Α. εμφανίζεται στις εκτιμήτριες ως ισχυρή συνέπεια.

Διάλεξη 7

Διαφορά σύγκλισης $\xrightarrow{p}$ και $\xrightarrow{\text{a.s.}}$

Φ_1	Φ_2	Φ_3	
1	1	1	...
1 2	1 2	1 2	...
1 2 3	1 2 3	1 2 3	...
1 2 3 4	1 2 3 4	1 2 3 4	...
$\vdots$	$\vdots$	$\vdots$	...
1 2 3 ... n	1 2 3 ... n	1 2 3 ... n	...
$\vdots$	$\vdots$	$\vdots$	...

$$X_n = \begin{cases} 1, & \text{αν επιλεγεί το } 1 \text{ στη } n\text{-οστή δοκιμή (κερδίζει)} \\ \\ 0, & \text{αν επιλεγεί οποιοσδήποτε άλλος απο } \{2, \dots, n\} \end{cases}$$

Τότε $X_n \sim \text{Be}\left(\frac{1}{n}\right), \ n = 1, 2, \dots$ αποτελούν ανεξάρτητες τ.μ.

Καθώς το $n \uparrow$ γίνεται όλο και πιο σπάνιο να κερδίσουμε.

$$\mathbb{P}(X_n = 1) = \frac{1}{n} \xrightarrow[n\to\infty]{} 0.$$

Έστω $\varepsilon > 0 \quad (0 < \varepsilon < 1)$.

$$\mathbb{P}\left(|X_n - 0| > \varepsilon\right) = \mathbb{P}(X_n > \varepsilon) = \mathbb{P}(X_n = 1) \longrightarrow 0.$$

Έχουμε $X_n \xrightarrow{p} 0$. Γίνεται όλο και πιο σπάνιο να βρεθούμε ε-**μακριά** από το 0 (καθώς $n \to \infty$).
Έχουμε πολλούς φοιτητές και αν πραγματοποιούσαν το πείραμα όλοι, θα διαπιστώναμε ότι το ποσοστό αυτών που κερδίζουν μικραίνει συνέχεια (ποσοστό$\to 0$, καθώς $n \to \infty$).

Αν $X_n \xrightarrow{\text{a.s.}} 0$, τότε θα έπρεπε ο κάθε φοιτητής πρακτικά να σταματήσει να κερδίζει από κάποια δοκιμή και πέρα ($X_n(\omega) \to 0$, σχεδόν $\forall \omega$).
$X_n(\omega) \in \{0, 1\}$, θα είχαμε $X_n(\omega) = 0$, τελικά $\forall n$.

Είναι πιο ισχυρή από την $X_n \xrightarrow{p} 0$, η οποία δεν αποκλείει ότι κάθε φοιτητής μπορεί να κερδίζει απείρως συχνά, απλά προφανώς θα κερδίζει όλο και πιο "αραιά".

Εδώ πράγματι $X_n \overset{\text{a.s.}}{\nrightarrow} 0$.

Η $\xrightarrow{p}$ δεν εξετάζει την εξέλιξη των δοκιμών για κάθε άτομο ξεχωριστά, μια "πληθυσμιακή" ιδιότητα που δεν ακολουθεί το "ιστορικό" του κάθε ατόμου.

Η $\xrightarrow{\text{a.s.}}$ αφορά τη συμπεριφορά των ακολουθιών $(X_n(\omega))$.
Κοιτάει κάθε "άτομο" και παρακολουθεί την $X_n(\omega)$, καθώς $n \to \infty$.

Συγκλίνει ή όχι;

Θεώρημα 0.0.3. (Ισχυρή συνέπεια εκτιμήτριας)

Μία εκτιμήτρια $\widehat{\theta}_n$ είναι ισχύρα συνεπής εκτιμήτρια του θ, αν

$$\widehat{\theta}_n \xrightarrow{\text{a.s.}} \theta, \quad \forall \theta \in \Theta.$$

Παρατήρηση 0.0.11. Όποιο και να είναι το θ, **τυπικά** ή **πρακτικά** κάθε μονοπάτι (διαδρομή) της εκτίμησης του θ, θα οδηγήσει στο θ $\left(\widehat{\theta}_n(\omega) \to \theta\right)$.

Συνέπεια εκτιμήτριας $(\widehat{\theta}_n \xrightarrow{p} \theta)$
Όποιο και να είναι το θ, πρακτικά το ποσοστό των μονοπατιών που θα βρεθούν ε-κοντά στο θ, αυξάνει απεριόριστα καθώς $n \to \infty$.

Χρήσιμο Κριτήριο της $\xrightarrow{\text{a.s.}}$

Αν $\forall \varepsilon > 0$,

$$\sum_{n \geq 1} \mathbb{P}\left(|X_n - x| > \varepsilon\right) < +\infty, \quad \text{(πλήρη σύγκλιση)}$$

τότε $X_n \xrightarrow{\text{a.s.}} x$. $\left[\not\Leftarrow \text{ είναι ισχυρότερη από την } \xrightarrow{\text{a.s.}}\right]$
Παρ' όλα αυτά αν η (X_n) είναι ακολ. ανεξ. τ.μ. τότε ισχύει η $\Leftrightarrow$.

Παράδειγμα 0.0.8. Αν $X_n \sim \text{Be}\left(\frac{1}{n}\right), \quad n \geq 1$, τότε

$$X_n \xrightarrow{p} 0, \quad \text{αλλά} \quad X_n \overset{\text{a.s.}}{\nrightarrow} 0.$$

Λύση

Δείξαμε ότι $X_n \xrightarrow{p} 0$. Έστω $\varepsilon > 0$.

$$\sum_{n\geq 1}\mathbb{P}\left(|X_n-0|>\epsilon\right)\overset{0<\varepsilon<1}{=}$$
$$\sum_{n\geq 1}\mathbb{P}(X_n=1)=$$
$$\sum_{n\geq 1}\frac{1}{n}=+\infty.$$

Γενικά αν σειρά $= +\infty \;\not\Rightarrow\; \overset{\text{a.s.}}{\not\to}$.

Εδώ όμως (X_n) είναι ανεξάρτητες τ.μ.. Άρα πράγματι $X_n \overset{\text{a.s.}}{\not\to} 0$.

Παράδειγμα 0.0.9. Έστω $X_1,\ldots,X_n$ ανεξάρτητες παρατηρήσεις από Be $\left(\frac{p}{n}\right)$, όπου $0<p\leq 1$ άγνωστο. Να δείξετε ότι η $\widehat{p}_n=\dfrac{X_1+X_2+\ldots+X_n}{1+\frac{1}{2}+\ldots+\frac{1}{n}}$ είναι ισχυρά συνεπής εκτιμήτρια του p. Συμπεράνετε ότι η $\widetilde{p}_n=\widehat{p}_n+X_n$ είναι συνεπής, αλλά όχι ισχυρά συνεπής εκτιμήτρια του p.

Λύση

Θα χρειαστούμε το ακόλουθο λήμμα.

Λήμμα 0.0.4. Αν (X_n) ανεξάρτητες τ.μ. με $\mathbb{E}(X_n)=0$ και $b_n\uparrow+\infty$ (ντετερμινιστική ακολουθία) με
$$\sum_{n=1}^{+\infty}\frac{\mathbb{V}(X_n)}{b_n^2}<+\infty\Rightarrow\frac{X_1+X_2+\ldots+X_n}{b_n}\xrightarrow{\text{a.s.}}0$$
(στηρίζεται, μεταξύ άλλων, στον I.N.M.A. του Kolmogorov για ανεξ. τ.μ.).

Επίσης θυμόμαστε το **Κριτήριο συμπύκνωσης του Cauchy** για σύγκλιση σειρών.

Πρόταση 0.0.6. Αν (a_n) ακολουθία θετικών αριθμών, τότε
$$\sum_n a_n<+\infty\Leftrightarrow\sum_n 2^n a_{2^n}<+\infty$$

Θέτουμε $b_n=1+\frac{1}{2}+\ldots+\frac{1}{n}$ και γνωρίζουμε ότι
$$b_n-\log n\to c$$
και επομένως
$$b_n\sim\log n\qquad\left(\lim_{n\to\infty}\frac{b_n}{\log n}=1\right)$$

και άρα μπορούμε να την αντικαταστήσουμε.

Θέτουμε $Y_n = X_n - \frac{p}{n}$, $\left(\mathbb{E}(X_n) = \frac{p}{n}\right)$, $n \geq 1$, έχουμε $\mathbb{E}(Y_n) = 0$, $\forall n \geq 1$ και η (Y_n) είναι ακολουθία ανεξάρτητων τ.μ..
Θα δείξουμε ότι

$$\sum_{n\geq 1} \frac{\mathbb{V}(Y_n)}{\log^2(n)} < +\infty \xRightarrow[\log n \sim b_n]{\text{Λήμμα}} \frac{Y_1 + Y_2 + \ldots + Y_n}{b_n} \xrightarrow{\text{a.s.}} 0.$$

$$\mathbb{V}(Y_n) = \mathbb{V}\left(X_n - \frac{p}{n}\right) = \mathbb{V}(X_n) \overset{X_n \sim \text{Be}\left(\frac{p}{n}\right)}{=\!=\!=} \frac{p}{n}\left(1 - \frac{p}{n}\right)$$

Άρα

$$\sum_{n=2}^{+\infty} \frac{\frac{p}{n}\left(1 - \frac{p}{n}\right)}{\log^2 n} \leq p \sum_{n=2}^{+\infty} \frac{1}{n \log^2 n} < +\infty,$$

αφού, από το κριτήριο σύγκλισης του Cauchy

$$\sum_{n=2}^{+\infty} \frac{2^n}{2^n \left(\log 2^n\right)^2} \overset{\log 2^n = n \log 2}{=\!=\!=} \sum_{n=1}^{+\infty} \frac{1}{n^2 \log^2 2} < +\infty.$$

Επομένως ισχύει το Λήμμα για $Y_n = X_n - \frac{p}{n}$.

Τελικά

$$\frac{X_1 + X_2 + \ldots + X_n - \mathbb{P}\left(1 + \frac{1}{2} + \ldots + \frac{1}{n}\right)}{1 + \frac{1}{2} + \ldots + \frac{1}{n}} = \frac{X_1 + \ldots + X_n}{1 + \ldots + \frac{1}{n}} - p \xrightarrow{\text{a.s.}} 0.$$

Άρα

$$\widehat{p}_n = \frac{X_1 + \ldots + X_n}{1 + \frac{1}{2} + \ldots + \frac{1}{n}} \xrightarrow{\text{a.s.}} p, \quad \forall p$$

άρα η $\widehat{p}_n$ είναι ισχυρά συνεπής εκτιμήτρια.

Όμως η $\widetilde{p}_n = \widehat{p}_n + X_n$, $X_n \sim \text{Be}\left(\frac{p}{n}\right)$, άρα $\left(\widehat{p}_n \xrightarrow{\text{a.s.}} p \Rightarrow \widehat{p}_n \xrightarrow{p} p\right)$

$$\left.\begin{array}{l} \widehat{p}_n \xrightarrow{p} p \\ \\ X_n \xrightarrow{p} 0 \end{array}\right\} \Rightarrow \widetilde{p}_n \xrightarrow{p} p \Rightarrow \text{η } \widetilde{p}_n \text{ είναι συνεπής.}$$

Από την άλλη όμως $X_n \overset{\text{a.s.}}{\nrightarrow} 0$.

Άρα $\widetilde{p}_n \xrightarrow{\text{a.s.}} p$, $p \in (0, 1]$.

Αν συνέκλινε, τότε $\widetilde{p}_n \xrightarrow{\text{a.s.}} 0$, $\widehat{p}_n \xrightarrow{\text{a.s.}} p \Rightarrow X_n \xrightarrow{\text{a.s.}} -p \Rightarrow$ **άτοπο**.

Τελικά η $\widetilde{p}_n$ **δεν είναι** ισχυρά συνεπής.

Ιδιότητες

Αν $X_n \xrightarrow{\text{a.s.}/p} X$ και $Y_n \xrightarrow{\text{a.s.}/p} Y$, τότε:

(i) $X_n + Y_n \xrightarrow{\text{a.s.}/p} X + Y$

(ii) $X_n Y_n \xrightarrow{\text{a.s.}/p} XY$

(iii) $X_n / Y_n \xrightarrow{\text{a.s.}/p} X/Y$ (όταν επιτρέπεται)

π.χ.
$\widetilde{p}_n = \widehat{p}_n + X_n \xrightarrow{p} p + 0 \quad \left(\widehat{p}_n \xrightarrow{p} p, \quad X_n \xrightarrow{p} 0\right)$

Παράδειγμα 0.0.10. Έστω $X_1, \ldots, X_n$ τυχαίο δείγμα από $\text{Exp}(\theta)$, $\theta > 0$. Θέλουμε να εκτιμήσουμε το θ, τη μέση τιμή και τη διασπορά της $\text{Exp}(\theta)$, και να εξετάσουμε ιδιότητες των εκτιμητριών (αμεροληψία, ΜΤΣ, L^2-συνέπεια, συνέπεια).

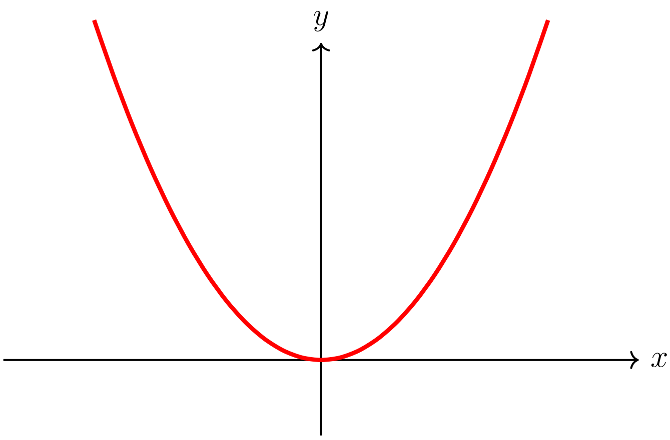
Λύση

$\overline{\theta}_n = \dfrac{1}{\overline{X}_n} = g\left(\overline{X}_n\right)$, όπου $g(x) = \frac{1}{x}$, $x > 0$ (g συνεχής και κυρτή).

Ποιοτικά μεροληψία

$$\mathbb{E}\left(\overline{\theta}_n\right) = \mathbb{E}\left(\frac{1}{\overline{X}_n}\right) = \mathbb{E}\left[g\left(\overline{X}_n\right)\right] \overset{g \text{ κυρτή}}{\geq} g\left[\mathbb{E}\left(\overline{X}_n\right)\right] = g\left(\frac{1}{\theta}\right) = \theta, \quad \forall \theta > 0.$$

Μάλιστα η ισότητα ισχύει αν η τ.μ. $\overline{X}_n$ είναι εκφυλισμένη ή αν είναι αφφινική $(ax + b)$.
Τελικά $\mathbb{E}\left(\overline{\theta}_n\right) > \theta$, $\forall \theta > 0$ δηλαδή έχει **θετική μεροληψία**!



$\mathbb{E}\left(g(x)\right) \geq g\left(\mathbb{E}(x)\right) = g(0)$

$\widehat{\theta}_n = \dfrac{1}{\overline{X}_n}$

Όμως

$$\overline{X}_n = \frac{X_1 + X_2 + \ldots + X_n}{n} \overset{X_i \sim \mathrm{Exp}(\theta)}{\sim} \frac{1}{n}\mathrm{G}(n,\theta) = \frac{1}{n\theta}\mathrm{G}(n,1), \quad \mathrm{G}(n,1) \equiv \mathrm{Erlang}(n,1).$$

Άρα

$$\widehat{\theta}_n = n\theta \frac{1}{\mathrm{G}(n,1)} = n\theta \frac{1}{x}, \quad x \sim \mathrm{G}(n,1). \tag{8}$$

$$\mathbb{E}\left(\frac{1}{x}\right)^{\kappa} = \mathbb{E}\left(x^{-\kappa}\right) =$$

$$\int_0^{+\infty} \frac{1}{\Gamma(n)} x^{-\kappa} x^{n-1} e^{-x} dx =$$

$$\int_0^{+\infty} \frac{1}{\Gamma(n)} x^{n-\kappa-1} e^{-x} dx =$$

$$\frac{\Gamma(n-\kappa)}{\Gamma(n)} \underbrace{\int_0^{+\infty} \frac{1}{\Gamma(n-\kappa)} x^{n-\kappa-1} e^{-x} dx}_{=1} \overset{\kappa \leq n}{=}$$

$$\frac{\Gamma(n-\kappa)}{\Gamma(n)} = \frac{(n-\kappa-1)!}{(n-1)!} = \frac{1}{(n-1)\ldots(n-\kappa)} \tag{9}$$

Επίσης

$$\mathbb{V}\left(\frac{1}{x}\right) = \mathbb{E}\left(\frac{1}{x}\right)^2 - \mathbb{E}\left(\frac{1}{x}\right) = \frac{1}{(n-1)(n-2)} - \frac{1}{(n-2)^2} =$$

$$\frac{1}{n-1}\left(\frac{1}{n-2} - \frac{1}{n-1}\right) = \frac{1}{(n-1)^2(n-2)}. \tag{10}$$

Τελικά από τις (8), (9) και (10)

$$\mathbb{E}(\widehat{\theta}_n) = n\theta \mathbb{E}\left(\frac{1}{x}\right) = \frac{n\theta}{n-1} = \theta + \frac{\theta}{n-1}$$

$$\mathbb{V}_\theta(\widehat{\theta}_n) = n^2\theta^2 \mathbb{V}\left(\frac{1}{x}\right) = \frac{n^2\theta^2}{(n-1)^2(n-2)} = \frac{n^2}{(n-1)^2(n-2)}\theta^2$$

Άρα

$$b_{\widehat{\theta}_n}(\theta) = \mathbb{E}_\theta(\widehat{\theta}_n) - \theta = \frac{\theta}{n-1}, \quad \forall \theta > 0.$$

$$\mathrm{MT\Sigma}_{\widehat{\theta}_n}(\theta) = b^2_{\widehat{\theta}_n}(\theta) + \mathbb{V}_\theta(\widehat{\theta}_n) = \frac{\theta^2}{(n-1)^2} + \frac{n^2\theta^2}{(n-1)^2(n-2)} = \frac{(n+2)\theta^2}{(n-1)(n-2)}, \quad \forall \theta > 0.$$

Έχουμε $\mathrm{MT\Sigma}_{\widehat{\theta}_n}(\theta) \xrightarrow[n\to\infty]{} 0 \Rightarrow$ η $\widehat{\theta}_n$ είναι L^2-**συνεπής** άρα και **συνεπής** εκτιμήτρια του θ. Μάλιστα $\widehat{\theta}_n = \dfrac{1}{\overline{X}_n} = g(\overline{X}_n) \xrightarrow{\text{a.s.}} g\left(\frac{1}{\theta}\right) = \theta, \quad \forall \theta > 0$ διότι η $g(x)$ είναι συνεχής και

$\overline{X}_n \xrightarrow{\text{a.s.}} \frac{1}{\theta}, \quad \forall \theta > 0$, από Ι.Ν.Μ.Α. και άρα εφαρμόζεται το Θ.Σ.Α.

Κάντε άσκηση μέση τιμή και διασπορά.

Παρατήρηση 0.0.12. Σε πολλά στατιστικά μοντέλα είναι χρήσιμο να κάνουμε **αναπαραμέτρηση** του μοντέλου έτσι ώστε η καινούρια παράμετρος να είναι πιο βολική για τους στατιστικούς μας σκοπούς. Λέμε αναπαραμέτρηση κάθε αλλαγή παραμέτρου μέσω μιας αμφιδιαφορίσιμης συνάρτησης με $\phi = \phi(\theta)$, όπου ϕ, ϕ^{-1} είναι διαφορίσιμες με

$$\phi : \Theta \to \Phi, \quad \text{όπου} \quad \Phi = \phi(\Theta).$$

Παράδειγμα 0.0.11. Η $\text{Exp}(\theta)$ έχει 2 τύπους:

- θ : παράμετρος **ρυθμού** (στις πιθανότητες).
- θ : παράμετρος **μέσης τιμής**.

Τότε αν θέσουμε $\mu = \mu(\theta) = \frac{1}{\theta}$, τότε η $\mu(\cdot)$ είναι αμφιδιαφορίσιμη και έτσι κάνουμε αναπαραμέτρηση

$$f_\theta(x) = \theta e^{-\theta x} \to \frac{1}{\mu} e^{-\frac{1}{\mu}x}, \quad x > 0$$

όπου $\mu = \frac{1}{\theta}$ η μέση τιμή της $\text{Exp}(\theta)$. Η “κακή” μορφή που παίρνει τώρα αντισταθμίζεται από το ότι

$$\mathbb{E}(X_1) = \mu.$$

$x \sim \text{Exp}(\theta)$

- Τύπου I $\quad \theta e^{-\theta x}$ (ρυθμό).
- Τύπου II $\quad \frac{1}{\theta} e^{-\frac{1}{\theta}x}$ (μέση τιμή).

Άσκηση

Δείξτε όλες τις ιδιότητες που κάναμε πριν (μέση τιμή, διασπορά).

$$\begin{aligned} \text{Τύπου II} \quad c \cdot \text{Exp}(\theta) &= \text{Exp}(c \cdot \theta) \\ c \cdot \text{G}(a, \theta) &= \text{G}(a, c \cdot \theta) \end{aligned}$$

Διάλεξη 8

Ορισμός 0.0.8. Έστω $X_1, X_2, \ldots, X_n$ τ.δ. από $f_\theta(x)$ και έστω U μία κλάση α.ε. κάποιας $g(\theta)$. Η εκτιμήτρια $u \in$ U λέγεται α.ε. **ελάχιστης διασποράς** (α.ε.ε.δ.) της $g(\theta)$ στην κλάση U, αν

$$\mathbb{V}_\theta(u) \leq \mathbb{V}_\theta(u'), \quad \forall \theta \in \Theta, \quad \forall u' \in \mathrm{U} \quad \text{(ομοιόμορφα ελάχιστης διασποράς)}$$

Παράδειγμα 0.0.12. Έστω $X_1, \ldots, X_n$ τ.δ. από κατανομή με $\mathbb{E}_\theta(X_1) = \mu(\theta) \in \mathbb{R}$.

(i) Να βρεθούν όλες οι **γραμμικές** α.ε. του $\mu(\theta)$.

(ii) Αν υποθέσουμε ότι $\mathbb{V}_\theta(X_1) = \sigma^2(\theta) < +\infty$, τότε να βρεθούν οι α.ε.ε.δ. στην κλάση των γραμμικών α.ε.

Λύση

(i) Οι γραμμικές εκτιμήτριες του $\mu(\theta)$ θα είναι $\sum_{i=1}^n \lambda_i X_i, \quad \lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$. Για να είναι α.ε. πρέπει

$$\mathbb{E}_\theta\left(\sum_{i=1}^n \lambda_i X_i\right) = \mu(\theta), \quad \forall \theta \in \Theta.$$

άρα

$$\sum_{i=1}^n \lambda_i \underbrace{\mathbb{E}_\theta(X_i)}_{\mu(\theta)} = \mu(\theta)$$

ή

$$\left(\sum_{i=1}^n \lambda_i\right) \mu(\theta) = \mu(\theta), \quad \forall \theta \in \Theta.$$

- Υπάρχει περίπτωση $\mu(\theta) = 0, \quad \forall \theta \in \Theta$.
  Τότε όλες είναι α.ε. του $\mu(\theta) \equiv 0$ (γνωστό)
  (π.χ. $\mathcal{U}[-\theta, \theta], \quad \theta > 0$) [δεν έχει ενδιαφέρον]

- Ενδιαφέρον περίπτωση: $\exists\theta : \mu(\theta) \neq 0$, και τότε $\sum_{i=1}^{n} \lambda_i = 1$
  Δείτε μάθημα 6 $\longrightarrow$Σημειώσεις $2017-18$ (4 λύσεις)

(ii) $f(\lambda) = \sum_{i=1}^{n} \lambda_i^2$

$\mathbb{V}_\theta\left(\sum_{i=1}^{n} \lambda_i X_i\right) \overset{\text{ανεξ.}}{=} \left(\sum_{i=1}^{n} \lambda_i^2\right)\sigma^2(\theta) = f(\lambda)\sigma^2(\theta), \quad (\sigma^2(\theta) > 0)$

Πρέπει να ελαχιστοποιήσουμε την $f(\lambda)$ υπό τον περιορισμό $\sum_{i=1}^{n} \lambda_i = 1 \Rightarrow \overline{\lambda} = \frac{1}{n}$.

$$f(\lambda) = n\frac{\sum_{i=1}^{n} \lambda_i^2}{n} = n\overline{\lambda^2} = n\left(\frac{\sum_{i=1}^{n}\left(\lambda_i - \overline{\lambda}\right)^2}{n} + \left(\overline{\lambda}\right)^2\right) =$$

$$\sum_{i=1}^{n}\left(\lambda_i - \frac{1}{n}\right)^2 + n\left(\frac{1}{n}\right)^2 = \sum_{i=1}^{n}\left(\lambda_i - \frac{1}{n}\right)^2 + \frac{1}{n} \geq \frac{1}{n}$$

$\Rightarrow$ για $\lambda_1 = \ldots = \lambda_n = \frac{1}{n}$, πετυχαίνουμε το ελάχιστο (μοναδικό), το $\frac{1}{n}$.

Τελικά η α.ε.ε.δ. του $\mu(\theta)$ από τις γραμμικές είναι η

$$\sum_{i=1}^{n} \frac{1}{n} X_i = \frac{X_1 + \ldots + X_n}{n} \longrightarrow \text{ ο δειγμ. μέσος}$$

Παρατήρηση 0.0.13. Η γνώση της κατανομής της εκτιμήτριας είναι πολύ σημαντική, όχι μόνο για να υπολογίζουμε ευκολότερα μέσες τιμές και διασπορές, αλλά και για να καθορίζουμε την ακρίβεια της εκτίμησης. Υπάρχουν περιπτώσεις που την υπολογίζουμε εύκολα π.χ. σε τ.δ. από $\mathcal{N}(\mu, \sigma^2)$

$$\overline{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

διότι άθροισμα **ανεξάρτητων κανονικών** $\mathcal{N}(\mu_i, \sigma_i^2)$

$$\sum_{i=1}^{n} \lambda_i X_i \sim \mathcal{N}\left(\sum_{i=1}^{n} \lambda_i \mu_i, \sum_{i=1}^{n} \lambda_i^2, \sigma_i^2\right).$$

Επιπλέον, αν

$$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2,$$

τότε

$$\frac{(n-1)S_n^2}{\sigma^2} \sim \chi_{n-1}^2 \Rightarrow S_n^2 \sim \frac{\sigma^2}{n-1}\chi_{n-1}^2$$

(και εδώ γνώση ακριβούς κατανομής).

Ιδιότητες χ_n^2

1. $\chi_n^2 \equiv G\left(\frac{n}{2}, \frac{1}{2}\right) \Rightarrow \mathbb{E}(\chi_n^2) \overset{a/\theta}{=} \frac{n}{2} 2 = n,$ το $\frac{1}{2}$ είναι παράμετρος "ρυθμού" (κλίμακας)

   $\mathbb{V}(\chi_n^2) \overset{a/\theta^2}{=} \frac{n}{2} 4 = 2n$

2. $\chi_{n_1}^2 + \chi_{n_2}^2 + \ldots + \chi_{n_\kappa}^2 \overset{\text{ανεξ.}}{=} \chi_{n_1+\ldots+n_\kappa}^2$

   $\left[\text{Από ιδιότητα Gamma } \mathrm{G}(a_1,\theta) + \mathrm{G}(a_2,\theta) + \ldots + \mathrm{G}(a_k,\theta) \overset{\text{ανεξ.}}{=} \mathrm{G}\left(\sum_{i=1}^k a_i, \theta\right)\right]$

3. $\chi_n^2 \cong \mathcal{N}(n, 2n)$, για μεγάλο n

   Αν $X \sim \chi_n^2$, τότε $X = \sum_{i=1}^n Z_i^2 \underset{\substack{\text{Κ.Ο.Θ.} \\ \text{(άθροισμα α.ι.τ.μ.)}}}{\overset{\text{μεγάλο } n}{\approx}} \mathcal{N}(n, 2n)$

Επιπλέον στο παράδειγμα

$$\theta = \left(\mu, \sigma^2\right)$$

$$\left(\widehat{\mu}_n, \widehat{\sigma_n^2}\right) = \left(\overline{X}_n, S_n^2\right) \sim ? \quad \text{(η από κοινού ?)}$$

Εδώ για τ.δ. απο κανονική η απάντηση είναι απλή $\longrightarrow$ είναι ανεξάρτητες τ.μ. $(\forall \theta)$.

Αρκεί να γνωρίζουμε τις περιθώριες κατανομές. Όμως αν δεν είναι από $\mathcal{N}(\mu, \sigma^2)$ γενικά είναι εξαρτημένες και θα ήθελε επιπλέον μελέτη. Κάποια θέματα **για διερεύνηση** (μέχρι τώρα):

- Υπάρχει γενικότερος τρόπος να βρίσκουμε α.ε.ε.δ.;
- Μπορούμε να έχουμε ασυμπτωτικές προσεγγίσεις κατανομών που να μας διευκολύνει τη μελέτη των εκτιμητριών για μεγάλο n; Π.χ. είδαμε χ_n^2 (σύγκλιση κατά κατανομή).
- Πώς αντιμετωπίζουμε από κοινού τις εκτιμήτριες όταν έχουμε πολυδιάστατη παράμετρο; $\longrightarrow$ **για ισχυρή συνέπεια + συνέπεια**.

Αν

$$X_n \xrightarrow{\text{a.s.}/p} X$$

και

$$Y_n \xrightarrow{\text{a.s.}/p} Y \Leftrightarrow (X_n, Y_n) \xrightarrow{\text{a.s.}/p} (X, Y)$$

Εφαρμογή

Αν

$$\widehat{\theta}_n = \left(\widehat{\theta}_{n,1}, \widehat{\theta}_{n,2}\right)$$

τότε

$$\widehat{\theta}_n \xrightarrow{\text{a.s.}/p} \underset{\substack{\| \\ (\theta_1,\theta_2)}}{\theta} \Leftrightarrow \widehat{\theta}_{n,1} \xrightarrow{\text{a.s.}/p} \theta_1, \quad \widehat{\theta}_{n,2} \xrightarrow{\text{a.s.}/p} \theta_2, \quad \forall \theta.$$

$\xrightarrow{d}$ Σύγκλιση κατά κατανομή

Έστω (X_n) μία ακολουθία τ.μ. και X άλλη μία τ.μ. τότε

$$X_n \xrightarrow{d} X \Leftrightarrow F_{X_n}(X) \xrightarrow[n\to\infty]{} F_X(X), \quad \forall X \text{ που είναι σημείο συνέχειας της } F_X.$$

Θεώρημα 0.0.5. Κεντρικό Οριακό Θεώρημα

Έστω (X_n) ακολουθία α.ι.τ.μ. με $0 < \mathbb{V}(X_1) < +\infty$. Τότε αν

$$S_n = \sum_{i=1}^{n} X_i \text{ και } \overline{X}_n = \frac{x_1 + \ldots + X_n}{n},$$

με $\mathbb{E}(X_1) = \mu$ και $\mathbb{V}(X_1) = \sigma^2$, τότε

$$\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \xrightarrow{d} \mathcal{N}(0,1) \quad (Z \sim \mathcal{N}(0,1))$$

ή

$$\sqrt{n}\frac{\overline{X}_n - \mu}{\sigma} \xrightarrow{d} \mathcal{N}(0,1) \quad (\mathcal{N}(0,0) \equiv 0)$$

$$\left[\frac{S_n - n\mu}{\sqrt{n}} \xrightarrow{d} \mathcal{N}(0,\sigma^2) \quad \text{ή} \quad \sqrt{n}(\overline{X}_n - \mu) \xrightarrow{d} \mathcal{N}(0,\sigma^2)\right]$$

Πρακτικά γράφουμε, για μεγάλο n:

$$S_n \approx \mathcal{N}(n\mu, n\sigma^2) \quad \text{ή} \quad \overline{X}_n \approx \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$$

Ειδικά

$$\chi_n^2 \cong \mathcal{N}(n, 2n) \quad \text{για μεγάλο } n.$$

$X = \sum_{i=1}^{n} Z_i^2 \quad (X = S_n \ \ \{\text{α.ι.τ.μ. } Z_i\})$

Π.χ.

$$\frac{\sum_{i=1}^{n} Z_i^2 - n}{\sqrt{2n}} \xrightarrow[\text{τυποποίηση}]{d} N(0,1)$$

Παρατήρηση 0.0.14. Προσοχή!

$\overline{X}_n \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$ όταν έχουμε τ.δ. $\mathcal{N}(\mu, \sigma^2)$. Γενικά $\overline{X}_n \nsim \mathcal{N}(\cdot)$ επομένως μπορούμε, για μεγάλο n, να το δεχτούμε ανεξάρτητα από ποια κατανομή προέρχονται τα X_i.

Π.χ. θυμηθείτε την $\mathcal{U}[0,\theta], \quad \theta > 0$.
Τότε

$$\overline{\theta}_n = 2\overline{X}_n = \overline{2X_n} = \overline{Y}_n$$

όπου $Y_n = 2X_n$, για κάθε $n \geq 1$.

$2X_i \sim \mathcal{U}[0, 2\theta], \quad 1 \leq i \leq n$.

$\left(\mu, \frac{\sigma^2}{n}\right)$ όπου $\mu = \mathbb{E}(\mathcal{U}[0,2\theta])$ και $\sigma^2 = \mathbb{V}(\mathcal{U}[0,2\theta])$.

$\overline{\theta}_n = \overline{Y}_n \nsim \mathcal{N}\left(\theta, \frac{\theta^2}{3n}\right)$, όμως για μεγάλο n, λόγω Κ.Ο.Θ.

$$\overline{\theta}_n \approx \mathcal{N}\left(\theta, \frac{\theta^2}{3n}\right)$$

$$\sqrt{n}\frac{\overline{\theta}_n - \theta}{\sqrt{\frac{\theta^2}{3}}} \xrightarrow[n\to+\infty]{d} \mathcal{N}(0,1) \quad \text{ή} \quad \sqrt{n}\left(\overline{\theta}_n - \theta\right) \xrightarrow[n\to+\infty]{d} \mathcal{N}\left(0, \frac{\theta^2}{3}\right).$$

Σχέση τρόπων σύγκλισης ακολουθιών τυχαίων μεταβλητών

$$\begin{matrix} \stackrel{\text{a.s.}}{\to} \\ \stackrel{L^2}{\to} \end{matrix} \quad \begin{matrix} \Rightarrow \\ \not\Leftarrow \end{matrix} \quad \stackrel{p}{\to} \quad \begin{matrix} \Rightarrow \\ \not\Leftarrow \end{matrix} \quad \stackrel{d}{\to}$$

Ειδική περίπτωση όταν $X_n \to c$ εκφυλισμένη τ.μ. δηλαδή $\mathbb{P}(X = c) = 1$ τότε

$$X_n \stackrel{p}{\to} c \Leftrightarrow X_n \stackrel{d}{\to} c$$

Θεώρημα 0.0.6. Θεώρημα Συνεχούς Απεικόνισης
Ήδη είχαμε δει ότι

$$X_n \xrightarrow{\text{a.s.}/p} X \to g(X_n) \xrightarrow{\text{a.s.}/p} g(X)$$

όταν g συνεχής.

(Αυτό γενικά ισχύει για g συνεχής σε σύνολο $D: \;\mathbb{P}(X \in D) = 1$.)
Επίσης ισχύει και για $\stackrel{d}{\to}$, άρα

$$X_n \stackrel{d}{\to} X \Rightarrow g(X_n) \stackrel{d}{\to} g(X).$$

Τελικά

$$X_n \xrightarrow{\text{a.s.}/p/d} X \Rightarrow g(X_n) \xrightarrow{\text{a.s.}/p/d} g(X).$$

Λήμμα 0.0.7. (Θεώρημα Slutsky)
Αν $X_n \xrightarrow{d} X$ και $Y_n \xrightarrow{p} c$, τότε

- $X_n + Y_n \xrightarrow{d} X + c$
- $X_n Y_n \xrightarrow{d} cX$
- $\dfrac{X_n}{Y_n} \xrightarrow[c\neq 0]{d} \dfrac{X}{c}$

Ενώ αν $X_n \xrightarrow{d} X$ και $Y_n \xrightarrow{d} Y \quad \overset{\text{γενικά}}{\not\Rightarrow}$
$X_n + Y_n \xrightarrow{d} X + Y, \quad X_n Y_n \xrightarrow{d} XY, \quad \dfrac{X_n}{Y_n} \xrightarrow{d} \dfrac{X}{Y}$
(ενώ στις $\xrightarrow{\text{a.s.}}$ και $\xrightarrow{p}$ ισχύει.)

Παράδειγμα 0.0.13. Όπως πριν.

$$\overline{\theta}_n = 2\overline{X}_n$$

$$\begin{aligned} \overline{X}_n &\approx \mathcal{N}\left(\frac{\theta}{2}, \frac{\theta^2}{12n}\right) \qquad \text{K.O.Θ.} \\ 2\overline{X}_n &\approx 2\mathcal{N}\left(\frac{\theta}{2}, \frac{\theta^2}{12n}\right) \\ &= \mathcal{N}\left(\theta, \frac{\theta^2}{3n}\right) \end{aligned}$$

Θεώρημα 0.0.8. Μέθοδος Δέλτα
Έστω (X_n) για την οποία υπάρχει $(r_n)\uparrow$ με $r_n(X_n - c) \xrightarrow{d} X$,

όπου $c \in \mathbb{R}$ σταθερά και X τ.μ.

Τότε αν πάρουμε κάποια g που είναι παραγωγίσιμη στο c, τότε

$$r_n\left(g(X_n) - g(c)\right) \xrightarrow{d} g'(c)X$$

Το θεώρημα ισχύει και για $g'(c) = 0$, αλλά έχει ενδιαφέρον για $g'(c) \neq 0$.

Εφαρμογή

Αν $\widehat{\theta}_n$ είναι εκτιμήτρια, $(r_n) \uparrow$ και $\theta \in \Theta$: $r_n(\widehat{\theta}_n - \theta) \xrightarrow{d} X, \quad \forall \theta \in \Theta$ και g παραγωγίσιμη στο θ, τότε

$$r_n\left(g(\widehat{\theta}_n) - g(\theta)\right) \xrightarrow{d} g'(\theta)X, \quad \forall \theta \in \Theta$$

(θα έχει ενδιαφέρον για $g'(\theta) \neq 0$).

Ειδικά για $r_n = \sqrt{n}, \quad X \sim \mathcal{N}(0, \sigma^2)$.

Αν

$$\sqrt{n}\left(\widehat{\theta}_n - \theta\right) \xrightarrow{d} \mathcal{N}(0, \sigma^2(\theta))$$

τότε

$$\sqrt{n}\left(g\left(\widehat{\theta}_n\right) - g(\theta)\right) \xrightarrow{d} \mathcal{N}\left(0, (g'(x))^2\sigma^2(\theta)\right).$$

Παράδειγμα 0.0.14. Έστω $X_1, X_2, \ldots, X_n$ τ.δ. $\underset{\text{ρυθμός}}{\underset{\downarrow}{\mathrm{Exp}(\theta)}}$

$$\overline{\theta}_n = \frac{1}{\overline{X}_n} = g(\overline{X}_n),$$

όπου $g(\theta) = \dfrac{1}{\theta}$, συνεχής στο $(0, +\infty)$ και παραγωγίσιμη.

Από το Κ.Ο.Θ.

$$\sqrt{n}\left(\overline{X}_n - \mathbb{E}(X_1)\right) \xrightarrow[n \to +\infty]{d} \mathcal{N}\left(0, \mathbb{V}_\theta(X_1)\right)$$

$$\sqrt{n}\left(\overline{X}_n - \frac{1}{\theta}\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{1}{\theta^2}\right) \quad \text{(μέθοδος Δέλτα)}$$

έχουμε

$$\sqrt{n}(\overline{\theta}_n - \theta) = \sqrt{n}\left(g(\overline{X}_n) - g\left(\frac{1}{\theta}\right)\right) \xrightarrow{d} \mathcal{N}\left(0, \left(g'\left(\frac{1}{\theta}\right)\right)^2 \frac{1}{\theta^2}\right)$$

$$g(\theta) = \frac{1}{\theta} \Rightarrow$$

$$g'(\theta) = -\frac{1}{\theta^2} \Rightarrow$$

$$\left(g'(\theta)\right)^2 = \frac{1}{\theta^4} \Rightarrow$$

$$\left(g'\left(\frac{1}{\theta}\right)\right)^2 = \theta^4 \Rightarrow$$

$$\sqrt{n}(\overline{\theta}_n - \theta) \xrightarrow{d} \mathcal{N}(0, \theta^2), \quad \forall \theta > 0.$$

Διάλεξη 9

Είδαμε με τη μέθοδο Δέλτα

$$r_n\left(\widehat{\theta}_n-\theta\right)\xrightarrow{d}X\sim K(\theta)\Rightarrow$$

$$r_n\left(g\left(\widehat{\theta}_n\right)-g(\theta)\right)\xrightarrow{d}g'(\theta)X,\quad \forall\theta\in\Theta.$$

Μας επιτρέπει να βρίσκουμε τις ασυμπτωτικές κατανομές εκτιμητριών, αλλά ταυτόχρονα θα μας επιτρέπει να ελέγξουμε ασυμπτωτικά το σφάλμα της εκτιμήτριας.

$$r_n\left(\widehat{\theta}_n-\theta\right)\xrightarrow{d}X\sim K(\theta),\quad \forall\theta.$$

$$\begin{aligned}
\widehat{\theta}_n &= \left(\widehat{\theta}_n-\theta\right)+\theta=\underset{\substack{\downarrow\\ 0}}{\frac{1}{r_n}}\underbrace{r_n\left(\widehat{\theta}_n-\theta\right)}_{\xrightarrow{d}X}+\underbrace{\theta}_{\text{σταθερό}}\\
&\Rightarrow \widehat{\theta}_n\xrightarrow{d}0X+\theta=\theta\\
&\Rightarrow \widehat{\theta}_n\xrightarrow{p}\theta,\quad \forall\theta\in\Theta\\
&\Rightarrow \text{η } \widehat{\theta}_n \text{ είναι συνεπής εκτιμήτρια του } \theta.
\end{aligned}$$

$$r_n\left(\widehat{\theta}_n-\theta\right)\xrightarrow{d}X\sim K(\theta)\Rightarrow\quad \text{η } \widehat{\theta}_n \text{ είναι συνεπής εκτιμήτρια του } \theta \quad\Leftrightarrow\quad \widehat{\theta}_n\xrightarrow{d}\theta,\quad \forall\theta\in\Theta.$$

$$X_n\xrightarrow{d}\underset{\substack{\downarrow\\ \text{σταθερό}}}{c}\Leftrightarrow X_n\xrightarrow{p}c$$

Ορισμός 0.0.9. Έστω T_n μια εκτιμήτρια του $g(\theta)$ σε κάποιο παραμετρικό μοντέλο με $\theta\in\Theta$. Αν $\exists(r_n)\to+\infty$ και μία οικογένεια κατανομών $K(\theta),\quad \theta\in\Theta$:

$$r_n(T_n-\theta)\xrightarrow{d}X\sim K(\theta),\quad \forall\theta\in\Theta,$$

τότε λέμε ότι η T_n έχει ασυμπτωτική κατανομή $K(\theta)$.

Παρατήρηση 0.0.15. Προσοχή!
Δείξαμε μέσω Slutsky ότι $r_n(T_n - \theta) \xrightarrow{d} K(\theta) \Rightarrow T_n \xrightarrow{d} \theta$
(όμως το θ εδώ εκφράζει τη συνέπεια της εκτιμήτριας και όχι την ασυμπτωτική της κατανομή).

Σχόλιο

Απαραίτητη προϋπόθεση της ύπαρξης ασυμπτωτικής κατανομής είναι η συνέπεια της αντίστοιχης εκτιμήτριας.

Προσοχή
Να μην εκφυλίζεται η ασυμπτωτική κατανομή $\left(g^{'}(\theta) \neq 0\right)$ όταν κοιτάμε συναρτήσεις $g\left(\widehat{\theta}_n\right)$.

Παράδειγμα 0.0.15. Ας υποθέσουμε ότι έχουμε ένα τ.δ. $X_1, \ldots, X_n$, όπου X_i έχουν άγνωστη κατανομή (δεν γνωρίζουμε την παραμετρική οικογένεια) με

$$\mathbb{V}_\theta(X_1) = \sigma^2(\theta) < +\infty, \text{ και έστω } \mathbb{E}_\theta(X_1) = \mu(\theta).$$

Μία α.ε. του $\mu = \mu(\theta)$ είναι ο δειγματικός μέσος $\overline{X}_n$ και μάλιστα ξέρουμε ότι είναι ισχυρά συνεπής, άρα και συνεπής (Ι.Ν.Μ.Α., Α.Ν.Μ.Α.).

$$\text{Δε γνωρίζουμε την κατανομή των } X_i \ \Rightarrow \text{Δε γνωρίζουμε την κατανομή του } \overline{X}_n$$

Τουλάχιστον μπορούμε να κάνουμε μία ασυμπτωτική προσέγγιση.
Εδώ είναι το ενδιαφέρον της ασυμπτωτικής κατανομής $\Rightarrow$ αποκτάμε πρόσβαση σε μία προσεγγιστική κατανομή του σφάλματος $\Rightarrow$ μπορούμε να ελέγξουμε το σφάλμα.

Πριν λοιπόν προχωρήσουμε στη μελέτη του προβλήματος ελέγχου του σφάλματος με ασυμπτωτικές μεθόδους, θα εστιάσουμε για εκτιμήτριες με γνωστή κατανομή.

Παράδειγμα 0.0.16. Έστω $X_1, \ldots, X_n$ τ.δ. από $\mathcal{N}(\mu, \sigma^2)$, όπου μ γνωστό και σ^2 άγνωστο. Εδώ ο δειγματικός μέσος $\overline{X}_n$ είναι καλή εκτιμήτρια του μ και $\overline{X}_n \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$.

Ο δειγματικός μέσος $\overline{X}_n$ είναι συνεπής εκτιμήτρια του μ.

$$\overline{X}_n \xrightarrow{p} \mu, \quad \forall \mu \in \mathbb{R}.$$

$$\forall \varepsilon > 0, \quad \underbrace{\mathbb{P}\left(\left|\overline{X}_n - \mu > \varepsilon\right|\right)}_{|\text{σφάλμα} > \varepsilon|} \xrightarrow[n\to\infty]{} 0.$$

Για να υπολογίσουμε αυτήν την πιθανότητα πρέπει να έχουμε γνώση της κατανομής του $\overline{X}_n$ (και του σφάλματος)

$$\overline{X}_n \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right) \Rightarrow \frac{\overline{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} = \sqrt{n}\frac{\overline{X}_n - \mu}{\sigma} \sim \mathcal{N}(0,1).$$

$$\begin{aligned}
\mathbb{P}_\mu\left(\left|\overline{X}_n - \mu\right| > \varepsilon\right) &= \mathbb{P}_\mu\left(\sqrt{n}\left|\frac{\overline{X}_n - \mu}{\sigma}\right| > \frac{\sqrt{n}\varepsilon}{\sigma}\right) = \mathbb{P}_\mu\left(|Z| > \frac{\sqrt{n}\varepsilon}{\sigma}\right) \\
&= \mathbb{P}\left(Z > \frac{\sqrt{n}\varepsilon}{\sigma}\right) + \mathbb{P}\left(Z < -\frac{\sqrt{n}\varepsilon}{\sigma}\right) = 1 - \Phi\left(\frac{\sqrt{n}\varepsilon}{\sigma}\right) + \Phi\left(-\frac{\sqrt{n}\varepsilon}{\sigma}\right) \\
&= 2 - 2\Phi\left(\frac{\sqrt{n}\varepsilon}{\sigma}\right) = 2\left(1 - \Phi\left(\frac{\sqrt{n}\varepsilon}{\sigma}\right)\right)
\end{aligned}$$

αφού $\Phi(-x) = 1 - \Phi(x)$ και $Z \sim \mathcal{N}(0,1)$.

Επομένως μπορούμε να το ελέγξουμε, εφ' όσον είναι γνωστό.

Το κλειδί σε αυτά είναι να φτιάχνουμε μία κατανομή που είναι ανεξάρτητη της παραμέτρου θ.

π.χ.

$$\underbrace{\frac{\overline{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}}}_{\text{τυχαία συνάρτηση του } \mu} \sim \mathcal{N}(0,1) \longrightarrow \text{ανεξάρτητη του } \mu$$

Ορολογία

- Θα λέμε ότι γίνεται ε-**σφάλμα**, αν

$$\omega \in \left\{\left|\widehat{\theta}_n - \theta\right| > \varepsilon\right\} \Leftrightarrow \left|\widehat{\theta}_n(\omega) - \theta\right| > \varepsilon$$

  δηλαδή αν το **σφάλμα ξεπέρασε** το ε στο θ, όπου θ (η γενικά άγνωστη τιμή) της παραμέτρου.

- Αν

$$\mathbb{P}\left(\left|\widehat{\theta}_n - \theta\right| > \varepsilon\right) \leq \delta,$$

  τότε θα λέμε ότι ένα ε-σφάλμα είναι/γίνεται δ-**σπάνιο** (δ μικρό).

- Θα λέμε ότι έχουμε/πετυχαίνουμε ε-**ακρίβεια**, αν

$$\omega \in \left\{\left|\widehat{\theta}_n - \theta\right| \leq \varepsilon\right\}$$

  ή αν

$$\left|\widehat{\theta}_n(\omega) - \theta\right| \leq \varepsilon.$$

  Αν έχουμε

$$\mathbb{P}\left(\left|\widehat{\theta}_n - \theta\right| \leq \varepsilon\right) \geq 1 - \delta,$$

  τότε πετυχαίνουμε/έχουμε ε-ακρίβεια, $(1-\delta)$-συχνά.

Γενικά η $\mathbb{P}\left(\left|\widehat{\theta}_n - \theta\right| \leq \varepsilon\right)$ ή $\mathbb{P}\left(\left|\widehat{\theta}_n - \theta\right| > \varepsilon\right)$ μπορεί να εξαρτάται από το θ.

Γενικά λοιπόν, πρέπει να μπορούμε να εξασφαλίσουμε φράγματα που είναι καθολικά, δηλαδή ανεξάρτητα της παραμέτρου θ. Θα μπορούσε να εκφραστεί ως εξής:

$$\sup_{\theta \in \Theta} \mathbb{P}_\theta\left(\left|\widehat{\theta}_n - \theta\right| > \varepsilon\right) \leq \delta \quad \text{(Θέλω το } \varepsilon\text{-σφάλμα να είναι } \delta\text{-σπάνιο για όλα τα } \theta \in \Theta.)$$

Αντίστοιχα

$$\inf_{\theta \in \Theta} \mathbb{P}_\theta\left(\left|\widehat{\theta}_n - \theta\right| \leq \varepsilon\right) \geq 1 - \delta \quad \text{(Θέλω } \varepsilon\text{-ακρίβεια } (1-\delta)\text{-συχνά για όλα τα } \theta \in \Theta.)$$

Συνέπεια Εκτιμήτριας

$$\forall \theta \in \Theta \quad \forall \varepsilon > 0 \quad \mathbb{P}_\theta\left(\left|\widehat{\theta}_n - \theta\right| > \varepsilon\right) \xrightarrow[n \to \infty]{} 0$$

$$\forall \theta \in \Theta \quad \forall \varepsilon > 0 \quad \forall \delta > 0 \quad \mathbb{P}_\theta\left(\left|\widehat{\theta}_n - \theta\right| > \varepsilon\right) \leq \delta \text{ τελικά } \forall n$$

για οποιαδήποτε παράμετρο

θα γίνει τελικά οσοδήποτε δ-σπάνιο

ένα οποιοδήποτε ε-σφάλμα

$$\boxed{\mathbb{P}_\theta\left(\left|\widehat{\theta}_n - \theta\right| \leq \varepsilon\right) \to 1}$$

$$\forall \theta \in \Theta \quad \forall \varepsilon > 0 \quad \forall \delta > 0 \quad \mathbb{P}_\theta\left(\left|\widehat{\theta}_n - \theta\right| \leq \varepsilon\right) \geq 1 - \delta$$

για οποιαδήποτε παράμετρο

μία οποιαδήποτε ε-ακρίβεια

πετυχαίνεται τελικά οσοδήποτε $(1-\delta)$-συχνά

Δεν έχουμε όμως εξασφάλιση ότι αυτό πετυχαίνεται ομοιόμορφα ως προς θ.

Η τεχνική που χρησιμοποιούμε είναι να διώχνουμε την εξάρτηση στο θ, είτε μέσω sup και inf, όπως είπαμε πριν, είτε στην καλύτερη περίπτωση βρίσκοντας κατανομές ελέυθερες παραμέτρων μέσω κατάλληλων μετασχηματισμών (ποσότητες οδηγούς) έχοντας εκτιμήσει οποιαδήποτε άλλη παράμετρο εμφανίζεται και είναι γενικά άγνωστη.

Στο αρχικό πρόβλημα δεν γνωρίζαμε την κατανομή του $\overline{X}_n$. Πώς μπορούμε να κάνουμε μία προσέγγιση;

Ένας δημοφιλής τρόπος είναι μέσω ασυμπτωτικών κατανομών ($\Rightarrow$ ασυμπτωτικές προσεγγίσεις). Το $\overline{X}_n$ εκτιμά το $\mu(\theta)$.

Θεωρούμε αρχικά ότι το $\sigma^2(\theta)$ είναι γνωστό.

$$\mathbb{P}_\theta\left(\left|\overline{X}_n - \mu(\theta)\right| > \varepsilon\right) = ?$$

ασυμπτωτική κατανομή

Το Κ.Ο.Θ. αν $\sigma^2(\theta) < +\infty$, τότε

$$\overline{X}_n \overset{\text{μεγάλο } n}{\approx} \mathcal{N}\left(\mu(\theta), \frac{\sigma^2(\theta)}{n}\right)$$

$$\frac{\overline{X}_n - \mu(\theta)}{\frac{\sigma(\theta)}{\sqrt{n}}} = \sqrt{n}\frac{\overline{X}_n - \mu(\theta)}{\sigma(\theta)} \approx \mathcal{N}(0,1).$$

Άρα

$$\mathbb{P}_\theta\left(\left|\overline{X}_n - \mu(\theta)\right| > \varepsilon\right) = \mathbb{P}_\theta\left(\left|\sqrt{n}\frac{\overline{X}_n - \mu(\theta)}{\sigma(\theta)}\right| > \frac{\sqrt{n}\varepsilon}{\sigma(\theta)}\right)$$

$$\overset{\text{για μεγάλο } n}{\approx} 2\left(1 - \Phi\left(\frac{\sqrt{n}\varepsilon}{\sigma(\theta)}\right)\right)$$

$\sigma(\theta) \to$ σταθερό

Υπολογίζουμε απευθείας θέτοντας

$$2\left(1-\Phi\left(\frac{\sqrt{n}\varepsilon}{\sigma}\right)\right)=\delta$$

και αν μας ζητείται να βρεθεί το μέγεθος του δείγματος που εξασφαλίζει ε-σφάλμα, δ-σπάνια (ε, δ γνωστά), τότε λύνουμε ως προς n.

- Για σταθερό γνωστό μέγεθος δείγματος n και επιθυμητό ε-σφάλμα, πόσο είναι το δ;

- Για σταθερό n και γνωστό δ, πόσο είναι το ε;

Διάλεξη 10

Στην αρχή της διάλεξης έγινε ο σχολιασμός των απαντήσεων στο Quiz 1

Άσκηση

Παρακολουθούμε την εξέλιξη μιας μετοχής Y_t στο χρόνο (ας πούμε κάθε μικρή χρονική περίοδο, π.χ. κάθε λεπτό, ανά ώρα, ανά μέρα) και ενδιαφερόμαστε να μοντελοποιήσουμε τις log-αποδόσεις (log-returns) δηλαδή τις

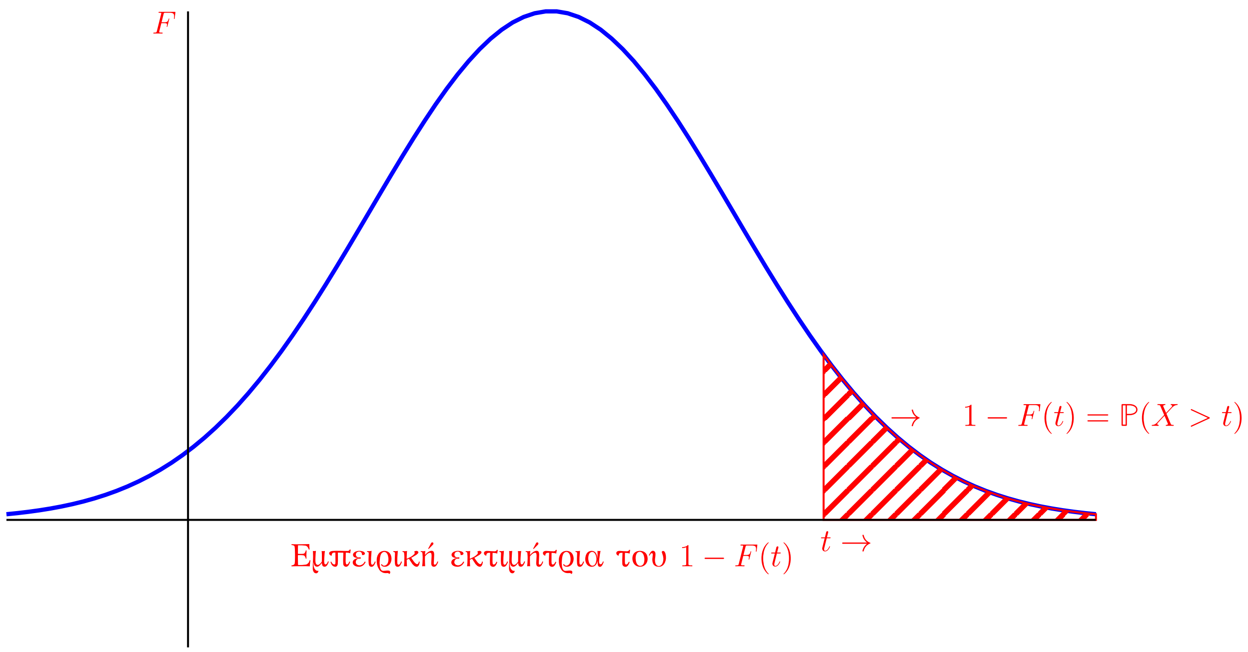
$$X_t = \log \frac{Y_{t+1}}{Y_t} \in \mathbb{R}, \quad t = 1, 2, \dots$$

- Μπορούμε να προσομοιώσουμε τιμές από τις (X_t) θεωρώντας ότι $Y_t \sim K(\theta)$. Επιλέγουμε εμείς την K, ενώ θ είναι οι παράμετροι υπό εκτίμηση.
- Είναι ενδιαφέρον να δείτε πως εξελίσσεται η συνάρτηση $(X_1, X_2, \dots, X_n)$

  $$t \longmapsto \frac{\#\{X_i > t\}}{n}$$

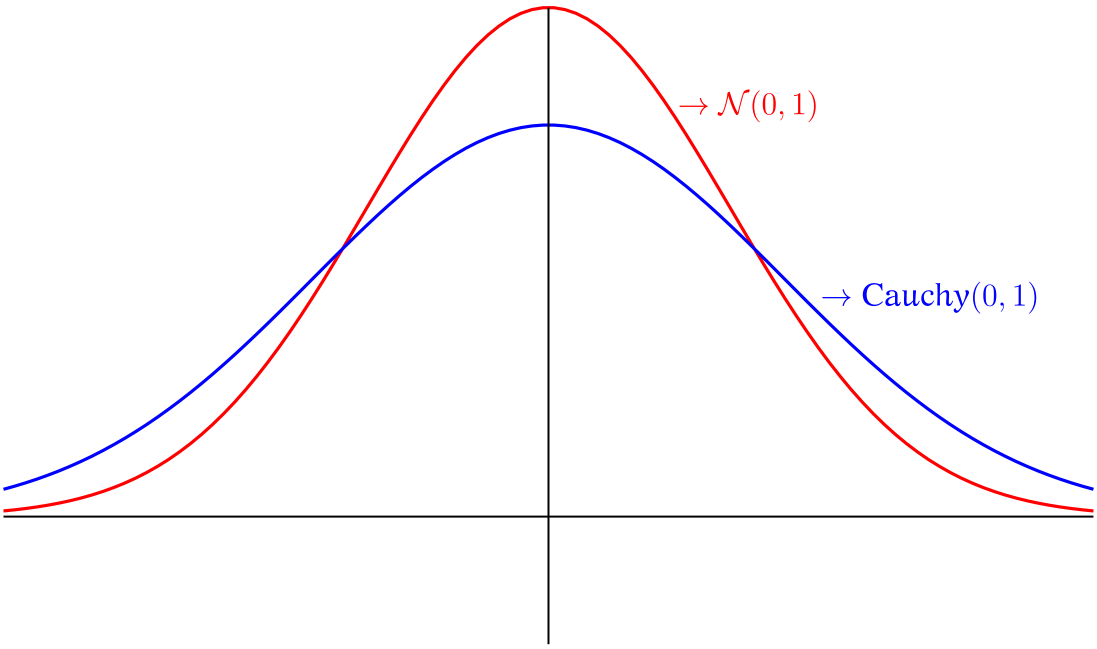
  καθώς το t μεγαλώνει $\equiv$ ποσοστό των παρατηρήσεων που είναι $> t$

  Εκτίμηση της $\overline{F}_{K_{(\theta)}}(t) = 1 - F_{K_{(\theta)}}(t)$



Σημείωση

Η διαφορά της Cauchy $(0,1)$ από την $\mathcal{N}(0,1)$ είναι ότι έχει πιο βαριές ουρές και επομένως μοντελοποιεί δεδομένα που προκύπτουν από κατανομές που δίνουν μεγαλύτερο βάρος σε μεγάλες τιμές.



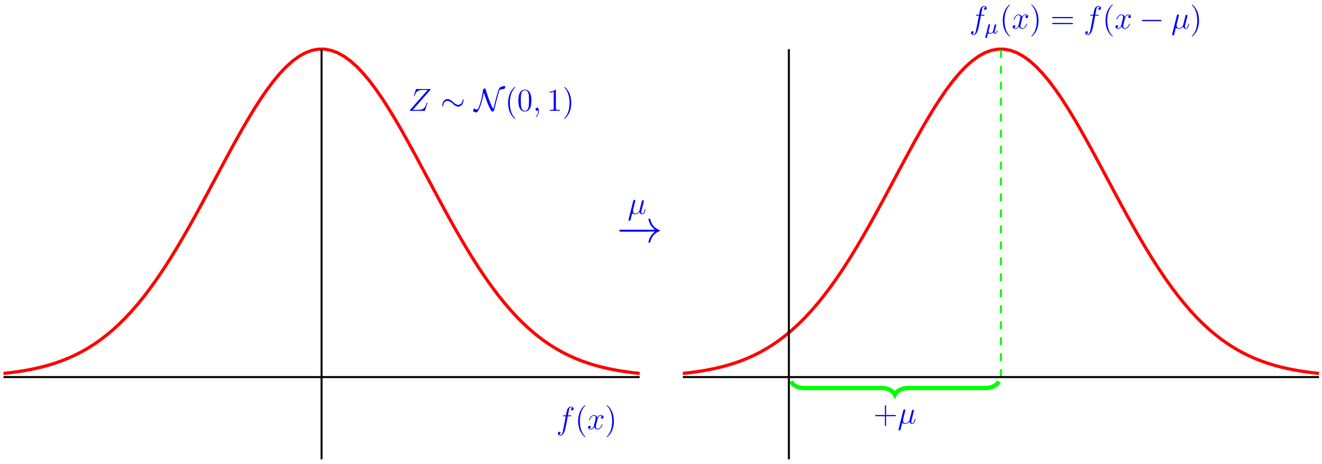
Οικογένεια θέσης - κλίμακας

Αν $Z \sim \mathcal{N}(0,1)$

$$X = \mu + \sigma \cdot Z, \quad \mu \in \mathbb{R},\ \sigma > 0$$

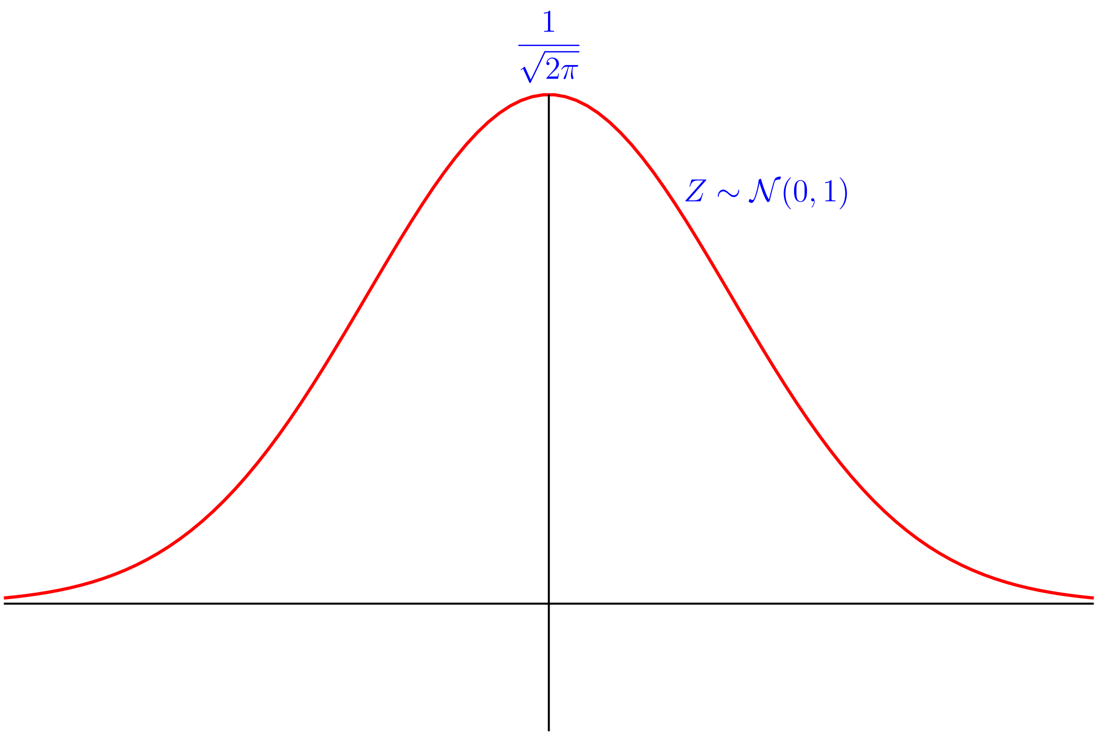
όπου μ: παράμετρος θέσης (μετατόπιση) και σ: παράμετρος κλίμακας (τέντωμα/ άπλωμα)

Οικογένεια θέσης



Οικογένεια κλίμακας

$$f = \phi \qquad \text{σ.π.π. της τυπικής κατανομής}$$



$$X = \frac{1}{2} \cdot Z$$

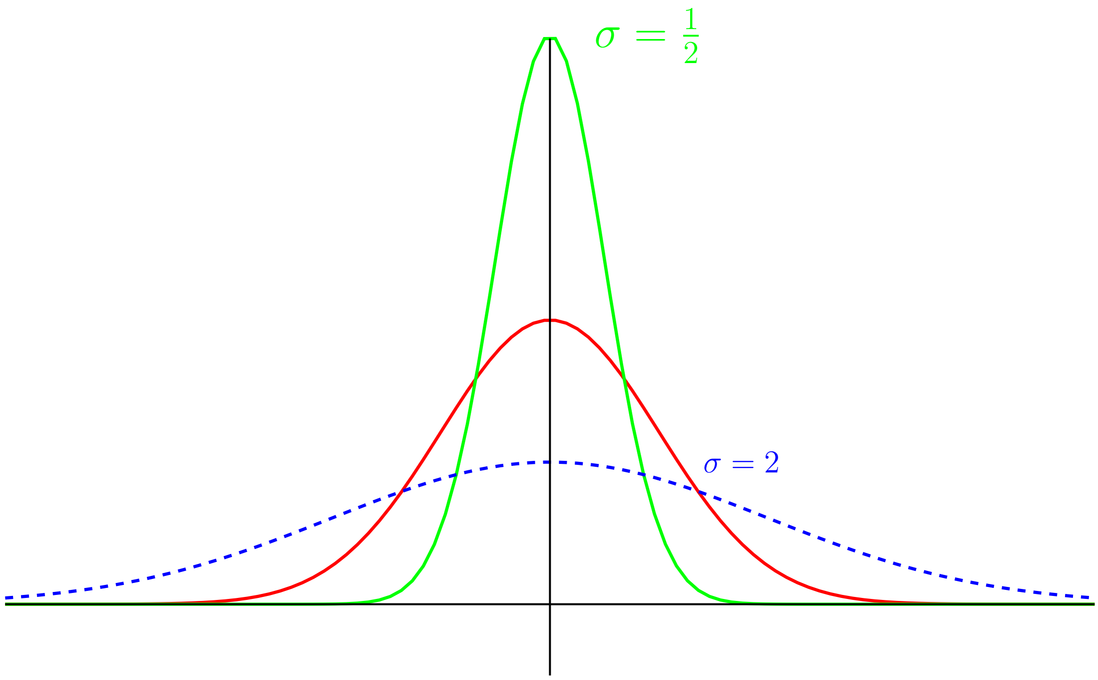
Είναι της μορφής $f_{\mu,\sigma}(x)$ και

$$f_{0,\frac{1}{2}}(x) = f_Z(2x) \cdot 2 = 2 \cdot \varphi(2x)$$

όπου $\varphi(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}$ και για $x = 0,\ \varphi(x) = \frac{1}{\sqrt{2\pi}}$.
Τελικά

$$f_{\mu,\sigma}(x) = \frac{1}{\sigma} \cdot f\left(\frac{X-\mu}{\sigma}\right)$$

Γενικά $Y = \mu + \sigma \cdot X$.


$\sigma = \frac{1}{2}$
$\sigma = 2$

Διάλεξη 11

$\rightarrow$ Λύση άσκησης 3 – διάλεξης 8.

(i) Έστω X = #αφίξεων ανά 5λεπτο. Θεωρούμε ότι διατέθουμε ένα τυχαίο δείγμα $X_1, X_2, \ldots, X_n$ από την κατανομή της X και θέτουμε $\pi = \mathbb{P}(X = 0)$.

Η παράμετρος π ερμηνεύεται ως η θεωρητική συχνότητα των 5-λέπτων που το σύστημα θα παραμένει εξ ολοκλήρου άδειο. Θα μπορούσε να είναι από μόνο του ένα χαρακτηριστικό ενδιαφέροντος του συστήματος. Στο τελευταίο ερώτημα δίνεται μία εξήγηση που το διαφοροποιεί από το θεωρητικό ποσοστό του χρόνου που το σύστημα παραμένει άδειο. Συνεχίζουμε με την υπόθεση ότι το π είναι η παράμετρος που θέλουμε να εκτιμήσουμε. Η εμπειρική εκτίμηση του $\pi = \mathbb{P}(X = 0)$, οδηγεί στην εκτιμήτρια

$$\widetilde{\pi}_n = \mathbb{P}(X^* = 0), \text{ με } X^* \sim \text{discr.Unif}\left(\{X_1, X_2, \ldots, X_n\}\right).$$

Επομένως,

$$\mathbb{P}(X^* = x) = \frac{\#\{X_i : X_i = x\}}{n} \Rightarrow \mathbb{P}(X^* = 0) = \frac{\sum_{i=1}^{n} \mathbb{1}_{\{X_i=0\}}}{n} = \widetilde{\pi}_n.$$

Έχουμε $\widetilde{\pi}_n = \overline{Y}_n$, όπου $Y_i = \mathbb{1}_{\{X_i=0\}}, i \geq 1$, α.ι.τ.μ. $\sim$ Be(π), διότι $\pi = \mathbb{P}(X = 0)$.

* **αμεροληψία**

$$\mathbb{E}\left(\widetilde{\pi}_n\right) = \mathbb{E}\left(\overline{Y}_n\right) = \mathbb{E}(\text{Be}(\pi)) = \pi \Rightarrow \widetilde{\pi}_n \text{ α.ε. του } \pi.$$

* **ΜΤΣ**

$$\text{ΜΤΣ}_{\widetilde{\pi}_n} = \mathbb{E}\left(\widetilde{\pi}_n - \pi\right)^2 \overset{\text{α.ε.}}{=} \mathbb{V}\left(\widetilde{\pi}_n\right) = \frac{\mathbb{V}(\text{Be}(\pi))}{n} = \frac{\pi(1-\pi)}{n},$$

διότι $\widetilde{\pi}_n = \overline{Y}_n$ και ξέρουμε ότι $\mathbb{V}\left(\overline{Y}_n\right) = \dfrac{\mathbb{V}(Y_1)}{n}$.

* **συνέπεια**
Από την παραπάνω σχέση του ΜΤΣ έχουμε ότι

$$\text{ΜΤΣ}_{\widetilde{\pi}_n} \underset{n\to\infty}{\longrightarrow} 0 \quad \Longrightarrow \quad \widetilde{\pi}_n \text{ είναι } L^2\text{-συνεπής}$$

Από την L^2-συνέπεια συμπεραίνουμε τη συνέπεια της εκτιμήτριας.

* **ισχυρή συνέπεια**
  Έχουμε
  $$\widetilde{\pi}_n = \overline{Y}_n \xrightarrow{\text{a.s.}} \mathbb{E}(Y_1) = \pi, \ \ (\forall \text{ δυνατό } p).$$
  Τελικά η $\widetilde{\pi}_n$ είναι ισχυρά συνεπής.

  Υπενθύμιση
  $$Y_1 = \mathbb{1}_{\{X_1=0\}} \in \{0,1\} \Rightarrow Y_1 \sim \mathrm{Be}(\pi), \text{ διότι } \mathbb{P}(Y_1=1) = \mathbb{P}(X_1=0) = \pi$$

(ii) $\sqrt{n}\,(\widetilde{\pi}_n - \pi)$ συγκλίνει; + οριακή κατανομή

[Διάλεξη 9 $\rightarrow$ αυτό καθορίζει την ασυμπτωτική κατανομή της εκτιμήτριας]

Η $(Y_n)_{n\geq 1}$ είναι ακολουθία α.ι.τ.μ. (αφού Y_n είναι συνάρτηση της X_n που είναι με τη σειρά της ακολουθία α.ι.τ.μ.), και επιπλέον $\mathbb{V}(Y_1) < +\infty$. Άρα ισχύουν οι προϋποθέσεις του Κ.Ο.Θ. και έχουμε ότι
$$\sqrt{n}\,(\widetilde{\pi}_n - \pi) = \sqrt{n}\left(\overline{Y}_n - \mathbb{E}(Y_1)\right) \xrightarrow[n\to\infty]{\text{d}} \mathcal{N}(0, \mathbb{V}(Y_1)) = \mathcal{N}\left(0, \pi(1-\pi)\right).$$
Αυτή είναι η οριακή κατανομή για κάθε δυνατό p.

(iii) Υποθέτουμε τώρα ότι $X_1, X_2, \ldots, X_n$ είναι ένα τυχαίο δείγμα από $\mathcal{P}(\lambda), \lambda > 0$. Έχουμε
$$\lambda = \mathbb{E}_\lambda(X_1), \text{ άρα } \widehat{\lambda}_n = \overline{X}_n \text{ (εκτιμήτρια ροπών)}$$
και
$$\pi = \mathbb{P}(X=0) = e^{-\lambda} = g(\lambda)$$
που προκύπτει αντικαθιστώντας με $x=0$ στη συνάρτηση πιθανότητας της Poisson: $\mathbb{P}(X=x) = e^{-\lambda}\dfrac{\lambda^x}{x!}$. Θέτουμε λοιπόν
$$\widehat{\pi}_n = g\left(\widehat{\lambda}_n\right) = g\left(\overline{X}_n\right) = e^{-\overline{X}_n},$$
που είναι η plug-in εκτιμήτρια του $\pi = g(\lambda)$.

(iv) (Αμεροληψία;)

Απαντάμε σε αυτό το ερώτημα πρώτα ποιοτικά. Έχουμε
$(e^{-x})' = -e^{-x}$, $(e^{-x})'' = e^{-x} > 0$ και άρα η $g(\lambda) = e^{-\lambda}$ είναι γνήσια κυρτή στο $(0,+\infty)$. Από την ανισότητα Jensen έχουμε
$$\mathbb{E}_\lambda\left(\widehat{\pi}_n\right) = \mathbb{E}_\lambda\left(g\left(\overline{X}_n\right)\right) \overset{\text{Jensen}}{\geq} g\left(\mathbb{E}\left(\overline{X}_n\right)\right) = e^{-\lambda} = \pi$$
και μάλιστα έχουμε ότι $\mathbb{E}_\lambda\left(\widehat{\pi}_n\right) > \pi$, διότι $\overline{X}_n$ **δεν είναι εκφυλισμένη τ.μ.** και η g δεν είναι **αφφινική** $(a\lambda + b)$. Συμπεραίνουμε ότι η $\widehat{\pi}_n$ είναι **θετικά μεροληπτική**. Μπορεί να υπολογιστεί η μεροληψία; Χρειαζόμαστε τον υπολογισμό της μέσης τιμής
$$\mathbb{E}_\lambda\left(e^{-\overline{X}_n}\right) = \ ?$$
Παρατηρούμε ότι $\overline{X}_n = \dfrac{1}{n}S_n$, όπου $S_n = \sum\limits_{i=1}^{n} X_i$ και $X_i \sim \mathcal{P}(\lambda) \Rightarrow S_n \sim \mathcal{P}(n\lambda)$ (άθροισμα n ανεξάρτητων $\mathcal{P}(\lambda)$). Ο παραπάνω υπολογισμός συσχετίζεται με τη γνώση της ροπογεννήτριας της Poisson. [αν δοθεί σε άσκηση η ροπογεννήτρια τότε μπορούμε να τη χρησιμοποιήσουμε]
$$\mathbb{E}_\lambda\left(e^{-\frac{1}{n}S_n}\right) = M_{S_n}\left(-\frac{1}{n}\right), S_n \sim \mathcal{P}(n\lambda)$$

Αν $M_X(t)$ είναι η ροπογεννήτρια κάποιας τ.μ. X, τότε $M_X(t) = \mathbb{E}\left(e^{tX}\right)$. Στην περίπτωση της Poisson, μπορεί να δειχθεί ότι αν $X \sim \mathcal{P}(\lambda)$, τότε

$$M_X(t) = e^{\lambda(e^t-1)}, \quad \forall t \in \mathbb{R}.$$

και

$$M_{S_n}\left(-\frac{1}{n}\right) = e^{n\lambda\left(e^{-\frac{1}{n}}-1\right)} = e^{-n\lambda\left(1-e^{-\frac{1}{n}}\right)} = e^{-\lambda\frac{1-e^{-1/n}}{1/n}}, \quad \left(\frac{1-e^{-x}}{x} \xrightarrow{x\to 0^+} 1\right)$$

Προσοχή γενικά και στο πεδίο σύγκλισης μιας ροπογεννήτριας, διότι αυτό δεν είναι πάντα το $\mathbb{R}$. Είναι φανερό λοιπόν ότι

$$\mathbb{E}_\lambda\left(\widehat{\pi}_n\right) = \mathbb{E}_\lambda\left(e^{-\overline{X}_n}\right) = \mathbb{E}\left(e^{-\frac{1}{n}S_n}\right) = M_{S_n}\left(-\frac{1}{n}\right).$$

Εφόσον $S_n \sim \mathcal{P}(n\lambda)$, αντικαθιστώντας στη ροπογεννήτριά της, συμπεραίνουμε ότι

$$\mathbb{E}_\lambda\left(\widehat{\pi}_n\right) = e^{n\lambda\left(e^{-\frac{1}{n}}-1\right)} = e^{-n\lambda\left(1-e^{-\frac{1}{n}}\right)} = e^{-\lambda\frac{1-e^{-1/n}}{1/n}},$$

και άρα η μεροληψία της προκύπτει αφαιρώντας το λ. Συνεχίστε τώρα την άσκηση υπολογίζοντας ακριβή τιμή για το ΜΤΣ.

Για να δειχθεί η ισχυρή συνέπεια παρατηρούμε πρώτα από τον Ι.Ν.Μ.Α. ότι

$$\overline{X}_n \xrightarrow{\text{a.s.}} \lambda, \quad \forall \lambda > 0.$$

Όμως

$$\widehat{\pi}_n = e^{-\overline{X}_n} = g\left(\overline{X}_n\right) \xrightarrow{\text{a.s.}} g(\lambda) = e^{-\lambda}, \quad \forall \lambda > 0,$$

από το Θ.Σ.Α., αφού η $g(\lambda) = e^{-\lambda}$ είναι συνεχής.

(v) Από το Κ.Ο.Θ. (ισχύουν οι προϋποθέσεις) έχουμε ότι

$$\sqrt{n}\left(\overline{X}_n - \lambda\right) = \sqrt{n}\left(\overline{X}_n - \mathbb{E}_\lambda(X_1)\right) \xrightarrow{d} \mathcal{N}(0, \mathbb{V}_\lambda(X_1)) = \mathcal{N}(0, \lambda).$$

Άρα επειδή η g είναι και παραγωγίσιμη με $g'(\lambda) = -e^{-\lambda}$ για κάθε $\lambda > 0$, με μία εφαρμογή της μεθόδου δέλτα, έχουμε ότι

$$\sqrt{n}\left(\widehat{\pi}_n - \pi\right) = \sqrt{n}\left(e^{-\overline{X}_n} - e^{-\lambda}\right) = \sqrt{n}\left(g\left(\overline{X}_n\right) - g(\lambda)\right) \xrightarrow[n\to\infty]{d} \mathcal{N}\left(0, (g'(\lambda))\cdot\lambda\right)$$

και

$$g'(\lambda) = -e - \lambda \Rightarrow \left(g'(\lambda)\right)^2 = e^{-2\lambda}$$

επομένως

$$\sqrt{n}\left(\widehat{\pi}_n - \pi\right) \xrightarrow[n\to\infty]{d} \mathcal{N}\left(0, \lambda\cdot e^{-2\lambda}\right), \ \forall \lambda$$

Λόγω της σχέσης $\pi = e^{-\lambda}$, αντικαθιστώντας όπου $\lambda = -\log\pi$, παίρνουμε τελικά ότι

$$\sqrt{n}\left(\widehat{\pi}_n - \pi\right) \xrightarrow[n\to\infty]{d} \mathcal{N}\left(0, \pi^2\log\left(\frac{1}{\pi}\right)\right).$$

Συνεχίστε την άσκηση συγκρίνοντας τις ασυμπτωτικές διασπορές της εμπειρικής μεθόδου και αυτής για την οποία υποθέσαμε ότι έχουμε τυχαίο δείγμα από Poisson.

Μέθοδος Μέγιστης Πιθανοφάνειας

MLE: Maximum Likelihood Estimation (Estimator). Είναι η πιο δημοφιλής μέθοδος εκτίμησης (Fisher 1912-1922). Αναπτύχθηκε ως βελτίωση της μεθόδου των ροπών (Pearson) και ελαχίστων τετραγώνων (Gauss - Legendre). Απαλλάσσεται από πολλά προβλήματα. Στηρίζεται στην **αρχή της πιθανοφάνειας**

Δεδομένα $x = (x_1, x_2, \ldots, x_n) \longrightarrow$ παραμετρικό μοντέλο με $\theta \in \Theta$ και $f(x_1, x_2, \ldots, x_n\ ;\ \theta)$ η από κοινού σ.π.π. ή σ.π. των δεδομένων για κάποιο θ

Αναρωτιόμαστε ποια είναι η τιμή του $\theta \in \Theta$ που κάνει πιο πιθανή ή πιο πυκνή την εμφάνιση των συγκεκριμένων $x_1, x_2, \ldots, x_n$ κάτω από το μοντέλο που έχουμε θεωρήσει.

$$x = (x_1, x_2, \ldots, x_n) \longmapsto L_x(\theta) = f(x_1, x_2, \ldots, x_n\ ;\ \theta)$$

$\forall x$, φτιάχνουμε τη συνάρτηση $L_x(\theta)$ (τη βλέπουμε ως συνάρτηση του θ)

Η τιμή εκείνη $\widehat{\theta}(x) = \max\limits_{\theta \in \Theta} = L_x(\theta)$ είναι η εκτίμηση που παίρνουμε για το θ και η εκτιμήτρια λέγεται εκτίμηση μέγιστης πιθανοφάνειας.

Η εκτιμήτρια φτιάχνεται $* \longmapsto X$ και $\widehat{\theta}(X)$ την εκτιμήτρια μέγιστης πιθανοφάνειας

Ορισμός 0.0.10. Συνάρτηση πιθανοφάνειας του $x = (x_1, x_2, \ldots, x_n)$ είναι η συνάρτηση

$$L(\theta) = L_x(\theta) = f(x_1, x_2, \ldots, x_n\ ;\ \theta)$$

Για κάθε x (δυνατή n-άδα) $\longmapsto$ συνάρτηση του θ

Αν $\widehat{\theta}(x) = \arg\max\limits_{\theta \in \Theta} L_x(\theta)$, τότε η εκτιμήτρια που προκύπτει

$$\widehat{\theta}(X) = \arg\max_{\theta \in \Theta} L_X(\theta)$$

λέγεται **εκτιμήτρια μέγιστης πιθανοφάνειας**.

Παρατήρηση 0.0.16.

1. Αν $X_1, \ldots, X_n$ τ.δ. $(X_i \sim f_\theta)$ τότε

$$L_x(\theta) = \prod_{i=1}^{n} f_\theta(x_i) = \prod_{i=1}^{n} f(x_i\ ;\ \theta) = \prod_{i=1}^{n} f(x_i \mid \theta)$$

2. $L\left(\widehat{\theta}\right) = \max_{\theta \in \Theta} L(\theta) \Leftrightarrow \log L\left(\widehat{\theta}\right) = \max_{\theta \in \Theta} \log L(\theta)$,
   $\log \nearrow$ (ισοδυναμία ως προς μεγιστοποίηση)

   Η συνάρτηση $\ell(\theta) = \log L(\theta)$ λέγεται συνάρτηση λογαριθμοπιθανοφάνειας ή λογαριθμοπιθανοφάνεια (που αντιστοιχεί στο x).

Παράδειγμα 0.0.17. Όταν υπάρχει και είναι μοναδική η ε.μ.π. Π.χ. Έστω $X_1, \ldots, X_n$ τ.δ. από $\operatorname{Exp}(\theta)$, τότε ακολουθούμε τα παρακάτω βήματα

1. Υπολογίζουμε την πιθανοφάνεια $(x_1, \ldots, x_n > 0)$

$$L(\theta) = \prod_{i=1}^{n} f_\theta(x_i) = \prod_{i=1}^{n} \theta e^{-\theta x_i} = \theta^n e^{-\theta \sum\limits_{i=1}^{n} x_i}$$

2. Υπολογίζουμε τη λογαριθμοπιθανοφάνεια

$$\ell(\theta) = \log L(\theta) = \sum_{i=1}^{n} \log\left(\theta e^{-\theta x_i}\right) = n \log\theta - \theta \sum_{i=1}^{n} x_i$$

3. Αν η $\ell(\theta)$ είναι παραγωγίσιμη βρίσκουμε τα στάσιμα σημεία της.

$$\ell'(\theta) = \frac{n}{\theta} - \sum_{i=1}^{n} x_i$$

Λύνουμε την **εξίσωση πιθανοφάνειας**

$$\ell'(\theta) = 0 \Leftrightarrow \frac{n}{\theta} = \sum_{i=1}^{n} x_i \Leftrightarrow \theta = \frac{n}{\sum\limits_{i=1}^{n} x_i} = \frac{1}{\overline{x}_n}$$

$\Rightarrow$ στάσιμο σημείο $\theta^* = \dfrac{1}{\overline{x}_n}$.

4. Διερεύνηση ως προς τη φύση του στάσιμου σημείου. (διάφοροι τρόποι) Έχουμε

$$\ell''(\theta) = -\frac{n}{\theta^2} < 0, \text{ για κάθε } \theta > 0$$

Άρα η ℓ είναι **γνήσια κοίλη**, άρα το στάσιμο σημείο είναι κατ' ανάγκη **ολικό μέγιστο**. Τελικά, η $\widehat{\theta} = \widehat{\theta}(X) = \dfrac{1}{\overline{X}_n}$ είναι η ε.μ.π. του θ.

Διάλεξη 12

Παράδειγμα 0.0.18. Έστω $X_1, \dots, X_n$ τ.δ. από $\text{Be}\,(p)$, $0 < p < 1$. Να βρεθεί η ε.μ.π. του p.

Λύση
Έστω $x = (x_1, \dots, x_n) \in \{0,1\}^n$. Υπολογίζουμε τη συνάρτηση πιθανοφάνειας.

$$L\,(p) = \prod_{i=1}^{n} f_p\,(x_i) = \prod_{i=1}^{n} p^{x_i}\,(1-p)^{1-x_i} = p^{\sum\limits_{i=1}^{n} x_i}\,(1-p)^{n-\sum\limits_{i=1}^{n} x_i}$$

- *Πρώτη περίπτωση*: $\sum\limits_{i=1}^{n} x_i = 0 \Leftrightarrow x_1 = x_2 = \cdots = x_n = 0$ (με θετική πιθανότητα). Τότε $L\,(p) = (1-p)^n$ συνεπώς δεν υπάρχει η ε.μ.π. για $p \in (0,1)$ δεν πετυχαίνεται το μέγιστο. Είναι καλύτερα να πάρουμε $p \in [0,1]$ ($\text{Be}(0) \equiv 0,\ \text{Be}(1) \equiv 1$) $\longrightarrow$ συμπαγοποίηση του παραμετρικού χώρου $\Theta = [0,1]$ τότε $\widehat{p} = 0$, είναι η ε.μ.π.

- *Δεύτερη περίπτωση:* Αν $\sum\limits_{i=1}^{n} x_i = n \Leftrightarrow x_1 = x_2 = \cdots = x_n = 1$ τότε $L\,(p) = p^n$ και άρα $\widehat{p} = 1$ (έχοντας κάνει επέκταση).

- *Τρίτη περίπτωση:* Έστω $0 < \sum\limits_{i=1}^{n} x_i < n$. Θέτουμε $s_n = \sum\limits_{i=1}^{n} x_i$ και τότε

$$\ell(p) = \log L\,(p) = s_n \log p + (n - s_n) \log (1-p)$$

Η $\ell(p)$ είναι παραγωγίσιμη στο $(0,1)$ και άρα σε μία θέση μεγίστου θα έχουμε $\ell'\,(p^*) = 0$. Λύνουμε την εξίσωση πιθανοφάνειας $\ell'(p) = 0$ (να βρούμε τα στάσιμα σημεία). Άρα

$$\ell'\,(p) = \frac{s_n}{p} - \frac{n - s_n}{1-p},\ p \in (0,1) = 0$$

$$\ell'(p) = 0 \Leftrightarrow \frac{s_n}{p} = \frac{n - s_n}{1-p} \Leftrightarrow (1-p)s_n = p(n - s_n) \Leftrightarrow p^* = \frac{s_n}{n} = \overline{x}_n$$

Η $\ell(p)$ έχει **μοναδικό στάσιμο σημείο**. Αν είναι τοπικό μέγιστο, τότε είναι **ολικο**. Αρχικά:

$$L\,(p) = p^{s_n}\,(1-p)^{n-s_n} \geq 0$$

$$L\,(0) = 0^{s_n}\,(1-0)^{n-s_n} = 0$$

$$L(1) = 0 \text{ δε χρειάζεται}$$

π.χ. με $L(p) = p(1-p)$ έχω για $n = 2, \quad (1,1)\ (0,2)\ (2,0)$.

$$S_2 = 0, \quad p^0(1-p)^2 = \underset{\substack{\| \\ 1}}{0^0} \cdot 1^2 = 1 \to L(0) = 1$$

$$\ell(p) = S_n \log p + (n - S_n) \log(1-p)$$

$$\log 0 = -\infty \qquad \left(\lim_{x \to 0^+} \log x\right)$$

$$0^0 = 1 \Longrightarrow 0 \cdot \underset{\substack{\| \\ -\infty}}{\log 0} = 0$$

Πάνω στο προηγούμενο μπορούμε να δούμε μια διαφορετική επιχειρηματολογία.

(i) $\ell(p) = \underset{>0}{s_n} \log p + (n - \underset{>0}{s_n}) \log(1-p)$

Η ℓ είναι άθροισμα γνησίων κοίλων συναρτήσεων και άρα και αυτή είναι γνήσια κοίλη μέ ένα στάσιμο σημείο , άρα έχει ολικό μέγιστο. Μάλιστα και στις δύο περιπτώσεις $\widehat{p} = 0 = \overline{x}_n \quad (x_1 = x_2 = \cdots = x_n = 0)$ και $\widehat{p} = 1 = \overline{x}_n \quad (x_1 = x_2 = \cdots = x_n = 1)$ έχουμε ίδιο αποτέλεσμα. Τελικά, η $\widehat{p} = \overline{x}_n$ είναι η ζητούμενη ε.μ.π.

(ii) $\ell'(p) = \dfrac{s_n}{p} - \dfrac{n - s_n}{1-p}$, $p^* = \overline{x}_n$.
Επαληθεύστε για

$$p < \overline{x}_n,\ \ell'(p) > 0 \Rightarrow \ell \nearrow \text{ γνησίως αύξουσα στο } (0, \overline{x}_n)$$

ενώ για

$$p > \overline{x}_n,\ \ell'(p) < 0 \Rightarrow \ell \searrow \text{ γνησίως φθίνουσα στο } (\overline{x}_n, 1)$$

(iii) $\ell''(p) = -\dfrac{s_n}{p^2} - \dfrac{n - s_n}{(1-p)^2} < 0$ (γνήσια κοίλη)

(iv) $\ell''(p^*) < 0$, τότε το p^* είναι θέση τοπικού μεγίστου και ως μοναδικό στάσιμο είναι ολικό μέγιστο. Αρκεί η δεύτερη παράγωγος να είναι < 0.

Στο εν λόγω παράδειγμα η $\overline{X}_n$ είναι εμπειρική, ε.ρ., ε.μ.π..

Παράδειγμα 0.0.19. Έστω $X_1, \ldots, X_n$ τ.δ. από $\mathcal{N}(\mu, \sigma^2)$, $\mu \in \mathbb{R}$, $\sigma^2 > 0$.

(i) Να βρεθεί η ε.μ.π. του μ, όταν το σ^2 είναι γνωστό.

(ii) Να βρεθεί η ε.μ.π. του σ^2, όταν το μ είναι γνωστό.

(iii) βρεθεί η ε.μ.π. του μ και σ^2, όταν τα μ και σ^2 είναι άγνωστα.

Λύση

(i) Έστω το διάνυσμα $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ (αφού το στήριγμα της $N(\mu, \sigma^2)$ είναι όλο το $\mathbb{R}$). Έχουμε

$$L(\mu) = \prod_{i=1}^{n} f_\mu(x_i) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x_i-\mu)^2} = \left(2\pi\sigma^2\right)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2}\sum\limits_{i=1}^{n}(x_i-\mu)^2}$$

$$\max_{\mu\in\mathbb{R}} L(\mu) \overset{\text{ως πρόβλημα}}{\Leftrightarrow} \min_{\mu\in\mathbb{R}} \sum_{i=1}^{n} (x_i - \mu)^2$$

Αν θέσουμε $f(\mu) = \sum\limits_{i=1}^{n} (\mu - x_i)^2 \rightarrow$ γνήσια κυρτή συνάρτηση, τότε

$$f'(\mu) = 2\sum_{i=1}^{n} (\mu - x_i) = 2\left(nm - \sum_{i=1}^{n} x_i\right)$$

και

$$f'(\mu) = 0 \Leftrightarrow \mu^* = \frac{\sum\limits_{i=1}^{n} x_i}{n} = \overline{x}_n$$

Άρα είναι και ολικό ελάχιστο $\Rightarrow$ ολικό μέγιστο στο αρχικό πρόβλημα, και τελικά

$$\widehat{\mu} = \overline{X}_n \quad \text{(ε.ε. + ε.ρ.)}$$

(ii) Έχουμε

$$L\left(\sigma^2\right) = \prod_{i=1}^{n} f_{\sigma^2}(x_i) = \prod_{i=1}^{n} \left(2\pi\sigma^2\right)^{-1/2} e^{-\frac{1}{2\sigma^2}(x_i-\mu)^2} = \left(2\pi\sigma^2\right)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum\limits_{i=1}^{n}(x_i-\mu)^2}$$

Θέτουμε $\theta = \sigma^2$ $\left(\text{απλά αλλαγή ονόματος θα μπορούσαμε } \theta = \frac{1}{\sigma^2}\right)$ τότε

$$L(\theta) = (2\pi\theta)^{-n/2} e^{-\frac{1}{2\theta}\sum\limits_{i=1}^{n}(x_i-\mu)^2}$$

$$\ell(\theta) = \log L(\theta) = -\frac{n}{2}\log(2\pi\theta) - \frac{1}{2\theta}\sum_{i=1}^{n}(x_i - \mu)^2$$

Αν $x_1 = x_2 = \cdots x_n = \mu$ τότε

$$\ell(\theta) = c - \frac{n}{2}\log\theta \xrightarrow[\theta\to 0^+]{} +\infty$$

Άρα η $\ell(\theta)$ όχι άνω φραγμένη $\Rightarrow$ $\nexists$ η ε.μ.π.

Επειδή $P_{\sigma^2}(X_1 = \mu, \dots, X_n = \mu) = \prod_{i=1}^{n} P(X_i = \mu) = 0, \forall \sigma^2 > 0$
(είναι μηδενικής πιθανότητας)

Σε οποιαδήποτε άλλη περίπτωση

$$\ell(\theta) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\theta - \frac{1}{2\theta}\sum_{i=1}^{n}(x_i - \mu)^2$$

$$\ell'(\theta) = -\frac{n}{2\theta} + \frac{1}{2\theta^2}\sum_{i=1}^{n}(x_i - \mu)^2$$

Επομένως

$$\ell'(\theta) = 0 \Leftrightarrow \frac{1}{\theta^2}\sum_{i=1}^{n}(X_i-\mu)^2 = \frac{n}{\theta} \Leftrightarrow \theta^* = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$$

το θ^* είναι μοναδικό στάσιμο σημείο.

Είναι

$$\ell''(\theta) = \frac{n}{2\theta^2} - \frac{1}{\theta^3}\sum_{i=1}^{n}(X_i-\mu)^2, \quad \forall\theta > 0$$

Άρα

$$\ell''(\theta^*) = \frac{1}{(\theta^*)^2}\left(\frac{n}{2} - \frac{1}{\theta^*}\sum_{i=1}^{n}(x_i-\mu)^2\right)$$

$$= \frac{1}{(\theta^*)^2}\left(\frac{n}{2} - n\right)$$

$$= -\frac{n}{2(\theta^*)^2} < 0$$

$\Rightarrow$ θ^* είναι τοπικό μέγιστο άρα ως μοναδικό στάσιμο σημείο $\Rightarrow$ ολικό μέγιστο.

Τελικά η

$$\widehat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2$$

είναι η ε.μ.π. του σ^2 (με πιθανότητα 1, $\forall\sigma^2 > 0$)

Σχόλιο Πολλές φορές αγνούμε τέτοιες περιπτώσεις.

(iii) Η κατάσταση είναι δυσκολότερη για περισσότερες άγνωστες παραμέτρους. Πιο ασφαλής λύση είναι να πραγματοποιείται διαδοχική μεγιστοποιήση ("για επίλυση στο χαρτί"). Κατ' αρχήν εδώ

$$L(\mu,\sigma^2) = (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

Υποθέτουμε ότι υπάρχει ένα $\theta^* = (\mu^*, (\sigma^2)^*)$ τέτοιο ώστε $L(\theta^*) \geq L(\theta)$ για κάθε $\theta \in \mathbb{R}\times(0,+\infty)$. Τότε μπορούμε να το βρούμε με διαδοχική μεγιστοποίηση

$$\max_{\mu,\sigma^2} L(\mu,\sigma^2) = \max_{\sigma^2}\max_{\mu} L(\mu,\sigma^2) \tag{11}$$

$$= \max_{\mu}\max_{\sigma^2} L(\mu,\sigma^2) \tag{12}$$

Παρατήρηση 0.0.17.

1 * Το πρόβλημα μεγιστοποίησης ως προς μ για σταθερό σ^2 ανάγεται στην εύρεση της ε.μ.π. του μ όταν το σ^2 είναι γνωστό. Έχει λυθεί.

* Το πρόβλημα μεγιστοποίησης ως προς σ^2 για σταθερό μ ανάγεται στην εύρεση της ε.μ.π. του σ^2 όταν το μ είναι γνωστό.

2 Αν κάποια από τα δύο προβλήματα (τα μονοπαραμετρικά) οδηγεί σε λύση που είναι ανεξάρτητη από την άλλη παράμετρο, τότε απλοποιείται το πρόβλημα και ξεκινάμε έτσι.

Στο συγκεκριμένο παράδειγμα

$$\begin{aligned}
\max_{\mu,\sigma^2} L\left(\mu,\sigma^2\right) &= \max_{\sigma^2}\max_{\mu} L\left(\mu,\sigma^2\right) \\
&= \max_{\sigma^2} L\left(\overline{x}_n,\sigma^2\right) \\
&\qquad \overline{x}_n \text{ είναι ανεξ. του } \sigma^2, \mu=\overline{x}_n \text{ (“γνωστό”)} \\
&= L\left(\overline{x}_n, \frac{1}{n}\sum_{i=1}^{n}\left(x_i-\overline{x}_n\right)^2\right)
\end{aligned}$$

Έχουμε $\widehat{\theta} = \left(\widehat{\mu}, \widehat{\sigma^2}\right) = \left(\overline{X}_n, \frac{1}{n}\sum_{i=1}^{n}\left(X_i-\overline{X}_n\right)^2\right)$

Αν κάποιος είχε επιλέξει

$$\begin{aligned}
\max_{\mu,\sigma^2} L\left(\mu,\sigma^2\right) &= \max_{\mu}\max_{\sigma^2} L\left(\mu,\sigma^2\right) \quad \text{(σαν το } \mu \text{ να είναι γνωστό)} \\
&= \max_{\mu} L\Bigg(\mu, \underbrace{\frac{1}{n}\sum_{i=1}^{n}\left(x_i-\mu\right)^2}_{\downarrow\ \left(\sigma^2\right)^*(\mu)}\Bigg) \\
&= \max_{\mu} (2\pi)^{-\frac{n}{2}}\left(\left(\sigma^2\right)^*(\mu)\right)^{-\frac{n}{2}} e^{-\frac{1}{2\left(\sigma^2\right)^*(\mu)}\sum_{i=1}^{n}(x_i-\mu)^2} \\
&= (2\pi)^{-\frac{n}{2}} \max_{\mu}\left(\frac{1}{n}\sum_{i=1}^{n}\left(x_i-\mu\right)^2\right)^{-\frac{n}{2}} e^{-\frac{n}{2}}
\end{aligned}$$

Συνεπάγεται ότι $\mu^* = \overline{X}_n$ $\left(\text{διότι προκύπτει πρόβλημα ελαχιστοποίησης του } \sum_{i=1}^{n}(x_i-\mu)^2\right)$ και φτάνουμε στη λύση αντικαθιστώντας

$$\left(\sigma^2\right)^*(\mu) \to \left(\sigma^2\right)^*(\mu^*) = \frac{1}{n}\sum_{i=1}^{n}\left(X_i-\overline{X}_n\right)^2$$

Γενικά Ισχύουν

$$\max_{\theta_1,\theta_2} L\left(\theta_1,\theta_2\right) = \max_{\theta_1}\max_{\theta_2} L\left(\theta_1,\theta_2\right) = \max_{\theta_1} L\left(\theta_1,\theta_2^*\left(\theta_1\right)\right) = L\left(\theta_1^*,\theta_2^*\left(\theta_1^*\right)\right)$$

$$\max_{\theta_1,\theta_2} L\left(\theta_1,\theta_2\right) = \max_{\theta_2}\max_{\theta_1} L\left(\theta_1,\theta_2\right) = L\left(\theta_1^*\left(\theta_2^*\right),\theta_2^*\right)$$

Τα προβλήματα μεγιστοποίησης με διανυσματική παράμετρο, για να λυθούν με το χέρι, είναι πολλές φορές ευκολότερο να λύνονται με διαδοχική μεγιστοποίηση. Παραπέμπουμε στις σημειώσεις 2017-18 για τρόπους αντιμετώπισης που στηρίζονται σε επίλυση των εξισώσεων πιθανοφάνειας π.χ. $\frac{\partial \ell}{\partial \theta_1}(\theta_1, \theta_2) = 0$, $\frac{\partial \ell}{\partial \theta_2}(\theta_1, \theta_2) = 0$ και στη συνέχεια διερεύνηση των στάσιμων σημείων (υπάρχουν διάφορα προβλήματα και θέλει αρκετή προσοχή).

Διάλεξη 13

Παράδειγμα 0.0.20 (Μη ύπαρξη της ε.μ.π. - Kiefer & Wolfowitz, 1956). Έστω $X_1, \ldots, X_n$ τ.δ. από μια κατανομή που αντιστοιχεί στη μίξη:

$$X_i \sim \begin{cases} \mathcal{N}(\mu, 1), & \text{με πιθ. } 1/2 \\ \\ \mathcal{N}(\mu, \sigma^2), & \text{με πιθ. } 1/2 \end{cases}$$

Να δείξετε ότι δεν υπάρχει η ε.μ.π. για καμία τιμή του x, όταν το $\theta = (\mu, \sigma)$ είναι άγνωστο.

Λύση

Αν ακολουθεί μια μίξη 2 κατανομών, τότε

$$f(x) = w_1 f_1(x) + w_2 f_2(x),$$

όπου $w_1 + w_2 = 1,\ 0 < w_1, w_2 < 1$ με w_1, w_2 τις μικτικές αναλογίες και f_i οι σ.π. ή σ.π.π. που αντιστοιχούν στις 2 κατανομές που αναμιγνύονται. (Προκύπτει και σε περιπτώσεις του τύπου, κατανομή ύψους (ή βάρους) ανδρών / γυναικών).

Εν προκειμένω,

$$f_\theta(x) = \frac{1}{2}\varphi(x-\mu) + \frac{1}{2\sigma}\varphi\left(\frac{x-\mu}{\sigma}\right), \forall x \in \mathbb{R},$$

όπου φ η σ.π.π. της $\mathcal{N}(0,1)$ με

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2},\ \forall\ x \in \mathbb{R}.$$

Υπενθυμίζουμε ότι αν $Y_1 = \mu + Z,\ Y_2 = \mu + \sigma Z$ και $Z \sim \mathcal{N}(0,1)$, τότε

$$f_1(x) = \varphi(x-\mu)$$

είναι ο μετασχηματισμός θέσης, και

$$f_2(x) = \frac{1}{\sigma}\varphi\left(\frac{x-\mu}{\sigma}\right)$$

είναι ο μετασχηματισμός θέσης (και) κλίμακας.

Θα δείξουμε ότι η $L(\theta)$ δεν είναι άνω φραγμένη, για καμία τιμή του x.

Έστω $x \in \mathbb{R}^n$, $\theta = (\mu, \sigma) \in \mathbb{R} \times (0, +\infty)$. Η συνάρτηση πιθανοφάνειας, που αντιστοιχεί στο $x = (x_1, x_2, \dots, x_n)$, δίνεται από τον τύπο:

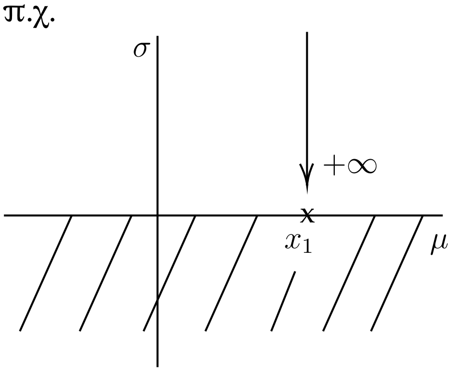
$$L(\theta) = \prod_{i=1}^{n} f_\theta(x_i) = \prod_{i=1}^{n} \left(\frac{1}{2}\varphi(x_i - \mu) + \frac{1}{2\sigma}\varphi\left(\frac{x_i - \mu}{\sigma}\right)\right)$$

Θα δείξουμε ότι

$$\lim_{\sigma \to 0} L(\underset{\substack{\| \\ x_1}}{\mu}, \sigma) = +\infty.$$

συνεπάγεται ότι η $L(\mu, \sigma)$ δεν είναι άνω φραγμένη, και άρα δεν υπάρχει η ε.μ.π. (σταθεροποιούμε την τιμή του μ στο x_1)

π.χ.



Έχουμε,

$$L(x_1, \sigma) = \left(\frac{1}{2}\varphi(0) + \frac{1}{2\sigma}\varphi(0)\right) \cdot \prod_{i=2}^{n} \left(\frac{1}{2}\varphi(x_i - x_1) + \frac{1}{2\sigma}\varphi\left(\frac{x_i - x_1}{\sigma}\right)\right)$$

όπου θέτοντας $z_i = x_i - x_1$, παίρνουμε ότι

$$L(x_1, \sigma) = \prod_{i: z_i = 0} \underset{(\text{ήδη } z_i = 0)}{\left(\frac{1}{2}\varphi(0) + \frac{1}{2\sigma}\varphi(0)\right)} \cdot \prod_{i: z_i \neq 0} \left(\frac{1}{2}\varphi(z_i) + \frac{1}{2\sigma}\varphi\left(\frac{z_i}{\sigma}\right)\right) \tag{13}$$

Παρατηρούμε ότι, για τα

- $i : z_i = 0 \Rightarrow$

$$\lim_{\sigma \to 0^+} \left(\frac{1}{2}\varphi(0) + \frac{1}{2\sigma}\varphi(0)\right) = +\infty \tag{14}$$

- $i : z_i \neq 0 \Rightarrow$

$$\lim_{\sigma \to 0^+} \left(\frac{1}{2}\varphi(z_i) + \frac{1}{2\sigma}\varphi\left(\frac{z_i}{\sigma}\right)\right) = \frac{1}{2}\varphi(z_i) + \frac{1}{2} \underbrace{\lim_{\sigma \to 0^+} \frac{1}{\sigma}\varphi\left(\frac{z_i}{\sigma}\right)}_{?} \tag{15}$$

Θέλουμε το όριο $\lim_{\sigma\to 0^+} \frac{1}{\sigma}\varphi\left(\frac{z}{\sigma}\right)$, για $z \neq 0$

$$\lim_{\sigma\to 0^+} \frac{1}{\sigma}\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2\sigma^2}} \overset{t=\frac{1}{\sigma}}{=} \lim_{t=+\infty} \frac{t}{\sqrt{2\pi}}e^{-\frac{z^2}{2}\cdot t^2}$$

$$= \frac{1}{\sqrt{2\pi}} \cdot \lim_{t\to+\infty} \frac{t}{e^{\frac{z^2}{2}\cdot t^2}}$$

$$\left[\frac{+\infty}{+\infty} \Rightarrow DLH\right] = \frac{1}{\sqrt{2\pi}} \cdot \lim_{t\to+\infty} \frac{1}{z^2 t \cdot e^{\frac{z^2}{2}\cdot t^2}}$$

$$= \frac{1}{+\infty} = 0$$

Έχουμε λοιπόν

$$\lim_{\sigma\to 0^+}\left(\frac{1}{2}\varphi(z_i) + \frac{1}{2\sigma}\varphi\left(\frac{z_i}{\sigma}\right)\right) \overset{(15)}{=} \frac{1}{2}\varphi(z_i) \tag{16}$$

Τελικά από τις (13), (14), (15) και (16) έχουμε

$$\lim_{\sigma\to 0^+} L(x_1,\sigma) = \underbrace{(+\infty)\cdots(+\infty)}_{\underset{(\geq 1)}{\#i}\ :z_i=0} \cdot \left(\prod_{i:z_i\neq 0} \underbrace{\frac{1}{2}\varphi(z_i)}_{>0}\right) = +\infty$$

Παράδειγμα 0.0.21. (Μοναδική ε.μ.π. - $L(\theta)$ όχι συνεχής)

Έστω $X_1,\ldots,X_n$ τ.δ. από $\mathcal{U}[0,\theta]$, $\theta > 0$. Να βρεθεί η ε.μ.π.

Λύση

Έστω $x = (x_1,\ldots,x_n)$, με $x_i \geq 0$, $\forall 1 \leq i \leq n$

$$f_\theta(x_i) = \begin{cases} \frac{1}{\theta}, & 0 \leq x_i \leq \theta \\ 0, & \text{αλλιώς} \end{cases} \Rightarrow f_\theta(x_i) = \frac{1}{\theta}\cdot \mathbb{1}_{[0,\theta]}(x_i)$$

Άρα,

$$L(\theta) = \prod_{i=1}^n f_\theta(x_i) = \prod_{i=1}^n \frac{1}{\theta}\cdot \mathbb{1}_{[0,\theta]}(x_i) = \left(\frac{1}{\theta}\right)^n \prod_{i=1}^n \mathbb{1}_{[0,\theta]}(x_i)$$

Τώρα,

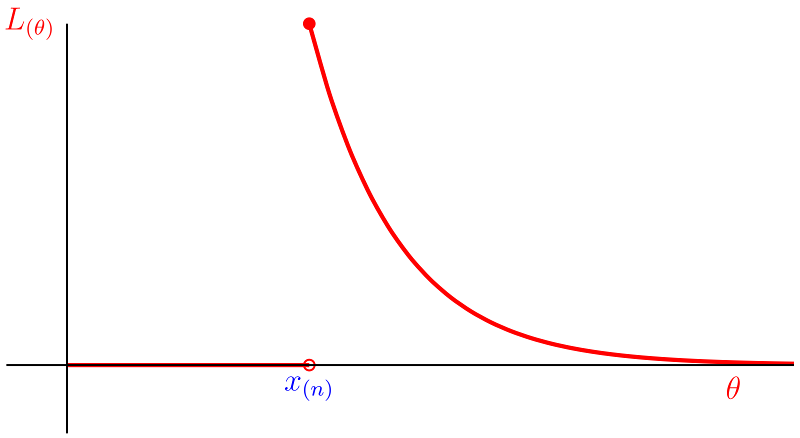
$$\mathbb{1}_{[0,\theta]}(x_i) = 1 \Leftrightarrow 0 \leq x_i \leq \theta,\ \forall 1 \leq i \leq n \Leftrightarrow x_{(n)} \leq \theta$$

Τελικά,

$$L(\theta) = \frac{1}{\theta^n}\cdot \mathbb{1}_{[x_{(n)},+\infty)}(\theta),\ \theta > 0$$

Έτσι,

$$L(\theta)=\begin{cases}0, & \theta < x_{(n)}\\ \dfrac{1}{\theta^n}, & \theta \geq x_{(n)}\end{cases}$$



Η $\dfrac{1}{\theta^n} \searrow$ είναι γνησίως φθίνουσα, άρα $\widehat{\theta}=X_{(n)}$

Θυμηθείτε ότι την είχαμε βρει ως εμπειρική εκτιμήτρια.

Παράδειγμα 0.0.22 (Παραλλαγή). Αν $X_1,\ldots,X_n$ είναι τ.δ. από $\mathcal{U}(0,\theta)$, τότε $L(\theta) = (1/\theta^n)\cdot 1_{(x_{(n)},+\infty)}(\theta)$, δηλαδή $L(\theta)=1/\theta^n$ για $\theta > x_{(n)}$, κι άρα δεν υπάρχει η ε.μ.π. - η L δεν έχει μέγιστο σημείο, κι ας είναι άνω φραγμένη!

Γενικότερα, η ύπαρξη της ε.μ.π. είναι "ευαίσθητη" στην επιλογή της σ.π.π. - ενώ αφορά την ίδια κατανομή, μπορεί είτε να υπάρχει ή να μην υπάρχει η ε.μ.π., ανάλογα με την έκδοση της σ.π.π. που επιλέγουμε. Μερικοί ορίζουν ε.μ.π. με τη βοήθεια του supremum, αν και πολλές φορές αποφεύγεται το παραπάνω πρόβλημα επιλέγοντας κλειστό διάστημα.

Παράδειγμα 0.0.23 (Μη μοναδική ε.μ.π. - άπειρες λύσεις). Έστω $_1,\ldots,X_n$ τ.δ. από $\mathcal{U}[\theta,\theta+1]$, για $\theta\in\mathbb{R}$. Να βρεθεί η ε.μ.π.

Έστω $x\in\mathbb{R}^n$ και $f(x_i)=\begin{cases}1, & \theta\leq x_i\leq\theta+1\\ 0, & \text{διαφορετικά}\end{cases}$

Άρα, $f_\theta(x_i)=1_{[\theta,\theta+1]}(x)$, και $L(\theta)=\prod_{i=1}^n f_\theta(x_i)=\prod_{i=1}^n 1_{[\theta,\theta+1]}(x_i)$

$$\begin{aligned}\theta\leq x_i\leq\theta+1 &\Leftrightarrow \forall i\in\{1,\ldots,n\}:\theta\leq x_i \text{ και } x_i\leq\theta+1\\ &\Leftrightarrow \theta\leq x_{(1)} \text{ και } x_{(n)}\leq\theta+1\\ &\Leftrightarrow x_{(n)}-1\leq\theta\leq x_{(1)}\end{aligned}$$

Τελικά $L(\theta)=1_{[x_{(n)}-1,x_{(1)}]}(\theta),\theta\in\mathbb{R}$. Κάνοντας τη γραφική της L, βλέπει κανείς ότι έχει άπειρα μέγιστα σημεία, και οποιοδήποτε $\widehat{\theta}\in[x_{(n)}-1,x_{(1)}]$ είναι ε.μ.π.

Παράδειγμα 0.0.24 (Μη μοναδική ε.μ.π., περιπτώσεις με 2 λύσεις). Έστω $X_1, X_2 \sim Cauchy(\mu, 1)$, με το μ είναι παράμετρος θέσης: $X = \mu + \sigma\mathcal{C}, \mathcal{C} \sim Cauchy(0, 1)$.

Έστω $x_1, x_2 \in \mathbb{R}$, $f_\mu(x_i) = \frac{1}{\pi} \cdot \frac{1}{1+(x_i-\mu)^2}, i = 1, 2$
$L(\mu) = \frac{1}{\pi^2} \cdot \frac{1}{(1+(x_1-\mu)^2)(1+(x_2-\mu)^2)}$. Επιλέγουμε $x_2 = 0$ - δουλέψτε τη γενική περίπτωση.
$l(\mu) = logL(\mu) = -2log\pi - log\left(1 + (\mu - x_1)^2\right) - log(1 + \mu)^2, \mu \in \mathbb{R}$
Λύνουμε $l'(\mu) = 0$, και βλέπουμε ότι:

- Αν $|x_1| < 2$, τότε η μοναδική λύση είναι $\mu^* = x_1/2$ με $l''(\mu*) < 0$, ολικό μέγιστο, άρα έχουμε μοναδική ε.μ.π.
- Αν $|x_1| = 2$, τότε έχουμε πάλι μοναδική λύση, τη $\mu^* = x_1/2$ αλλά $l''(\mu) = 0$ - παρόλα αυτά, προκύπτει ε.μ.π. - γιατί;
- Αν $|x_1| > 2$, τότε
  - $\mu_1^* = x_1/2$, ενώ $l''(\mu^*) > 0$, άρα το μ^* γίνεται τοπικό ελάχιστο.
  - $\mu_{2,3}^* = \frac{x_1 \pm \sqrt{x_1^2 - 4}}{2}$ διακεκριμένες ρίζες με $l''(\mu_{2,3}^*) < 0$ τα τοπικά μέγιστα, (που βγαίνουν ολικά κι άρα), έχουμε διπλή ε.μ.π.

Μια διαφορετική ματιά στην κατανομή κατανομή Από τότε που εμφανίστηκε η κανονική κατανομή, είχε συνδεθεί με την κατανομή των σφαλμάτων γύρω από μια θεωρητική τιμή ή μια θεωρητική ιδανική καμπύλη:

$$X = \mu + \sigma Z = \mu + \mathcal{E}, \mathcal{E} \sim \mathcal{N}(0, \sigma^2),$$

όπου το μ είναι η μέση τιμή ή η πλέον αντιπροσωπευτική τιμή του χαρακτηριστικού σε έναν πληθυσμό, και το $\mathcal{E}$ είναι ένα τυχαίο σφάλμα ή τυχαία απόκλιση από τη θεωρητική μέση τιμή: $\mathcal{E}_i = X_i - \mu$ είναι το σφάλμα της i-οστής παρατήρησης. Αν έχουμε ένα τ.δ. $_1, \ldots, X_n$, τότε $X_i = \mu + \mathcal{E}_i$. Γενικά το μ είναι άγνωστο, άρα δεν γνωρίζουμε το $\mathcal{E}_i$.
Εκτιμήτρια
Ελαχίστων Τετραγώνων
Λέμε του μ, την εκτιμήτρια που προκύπτει ως λύση της ελαχιστοποίησης του $\sum_i (\underbrace{X_i - \mu}_{\mathcal{E}_i})^2$, δηλ. του αθροίσματος των τετραγώνων σφαλμάτων.
Έτσι, έχουμε

$$\widehat{\mu} = \arg\min_{\mu \in \mathbb{R}} \sum_{i=1}^{n} (X_i - \mu)^2 = \overline{X_n},$$

όπου, arg(ument) είναι το όρισμα για το οποίο επιτυγχάνεται η τιμή στα δεξιά του arg. Εν προκειμένω, η εκτιμήτρια συμπίπτει με την ε.μ.π.

Διάλεξη 14

Παρατηρήσεις:

1. Δείξαμε ότι η λύση που προκύπτει από την ελαχιστοποίηση του $\sum_{i=1}^{n}(X_i-\mu)^2$, μας δίνει την εκτιμήτρια $\widehat{\mu}=\overline{X_n}$. Έτσι, συμπαιρένουμε ότι η εκτιμήτρια ελαχίστων τετραγώνων, συμπίπτει με την εκτιμήτρια μέγιστης πιθανοφάνειας και την εκτιμήτρια ροπών.

2. Όταν έχουμε ένα μονοπαραμετρικό πρόβλημα και θέλουμε να βρούμε την ε.ε.τ., τότε θα πρέπει να λυθεί η $\sum_{i=1}^{n}(X_i-\mu(\theta))^2$ - δηλαδή να γίνει η ελαχιστοποίηση ως προς θ, ενώ $\mu(\theta)=\mathbb{E}_\theta(X_1)$. Αν η $\mu(\theta)$ είναι παραγωγίσιμη με $\mu'(\theta)\neq 0$, τότε καταλήγουμε στο ότι $\mathbb{E}_\theta(X_1)=\mu(\theta)=\overline{X_n}$

   Δηλαδή, η ε.ε.τ. αντιστοιχεί στην εκτιμήτρια ροπών.

3. Τονίζουμε πως δεν έχουμε κάνει κάποια υπόθεση για την κατανομή των X_i, παρά μόνο μια υπόθεση για τη μέση τιμή - επομένως, μοντελοποιούμε μόνο τη μέση τιμή.

4. Στη μέθοδο των ελαχίστων τετραγώνων, η εκτίμηση της διασποράς γίνεται εμπειρικά μέσω της σχέσης

$$\widehat{\sigma^2}=\frac{\sum_{i=1}^{n}(X_i-\widehat{\mu})^2}{n}=\frac{\sum_{i=1}^{n}(X_i-\overline{X_n})^2}{n}=M_2$$

   Γράφοντας $\widehat{\mathcal{E}_i}=X_i-\widehat{\mu}$ (γνωστά και ως κατάλοιπα, residuals)[1], τότε

$$\widehat{\sigma^2}=\frac{\sum_{i=1}^{n}(\mathcal{E}_i)^2}{n}$$

5. Ένα πολύ ενδαιφέρον θέμα στη Στατιστική, είναι η *πρόβλεψη* μιας καινούριας παρατήρης. Σε αυτό το μοντέλο, αν X_{n+1} είναι μια νέα παρατήρηση (που ακόμα δεν έχουμε δει) και θέλουμε ενα την "προβλέψουμε", τι θα προτείναμε ως τιμή, στη βάση του τ.δ. που έχουμε ήδη πρόσβαση, στη βάσω δηλαδή των $X_1,\ldots,X_n$:

$$\widehat{X}_{n+1}=g(X_1,\ldots,X_n)$$

Απλή γραμμική παλινδρόμηση

Μεγάλο ενδιαφέρον παρουσιάζεται όταν - εκτός από τις παρατηρήσεις που διαθέτουμε - έχουμε επιπλέον πληροφορία, που θα μπορούσε να αξιοποιηθεί για να προβλέψουμε ακόμα καλύτερα, παρατηρήσεις που δεν έχουμε. Αυτό οδηγεί (και) στη μελέτη

[1] Περισότερα πράγματα επί τούτου, στα Γραμμικά Μοντέλα.

της στατιστικής εξάρτησης που έχουν δύο χαρακτηριστικά ενδιαφέροντος μεταξύ τους. Παραδείγματος χααρη, δύο τέτοια χαρακτηριστικά μπορούν να είναι το ύψος και το βάρος ενός ατόμου, η κατανάλωση αλκοόλ με την εμφάνιση συγκεκριμένων ασθενειών, και τα λοιπά. Σχετικό είναι, το επόμενο παράδειγμα:

Το 1929 ο Edwin Hubble μελέτησε τη σχέση μεταξύ της απόστασης από τη γη, που έχουν διάφορα νεφελώματα έξω από το Γαλαξία μας, με την ταχύτητα που απομακρύνονται από τη Γη.

Χρησιμοποίησε, για αυτό το σκοπό, 24 ουράνια αντικείμενα, για τα οποία μπορούσε να εκτιμήσει αυτά τα χαρακτηριστικά, μέσω κάποιων μεθόδων στη φυσική. Η απόσταση μετριέται σε Mpc όπου $1Mpc \cong 3,08 \cdot 10^{22}m$ και $1pc \cong 3,26$ έτη φωτός.

Τα δεδομένα είναι διαθέσιμα μέσω της R, και θα τα δούμε.

Μια απλή παρατήρηση: όσο αυξάνεται η απόσταση, τα αντικείμενα απομακρύνονται με μεγαλύτερη ταχύτητα. Αυτό το γεγονός επιβεβαιώνει πειραματικά προηγούμενες προβλέψεις, ότι το σύμπαν διαστέλλεται! Η σταθερά που αντιστεί στο συντελεστή διεύθυνση της ευθείας, είναι γνωστή ως σταθερά του $Hubble$.

Υποψήφια μοντέλα:

1. $Y_i = \beta \cdot X_i + \mathcal{E}_i, \forall 1 \leq i \leq n$, όπου τα $\mathcal{E}_i \sim \mathcal{N}(0, \sigma^2)$ ανεξάρτητα.

   Ισοδύναμα, $Y_i \sim \mathcal{N}(\beta \cdot X_i, \sigma^2)$. Έτσι, τα $Y_1, \dots, Y_n$ είναι ανεξάρτητα αλλά όχι ισόνομα, καθώς αλλάζει η μέση τιμή.

2. $Y_i = \beta_0 + \beta_1 \cdot X_i + \mathcal{E}_i, \forall 1 \leq i \leq n$, με τα $\mathcal{E}_i \sim \mathcal{N}(0, \sigma^2)$ ανεξάρτητα.

   Ισοδύναμα, $Y_i \sim \mathcal{N}(\beta_0 + \beta_1 \cdot X_i, \sigma^2)$, με τα $Y_1, \dots, Y_n$ να είναι ανεξάρτητα.

3. $Z_i = Y_i / X_i = \mu + \mathcal{E}_i$, όπου τα $\mathcal{E}_i \sim \mathcal{N}(0, \sigma^2)$ ανεξάρτητα. Εδώ, οι $Z1, \dots, Z_n$ είναι τ.μ. $\mathcal{N}(\mu, \sigma^2)$. Λύνοντας ως προς Y_i, παίρνουμε $Y_i = \mu X_i + X_i \mathcal{E}_i \sim \mathcal{N}(\mu X_i, X_i^2 \sigma^2 2)$

Εύρεση της ε.μ.π.

Για το πρώτο μοντέλο, $Y_i = bX_i + \mathcal{E}_i$, έχουμε:

$$L(\beta, \sigma^2) = \prod_{i=1}^{n} f_{\beta, \sigma^2}(y_i) = (2\pi\sigma^2)^{-n/2} \exp\left\{ \frac{-1}{2\sigma^2} \cdot \sum_{i=1}^{n} (y_i - \beta x_i)^2 \right\}$$

Μπορούμε πρώτα να μεγιστοποιήσουμε πρώτα ως προς β:

$$\arg\max_{\beta \in \mathbb{R}} L(\beta, \sigma^2) = \arg\min_{\beta \in \mathbb{R}} \sum_{i=1}^{n} (y_i - \beta x_i)^2$$

Έτσι, όριζοντας $f(\beta) = \sum_i (x_i \beta - y_i)^2$, βλέπουμε ότι η συνάρτηση ενδιαφέροντος είναι (γνήσια) κυρτή (εκτός από το $x_1 = \cdots = x_n = 0$). Επιπλέον, $f'(\beta) = 2 \cdot \sum_i x_i \cdot (x_i \beta - y_i)$. Βλέπουμε ότι $f'(\beta) = 0 \Leftrightarrow \beta^* = \dfrac{\sum_i x_i y_i}{\sum_i X_i^2}$, το μοναδικό στάσιμο σημείο της $f\cdot$ άρα, είναι και ολικό ελάχιστο. Έτσι, έχουμε την ε.μ.π.

$$\left(\widehat{\beta}, \widehat{\sigma^2}\right) = \left(\frac{\sum_{i=1}^{n} X_i Y_i}{\sum_{i=1}^{n} X_i^2}, \frac{\sum_{i=1}^{n} (Y_i - \widehat{\beta} X_i)^2}{n} \right)$$

Για το δεύτερο μοντέλο, ακολουθόντας τα ίδια βήματα παίρνουμε τη συνάρτηση πιθανοφάνειας:

$$L(\beta_0, \beta_1, \sigma^2) = (2\pi\sigma^2)^{-n/2} \exp\left\{ \frac{-1}{2\sigma^2} \cdot \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2 \right\}$$

έτσι, όπως και πριν, $\arg\max_{\beta_0,\beta_1\in\mathbb{R}} L(\beta_0,\beta_1,\sigma^2) = \arg\min_{\beta_0,\beta_1\in\mathbb{R}} \sum_{i=1}^n (y_i-\beta_0-\beta_1 x_i)^2$.

Παρατηρούμε ότι οδηγούμαστε σε (γνήσια) κυρτή - βλ. μερικές παραγώγους, $\partial f/\partial\beta_0 = \partial f/\partial\beta_1 = 0$

Για
$$f(\beta) = \sum_i (y_i-\beta_0-\beta_1 x_i)^2,$$
βλέπουμε ότι $f(\beta) = \sum_i (y_i(\beta_1)-\beta_0)^2$, όπου $y_i(\beta_1) = y_i - \beta_1 x_i$.

Έτσι,
$$\beta^*(\beta_1) = \sum_i y_i(\beta_1)/n = \overline{y} - \overline{x}\cdot\beta_1$$
Άρα, παίρνουμε την έκφραση
$$\min_{\beta_1} \sum_i (y_i - \beta_0^*(\beta_1) - \beta_1 x_i)^2 = \sum_i \left[(y_i - \overline{y_i} - \beta_1(x_i-\overline{x})\right]^2 =: \sum_i (y_i^0 - \beta_1 x_i^0)^2$$
- αναγόμαστε δηλαδή, στη λύση του πρώτου μοντέλου.

Έτσι,
$$\widehat{\beta_1} = \frac{\sum_i x_i^0 y_i^0}{\sum (x_i^0)^2} = \frac{\sum_{y=1}^n (x_i-\overline{x})(y_i-\overline{y})}{\sum_{i=1}^n (x_i-\overline{x})(x_i-\overline{x})}$$
Αν χρησιμοποιήσουμε τη δειγματική συνδιακύμανση, τότε έχουμε τις εκφράσεις
$$S_X = \sqrt{\frac{\sum(x_i-\overline{x})^2}{n}}, S_Y = \sqrt{\frac{\sum(y_i-\overline{y})^2}{n}}, S_{X,Y} = \frac{\sum(x_i-\overline{x})(y_i-\overline{y})}{n},$$
$$Cov(X,Y) = \mathbb{E}\left[(X-\mu_X)(Y-\mu_Y)\right] = \rho\cdot\sigma_X\cdot\sigma_Y$$
$$\widehat{\beta_1} = S_{X,Y}/S_X^2 = r\cdot\frac{S_X\cdot S_Y}{S_X^2} = r\cdot\frac{S_Y}{S_X}$$
$$\widehat{\beta_0} = \overline{Y} - \overline{X\beta_1}$$
$$\widehat{\sigma^2} = \frac{\sum_i (Y_i - \widehat{\beta_0} - \widehat{\beta_1}X_i)^2}{n}$$

Η ευθεία παλινδρόμησης είναι $y = \widehat{\beta_0} + \widehat{\beta_1}x = \overline{y} - \widehat{\beta_1}\overline{x} \Rightarrow \boxed{y-\overline{y} = \widehat{\beta_1}(x-\overline{x})}$

Όσον αφορά το τρίτο μοντέλο, έχουμε συναντήσει κανονικές κατανομές στο παρελθόν, κι ως εκ τούτου,
$$\left(\widehat{\mu},\widehat{\sigma^2}\right) = \left(\frac{\sum_i Y_i/X_i}{n}(=\overline{Z_n}), \frac{\sum_i (Z_i-\overline{Z_n})^2}{n}\right)$$

Διάλεξη 15

Τα δύο σημαντικότερα και πιο συχνά χρησιμοποιούμενα είναι:

$$= -2log\widehat{L} + 2k, \tag{17}$$

όπου $\widehat{L} = L(\widehat{\theta})$ - η τιμή της πιθανοφάνειας στη θέση $\theta = \widehat{\theta}$ (η ε.μ.π.) και k είναι το πλήθος των εκτιμούμενων παραμέτρων.

$$= -2log\widehat{k} + k \cdot logn, \tag{18}$$

όπου $\widehat{L}$ και k όπως πριν, ενώ n είναι το μέγεθος του δείγματος.

Το μοντέλο που έχει τη μικρότερη τιμή ως προς αυτά τα κριτήρια, είναι το καλύτερο. Μπορεί όμως, να προκύψει ένα μοντέλο που να είναι καλύτερο ως προς το AIC, ενώ ένα άλλο να είναι καλύτερο ως προς το BIC. Σε αυτήν την περίπτωση, χρειάζεται περαιτέρω διερεύνηση, ή να τα κρατήσουμε όσα κι αν είναι.

17: AIC - Akaike Information Criterion

Ο Ιάπωνας μαθηματικός H. Aikake θεμελίωσε αυτό το κριτήριο μέσω της θεωρίας πληροφορίας (Information theory), μέσω ενός πολύ σημαντικού άρθρου του.[2]

Τα κριτήρια αυτά, μαζί με το BIC, είναι έγκυρα ασυμπτωτικά (για μεγάλο n) και δίνουν μια εκτίμηση του σφάλματος πρόβλεψης ενός μοντέλου, όταν καινούργιες παρατηρήσεις θα είναι διαθέσιμες κάτω από διαφορετικές υποθέσεις.[3]

Μια διόρθωση του AIC έχει προταθεί για γραμμικά μοντέλα και μικρό μέγεθος δείγματος. Δίνεται από τη σχέση:

$$AIC_C = AIC + \frac{2k^2 + 2k}{n - k - 1}, \tag{19}$$

με k, n να είναι το πλήθος των εκτιμούμενων παραμέτρων και το μέγεθος δείγματος αντίστοιχα.

Παρατήρηση 0.0.18. Η διόρθωση του AIC σε AIC_C ισχύει για γραμμικά μοντέλα, με κανονικά σφάλματα, κι άρα πρέπει να αποφεύγεται η γενικευμένη χρήση του. Άλλες διορθώσεις έχουν προταθεί για διαφορετικά μοντέλα, και χρειάζεται βιβλιογραφική αναζήτηση.

[2] A new look at the Statistical Model Identification - 1974

[3] out-of-sample prediction error

18: BIC - Bayesian Information Criterion

Λέγεται και κριτήριο του Schawrz (1978). Εξάγεται μέσα από την επιλογή κατάλληλων prior για τις παραμέτρους του μοντέλου, αλλά χρησιμοποιείται στην κλασική στατιστική.

Η βασική διαφορά με το AIC είναι η ποινή (penalty) που δίνεται σε ένα μοντέλο με k παραμέτρους: δίνει μεγαλύτερη που είναι $logn$ αντί για 2, κι άρα "ευνοεί" μοντέλα με μικρότερο πλήθος παραμέτρων.

Ιδιότητα του αναλλοίωτου της ε.μ.π.

Έστω Θ ο παραμετρικός χώρος ενός μοντέλου, και $\widehat{\theta}$ η ε.μ.π. του θ.

Πρόταση 0.0.7. Αν $\phi = \varphi(\theta)$ με φ να είναι "1-1" συνάρτηση (συνήθως θα είναι αναπαραμέτρηση, κι άρα θα έχουμε μια αμφιδιαφορίσιμη φ), τότε η $\widehat{\phi} = \varphi(\widehat{\theta})$ είναι η ε.μ.π. του $\phi \in \Phi = \varphi(\Theta)$. [4]

Απόδειξη: Έχουμε $\phi = \varphi(\theta) \overset{1-1}{\Leftrightarrow} \theta = \varphi^{-1}(\phi)$ Επίσης, έχουμε

$$L(\theta) = \prod_{i=1}^{n} f_\theta(x_i)$$

Θέτουμε $f_\phi^* = f_{\varphi^{-1}(\phi)}(x_i)$ κι αντικαθιστούμε για να πάρουμε:

$$L^*(\phi) = \prod_{i=1}^{n} f_\phi^*(x_i) = \prod_{i=1}^{n} f_{\varphi^{-1}(\phi)}(x_i) = L\left(\varphi^{-1}(\phi)\right) \tag{20}$$

Από την υπόθεση, έχουμε $L(\widehat{\theta}) = \max_{\theta\in\Theta} L(\theta)$, άρα

$$\begin{aligned} L^*(\phi^*) &\overset{20}{=} L\left(\varphi^{-1}(\phi^*)\right) = L\left(\varphi^{-1}(\varphi(\widehat{\theta}))\right) = L(\widehat{\theta}) \\ &= \max_{\theta\in\Theta} L(\theta) = \max_{\varphi^{-1}(\phi)\in\Theta} L(\varphi^{-1}(\phi)) = \max_{\phi\in\Phi} L^*(\phi), \end{aligned}$$

άρα η $\widehat{\phi}$ είναι η ε.μ.π. του ϕ.

Παράδειγμα 0.0.25. Έστω $X_1, \ldots, X_n$ τ.δ. από $Exp(\theta), \theta > 0$ η παράμετρος ρυθμού. Δείξαμε ότι η $\widehat{\theta} = \dfrac{1}{\overline{X_n}}$ είναι η ε.μ.π. του θ.

Έστω $X_1, \ldots, X_n$ τ.δ. από $Exp(\theta), \theta > 0$ η παράμετρος κλίμακας (το θ είναι η μέση τιμή). Ποια είναι η ε.μ.π. του θ;

Απόδειξη: Δεν χρειάζεται να λυθεί· θα βγει από το αναλλοίωτο της ε.μ.π!

Έστω $\phi = \varphi(\theta) = \dfrac{1}{\theta}$. Τότε, το ϕ είναι η παράμετρος ρυθμού, αφού $\phi = \dfrac{1}{\mathbb{E}(X_1)}$

Έχουμε δείξει ότι $\widehat{\phi} = \dfrac{1}{\overline{X_n}}$. Άρα, εφόσον $\theta = \dfrac{1}{\phi}$, τότε η $\widehat{\theta} = \dfrac{1}{\widehat{\phi}} = \overline{X_n}$ είναι η ε.μ.π. του θ.

(Εδώ η $1/\phi$ είναι “1-1”, αλλά δεν χρειάζεται.)

Παρατήρηση 0.0.19. Ενώ η ε.μ.π. “απολαμβάνει” την ιδιότητα του αναλλοίωτου, δεν συμβαίνει το ίδιο με τις εκτιμήτριες ροπών. Αυτό δείνιε ένα σημαντικό πλεονέκτημα στις ε.μ.π.

Παρατήρηση 0.0.20. Με κατάλληλη επέκταση της έννοιας της πιθανοφάνειας, στην έννοια της επαγόμενης συνάρτησης πιθανοφάνειας[5], η ιδιότητα του αναλλοίτωτου επεκτείνεται και σε συναρτήσεις που δεν είναι “1-1”.

Έτσι, αν $\widehat{\theta}$ είναι η ε.μ.π. του θ, μπορούμε να λέμε ότι η $g(\widehat{\theta})$ είναι η ε.μ.π. του $g(\theta)$ χωρίς η g να είναι κατ' ανάγκη “1-1”.

Παράδειγμα 0.0.26. Αν $\theta = (\mu, \sigma^2)$ σε τ.δ. από (μ, σ^2) όπου μ, σ^2 άγνωστα, και

$$\widehat{\theta} = \widehat{(\mu, \sigma^2)} = \left(\overline{X_n}, M_2\right)$$

Πώς θα λέγαμε ότι η $\widehat{\mu} = \overline{X_n}$?

<u>**Απάντηση:**</u> Εφόσον $\widehat{\theta}$ είναι η ε.μ.π., τότε $\boxed{\mu = pr_1(\theta)}$.

Άρα

$$\widehat{\mu} = pr_1(\widehat{\theta}) = \overline{X_n}$$

Αντίστοιχα,

$$\sigma^2 = pr_2(\theta) \Rightarrow \widehat{\sigma^2} = pr_2(\theta) = M_2$$

Ένα άλλο παράδειγμα είναι, σε ένα $X_1, \ldots, X_n$ τ.δ. από $Be(p)$ με $\widehat{p} = \overline{X_n}$ και θέλουμε να εκτιμήσουμε τη διασπορά $g(p) = p(1-p)$ - δεν είναι “1-1”!

Παρόλα αυτά, $\widehat{g(p)} = g(\widehat{p}) = \widehat{p}(1 - \widehat{p})$

Ασκήσεις

1) Έστω $X_1, ..., X_n$ τ.δ. από $\mathcal{U}[\theta_1, \theta_2]$ με $\theta_1 < \theta_2 \in \mathbb{R}$. Να βρεθεί η ε.μ.π. του θ και η ε.μ.π. του εύρους $\theta_2 - \theta_1$.

<u>**Λύση:**</u>

Έστω $x = (x_1, ..., x_n) \in \mathbb{R}^n$. Τότε

$$f_\theta(x_i) = \frac{1}{\theta_2 - \theta_1} \cdot 1_{[\theta_1, \theta_2]}(x_i),$$

κι έτσι

$$L(\theta)=\prod_{i=1}^{n} f_\theta(x_i)=\prod_{i=1}^{n}\frac{1}{\theta_2-\theta_1}\cdot 1_{[\theta_1,\theta_2]}(x_i)=\frac{1}{(\theta_2-\theta_1)^n}\prod_{i=1}^{n}1_{[\theta_1,\theta_2]}(x_i)$$

Τότε,

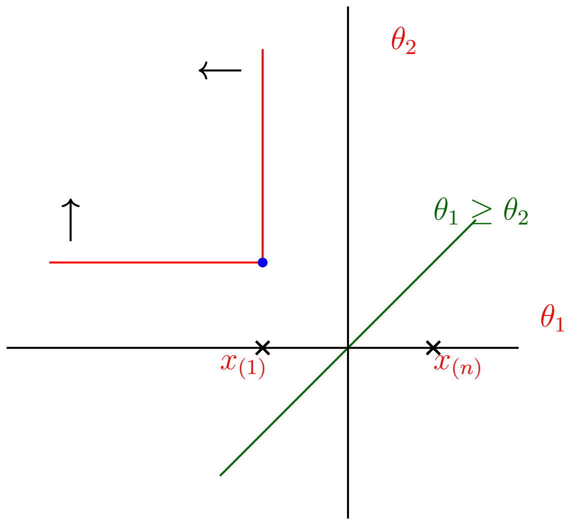
$$\begin{aligned}\forall i\in\{1,...,n\}:\theta_1\leq x_i\leq\theta_2 &\Leftrightarrow \theta_1\leq x_i \text{ και } x_i\leq\theta_2, \forall i\in\{1,...,n\}\\ &\Leftrightarrow \theta_1\leq x_{(1)} \text{ και } x_{(n)}\leq\theta_2\end{aligned}$$

Άρα,

$$L(\theta)=\frac{1}{(\theta_2-\theta_1)^n}\cdot 1_{(-\infty,x_{(1)}]}(\theta_1)\cdot 1_{[x_{(n)},+\infty)}(\theta_2), \theta_1<\theta_2$$

Υπό μία τετριμμένη περίπτωση - $x_1=\cdots=x_n$ - έχουμε $x_{(1)}=x_{(n)}$ κι άρα $L(\theta)$ μη φραγμένη, καθώς $\theta_2-\theta_1\to 0$

Ενδιαφέρουσα είναι η περίπτωση $x_{(1)}<x_{(n)}$, η οποία έχει πιθανότητα 1.



Τότε, θα έχουμε

Διάλεξη 16

Συνέχεια άσκησης: τ.δ. $\mathcal{U}[-\theta, \theta^2]$, $\theta > 0$.

$$\left.\begin{array}{r} -x_{(1)} \leq \theta \\ \\ x_{(n)} \leq \theta^2 \end{array}\right\} \Leftrightarrow \quad ?$$

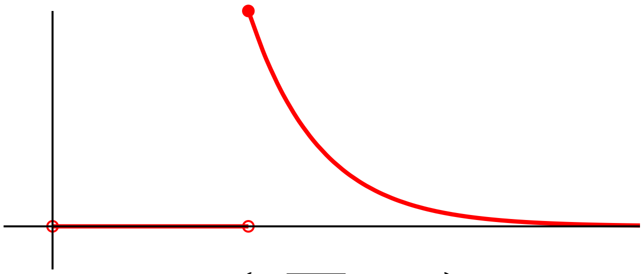
Προσοχή στα x_i (παίρνει και αρνητικές και θετικές τιμές)

- αν $x_{(n)} < 0$, τότε $x_{(1)} \leq x_{(n)} < 0 \;\; \Rightarrow \theta \geq -x_{(1)}$ (ο μόνος περιορισμός)
- αν $x_{(n)} \geq 0$ τότε

$$\left.\begin{array}{r} -x_{(1)} \leq \theta \\ \\ x_{(n)} \leq \theta^2 \end{array}\right\} \Leftrightarrow \quad \left.\begin{array}{r} -x_{(1)} \leq \theta \\ \\ \sqrt{x_{(n)}} \leq \theta \end{array}\right\} \Leftrightarrow \quad \theta \geq \max\left\{\sqrt{x_{(n)}}, -x_{(1)}\right\}$$

Ενοποιώντας τις 2 λύσεις έχουμε τελικά $\theta \geq \max\left\{\sqrt{x^+_{(n)}}, -x_{(1)}\right\}$ όπου $x^+ = \max(0, x)$.

Τελικά $L(\theta) = \dfrac{1}{(\theta^2+\theta)^n} \cdot \mathbb{1}_{\left[\max\left\{\sqrt{x^+_{(n)}}, -x_{(1)}\right\}, +\infty\right)}(\theta), \quad \theta > 0.$



Τελικά $\widehat{\theta} = \max\left\{\sqrt{X^+_{(n)}}, -X_{(1)}\right\}$ είναι η ε.μ.π. του θ.

Ορισμός 0.0.11. Έστω $X_1, \ldots, X_n$ τ.δ. με κοινή κατανομή $K(\theta)$. Το μοντέλο που προκύπτει λέγεται **ταυτοποιήσιμο** ή **αναγνωρίσιμο** (identifiable), αν για $\theta_1, \theta_2 \in \Theta$ με $\theta_1 \neq \theta_2 \Rightarrow K(\theta_1) \neq K(\theta_2)$.

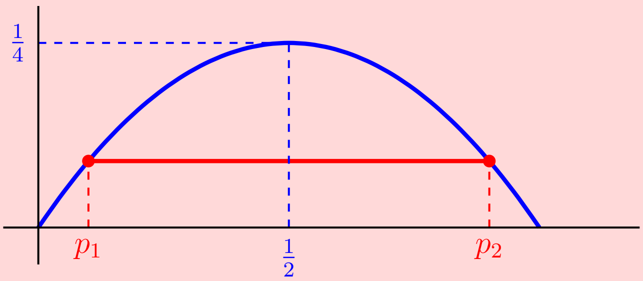
* Σε αντίθετη περίπτωση έχουμε για διαφορετικές παραμέτρους την ίδια κατανομή, και έτσι δεν θα μπορούσαμε να ξεχωρίσουμε αυτά τα θ.

Παρατήρηση 0.0.21. Υπάρχει και ένα θέμα πρακτικής ταυτοποιησιμότητας που συνήθως μας οδηγεί στη συνθήκη ότι το πλήθος των δεδομένων $\geq$ το πλήθος των παραμέτρων που θέλουμε να εκτιμήσουμε.

Παράδειγμα 0.0.27. Έστω $X_1, \dots, X_n$ τ.δ. $\text{Be}(p),\ 0 \leq p \leq 1$. Τότε το μοντέλο αυτό είναι ταυτοποιήσιμο.

Πράγματι, αρκεί για $p_1 \neq p_2 \Rightarrow \text{Be}(p_1) \neq \text{Be}(p_2)$. Αυτό ισχύει διότι αλλάζει η πιθανότητα της επτυχίας $\mathbb{P}_p(X = 1)$, άρα αλλάζει η κατανομή.

$\rightarrow$ Ας υποθέσουμε ότι στο ίδιο πρόβλημα θέλουμε να παραμετροποιήσουμε το μοντέλο με τη διασπορά. Εδώ $g(p) = p(1 - p),\ 0 \leq p \leq 1$.



Το μοντέλο δεν είναι ταυτοποιήσιμο. Για όλα τα $0 \leq \sigma^2 < \dfrac{1}{4}$, έχουμε δύο διαφορετικά p, το p και το $1 - p$.

Πρόταση 0.0.8. Κάθε αναπαραμέτρηση ενός ταυτοποιήσιμου μοντέλου, οδηγεί σε ταυτοποιήσιμο μοντέλο.

$\rightarrow$ Πράγματι, έχουμε "1-1" συνάρτηση, άρα διατηρείται η ταυτοποιησιμότητα.

Παράδειγμα 0.0.28. Έστω $X = \mu + \sigma \cdot Z,\ \mu \in \mathbb{R},\ \sigma \in \mathbb{R}$, όπου η Z έχει συμμετρική κατανομή, δηλ. $Z \stackrel{d}{=} -Z$, π.χ. $Z \sim \mathcal{N}(0, 1)$. Τότε το μοντέλο που προκύπτει με παραμέτρους το μ και το σ δεν είναι ταυτοποιήσιμο.

$\rightarrow$ Πράγματι, αν $\sigma = 1$ ή $\sigma = -1$, τότε $\mu + Z \stackrel{d}{=} \mu - Z$, δηλαδή έχουμε δύο διαφορετικά σ με κοινή κατανομή.

Διαστήματα Εμπιστοσύνης

Προς το παρόν έχουμε ορίσει εκτιμήτριες άγνωστων παραμέτρων μέσω σημειακής εκτίμησης, δηλαδή το αποτέλεσμα $\widehat{\theta}(x)$ της εκτιμήτριας $\widehat{\theta}(X)$ σε κάθε συγκεκριμένο δείγμα, είναι ένας αριθμός, άρα ένα σημείο του παραμετρικού χώρου Θ. Αν π.χ., η άγνωστη κατανομή είναι συνεχής, τότε γενικά έχουμε ότι

$$P_\theta(\widehat{\theta} = \theta) = 0, \ \forall \theta \in \Theta,$$

ενώ ακόμα και για διακριτή κατανομή γενικά θα είναι πολύ μικρή η πιθανότητα. Εξάλλου, η αρχική ιδέα της σημειακής εκτίμησης δεν είναι να συμπέσουμε ακριβώς στο άγνωστο θ, αλλά να είμαστε αρκετά κοντά.

Θα χρειαστούμε μία μέθοδο αποτίμησης του σφάλματος που κάνουμε με πιθανοτικούς όρους και αυτό επιτυγχάνεται συνήθως με τον καθορισμό ενός τυχαίου διαστήματος (τα άκρα του οποίου θα μεταβάλλονται από δείγμα σε δείγμα, ανάλογα με τα δεδομένα που έχουμε διαθέσιμα) ώστε να εξασφαλίσουμε στατιστικά ότι με μια συγκεκριμένη πιθανότητα (όσο μεγάλη εμείς επιθυμούμε) το διάστημα αυτό καλύπτει την παράμετρο που θέλουμε να εκτιμήσουμε ή ισοδύναμα η άγνωστη τιμή της παραμέτρου περιλαμβάνεται μέσα στο διάστημα αυτό.

Έτσι οδηγούμαστε στις διαστηματικές εκτιμήτριες.

Ορισμός 0.0.12. Έστω $L(X)$ και $U(X)$ δύο στατιστικές συναρτήσεις με $L(x) \le U(x), \ \forall x$ και $L(x), U(x) \in \overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$.

Λέμε διαστηματική εκτιμήτρια του θ (interval estimator) κάθε διάστημα της μορφής

$$\underset{\substack{\downarrow \\ Interval}}{I}(X) = (\underset{\substack{\downarrow \\ Lower}}{L}(X), \underset{\substack{\downarrow \\ Upper}}{U}(X)) \ \text{ ή } I(X) = [L(X), U(X)]$$

Συνήθως $I(X) \subset \Theta$, και άρα καθορίζει ένα σύνολο υποψήφιων τιμών της άγνωστης παραμέτρου.

Γράφουμε $\{\theta \in I(X)\}$ για να δηλώσουμε το ενδεχόμενο το $I(X)$ να περιλαμβάνει το θ και όχι $I(X) \ni \theta$ (αυτό θα δημιουργούσε λιγότερη σύγχυση αλλά δε συνηθίζεται).

Δε θα πρέπει επομένως να συγχέεται με το ότι η θ είναι τυχαία μεταβλητή. ΔΕΝ ΕΙΝΑΙ.

- Άρα $\mathbb{P}_\theta\left(\theta \in I(X)\right)$, δηλώνει την πιθανότητα το τυχαίο διάστημα $I(X)$ να περιλαμβάνει το θ.

Παράδειγμα 0.0.29. Για τ.δ. από $\mathcal{N}(\mu, \sigma^2)$ με σ^2 γνωστό, είχαμε βρει ότι

$$\mathbb{P}_\mu(\left|\overline{X}_n - \mu\right| > \varepsilon) = 2\left[1 - \Phi\left(\frac{\sqrt{n}\cdot\varepsilon}{\sigma}\right)\right]$$

Εδώ μας ενδιέφερε η πιθανότητα ενός ε-σφάλματος. Το ίδιο μας δίνει και μια διαστηματική εκτιμήτρια του μ, μέσα από την ακρίβεια.

$$\Rightarrow \mathbb{P}_\mu(\left|\overline{X}_n - \mu\right| \leq \varepsilon) = 1 - \left(\mathbb{P}_\mu\left|\overline{X}_n - \mu\right| > \varepsilon\right) = 2\Phi\left(\frac{\sqrt{n}\cdot\varepsilon}{\sigma}\right) - 1$$

$$\Rightarrow \mathbb{P}_\mu(\overline{X}_n - \varepsilon \leq \mu \leq \overline{X}_n + \varepsilon) = \mathbb{P}_\mu(\overline{X}_n - \varepsilon < \mu < \overline{X}_n + \varepsilon) = 2\Phi\left(\frac{\sqrt{n}\cdot\varepsilon}{\sigma}\right) - 1$$

και άρα αν

$$I(X) = (\overline{X}_n - \varepsilon,\ \overline{X}_n + \varepsilon) = \overline{X}_n \pm \varepsilon$$

είναι ένα τυχαίο διάστημα, τότε ως διαστηματική εκτιμήτρια έχουμε ότι η πιθανότητα να καλύπτει το μ είναι

$$2\Phi\left(\frac{\sqrt{n}\cdot\varepsilon}{\sigma}\right) - 1$$

Ορισμός 0.0.13. Για μια διαστηματική εκτιμήτρια $I(X)$ μιας παραμέτρου θ, λέμε **πιθανότητα κάλυψης** (coverage probability) του $I(X)$, την πιθανότητα

$$\mathbb{P}_\theta\left(\theta \in I(X)\right) \quad \text{(συνάρτηση του } \theta)$$

που αντιστοιχεί στην πιθανότητα που έχει το $I(X)$ να συμπεριλαμβάνει το θ, δηλαδή το άγνωστο θ να εμπεριέχεται μέσα στις τιμές του $I(X)$.

Πίσω στο παράδειγμα

Για δεδομένο $\varepsilon > 0$, το $I(X) = (\overline{X}_n - \varepsilon,\ \overline{X}_n + \varepsilon)$ έχει πιθανότητα κάλυψης

$$2\Phi\left(\frac{\sqrt{n}\cdot\varepsilon}{\sigma}\right) - 1,$$

π.χ. αν επιλέξουμε $\varepsilon = 1,\ \sigma = 1,\ n = 4$, τότε

$$\mathbb{P}\left[\mu \in (\overline{X} - 1,\ \overline{X} + 1)\right] = \mathbb{P}(\overline{X} - 1 < \mu < \overline{X} + 1) = 2\Phi(2) - 1 \cong 0,9544$$

(παρατηρείστε ότι εδώ η πιθανότητα κάλυψης βγαίνει ανεξάρτητη του μ).

Ορισμός 0.0.14. Για μια διαστηματική εκτιμήτρια $I(X)$ του θ, ορίζουμε ως **συντελεστή εμπιστοσύνης** γ του $I(X)$, τον αριθμό

$$\gamma := \inf_{\theta \in \Theta} \mathbb{P}_\theta(\theta \in I(X)).$$

Συνήθως θέτουμε $\gamma = 1 - \alpha$ και το λέμε $(1 - \alpha)$-Δ.Ε. (Διάστημα Εμπιστοσύνης).

Παρατήρηση 0.0.22. Ένα τυχαίο διάστημα $I(X)$ με σ.ε. (συντελεστή εμπιστοσύνης) $(1 - \alpha)$ λέγεται $(1 - \alpha)$-Δ.Ε. (confidence interval) και η συνήθης τακτική κατασκευής Δ.Ε. είναι να προκαθορίζουμε το σ.ε. $(1 - \alpha)$ για κάποιες standard τιμές του α, και στη συνέχεια αναζητάμε διαστήματα $I_{1-\alpha}(X)$ ή $I^{\theta}_{1-\alpha}(X)$ που να ικανοποιούν τη συνθήκη του σ.ε.

Τα λέμε και $100 \times (1 - \alpha)\%$-Δ.Ε. Συνήθεις τιμές του $\alpha \in \{0.1,\ 0.05,\ 0.01,\ 0.001\}$. Αντίστοιχα μιλάμε για $90\%,\ 95\%,\ 99\%$, ή $99,9\%$-Δ.Ε.

Επιπλέον πολλές φορές μας προκύπτει ότι $\mathbb{P}_\theta(\theta \in I(X)) = 1 - \alpha,\ \forall \theta \in \Theta$.

- Τότε πολλοί αναφέρονται σε αυτήν την περίπτωση ως ακριβή $(1 - \alpha)$-Δ.Ε. (όπως το παράδειγμα πριν).
- Άλλοι αναφέρονται ως ακριβές $(1 - \alpha)$-Δ.Ε. όταν ικανοποιείται ο ορισμός, δηλαδή όταν
$$\inf_{\theta \in \Theta} \mathbb{P}_\theta(\theta \in I(X)) = 1 - \alpha$$
άρα στην ουσία ταυτίζεται η έννοια του σ.ε. με τη δημιουργία ενός ακριβούς $(1-\alpha)$-Δ.Ε.
- Σε κάποιες άλλες περιπτώσεις δεν μπορούμε να φτιάξουμε ακριβή $(1 - \alpha)$-Δ.Ε., όπως π.χ. γενικά όταν έχουμε διακριτές κατανομές (εκεί θα χρησιμοποιήσουμε και τυχαιοποίηση) ή ακόμα και αν είναι εφικτό, πολλές φορές προτιμάμε προσεγγιστικά $(1 - \alpha)$-Δ.Ε. (π.χ. ασυμπτωτικά).

Διάλεξη 17

Πρόβλημα 1 προσδιορισμός ενός $(1-\alpha)$-Δ.Ε. για τη μέση τιμή μ ενός τ.δ. $\mathcal{N}(\mu, \sigma^2)$, με σ^2 γνωστό.

Λύση

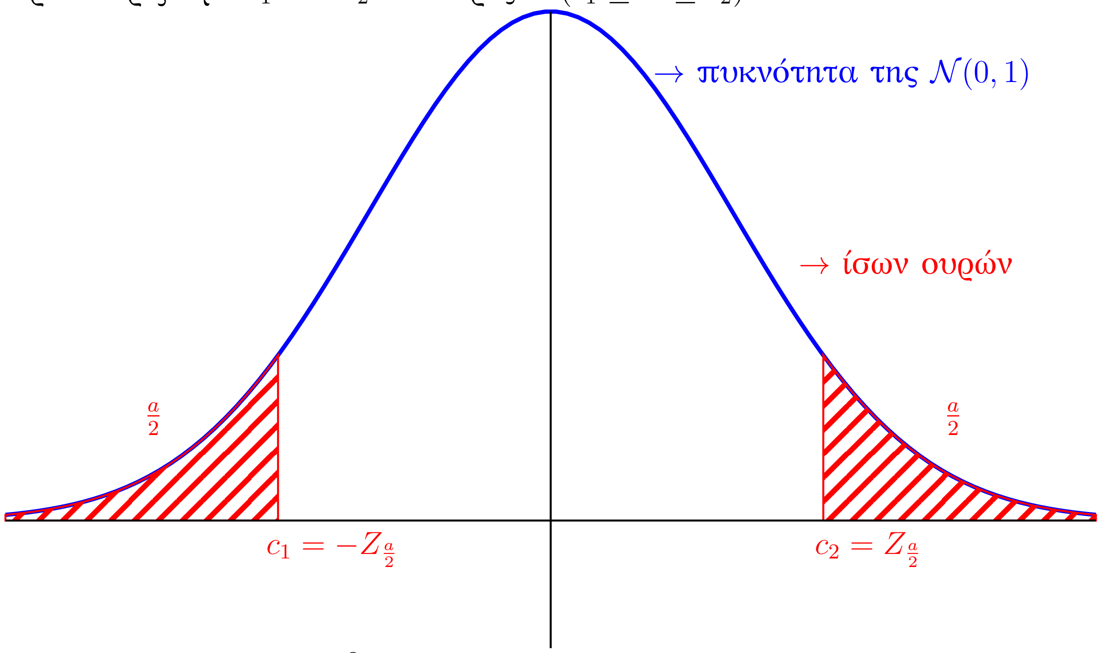
- Ξεκινάμε από μία εκτιμήτρια. Εδώ $\overline{X}_n$ είναι εκτιμήτρια του μ.

$$\overline{X}_n \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$$

- Φτιάχνουμε μια ποσότητα οδηγό (pivot) με κατανομή ανεξάρτητη του μ.

$$Z = \sqrt{n}\,\frac{\overline{X}_n - \mu}{\sigma} \sim \mathcal{N}(0,1)$$

- Προσδιορίζουμε c_1 και c_2 σταθερές: $\mathbb{P}(c_1 \le Z \le c_2) = 1-\alpha$



όπου $Z_{\frac{\alpha}{2}}$ είναι το άνω $\frac{\alpha}{2}$-ποσοστημόριο της $\mathcal{N}(0,1)$ και

$$Z_{\frac{\alpha}{2}} : \ \mathbb{P}\left(Z > Z_{\frac{\alpha}{2}}\right) = \frac{\alpha}{2}$$

Λέγεται Δ.Ε. ίσων ουρών γιατί επιλέξαμε $\left(\frac{\alpha}{2}, \frac{\alpha}{2}\right)$. Είναι ο πιο απλός τρόπος επιλογής.

$$\left[\mathbb{P}(c_1 \le Z \le c_2) = \Phi(c_2) - \Phi(c_1) = \Phi\left(Z_{\frac{\alpha}{2}}\right) - \Phi\left(-Z_{\frac{\alpha}{2}}\right) = 1 - \frac{\alpha}{2} - \frac{\alpha}{2} = 1-\alpha\right]$$

- $$-Z_{\frac{\alpha}{2}} \leq \sqrt{n} \cdot \frac{\overline{X} - \mu}{\sigma} \leq Z_{\frac{\alpha}{2}}$$

  $$\Leftrightarrow \overline{X} - Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} + Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

  $$\Leftrightarrow \mu \in \left[\overline{X} - Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}},\ \overline{X} + Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}\right]$$

  ή πιο σύντομα γράφουμε $\mu \in \underset{\text{εκτιμήτρια}}{\overline{X}} \pm \underset{\text{ποσοστημόριο}}{Z_{\frac{\alpha}{2}}} \cdot \underset{\sigma_{\overline{X}}}{\frac{\sigma}{\sqrt{n}}}$

Το σημαντικότερο βήμα εδώ είναι η εύρεση μιας κατάλληλης ποσότητας οδηγού (αντιστρεπτή ποσότητα), δηλαδή μιας ποσότητας $Q(\underset{\text{τ.δ.}}{X},\ \underset{\text{παράμετρος}}{\theta})$ έτσι ώστε να έχει κατανομή ανεξάρτητη της παραμέτρου θ.

ειδική περίπτωση: ancillary statistic (συμπληρωματική σ.σ.) που είναι ποσότητα οδηγός η έκφραση της οποίας δεν εξαρτάται από την παράμετρο θ (έχουν κατανομή ανεξάρτητη του θ), αλλά συμπληρωματικά με άλλες σ.σ. μπορούν να δώσουν πληροφορία για το θ.

Τι τρόποι εύρεσης υπάρχουν;

Προς το παρόν θα χρησιμοποιήσουμε **εκτιμήτριες** (άρα θα σκεφτόμαστε κατάλληλες εκτιμήτριες), αργότερα θα εμπλακεί και η έννοια της **επαρκούς σ.σ.**, δηλαδή σ.σ. που θα "κουβαλάνε" σχεδόν ό,τι χρειαζόμαστε για να εξάγουμε οτιδήποτε συμπέρασμα θέλουμε για το θ.

Σε κάποιες περιπτώσεις θα μας προκύπτουν φυσιολογικά από τα μοντέλα κατανομές ανεξάρτητες της παραμέτρου θ, π.χ. οικογένειες θέσης, κλίμακας κ.λ.π.

Πρόβλημα 2 Τι θα γινόταν αν και το σ^2 είναι άγνωστο; Πώς θα φτιάχναμε $(1-\alpha)$-Δ.Ε. για το μ, αλλά και για το σ^2;

Λύση

- ο δ.μ. $\overline{X}$ είναι εκτιμήτρια του μ

  $$\sqrt{n} \cdot \frac{\overline{X} - \mu}{\sigma} \sim \mathcal{N}(0,1), \text{ με } \sigma \text{ άγνωστο.}$$

  (παρατηρούμε ότι πάλι $\mathcal{N}(0,1)$ είναι!)

  Πρέπει να εκτιμηθεί το σ. Θα βάλουμε $\widehat{\sigma} = S = \sqrt{S^2}$, όπου S^2 η α.ε. του σ^2.

Θέτουμε $T = \sqrt{n} \cdot \dfrac{\overline{X} - \mu}{S}$. Ποια είναι η κατανομή της T;

Είναι ποσότητα οδηγός και θα χρησιμοποιηθεί για την κατασκευή Δ.Ε. Παρατηρείστε λοιπόν ότι

$$T = \frac{\dfrac{\dfrac{\overline{X} - \mu}{\sigma}}{\sqrt{n}}}{\dfrac{S}{\sigma}} = \frac{\dfrac{\dfrac{\overline{X} - \mu}{\sigma}}{\sqrt{n}}}{\sqrt{\dfrac{\left(\dfrac{(n-1)S^2}{\sigma^2}\right)}{n-1}}} = \frac{Z}{\sqrt{\dfrac{Y}{n-1}}}$$

όπου $Z \sim \mathcal{N}(0,1)$ και $Y \sim \chi^2_{n-1}$ και επειδή το $\overline{X}$ είναι ανεξάρτητο από το S^2, έχουμε επίσης ότι Z και Y είναι ανεξάρτητες τ.μ.

Ορισμός 0.0.15. Έστω $T = \dfrac{Z}{\sqrt{\dfrac{Y}{n}}}$ όπου $Z \sim \mathcal{N}(0,1),\ Y \sim \chi^2_n$ και $Z \perp\!\!\!\perp Y$.

Τότε $T \sim t_n$ που λέγεται **κατανομή Student** με n βαθμούς ελευθερίας.

Γυρνώντας πίσω στο πρόβλημα έχουμε ότι $\boxed{T = \sqrt{n} \cdot \dfrac{\overline{X} - \mu}{S} \sim t_{n-1}}$

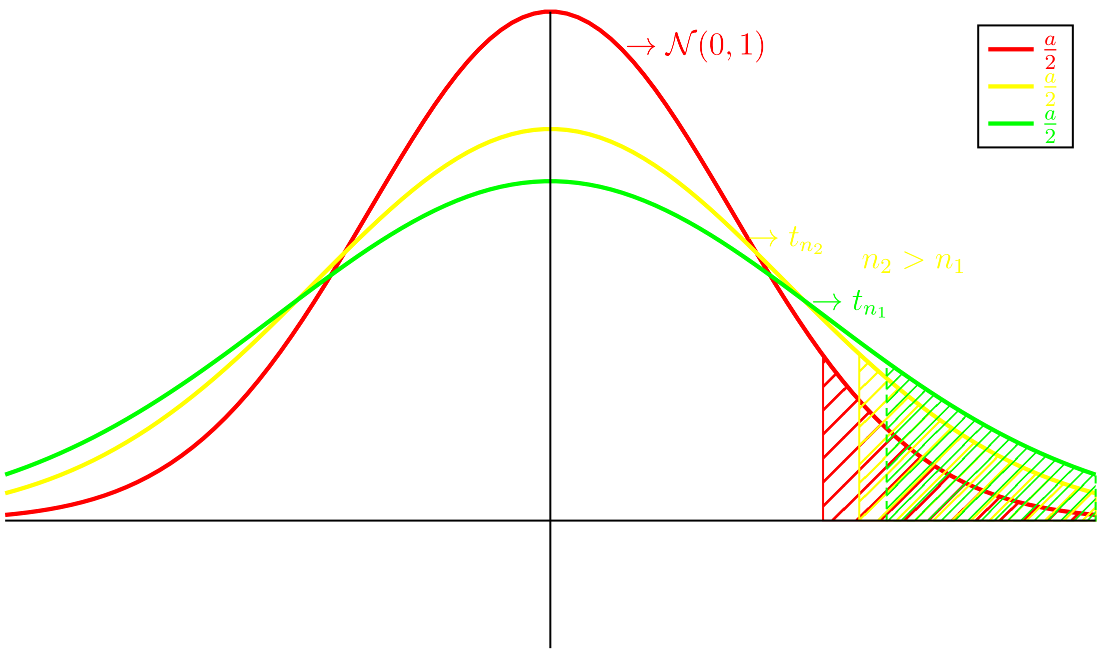
Ιδιότητες της t_n

$$t_n = \frac{\mathcal{N}(0,1)}{\sqrt{\dfrac{\chi^2_n}{n}}} + \text{ ανεξ.}$$

Άρα μπορούμε να γράψουμε μία $T \sim t_n$ ως

$$T = \frac{Z}{\sqrt{\dfrac{\sum_{i=1}^{n} Z_i^2}{n}}}, \quad \text{με } Z, Z_i \sim \mathcal{N}(0,1), \text{ και } \{Z, Z_1, Z_2, \ldots\} \text{ ανεξάρτητες τ.μ.}$$

Σχήμα της πυκνότητας της t_n



- Παρατηρούμε ότι έχει συμμετρική κατανομή, δηλαδή $T \stackrel{d}{=} -T$ (υπενθύμιση: $Z \stackrel{d}{=} -Z$) και ότι έχει πιο βαριές ουρές από αυτές της $\mathcal{N}(0,1)$.
  - Για $n = 1$,

$$T = \frac{Z_1}{\sqrt{Y}}, \quad Y \sim \chi_1^2 = \mathcal{N}^2(0,1)$$

$$= \frac{Z_1}{Z_2}, \quad \text{με } Z_1 \perp\!\!\!\perp Z_2 \text{ και } Z_1, Z_2 \sim \mathcal{N}(0,1)$$

$$\Rightarrow \boxed{T \sim \text{Cauchy}(0,1)}$$

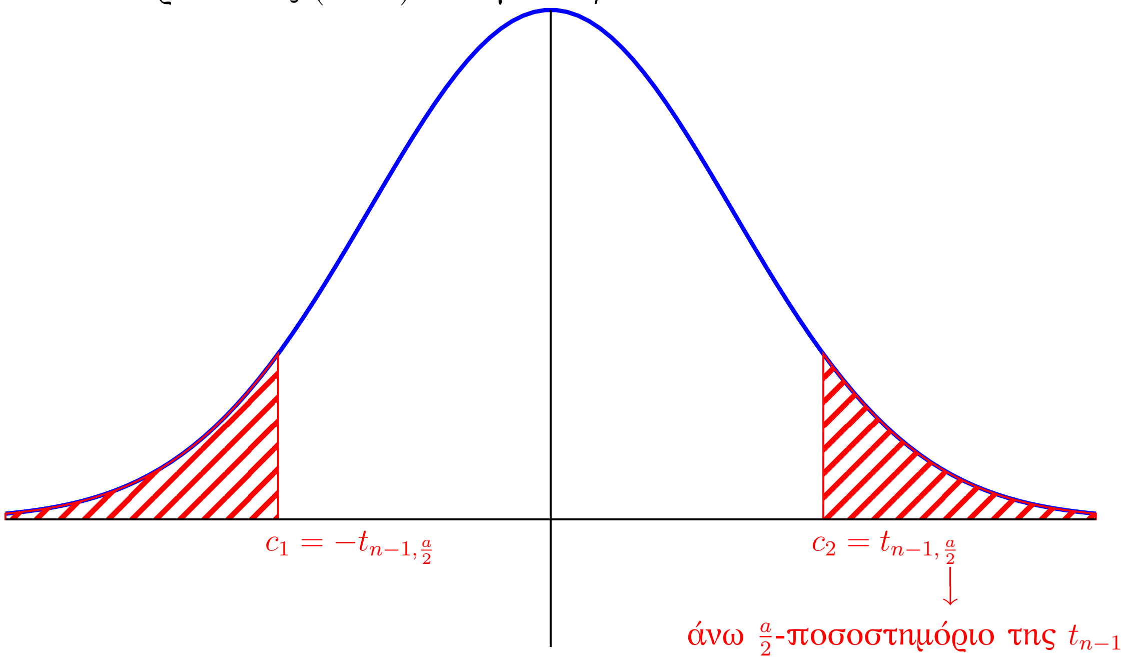
    Βλέπουμε ότι για $n = 1$, δεν έχει μέση τιμή.

  - για $n \geq 2$, $\mathbb{E}(T) = 0$ (διότι είναι συμμετρική)

  - για $n \geq 3$, $\mathbb{V}(T) = \dfrac{n}{n-2}$

Μετά θα δούμε ότι για μεγάλο n, $t_n \cong \mathcal{N}(0,1)$.

- Για την εύρεση ενός $(1-\alpha)$-Δ.Ε. για το μ:



- Άρα

$$\mu \in \left[\overline{X} - t_{n-1,\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}},\ \overline{X} + t_{n-1,\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}}\right]$$

ή σύντομα

$$\boxed{\mu \in \overline{X} \pm t_{n-1,\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}}}$$

Άρα βγάζει μεγαλύτερα διαστήματα από αυτό που θα προέκυπτε αν στη θέση του $t_{n-1,\frac{\alpha}{2}}$ αντικαταστήσουμε με $Z_{\frac{\alpha}{2}}$.

Έχουμε λοιπόν,

$$t_{n-1,\frac{\alpha}{2}} > Z_{\frac{\alpha}{2}}$$

Πρόταση 0.0.9. Αν $T_n \sim t_n$, τότε $T_n \xrightarrow{d} Z \sim \mathcal{N}(0,1) \Rightarrow$ ασυμπτωτικά $t_n \cong \mathcal{N}(0,1)$.

Απόδειξη

Θέτουμε

$$T_n = \frac{Z}{\sqrt{\frac{\sum\limits_{i=1}^{n} Z_i^2}{n}}}, \qquad \text{όπου } \{Z, Z_1, Z_2, \ldots\} \text{ ανεξάρτητες τ.μ..}$$

Τότε $T_n \sim t_n$.

Έχουμε

$$\frac{\sum\limits_{i=1}^{n} Z_i^2}{n} = \overline{Z_n^2} \xrightarrow{a.s.} \mathbb{E}\left(Z^2\right)$$

όπου

$$\mathbb{E}\left(Z^2\right) = \mathbb{V}(Z) + \underbrace{\mathbb{E}^2(Z)}_{=0} = 1$$

Από το Θεώρημα Slutsky

$$T_n = \frac{Z}{\sqrt{\overline{Z_n^2}}} \xrightarrow{d} Z \sim \mathcal{N}(0,1)$$

αφού $\sqrt{\overline{Z_n^2}} \xrightarrow{a.s.} 1$

Μάλιστα μπορούμε να αντικαταστήσουμε $t_{n-1,\frac{\alpha}{2}}$ για μεγάλο n με το $Z_{\frac{\alpha}{2}}$.

Πρόταση 0.0.10.

$$F_n^{-1}(\alpha) \xrightarrow[n\to+\infty]{} \Phi^{-1}(\alpha)$$

ή ισοδύναμα

$$q_\alpha^{t_n} \longrightarrow q_\alpha^{\mathcal{N}(0,1)}$$

Απόδειξη

Έχουμε ήδη δείξει ότι

$$T_n \xrightarrow{d} Z \sim \mathcal{N}(0,1) \Rightarrow F_n(x) \xrightarrow[n\to+\infty]{} \Phi(x), \quad \forall x \in \mathbb{R},$$

διότι η Φ είναι συνεχης.

$$F_n(Z) \xrightarrow{a.s.} \Phi(Z), \quad \text{διότι } F_n\left(Z\left(\omega\right)\right) \to \Phi(Z(\omega)),\ \forall\omega$$

$$\Rightarrow F_n(Z) \xrightarrow{d} \Phi(Z)$$

Έστω $0 < \alpha < 1$.

$$\mathbb{P}\left(F_n(Z) \leq \alpha\right) \xrightarrow[n\to+\infty]{} \mathbb{P}\left(\Phi(Z) \leq \alpha\right)$$

$$\text{ή} \quad \mathbb{P}\left(Z \leq F_n^{-1}(\alpha)\right) \xrightarrow[n\to+\infty]{} \mathbb{P}\left(Z \leq \Phi^{-1}(\alpha)\right)$$

$$\text{ή} \quad \Phi\left(F_n^{-1}(\alpha)\right) \xrightarrow[n\to+\infty]{} \Phi\left(\Phi^{-1}(\alpha)\right) = \alpha$$

$$\overset{\Phi^{-1}\ \text{συνεχής}}{\Rightarrow} F_n^{-1}(\alpha) \xrightarrow[n\to+\infty]{} \Phi^{-1}(\alpha)$$

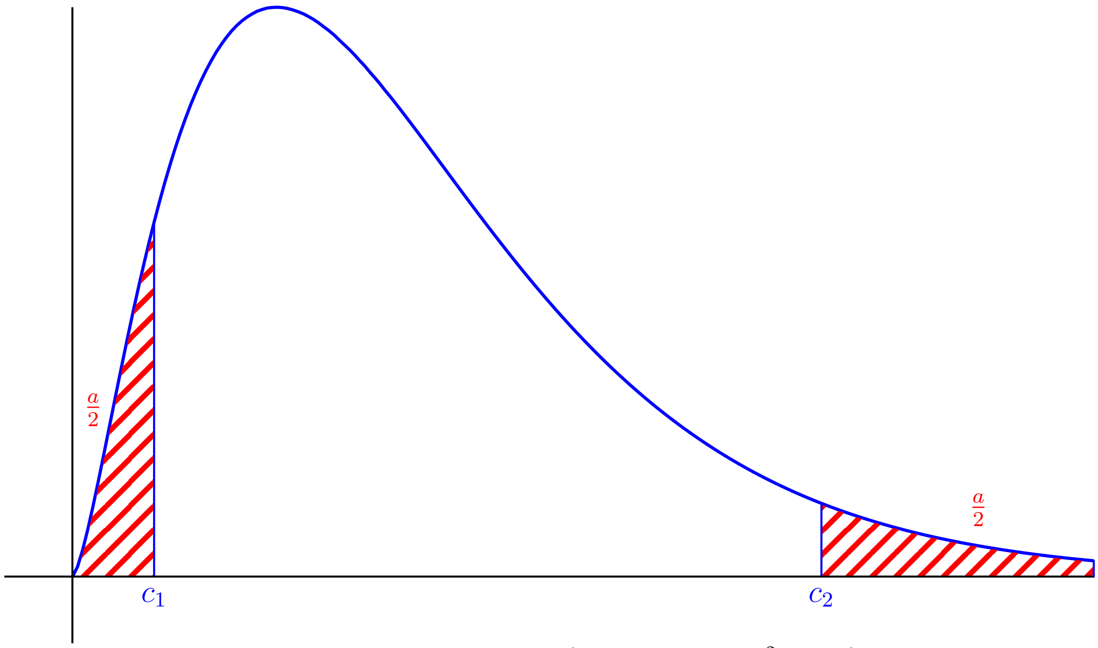
$$\text{ή} \quad \boxed{q_\alpha^{t_n} \longrightarrow q_\alpha^{\mathcal{N}(0,1)}}$$

$$\text{ή} \quad \boxed{t_{n,\alpha} \xrightarrow[n\to+\infty]{} Z_\alpha}$$

Άρα πρακτικά για μεγάλο n

$$\overline{X}_n \pm t_{n-1,\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}} \cong \overline{X}_n \pm Z_{\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}}$$

Πώς φτιάχνουμε $(1-\alpha)$-Δ.Ε. για το σ^2 ?

- S^2 εκτιμήτρια του σ^2
- ποσότητα οδηγό
$$\boxed{\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}} \to \text{ ανεξάρτητη του } \sigma^2$$
- 

όπου $c_1 = \chi^2_{n-1,1-\frac{\alpha}{2}}$, $c_2 = \chi^2_{n-1,\frac{\alpha}{2}}$ και $\mathbb{P}\left(c_1 \leq \frac{(n-1)S^2}{\sigma^2} \leq c_2\right) = 1-\alpha.$

Έχουμε

$$c_1 \leq \frac{(n-1)S^2}{\sigma^2} \leq c_2 \Leftrightarrow$$

$$\frac{1}{c_2} \leq \frac{\sigma^2}{(n-1)S^2} \leq \frac{1}{c_1} \Leftrightarrow$$

$$\frac{(n-1)S^2}{c_2} \leq \sigma^2 \leq \frac{(n-1)S^2}{c_1} \Leftrightarrow$$

$$\boxed{\sigma^2 \in \left[\frac{(n-1)S^2}{\chi^2_{n-1,\frac{\alpha}{2}}}, \frac{(n-1)S^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}}\right]}$$

Άσκηση

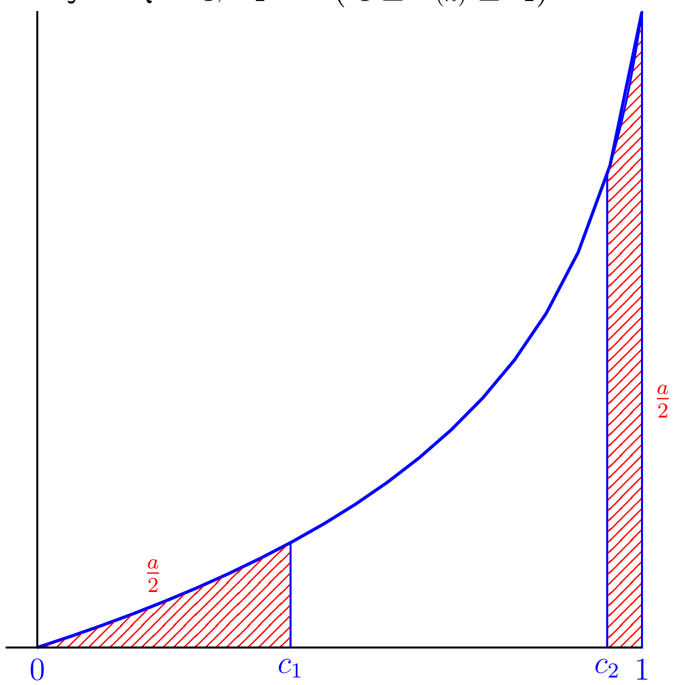
Έστω $X_1, \ldots, X_n$ τ.δ. $\mathcal{U}[0,\theta]$, $\theta > 0$. Να βρεθεί $(1-\alpha)$-Δ.Ε. για το θ.

Λύση

Εκτιμήτρια του θ ?

Έχουμε δει ότι $\widehat{\theta}_n = X_{(n)}$ και είχαμε δείξει ότι $\boxed{X_{(n)} = \theta \cdot U_{(n)}}$ όπου $U_{(n)}$ το μέγιστο από n ομοιόμορφες στο [0,1].

- $\dfrac{X_{(n)}}{\theta} = U_{(n)}$ είναι ποσότητα οδηγός
- Αναζητούμε $c_1,\ c_2 :\ \mathbb{P}\left(c_1 \leq U_{(n)} \leq c_2\right) = 1-\alpha$.



Έχουμε δείξει ότι $U_{(n)} \sim \mathrm{B}(n,1)$ και θέλουμε $F_{U_{(n)}}(u) = u^n,\ 0 < u < 1$.

$$\left.\begin{array}{ll} c_1 : & F_{U_{(n)}}(c_1) = \dfrac{\alpha}{2} \\[2ex] c_2 : & F_{U_{(n)}}(c_2) = 1 - \dfrac{\alpha}{2} \end{array}\right| \Leftrightarrow \left.\begin{array}{l} c_1^n = \dfrac{\alpha}{2} \\[2ex] c_2^n = 1 - \dfrac{\alpha}{2} \end{array}\right| \Leftrightarrow \begin{array}{l} c_1 = \sqrt[n]{\dfrac{\alpha}{2}} \\[2ex] c_2 = \sqrt[n]{1 - \dfrac{\alpha}{2}} \end{array}$$

-

$$c_1 \leq \underset{\displaystyle\frac{X_{(n)}}{\theta}}{\underset{\shortparallel}{U_{(n)}}} \leq c_2 \Leftrightarrow$$

$$\frac{1}{c_2} \leq \frac{\theta}{X_{(n)}} \leq \frac{1}{c_1} \Leftrightarrow$$

$$\frac{X_{(n)}}{c_2} \leq \theta \leq \frac{X_{(n)}}{c_1}$$

Άρα

$$I_{1-\alpha}^{\theta}(X) = \left[\frac{X_{(n)}}{\sqrt[n]{1-\dfrac{\alpha}{2}}},\ \frac{X_{(n)}}{\sqrt[n]{\dfrac{\alpha}{2}}} \right]$$

Απορίες στην ύλη του Quiz 2

① Συνάρτηση Πιθανοφάνειας με 3 παραμέτρους?

-Ίδια τεχνική

- ή κάνουμε 3 μονοδιάστατες μεγιστοποιήσεις (σταθεροποιώντας πρώτα τις δύο, μετά κάποια άλλη)
- ή λύνουμε εξισώσεις πιθανοφάνειας εφόσον υπάρχουν οι μερικές παράγωγοι

② Μετασχηματισμοί Κλίμακας

Εκθετική

- θ: παράμετρος ρυθμού

$$\theta \cdot e^{-\theta x},\ x > 0$$

$$\mathbb{E}(X_1) = \frac{1}{\theta} \quad \left(\text{γιατί } \theta = \frac{1}{\mathbb{E}(X_1)}, \text{ όταν } \theta \text{ παράμετρος ρυθμού}\right).$$

- θ: παράμετρος κλίμακας $\to$ Λειτουργεί πολλαπλασιαστικά, ως $\theta\cdot$ (τ.μ.) (Μπορώ να σκέφτομαι $\theta = \mathbb{E}(X_1)$)

$$X \sim \text{Exp}(\theta)$$

$$X = \theta \cdot Y,\ Y \sim \text{Exp}(1),\ \mathbb{E}(Y) = 1 \ \left(\text{αφού θα ήταν } \int x \cdot e^{-x} dx = 1\right).$$

Γάμμα

- $X \sim \text{G}(\alpha, \theta)\ \to$ δεν είναι κατευθείαν ρυθμός

$$\mathbb{E}(X) = \frac{\alpha}{\theta} \quad \text{(σαν να είναι ρυθμού)}$$

$$\frac{\theta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\theta x}$$

και

$$\mathrm{G}(\alpha, \theta) = \frac{1}{\theta} \cdot \mathrm{G}(\alpha, 1)$$

όπου με $\alpha = \dfrac{n}{2}$ έχουμε

$$\begin{aligned}
\mathrm{G}\left(\frac{n}{2}, \theta\right) &= \frac{1}{\theta} \cdot \mathrm{G}\left(\frac{n}{2}, 1\right) \\
&= \frac{1}{2\theta}\mathrm{G}\left(\frac{n}{2}, \frac{1}{2}\right) \\
&= \frac{1}{2\theta} \cdot \chi_n^2
\end{aligned}$$

αφού $\chi_n^2 \equiv \mathrm{G}\left(\dfrac{n}{2}, \dfrac{1}{2}\right)$.

- $X \sim \mathrm{G}(\alpha, \theta) \ \rightarrow$ παράμετρος κλίμακας

$$X \sim \theta \cdot \mathrm{G}(\alpha, 1) = \mathrm{G}(\alpha, \theta)$$

όπου

$$\mathbb{E}(\mathrm{G}(\alpha, \theta)) = \alpha \cdot \theta$$

$$\mathbb{V}(\mathrm{G}(\alpha, \theta)) = \alpha \cdot \theta^2$$

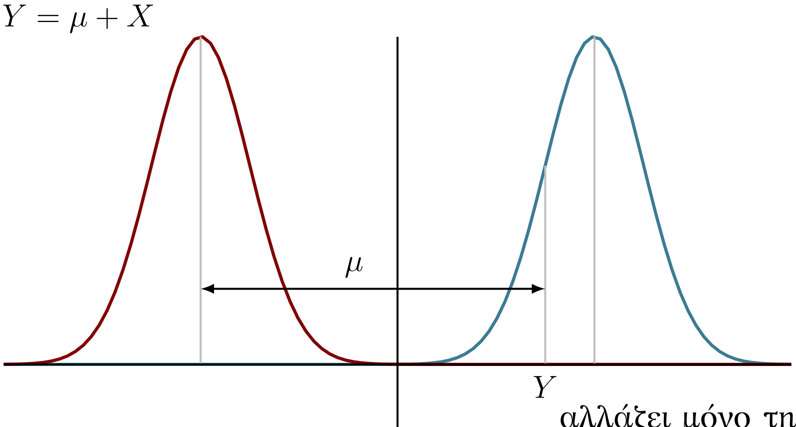
Εδώ $\chi_n^2 \equiv \mathrm{G}\left(\dfrac{n}{2}, 2\right)$.

Παράμετρος θέσης $\rightarrow$ μετατόπιση $\searrow$

δηλαδή προσθέτουμε μια σταθερά

μετασχηματισμός θέσης

$Y = \mu + X$



αλλάζει μόνο τη θέση κατανομής

$f_Y(y) = f_X(Y - \mu)$

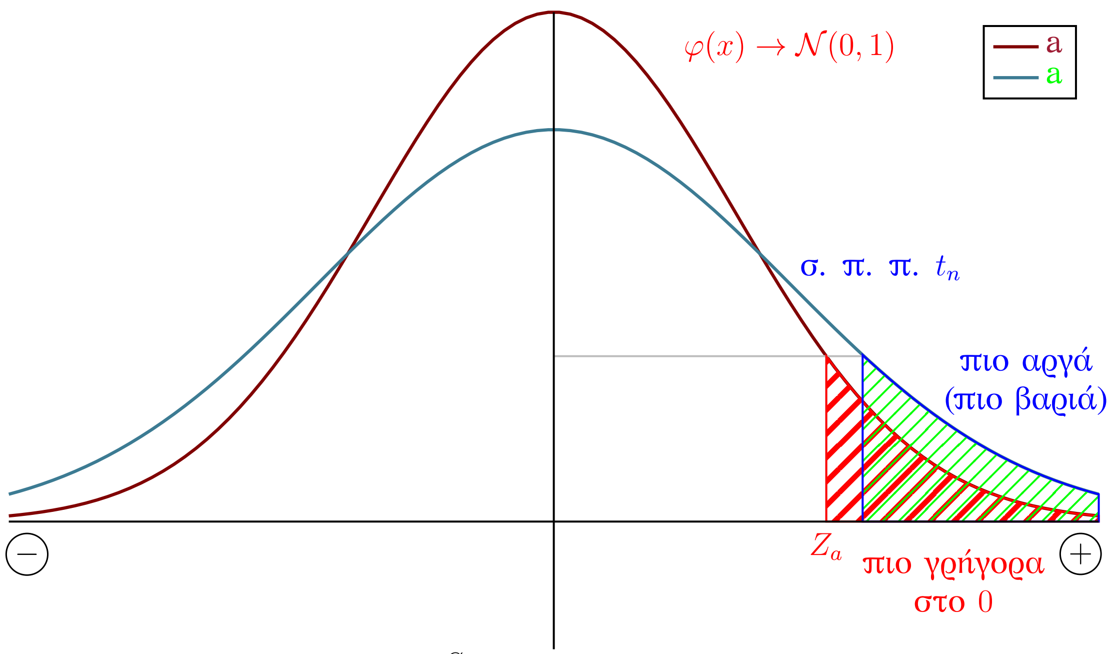
μετασχηματισμός κλίμακας

$$Y = \sigma \cdot X,\ \sigma > 0 \ \rightarrow\ f_Y(y) = \frac{1}{\sigma} \cdot f_X\left(\frac{Y}{\sigma}\right).$$

3 Ποσότητα οδηγός με μ και σ^2 άγνωστοι σε τ.δ. $\mathcal{N}(\mu, \sigma^2)$

- Όταν το σ^2 είναι γνωστό, έχουμε κανονική.
- Όταν το σ^2 είναι άγνωστο, έχουμε *Student*.

$$T \sim t_{n-1} \xrightarrow{\text{τυποποίηση}} T = \frac{\overline{X} - \mu}{\frac{S}{\sqrt{n}}} = \boxed{\sqrt{n}\frac{\overline{X} - \mu}{S}}.$$

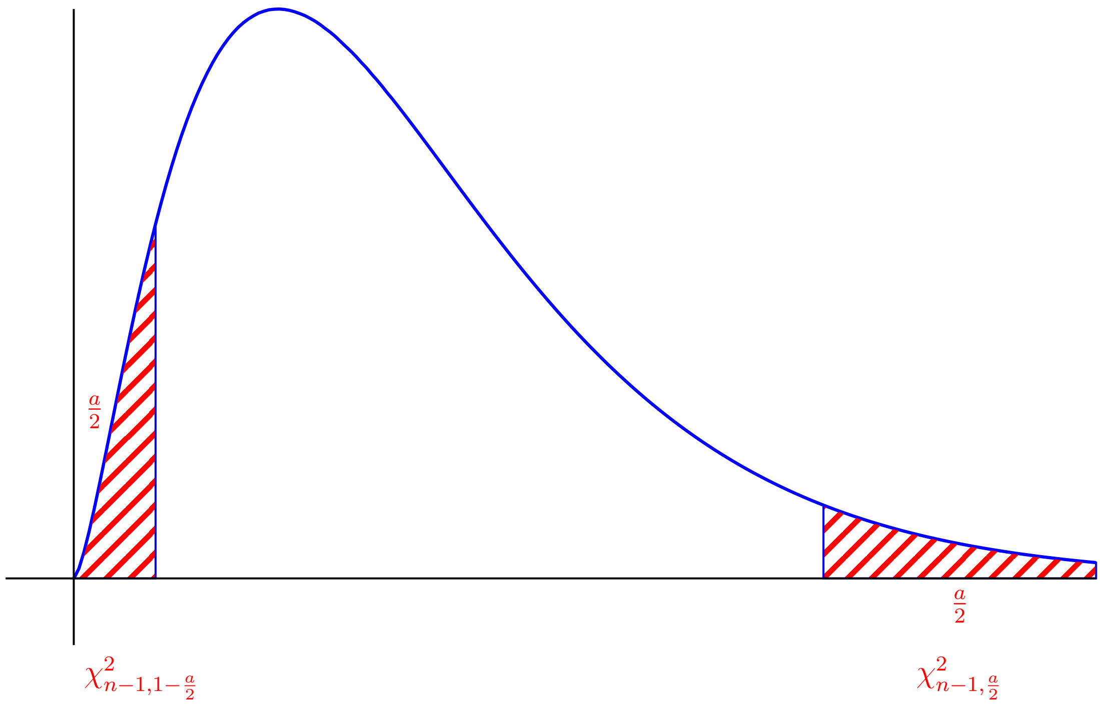


$$(1-\alpha) - \underset{\text{(ακριβή)}}{\text{Δ.Ε.}} \quad \overline{X} \pm t_{n-1,\ \frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}}$$

$$(1-\alpha) - \underset{\text{(προσεγγιστικά)}}{\text{Δ.Ε.}} \quad \overline{X} \pm Z_{\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}}$$

④

$$\boxed{\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}}$$



$$c_1 \leq \frac{(n-1)S^2}{\sigma^2} \leq c_2$$

⑤ plug-in για ε.μ.π. όχι 1-1

Έχουμε $\theta = (\theta_1, \theta_2, \ldots, \theta_d)$ και $\widehat{\theta}$ ε.μ.π.. Τότε,

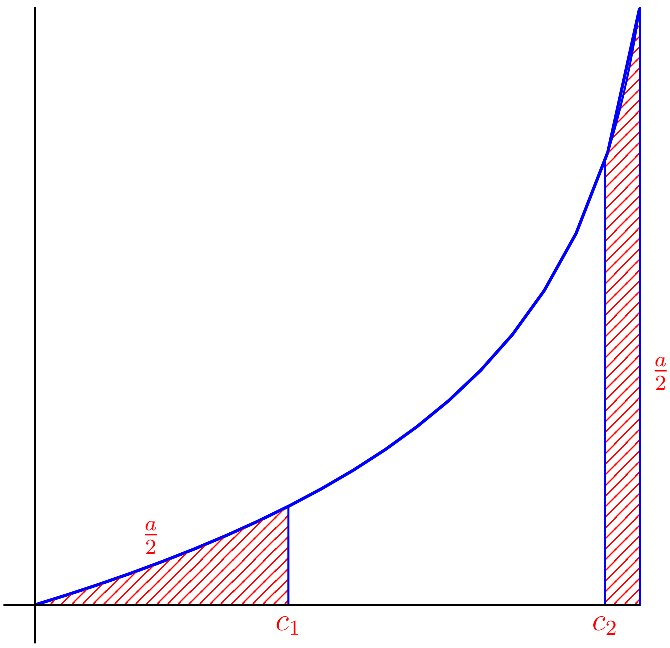
$$\varphi = g(\theta) \xrightarrow{plug-in} \widehat{\varphi} = g(\widehat{\theta})$$

όπου φ η καινούργια παράμετρος που θέλω να εκτιμήσω και $\widehat{\varphi}$ μία εκτιμήτρια του φ.

Αν $\widehat{\theta}$ είναι ε.μ.π. για οποιαδήποτε g, δεχόμαστε ότι είναι ε.μ.π. του φ

$\Rightarrow$ $\widehat{\theta}_i$ είναι ε.μ.π. του θ_i, $\forall\ 1 \leq i \leq d$.

⑥



$$U_{(n)} = \frac{X_{(n)}}{\theta}$$

$$F(c_1) = \frac{\alpha}{2}$$

$$F(c_2) = 1 - \frac{\alpha}{2}$$

⑦

$$\sqrt{n}\left(\overline{X}_n - p\right) \xrightarrow{d} \mathcal{N}(0, p(1-p))$$

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

με $\sigma(p) = \sqrt{p(1-p)}$ και $\widehat{\sigma(p)} = \sigma(\widehat{p})$.

Θέλω $\boxed{\sigma\left(\widehat{p}\right) \xrightarrow{p} \sigma(p)}$.

$$\sqrt{n}\frac{\overline{X}_n - p}{\sigma\left(\widehat{p}\right)} = \sqrt{n}\frac{\overline{X}_n - p}{\sigma(p)} \cdot \frac{\sigma(p)}{\sigma\left(\widehat{p}\right)},$$

όπου $\sqrt{n}\frac{\overline{X}_n - p}{\sigma(p)} \xrightarrow{d} \mathcal{N}(0,1)$ και $\frac{\sigma(p)}{\sigma\left(\widehat{p}\right)} \xrightarrow{p} 1$.

ΜΑΘΗΜΑΤΙΚΗ ΣΤΑΤΙΣΤΙΚΗ - QUIZ 2, 20 Νοέμβρη 2020

<u>ΕΡΩΤΗΣΕΙΣ:</u>

1. Έστω $t_{n,a}$ και z_a τα άνω a-ποσοστημόρια της t_n και της $N(0,1)$. Τί είναι σωστό ?
   ◯ $t_{n,a} < z_a$ ◯ $t_{n,a} = z_a$ ◯ $t_{n,a} > z_a$
2. Έστω $X \sim \mathcal{N}\big(0, \sigma^2(\theta)\big)$ με $\sigma^2(\theta) > 0$. Σε ποιά περίπτωση η εκτιμήτρια $\widehat{\theta}_n$ του θ είναι συνεπής και ασυμπτωτικά κανονική ?
   ◯ $\widehat{\theta}_n \xrightarrow[n\to\infty]{d} X \ (\forall\theta)$ ◯ $\sqrt{n}(\widehat{\theta}_n - \theta) \xrightarrow[n\to\infty]{d} X \ (\forall\theta)$ ◯ $\widehat{\theta}_n - \theta \xrightarrow[n\to\infty]{d} X \ (\forall\theta)$
3. Αν $r_n(\widehat{\theta}_n - \theta) \xrightarrow[n\to\infty]{d} \mathcal{N}\big(0, \sigma^2(\theta)\big)$, τότε η δέλτα μέθοδος με κατάλληλες υποθέσεις μας εξασφαλίζει ότι η ακολουθία $r_n\big(g(\widehat{\theta}_n) - g(\theta)\big)$ συγκλίνει κατά κατανομή στην:
   ◯ $\mathcal{N}\big(0, [g'(\theta)]^2\,\sigma^2(\theta)\big)$ ◯ $\mathcal{N}\big(0, g^2(\theta)\,\sigma^2(\theta)\big)$ ◯ $\mathcal{N}\big(0, g'(\theta)\,\sigma^2(\theta)\big)$
4. Μία εκτιμήτρια μέγιστης πιθανοφάνειας ικανοποιεί πάντα την ιδιότητα:
   ◯ της ύπαρξης ◯ της μοναδικότητας ◯ του αναλλοίωτου
5. Μία μοναδική ρίζα της $\ell'(\theta) = 0$, με την $\ell(\theta)$ ορισμένη σε ανοικτό διάστημα του $\mathbb{R}$, που είναι θέση τοπικού μεγίστου της $\ell(\theta)$, είναι επίσης και
   ◯ σημείο καμπής ◯ θέση ολικού μεγίστου ◯ εκτιμήτρια ροπών
6. Σε ένα τ.δ. από ομοιόμορφη U$[0,\theta]$, $\theta > 0$, η ε.μ.π. είναι:
   ◯ $2\overline{X}$ ◯ $X_{(n)}(n+1)/n$ ◯ $X_{(n)}$
7. Αν αυξηθεί ο συντελεστής εμπιστοσύνης ενός Δ.Ε., τότε γενικά το μήκος του:
   ◯ ελαττώνεται ◯ παραμένει σταθερό ◯ αυξάνεται
8. Αν αυξηθεί το μέγεθος του δείγματος, τότε γενικά το μήκος ενός Δ.Ε.:
   ◯ ελαττώνεται ◯ παραμένει σταθερό ◯ αυξάνεται
9. Σε ένα τυχαίο δείγμα από $\mathcal{N}(\mu, \sigma^2)$ με άγνωστα μ, σ^2 ένα 95%-Δ.Ε. για το μ είναι το
   ◯ $\overline{X} \pm z_{0.05}\, S/\sqrt{n}$ ◯ $\overline{X} \pm t_{0.025}\, S/\sqrt{n}$ ◯ $\overline{X} \pm t_{0.05}\, S/\sqrt{n}$
10. Σε τ.δ. από $\mathcal{N}(\mu, \sigma^2)$, άγνωστα μ, σ^2, το $\overline{X} \pm z_{0.05}\, S/\sqrt{n}$ έχει πιθανότητα κάλυψης
   ◯ > 0.9 ◯ > 0.95 ◯ < 0.9

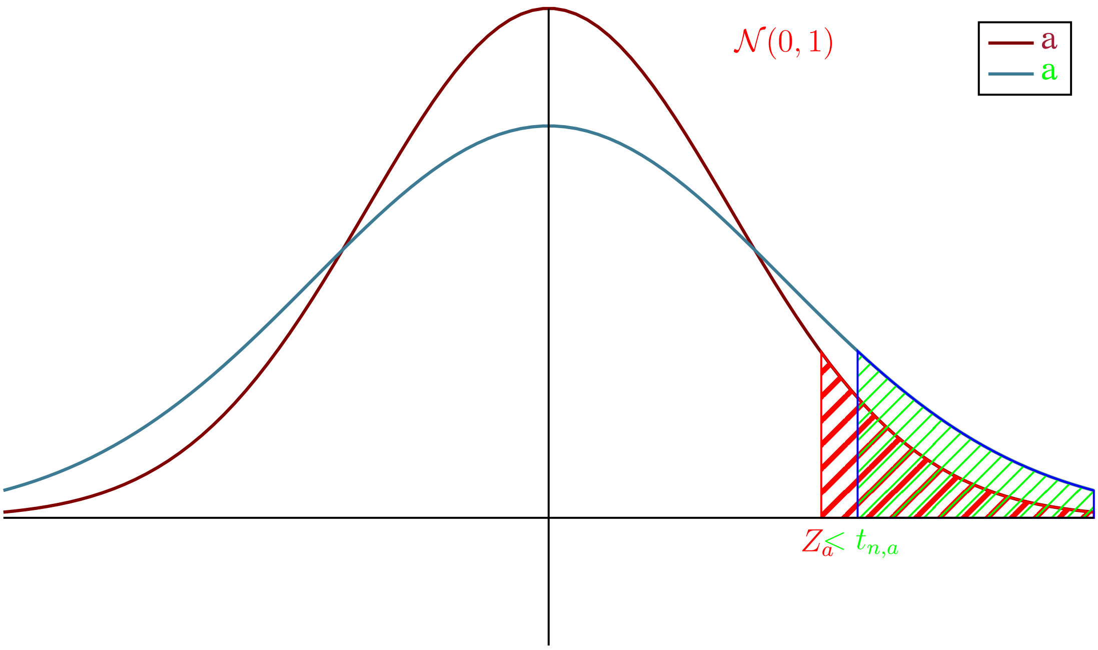
ΚΑΛΗ ΕΠΙΤΥΧΙΑ!

Απαντήσεις

1	Γ
2	Β
3	Α
4	Γ
5	Β
6	Γ
7	Γ
8	Α
9	Β
10	Γ

Σχολιασμός του Quiz 2

① $t_{n,\alpha} > Z_\alpha$



Η *Student* γενικά έχει πιο βαριές ουρές σε σχέση με την κανονική. Το γράφημα της καμπύλης κατεβαίνει είτε δεξιά είτε αριστερά πιο αργά στο 0 σε σχέση με το άλλο, με αποτέλεσμα να έχει μεγαλύτερη πιθανότητα να πάρει μεγαλύτερες τιμές κι αυτό έχει σαν αποτέλεσμα $t_{n,\alpha} > Z_\alpha$, για να πετύχουμε το ίδιο εμβαδόν.

②

- Γενικά για μια εκτιμήτρια θυμίζουμε ότι η συνέπεια εκφράζεται ως

$$\widehat{\theta}_n \xrightarrow{p} \theta, \ \forall \theta \in \Theta.$$

- Η ασυμπτωτική κανονικότητα και γενικότερα το να έχει κάποια ασυμπτωτική κατανομή το έχουμε διατυπώσει ως εξής:

$$r_n\left(\widehat{\theta}_n - \theta\right) \xrightarrow{d} X \text{ όπου } X \text{ μη εκφυλισμένη και } r_n \to +\infty.$$

  Στην περίπτωση της ασυμπτωτικής κανονικότητας έχουμε

$$\boxed{r_n\left(\widehat{\theta}_n - \theta\right) \xrightarrow{d} \mathcal{N}\left(0, \sigma^2(\theta)\right)}$$

Είχαμε πει ότι αυτό εξασφαλίζει και τη συνέπεια

$$\widehat{\theta}_n - \theta = \underbrace{\frac{1}{\sqrt{n}}}_{\downarrow\, 0} \underbrace{\sqrt{n}\left(\widehat{\theta}_n - \theta\right)}_{\downarrow d\; X} \xrightarrow{p} 0$$

$$\Rightarrow \widehat{\theta}_n \xrightarrow{p} \theta$$

(3) Εφαρμογή της Μεθόδου Δέλτα. Έχουμε ήδη μια ασυμπτωτική κανονικότητα της $\widehat{\theta}_n$, αφού από υπόθεση

$$r_n\left(\widehat{\theta}_n - \theta\right) \xrightarrow[n\to\infty]{d} \mathcal{N}\left(0, \sigma^2(\theta)\right)$$

και

$$r_n\left(\widehat{\theta}_n - \theta\right) \xrightarrow{d} \mathcal{N}\left(0, \sigma^2(\theta)\right)$$

$$\Rightarrow r_n\left(g\left(\widehat{\theta}_n\right) - g(\theta)\right) \xrightarrow{d} \mathcal{N}\left(0, \left[g'(\theta)\right]^2\sigma^2(\theta)\right).$$

Θα μπορούσαμε να το γράψουμε και ως εξής:

$$r_n\left(\widehat{\theta}_n - \theta\right) \xrightarrow{d} X$$

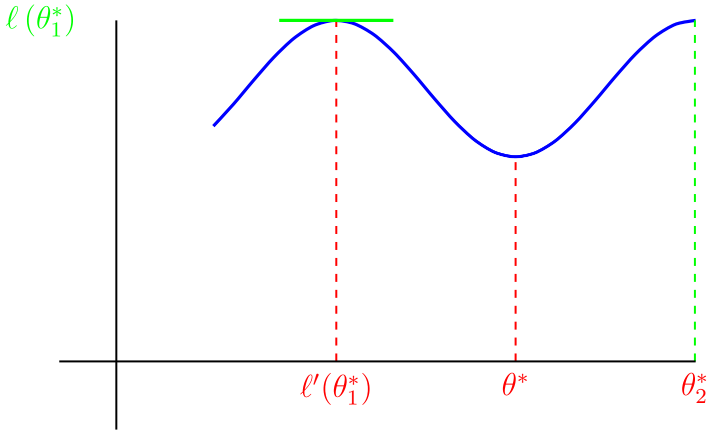
$$\Rightarrow r_n\left(g\left(\widehat{\theta}_n\right) - g(\theta)\right) \xrightarrow{d} g'(\theta)\cdot X.$$

(4) Μία ε.μ.π. έχει την ιδιότητα του αναλλοίωτου που σημαίνει ότι αν $\widehat{\theta}_n$ είναι ε.μ.π. του θ τότε $\widehat{g(\theta)} = g(\widehat{\theta})$ (plug-in).

Την ε.μ.π. την καθορίζουμε μέσα από την μεγιστοποίηση της πιθανοφάνειας. Άρα λοιπόν όταν βρεθούμε σε καταστάσεις που δεν είναι 1-1 θα πρέπει να επεκταθεί η έννοια της πιθανοφάνειας. Για το λόγο αυτό, υπάρχει η επαγόμενη συνάρτηση πιθανοφάνειας στην οποία κατοχυρώνεται και αυτό, δηλαδή να μπορούμε να πούμε αυτού του είδους την εκτιμήτρια του $g(\theta)$ ($\rightarrow$ άρθρο του Zehna που εξηγεί αυτή τη μέθοδο).

Έχουμε δει παραδείγματα περί μη ύπαρξης πάντα (να υπάρχει με θετική πιθανότητα, αλλά όχι με πιθανότητα 1) και μη μοναδικότητας πάντα (έχουμε δει περιπτώσεις όπου έχει άπειρες λύσεις όπως στην $\mathcal{U}[\theta, \theta+1]$, περίπτωση όπου έχει σε κάποιες περιπτώσεις μοναδική αλλά σε κάποιες άλλες έχει διπλή λύση, όπως ήταν η περίπτωση της Cauchy).

⑤ $\ell'(\theta) = 0 \to$ μοναδική ρίζα $\Rightarrow$ μοναδικό στάσιμο σημείο που από υπόθεση είναι και τοπικό μέγιστο $\Rightarrow$ ολικό μέγιστο.



Ας υποθέσουμε ότι δεν είναι ολικό μέγιστο, άρα $\exists\ \theta_2^* : \ell(\theta_1^*) = \ell(\theta_2^*)$.
Τότε από το Θεώρημα Rolle $\exists\ \theta^* \in (\theta_1^*, \theta_2^*) : \ell'(\theta^*) = 0$, άτοπο, γιατί έχουμε υποθέσει μοναδική ρίζα.

⑥ Σε τ.δ. $\mathcal{U}[0, \theta]$, έχουμε ότι το $X_{(n)}$ είναι ε.μ.π. του θ.

Θυμόμαστε ότι εφόσον όλες οι παρατηρήσεις είναι $\leq \theta$, έτσι και $X_{(n)} \leq \theta$.

⑦

$\nearrow$ το $1 - \alpha$, δηλ. ο συντελεστής εμπιστοσύνης $\Rightarrow$

$\nearrow$ το μήκος του Δ.Ε.

(Το να αυξηθεί ο σ.ε. σημαίνει ότι έχουμε μεγαλύτερη σιγουριά ότι έχουμε πιάσει την παράμετρο, και για να συμβεί αυτό θα πρέπει να αυξηθεί το μήκος του Δ.Ε.)

Αυτό αντίστοιχα σημαίνει:

$$\alpha \downarrow \Rightarrow \uparrow \text{μήκος}$$

$$\alpha \uparrow \Rightarrow \downarrow \text{μήκος}.$$

⑧ αύξηση του μεγέθους του δείγματος (για ίδιο σ.ε.) $\Rightarrow$ ελάττωση του μήκους Δ.Ε.

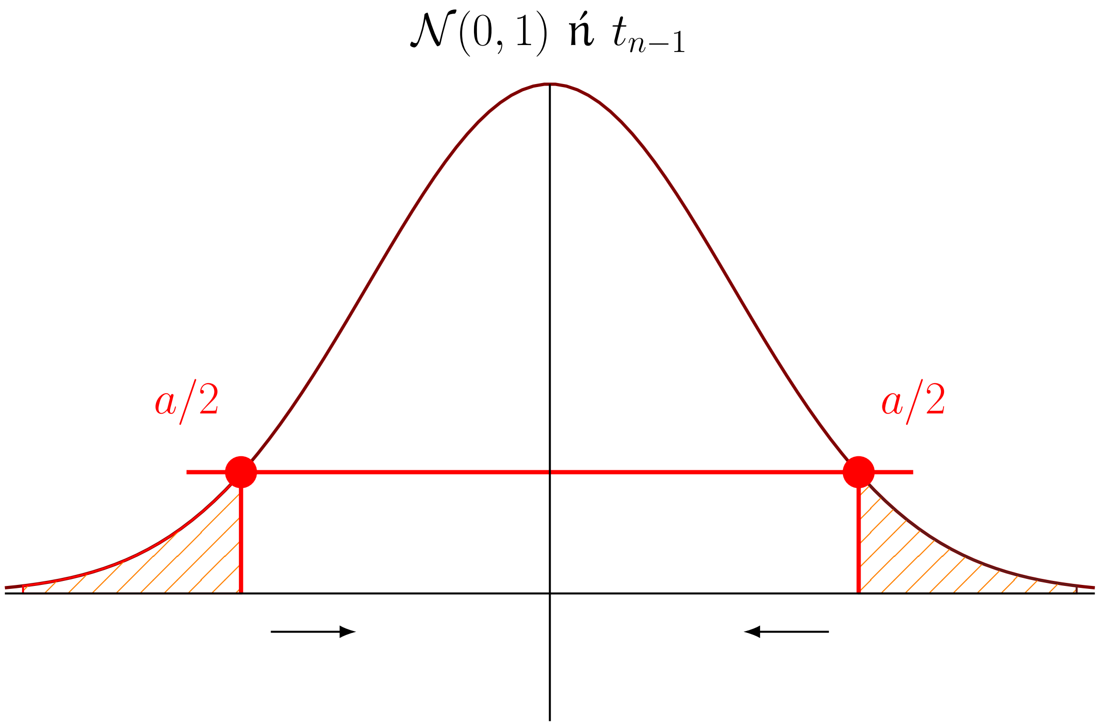
⑨ $\overline{X} \pm t_{0.025} \dfrac{S}{\sqrt{n}}$ (με σ^2 άγνωστο).

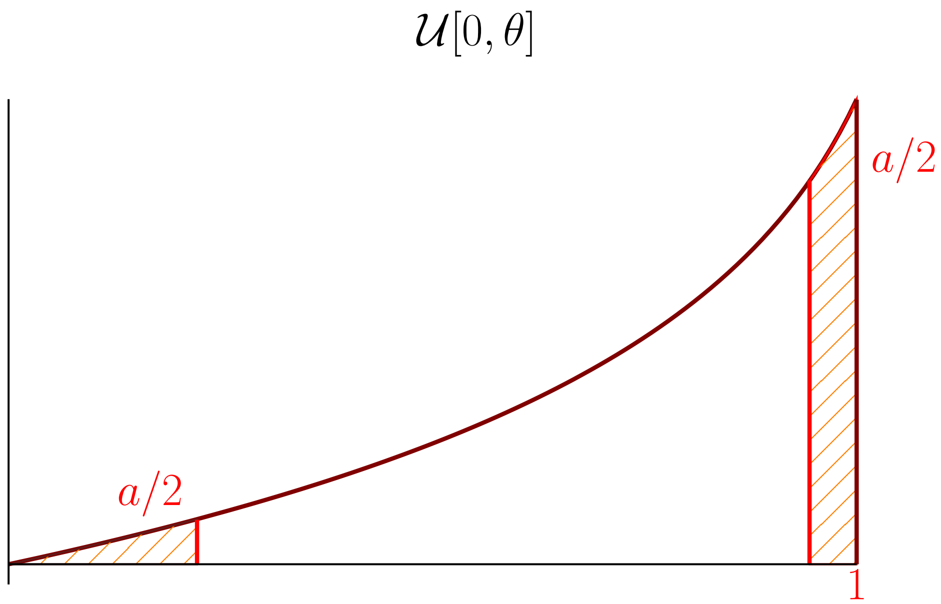
⑩ <0.90,

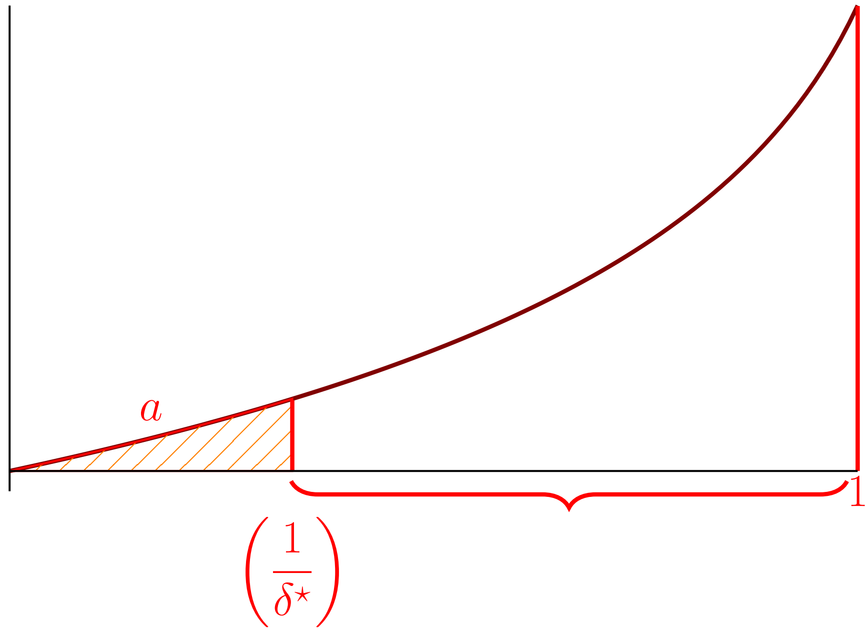
διότι το ακριβές είναι $\overline{X} \pm t_{0.05} \dfrac{S}{\sqrt{n}}$.

Άρα αφού $Z_{0.05} < t_{0.05} \Rightarrow$ μικρότερο μήκος έχει το $\overline{X} \pm Z_{0.05} \dfrac{S}{\sqrt{n}}$ από το ακριβές, άρα πιθ. κάλυψης <0.90.

Διάλεξη 18







Πρόβλημα:

Εύρεση βέλτιστου $(1-a)-$Δ.Ε. για τ.δ. από $\mathcal{U}[0,\theta]$, $\;\theta>0$, όπου το βέλτιστο νοείται ως προς την ελαχιστοποίηση του μήκους του Δ.Ε. (μεγαλύτερη ακρίβεια για δεδομένο σ.ε.).

Λύση

Θα αναζητήσουμε λύση της μορφής

$$I^{\star}_{1-a}(x)=\left[\frac{X_{(n)}}{\sqrt[n]{1-\frac{a}{2}}},\frac{X_{(n)}}{\sqrt[n]{\frac{a}{2}}}\right]=[\gamma_1 X_{(n)},\delta_1 X_{(n)}]$$

Θέλουμε να ελαχιστοποιήσουμε το $l\left(I^{\theta}_{1-a}(x)\right)\longrightarrow \min\limits_{\gamma,\delta}(\delta X_{(n)}-\gamma X_{(n)})=\min\limits_{\gamma,\delta}(\delta-\gamma)X_{(n)}$, όταν ικανοποιούνται οι περιορισμοί:

(i) $\mathbb{P}_{\theta}\left[\theta\in(\gamma X_{(n)},\delta X_{(n)})\right]=\mathbb{P}\left(\gamma X_{(n)}<\theta<\delta X_{(n)}\right)=1-a,\quad\forall\theta>0$

(ii) $\gamma,\delta\geq 1,\quad(\gamma<\delta)$

Αν υποθέσουμε ότι $\gamma<1$, τότε $\dfrac{1}{\gamma}>1$ και άρα

$$F_{U_{(n)}}\left(\frac{1}{\gamma}\right)=F_{U_{(n)}}(1)=1,\quad \text{διότι } 0\leq U_{(n)}\leq 1$$

άρα αν είχε επιλεγεί $\gamma<1$, τότε

$$\mathbb{P}\left(\theta\in[\gamma X_{(n)},\delta X_{(n)}]\right)=$$
$$\mathbb{P}\left(\theta\in[X_{(n)},\delta X_{(n)}]\right)=1-a$$

και συμπεραίνουμε ότι το τελευταίο, εφ'όσον έχει μικρότερο μήκος, είναι καλύτερο.

$$\left(\mathbb{P}\left(\gamma X_{(n)} \leq \theta\right) = \mathbb{P}\left(\underbrace{\frac{X_{(n)}}{\theta}}_{=U_{(n)}} \leq \underbrace{\frac{1}{\gamma}}_{>1}\right) = \mathbb{P}\left(\frac{X_{(n)}}{\theta} \leq 1\right)\right)$$

Άρα μπορούμε να περιοριστούμε σε $\gamma, \delta \geq 1$.
Επομένως

$$\mathbb{P}_\theta\left(\frac{1}{\delta} \leq \underbrace{\frac{X_{(n)}}{\theta}}_{=U_{(n)}} \leq \frac{1}{\gamma}\right) = 1 - a, \quad \forall \theta > 0$$

και από εδώ παίρνουμε

$$F_{U_{(n)}}\left(\frac{1}{\gamma}\right) - F_{U_{(n)}}\left(\frac{1}{\delta}\right) = 1 - a.$$

Πλέον αρκεί να ελαχιστοποιήσουμε το $(\delta - \gamma)$, όταν

(i) $\gamma^{-n} - \delta^{-n} = 1 - a$

(ii) $\gamma, \delta \geq 1$

Τελικά (δείτε σημ. 2017) $\gamma_\star = 1$, $\delta_\star = a^{-\frac{1}{n}} = \dfrac{1}{\sqrt[n]{a}}$.
Έτσι οδηγούμαστε

$$\boxed{I^\star_{1-a} = \left[X_{(n)}, a^{-\frac{1}{n}} X_{(n)}\right]}$$

$$\left[\begin{array}{cc} X_{(n)} \leq \theta \leq \delta_\star X_{(n)} & \Leftrightarrow \\ U_{(n)} \leq 1 \ \wedge \ U_{(n)} \geq \dfrac{1}{\delta_\star} & \Leftrightarrow \\ \dfrac{1}{\delta_\star} \leq U_{(n)} \leq 1 & \end{array}\right]$$

Ασυμπτωτικά Δ.Ε.

Ειδική περίπτωση προσεγγιστικών Δ.Ε. που κάποια ποσότητα $Q(X, \theta)$, αντί να είναι ποσότητα οδηγός για το πεπερασμένο n, είναι ασυμπτωτικά οδηγός (ή χρησιμοποιούμε για λόγους πρακτικούς την ασυμπτωτική του κατανομή).

Παράδειγμα 0.0.30. Έστω $X_1, X_2, \ldots, X_n$ τ.δ. $\mathcal{N}(\mu, \sigma^2)$ με μ, σ^2 άγνωστα και δείξαμε ότι

$$T_n = \sqrt{n}\frac{\overline{X}_n - \mu}{S} \sim t_{n-1}$$

και είναι ποσότητα οδηγός.
Ας το αγνοήσουμε!

Έχουμε δείξει ότι $T_n \xrightarrow[n\to\infty]{d} \mathcal{N}(0,1)$.
Άρα

$$\lim_{n\to\infty} \mathbb{P}\left(-Z_{\frac{a}{2}} \leq T_n \leq Z_{\frac{a}{2}}\right) =$$

$$\mathbb{P}\left(-Z_{\frac{a}{2}} \leq Z \leq Z_{\frac{a}{2}}\right) = 1 - a$$

$$\Rightarrow \quad \widetilde{I}^{\mu}_{1-a}(x) = \overline{X} \pm Z_{\frac{a}{2}}\frac{S}{\sqrt{n}} \longrightarrow \begin{array}{c}\text{ασυμπτωτικό}\\ (1-a)\text{−Δ.Ε.}\end{array}$$

$$\big\downarrow$$

αντικαταστ.
το $t_{n-1,\frac{a}{2}}$

$$T_n\sqrt{n}\frac{\overline{X}_n - \mu}{S_n} = \sqrt{n}\frac{\overline{X}_n - \mu}{\sigma} \cdot \frac{\sigma}{S_n} \xrightarrow[n\to\infty]{\text{Slutsky}\ d} \mathcal{N}(0,1)$$

$$\frac{\sigma}{S_n} \xrightarrow{p} 1$$

$$\sqrt{n}\frac{\overline{X}_n - \mu}{\sigma} \xrightarrow[d]{\text{Κ.Ο.Θ.}} \mathcal{N}(0,1)$$

Αν συγκρίνουμε τα 2 Δ.Ε.

$\boxed{\text{ασυμπτ. προσέγγ.}}$

$$\underbrace{\overline{X}_n \pm t_{n-1,\frac{a}{2}}}_{\text{ακριβές}}\frac{S}{n}\ ,\quad \underbrace{\overline{X}_n \pm Z_{\frac{a}{2}}}_{\text{προσεγγιστικό}}\frac{S}{n}$$

$$t_{n-1,\frac{a}{2}} > Z_{\frac{a}{2}} \quad \Rightarrow \quad \begin{array}{c}\text{το προσεγγιστικό}\\ \text{έχει μικρότερο μήκος}\\ \text{(έχει μικρότερο συντελεστή)}\end{array}$$

Άρα το $\overline{X}_n \pm Z_{\frac{a}{2}}\frac{S}{\sqrt{n}}$ είναι ανεκτικό $\left(\begin{array}{l}\text{permissive}\\ \text{anti-conservative}\end{array}\right)$ που σημαίνει ότι έχει πραγματικά μικρότερο συντελεστή από αυτό που θα έπρεπε. Σε αντίθετη περίπτωση λέγεται συντηρητικό ($> 1-a$) (conservative).
Γενικά προτιμάμε το ακριβές, εκτός αν έχει πολύ μεγάλο n [παλιά $n \geq 30 \to$ ασυμπτ. προσέγγιση].

Άσκηση

Έστω $X_1, \ldots, X_n$ τ.δ. $\mathrm{Exp}(\theta)$, $\theta > 0$ όπου θ είναι παράμετρος κλίμακας (θ : μέση τιμή).

(i) Θα βρούμε $(1-a)-$Δ.Ε. για το θ (ακριβές).

(ii) Θα βρούμε διαφορετικά ασυμπτωτικά $(1-a)-$Δ.Ε.

Λύση

(i) Έχουμε κατ'ευθείαν ότι ο $\overline{X}_n$ είναι μία εκτιμήτρια του θ.

$$\overline{X}_n = \frac{S_n}{n} \sim \frac{1}{n} G(n, \theta) \qquad \left(\begin{array}{l} \sum\limits_{i=1}^{n} X_i \\ \text{άθροισμα } n \text{ ανεξ. } \mathrm{Exp}(\theta) \end{array} \right)$$

Τότε

$$\begin{aligned} n\overline{X}_n \sim G(n,\theta) &= \theta G(n,1) \\ &= \frac{\theta}{2} G(n,2) \\ &= \frac{\theta}{2} G\left(\frac{2n}{2}, 2\right) \\ &= \frac{\theta}{2} \chi^2_{2n} \end{aligned}$$

$\chi^2_n \equiv G\left(\frac{n}{2}, \overset{\overset{\theta}{\uparrow}}{2}\right) \longrightarrow$ Όταν το θ είναι παράμετρος κλίμακας.

Τελικά έχουμε $\dfrac{2S_n}{\theta}$ είναι ποσότητα οδηγός.

Προσοχή! 2 συμβολισμοί (τύποι) *Gamma*.

$G(a,\theta)$	$G(a,\theta)$
$\parallel$	$\parallel$
$\frac{1}{\theta} G(a,1)$	$\theta G(a,1)$
Ειδικά για $a = n$	$\parallel$
$\frac{1}{\theta} G(n,1) =$	$\frac{\theta}{2} G(a,2)$
$\frac{1}{2\theta} G\left(\frac{2n}{2}, \frac{1}{2}\right) =$	$\downarrow a = n$
$\frac{1}{2\theta} \chi^2_{2n}$	$\frac{\theta}{2} G(n,2) \equiv \frac{\theta}{2} \chi^2_{2n}$

Έχουμε $\dfrac{2S_n}{\theta} \sim \chi^2_{2n}$ και άρα

$$c_1=\chi^2_{2n,1-\frac{a}{2}}\ ,\ c_2=\chi^2_{2n,\frac{a}{2}}\Rightarrow I^{\theta}_{1-a}(x)=\left[\frac{2S_n}{\chi^2_{2n,\frac{a}{2}}}\ ,\ \frac{2S_n}{\chi^2_{2n,1-\frac{a}{2}}}\right]=\underbrace{\boxed{\left[\frac{2\overline{X}_n}{\left(\frac{\chi^2_{2n,\frac{a}{2}}}{n}\right)}\ ,\ \frac{2\overline{X}}{\left(\frac{\chi^2_{2n,1-\frac{a}{2}}}{n}\right)}\right]}}_{(1-a)-\text{ακριβές}}$$

(ii) (α′ τρόπος)

Αναζητώ ασυμπτωτική κατανομή της εκτιμήτριας.

$$\sqrt{n}\left(\overline{X}_n-\underbrace{\theta}_{\mathbb{E}_\theta(X_1)}\right)\xrightarrow[n\to\infty]{d}\mathcal{N}\left(0,\underbrace{\sigma^2(\theta)}_{\mathbb{V}_\theta(X_1)}\right)\quad \text{(από το Κ.Ο.Θ.)}$$

δηλαδή

$$\sqrt{n}\left(\overline{X}_n-\theta\right)\xrightarrow{d}\mathcal{N}\left(0,\underbrace{\theta^2}_{\boxed{\text{παράμετρος κλίμακας}}}\right)\Rightarrow\sqrt{n}\frac{\overline{X}_n-\theta}{\theta}\mathcal{N}(0,1)$$

Ασυμπτωτικός οδηγός

$$Z=\sqrt{n}\frac{\overline{X}_n-\theta}{\underbrace{\widehat{\theta}}_{(=\overline{X}_n)}}\xrightarrow{d}\mathcal{N}(0,1)\qquad \widehat{\sigma(\theta)}=\sigma(\widehat{\theta})\ \ [\text{τεχνική } \sigma(\theta)\to\sigma(\widehat{\theta})]$$

Πράγματι

$$\sqrt{n}\frac{\widehat{\theta}_n-\theta}{\widehat{\theta}_n}=\overbrace{\sqrt{n}\frac{\widehat{\theta}_n-\theta}{\theta}}^{\xrightarrow{d}\mathcal{N}(0,1)}\cdot\overbrace{\frac{\theta}{\widehat{\theta}_n}}^{\xrightarrow{p}1}\xrightarrow[d]{\text{(Θ.Σ.Α.) Slutsky}}\mathcal{N}(0,1)$$

Άρα $\widetilde{I}^1_{1-a}(x)=\underbrace{\widehat{\theta}_n}_{\boxed{\text{εκτιμήτρια}}}\pm\underbrace{Z_{\frac{a}{2}}}_{\boxed{\substack{\text{ποσοστ.}\\ \text{ασυμπτ.}\\ \text{καταν.}}}}\cdot\underbrace{\frac{\widehat{\theta}_n}{\sqrt{n}}}_{\theta=\sigma_{X_1},\ \ \widehat{\sigma}_{\widehat{\theta}_n}=\frac{\widehat{\sigma}_{X_1}}{\sqrt{n}}}=\widehat{\theta}_n\left(1\pm Z_{\frac{a}{2}}\frac{1}{\sqrt{n}}\right).$

Τελικά

$$\widetilde{I}^1_{1-a}(x) = \left[\left(1 - \frac{Z_{\frac{a}{2}}}{\sqrt{n}}\right)\overline{X}_n \, , \, \left(1 + \frac{Z_{\frac{a}{2}}}{\sqrt{n}}\right)\overline{X}_n\right] \to \text{συμμετρικό}$$

(β′ τρόπος) $\left(\begin{array}{c}\text{Variance stabilizing transformations}\\ \text{μετασχηματισμοί που}\\ \text{σταθεροποιούν τη διασπορά.}\end{array}\right)$

vspace3mm

$$\sqrt{n}\left(\widehat{\theta}_n - \theta\right) \xrightarrow{d} \mathcal{N}\left(0, \sigma^2(\theta)\right)$$

Αναζητούμε $\varphi = g(\theta)$ (αναπαραμέτρηση):

$$\sqrt{n}\left(\widehat{\varphi}_n - \varphi\right) = \sqrt{n}\left(g(\widehat{\theta}_n) - g(\theta)\right) \xrightarrow[n\to\infty]{d} \mathcal{N}(0, c)$$

$g:$ διαφορίσιμη

Άρα για να τη βρούμε λύνουμε

$$\left(g'(\theta)\right)^2 \sigma^2(\theta) = c \quad \text{ή } (c=1) \quad g'(\theta)\sigma(\theta) = 1 \quad \text{ή} \quad g'(\theta) = \frac{1}{\sigma(\theta)}$$

$$\Rightarrow g(\theta) = \int\limits_{\theta\in\Theta} \frac{1}{\sigma(\theta)} d\theta \qquad \text{(αγνοούμε τη σταθερά)}$$

Επιπλέον από τη σχέση

$$g'(\theta) = \frac{1}{\sigma(\theta)} > 0 \;\Rightarrow g'(\theta) > 0 \;\Rightarrow\; \text{η } g \text{ είναι } \uparrow .$$

Εδώ έχουμε $g'(\theta) = \dfrac{1}{\theta} \Rightarrow g(\theta) = \log(\theta), \;\; \theta > 0$

Άρα

$$\sqrt{n}\left(\log\left(\widehat{\theta}_n\right) - \log(\theta)\right) \xrightarrow[n\to\infty]{d} \mathcal{N}(0,1) \Rightarrow$$
$$\widetilde{I}^{2,\log\theta}_{1-a} = \left[\log\left(\widehat{\theta}_n\right) - \frac{Z_{\frac{a}{2}}}{\sqrt{n}} \, , \, \log\left(\widehat{\theta}_n\right) + \frac{Z_{\frac{a}{2}}}{\sqrt{n}}\right] \xRightarrow{\log\theta\uparrow}$$
$$\widetilde{I}^2_{1-a} = \left[e^{-\frac{Z_{a/2}}{\sqrt{n}}}\widehat{\theta}_n \, , \, e^{\frac{Z_{a/2}}{\sqrt{n}}}\widehat{\theta}_n\right]$$

$$\sqrt{n}\frac{\widehat{\theta}_n - \theta}{\theta} \xrightarrow{d} \mathcal{N}(0,1) \Rightarrow$$
$$\sqrt{n}\left(\frac{\widehat{\theta}_n}{\theta} - 1\right) \xrightarrow{d} \mathcal{N}(0,1)$$

Άρα

$$\mathbb{P}\left(-Z_{\frac{a}{2}} \leq \sqrt{n}\left(\frac{\widehat{\theta}_n}{\theta} - 1\right) \leq Z_{\frac{a}{2}}\right) \overset{\text{ασυμπτωτ.}}{=} 1 - a \Leftrightarrow$$

$$1 - \frac{Z_{\frac{a}{2}}}{\sqrt{n}} \leq \frac{\widehat{\theta}_n}{\theta} \leq 1 + \frac{Z_{\frac{a}{2}}}{\sqrt{n}} \Leftrightarrow$$

$$\widetilde{I}^3_{1-a} = \left[\frac{1}{1 + \frac{Z_{a/2}}{\sqrt{n}}}\widehat{\theta}_n \; , \; \frac{1}{1 - \frac{Z_{a/2}}{\sqrt{n}}}\widehat{\theta}_n\right]$$

Παρατήρηση 0.0.23. Το ακριβές $(1-a)$–Δ.Ε. εξασφαλίζει ακριβώς πιθανότητα $1-a$.
Τα άλλα θα έχουν κάποια διαφορά (δουλέψτε).
Όλα τα ασυμπτωτικά είναι της μορφής

$$\left[\gamma\overline{X}_n \; , \; \delta\overline{X}_n\right] \to \text{όπως και το ακριβές (τελικά συγκρίνουμε } \delta - \gamma\text{)}$$

Παρατήρηση 0.0.24.

$$\begin{aligned} e^{-x} &\cong 1 - x & \text{από} \quad & \widetilde{I}^2_{1-a} \longrightarrow \widetilde{I}^1_{1-a} \\ e^{-x} = \frac{1}{e^x} &\cong \frac{1}{1+x} & \text{από} \quad & \widetilde{I}^2_{1-a} \longrightarrow \widetilde{I}^3_{1-a} \end{aligned}$$

Διάλεξη 19

Παράδειγμα 0.0.31. Πραγματοποιούμε δημοσκόπηση για εκλογές σε δείγμα 400 ατόμων. Το αποτελέσματα δίνουν 212 προθέσεις ψήφου υπέρ του υποψήφιου A και 188 υπέρ του B (π.χ. 2ος γύρος δημοτικών εκλογών).

(i) Δώστε ένα 95%-Δ.Ε. για τις προθέσεις ψήφου υπέρ του A σε όλο το εκλογικό σώμα (υποθέτουμε αρκετά μεγάλο).

(ii) Αν το ποσοστό των ψήφων που πήρε από το συγκεκριμένο δείγμα θεωρήσουμε ότι παραμένει αμετάβλητο, ποιό είναι το ελάχιστο μέγεθος δείγματος έτσι ώστε με σ.ε. 0.95, το διάστημα να μην περιέχει την τιμή 0.5;

<u>Λύση</u>

(i) Στον πληθυσμό μπορούμε να θεωρήσουμε ότι p = ποσοστό των ατόμων που ψηφίζουν τον A (άγνωστο). Πρέπει να εκτιμηθεί. Επιλέγουμε τυχαία n-άτομα από τον πληθυσμό και

$$X_i \begin{cases} 1 & \text{, αν το } i \text{ θα ψήφιζε τον } A \\ 0 & \text{, διαφορετικά} \end{cases}$$

Μπορούμε να υποθέσουμε ότι $X_i \sim \text{Be}(p), \quad 1 \leq i \leq n$ και αποτελούν τυχαίο δείγμα.
Έχουμε δεί πως ο δειγματικός μέσος $\overline{X}_n$ είναι ε.μ.π. και ε.ρ. του p και θέλουμε $(1-a)$-Δ.Ε.
Δεν είναι απλό να βρεθεί ακριβές και φτιάχνουμε ασυμπτωτικό $(1-a)$-Δ.Ε.
Χρειαζόμαστε ασυμπτωτική κατανομή του $\overline{X}_n$.

$$\sqrt{n}\left(\overline{X}_n - \underset{\substack{\downarrow \\ \mathbb{E}(X_1)}}{p}\right) \xrightarrow{d} \mathcal{N}\left(0, \underset{\substack{\downarrow \\ p(1-p)}}{\mathbb{V}(X_1)}\right) \Rightarrow \sqrt{n}\frac{\overline{X}_n - p}{\sqrt{p(1-p)}} \xrightarrow{d} \mathcal{N}(0,1).$$

Από αυτά που είπαμε πριν

$$\sqrt{n}\frac{\overline{X}_n - p}{\sqrt{\overline{X}_n\left(1-\overline{X}_n\right)}} \xrightarrow{d} \mathcal{N}(0,1)$$

Θέτουμε $Z = \sqrt{n}\frac{\overline{X}_n - p}{\sqrt{\overline{X}_n\left(1-\overline{X}_n\right)}}$ και φτιάχνουμε στο ασυμπτωτικό $(1-a)$-Δ.Ε.

$$\boxed{\widetilde{I}^{p}_{1-a}(x) = \overline{X}_n \pm Z_{\frac{a}{2}}\frac{\sqrt{\overline{X}_n\left(1-\overline{X}_n\right)}}{\sqrt{n}}}$$

Για τα δεδομένα του προβλήματος

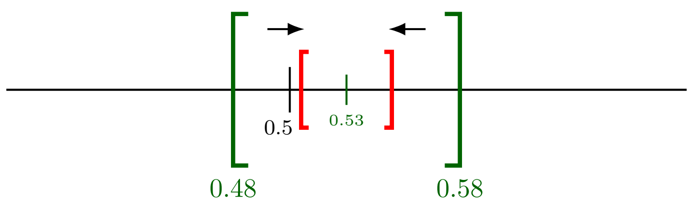
$$\widetilde{I}_{0.95} \cong [0.48, 0.58].$$

Άρα με σ.ε. 0.95 δεν μπορούμε να πούμε ότι ο A θα είναι ο νικητής, αφού βλέπουμε ότι μέσα στο Δ.Ε. έχουμε τιμές ≤ 0.5, $\boxed{\widehat{p}_{400} = 0.53}$.

(ii) Αν $\widehat{p}_n = 0.53$ (αμετάβλητο για το n). Όσο το n $\uparrow$, ελαττώνεται το μήκος του $(1-a)$-Δ.Ε. (αυξάνεται η ακρίβεια). Ποιο είναι το ελάχιστο n : $0.5 \notin \widetilde{I}_{1-a}$. Πρέπει να λύσουμε

$$\widehat{p}_n - Z_{\frac{a}{2}}\sqrt{\frac{\widehat{p}_n\left(1-\widehat{p}_n\right)}{n}} \;>\; 0.5$$

αριστερό άκρο



Άρα

$$0.53 - 1.96\sqrt{\frac{0.53\cdot 0.47}{n}} \;>\; 0.5 \Leftrightarrow 1.96\sqrt{\frac{0.53\cdot 0.47}{n}} \;<\; 0.03 \Leftrightarrow n \;>\; \frac{\left(\frac{0.03}{1.96}\right)^2}{0.53\cdot 0.47} \cong 1064$$

Έλεγχοι Υποθέσεων

Υποθέτουμε ότι μια μεγάλη φαρμακευτική εταιρεία έχει διεξάγει στατιστικές αναλύσεις και μπορεί να υποθέσει για το 2019 ότι η κατανομή των ημερήσιων κερδών της μπορεί να υποτεθεί $\mathcal{N}(10,1)$ (σε εκατομμύρια €). Το Γενάρη του 2020 υπάρχουν ενδείξεις για μείωση των κερδών και στις 21 Γενάρη ο Γενικός Διευθυντής καλεί συμβούλιο για να αποφασιστεί αν θα ξεκινήσει διαφημιστική καμπάνια για την προβολή της εταιρείας.
Αν τα δεδομένα του Γενάρη δεν δίνουν πολύ ισχυρές ενδείξεις ότι το μέσο κέρδος μ έχει μετακινηθεί σε $\mu < 10$, τότε ο Διευθυντής δεν θα προβεί σε πολυδάπανη διαφημιστική καμπάνια.

Σχόλια:

Υπάρχει μία αρχική πίστη για το μέσο κέρδος $\mu = 10$. Ελλείψει δεδομένων δεν θα γινόταν καμία κίνηση, αφού προφανώς δεν θα υπήρχε λόγος μετακίνησης (χωρίς ενδείξεις). Θα μπορούσε λόγω τυχαιότητας $\overline{X} < 10$,

$$\underset{\substack{\downarrow \\ \text{εκτίμηση του } \mu}}{\overline{X}} = \frac{X_1 + X_2 + \ldots + X_{30}}{30} \qquad \text{(δεδομένα του Γενάρη)}$$

Πρέπει λοιπόν να καθοριστεί ένα όριο:
αν $\overline{X} < c$, τότε να «μετακινηθούμε» από το $\mu = 10$ και να αποφασίσουμε ότι $\mu < 10$ (μείωση του μέσου κέρδους).
Λέμε:
$H_0: \quad \mu = 10 \longrightarrow$ μηδενική υπόθεση(:null hypothesis). Εκφράζει την αρχική μας πίστη ή αυτό που θα δεχόμασταν αν δεν υπήρχαν δεδομένα.
$H_1: \quad \mu < 10 \longrightarrow$ εναλλακτική υπόθεση(:alternative hypothesis)
Για να πάρουμε μία απόφαση αν θα απορρίψουμε την H_0 έναντι της H_1, διεξάγουμε έναν έλεγχο υποθέσεων με τη βοήθεια του τ.δ. $X_1, X_2, \ldots, X_{30}$ των κερδών στις 30 μέρες του Γενάρη χρειαζόμαστε μία κατάλληλη $T = T(x)$ για να αποφασίσουμε (ελεγχοσυνάρτηση ή κρίνουσα). Εδώ φυσιολογικά επιλέγουμε $T(x) = \overline{X}$ (δεν είναι απαραίτητο).
Ανάλογα με την τιμή $T(x) = t$ θα αποφασίσουμε αν απορρίπτεται η H_0 (έναντι της H_1) ή όχι (δηλαδή παραμένουμε στην H_0). Διαισθητικά εδώ αν $T(x)$ αρκετά μικρό, απορρίπτουμε, διαφορετικά όχι.

Κρίσιμη περιοχή $\equiv$ περιοχή απόρριψης της H_0
Κρίσιμη τιμή, για $T(x) < c_a$, απορρίπτουμε όπου c_a η κρίσιμη τιμή όπου a λέγεται ε.σ.σ. (επίπεδο στατιστικής σημαντικότητας).

Εδώ το $c_a: \quad (\mu_0 = 10)$

$$\mathbb{P}_{H_0}\left(\overline{X} < c_a\right) = \mathbb{P}_{\mu_0}\left(\overline{X} < c_a\right) = a$$

δηλαδή εδώ π.χ. για $a = 0.05$.

$$\mathbb{P}_{10}\big(\overline{X} < c_{0.05}\big) = 0.05$$

η H_0 είναι αληθής ε.σ.σ.

περιοχή απόρριψης της $H_0 \equiv$ κρίσιμη περιοχή του ελέγχου

Γενικότερα

$$\boxed{\mathbb{P}(\text{λανθασμένης απόρριψης της } H_0) = a}$$

Λέμε:

σφάλμα τύπου-I: λανθασμένη απόρριψη της H_0,
δηλαδή η H_0 είναι αληθής και εμείς την απορρίπτουμε.

Σχόλιο:

Ένα αποτέλεσμα λέμε ότι είναι στατιστικά σημαντικό όταν απορρίπτεται η H_0. Το ενδεχόμενο απόρριψης της H_0 είναι η αιτία διεξαγωγής του ελέγχου. Είναι η απόρριψη της H_0 που θα επιφέρει γενικά μεταβολές. Έτσι «κρατάμε» ότι στατιστικά σημαντικό είναι αυτό που θα αλλάζει την αρχική μας θέση (γενικά πιο επιφυλακτικά).

π.χ.

- H_0 : «Ο κατηγορούμενος είναι αθώος»
- H_1 : «Ο κατηγορούμενος είναι ένοχος»

Χωρίς στοιχεία το νομικό καθεστώς (σε δημοκρατία) προστατεύει την αθωότητα.

- H_0 : «Το υπάρχον φάρμακο είναι καλύτερο από κάποιο πειραματικό»
- H_1 : «Το αντίθετο»

Για να γίνει αντικατάσταση πρέπει να έχουμε ισχυρές ενδείξεις υπέρ του πειραματικού.

Εύρεση κρίσιμης περιοχής $\Leftrightarrow$ εύρεση κρίσιμης τιμής.

$$\overline{X} \overset{\mu_0}{\sim} \mathcal{N}\left(10, \frac{1}{n}\right) \Rightarrow Z = \frac{\overline{X} - 10}{\sqrt{\frac{1}{n}}} = \sqrt{n}\left(\overline{X} - 10\right) \sim \mathcal{N}(0,1).$$

Άρα

$$\mathbb{P}_{10}\left(\overline{X} < c_a\right) = \mathbb{P}_{10}\left(\sqrt{n}\left(\overline{X} - 10\right) < \sqrt{n}(c_a - 10)\right) = \Phi\left(\sqrt{n}(c_a - 10)\right).$$

Έχουμε

$$\begin{aligned} \Phi\left(\sqrt{n}(c_a - 10)\right) = a = \Phi(Z_{1-a}) &\Leftrightarrow \\ \sqrt{n}(c_a - 10) = Z_{1-a} &\Leftrightarrow \\ c_a = 10 + \frac{Z_{1-a}}{n} &= \qquad\qquad (Z_{1-a} = -Z_a) \\ 10 - \frac{Z_a}{\sqrt{n}} & \end{aligned}$$

Κρίσιμη περιοχή :$\longrightarrow$ περιοχή απόρριψης της H_0.

$$C_a = \left\{\overline{X} < 10 - \frac{Z_a}{\sqrt{n}}\right\}$$

Εναλλακτικά $\overline{X} \to Z$

$$\{Z < Z_{1-a}\} = \{Z < -Z_a\}.$$

Ορισμός 0.0.16. Έλεγχος Υποθέσεων (Hypothesis Testing)
Έστω H_0 μία μηδενική υπόθεση και H_1 μία εναλλακτική υπόθεση. Λέμε έλεγχο υποθέσεων της H_0 έναντι της H_1, τη διαδικασία κατά την οποία απορρίπτουμε ή όχι την H_0 έναντι της H_1, που καθορίζει το κριτήριο, βάσει του οποίου ο δ.χ. διαμερίζεται σε 2 περιοχές, στην περιοχή απόρριψης C της H_0, που λέγεται κρίσιμη περιοχή του ελέγχου και στην περιοχή αποδοχής της C^c.

Σχόλιο:
Είναι λοιπόν συνήθης η τακτική να προκαθορίζουμε ένα a, το οποίο λέγεται ε.σ.σ. του ελέγχου και θέτει ένα άνω φράγμα στην πιθανότητα σφάλματος τύπου-I, δηλάδή της λανθασμένης απόρριψης της H_0.

Ορισμός 0.0.17. Αν $H_0: \ \theta = \theta_0$ (συγκεκριμένη τιμή), δηλαδή έχουμε μονοσύνολο, τότε η υπόθεση αυτή λέγεται απλή υπόθεση και τότε η κατανομή των X_i είναι πλήρως καθορισμένη.
Αν $H_0: \ \theta \in \Theta_0 \subseteq \Theta$ και Θ_0 όχι μονοσύνολο, τότε η H_0 λέγεται σύνθετη υπόθεση. Ανάλογα και η εναλλακτική (χωρίζεται σε απλή και σύνθετη).

Καθορισμός της Κρίσιμης Περιοχής

Αν η H_0 είναι απλή $\theta = \theta_0$, τότε θέλουμε

$$\mathbb{P}_{\theta_0}(C) \leq \underset{\substack{\uparrow \\ \text{άνω φράγμα}}}{a}$$

Αν η H_0 είναι σύνθετη

$$\sup_{\theta \in \Theta_o} \mathbb{P}_\theta(C) \leq a$$

Όταν $\mathbb{P}_{\theta_0}(C) = a$, τότε το a λέγεται μέγεθος του ελέγχου ή αντίστοιχα

$$\underset{\substack{\downarrow \\ \equiv \text{ μέγεθος}}}{\sup_{\theta \in \Theta_0}} \mathbb{P}_\theta(C) = a$$

Λέμε:

σφάλμα τύπου II: Λανθασμένη αποδοχή της H_0, δηλαδή δεχόμαστε την H_0, ενώ ισχύει η H_1 (φυσιολογικά μπορούμε να υποπέσουμε και σε αυτό το σφάλμα).
Αν π.χ.

$$H_0: \quad \theta = \theta_0 \quad vs \quad H_1: \quad \theta = \theta_1$$

τότε σφάλμα τύπου II: δεχόμαστε ότι $\theta = \theta_0$, ενώ $\theta = \theta_1$.

$$\mathbb{P}(\text{σφάλμα τύπου II}) = \mathbb{P}_{\theta_1}(C^c) = 1 - \mathbb{P}_{\theta_1}(C) = 1 - \pi(\theta_1),$$

όπου το $\pi(\theta_1)$ λέγεται ισχύς του ελέγχου.

Γενικότερα, ορίζουμε ως συνάρτηση ισχύος του ελέγχου, τη συνάρτηση $\pi(\theta) = \mathbb{P}_\theta(C)$.

Για $H_1: \ \theta = \theta_1$, τότε ισχύς του ελέγχου είναι το $\pi(\theta_1)$ που λέγεται και

$$\begin{aligned} \pi = \mathbb{P}_{\theta_1}(C) &= \mathbb{P}_{\theta_1}(\text{αποδοχή της } H_1) \\ &= \mathbb{P}(\text{σωστή αποδοχή της} H_1) \end{aligned}$$

Μάλιστα λέγεται και διακριτική ικανότητα του ελέγχου και εκφράζει την πιθανότητα να «διακρίνουμε» σωστά ότι η H_1 είναι αληθής.

Περιληπτικά:

Η H_0 είναι / Απόφαση	Αληθής	Ψευδής (ισχύει η H_1)
Απόρριψη	Σφάλμα τύπου I (False positive)	Σωστή απόφαση (True positive)
Αποδοχή	Σωστή απόφαση (True negative)	Σφάλμα τύπου II (False negative)

Διάλεξη 20

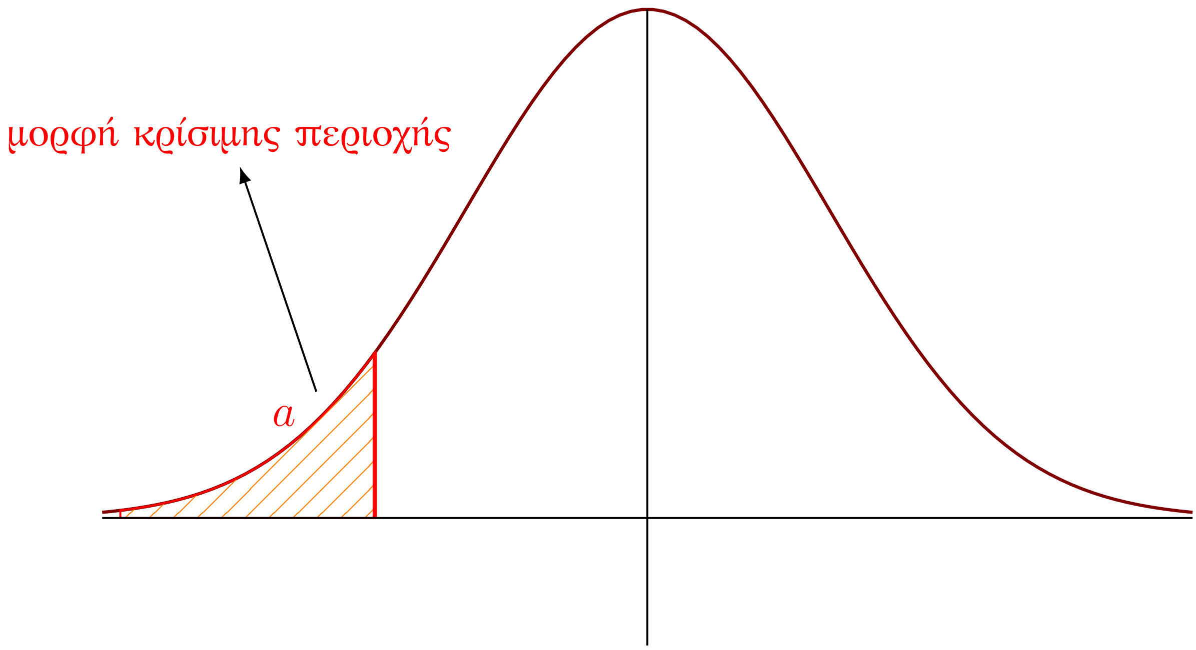
Θυμίζουμε το παράδειγμα με την μεγάλη πολυεθνική:

$$
\begin{array}{l}
H_0: \quad \mu = 10 \\
\text{vs} \\
H_1: \quad \mu < 10
\end{array}
$$

ελεγχοσυνάρτηση $\overline{X}$

κρίσιμη περιοχή $\mathbb{P}_{10}\left(\underbrace{\overline{X} < c_a}_{\text{κρίσιμη περιοχή}}\right) = \underbrace{a}_{\text{ε.σ.σ.}}$

$\mathbb{P}$(σφάλμα τύπου I)



Θα μπορούσαμε να έχουμε

$$\mu_1 < \mu_0, \quad H_1: \quad \mu = \mu_1$$

τότε δε θα άλλαζε τίποτα στη μορφή της κρίσιμης περιοχής. Θα παίρναμε την ίδια.

$$\overline{X} \longrightarrow Z = \sqrt{n}\frac{\overline{X} - 10}{1(\sigma = 1)}$$

(Δεν είναι μοναδικός ο τρόπος που επιλέγουμε ελεγχοσυναρτήσεις.)

$Z = \sqrt{n}\left(\overline{X} - 10\right) \qquad Z < z_a$

Θέσουμε $Z < z_a \Leftrightarrow \overline{X} < c_a$

Παρόμοιος έλεγχος υποθέσεων

$H_0: \quad \mu = \mu_0$
vs
$H_1: \quad \mu > \mu_0$

(αλλάζει η μορφή της κρίσιμης περιοχής)

π.χ.

πιστεύουμε ότι το μέσο κέρδος $\mu = 10$, αλλά αν υπάρχουν δεδομένα που μας κάνουν να πιστεύουμε ότι $\mu > 10$ τότε π.χ. θα μπορούμε να κινητοποιήσει το ενδιαφέρον να κάνουμε μία επένδυση.

Μορφή κρίσιμης περιοχής

$$\boxed{\overline{X} > c_a \Leftrightarrow Z > z_a}$$

$$\mathbb{P}_{\mu_0}(Z > z_a) = a$$

δείγμα

- $Z = z$
  - $z > z_a$ Απόφαση: Απόρριψη της H_0.
  - $z \leq z_a$ Απόφαση: Δέχομαι την H_0.

Παρόμοια

$H_0: \quad \mu = \mu_0$
vs
$H_1: \quad \mu = \mu_1$
$(\mu_1 > \mu_0)$

$\Rightarrow$ ίδια μορφή κρίσιμης περιοχής

- $H_0: \quad \mu = \mu_0$
  vs
  $H_1: \quad \mu \geq \mu_1$
  $(\mu_1 > \mu_0)$

  π.χ. $\mu_1 = 15$, τότε κάνω επένδυση πάλι έχουμε ίδια μορφή κρίσιμης περιοχής.

- $H_0: \quad \mu = \mu_0$
  vs
  $H_1: \quad \mu \neq \mu_0$

  Εδώ ελέγχουμε αν υπάρχει μεταβολή του μ.

(Η μεταβολή του μ κινητοποιεί π.χ. τη διεξαγωγή κάποιας διορθωτικής κίνησης ή κάποιας μεταβολής στο επενδυτικό πρόγραμμα της εταιρείας.)

π.χ. Σε ένα εργοστάσιο παραγωγής ενός εξαρτήματος με συγκεκριμένες προδιαγραφές δεδομένης διαμέτρου (προδιαγραφές).

π.χ.

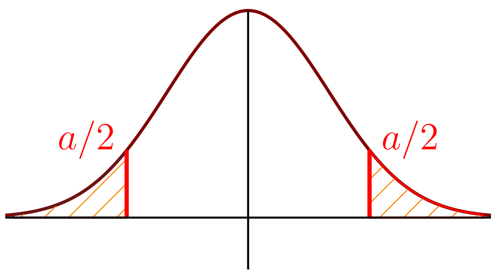
$$\begin{array}{l} H_0: \quad \mu = \mu_0 \\ \text{vs} \\ H_1: \quad \mu \neq \mu_0 \end{array}$$

κάτι χάλασε στην παραγωγική διαδικασία

Ελεγχοσυνάρτηση $\overline{X}$ εκτιμά το μ.

$$Z = \sqrt{n}\frac{\overline{X} - \mu_0}{\sigma}$$

όταν έχουμε σταθερό γνωστό σ



$$C = \left\{ |Z| > z_{\frac{a}{2}} \right\}, \quad \mathbb{P}(X \in C) = a$$

Έχουμε ισοδύναμα

$$\boxed{\left|\overline{X} - \mu_0\right| > c_a \Leftrightarrow |Z| > z_{\frac{a}{2}}}$$

$$
\begin{aligned}
Z = \sqrt{n}\frac{\overline{X}-\mu_0}{\sigma} \quad &\Rightarrow \left|\sqrt{n}\frac{\overline{X}-\mu_0}{\sigma}\right| > Z_{\frac{a}{2}} \Leftrightarrow \\
&\left|\overline{X}_n - \mu_0\right| > \frac{\sigma}{\sqrt{n}} Z_{\frac{a}{2}} \Leftrightarrow \\
\text{ή} \quad &\left.\begin{aligned} \overline{X}_n - \mu_0 &> \frac{\sigma}{\sqrt{n}} Z_{\frac{a}{2}} \\ \overline{X}_n - \mu_0 &< -\frac{\sigma}{\sqrt{n}} Z_{\frac{a}{2}} \end{aligned}\right\} \Leftrightarrow \\
\text{ή} \quad \circledast &\boxed{\begin{aligned} \overline{X}_n &> \mu_0 + Z_{\frac{a}{2}}\frac{\sigma}{\sqrt{n}} \\ \overline{X}_n &< \mu_0 - Z_{\frac{a}{2}}\frac{\sigma}{\sqrt{n}} \end{aligned}} \Leftrightarrow \\
&\overline{X}_n \notin \mu_0 \pm Z_{\frac{a}{2}}\frac{\sigma}{\sqrt{n}} \\
&c_a = Z_{\frac{a}{2}}\frac{\sigma}{\sqrt{n}}
\end{aligned}
$$

Στην περίπτωση που έχουμε δίπλευρη εναλλακτική (δύο ξένα διαστήματα απόρριψης εκατέρωθεν του μ_0).
Εναλλακτικά τις ανισότητες $\circledast$.

$$
\overset{\circledast}{\Leftrightarrow} \left.\begin{aligned} \mu_0 &< \overline{X}_n - Z_{\frac{a}{2}}\frac{\sigma}{\sqrt{n}} \\ \mu_0 &> \overline{X}_n + Z_{\frac{a}{2}}\frac{\sigma}{\sqrt{n}} \end{aligned}\right\} \Leftrightarrow
$$

$$
\mu_0 \notin \left[\overline{X}_n - Z_{\frac{a}{2}}\frac{\sigma}{\sqrt{n}}, \overline{X}_n + Z_{\frac{a}{2}}\frac{\sigma}{\sqrt{n}}\right]
$$

$$
\overline{X}_n \pm Z_{\frac{a}{2}}\frac{\sigma}{\sqrt{n}} \quad \longrightarrow (1-a) - \text{Δ.Ε. για το } \mu \text{ της } \mathcal{N}(\mu, \sigma^2), \quad \sigma^2 \text{ γνωστό}
$$

Συμπέρασμα

Δυϊκότητα ελέγχου υποθέσεων + διαστημάτων εμπιστοσύνης.

Για να ελέγξουμε την υπόθεση

$$H_0: \quad \mu = \mu_0 \quad \text{vs} \quad H_1: \quad \mu \neq \mu_0$$

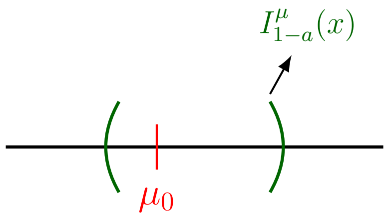
$$\Leftrightarrow$$

να κατασκευάσουμε το $(1-a)-$ Δ.Ε. για το μ.

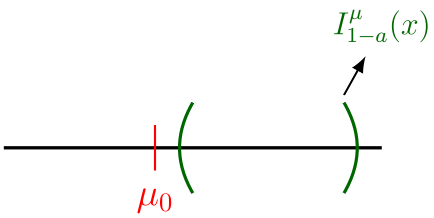
Πρόταση 0.0.11.

- Η H_0 απορρίπτεται σε ε.σ.σ. $a \quad \Leftrightarrow \quad \mu_0 \notin I^{\mu}_{1-a}(x)$.
- Η H_0 γίνεται δεκτή σε ε.σ.σ. $a \quad \Leftrightarrow \quad \mu_0 \notin I^{\mu}_{1-a}(x)$.

Δύο Σενάρια



$\Rightarrow$ δεν απορρίπτεται η H_0 vs H_1



$\Rightarrow$ απορρίπτεται η H_0 vs H_1 (στατιστικά σημαντικό αποτέλεσμα)

Γενικεύεται για $H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$ και μπορούμε να πραγματοποιήσουμε τον έλεγχο σε ε.σ.σ. a [για δίπλευρες εναλλακτικές] κοιτάζοντας μόνο το $(1-a)$–Δ.Ε.

π.χ.

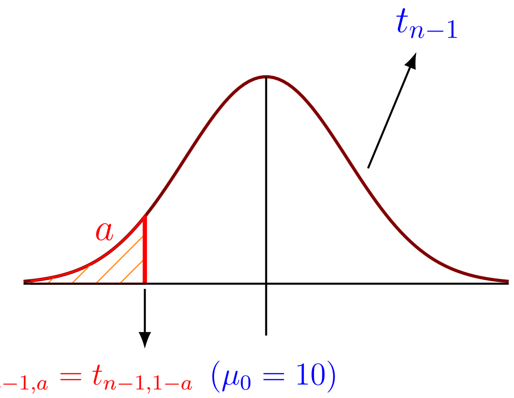
$H_0 : \quad \mu = \mu_0$
vs
$H_1 : \quad \mu < \mu_0$

(+ υποθέτουμε άγνωστη διασπορά.)
[$\rightarrow$ H_0 παύει να είναι απλή $\rightarrow$ σύνθετη]

$$\Theta_0 = \{(\mu_0, \sigma^2) : \sigma^2 > 0\}$$

Θέλουμε να αποφασίσουμε υπέρ της H_0 ή όχι.

Πριν $Z = \sqrt{n}\dfrac{\overline{X} - \mu_0}{\underset{\downarrow \text{ γνωστό}}{\sigma}}$, τώρα $T = \sqrt{n}\dfrac{\overline{X} - \mu_0}{S} \sim t_{n-1}$.



Μορφή της κρίσιμης περιοχής $\boxed{T < -t_{n-1,a}}$ τότε $\mathbb{P}_{\mu_0}(X \in C) = \mathbb{P}(T < -t_{n-1,a}) = a$.

$$\begin{aligned} T &< -t_{n-1,a} \Leftrightarrow \\ \sqrt{n}\frac{\overline{X} - \mu_0}{S} &< -t_{n-1,a} \Leftrightarrow \\ &\boxed{\overline{X} < \mu_0 - t_{n-1,a}\frac{S}{\sqrt{n}}}. \end{aligned}$$

Εντελώς ανάλογα

$$\begin{aligned} H_0 :& \quad \mu = \mu_0 \\ \text{vs}& \\ H_1 :& \quad \mu > \mu_0 \\ \text{ή}& \\ H_1 :& \quad \mu = \mu_1 \end{aligned}$$

Θα δώσουν κρίσιμη περιοχή

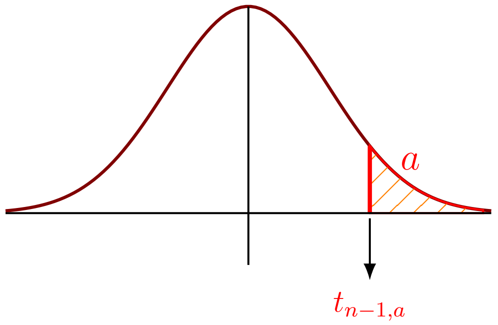
$$\boxed{T > t_{n-1,a} \Leftrightarrow \overline{X} > \mu_0 + t_{n-1,a}\frac{S}{\sqrt{n}}}$$

$\downarrow$

ισόνομες αναπαραστάσεις της κρίσιμης περιοχής C_a

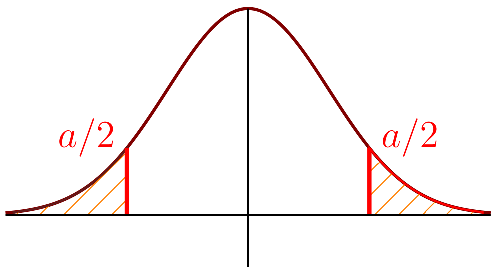
$\downarrow$

ε.σ.σ.



$$
\begin{array}{ll}
H_0: & \mu = \mu_0 \\
\text{vs} & \\
H_1: & \mu \neq \mu_0 \\
 & (\sigma^2 \text{ άγνωστο})
\end{array}
$$

$$\boxed{T = \sqrt{n}\frac{\overline{X} - \mu_0}{S}}$$



Κρίσιμη περιοχή C_a.

Ως προς $\overline{X}$. $\longleftarrow \boxed{|T| > t_{n-1,\frac{a}{2}}}$

$(T < -t_{n-1,\frac{a}{2}}$ ή $T > t_{n-1,\frac{a}{2}})$

Εναλλακτικός Τρόπος Διεξαγωγής Ελέγχου

Ας υποθέσουμε ότι από το συγκεκριμένο δείγμα πήραμε $\overline{X} = \overline{x}$ $(\overline{X}(\omega) = \overline{x})$. Τότε μπορούμε να υπολογίσουμε απευθείας:

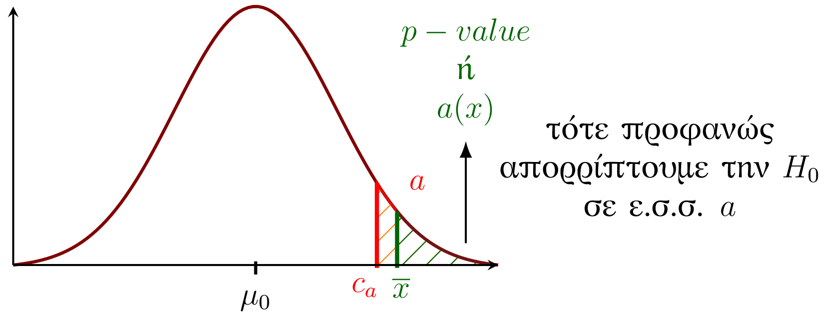
$$H_0: \ \mu = \mu_0 \quad \text{vs} \quad H_1: \ \mu > \mu_0 \longrightarrow \begin{array}{c}\text{κρίσ. περιοχής} \\ \overline{X} > C_a\end{array}$$

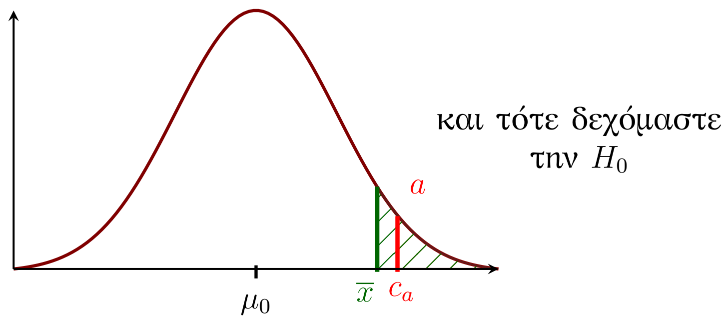
$$\boxed{\mathbb{P}_{\mu_0}\left(\overline{X} > \overline{x}\right) \equiv p - value}$$

και θα λέγεται παρατηρούμενο επίπεδο σημεντικότητας (observed significance level). Το γράφουμε και $a(x)$.

Έχουμε από κατασκευή του ελέγχου υποθέσεων ότι
$\mathbb{P}_{\mu_0}\left(\overline{X} > c_a\right) = a, \quad \overline{X} \sim \mathcal{N}\left(\mu_0, \frac{\sigma^2}{n}\right).$

- $\overline{x} > c_a \Leftrightarrow \mathbb{P}_{\mu_0}\left(\overline{X} > \overline{x}\right) < \mathbb{P}_{\mu_0}\left(\overline{X} > c_a\right) = a$





- $\overline{x} \leq c_a \Leftrightarrow \mathbb{P}_{\mu_0}\left(\overline{X} > \overline{x}\right) \geq a$

Άρα ισοδύναμα διεξάγεται ο έλεγχος υποθέσεων ως εξής:

- Αν $p - value < a$, τότε απορρίπτουμε την H_0.
- Αν $p - value \geq a$, τότε δεν απορρίπτουμε την H_0 (δηλ. την δεχόμαστε).

Μνημονικός κώδικας

Όταν χρησιμοποιούμε την $p - value$ να θυμόμαστε:

- Αν η $p - value$ είναι αρκετά μικρή, τότε είναι αρκετά «απίθανο» να ισχύει η H_0 (σε σχέση με το a).
- Αν η $p - value$ είναι μεγάλη, τότε είναι «πιο πιθανό» να ισχύει η H_0. Είναι ένα μέτρο που βοηθάει στη διεξαγωγή του ελέγχου και όχι $\mathbb{P}$(«να ισχύει η H_0»).

Διάλεξη 21

Ορισμός 0.0.18. Έστω ο έλεγχος $H_0 : \theta = \theta_0$ vs $H_1 : \theta = \theta_1, \ \theta_0 \neq \theta_1$ και C μία κρίσιμη περιοχή μεγέθους α. Η C λέγεται **ισχυρότατη κρίσιμη περιοχή** (Ι-κ.π.) αν για κάθε άλλη κ.π. D μεγέθους α, ισχύει

$$\mathbb{P}_{\theta_1}(C) \geq \mathbb{P}_{\theta_1}(D)$$

Παράδειγμα 0.0.32. Έστω ότι στην αγορά κυκλοφορούν δύο τύποι νομισμάτων, κανονικά & κίβδηλα. Σε μία ρίψη ενός νομίσματος, αν $p = \mathbb{P}(\text{“γράμματα”})$ τότε υποθέτουμε αν το νόμισμα είναι κανονικό τότε $p_0 = \frac{1}{2}$, διαφορετικά αν είναι κίβδηλο $p_1 = \frac{1}{4} < p_0$.

Θέλουμε να ελέγξουμε με 4 ανεξάρτητες ρίψεις του νομίσματος, αν ένα δεδομένο νόμισμα είναι κίβδηλο.

$$H_0 : p = p_0 = \frac{1}{2} \quad \text{vs} \quad H_1 : p = p_1 = \frac{1}{4} \text{ (κίβδηλο)} \quad (p_1 < p_0)$$

Ελεγχοσυνάρτηση: $\boxed{S = \sum_{i=1}^{n} X_i} \overset{(n=4)}{\sim} \text{Bin}(4, p), \ p \in \left\{\frac{1}{4}, \frac{1}{2}\right\}$
(Διαισθητικά απορρίπτουμε αν $S < C_\alpha$)

Πρόβλημα Καταρχάς πόσο να επιλέξουμε το α?

στοιχειώδη ενδεχόμενα: $\{S = 0\}, \{S = 1\}, \{S = 2\}, \{S = 3\}, \{S = 4\}$.
κρίσιμες περιοχές $\to$ θα πρέπει να ενώσουμε στοιχειώδη ενδεχόμενα.

Έχουμε αν $C_1 = \{0\}$ και $C_2 = \{4\}$ ότι

$$\mathbb{P}_{\frac{1}{2}}(x \in C_1) = \mathbb{P}_{\frac{1}{2}}(S = 0) = \mathbb{P}_{\frac{1}{2}}(S = 4) = \mathbb{P}_{\frac{1}{2}}(x \in C_2) = \frac{1}{16}$$

Θέτουμε $\alpha = \frac{1}{16}$ (για να μπορούμε να το πετύχουμε)

$$\left\{\underset{S=0}{\frac{1}{16}}, \quad \underset{S=1}{\frac{4}{16}}, \quad \underset{S=2}{\frac{6}{16}}, \quad \underset{S=3}{\frac{4}{16}}, \quad \underset{S=4}{\frac{1}{16}}\right\}$$

Μόνο δύο δυνατές κρίσιμες περιοχές με $\alpha = \frac{1}{16}$ (συμπίπτει με το μέγεθος).

Συγκρίνουμε την ισχύ (όταν η εναλλακτική είναι απλή)

$$\pi_1 = \mathbb{P}_{\frac{1}{4}}(S=0) = \left(\frac{3}{4}\right)^4 = \frac{81}{256} \cong 0,316$$

$$\pi_2 = \mathbb{P}_{\frac{1}{4}}(S=4) = \left(\frac{1}{4}\right)^4 = \frac{1}{256} \cong 0,004$$

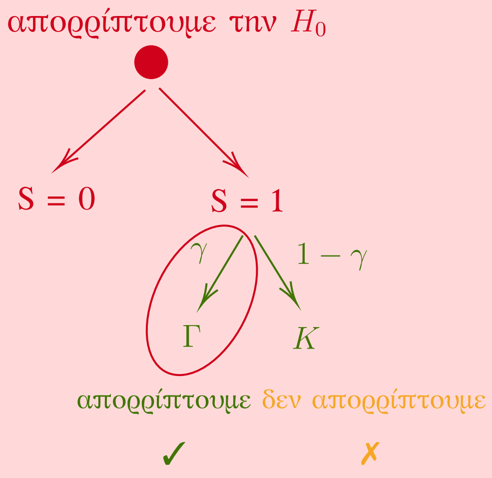
Συμπέρασμα το "λογικό" και το "παράλογο" έχει μαθηματικό αντίκτυπο και φαίνεται στην ισχύ του ελέγχου (το λογικό οδηγεί σε πολύ ισχυρότερο έλεγχο).

Εδώ η C_1 είναι η I-κ.π. μεγέθους $\frac{1}{16} = 0,0625$

$\rightarrow$ Τι θα κάναμε αν ζητούσαμε $\alpha = 0,1$?

Αν προσθέταμε το $\{S=1\}$ και είχαμε κρίσιμη περιοχή $S \in \{0,1\}$,
τότε $\alpha = \mathbb{P}_{\frac{1}{2}}(S \leq 1) = \frac{5}{16} = 0,3125 > 0,1$.

Σκέψη (να κάνουμε τυχαιοποίηση)
$\rightarrow$ Ρίχνουμε ένα νόμισμα δικό μας, που έχει συγκεκριμένη πιθανότητα γ να φέρει Γ (Γράμματα) και αν έρθουν Γ απορρίπτουμε την H_0, διαφορετικά τη δεχόμαστε.



$\rightarrow$ πώς επιλέγουμε το γ (όταν δεν κάνουμε τυχαιοποίηση)

$$R(X) = \begin{cases} 1 & , \ X \in C \\ 0 & , \ \text{διαφορετικά} \end{cases}$$

$\mathbb{E}_{\frac{1}{2}}\left(R(X)\right) = \mathbb{P}_{\frac{1}{2}}(X \in C) = \frac{1}{16}$ στο προηγούμενο.

Θέτουμε

$$R(X) = \begin{cases} 1 & , \ S(X) = 0 \\ \gamma & , \ S(X) = 1 \\ 0 & , \ 2 \leq S(X) \leq 4 \end{cases}$$

με $R(X) \to$ πιθανότητα απόρριψης.

$$\mathbb{E}_{\frac{1}{2}}\left[R(X)\right] = 1 \cdot \mathbb{P}_{\frac{1}{2}}(S = 0) + \gamma \cdot \mathbb{P}_{\frac{1}{2}}(S = 1)$$

και εμείς ζητάμε

$$\mathbb{E}_{\frac{1}{2}}\left[R(X)\right] = \underset{\substack{\downarrow \\ \textbf{ε.σ.σ.}}}{\alpha}$$

$$R = R(X) \quad (r(X))$$

$\hookrightarrow$ rejection (απόρριψη)

γράφεται και $[\varphi \ x]$

Στο συγκεκριμένο

$$\frac{1}{16} + \gamma \cdot \frac{4}{16} = \alpha \Rightarrow \boxed{\gamma = 0,15} \to 0,15 \text{ πιθανότητα να φέρουμε } \Gamma$$

Ομοιόμορφα Ισχυρότατοι Έλεγχοι

Παρουσιάζονται στοιχεία από μια γενική θεωρία κατασκευής ελέγχων υποθέσεων που δίνουν θεωρητική βάση στις διαισθήσεις μας.

Υποθέτουμε

$$H_0 : \theta \in \Theta_0 \quad \text{vs} \quad H_1 : \theta \in \Theta_1$$

όπου $\Theta_0 \cap \Theta_1 = \emptyset$ και (συνήθως) $\Theta = \Theta_0 \cup \Theta_1$ (αλλά όχι πάντα).

Ορισμός 0.0.19. Μία διεξαγωγή ελέγχου υποθέσεων της H_0 vs H_1, γίνεται μέσα από τον καθορισμό μιας σ.σ. $R = R(X)$ (ή $r(X)) \in [0,1]$. Αν $X = x$, είναι το δείγμα μετά την παρατήρησή του, τότε **απορρίπτουμε την** H_0 **με πιθανότητα** $R(x)$ και δεχόμαστε την H_0 με πιθανότητα $1 - R(x)$.

Αν

(i) $R(x) \in \{0,1\}$, τότε λέγεται **μη τυχαιοποιημένος έλεγχος** και

(ii) υπάρχει $x : 0 < R(x) < 1$ (με θετική πιθανότητα), τότε λέγεται **τυχαιοποιημένος έλεγχος**

- Για ένα δεδομένο έλεγχο ορίζουμε τη **συνάρτηση ισχύος** του $R(X)$, τη συνάρτηση

$$\pi_R(\theta) := \mathbb{E}_\theta(R), \quad \forall \theta \in \Theta.$$

Παρατήρηση 0.0.25.

① Για $R(X) \in \{0,1\}$ (μη τυχαιοποιημένος έλεγχος) έχουμε
$\pi_R(\theta) = \mathbb{P}_\theta(R=1) = \mathbb{P}_\theta(\text{απόρριψης της } H_0)$.

Άρα αν

- $\theta \in \Theta_0, \quad \pi_R(\theta) = \mathbb{P}_\theta(\text{σφάλμα τύπου I})$
- $\theta \in \Theta_1, \quad \pi_R(\theta) = \mathbb{P}_\theta(\text{σωστή αποδοχή της } H_1) = 1 - \mathbb{P}_\theta(\text{σφάλμα τύπου II})$.

② Με ένα δείγμα σταθερού μεγέθους, δεν μπορούμε να ελαχιστοποιήσουμε ταυτόχρονα τους 2 τύπους σφαλμάτων.

Η στρατηγική λοιπόν είναι να μεγιστοποιήσουμε την ισχύ του ελέγχου $\pi_R(\theta)$, για όλα τα $\theta \in \Theta_1$ και τα R που ικανοποιούν

$$\sup_{\theta \in \Theta_0} \pi_R(\theta) \leq \alpha,$$

όπου α είναι ένα ε.σ.σ.

Η ποσότητα $\sup\limits_{\theta \in \Theta_0} \pi_R(\theta)$ λέγεται **μέγεθος του ελέγχου**.

Ορισμός 0.0.20. Ένας έλεγχος R^* μεγέθους α, λέγεται **Ομοιόμορφα Ισχυρότατος Έλεγχος (Ο.Ι.Ε.)**, αν

$$\pi_{R^*}(\theta) \geq \pi_R(\theta), \quad \forall \theta \in \Theta, \quad \forall R \text{ με ε.σ.σ. } \alpha$$

Μορφή Ο.Ι.Ε. για H_0, H_1 απλές υποθέσεις

Θεώρημα 0.0.9. (Λήμμα Neyman-Pearson) Ας υποθέσουμε ότι $\Theta_0 = \{\theta_0\}$, $\Theta_1 = \{\theta_1\}$ και $L_x(\theta)$, $\theta \in \{\theta_1, \theta_2\}$ είναι η συνάρτηση πιθανοφάνειας τότε ισχύουν τα εξής

(i) (ύπαρξη) $\forall \alpha \in [0,1]$ **υπάρχει** ένας Ο.Ι.Ε. μεγέθους α της μορφής

$$R^*(X) = \begin{cases} 1 & , \text{αν } L_X(\theta_1) > c \cdot L_X(\theta_0) \\ \gamma & , \text{αν } L_X(\theta_1) = c \cdot L_X(\theta_0) \\ 0 & , \text{αν } L_X(\theta_1) < c \cdot L_X(\theta_0) \end{cases}$$

όπου $\gamma \in [0,1]$ και $c \geq 0$ κατάλληλα επιλεγμένες σταθερές τ.ω. $E_{\theta_0}[R^*(X)] = \alpha$ ($c = +\infty$, επιτρέπεται)

(ii) (μοναδικότητα) Αν R^{**} είναι ένας Ο.Ι.Ε. μεγέθους α, τότε

$$R^{**}(X) = \begin{cases} 1 & , \text{αν } L_X(\theta_1) > c \cdot L_X(\theta_0) \quad \text{(απορρίπτει)} \\ 0 & , \text{αν } L_X(\theta_1) < c \cdot L_X(\theta_0) \quad \text{(δέχεται)} \end{cases}$$

Συμπεραίνουμε ότι πάντα υπάρχει ένας Ο.Ι.Ε. για απλή μηδενική και εναλλακτική, που καθορίζεται μοναδικά εκτός ίσως από το σύνολο $\Gamma = \{x : L_X(\theta_1) = c \cdot L_X(\theta_0)\}$.

Αν το Γ είναι μηδενικής πιθανότητας τότε έχουμε μοναδικό μη τυχαιοποιημένο έλεγχο. Σε διαφορετική περίπτωση, οι Ο.Ι.Ε. είναι τυχαιοποιημένοι στο σύνολο Γ, για να πετύχουμε το επιθυμητό α, και μπορούν να επιλεγούν σταθεροί στο Γ.

Παρατήρηση 0.0.26.

1. Ένας Ο.Ι.Ε. απορρίπτει την H_0, όταν $\dfrac{L_X(\theta_1)}{L_X(\theta_0)} > c$ (για $c > 0$).

2. Αν $c \geq 1$, τότε το παραπάνω γράφεται ισοδύναμα

$$\frac{L_X(\theta_0)}{\max\{L_X(\theta_0), L_X(\theta_1)\}} < c'$$

όπου $\max\{L_X(\theta_0), L_X(\theta_1)\} = L\left(\widehat{\theta}\right) \quad (\theta = \{\theta_0, \theta_1\})$.

Αυτό θα δώσει αργότερα την ιδέα για επέκταση του λόγου πιθανοφάνειας για μη απλές υποθέσεις.

Διάλεξη 22

Είχαμε δει το λήμμα $N-P$. Μπορούμε να γράψουμε

$$\lambda(x) = \frac{L_x(\theta_1)}{L_x(\theta_0)} > c$$

όπου c οδηγεί στην κρίσιμη περιοχή.

Μπορούμε επίσης να χρησιμοποιήσουμε

$$\ell_x(\theta_1) - \ell_x(\theta_0) > c'$$

- Το Λήμμα $N-P$ εφαρμόζεται γενικότερα για τη σύγκριση 2 κατανομών που έχουν πυκνότητα ως προς κάποιο κυρίαρχο σ-πεπερασμένο μέτρο.

$$\xrightarrow[\text{ένεση θεωρίας μέτρου}]{} p_0, p_1 \ll \mu \begin{cases} \mu = \nu \text{ αριθμ. μέτρο} & f_0, f_1 \text{ συναρτ. πιθανότητας} \\ \mu = \lambda \text{ μέτρο Lebesgue} & f_0, f_1 \text{ συναρτ. πυκν. πιθανότητας} \end{cases}$$

$$H_0 : p = p_0 \quad \text{vs} \quad H_1 : p = p_1 \Rightarrow \boxed{H_0 : f = f_0 \quad \text{vs} \quad H_1 : f = f_1}$$

$\boxed{N-P}$ (απόδειξη σελ. 410 Shao)

(i) (ύπαρξη) $\forall \alpha \in [0,1]$, $\exists$ O.I.E. μεγέθους α

$$R^*(X) = \begin{cases} 1 & \text{, αν } f_1(X) > c \cdot f_0(X) \\ \gamma & \text{, αν } f_1(X) = c \cdot f_0(X) \\ 0 & \text{, αν } f_1(X) < c \cdot f_0(X) \end{cases}$$

όπου $0 \leq \gamma \leq 1$ και $c \geq 0$ (ή $+\infty$): $\underbrace{\boxed{\mathbb{E}_0 \left[R^*(X)\right] = \alpha}}_{\text{μέγεθος}}$

(ii) (μοναδικότητα) Αν R^{**} είναι ένας O.I.E. μεγέθους α, τότε

$$R^{**}(X) = \begin{cases} 1 & \text{, αν } f_1(X) > c \cdot f_0(X) \\ 0 & \text{, αν } f_1(X) < c \cdot f_0(X) \end{cases}$$

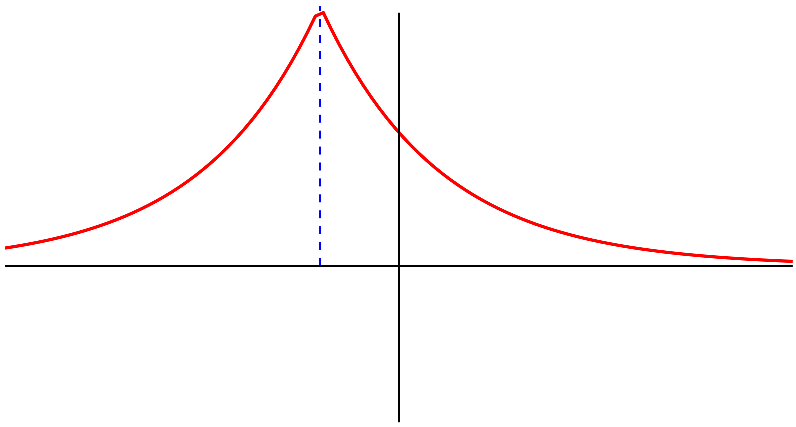
Παράδειγμα 0.0.33. Έστω X μία μόνο παρατήρηση. Ελέγχουμε

$$H_0 : X \sim \mathcal{N}(0,1) \quad \text{vs} \quad H_1 : X \sim \text{Laplace}(0,2)$$

όπου αν $X \sim \text{Laplace}(\mu, b)$ (λέγεται και διπλή εκθετική, double exponential) τότε

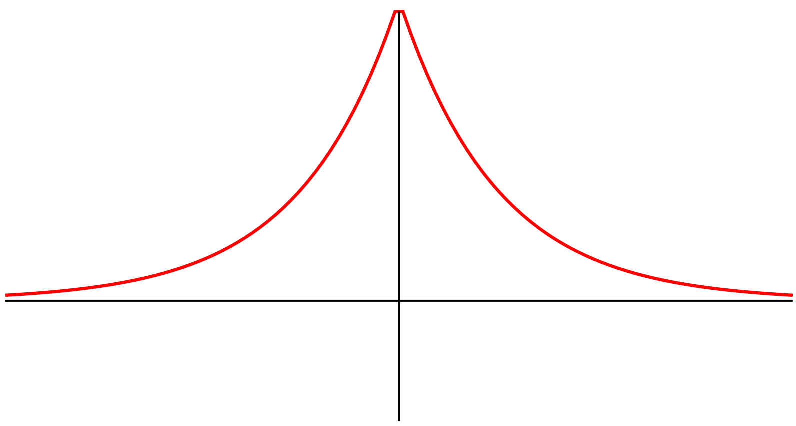
$$f_X(x) = \frac{1}{2b} \cdot e^{-\frac{|x-\mu|}{b}}, \quad x \in \mathbb{R}$$

διπλή εκθετική για $\mu = -2$ και $b = 3$



Για $\mu = 0,\ b = 2 \to f_X(x) = \dfrac{1}{4} \cdot e^{-\frac{|x|}{2}}, \quad x \in \mathbb{R}$

διπλή εκθετική για $\mu = 0$ και $b = 2$



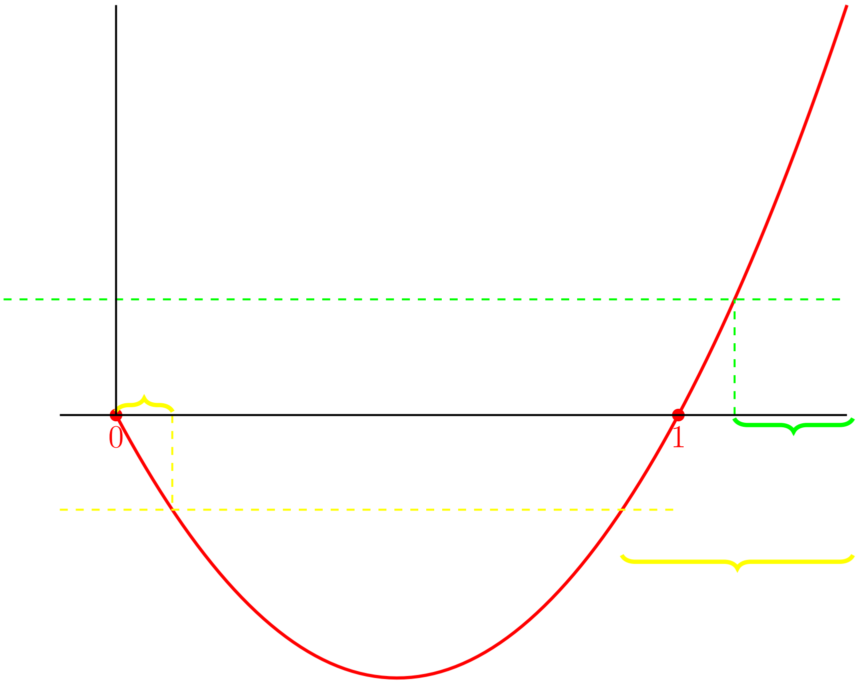
Είναι συνεχείς κατανομές $\Rightarrow$ $\exists$ μοναδικός μη τυχαιοποιημένος Ο.Ι.Ε.

Για να βρούμε την κρίσιμη περιοχή λύνουμε

$$\begin{aligned} f_1(x) > c \cdot f_0(x) &\Leftrightarrow \frac{1}{4} \cdot e^{-\frac{|x|}{2}} > c \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}} \\ &\Leftrightarrow e^{\frac{x^2-|x|}{2}} > c' \\ &\Leftrightarrow \boxed{x^2 - |x| > c''} \end{aligned}$$

Θέτουμε $y = |x|$. Τότε $y^2 - y > c'', \quad y \geq 0$.

$$y^2 - y > c'',\ \ y \geq 0$$



Αν $c'' < 0$,

$$\mathbb{P}_0(\text{απόρριψης της } H_0) > \underbrace{\mathbb{P}_0\left(|x| > 1\right)}_{\mathbb{P}_0(x<-1)+\mathbb{P}_0(x>1)} = 2\Phi(-1) = 2\left(1 - \Phi(1)\right) \cong 0,3374$$

$$(\text{υπενθ.: } \Phi(x) = 1 - \Phi(x))$$

Τα συνήθη ε.σ.σ. $\alpha \leq 0,01$.

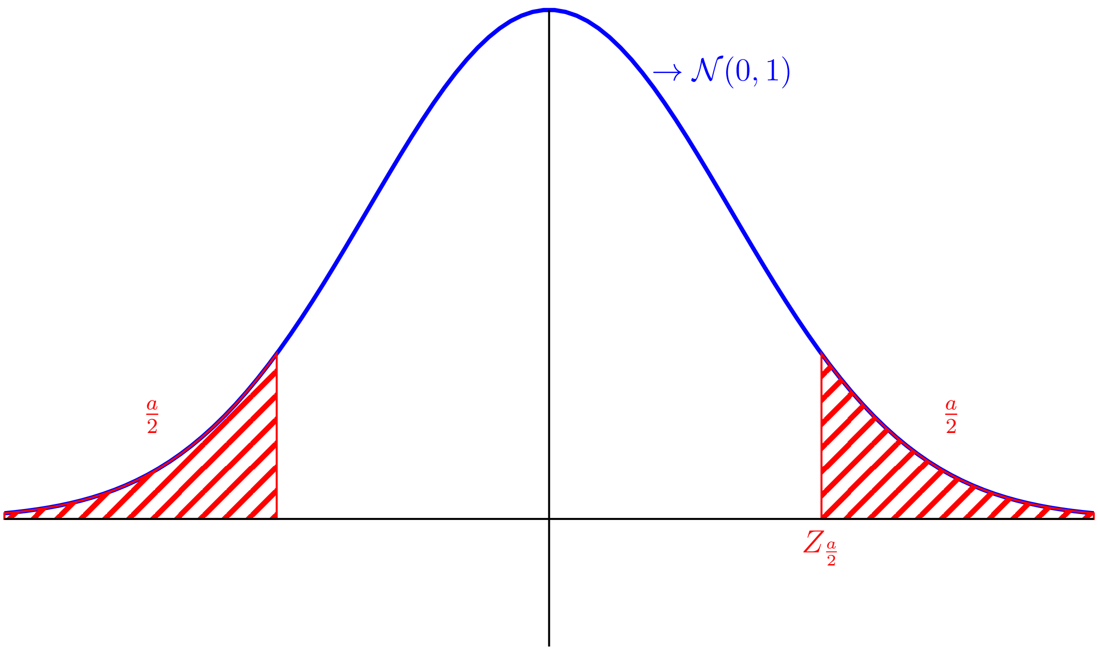
Το ενδιαφέρον κομμάτι είναι για $c'' > 0$.

Κρίσιμη περιοχή

$$\{|X| > c_a\} \quad (c_a > 1) \text{ όπου } X \underset{H_0}{\sim} \mathcal{N}(0,1).$$

Θέλουμε

$$\mathbb{P}_0\big(|X| > \underbrace{c_a}_{=Z_{\frac{a}{2}}}\big) = a$$



$c_a = \left(-\infty, -Z_{\frac{a}{2}}\right) \cup \left(Z_{\frac{a}{2}}, +\infty\right)$ και απορρίπτουμε αν $|X| > Z_{\frac{a}{2}}$.

- ισχύς του ελέγχου? $H_1 \to X \sim \text{Laplace}(0,2)$.

$$\pi_1 = \mathbb{E}_1\left(R^*(X)\right) = \mathbb{P}_1\left(|X| > Z_{\frac{a}{2}}\right) = 1 - \mathbb{P}\left(-Z_{\frac{a}{2}} \leq X \leq Z_{\frac{a}{2}}\right)$$

$$= 1 - \frac{1}{4}\int_{-Z_{\frac{a}{2}}}^{Z_{\frac{a}{2}}} e^{-\frac{|x|}{2}}\, dx = 1 - \frac{1}{2}\int_{0}^{Z_{\frac{a}{2}}} e^{-\frac{x}{2}}\, dx$$

$$= 1 - \left(1 - e^{-\frac{Z_{\frac{a}{2}}}{2}}\right) = \boxed{e^{-\frac{Z_{\frac{a}{2}}}{2}}}$$

Παράδειγμα 0.0.34. Έστω $X_1, \ldots, X_n$ τ.δ. από Be(p). Θέλουμε $H_0 : p = p_0$ vs $H_1 : p = p_1,\ p_1 < p_0$ (π.χ. εξαπάτηση σε τυχερό παιχνίδι). Να βρεθεί Ο.Ι.Ε. μεγέθους a.

Λύση

Από $N - P$ έχουμε

$$R^*(X) = \begin{cases} 1 & , \text{αν } \lambda(X) > c \\ \gamma & , \text{αν } \lambda(X) = c \\ 0 & , \text{αν } \lambda(X) < c \end{cases}$$

όπου $\lambda(X) = \dfrac{L_X(\theta_1)}{L_X(\theta_0)}$.

Παρατηρούμε ότι

$$\lambda(X) = \frac{L_X(p_1)}{L_X(p_0)} = \frac{p_1^{\sum_{i=1}^{n} X_i}}{p_0^{\sum_{i=1}^{n} X_i}} \cdot \frac{(1-p_1)^{n-\sum_{i=1}^{n} X_i}}{(1-p_0)^{n-\sum_{i=1}^{n} X_i}}$$

$$= \frac{\left(\frac{p_1}{1-p_1}\right)^{\sum_{i=1}^{n} X_i}}{\left(\frac{p_0}{1-p_0}\right)^{\sum_{i=1}^{n} X_i}} \cdot \left(\frac{1-p_1}{1-p_0}\right)^n$$

$$= \left(\frac{odds(p_1)}{odds(p_0)}\right)^{\sum_{i=1}^{n} X_i} \cdot \left(\frac{1-p_1}{1-p_0}\right)^n$$

Υπενθύμιση

$$\frac{p}{1-p} = odds(p)$$

$$\frac{\mathbb{P}(A)}{\mathbb{P}(A^c)} \rightarrow \text{ πόσο πιο πιθανό είναι να συμβεί από το να μη συμβεί}$$

* $odds(p)$ είναι $\nearrow$ στο (0,1).

$$\text{αν } p_1 < p_0 \Rightarrow odds(p_1) < odds(p_0) \Rightarrow \frac{odds(p_1)}{odds(p_0)} < 1$$

$$\Rightarrow \text{αν } S(X) = \sum_{i=1}^{n} X_i, \text{ τότε } \lambda(X) \text{ είναι φθίνουσα ως προς } S(X)$$

$$\Rightarrow \lambda(X) > c \Leftrightarrow \underbrace{S(X)}_{\sum_{i=1}^{n} X_i} < c'$$

Τελκά αφού $S(X) \in \{0, 1, 2, \ldots, n\}$ με $S(X) = \sum_{i=1}^{n} X_i \overset{p=p_0}{\sim} \text{Bin}(n, p_0)$

$$R^*(X) = \begin{cases} 1 & , \text{αν } S < m \\ \gamma & , \text{αν } S = m \\ 0 & , \text{αν } S > m, \end{cases}$$

όπου m, γ προσδιορίζονται

$$a = \mathbb{E}_0[R^*] = \mathbb{P}_0(S < m) + \gamma \cdot \mathbb{P}_0(S = m)$$

$\Leftrightarrow$ βρίσκουμε m, γ

$$F_0(m-1) \leq a < F_0(m) \quad \text{και} \quad \gamma = \frac{a - F_0(m-1)}{\mathbb{P}_0(S = m)}$$

όπου $\mathbb{P}_0(S = m) = F_0(m) - F_0(m-1)$.

Συνήθως ο R^* είναι τυχαιοποιημένος έλεγχος εκτός εάν $a_0 = F_0(m-1)$ (δηλ. συμπίπτει με τιμή της σ.κ. κάτω από την H_0).

Παρατήρηση 0.0.27. Στον παραπάνω έλεγχο είναι φανερό ότι ο Ο.Ι.Ε. δεν εξαρτάται από το p_1, $\forall p_1 < p_0$. Οδηγούμαστε έτσι σε Ο.Ι.Ε. για $H_0 : p = p_0$ vs $H_1 : p < p_0$.

Λήμμα 0.0.10. Αν υποθέσουμε ότι υπάρχει έλεγχος R^* μεγέθους a, έτσι ώστε $\forall \theta_1 \in \Theta_1$, ο R^* είναι Ο.Ι.Ε. για την $H_0 : \theta \in \Theta_0$ vs $H_1 : \theta = \theta_1$, τότε ο R^* είναι Ο.Ι.Ε. για την H_0 vs $H_1 : \theta \in \Theta_1$.

Απόδειξη

Έστω R έλεγχος με ε.σ.σ. a για την H_0 vs $H_1 : \theta \in \Theta_1$. Τότε αν $\theta_1 \in \Theta_1$, ο R είναι πάλι με ε.σ.σ. a για την H_0 vs $H_1 : \theta = \theta_1$ (δεν αλλάζει το H_0). Ο R^* είναι μεγέθους a & είναι Ο.Ι.Ε. H_0 vs $H_1 : \theta = \theta_1 \Rightarrow \pi_{R^*}(\theta) \geq \pi_R(\theta)$ $\quad$ (R^* Ο.Ι.Ε.).

Συμπερασματικά,

$$\left.\begin{array}{r} \forall R \text{ με ε.σ.σ. } a \\ \\ \forall \theta \in \Theta_1 \end{array}\right\} \Rightarrow \quad \text{είναι Ο.Ι.Ε. } H_0 \quad \text{vs} \quad H_1 : \theta \in \Theta_1$$

$\rightarrow$ θα γίνει επέκταση της έννοιας του Ο.Ι.Ε. και σε γενικότερες υποθέσεις σύνθετων υποθέσεων.

Ορισμός 0.0.21. Υποθέτουμε $\Theta \subset \mathbb{R}$ και έστω f_θ σ.π. / σ.π.π. Λέμε ότι η

$$\{f_\theta : \theta \in \Theta\}$$

έχει (την ιδιότητα του) Μ.Λ.Π. (μονότονο λόγο πιθανοφάνειας) στην $T = T(X)$ (μονοδιάστατο) αν $\forall \theta_1 < \theta_2 \Rightarrow \dfrac{f_{\theta_2}(x)}{f_{\theta_1}(x)}$ είναι αύξουσα συνάρτηση του $T(X)$ στο $\{x :$ τουλάχιστον ένα από τα $f_{\theta_2}(x)$ ή $f_{\theta_1}(x) > 0\}$ $\quad$ (επιτρέπουμε για $a > 0$, $\dfrac{a}{0} = +\infty$).

Θεώρημα 0.0.11. (Karlin-Rubin) (απόδειξη σελ. 66 Lehmann-Romano, σελ. 399 Shao)

Ας υποθέσουμε ότι το X έχει κατανομή που έχει Μ.Λ.Π. στην $T(X)$.
Αν $H_0 : \theta \leq \theta_0 \quad$ vs $\quad H_1 : \theta > \theta_0$, τότε

(i) υπάρχει Ο.Ι.Ε. μεγέθους a, που δίνεται από

$$R^*(X) = \begin{cases} 1 & , \text{αν } T(X) > c \\ \gamma & , \text{αν } T(X) = c \\ 0 & , \text{αν } T(X) < c \end{cases}$$

όπου $c, \ \gamma$ τ.ω $\pi_{R^*}(\theta_0) = a$.

(ii) $\pi_{R^*}(\theta) \nearrow \ \forall \theta : \ 0 < \pi_{R^*}(\theta) < 1$

(iii) $\forall \theta < \theta_0$, το R^* ελαχιστοποιεί την $\pi_R(\theta)$ (δηλ. την πιθανότητα σφάλματος τύπου I) για R με $\pi_R(\theta_0) = a$.

(iv) $\forall$ σταθερό θ_1, ο R^* είναι Ο.Ι.Ε. για τον έλεγχο

$$H_0 : \theta \leq \theta_1 \quad \text{vs} \quad H_1 : \theta > \theta_1$$

με μέγεθος $\pi_{R^*}(\theta_1)$

Παρατήρηση 0.0.28. Από το προηγούμενο Θεώρημα συμπεραίνουμε ότι μπορούμε να πάρουμε Ο.Ι.Ε. μεγέθους a για την $H_0 : \theta \geq \theta_0$ vs $H_1 : \theta < \theta_0$, αντιστρέφοντας τις ανισότητες στον R^*.

Παράδειγμα 0.0.35. Έστω $X_1, X_2, \ldots, X_n$ τ.δ. από $\mathcal{N}(\mu, \sigma^2)$ με μ άγνωστο και σ^2 γνωστό. Έστω $H_0 : \mu \leq \mu_0$ vs $H_1 : \mu > \mu_0$ (όπου μ_0 σταθερά $\in \mathbb{R}$).

Επαληθεύστε ότι

$$c_a = \left\{ x \in \mathbb{R}^n : \overline{X} > c_a \right\} = \left\{ x \in \mathbb{R}^n : Z\left(\overline{X}\right) > Z_a \right\}$$

όπου $Z\left(\overline{X}\right) = \sqrt{n} \cdot \dfrac{\overline{X} - \mu_0}{\sigma}$ τυποποίηση για $\mu = \mu_0$.

Λύση

Θ.δ.ο. το X έχει την ιδιότητα Μ.Λ.Π. για την $T(X)=\sum\limits_{i=1}^{n}X_i$. Έστω $\mu_1<\mu_2$. Τότε

$$\frac{L(\mu_2)}{L(\mu_1)}=\frac{(2\pi\sigma^2)^{-\frac{n}{2}}\,e^{-\frac{1}{2\sigma^2}\sum\limits_{i=1}^{n}(X_i-\mu_2)^2}}{(2\pi\sigma^2)^{-\frac{n}{2}}\,e^{-\frac{1}{2\sigma^2}\sum\limits_{i=1}^{n}(X_i-\mu_1)^2}}$$

$$=e^{\frac{\mu_2-\mu_1}{2\sigma^2}\left(2\sum\limits_{i=1}^{n}X_i-n(\mu_1+\mu_2)\right)}\quad \text{με } \mu_2>\mu_1$$

$$=e^{a\cdot\sum\limits_{i=1}^{n}X_i+\beta},\ a>0$$

$$\Rightarrow \frac{L(\mu_2)}{L(\mu_1)}\to\frac{f(X\ ;\ \mu_2)}{f(X\ ;\ \mu_1)} \text{ που είναι αύξουσα συνάρτηση του } T(X)$$

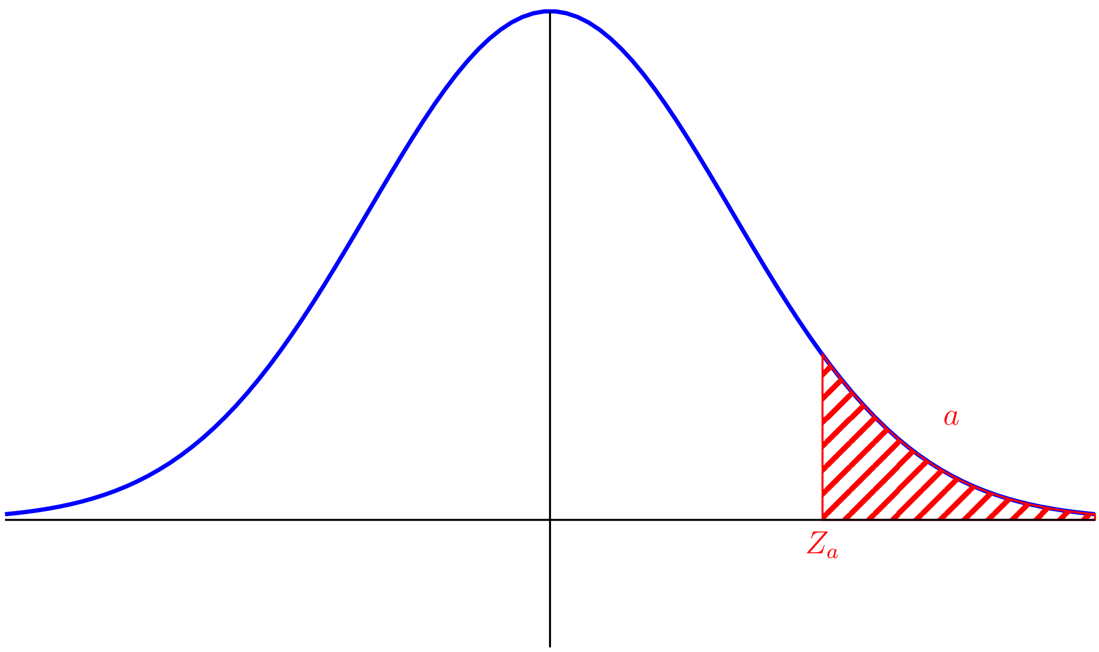
όπου $T(X)=\sum\limits_{i=1}^{n}X_i \Rightarrow$ έχει το Μ.Λ.Π. και άρα καταλήγουμε στον Ο.Ι.Ε. μέσω Karlin-Rubin.

Επιπλέον, είναι μη τυχαιοποιημένος ο Ο.Ι.Ε. και η κρίσιμη περιοχή $\sum\limits_{i=1}^{n}X_i>c$ και προσδιορίζουμε το c_a για $\mu\le\mu_0$ μέσω της

$$\mathbb{P}_{\mu_0}\left(\sum_{i=1}^{n}X_i>c_a\right)=\mathbb{P}_{\mu_0}\left(\overline{X}>c'_a\right)$$

$\hookrightarrow$ ή Z

και καταλήγουμε στο ότι $c_a=\{x\in\mathbb{R}^n : Z(X)>Z_a\}$.



Από το Z_a υπολογίζουμε το $c_a = \mu_0 + Z_a \cdot \dfrac{\sigma}{\sqrt{n}}$ το οποίο χρησιμοποιείται στο $\overline{X} > c_a$ ή αλλιώς χρησιμοποιούμε $\sum\limits_{i=1}^{n} X_i > \dots$

$$\left(\text{υπενθύμιση: } Z = \sqrt{n} \cdot \frac{\overline{X} - \mu_0}{\sigma}\right)$$

Ισοδύναμα, χρησιμοποιούμε κάποιο απο τα $\overline{X},\ \sum\limits_{i=1}^{n} X_i,\ Z$ εκ των οποίων το Z μας βοηθάει να τα υπολογίζουμε πιο εύκολα.

Ανάλογα, αν είχαμε $H_0 : \mu \geq \mu_0$ vs $H_1 : \mu < \mu_0$ θα παίρναμε Ο.Ι.Ε. μεγέθους a με τις ανισώσεις ανάποδα.

Διάλεξη 23

Μια κατηγορία κατανομών που έχουν την ιδιότητα του Μ.Λ.Π. προκύπτουν μέσα από εκθετικές οικογένειες κατανομών (Ε.Ο.Κ.).

Ορισμός 0.0.22. Έστω $x = (x_1, \ldots, x_n)$ ένα τυχαίο διάνυσμα με σ.π. / σ.π.π.

$$f_\theta(x), \;\; \theta \in \Theta \subseteq \mathbb{R}^S.$$

Η οικογένεια $\{f_\theta(x) : \theta \in \Theta\}$ λέμε ότι είναι / αποτελεί Ε.Ο.Κ.
(εκθετική οικογένεια κατανομών), αν

(i) $S_f = \{x \in \mathbb{R}^n : f_\theta(x) > 0\}$ είναι ανεξάρτητο του θ.

$$\text{(ii)} \quad f_\theta(x) = e^{\sum\limits_{\kappa=1}^{S} \varphi_\kappa(\theta)\cdot T_\kappa(x) - B(\theta)} \cdot h(x), \quad \forall x \in S_f$$

$$= \beta(\theta) \cdot e^{\sum\limits_{\kappa=1}^{S} \varphi_\kappa(\theta)\cdot T_\kappa(x)} \cdot h(x), \quad \forall x \in S_f$$

Παρατήρηση 0.0.29.

1. Όταν $\theta = (\theta_1, \ldots, \theta_S)$ με $S \geq 2$, τότε λέμε την προκύπτουσα Ε.Ο.Κ. S-παραμετρική, για $S = 2$ τη λέμε και διπαραμετρική, ενώ για $S = 1$ τη λέμε μονοπαραμετρική.
2. Όταν το $n > 1$, τότε λέγεται και n-διάστατη Ε.Ο.Κ. (πολυδιάστατη), ενώ για $n = 1$ λέγεται μονοδιάστατη.
3. Θα λέμε ότι το X (τυχαίο διάνυσμα) ή η X (τυχαία μεταβλητή) ανήκει σε Ε.Ο.Κ.(θ).

Κανονική μορφή μιας Ε.Ο.Κ.

Αν θέσουμε $\phi = \big(\underbrace{\varphi_1(\theta)}_{\phi_1}, \ldots, \underbrace{\varphi_S(\theta)}_{\phi_S}\big)$, τότε μια Ε.Ο.Κ. μπορεί να τεθεί με αναπαραμέτρηση στην **κανονική μορφή**

- $$f(x\ ;\ \phi) = e^{\sum\limits_{\kappa=1}^{S} \phi_\kappa \cdot T_\kappa(x) - A(\phi)} \cdot h(x), \quad \forall x \in S_f,\ \phi \in \Phi$$

  όπου $\phi = (\phi_1, \ldots, \phi_S)$ λέγεται **φυσική παράμετρος** και το Φ λέγεται **φυσικός παραμετρικός χώρος** και προσδιορίζεται από τη σχέση

  $$\Phi = \begin{cases} \phi \in \mathbb{R}^S : \displaystyle\int\limits_{S_f \subset \mathbb{R}^n} e^{\sum\limits_{\kappa=1}^{S} \phi_\kappa \cdot T_\kappa(x)} \cdot h(x)\ dx < +\infty, & \text{X συνεχής} \\ \\ \phi \in \mathbb{R}^S : \displaystyle\sum_{x \in S_f} e^{\sum\limits_{\kappa=1}^{S} \phi_\kappa \cdot T_\kappa(x)} \cdot h(x), & \text{X διακριτή} \end{cases}$$

- Ο χώρος αυτός περιλαμβάνει καταρχήν όλα τα $\varphi(\theta),\ \theta \in \Theta$ και επιπλέον μπορεί να είναι και μεγαλύτερος, δηλαδή

  $$\varphi(\Theta) \subset \Phi \Rightarrow \Theta \subset \varphi^{-1}(\Phi)$$

  όπου $\phi^{-1}(\Phi) = \widetilde{\Theta}$ και $\widetilde{\Theta}$ επέκταση του Θ.

- $$e^{A(\phi)} = \int\limits_{S_f} \underbrace{e^{\phi^t \cdot T(x)} \cdot h(x)}_{g(x\ ;\ \phi)}\ dx \Rightarrow \underset{\substack{\downarrow \\ f_\phi(x)}}{f(x\ ;\ \phi)} = e^{-A(\phi)} \cdot g(x\ ;\ \phi) = \frac{g(x\ ;\ \phi)}{e^{A(\phi)}}$$

  όπου $e^{A(\phi)} = c(\phi)$, δηλαδή είναι μια σταθερά κανονικοποίησης της g.

Παράδειγμα 0.0.36. Κατανομή Poisson$(\lambda),\ \lambda > 0$. **Θ.δ.ο.** η $X \sim \mathcal{P}(\lambda),\ \lambda > 0$ αποτελεί Ε.Ο.Κ.

Λύση

$$f_\lambda(x) = e^{-\lambda} \cdot \frac{\lambda^x}{x!}, \quad x = 0, 1, 2, \ldots$$

(i) $S_{f_\lambda} = \mathbb{N} \Rightarrow$ ανεξάρτητο του λ.

(ii) $f_\lambda(x) = e^{-\lambda} \cdot \lambda^x \cdot \dfrac{1}{x!}$, όπου

$$e^{-\lambda} = \beta(\lambda), \quad \lambda^x = e^{(\log\lambda)\cdot x} = e^{\varphi(\lambda)\cdot T(x)}, \quad \frac{1}{x!} = h(x)$$

με $\varphi(\lambda) = \log\lambda,\ T(x) = x$. Πράγματι αποτελεί Ε.Ο.Κ.

- Να τεθεί και σε κανονική μορφή.

  Έχουμε

$$f_\phi(x) = e^{\phi \cdot T(x) - e^\phi} \cdot h(x), \quad \underbrace{\forall x \in \mathbb{N}}_{\text{στήριγμα}}$$

όπου $A(\phi) = e^\phi$ και $\lambda > 0 \Rightarrow \boxed{\varphi \in \mathbb{R}}$, όπου $\varphi = \log \lambda$.

Έχουμε $e^{\phi \cdot T(x) - A(\phi)}$, άρα

$$A(\phi) = B\left(\theta(\phi)\right)$$

$$B(\lambda) = \lambda \underset{\lambda \to \lambda(\phi)}{\Rightarrow} A(\phi) = B\left(\lambda(\phi)\right) = \lambda(\phi) = e^\phi$$

$\Rightarrow$ επέκταση για τυχαίο δείγμα από $\mathcal{P}(\lambda),\ \lambda > 0$.

μονοδιάστατη Ε.Ο.Κ. $\Rightarrow$ n-διάστατη Ε.Ο.Κ.

Για $x = (x_1, x_2, \ldots, x_n)$

(i) Στήριγμα $S_f = (\mathbb{N})^n$ ανεξ. του λ.

(ii) Για $x \in \mathbb{R}^n$,

$$f_\lambda(x) = \prod_{i=1}^n f_\lambda(x_i) = \beta^n(\lambda) \cdot e^{\phi(\lambda) \cdot \sum\limits_{i=1}^n T(x_i)} \cdot \prod_{i=1}^n h(x_i)$$

$$= \underbrace{\beta^*(\lambda)}_{\text{εξ. μόνο από } \lambda} \cdot e^{\phi(\lambda) \cdot T(x)} \cdot h(x)$$

όπου $\beta^*(\lambda) = \beta^n(\lambda),\ T(x) = \sum\limits_{i=1}^n T(x_i)$ και $h(x) = \prod\limits_{i=1}^n h(x_i)$.

Τελικά Ε.Ο.Κ.

Διάλεξη 24

Οι Ε.Ο.Κ. που θεωρούμε εδώ είναι **πλήρους τάξης**, όπου $\dim \Theta = \dim \Phi = s$.
(υπάρχει περίπτωση $\dim \Theta < \dim \Phi \to$ δείτε καμπυλωμένες εκθετικές οικογένειες κατανομών (curved))

Ειδικότερα, έχουμε επίσης ότι $\emptyset \subset \Theta^{\text{o}} \subset \mathbb{R}^s$.

Παράδειγμα 0.0.37. Έστω $X \sim \mathcal{U}[0,\theta],\ \theta > 0$. Θ.δ.ο. $X \notin$ Ε.Ο.Κ.(θ)

- $f_\theta(x) = \dfrac{1}{\theta} \cdot \mathbb{1}_{[0,\theta]}(x) \Rightarrow S_{f_\theta} = [0,\theta]$, εξαρτάται από το $\theta \Rightarrow$ δεν είναι Ε.Ο.Κ.

Παράδειγμα 0.0.38. Έστω $X \sim \mathcal{N}(\mu,\sigma^2),\ \mu,\sigma^2$ άγνωστα. Αποτελεί Ε.Ο.Κ.;

$\theta = (\mu,\sigma^2),\ \mu \in \mathbb{R},\ \sigma^2 > 0.$

- $f_\theta(x) = \dfrac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, \quad x \in \mathbb{R}.$
  (i) $S_{f_\theta} = \mathbb{R}$ ανεξάρτητο του θ
  (ii) $f_\theta(x) = \dfrac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\mu^2}{2\sigma^2}} \cdot \ e^{\frac{\mu}{\sigma^2}\cdot x - \frac{1}{2\sigma^2}\cdot x^2}$

  με
  $$\beta(\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\mu^2}{2\sigma^2}}$$

  $$e^{\varphi_1(\theta)\cdot T_1(x) + \varphi_2(\theta)\cdot T_2(x)} = e^{\frac{\mu}{\sigma^2}\cdot x - \frac{1}{2\sigma^2}\cdot x^2}$$

  (Αντί του η που συναντάμε στη βιβλιογραφία θα χρησιμοποιούμε το φ, φυσική παράμετρος, ενώ το ϕ θα μας οδηγεί στην παράμετρο)
  όπου
  $$\varphi_1(\theta) = \frac{\mu}{\sigma^2}, \quad T_1(x) = x \quad \text{και} \quad \varphi_2(\theta) = -\frac{1}{2\sigma^2}, \quad T_2(x) = x^2$$

  Συμπεραίνουμε ότι αποτελεί Ε.Ο.Κ. Επιπλέον $\phi_1 = \dfrac{\mu}{\sigma^2} \in \mathbb{R},\ \phi_2 = -\dfrac{1}{2\sigma^2} < 0.$

Κανονική Μορφή

$$f_\phi(x) = e^{\phi_1 \cdot T_1(x) + \phi_2 \cdot T_2(x) - A(\phi)} \cdot h(x), \quad \text{όπου } h(x) \equiv 1,\ \forall x \in \mathbb{R}$$

$$\beta(\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\mu^2}{2\sigma^2}} = e^{-\frac{1}{2}\left(log(2\pi\sigma^2) + \frac{\mu^2}{\sigma^2}\right)}$$

$$= e^{-B(\theta)}$$

Άρα

$$\left.\begin{aligned} B(\theta) &= \frac{1}{2} log(2\pi\sigma^2) + \frac{1}{2} \cdot \frac{\mu^2}{\sigma^2} \\ A(\phi) &= B(\theta(\phi)) \end{aligned}\right| \Leftrightarrow \left.\begin{aligned} \phi_1 &= \frac{\mu}{\sigma^2} \\ \phi_2 &= -\frac{1}{2\sigma^2} \end{aligned}\right| \Leftrightarrow \begin{aligned} \mu &= -\frac{\phi_1}{2\phi_2} \\ \sigma^2 &= -\frac{1}{2\phi_2} \end{aligned}$$

Άρα

$$A(\phi) = \frac{1}{2} \log\left(-\frac{\pi}{\phi_2}\right) + \frac{1}{2} \cdot \frac{\frac{\phi_1^2}{4\phi_2^2}}{-\frac{1}{2\phi_2}} = \frac{1}{2} \log\left(-\frac{\pi}{\phi_2}\right) - \frac{\phi_1^2}{4\phi_2}$$

Προσέξτε ότι

- $\frac{\partial A}{\partial \phi_1}(\phi) = -\frac{2\phi_1}{4\phi_2} = -\frac{\phi_1}{2\phi_2} = \mu$

  $T_1(x) = x, \quad \mathbb{E}(T_1(X)) = \mathbb{E}(X) = \mu$

- $\frac{\partial^2 A}{\partial \phi_1^2}(\phi) = -\frac{1}{2\phi_2} = \sigma^2$

  $T_1(x) = x, \quad \mathbb{V}(T_1(X)) = \mathbb{V}(X) = \sigma^2$

Πρόταση 0.0.12. Έστω X σε Ε.Ο.Κ. με κανονική μορφή

$$f_\phi(x) = e^{\sum\limits_{i=1}^{S} \phi_\kappa T_\kappa(x) - A(\phi)} \cdot h(x), \quad x \in S_f$$

Τότε

- $\mathbb{E}_\phi\left[T_\kappa(X)\right] = \frac{\partial A}{\partial \phi_\kappa}(\phi),\ \forall 1 \leq \kappa \leq S.$

- $\mathbb{C}_\phi\left[T_\kappa(X), T_\ell(X)\right] = \frac{\partial^2 A}{\partial \phi_\kappa \partial \phi_\ell}(\phi),\ \forall 1 \leq \kappa, l \leq S.$

  Ειδικά για $\kappa = \ell$,

- $\mathbb{V}_\phi\left[T_\kappa(X)\right] = \frac{\partial^2 A}{\partial \phi_\kappa^2}(\phi),\ \forall 1 \leq \kappa \leq S.$

Αν $s = 1$, μονοπαραμετρική Ε.Ο.Κ.

$$\mathbb{E}_{\phi}\left[T(X)\right] = A'(\phi)$$

$$\mathbb{V}_{\phi}\left[T(X)\right] = A''(\phi)$$

Η πρόταση αυτή μας επιτρέπει να βρίσκουμε μέσες τιμές, διασπορές και συνδιακυμάνσεις των σ.σ. $T_{\kappa}(X)$.

Παρατήρηση 0.0.30. Ειδικά στο προηγούμενο παράδειγμα, θα βρίσκαμε, αφού $T_1(X) = X \sim \mathcal{N}(\mu, \sigma^2)$ αυτά που είπαμε πριν, $\mathbb{E}(X) = \mu$, $\mathbb{V}(X) = \sigma^2$ κι επίσης αφού $T_2(X) = X^2$, μπορεί να βρεθεί

$$\underbrace{\mathbb{E}(X^2), \quad \mathbb{V}(X^2), \quad \mathbb{C}(X, X^2)}_{\text{κάντε practice}}$$

Πρόταση 0.0.13. Αν $X = (X_1, \ldots, X_n)$ ένα τ.δ. από μονοδιάστατη s-παραμετρική Ε.Ο.Κ.(θ), με $T(X_i) = (T_1(X_i), \ldots, T_s(X_i))$, τότε το X ανήκει σε n-διάστατη s-παραμετρική Ε.Ο.Κ. με

$$T(X) = \sum_{i=1}^{n} T(X_i),$$

όπου $T(X) = (T_1(X), \ldots, T_s(X))$ και $T_{\kappa}(X) = \sum_{i=1}^{n} T_{\kappa}(X_i)$.

Απόδειξη

Έχουμε $X_i \in$ Ε.Ο.Κ.(θ), $\theta = (\theta_1, \ldots, \theta_s)$ και $T(X_i) = (T_1(X_i), \ldots, T_s(X_i))$ στην αναπαράστασή της στην Ε.Ο.Κ.

Ισχύει αν $x = (x_1, \ldots, x_n)$

$$f_{\theta}(x) = \prod_{i=1}^{n} f_{\theta}(x_i),$$

όπου

$$x \in S_n = \{x \in \mathbb{R}^n : f_{\theta}(x) > 0\} = S_1^n$$

όπου $S_1 = \{x_i \in \mathbb{R} : f_{\theta}(x_i) > 0\}$

και επιπλέον είναι ανεξ. του θ, λόγω του ότι η μονοδιάστατη είναι σε Ε.Ο.Κ.

$\Rightarrow S_n$ ανεξ. του θ.

$$f_\theta(x) = \prod_{i=1}^{n} f_\theta(x_i) = \prod_{i=1}^{n} \beta(\theta) \cdot e^{\sum\limits_{\kappa=1}^{S} \varphi_\kappa(\theta) \cdot T_\kappa(x_i)} \cdot h(x_i)$$

$$= \beta^n(\theta) \cdot e^{\sum\limits_{\kappa=1}^{S} \varphi_\kappa(\theta) \cdot \sum\limits_{i=1}^{n} T_\kappa(x_i)} \cdot \prod_{i=1}^{n} h(x_i)$$

$$\beta^*(\theta) \cdot e^{\sum\limits_{\kappa=1}^{S} \varphi_\kappa(\theta) \cdot T_\kappa(x)} \cdot h(x),$$

όπου $T_\kappa(x) = \sum\limits_{i=1}^{n} T_\kappa(x_i)$.

Παρατήρηση 0.0.31. Αν θέλετε να δείξετε ότι ένα τ.δ. είναι σε n-διάστατη Ε.Ο.Κ. τότε δείχνετε ότι η μονοδιάστατη είναι σε Ε.Ο.Κ. και εφαρμόζετε την Πρόταση

↳ ίδιο $\varphi(\theta)$, $\quad \beta_n(\theta) = \beta^n(\theta)$, $\quad T(x) = \sum\limits_{i=1}^{n} T(x_i)$.

π.χ. σε κανονική μορφή $A_n(\phi) = n \cdot A(\phi)$.

Υπενθύμιση (Karlin-Rubin)

Αν X κατανομή με Μ.Λ.Π., τότε για την $H_0 : \theta \leq \theta_0 \quad$ vs $\quad H_1 : \theta > \theta_0$, τότε υπάρχει Ο.Ι.Ε. μεγέθους a, που δίνεται από

$$R^*(X) = \begin{cases} 1 & \text{, αν } T(X) > c \\ \gamma & \text{, αν } T(X) = c \\ 0 & \text{, αν } T(X) < c \end{cases}$$

$$\left(\text{για } \Theta \subset \mathbb{R}, \quad \text{Μ.Λ.Π.: } \theta_1 < \theta_2 \Rightarrow \frac{f_{\theta_2}(x)}{f_{\theta_1}(x)} \text{ αύξουσα στο } T(x)\right)$$

Λήμμα 0.0.12. Έστω $\Theta \subset \mathbb{R}$ και $\phi(\theta)$ αύξουσα συνάρτηση του θ. Τότε η μονοπαραμετρική Ε.Ο.Κ. (n-διάστατη) με

$$f_\theta(x) = e^{\phi(\theta)T(x) - B(\theta)} \cdot h(x), \quad x \in S_f$$

έχει Μ.Λ.Π. στην $T(x)$.

Απόδειξη

Έστω $\theta_1 < \theta_2$

$$\frac{f_{\theta_2}(x)}{f_{\theta_1}(x)} = \frac{e^{\phi(\theta_2)T(x)-B(\theta_2)} \cdot h(x)}{e^{\phi(\theta_1)T(x)-B(\theta_1)} \cdot h(x)} = e^{[\overbrace{\phi(\theta_2)-\phi(\theta_1)}^{\geq 0}]\cdot T(x)-B(\theta_2)+B(\theta_1)}$$

$$\Rightarrow \frac{f_{\theta_2}(x)}{f_{\theta_1}(x)} \text{ είναι αύξουσα συνάρτηση του } T(x).$$

Λόγω Karlin-Rubin

Πόρισμα 0.0.13. Η μονοπαραμετρική Ε.Ο.Κ. με $\phi(\theta)$ γνήσια αύξουσα ή φθίνουσα μας δίνει Ο.Ι.Ε. μεγέθους a

$$H_0 : \theta \leq \theta_0 \quad \text{vs} \quad H_1 : \theta > \theta_0$$

με

$$\begin{array}{cc} \phi \text{ γνήσια αύξουσα} & \phi \text{ γνήσια φθίνουσα} \end{array}$$

$$R^*(X) = \left\{ \begin{array}{lll} 1 & , & \text{αν } T(X) > c \\ \gamma & , & \text{αν } T(X) = c \\ 0 & , & \text{αν } T(X) < c \end{array} \right. \Bigg| \left\{ \begin{array}{lll} 1 & , & \text{αν } T(X) < c \\ \gamma & , & \text{αν } T(X) = c \\ 0 & , & \text{αν } T(X) > c \end{array} \right.$$

Αν ο έλεγχος είναι της μορφής

$$H_0 : \theta \geq \theta_0 \quad \text{vs} \quad H_1 : \theta < \theta_0$$

τότε οι παραπάνω ανισότητες αντιστρέφονται.

Σε κάθε περίπτωση οι σταθερές γ και a, καθορίζονται $\boxed{\Pi_{R^*}(\theta_0) = a}$

Παραδείγματα Κατανομών στην μονοπαραμετρική Ε.Ο.Κ.					
Κατανομή	$\varphi(\cdot)$	$T(X_i)$	$A(\phi)$	$T(X)$	Ο.Ι.Ε. $(H_0 : \theta \leq \theta_0$ vs $H_1 : \theta > \theta_0)$ Κρίσιμη περιοχή
Be(p)	$\log \frac{p}{1-p}$	X_i	$\log\big(1+e^{\phi}\big)$	$\sum_{i=1}^{n} X_i$	$> c$
Bin(N,p)	$\log \frac{p}{1-p}$	X_i	$N \cdot \log\big(1+e^{\phi}\big)$	$\sum_{i=1}^{n} X_i$	$> c$
$\mathcal{P}(\lambda)$	$\log \lambda$	X_i	e^{ϕ}	$\sum_{i=1}^{n} X_i$	$> c$
NegBin(r,p)	$\log p$	X_i	$-r \cdot \log\big(1-e^{\phi}\big)$	$\sum_{i=1}^{n} X_i$	$> c$
Exp(θ)	$-\theta$	X_i	$-\log(-\phi)$	$\sum_{i=1}^{n} X_i$	$< c$
G(a,θ)	?	?	?	?	?
Weibull$(\cdot,\cdot)$?					
Laplace$(\cdot,\cdot)$					
B(a,β)					

Διάλεξη 25

Παράδειγμα 0.0.39. Έστω $X_1, \dots, X_n$ τ.δ. από $\mathcal{N}(\mu, \sigma^2)$, όπου σ^2 γνωστό $\mu \in \mathbb{R}$.

Για τον έλεγχο

$$H_0 : \mu \leq \mu_0 \quad \text{vs} \quad H_1 : \mu > \mu_0$$

έχουμε ότι το τ.δ. $X = (X_1, \dots, X_n)$ ανήκει σε μονοπαραμετρική Ε.Ο.Κ. με $T(X) = \overline{X}$

και $\varphi(\mu) = \dfrac{n\mu}{\sigma^2} \nearrow$ ως προς μ.

Άρα $\exists$ Ο.Ι.Ε. μεγέθους a που είναι της μορφής $\boxed{R_a^* = \mathbb{1}_{(c_a, +\infty)}\left(\overline{X}\right)}$

Επιπλέον βρίσκουμε το c_a μέσα από τη σχέση

$$\mathbb{P}_{\mu_0}\left(\overline{X} > c_a\right) = a \tag{$\star$}$$

Ποιά είναι η κατανομή του $\overline{X}$ όταν $\mu = \mu_0$;

$$\overline{X} \overset{\mu_0}{\sim} \mathcal{N}\left(\mu_0, \frac{\sigma^2}{n}\right) \Rightarrow \mathbb{P}_{\mu_0}\left(\overline{X} > c_a\right) = \mathbb{P}_{\mu_0}\left(\underbrace{\frac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}}_{=Z \sim \mathcal{N}(0,1)} > \frac{c_a - \mu_0}{\frac{\sigma}{\sqrt{n}}}\right)$$

$$\overset{(\star)}{\Rightarrow} \frac{c_a - \mu_0}{\frac{\sigma}{\sqrt{n}}} = z_a \Rightarrow \boxed{c_a = \mu_0 + z_a \frac{\sigma}{\sqrt{n}}} \rightarrow \text{εδώ έχουμε μέγεθος } a \text{ στον έλεγχο}$$

Παρατήρηση 0.0.32. (ισοδυναμία ελεγχοσυναρτήσεων)

$$Z > z_a \Leftrightarrow \overline{x} > c_a$$

ή ισοδύναμα

$$g\left(\overline{x}\right) > z_a \quad \Leftrightarrow \quad \overline{x} > g^{-1}(z_a) \quad \text{με } g\left(\overline{x}\right) = \frac{\overline{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

όπου $g(\overline{X})$ ο μετασχηματισμός της τυποποίησης.

→ Καθορισμός ισχύος? Έχει ενδιαφέρον για $\mu > \mu_0$.

Εύρεση ισχύος

$$\pi_{R^*}(\mu) = \mathbb{P}_\mu\left(\overline{x} > c_a\right) = \mathbb{P}_\mu\left(\underbrace{\frac{\overline{x}-\mu}{\frac{\sigma}{\sqrt{n}}}}_{=Z\sim\mathcal{N}(0,1)} > \frac{c_a-\mu}{\frac{\sigma}{\sqrt{n}}}\right) \qquad (\text{όπου } c_a \text{ έχει προσδιοριστεί})$$

$$= 1-\Phi\left(\frac{c_a-\mu}{\frac{\sigma}{\sqrt{n}}}\right) = \Phi\left(\frac{\mu-c_a}{\frac{\sigma}{\sqrt{n}}}\right) \overset{c_a=\mu_0+z_a\frac{\sigma}{\sqrt{n}}}{=\!=\!=\!=} \Phi\left(\frac{\mu-\mu_0-z_a\frac{\sigma}{\sqrt{n}}}{\frac{\sigma}{\sqrt{n}}}\right)$$

$$= \Phi\left(\frac{\mu-\mu_0}{\frac{\sigma}{\sqrt{n}}} - z_a\right)$$

→ Ο.Ι.Ε. έξω από την Ε.Ο.Κ.!

Παράδειγμα 0.0.40. Έστω $X_1, X_2, \ldots, X_n$ τ.δ. από $\mathcal{U}[0,\theta]$, $\theta > 0$. Θέλουμε

$$H_0 : \theta \leq \theta_0 \quad \text{vs} \quad H_1 : \theta > \theta_0$$

Θ.δ.ο. το τ.δ. $X = (X_1, \ldots, X_n)$ έχει Μ.Λ.Π. στο $X_{(n)}$.

Πράγματι,

- έστω $\theta_1 < \theta_2$ και θα θυμηθούμε ότι $f_\theta(x) = \frac{1}{\theta^n} \cdot \mathbb{1}_{[0,\theta]}(x_{(n)})$

$$\frac{f_{\theta_2}(x)}{f_{\theta_1}(x)} = \frac{\theta_1^n}{\theta_2^n} \cdot \frac{\mathbb{1}_{[0,\theta_2]}(x_{(n)})}{\mathbb{1}_{[0,\theta_1]}(x_{(n)})}$$

Θέλουμε αυτό να είναι $\nearrow$ στο $x_{(n)}$ στα x για τα οποία τουλάχιστον ένα από τα

$$\underbrace{f_{\theta_1}(x) \text{ ή } f_{\theta_2}(x) > 0}_{\text{περιορισμός}}.$$

Για $\theta_1 < \theta_2$,

- αν $x_{(n)} > \theta_2$ τότε $f_{\theta_1}(x) = f_{\theta_2}(x) = 0$.
- αν $x_{(n)} \leq \theta_2$ τότε $f_{\theta_2}(x) > 0 \Rightarrow$ τουλάχιστον ένα είναι > 0.

Συμπέρασμα ο περιορισμός $\Leftrightarrow x_{(n)} \leq \theta_2$.

Άρα αν $x_{(n)} \leq \theta_2 \Rightarrow$

$$\frac{f_{\theta_2}(x)}{f_{\theta_1}(x)} = \left(\frac{\theta_1}{\theta_2}\right)^n \cdot \frac{1}{\mathbb{1}_{[0,\theta_1]}(x_{(n)})} = \begin{cases} \frac{\theta_1^n}{\theta_2^n} & , 0 < x_{(n)} \leq \theta_1 \\ +\infty & , \theta_1 < x_{(n)} \leq \theta_2 \end{cases}$$

Συμπεραίνουμε ότι είναι πράγματι αύξουσα στο $x_{(n)}$

$\Rightarrow$ έχει Μ.Λ.Π. στο $x_{(n)}$

$\Rightarrow +x_{(n)}$ είναι συνεχής, άρα από Karlin-Rubin $\exists$ Ο.Ι.Ε.

$$R^* = \mathbb{1}_{(c,+\infty)}(x_{(n)})$$

και στη συνέχεια προσδιορίζουμε το c_a μέσω της συνθήκης

$$\mathbb{P}_{\theta_0}(x_{(n)} > c_a) = a \tag{$\star$}$$

Έχουμε δείξει ότι $F_{X_{(n)}}(x) \overset{0\leq x\leq\theta}{=\!=\!=} \left(\frac{x}{\theta}\right)^n$.

Άρα $\mathbb{P}_{\theta_0}(x_{(n)} > c_a) = 1 - \mathbb{P}_{\theta_0}(x_{(n)} \leq c_a) = 1 - \left(\frac{c_a}{\theta_0}\right)^n$ και θέλουμε $\overset{(\star)}{=} a$.

Άρα

$$1 - \left(\frac{c_a}{\theta_0}\right)^n = a \Leftrightarrow \frac{c_a^n}{\theta_0^n} = 1 - a \Leftrightarrow c_a^n = \theta_0^n(1-a) \Leftrightarrow c_a = \theta_0 \sqrt[n]{1-a}$$

Εύρεση ισχύος ($\theta > \theta_0$)

$$\pi_{R^*}(\theta) = \mathbb{P}_\theta(x_{(n)} > c_a) = 1 - \mathbb{P}_\theta(x_{(n)} \leq c_a) = 1 - \left(\frac{c_a}{\theta}\right)^n \overset{c_a=\theta_0\sqrt[n]{1-a}}{=\!=\!=\!=\!=}$$

$$= 1 - \left(\frac{\theta_0}{\theta}\right)^n (1-a)$$

$\rightarrow$ Ο.Ι.Ε. δεν υπάρχουν πάντα.

Θεωρούμε τώρα τη γενική περίπτωση όπου

$$H_0 : \theta \in \Theta_0 \quad \text{vs} \quad H_1 : \theta \in \Theta_1$$

όπου $\Theta = \Theta_0 \cup \Theta_1$, με $\Theta_0 \cap \Theta_1 = \emptyset$.

Ορισμός 0.0.23. Ένας έλεγχος **λόγου πιθανοφάνειας** (λέγεται και γενικευμένου) είναι ένας έλεγχος που απορρίπτει την H_0, αν $\lambda(x) < c$, όπου

$$\lambda(x) = \frac{\sup\limits_{\theta\in\Theta_0} L(\theta)}{\sup\limits_{\theta\in\Theta} L(\theta)}$$

για κατάλληλο $c: \ 0 \leq c \leq 1$.

Παρατήρηση 0.0.33.

(i) όταν το $\lambda(x)$ είναι καλά ορισμένο, τότε $\lambda(x) \leq 1$.

(ii) αν η H_0 είναι αληθής, τότε το $\lambda(x)$ τείνει να είναι κοντά στο 1, ενώ αν η H_1 είναι αληθής, τότε τείνει να είναι μακριά από το 1.

(iii) αν υπάρχουν οι ε.μ.π. $\widehat{\theta}$ και $\widehat{\theta_0}$ στον Θ και στον Θ_0 αντίστοιχα, τότε

$$\lambda(x) = \frac{L\left(\widehat{\theta_0}\right)}{L\left(\widehat{\theta}\right)}$$

όπου $L\left(\widehat{\theta_0}\right)$ περιορισμένος (restricted) στο Θ_0.

(iv) αν $\Theta_0 = \{\theta_0\},\ \Theta_1 = \{\theta_1\}$, τότε ο έλεγχος-$LR$ (Likehood Ratio) είναι ισοδύναμος με αυτόν που προκύπτει από $N-P$, $\dfrac{L(\theta_1)}{L(\theta_0)} > c_0$ για $c_0 \geq 1$.

(v) για $0 < a < 1$, αν $\exists c_a \in [0,1] : \sup\limits_{\theta \in \Theta_0} \mathbb{P}_\theta(\lambda(X) < c_a) = a$, τότε μπορούμε να φτιάξουμε έλεγχο-LR μεγέθους a.

Στη γενική περίπτωση, ακόμα και αν το $\lambda(x)$ είναι συνεχής τ.μ. ή/και αν γίνει τυχαιοποίηση μπορεί να μην υπάρχει c_a όπως παραπάνω.

Πρόταση 0.0.14. (απόδειξη π.χ. σελ. 429 Shao)
Ας υποθέσουμε ότι το τ.δ. X ανήκει σε μονοπαραμετρική Ε.Ο.Κ. με $\varphi(\theta) \nearrow$ και διαφορίσιμη στο θ. Τότε

(i) Για τον έλεγχο

$$H_0 : \theta \leq \theta_0 \quad \text{vs} \quad H_1 : \theta > \theta_0,$$

υπάρχει έλεγχος-LR με περιοχή απόρριψης που συμπίπτει με τον Ο.Ι.Ε.

$$R^*(X) = \begin{cases} 1 & , \text{αν } T(X) > c \\ \gamma & , \text{αν } T(X) = c \\ 0 & , \text{αν } T(X) < c \end{cases}$$

(ii) Για τον έλεγχο

$$H_0 : \theta \leq \theta_1 \text{ ή } \theta \geq \theta_2 \quad \text{vs} \quad H_1 : \theta_1 < \theta < \theta_2$$

υπάρχει έλεγχος-LR με κρίσιμη περιοχή που δίνεται

$$R^*(X) = \begin{cases} 1 & , \text{αν } c_1 < T(X) < c_2 \\ \gamma_i & , \text{αν } T(X) = c_i,\ i = 1,2 \\ 07, \text{αν } T(X) < c_1 \text{ ή } T(X) > c_2 & \end{cases}$$

όπου τα γ_i και c_i προσδιορίζονται από

$$\pi_{R^*}(\theta_1) = \pi_{R^*}(\theta_2) = a \qquad \text{(μάλιστα εδώ προκύπτει Ο.Ι.Ε.)}$$

(iii) Για ελέγχους

$$H_0 : \theta_1 \leq \theta \leq \theta_2 \quad \text{vs} \quad H_1 : \theta < \theta_1 \text{ ή } \theta > \theta_2 \quad (\theta_1 < \theta_2)$$

$$H_0 : \theta = \theta_0 \quad \text{vs} \quad H_1 : \theta \neq \theta_0$$

υπάρχει έλεγχος-LR με κρίσιμη περιοχή που δίνεται από τη σχέση $T(X) < c_1$ ή $T(X) > c_2$ με σταθερές όπως παραπάνω.

Πόρισμα 0.0.14. Σε τ.δ. από $\mathcal{N}(\mu, \sigma^2)$ με σ^2 γνωστό, τότε για

$$H_0 : \mu = \mu_0 \quad \text{vs} \quad H_1 : \mu \neq \mu_0$$

υπάρχει έλεγχος-LR με κρίσιμη περιοχή που δίνεται από τη σχέση $\overline{X} < c_1$ ή $\overline{X} > c_2$.

Ειδικά μπορούμε να τα επιλέξουμε με τρόπο ώστε η κρίσιμη περιοχή να είναι της μορφής

$$|Z| > z_{\frac{a}{2}},$$

οπότε τότε

$$c_1 = \mu_0 - Z_{\frac{a}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

$$c_2 = \mu_0 + Z_{\frac{a}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

Εδώ πέρα προσοχή. Συνήθως με δίπλευρες εναλλακτικές δεν έχουμε Ο.Ι.Ε.

Παράδειγμα 0.0.41. (εφαρμογή-LR)

Σε ένα τ.δ. $X_1, \ldots, X_n$ από $\mathcal{N}(\mu, \sigma^2)$ όπου $\mu \in \mathbb{R}$, $\sigma^2 > 0$ είναι άγνωστα. Θέλουμε να πραγματοποιήσουμε έλεγχο

$$H_0 : \mu = \mu_0 \quad \text{vs} \quad H_1 : \mu \neq \mu_0,$$

με σ^2 άγνωστο. Πώς κάνουμε έναν τέτοιο έλεγχο με σ^2 άγνωστο;

Λύση

Θα χρησιμοποιήσουμε έλεγχο-LR

$$\lambda(x) = \frac{\sup\limits_{\theta\in\Theta_0} L(\theta)}{\sup\limits_{\theta\in\Theta} L(\theta)}$$

Καταρχήν για το θ υπάρχει η ε.μ.π. $\widehat{\theta} = \left(\widehat{\mu}, \widehat{\sigma^2}\right) = \left(\overline{X}, M_2\right)$ άρα

$$\sup_{\theta\in\Theta} = L\left(\widehat{\theta}\right) = \left(2\pi\widehat{\sigma^2}\right)^{-\frac{n}{2}} \; e^{-\frac{1}{2\widehat{\sigma^2}} \sum\limits_{i=1}^{n}\left(X_i-\overline{X}\right)^2}$$

Όμως

$$M_2 = \frac{1}{n}\cdot\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2 \Rightarrow L\left(\widehat{\theta}\right) = \left(2\pi\widehat{\sigma^2}\right)^{-\frac{n}{2}} \; e^{-\frac{n}{2}} \tag{21}$$

Αντίστοιχα στον Θ_0, το $\mu = \mu_0$ και σ^2 άγνωστο, τότε έχουμε βρει για μ σταθερό ότι

$$\widehat{\sigma_0^2} = \frac{1}{n}\cdot\sum_{i=1}^{n}(X_i-\mu_0)^2 \Rightarrow \sup_{\theta\in\Theta_0} L(\theta) = L\left(\widehat{\theta_0}\right) = \left(2\pi\widehat{\sigma_0^2}\right)^{-\frac{n}{2}} \; e^{-\frac{n}{2}} \tag{22}$$

Άρα

$$\lambda(x) = \frac{L\left(\widehat{\theta_0}\right)}{L\left(\widehat{\theta}\right)} = \left(\frac{2\pi\widehat{\sigma_0^2}}{2\pi\widehat{\sigma^2}}\right)^{-\frac{n}{2}} = \left(\frac{\widehat{\sigma_0^2}}{\widehat{\sigma^2}}\right)^{-\frac{n}{2}} \tag{23}$$

Επίσης έχουμε $\widehat{\sigma^2} < \widehat{\sigma_0^2}$, διότι

$$\widehat{\sigma_0^2} = \frac{1}{n}\cdot\sum_{i=1}^{n}(X_i-\mu_0)^2 = \mathbb{E}\left(X^*-\mu_0\right)^2 = \left(\overline{X}-\mu_0\right)^2 + \underbrace{\mathbb{V}(X^*)}_{\widehat{\sigma^2}} \tag{24}$$

Άρα επειδή $\widehat{\sigma_0^2} > \widehat{\sigma^2}$

$$\frac{L\left(\widehat{\theta_0}\right)}{L\left(\widehat{\theta_1}\right)} < c \overset{\text{αρν.εκθ.}}{\underset{(23)}{\Leftrightarrow}} \frac{\widehat{\sigma_0^2}}{\widehat{\sigma^2}} > c' \underset{(24)}{\Leftrightarrow} \frac{\left(\overline{X}-\mu_0\right)^2}{\widehat{\sigma^2}} > c''$$

$$\Leftrightarrow \frac{\left(\overline{X}-\mu_0\right)^2}{\frac{S^2}{n}} > c'' \Leftrightarrow \underbrace{\left|\frac{\overline{X}-\mu_0}{\frac{S}{\sqrt{n}}}\right| > c_a}_{\text{μορφή κρίσιμης περιοχής}}$$

όταν θέλουμε μέγεθος a.

Όμως αν

$$T = \frac{\overline{X}-\mu_0}{\frac{S}{\sqrt{n}}} \overset{\mu_0}{\sim} t_{n-1}$$

τότε γνωρίζουμε την κατανομή του T και

$$c_a = t_{n-1,\frac{a}{2}} : \; \mathbb{P}\left(|T| > t_{n-1,\frac{a}{2}}\right) = a$$

Διάλεξη 26

Εφαρμογές

① Σώζουν οι αερόσακοι ζωές;

Υποθέτουμε ότι θάνατοι σε τροχαία ατυχήματα, για μία συγκεκριμένη μάρκα αυτοκινήτου, συμβαίνουν κατά μέσο όρο 6 / εβδομάδα και η εταιρεία εισάγει αερόσακους.

Θέλει η εταιρεία με τα δεδομένα της επόμενης χρονιάς (52 εβδομάδες) να ελέγξουν την αποτελεσματικότητά τους.

Υποθέτουμε ότι $X_i = \#$ θανάτων την i-εβδομάδα που θεωρούμε ότι $X_i \sim \mathcal{P}(\lambda),\ \lambda > 0$.

Πώς να σχεδιάσουμε έναν κατάλληλο έλεγχο για να ελέγξουμε την αποτελεσματικότητά τους.

$$H_0 : \lambda = 6 \quad \text{vs} \quad H_1 : \lambda < 6$$

θα διεξαχθεί σε κάποιο επίπεδο στατιστικής σημαντικότητας α, π.χ. $\alpha = 0.05$.

Έχουμε δείξει ότι ένας Ομοιόμορφα Ισχυρότατος Έλεγχος μεγέθους $\alpha = 0.05$, θα είναι ο

$$R^* = \begin{cases} 1 & , \ T < c_{0.05} \\ \gamma & , \ T = c_{0.05} \\ 0 & , \ T > c_{0.05} \end{cases}$$

όπου $T = \sum_{i=1}^{n} X_i \underset{n=52}{\overset{\lambda=6}{\sim}} \mathcal{P}(\underbrace{312}_{n\lambda})$, ως άθροισμα n ανεξάρτητων $\mathcal{P}(\lambda)$, και εφόσον η $\mathcal{P}$ ανήκει σε μονοπαραμετρική Εκθετική Ομάδα Κατανομών και η $H_0 : \lambda = \lambda_0 \ \text{ vs } \ H_1 : \lambda < \lambda_0$ $\big[$διαισθητικά η H_1 ευνοείται όταν το $\sum X_i$ είναι επαρκώς μικρό$\big]$

όπου τα $c_{0.05}$ και γ προσδιορίζονατι από τη σχέση

$$F_0(c_{0.05} - 1) < 0.05 \leq F_0(c_{0.05})$$

και

$$\gamma = \frac{0.05 - F(c_{0.05} - 1)}{\mathbb{P}_0(T = c_{0.05})}$$

$$\left.\begin{aligned} F_0(282) &\cong 0.0457 \\ F_0(283) &\cong 0.0517 \end{aligned}\right\} \Rightarrow c_{0.05} = 283 \quad \gamma \cong 0.727$$

$\rightarrow$ Εδώ $n = 52$ και θα μπορούσαμε να πραγματοποιήσουμε έναν ασυμπτωτικό έλεγχο σε επίπεδο στατιστικής σημαντικότητας $\alpha = 0.05$, βρίσκοντας την ασυμπτωτική κατανομή της ελεγχοσυνάρτησης.

Εδώ $T = \sum_{i=1}^{52} X_i \approx \mathcal{N}(\underset{\substack{\| \\ 312}}{n\lambda}, \underset{\substack{\| \\ 312}}{n\lambda})$ (με εφαρμογή του Κεντρικού Οριακού Θεωρήματος)

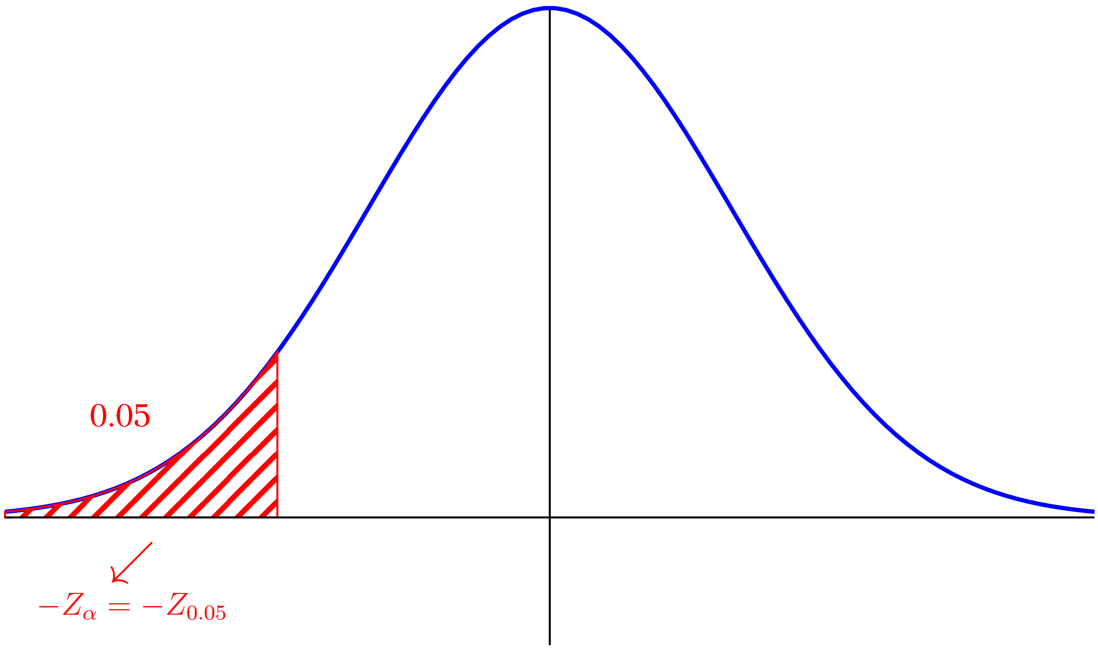
$\rightarrow$ λύνουμε

$$\mathbb{P}(T < c_\alpha) = \alpha$$

$$(\text{τυποπ.}) \ \Updownarrow \ \text{ασυμπτ. προσεγ.}$$

$$\mathbb{P}\left(Z < \frac{c_\alpha - 312}{\sqrt{312}}\right) = \Phi\left(\frac{c_\alpha - 312}{\sqrt{312}}\right)$$

Άρα λύνουμε $\Phi\left(\frac{c_\alpha - 312}{\sqrt{312}}\right) = \alpha = 0.05$



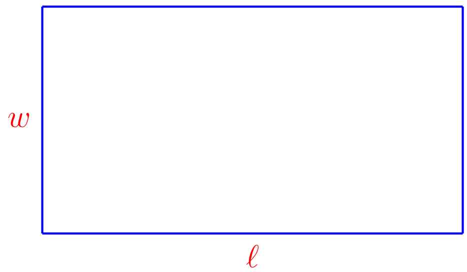
$\Rightarrow \dfrac{c_\alpha - 312}{\sqrt{312}} = -Z_{0.05} \cong -1.645$

Τελικά $c_\alpha = 312 - 1.645\sqrt{312} \cong 282.94$

Τελικά κάνοντας στρογγυλοποίηση $c_{0.05} = 283$! (ακριβώς ίδιο με το “ακριβές”)

2 Οι Ορθογώνιες χάντρες των Ινδιάνων Shoshoni

Τα περισσότερα άτομα, αν τα ρωτήσουμε να σχεδιάσουν ένα ορθογώνιο θα κάνουν κάτι



κάτι που "ευχαριστεί το μάτι" (όχι τόσο τετράγωνο, όχι τόσο πεπλατυσμένο).

Οι Αρχαίοι Έλληνες έλεγαν το ορθογώνιο χρυσό, αν ο λόγος

$$\frac{\text{πλάτος}}{\text{μήκος}} = \frac{w}{\ell} = \frac{1}{2}(\sqrt{5}-1) \cong 0.618,$$

συμπίπτει με τη χρυσή αναλογία $\dfrac{1}{\varphi} = \varphi - 1$.

$\left(\text{χρυσός λόγος: } \dfrac{w+\ell}{\ell} = \dfrac{\ell}{w} \to \varphi \cong 1.618\right)$

Οι Ινδιάνοι χρησιμοποιούσαν ορθογώνιες χάντρες για να διακοσμήσουν τα δερμάτινα αγαθά τους. Στην πειραματική αισθητική (experimental aesthetics) έγινε έρευνα με πραγματικά δεδομένα, μέτρησαν 20 λόγους από αυτά.

0.693	0.670	0.654	0.749
0.606	0.553	0.601	0.609
0.672	0.662	0.606	0.615
0.844	0.570	0.903	0.576
0.688	0.628	0.690	0.611

δεδομένα απο Lowie's Selected Papers in Anthropology ed Dubois,
Univ. of California Press

→ Μήπως οι Shoshoni ενστικτωδώς κάνουν τα ορθογώνιά τους να είναι συμβατά με το χρυσό λόγο;

$$H_0 : \theta = 0.618 \quad \text{vs} \quad H_1 : \theta \neq 0.618$$

Υποθέτουμε ότι είναι "κανονικά" κατανεμημένα (χρειάζεται επιπλέον έλεγχο).

Άρα $X_i \sim \mathcal{N}(\theta, \sigma^2)$, όπου θ, σ^2 άγνωστα.

Πώς πραγματοποιούμε τον έλεγχο;

Είναι από τ.δ. $\mathcal{N}(\theta, \sigma^2)$ με μ, σ^2 άγνωστα. Θα χρησιμοποιήσουμε

$$T \overset{H_0}{=} \sqrt{n} \cdot \frac{\overline{x} - 0.618}{S} \sim t_{19} \quad (t_{n-1}, n = 20)$$

$\overline{x} = 0.660,\ S = 0.093 \to t = 2.019$

p-value $\equiv a(x) = \mathbb{P}(|T| > t) = 0.058$
$$\downarrow$$
Δεν απορρίπτεται η H_0 σε ε.σ.σ. $a = 0.05$

Έλεγχοι Διασποράς σε τ.δ. από $\mathcal{N}(\mu, \sigma^2)$

Έστω $X = (X_1, \ldots, X_n)$ τ.δ. από $\mathcal{N}(\mu, \sigma^2)$. Θέλουμε να ελέγξουμε:

(i) $H_0 : \sigma^2 \leq \sigma_0^2$ vs $H_1 : \sigma^2 > \sigma_0^2$

(ii) $H_0 : \sigma^2 \geq \sigma_0^2$ vs $H_1 : \sigma^2 < \sigma_0^2$

(iii) $H_0 : \sigma^2 = \sigma_0^2$ vs $H_1 : \sigma^2 \neq \sigma_0^2$

Να βρεθούν Ο.Ι.Ε. μεγέθους a εκεί που υπάρχουν, διαφορετικά να γίνει χρήση του ελέγχου-LR για την εύρεση ελέγχου μεγέθους a, όταν

(A) το μ είναι γνωστό

(B) το μ είναι άγνωστο
(για 2 παραμέτρους άγνωστες χρησιμοποιούμε έλεγχο-LR)

Παράδειγμα B-(i) [τα άλλα να λυθούν]

Έχουμε 2 παραμέτρους άγνωστες $\Rightarrow$ LR για την εύρεση ελέγχου μεγέθους a.

Θέτουμε
$$\lambda(x) = \frac{\sup\limits_{\theta \in \Theta_0} L(\theta)}{\sup\limits_{\theta \in \Theta} L(\theta)}, \quad \theta = (\mu, \sigma^2)$$

και λύνουμε $\lambda(x) < c$

Καταρχάς γνωρίζουμε ότι ανεξάρτητα από την τιμή του σ^2, $\boxed{\widehat{\mu} = \overline{X}}$.

Άρα για $\Theta_0 = \mathbb{R} \times [0, \sigma_0^2]$ $\quad [H_0 : \sigma^2 \leq \sigma_0^2] \to \mu$ nuisance parameter
(παράμετρος θορύβου)

και
$$\sup_{\substack{\mu \in \mathbb{R} \\ \sigma^2 \leq \sigma_0^2}} L(\mu, \sigma^2) = \sup_{\sigma \leq \sigma_0^2} \sup_{\mu \in \mathbb{R}} L(\mu, \sigma^2) = \sup_{\sigma \leq \sigma_0^2} L(\overline{x}, \sigma^2)$$

$$L(\overline{x}, \sigma^2) = \left(2\pi\sigma^2\right)^{-\frac{n}{2}} \ e^{-\frac{1}{2\sigma^2} \ \sum\limits_{i=1}^{n} (x_i - \overline{x})^2}$$

$\Rightarrow$ αν θέσουμε $\widehat{\sigma^2} = \dfrac{1}{n} \sum\limits_{i=1}^{n} (x_i - \overline{x})^2 \quad (= M_2)$ τότε

$$L(\overline{x}, \sigma^2) = \left(2\pi\sigma^2\right)^{-\frac{n}{2}} \ e^{-\frac{n}{2} \cdot \frac{\widehat{\sigma^2}}{\sigma^2}} \tag{25}$$

$$\overset{\log L}{\Rightarrow} \ell(\overline{x}, \sigma^2) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\sigma^2 - \frac{n}{2}\frac{\widehat{\sigma^2}}{\sigma^2}$$

με $\theta = \sigma^2$ γίνεται,

$$\ell(\overline{x}, \theta) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\theta - \frac{n}{2}\frac{\widehat{\theta}}{\theta}$$

$$\Rightarrow \ell'(\overline{x}, \theta) = -\frac{n}{2\theta} + \frac{n}{2}\frac{\widehat{\theta}}{\theta^2} = \boxed{\frac{n}{2\theta}\left(\frac{\widehat{\theta}}{\theta} - 1\right)}$$

$$\Rightarrow \ell'(\overline{x}, \theta) > 0 \Leftrightarrow \theta < \widehat{\theta}$$

$$\ell'(\overline{x}, \theta) = 0 \Leftrightarrow \theta = \widehat{\theta}$$

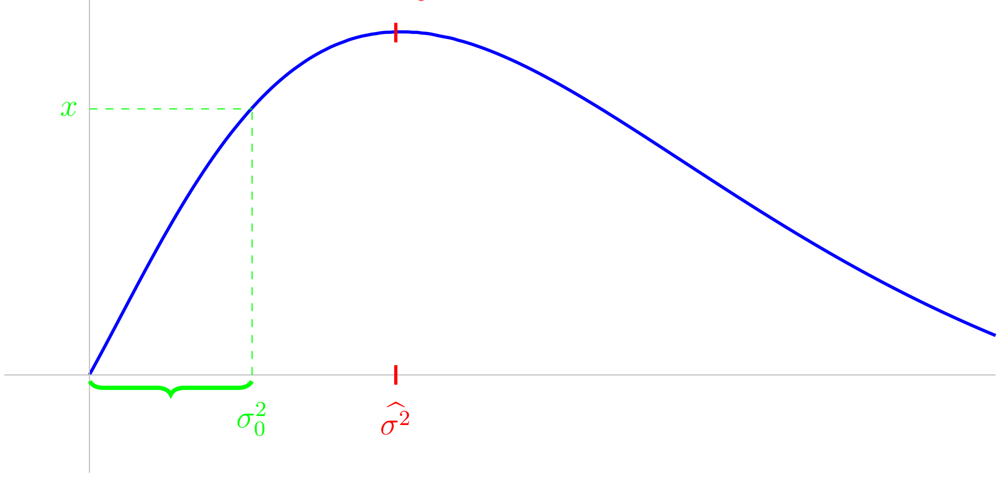
$$\overset{\theta \to \sigma^2}{\Rightarrow} L(\overline{x}, \sigma^2) \nearrow \text{ για } \sigma^2 < \widehat{\sigma^2}$$

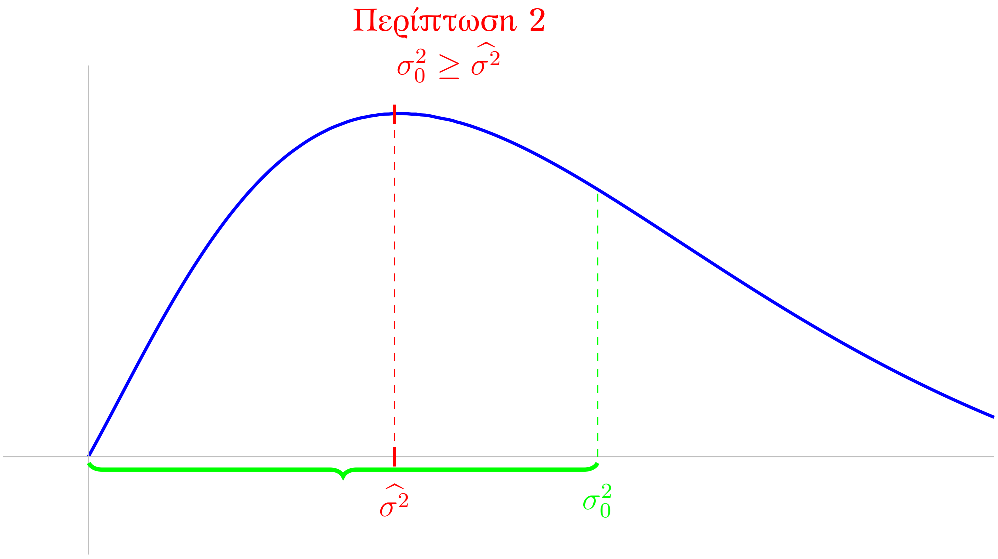
$$L(\overline{x}, \sigma^2) \searrow \text{ για } \sigma^2 > \widehat{\sigma^2}$$

$+$ παρεμβολή του σ_0^2

Περίπτωση 1

$\sigma_0^2 < \widehat{\sigma^2}$





$\sigma^2 \leq \sigma_0^2$ $\max\limits_{\sigma^2<\sigma_0^2} L(\overline{x},\theta) = L(\overline{x},\sigma_0^2)$

μέγιστο $L\left(\overline{x},\widehat{\sigma}^2\right) \Rightarrow \lambda(x) = \dfrac{L\left(\overline{x},\widehat{\sigma}^2\right)}{L\left(\overline{x},\widehat{\sigma}^2\right)} = 1 \longrightarrow$ δεχόμαστε την H_0. Τι γίνεται για $\sigma_0^2 < \widehat{\sigma}^2$;

Αναλύουμε την Περίπτωση 1.

$$\lambda(x) = \frac{L\left(\overline{x},\sigma_0^2\right)}{L\left(\overline{x},\widehat{\sigma}^2\right)} \overset{(25)}{=} \left(\frac{\sigma_0^2}{\widehat{\sigma}^2}\right)^{-\frac{n}{2}} e^{-\frac{n}{2}\left(\frac{\widehat{\sigma}^2}{\sigma_0^2}-1\right)}$$

$$= e^{\frac{n}{2}} \cdot \left(\frac{\widehat{\sigma}^2}{\sigma_0^2}\right)^{\frac{n}{2}} \cdot e^{-\frac{n}{2}\cdot\frac{\widehat{\sigma}^2}{\sigma_0^2}}$$

όπου $\widehat{\sigma}^2 = \widehat{\sigma}^2(x)$

Άρα $\lambda(x) < c \Leftrightarrow \left(\dfrac{\widehat{\sigma}^2}{\sigma_0^2}\right)^{\frac{n}{2}} \cdot e^{-\frac{n}{2}\cdot\frac{\widehat{\sigma}^2}{\sigma_0^2}} < c'$

Θέτουμε $g(x) = x^{\frac{n}{2}} e^{-\frac{n}{2}x}$, όπου $x = \dfrac{\widehat{\sigma}^2}{\sigma_0^2} > 1$

$$g'(x) = \frac{n}{2} x^{\frac{n}{2}-1} e^{-\frac{n}{2}x} - \frac{n}{2} x^{\frac{n}{2}} e^{-\frac{n}{2}x}$$

$$= \frac{n}{2} x^{\frac{n}{2}-1} e^{-\frac{n}{2}x}(1-x) < 0, \quad \text{επειδή } x > 1$$

$\Rightarrow g(x) \searrow$ και συμπεραίνουμε ότι

$$g(x) < c' \overset{x=\frac{\widehat{\sigma}^2}{s_0^2}}{\Leftrightarrow} \frac{\widehat{\sigma}^2}{s_0^2} > c'' \overset{\widehat{\sigma}^2 \to S^2}{\Leftrightarrow} \frac{S^2}{\sigma_0^2} > c'''$$

Συμπεραίνουμε τελικά ότι ο έλεγχος-LR μας οδηγεί σε μορφή ελέγχου $R_\alpha = \mathbb{1}_{(c_\alpha,+\infty)}\left(\frac{S^2}{\sigma_0^2}\right)$

Παρατήρηση 0.0.34. Για τα συνήθη ε.σ.σ. $c_\alpha \gg 1$.

- Καθορισμός του c_α για μέγεθος α.

Θέλουμε, για $\mu \in \mathbb{R},\ \sigma^2 \leq \sigma_0^2$

$$\sup \mathbb{P}_{\mu,\sigma^2}\left(\frac{S^2}{\sigma_0^2} > c_\alpha\right) = \alpha$$

Καταρχήν

$$\frac{(n-1)S^2}{\sigma^2} \overset{\sigma^2}{\sim} \chi^2_{n-1}, \quad \text{ανεξ. του } \mu \in \mathbb{R}\ (\forall\ \mu \in \mathbb{R})$$

Άρα

$$\sup_{\substack{\mu\in\mathbb{R}\\ \sigma^2\leq\sigma_0^2}} \mathbb{P}_{\mu,\sigma^2}\left(\frac{S^2}{\sigma_0^2} > c_\alpha\right) \overset{!}{=} \sup_{\sigma^2\leq\sigma_0^2} \mathbb{P}_{\sigma^2}\left(\frac{(n-1)S^2}{\sigma^2} > \frac{\sigma_0^2}{\sigma^2}(n-1)\,c_\alpha\right)$$

όπου το $\left(\frac{(n-1)S^2}{\sigma^2}\right)$ είναι μια ελεγχοσυνάρτηση $Q \overset{\sigma^2}{\sim} \chi^2_{n-1}$.

Επομένως

$$\sup_{\sigma^2\leq\sigma_0^2} \underset{\substack{\| \\ 1-F}}{\overline{F}_{\chi^2_{n-1}}}\left(\frac{\sigma_0^2}{\sigma^2}(n-1)\,c_\alpha\right) = \sup_{\sigma^2\leq\sigma_0^2} \overline{F}\left(g\left(\sigma^2\right)\right) \quad (\text{συνάρτηση επιβίωσης})$$

Έχουμε $g(\sigma^2)\ \searrow$ καθώς $\sigma^2 \nearrow \sigma_0^2$

$$\Rightarrow \overline{F}(g(\sigma^2)) \nearrow \qquad (\text{Γενικά } \overline{F}(x) \searrow \text{ καθώς } x \nearrow)$$

$$\Rightarrow \sup_{\sigma^2\leq\sigma_0^2}(\cdot) = \mathbb{P}_{\sigma_0^2}\left(\frac{(n-1)S^2}{\sigma_0^2} > (n-1)c_a\right)$$

όπου $\frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2_{n-1}$

Άρα θέτοντας το παραπάνω $= a \Rightarrow$

$$(n-1)c_a = \chi^2_{n-1,a} \Rightarrow \boxed{c_a = \frac{\chi^2_{n-1,a}}{n-1}}$$

Παρατήρηση 0.0.35. Αποδεικνύεται μάλιστα ότι είναι Ο.Ι.Ε. (δείτε σελ. 87 Lehmann-Romano) και μάλιστα κανένας άλλος έλεγχος με 2 άγνωστες παραμέτρους στην $\mathcal{N}(\mu,\sigma^2)$ δεν είναι Ο.Ι.Ε. στα συνήθη ε.σ.σ..

Απορίες στην ύλη του Quiz 3

Από το παράδειγμα της επιλογής 2 υποψηφίων (π.χ. δημάρχων) A ή B από δείγμα 400 ατόμων από τα οποία τα 212 επιλέγουν τον A και τα 188 τον B.

Είχαμε δει $\underset{\overline{x}}{\underset{\parallel}{\widehat{p}}} = 0.53$ $\quad x_i = \begin{cases} 1 & ,i\text{-άτομο επιλέγει τον A} \\ \\ 0 & ,\text{διαφορετικά} \end{cases}$ ασυμπτωτικό $(1-a)$-Δ.Ε.

$$\widetilde{I}^p_{1-a} = \widehat{p} \pm z_{\frac{a}{2}} \cdot \frac{\sqrt{\widehat{p}(1-\widehat{p})}}{\sqrt{n}}$$

Τότε είδαμε διαισθητικά ότι $0.5 \in \widetilde{I}^p_{0.95}$ άρα δε φαίνεται να εξασφαλίζεται με 95% σιγουριά νίκη του A.

$\rightarrow$ Ας υποθέσουμε ότι θέλαμε να ελέγξουμε νίκη του A.

Έχουμε $\widehat{p} =$ ποσοστό ψήφων του A.

$$\underset{\substack{\downarrow \\ \text{ισοψηφία} \\ \text{ή} \\ \text{νίκη του B}}}{H_0 : p \leq 0.5} \qquad \text{vs} \qquad \underset{\substack{\downarrow \\ \text{νίκη του A}}}{H_1 : p > 0.5}$$

ασυμπτωτικός έλεγχος (κανονική)

$\rightarrow$ περιοχή απόρριψης: $Z > z_a$ σε ε.σ.σ. a

Γιατί Z;

Είπαμε $\widehat{p} = \overline{X}$ και H_0 σύνθετη, $p \leq 0.5$ Αφαιρώ μέση τιμή και διαιρώ με τυπική απόκλιση. Εδώ,

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}, \quad X_i \sim \text{Be}(p)$$

άρα

$$\mathbb{E}(\overline{X}) = p, \quad \mathbb{V}(\overline{X}) = \frac{p(1-p)}{n}.$$

Άρα

$$\frac{\overline{X}-p}{\sqrt{\frac{p(1-p)}{n}}} \overset{+n \text{ μεγάλο}}{\underset{p}{\approx}} \mathcal{N}(0,1)$$

Χρησιμοποιώ τη Z που είναι ασυμπτωτικά οδηγός και τη χρησιμοποιώ ως ελεγχοσυνάρτηση. Σε ελέγχους της μορφής

$$H_0 : p \leq p_0 \quad \text{vs} \quad H_1 : p > p_0,$$

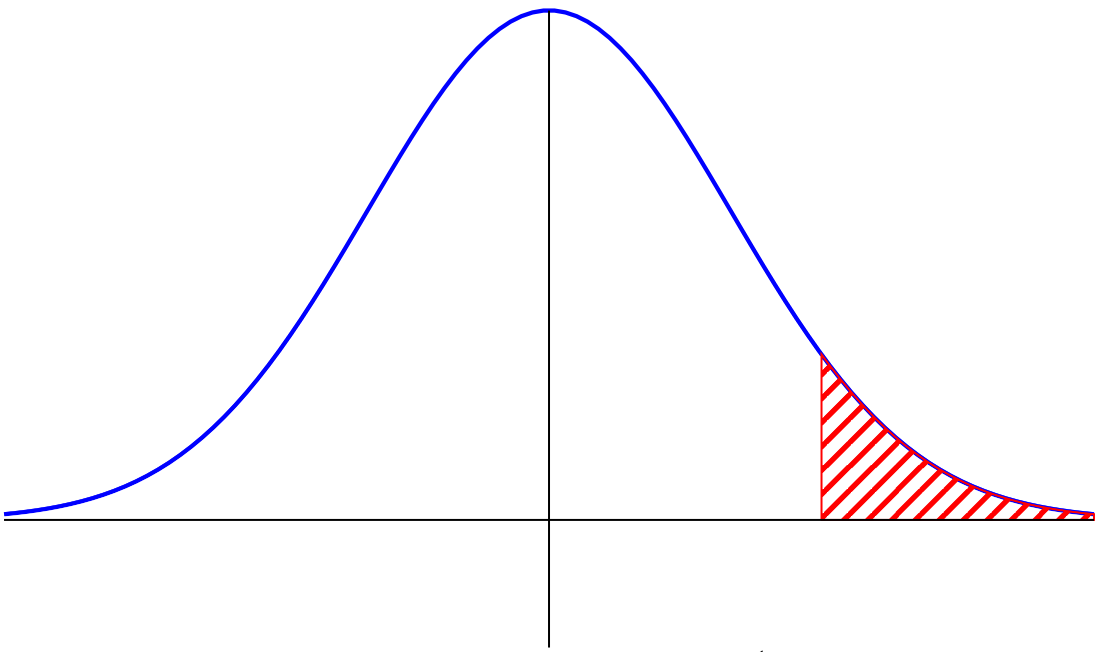
θέλουμε

$$\sup_{0 \leq p \leq p_0} \mathbb{P}_p(X \in C) = a$$

Τελικά επιλέγουμε το ακραίο σημείο και άρα

$$Z = \sqrt{n}\frac{\widehat{p}-\frac{1}{2}}{\sqrt{\frac{1}{2}\cdot\frac{1}{2}}} = 2\sqrt{n}\left(\widehat{p}-\frac{1}{2}\right)$$

και για έλεγχο σε ε.σ.σ. a



$\to$ κοιτάμε τη μορφή της εναλλακτικής $\quad p > p_0 = \frac{1}{2}$ άρα $Z > z_a$

Αν θέλουμε με το $\widehat{p}$ θα πρέπει να πούμε ότι αυτό θα αντιστοιχεί σε

$$c_a = \frac{1}{2} + z_a \cdot \frac{\sqrt{\frac{1}{4}}}{\sqrt{n}} = \frac{1}{2} + z_a \cdot \frac{1}{2\sqrt{n}}$$

$\widehat{p} = \overline{X} > c_a$ ή $Z > z_a$, όπου $c_a = p_0 + z_a \frac{\sqrt{p_0(1-p_0)}}{\sqrt{n}}$

$\to$ Θέλουμε να ελέγξουμε τη νίκη του Β

$$H_0 : p \geq 0.5 \quad \text{vs} \quad H_1 : p < 0.5$$

κρίσιμη περιοχή με έλεγχο σε ε.σ.σ. a

$$Z < z_a \Leftrightarrow \widehat{p} < \frac{1}{2} - z_a \cdot \frac{1}{2\sqrt{n}}$$

Όταν έχουμε

$$\begin{aligned} &H_0 : \theta \leq \theta_0 \quad \text{vs} \quad H_1 : \theta > \theta_0 \\ \text{ή} \quad &H_0 : \theta \geq \theta_0 \quad \text{vs} \quad H_1 : \theta < \theta_0 \end{aligned}$$

και έχουμε μία μόνο παράμετρο για εκτίμηση, τα πράγματα είναι πολύ απλά για τους γνωστούς ελέγχους με κανονική κατανομή.

Αν είχαμε

$$H_0 : p = \frac{1}{2} \quad \text{vs} \quad H_1 : p \neq \frac{1}{2}$$

Τι ελέγχουμε;

Ελέγχουμε αν τελικά τα δεδομένα μας δίνουν ενδείξεις σε ε.σ.σ. a ότι κάποιος από τους Α και Β θα κερδίσει.

Για να μην υπάρχει αντίφαση λέμε ότι ελέγχουμε αυτό που βρίσκεται στην εναλλακτική.

Μπορούμε εδώ να χρησιμοποιήσουμε $(1-a)$-Δ.Ε.?

Ναι! Γιατι? $\rightarrow$ Γιατί είναι αμφίπλευρος

$$H_1 : \theta < \theta_0 \quad \text{ή} \quad \theta > \theta_0 \quad \text{(αλλάζει η φορά της ανίσωσης)}$$

Προσοχή Δεν μπορούμε να χρησιμοποιήσουμε $(1-a)$-Δ.Ε. με μονόπλευρο έλεγχο!

$\rightarrow$ Έλεγχος μέσης τιμής κανονικής με άγνωστη διασπορά

$$H_0 : \mu = \mu_0 \quad \text{vs} \quad H_1 : \mu > \mu_0$$

$$T = \frac{\overline{X} - \mu_0}{\frac{S}{\sqrt{n}}} \overset{\mu_0}{\sim} t_{n-1}, \quad \text{όπου } S = \sqrt{S^2} \text{ και } S^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2$$

και κρίσιμη περιοχή $\rightarrow (t_a, +\infty)$.

Αν χρησιμοποιήσουμε $\overline{X}$ θα έχουμε

$$\mu_0 + t_a \cdot \frac{S}{\sqrt{n}}$$

και απορρίπτουμε την H_0 εντός κρίσιμης περιοχής

$$T > t_a$$

$$\overline{X} > \mu_0 + t_a \frac{S}{\sqrt{n}}$$

ΜΑΘΗΜΑΤΙΚΗ ΣΤΑΤΙΣΤΙΚΗ - QUIZ 3, 11 Δεκέμβρη 2020

<u>ΕΡΩΤΗΣΕΙΣ:</u>

1. Έστω $H_0 : \mu = 0$ vs $H_1 : \mu > 0$ σε τ.δ. $N(\mu, 1)$. Απορρίπτουμε την H_0 σε ε.σ.σ. a, όταν

   ◯ $Z > z_a$ ◯ $Z < z_a$ ◯ $|Z| > z_{a/2}$

2. Έστω $H_0 : \mu = 0$ vs $H_1 : \mu < 0$ σε τ.δ. $N(\mu, 1)$. Απορρίπτουμε την H_0 σε ε.σ.σ. a, όταν

   ◯ $\overline{X} > c_a$ ◯ $\overline{X} < c_a$ ◯ $|\overline{X}| > c_a$

3. Έστω $H_0 : \mu = 0$ vs $H_1 : \mu \neq 0$ σε τ.δ. $N(\mu, 1)$. Απορρίπτουμε την H_0 σε ε.σ.σ. a, όταν

   ◯ $\overline{X} > c_a$ ◯ $\overline{X} < c_a$ ◯ $|\overline{X}| > c_a$

4. Έστω $H_0 : \mu = 0$ vs $H_1 : \mu > 0$ σε τ.δ. $N(\mu, \sigma^2)$, με σ^2 άγνωστο. Απορρίπτουμε την H_0 σε ε.σ.σ. a, όταν

   ◯ $\overline{X} > t_{n-1,a} S/\sqrt{n}$ ◯ $\overline{X} > t_{n,a} S/\sqrt{n}$ ◯ $\overline{X} > t_{n,a} S^2/\sqrt{n}$

5. Έστω $H_0 : \mu = 0$ vs $H_1 : \mu > 0$ σε τ.δ. $N(\mu, \sigma^2)$, με σ^2 άγνωστο. Αν απορρίπτουμε την H_0 για $\overline{X} > z_a S/\sqrt{n}$, τότε το μέγεθος του ελέγχου είναι

   ◯ $< a$ ◯ $> a$ ◯ $= a$

6. Έστω $H_0 : \mu = 0$ vs $H_1 : \mu \neq 0$ σε τ.δ. $N(\mu, \sigma^2)$, με σ^2 άγνωστο. Αν I_{1-a} είναι ένα $(1-a)$-Δ.Ε. για το μ, τότε απορρίπτουμε την H_0 σε ε.σ.σ. a, όταν

   ◯ $0 \in I_{1-a}$ ◯ $\overline{X} \in I_{1-a}$ ◯ $0 \notin I_{1-a}$

7. Αν διεξάγεται ένας έλεγχος υποθέσεων σε ε.σ.σ. a με τη χρήση της p-value, τότε απορρίπτουμε την H_0, όταν

   ◯ p-value < a ◯ p-value = a ◯ p-value > a

8. Μία κρίσιμη περιοχή σε έναν έλεγχο υποθέσεων αντιστοιχεί στην περιοχή

   ◯ αποδοχής της H_0 ◯ απόρριψης της H_0 ◯ τυχαιοποίησης της H_0

9. Έστω $H_0 : \theta = \theta_0$ vs $H_1 : \theta = \theta_1$ με $\theta_0 \neq \theta_1$. Αν $A = \{X \in C\}$ είναι το ενδεχόμενο απόρριψης της H_0, τότε η ισχύς του ελέγχου είναι η

   ◯ $\mathbb{P}_{\theta_0}(A)$ ◯ $\mathbb{P}_{\theta_1}(A) - \mathbb{P}_{\theta_0}(A)$ ◯ $\mathbb{P}_{\theta_1}(A)$

10. Στον παραπάνω έλεγχο υποθέσεων, αν $L_x(\theta)$ είναι η συνάρτηση πιθανοφάνειας, τότε ο Ο.Ι.Ε. μεγέθους a (κατά Neyman-Pearson) διαμορφώνεται με κρίσιμη περιοχή που καθορίζεται από τη σχέση

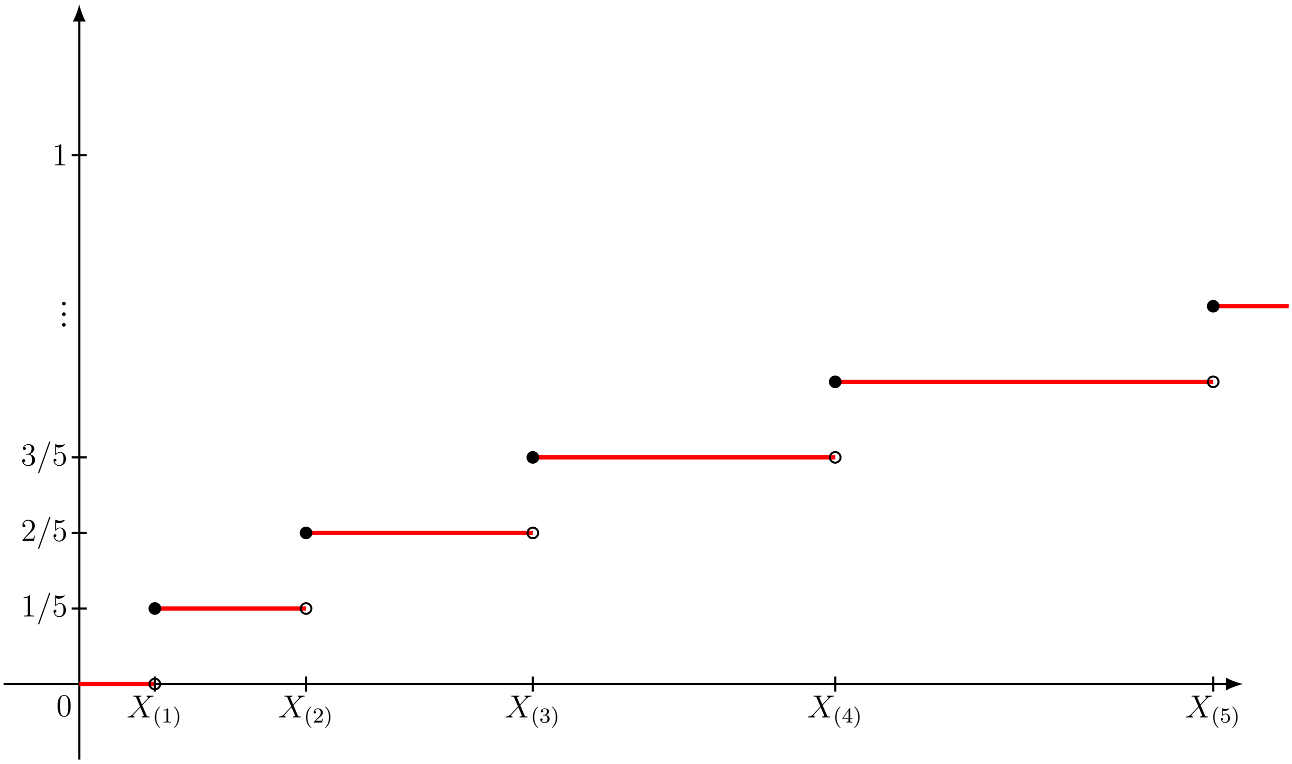
    ◯ $L_x(\theta_1)/L_x(\theta_0) < c$ ◯ $L_x(\theta_1)/L_x(\theta_0) = c$ ◯ $L_x(\theta_1)/L_x(\theta_0) > c$

ΚΑΛΗ ΕΠΙΤΥΧΙΑ!

Απαντήσεις

1	Α
2	Β
3	Γ
4	Α
5	Β
6	Γ
7	Α
8	Β
9	Γ
10	Γ

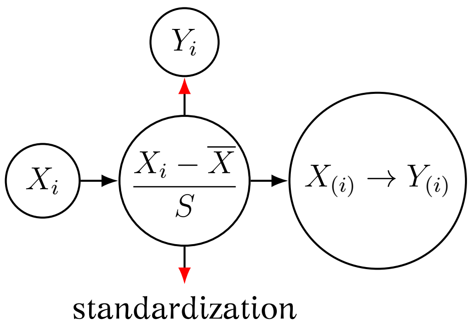
Διάλεξη 27



εμπειρική

$$\boxed{X_{(i)} = q^\star\left(\frac{i-0.5}{5}\right)}$$

$(F^\star)^{-1}$



Γενικά

$$X^\star = \begin{cases} X_{(1)} & \to & \frac{1}{5} \\ X_{(2)} & \to & \frac{1}{5} \\ & \vdots & \\ X_{(5)} & \to & \frac{1}{5} \end{cases}$$

$$F_{X^\star}(x) \equiv F^\star(x) = \begin{cases} 0 & , x < X_{(1)} \\ \frac{1}{5} & , X_{(1)} \le X < X_{(2)} \\ \frac{2}{5} & , X_{(2)} \le X < X_{(3)} \\ & \vdots \\ 1 & , x \ge X_{(5)} \end{cases}$$

εμπειρ. συν. κατ.

Πρόβλημα με 2 τυχαία δείγματα

Θα υποθέσουμε ότι έχουμε 2 ανεξάρτητα τυχαία δείγματα $X_1, \ldots, X_{n_x}$, $\quad Y_1, \ldots, Y_{n_y}$ που προέρχονται από $\mathcal{N}(\mu_x, \sigma_x^2)$, $\quad \mathcal{N}(\mu_y, \sigma_y^2)$ (διαφορετικά άτομα). Θα κάνουμε διάφορους ελέγχους υποθέσεων με διαφορετικές υποθέσεις. Πρώτα μας απασχολούν οι μέσες τιμές.

$\mu_x = \mu_y \Leftrightarrow \mu_x - \mu_y = 0$

$$H_0: \ \mu_x - \mu_y = 0 \ \textbf{vs} \ H_1: \begin{cases} \mu_x - \mu_y > 0 \\ \mu_x - \mu_y < 0 \\ \mu_x - \mu_y \neq 0 \end{cases}$$

όταν

(i) $\sigma_x^2, \ \sigma_y^2$ γνωστά.

(ii) ένα γνωστό και ένα άγνωστο από τα $\sigma_x^2, \ \sigma_y^2$.

(iii) όταν $\sigma_x^2, \ \sigma_y^2$ και τα δύο άγνωστα και

- $\sigma_x^2 = \sigma_y^2 = \sigma^2$

  ή

- γενικά (χωρίς κατ' ανάγκη ισότητα)
  Behrens-Fisher πρόβλημα.

Λύση

Έχουμε $\sigma_x^2 = \sigma_y^2 = \sigma^2$ (μοιράζονται μία κοινή παράμετρο).
Ποιά είναι η ε.μ.π.;

Εκτιμήτρια $\overline{X} - \overline{Y}$ εκτιμά το $\mu_x - \mu_y$

$$\left.\begin{aligned} \overline{X} &\sim \mathcal{N}\left(\mu_x, \frac{\sigma^2}{n_x}\right) \\ \overline{Y} &\sim \mathcal{N}\left(\mu_y, \frac{\sigma^2}{n_y}\right) \end{aligned}\right\} \xRightarrow{\text{ανεξ. τ.δ.}} \overline{X} - \overline{Y} \sim \mathcal{N}\left(\mu_x - \mu_y, \sigma^2\left(\frac{1}{n_x} + \frac{1}{n_y}\right)\right)$$

Η διασπορά είναι άγνωστη και πρέπει να εκτιμηθεί. Διαισθητικά παίρνουμε πληροφορίες και από τα 2 δείγματα και συνδυάζουμε τις 2 εκτιμήτριες:

$$S_x^2 = \frac{1}{n_x - 1} \sum_{i=1}^{n_x} \left(X_i - \overline{X}\right)^2 \tag{26}$$

$$S_y^2 = \frac{1}{n_y - 1} \sum_{i=1}^{n_y} \left(Y_i - \overline{Y}\right)^2 \tag{27}$$

Σταθμίζουμε τις εκτιμήτριες συναρτήσει των βαθμών ελευθερίας.

$$S^2 = \underbrace{\frac{n_x - 1}{n_x + n_y - 2}}_{w_x} S_x^2 + \underbrace{\frac{n_y - 1}{n_x + n_y - 2}}_{w_y} S_y^2 \qquad (w_x + w_y = 1) \tag{28}$$

pooled variance

Από τις (26) και (27):

$$\left.\begin{aligned} \frac{(n_x - 1)S_x^2}{\sigma^2} &\sim \chi^2_{n_x - 1} \\ \frac{(n_y - 1)S_y^2}{\sigma^2} &\sim \chi^2_{n_y - 1} \end{aligned}\right\} \xRightarrow{+\text{ανεξ.}} \frac{(n_x - 1)S_x^2 + (n_y - 1)S_y^2}{\sigma^2} \sim \chi^2_{n_x + n_y - 2}$$

και με την (28) συνεπάγεται:

$$\frac{(n_x + n_y - 2)S^2}{\sigma^2} \overset{\sigma^2}{\sim} \chi^2_{n_x + n_y - 2} \tag{29}$$

Τελικά αν θέσουμε

$$T = \frac{\overline{X} - \overline{Y} - (\mu_x - \mu_y)}{S\sqrt{\frac{1}{n_x} + \frac{1}{n_y}}} \sim t_{n_x + n_y - 2}$$

όπου $S = \sqrt{S^2}$

Ένας έλεγχος μεγέθους a προκύπτει ως:

$$R_a = \mathbf{1}_{\left(t_{n_x + n_y - 2, \frac{a}{2}}, +\infty\right)}(|T|)$$

και απορρίπτεται σε ε.σ.σ. a, αν

$$|t| > t_{n_x + n_y - 2, \frac{a}{2}}$$

Οι έλεγχοι αυτοί λέγονται 2-sample t-test. Ισοδύναμα με p-value.

$$\text{p-value: } \mathbb{P}\left(|T| > |t|\right) = 2\mathbb{P}\left(T > |t|\right)$$

Αν p-value$< a$, απορρίπτουμε την H_0.

Παρατήρηση 0.0.36. Ένας έλεγχος αυτής της μορφής γίνεται ισοδύναμα με το $(1-a)$-Δ.Ε. για το $\mu_x - \mu_y$. Εδώ έχουμε τον έλεγχο. Πώς βρίσκουμε το I_{1-a};

$$T = \frac{\overline{X} - \overline{Y} - (\mu_x - \mu_y)}{S\sqrt{\frac{1}{n_x} + \frac{1}{n_y}}} \sim t_{n_x+n_y-2}.$$

αντιπροσ. κάθε έλεγχο $\mu_x - \mu_y = c \quad (c \in \mathbb{R})$ γίνεται με τον ίδιο τρόπο. Απορρίπτεται σε

ε.σ.σ. $a \quad \Leftrightarrow |T| > t_{n_x+n_y-2}$. Ισοδύναμα δεχόμαστε σε ε.σ.σ. $a \Leftrightarrow |T| \leq t_{n_x+n_y-2}$.
Όμως

$$T = \left| \frac{\overline{X} - \overline{Y} - (\mu_x - \mu_y)}{S\sqrt{\frac{1}{n_x} + \frac{1}{n_y}}} \right| \Leftrightarrow \mu_x - \mu_y \in I_{1-a}^{\mu_x - \mu_y}$$

$\left| \dfrac{\widehat{\theta} - \theta}{\widehat{\sigma}_{\widehat{\theta}}} \right| \leq c_a \longrightarrow$ λύνουμε ως προς θ.

Τελικά

$$\boxed{I_{1-a}^{\mu_x - \mu_y} = \overline{X} - \overline{Y} \pm t_{n_x+n_y-2,\frac{a}{2}} S\sqrt{\frac{1}{n_x} + \frac{1}{n_y}}}$$

Διάλεξη 28

Μέσα στους ελέγχους ισότητας μέσων που δόθηκαν, αφήσαμε για μετά την περίπτωση που σ_x^2, σ_y^2, όχι κατ'ανάγκη ίσα. Θα πρέπει να εισάγουμε την F-κατανομή.

Ορισμός 0.0.24. Μία τυχαία μεταβλητή F λέμε ότι ακολουθεί την F-κατανομή (Fisher-Snedecor) με n_1 και n_2 βαθμούς ελευθερίας και γράφουμε $F \sim F_{n_1,n_2}$, αν

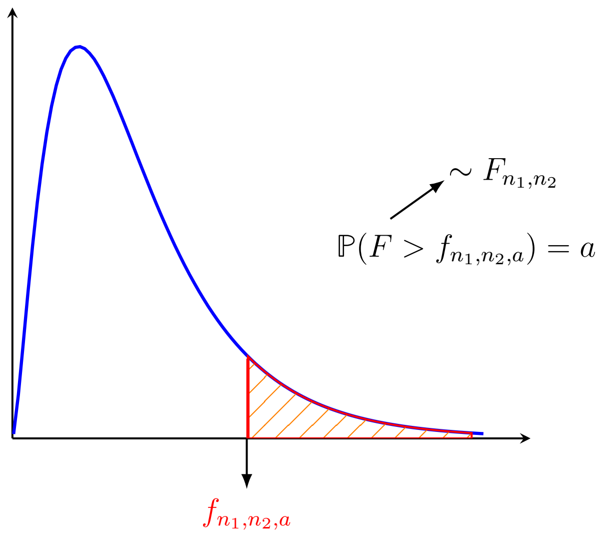
$$F = \frac{\frac{X_1}{n_1}}{\frac{X_2}{n_2}},$$

αν $\begin{matrix} X_1 \sim \chi^2_{n_1} \\ X_2 \sim \chi^2_{n_2} \end{matrix}$ και $X_1 \perp\!\!\!\perp X_2$.

Συμβολικά

$$F_{n_1,n_2} = \frac{\chi^2_{n1/n_1}}{\chi^2_{n_2/n_2}} \quad + \text{ανεξ.}$$

Συμβολισμός ποσοστημορίου: $f_{n_1,n_2,a}$ ή $f_{n_1,n_2;a}$



Άσκηση (για τους μπλε)

Wikipedia: Θα δείτε τη σ.π.π. της $f_{n_1,n_2} \rightarrow$ για (n_1, n_2). Πάτε στην R και αναπαραστήστε τις καμπύλες (+ δικές σας επιλογές).

Απλές Ιδιότητες

(i) $T \sim t_n \Rightarrow \boxed{F = T^2 \sim F_{1,n}}$

Απόδειξη:

$$t_n = \frac{\mathcal{N}(0,1)}{\sqrt{\frac{\chi_n^2}{n}}} \quad + \text{ανεξ.} \Rightarrow$$

$$t_n^2 = \frac{\mathcal{N}(0,1)}{\frac{\chi_n^2}{n}} \quad + \text{ανεξ.} \Rightarrow$$

$$t_n^2 = \frac{\chi_{1/1}^2}{\frac{\chi_n^2}{n}} \quad + \text{ανεξ.} \equiv F_{1,n} \qquad \left(\mathcal{N}^2(0,1) \equiv \chi_1^2\right) \blacksquare$$

Άσκηση (για τους κόκκινους)

Βρείτε τη σ.π.π. της $t_n \rightarrow$ από το πηλίκο.
Βρείτε τη σ.π.π. του T όπου $T \sim t_n$ και επαληθεύστε ότι συμπίπτει με $F_{1,n}$.

(ii) $F \sim F_{n_1,n_2} \Rightarrow \dfrac{1}{F} \sim F_{n_2,n_1}$

Απόδειξη:

$$F_{n_1,n_2} = \frac{\chi_{n_1/n_1}^2}{\chi_{n_2/n_2}^2} \quad + \text{ανεξ.} \Rightarrow$$

$\boxed{\text{inverse Fisher} \rightarrow \text{Fisher}}$

$$F_{n_1,n_2}{}^{-1} = \frac{\chi_{n_2/n_2}^2}{\chi_{n_1/n_1}^2} \quad + \text{ανεξ.} \Rightarrow$$

$$\frac{1}{F_{n_1,n_2}}$$

$$F_{n_1,n_2}^{-1} = F_{n_2,n_1} \quad \blacksquare$$

Εφαρμογή: (Έλεγχος ισότητας διασπορών σε 2 ανεξ. τ.δ. $\mathcal{N}(\mu_x, \sigma_x^2), \mathcal{N}(\mu_y, \sigma_y^2)$)

Για να αλλάξουμε $\rightarrow$ θα πραγματοποιήσουμε τον έλεγχο μέσω κατασκευής Δ.Ε.
Περιπτώσεις:

(i) μ_x, μ_y γνωστά.

(ii) Ένα από τα μ_x, μ_y είναι γνωστό και το άλλο άγνωστο.

(iii) μ_x, μ_y και τα δύο άγνωστα.

[Άσκηση: (i) + (ii)]

Λύση (iii):

Θέτουμε $\lambda = \dfrac{\sigma_x^2}{\sigma_y^2}$

$$H_0: \quad \sigma_x^2 = \sigma_y^2 \quad \Leftrightarrow \frac{\sigma_x^2}{\sigma_y^2} = 1 \quad \Leftrightarrow \lambda = 1$$

$$H_1: \quad \sigma_x^2 \neq \sigma_y^2 \quad \Leftrightarrow \frac{\sigma_x^2}{\sigma_y^2} \neq 1 \quad \Leftrightarrow \lambda \neq 1$$

Απορρίπτουμε (την H_0) σε ε.σ.σ. $a \quad \Leftrightarrow \quad 1 \notin I_{1-a}^{\lambda}$
[Δίπλευρος έλεγχος]

Εκτιμήτρια του λ;

$$\lambda = \frac{\sigma_x^2}{\sigma_y^2} = g\left(\sigma_x^2, \sigma_y^2\right)$$

Μέσω $plug-in$: S_x^2 για την σ_x^2, S_y^2 για την σ_y^2.

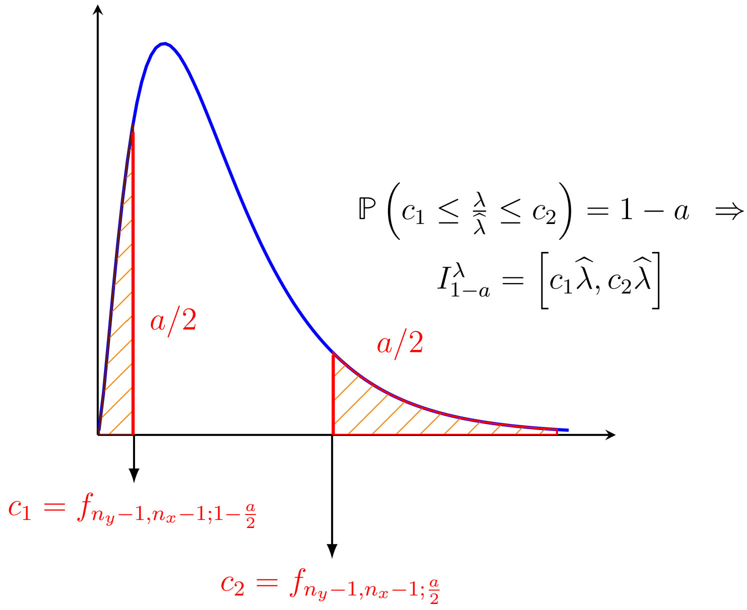
$$\boxed{\widehat{\lambda} = \frac{S_x^2}{S_y^2}}$$

$$S_x^2 = \frac{1}{n_x - 1}\sum_{i=1}^{n_x}\left(X_i - \overline{X}\right)^2$$

$$S_y^2 = \frac{1}{n_y - 1}\sum_{i=1}^{n_y}\left(Y_i - \overline{Y}\right)^2$$

$$\left.\begin{array}{c} \dfrac{(n_x-1)S_x^2}{\sigma_x^2} \sim \chi^2_{n_x-1} \\[2ex] \dfrac{(n_y-1)S_y^2}{\sigma_y^2} \sim \chi^2_{n_y-1} \\[2ex] S_x^2 \perp\!\!\!\perp S_y^2 \\ \text{(διότι ανεξ. τ.δ.)} \end{array}\right\} \Rightarrow \begin{array}{l} \dfrac{S_x^2}{\sigma_x^2} \sim \dfrac{1}{n_x-1}\chi^2_{n_x-1} \\[2ex] \dfrac{S_y^2}{\sigma_y^2} \sim \dfrac{1}{n_y-1}\chi^2_{n_y-1} \\[2ex] + \text{ ανεξ.} \end{array}$$

$$\Rightarrow \frac{S_x^2/\sigma_x^2}{S_y^2/\sigma_y^2} = \frac{S_x^2/S_y^2}{\sigma_x^2/\sigma_y^2} \Rightarrow \boxed{\frac{\widehat{\lambda}}{\lambda} \sim F_{n_x-1, n_y-1}} \Rightarrow \boxed{\frac{\lambda}{\widehat{\lambda}} = F_{n_y-1, n_x-1}}$$



Προσοχή!

Στην πράξη η εφαρμογή αυτού του ελέγχου συνιστάται μόνο στην περίπτωση που είμαστε σίγουροι για την κανονικότητα των δεδομένων. Ο έλεγχος αυτός δεν είναι ευσταθής (robust) όταν αποκλίνουμε από κανονική κατανομή και επομένως γενικά συνιστώνται άλλοι έλεγχοι για ισότητα διασπορών, π.χ. ο έλεγχος Levene.
Επιπλέον, δεν πρέπει να προηγείται ενός ελέγχου μέσων τιμών (π.χ. ισότητας μέσων τιμών σε 2 ανεξ. τ.δ.). όταν οι διασπορές είναι άγνωστες και εμείς θα θέλαμε να ελέγξουμε αν είναι ίσες, ώστε να επιλέξουμε ανάμεσα σε έλεγχο μέσων τιμών με ίσες ή όχι κατ'ανάγκη ίσες μέσες τιμές.

ΔΕΝ ΣΥΝΙΣΤΑΤΑΙ ΑΥΤΗ Η ΠΡΑΚΤΙΚΗ
[Π.χ. Zimmerman; A note on preliminary tests of equality of variancies; 2004.]

Aspin-Welch Unequal Variance t-test

Εδώ παρουσιάζεται η περίπτωση που δεν υποθέτουμε κατ'ανάγκη ότι $\sigma_x^2 = \sigma_y^2$. Η περίπτωση αυτή συνδέεται με το πρόβλημα Behrens-Fisher. Οι πιο συνηθισμένοι έλεγχοι είναι της μορφής:

$$H_0: \quad \mu_x = \mu_y \qquad (\mu_x - \mu_y = 0)$$

vs

$$H_1: \begin{cases} \mu_x \neq \mu_y & (\mu_x - \mu_y \neq 0) \\ \mu_x < \mu_y & (\mu_x - \mu_y < 0) \\ \mu_x > \mu_y & (\mu_x - \mu_y > 0) \end{cases}$$

Το πρώτο μέρος εύρεσης κατάλληλης ελεγχοσυνάρτησης είναι παρόμοιο με προφανείς τροποποιήσεις.

Αν $\mu = \mu_x - \mu_y \Rightarrow \widehat{\mu} = \overline{X} - \overline{Y} \qquad (\overline{X}_{n_x}, \overline{Y}_{n_y})$

$$\overline{X} - \overline{Y} \sim \mathcal{N}\left(\mu_x - \mu_y, \frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}\right)$$
$$\overset{H_0\ (\mu_x=\mu_y)}{\sim} \mathcal{N}\left(0, \frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}\right)$$

Άρα

$$\boxed{T = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{S_x^2}{n_x} + \frac{S_y^2}{n_y}}}} \longrightarrow$$

Φαίνεται αντίστοιχη ελεγχοσυνάρτηση με τις προηγούμενες [όμως δε βγαίνει *Student*!]

$$\boxed{\sim \text{Behrens-Fisher}}$$

Η εύρεση της ακριβής κατανομής της ελεγχοσυνάρτησης αυτής είναι το πρόβλημα Behrens-Fisher!

Έχουν βρεθεί προσεγγιστικές λύσεις. Συγκεκριμένα, χρησιμοποιείται πιο συχνά $T \approx t_d$, όπου d είναι τυχαία μεταβλητή!

ακολουθεί προσεγγιστικά

$$d = \frac{\left(S_x^2/n_x + S_y^2/n_y\right)^2}{\frac{1}{n_x-1}\left(S_x^2/n_x\right)^2 + \frac{1}{n_y-1}\left(S_y^2/n_y\right)^2}$$

Συνήθως χρησιμοποιείται ο πλησιέστερος ακέραιος στη θετική πραγματική τιμή του d που δίνεται στη δαφάνεια 14.

Τελικά, χρησιμοποιώντας αυτήν την προσέγγιση διεξάγεται το Welch t-test, όπου απορρίπτεται η H_0 σε ε.σ.σ. a, αν $|t| > t_{d,\frac{a}{2}}$ ή p-value$< a$.

Παρατήρηση 0.0.37. Αν δεν υπάρχει ισχυρή πεποίθηση ότι υπάρχει ισότητα διασπορών στα 2 τυχαία δείγματα, τότε συνίσταται η χρήση του Welch t-test, αντί του κλασικού t-test με $\sigma_x^2 = \sigma_y^2 = \sigma^2$ και έχει δειχθεί πειραματικά ότι δεν πρέπει να γίνεται προκαταρκτικά έλεγχος ισότητας διασποράς. Αν υπάρχουν αμφιβολίες για $\sigma_x^2 = \sigma_y^2 \Rightarrow$ Welch t-test.

Ο έλεγχος Levene για ισότητα διασπορών

Ο έλεγχος αυτός χρησιμοποιείται για να ελεγχθεί αν k δείγματα (ομάδες) έχουν ίσες διασπορές. Αναφέρεται ως έλεγχος ομογένειας και εμφανίζεται στην ανάλυση διασποράς (ANOVA)→ Analysis Of Variance. Ο έλεγχος Levene είναι λιγότερο ευαίσθητος από τον έλεγχο Bartlett όταν τα δεδομένα είναι μη κανονικά, αλλά ο τελευταίος είναι ισχυρότερος όταν είμαστε κοντά σε κανονική κατανομή.

Για τον Levene:
Για k-ομάδες:

$H_0: \quad \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2$

vs

$H_1: \qquad \sigma_i^2 \neq \sigma_j^2 \qquad$ για τουλάχιστον ένα ζεύγος i, j με $i \neq j$

[Δείτε ANOVA για γενική μορφή, εμείς τον εξειδικεύουμε για $k = 2$.]

$H_0: \quad \sigma_x^2 = \sigma_y^2$

vs

$H_1: \quad \sigma_x^2 \neq \sigma_y^2$

Ελεγχοσυνάρτηση: (για $k = 2$)

Δημιουργούνται κάποια σκορ

$$\begin{array}{l} Z_i^x = \left|X_i - \overline{X}\right|, \quad i = 1, 2, \ldots, n_x \\ Z - i^y = \left|Y_i - \overline{Y}\right|, \quad i = 1, 2, \ldots, n_y \end{array}$$

$$\begin{aligned} \overline{Z}^x &= \left(Z_1^x + \ldots + Z_{n_x}^x\right)/n_x \quad && \to \text{μέσος του } x - group \\ \overline{Z}^y &= \left(Z_1^y + \ldots + Z_{n_y}^y\right)/n_y \quad && \to \text{μέσος του } y - group \\ \overline{Z}^{x,y} &= \frac{Z_1^x + \ldots + Z_{n_x}^x + Z_1^y + \ldots + Z_{ny}^y}{n_x + n_y} \quad && \to \text{μέσος όλων } xy - group \end{aligned}$$

$$W = (n_x + n_y - 2)\frac{n_x\left(\overline{Z}^x - \overline{Z}^{x,y}\right)^2 + n_y\left(\overline{Z}^y - \overline{Z}^{x,y}\right)^2}{\sum\limits_{i=1}^{n_x}\left(Z_i^x - \overline{Z}^x\right)^2 + \sum\limits_{i=1}^{n_y}\left(Z_i^y - \overline{Z}^y\right)^2}$$

Προκύπτει ότι κάτω από την $H_0: \quad \sigma_x^2 = \sigma_y^2$

$$W \approx F_{1, n_x + n_y - 2}$$

Άσκηση

1. Μέσα από προσομοίωση `R` να γίνει qq-plot σύγκρισης εμπειρικής vs θεωρητικής.
2. Μπορείτε ν.δ.ο. για σχετικά μεγάλα n_x, n_y: $W_{n_x, n_y} \overset{d}{\approx} F_{1, n_x + n_y - 2}$;

Διάλεξη 29

Αρχή της επάρκειας: Έστω $X = (X_1, \ldots, X_n)$ τ.δ. και $T = T(x)$ μία σ.σ. Τότε αν οποιοδήποτε συμπέρασμα για το θ, εξαρτάται από το τ.δ. X, μόνο μέσω του $T(X)$, τότε η $T(X)$ είναι επαρκής σ.σ.

Ορισμός 0.0.25. (τυπικός ορισμός Επάρκειας)
Έστω $X = (X_1, \ldots, X_n)$ τ.δ. από παραμετρική οικογένεια κατανομών $\mathcal{F} = \{F_\theta : \theta \in \Theta\}$. Μία σ.σ. $T = T(X)$ λέγεται επαρκής σ.σ. για το θ ή για την οικογένεια κατανομών $\mathcal{F}$, αν η δεσμευμένη κατανομή της X δοθείσης της $T = t$, για όλα τα "επιτρεπτά" t, είναι ανεξάρτητη του θ.

<u>**Σχόλιο**</u> Με άλλα λόγια σημαίνει ότι αν ξέρω την τιμή t, οποιαδήποτε άλλη γνώση δε μας δίνει πληροφορία για το θ.
Επιπλέον μπορούμε να γράφουμε $[X|\ T = t]$ για την τ.μ. που προκύπτει αν δεσμεύσουμε στην τιμή $T = t$.
Επίσης αν η T είναι επαρκής σ.σ. τότε από το παραπάνω, οποιαδήποτα άλλη $U = U(X)$, αν τη δεσμεύσουμε στην T, θα έχει κατανομή ανεξάρτητη του θ και θα γράφουμε $[U(X)|\ T = t]$ έχει κατανομή ανεξάρτητη του θ.

Παρατήρηση 0.0.38.

① Το ίδιο το τ.δ. X είναι προφανώς επαρκής σ.σ. για το θ, αφού

$$[X|X = x] = x$$

$\Rightarrow$ η τιμή σταθεροποιείται και άρα είναι ανεξάρτητο του θ (τυχαίο διάνυσμα που είναι σταθερό).

Όμως το ενδιαφέρον εστιάζεται σε σ.σ. που επιφέρουν ελάττωση της διάστασης των δεδομένων. Αναζητούμε $T(X) : \dim T(X) < \dim X = n$ ($\Rightarrow$ ελάττωση της διάστασης των δεδομένων).

② Η κατανομή της $[X|\ T = t]$ είναι ανεξάρτητη του θ, εμπεριέχει ανεξαρτησία του θ και στο στήριγμα της παραπάνω,
π.χ. αν υποθέσουμε ότι έχουμε μια έκφραση της σ.π.π. της $[X|\ T = t]$, τότε ελέγχουμε και τη συναρτησιακή της μορφή όταν παίρνει θετικές τιμές, αλλά και το σύνολο που τις παίρνει. Πρέπει και τα δύο να είναι ανεξάρτητα του θ.

Παράδειγμα 0.0.42. Έστω $X = (X_1, X_2)$ τ.δ. από Be$(p), 0 \leq p \leq 1$. Θέτουμε $U(X) = X_1$ και $T(X) = X_1 + X_2$.
Ερώτημα: Να εξεταστεί αν είναι επαρκείς σ.σ.

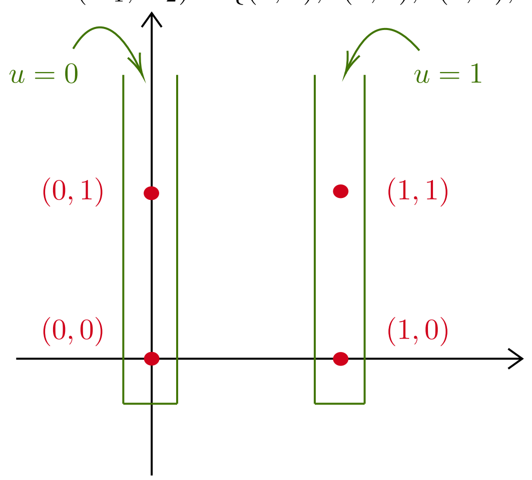
Λύση

Διαισθητικά: Φαίνεται το $X_1 + X_2$ να είναι επαρκής σ.σ. (συνδέεται άμεσα με το $\overline{X}$ που εκτιμά "αποτελεσματικά" το p), ενώ το X_1 δεν περιλαμβάνει πληροφορία από τη 2η παρατήρηση $\Rightarrow$ απώλεια πληροφορίας για το p.

Θα τις εξετάσουμε με τον ορισμό της επάρκειας.

α΄) Εξετάζουμε την $U = X_1$. Πρέπει $[X|\ \underset{\substack{\downarrow \\ \text{επιτρεπτά } u}}{U = u}]$ να έχει κατανομή ανεξάρτητη του p

$X = (X_1, X_2) \in \{(0,0),\ (0,1),\ (1,0),\ (1,1)\}$



Είναι φανερό ότι $\underset{\substack{\downarrow \\ \text{επιτρεπτές του } U}}{U \in \{0,1\}}$.

$\rightarrow$ αν $u = 0$ γράφουμε $S_{x|0} = \{(0,0),\ (0,1)\}$

$\rightarrow$ αν $u = 1$ γράφουμε $S_{x|1} = \{(1,0),\ (1,1)\}$

Έστω $(x_1, x_2) \in \{0,1\}^2$ $[= \{0,1\} \times \{0,1\}]$ τότε

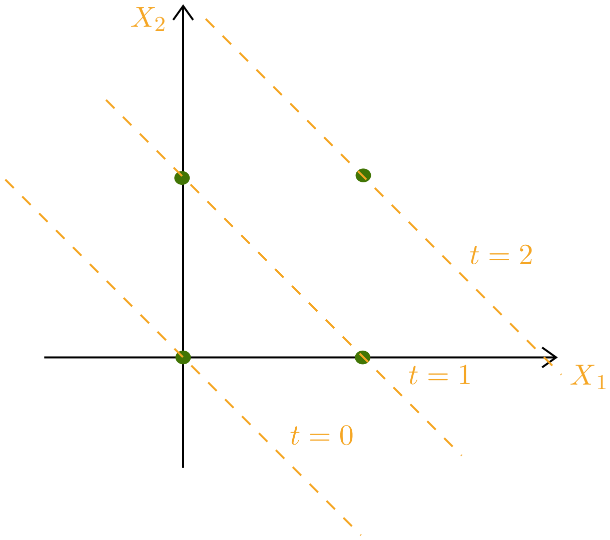
$$\mathbb{P}\left(X_1 = x_1,\ X_2 = x_2 \mid U = u\right) = \mathbb{P}\left(X_1 = x_1,\ X_2 = x_2 \mid X_1 = u\right) =$$

$$\mathbb{P}\left(X_1 = x_1 \mid X_1 = u\right)\mathbb{P}\left(X_2 = x_2 \mid X_1 = u\right) \overset{X_1 \perp\!\!\!\perp X_2}{=}$$

$$\mathbb{P}(X_2 = x_2) \cdot \mathbb{1}_{\{X_1 = u\}} = \underbrace{p^{x_2}(1-p)^{1-x_2}}_{\downarrow} \cdot \mathbb{1}_{\{X_1 = u\}}$$

εξαρτάται από το $p \Rightarrow$ δεν είναι επαρκής σ.σ.

β′) Εξετάζουμε την επάρκεια της T.



$$T = X_1 + X_2 \text{ άρα } T \in \{0,1,2\}$$

$$S_{x|t} = \{x \in \{0,1\}^2 : x_1 + x_2 = t\}$$

$$S_{x|0} = \{(0,0)\}$$

$$S_{x|1} = \{(0,1),\ (1,0)\}$$

$$S_{x|2} = \{(1,1)\}$$

Το στήριγμα $S_{x|t}$ είναι ανεξάρτητο του p.

$$\mathbb{P}(X_1 = x_1,\ X_2 = x_2 \mid T = t) = \frac{\mathbb{P}(X_1 = x_1,\ X_2 = x_2,\ X_1 + X_2 = t)}{\mathbb{P}(X_1 + X_2 = t)} \overset{X_1+X_2\sim\text{Bin}(2,p)}{=\!=\!=\!=\!=}$$

$$= \frac{\mathbb{P}(X_1 = x_1,\ X_2 = x_2)}{\mathbb{P}(X_1 + X_2 = t)} \cdot \mathbb{1}_{\{x_1+x_2=t\}}$$

$$= \frac{p^{x_1}(1-p)^{1-x_1} p^{x_2}(1-p)^{1-x_2}}{\binom{2}{t} p^t (1-p)^{2-t}} \cdot \mathbb{1}_{\{x_1+x_2=t\}}$$

$$= \frac{p^t(1-p)^{2-t}}{\binom{2}{t} p^t (1-p)^{2-t}} \cdot \mathbb{1}_{\{x_1+x_2=t\}}$$

$$= \frac{1}{\binom{2}{t}} \cdot \mathbb{1}_{\{x_1+x_2=t\}}$$

$$\Longrightarrow [X \mid T = t] \text{ έχει κατανομή ανεξ. του } p.$$

Άρα

$$[X \mid T=t] \sim \text{discr.Unif}\big(\{x \in \{0,1\}^2 : x_1 + x_2 = t\}\big)$$

με $t \in \{0,1,2\} \to 3$ διαφορετικές κατανομές.

Γενίκευση για μέγεθος δείγματος n

Αν X_i τ.δ. Be(p), τότε η $T = \sum_{i=1}^{n} X_i$ είναι επαρκής σ.σ. για το p και μάλιστα:

$$[X \mid T=t] \sim \text{discr.Unif}\left(\left\{x \in \{0,1\}^n : \sum_{i=1}^{n} X_i = t\right\}\right), \quad t = 0,1,2,\ldots,n$$

και

$$\mathbb{P}(X = x \mid T = t) = \frac{1}{\binom{n}{t}}, \quad x \in S_{x|t},$$

όπου $\binom{n}{t} = \#$ σημείων $\sum_{i=1}^{n} X_i = t$.

Παρατήρηση 0.0.39. Προφανώς μια επαρκής σ.σ. δεν είναι κατ' ανάγκη εκτιμήτρια μιας παραμέτρου,
π.χ. πριν είχαμε $T = \sum_{i=1}^{n} X_i \in \{0,1,\ldots\}$ ενώ μία $\widehat{p} \in [0,1]$.

Χαρακτηρισμός Επάρκειας

Θεώρημα 0.0.15. (Παραγοντικό Κριτήριο Neyman) [Π.Κ.Ν.]

Έστω $X = (X_1,\ldots,X_n)$ ένα τ.δ. από κατανομή με σ.π./ σ.π.π. $f_\theta(x_i)$, όπου $\theta = (\theta_1,\ldots,\theta_s)$.

Τότε η σ.σ. $T = T(X)$ είναι επαρκής σ.σ. για το θ

$\Leftrightarrow \exists$ συναρτήσεις $g(t\ ;\ \theta)$ και $h(x) : f_\theta(x) = g(T(x)\ ;\ \theta)\cdot h(x), \quad \forall x \in \mathbb{R}^n,\ \forall \theta \in \Theta.$

Απόδειξη (για διακριτές τ.μ.)

Κατ' αρχήν λόγω του ότι οι X_i είναι διακριτές τ.μ., έχουμε ότι και η $T = T(X)$ είναι και αυτή διακριτή και ισχύει ότι

$$S_{T;\theta} = T(S_{X;\theta})$$

Άρα συμπεραίνουμε ότι $T(x) \in S_{T;\theta} \Leftrightarrow x \in S_{X;\theta}$.

Απόδειξη

$\Rightarrow$) [T επαρκής $\Rightarrow f_\theta(x) = \ldots$]

Έστω $x \in \mathbb{R}^n : T(x) \in S_{T;\theta}$ $(\Leftrightarrow x \in S_{x;\theta})$.
Τότε

$$f_\theta(x) = \mathbb{P}(X = x) = \mathbb{P}_\theta(X = x,\ T(X) = T(x)) = \dots$$

$$\left[\{X \underset{A}{=} x\} \subset \{T(X) \underset{B}{=} T(x)\} \ \Rightarrow\ \{X \underset{A}{=} x\} = \{\underset{A}{X = x,} \underset{\cap}{\ } \underset{B}{T(X) = T(x)}\}\right]$$

$$\dots = \mathbb{P}_\theta(T(X) = T(x)) \cdot \mathbb{P}_\theta(X = x \mid T(X) = T(x)) \overset{\text{υποθ. } T \text{ επαρκής}}{=}$$

$$\overset{\mathbb{P}_{\not\theta}}{=} \mathbb{P}_\theta(T(X) = T(x)) \cdot \mathbb{P}(X = x \mid T(X) = T(x)) = g^*(T(x)\ ;\ \theta) \cdot h(x)$$

Άρα $\forall x \in \mathbb{R}^n$,

$$f_\theta(x) = g^*(T(x)\ ;\ \theta) \cdot h(x) \cdot \mathbb{1}_{S_{T;\theta}}(T(x))$$

$$= g(T(x)\ ;\ \theta) \cdot h(x)$$

όπου $g(T(x)\ ;\ \theta) = g^*(T(x)\ ;\ \theta) \cdot \mathbb{1}_{S_{T;\theta}}(T(x))$.

Άρα δείξαμε το $(\Rightarrow)$ και το $(\Leftarrow)$ στο επόμενο μάθημα.

$$\left[\mathbb{1}_{S_{T;\theta}}(T(x)) = \mathbb{1}_{S_{X;\theta}}(x)\right]$$

Διάλεξη 30

Π.Κ.Ν.: T επαρκής για το $\theta \Leftrightarrow f_\theta(x) = g\left(T(x)\ ;\ \theta\right) \cdot h(x),\ \forall x \in \mathbb{R}^n$

Απόδειξη

($\Rightarrow$) έγινε (για τη διακριτή περίπτωση)

($\Leftarrow$) Θέλουμε ν.δ.ο. $\mathbb{P}_\theta\left(X = x \mid T(X) = t\right)$ είναι ανεξάρτητο του θ

$$\mathbb{P}_\theta\left(X = x \mid T(x) = t\right) = \frac{\mathbb{P}_\theta\left(X = x,\ T(X) = t\right)}{\mathbb{P}_\theta(T(X) = t)} = \frac{\mathbb{P}_\theta\left(X = x,\ T(X) = x,\ T(X) = t\right)}{\mathbb{P}_\theta(T(X) = t)}$$

$$= \frac{\mathbb{P}_\theta\left(X = x,\ T(X) = T(x)\right)}{\mathbb{P}_\theta\left(T(X) = t\right)} \cdot \mathbb{1}_{\{T(x)=t\}}$$

$$= \frac{\mathbb{P}_\theta(X = x)}{\mathbb{P}_\theta(T(X) = t)} \cdot \mathbb{1}_{\{T(x)=t\}}$$

$$= \frac{g(\overset{=t}{T(x)}\ ;\ \theta) \cdot h(x)}{\sum\limits_{x:T(x)=t} \underbrace{\mathbb{P}_\theta(X = x)}_{g(t\ ;\ \theta)\cdot h(x)}} \cdot \mathbb{1}_{\{T(x)=t\}}, \quad \text{κάνοντας την απλοποίηση}$$

$$= \frac{h(x)}{\sum\limits_{x:T(x)=t} h(x)} \cdot \mathbb{1}_{\{T(x)=t\}} \Rightarrow \text{ανεξάρτητο του } \theta.$$

Τελικά έχουμε $\mathbb{P}_\theta\left(X = x \mid T(X) = t\right)$ είναι ανεξ. του $\theta \Rightarrow$ η T είναι επαρκής για το θ.

Παρατήρηση 0.0.40. Από το Π.Κ.Ν. παίρνουμε και ένα κριτήριο μη επάρκειας μιας ε.σ.σ. T.

$$T \text{ επαρκής για το } \theta \overset{\text{Π.Κ.Ν.}}{\Leftrightarrow} f_\theta(x) = g(T(x)\ ;\ \theta) \cdot h(x),\ \forall x \in \mathbb{R}^n$$

Αν πάρουμε x, y με $x \neq y:\ T(x) = T(y)$ με $f_\theta(y) > 0$, τότε

$$\frac{f_\theta(x)}{f_\theta(y)} = \frac{\cancel{g\left(T(x)\ ;\ \theta\right)} \cdot h(x)}{\cancel{g\left(T(y)\ ;\ \theta\right)} \cdot h(y)} = \frac{h(x)}{h(y)}$$

αρά ανεξάρτητο του θ.

Συμπεραίνουμε ότι αν μια U είναι υποψήφια επαρκής σ.σ. και

$\exists\ x, y \in \mathbb{R}^n$ με $U(x) = U(y)$ και $\dfrac{f_\theta(x)}{f_\theta(y)}$ εξαρτάται από το θ,

τότε η U δεν είναι επαρκής σ.σ.

Παράδειγμα 0.0.43. Έστω $X_1, \ldots, X_n$ τ.δ. $\mathcal{N}(\mu, \sigma^2)$, όπου σ^2 γνωστό.

(i) Να βρεθεί επαρκής σ.σ. για το μ.

(ii) Ν.δ.ο. η $U = \sum_{i=1}^{n} X_i^2$ δεν είναι επαρκής για το μ.

<u>Λύση</u>

(i)

$$f_\mu(X) \equiv L(\mu) = \left(2\pi\sigma^2\right)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2}$$

$$= \underbrace{\left(2\pi\sigma^2\right)^{-\frac{n}{2}} e^{\frac{\mu}{\sigma^2}\sum_{i=1}^{n} X_i - \frac{n}{2\sigma^2}\mu^2}}_{g\left(\sum_{i=1}^{n} X_i\ ;\ \mu\right)} \cdot \underbrace{e^{-\frac{\sum_{i=1}^{n} X_i^2}{2\sigma^2}}}_{h(x)}$$

(η παραγοντοποίηση δεν είναι μοναδική)

όπου η επαρκής σ.σ. κατά *Neyman* είναι η $T(X) = \sum_{i=1}^{n} X_i$

Παρατήρηση 0.0.41. $\Rightarrow$ ο δ.μ. $\overline{X}$ είναι επίσης επαρκής σ.σ., αφού

$$g\left(\sum_{i=1}^{n} X_i\ ;\ \mu\right) = \left(2\pi\sigma^2\right)^{-\frac{n}{2}} e^{\frac{\mu}{\sigma^2}\sum_{i=1}^{n} X_i - \frac{n}{2\sigma^2}\mu^2}$$

$$= \left(2\pi\sigma^2\right)^{-\frac{n}{2}} e^{\frac{n\mu}{\sigma^2}\cdot\frac{\sum_{i=1}^{n} X_i}{n} - \frac{n}{2\sigma^2}\mu^2} = g^*\left(\frac{\sum_{i=1}^{n} X_i}{n}\ ;\ \mu\right) = g^*\left(\overline{X}\ ;\ \mu\right)$$

όπου η επαρκής σ.σ. κατά *Neyman* είναι η $T(X) = \overline{X}$

(ii) Θέλουμε ν.δ.ο. η $U = \sum_{i=1}^{n} X_i^2$ δεν είναι επαρκής για το μ.

Αναζητούμε x, y με $x \neq y$ και $U(x) = U(y)$ έτσι ώστε $\dfrac{f_\mu(x)}{f_\mu(y)}$ να εξαρτάται από το μ.

Πρακτικός κανόνας:

Βρίσκουμε x, y με $x \neq y$, $U(x) = U(y)$ και $T(x) \neq T(y)$ (με T επαρκής)

αν εδώ π.χ. $x = (0, \ldots, 0, 1)$, $y = (0, \ldots, 0, -1)$ τότε

$$U(x) = U(y) = 1 \text{ και } T(x) = 1 \neq T(y) = -1$$

οπότε

$$\frac{f_\mu(x)}{f_\mu(y)} = \frac{e^{\frac{\mu}{\sigma^2} - \frac{n}{2\sigma^2}\mu^2}}{e^{-\frac{\mu}{\sigma^2} - \frac{n}{2\sigma^2}\mu^2}} = e^{\frac{2\mu}{\sigma^2}} \to \text{ εξαρτάται από το } \mu$$

$\Rightarrow$ η $U = \sum_{i=1}^{n} X_i^2$ δεν είναι επαρκής για το μ.

Παράδειγμα 0.0.44. Έστω $X_1, \ldots, X_n$ τ.δ. από $\mathcal{U}[a, \beta]$, όπου $-\infty < a < \beta < +\infty$. Να βρεθεί επαρκής σ.σ. για το ζεύγος $\theta = (a, \beta)$.

Λύση

Έστω $x \in \mathbb{R}^n$. Τότε

$$f_\theta(x) = \left(\frac{1}{\beta - a}\right)^n \cdot \mathbb{1}_{\left(-\infty, X_{(1)}\right]}(a) \cdot \mathbb{1}_{\left[X_{(n)}, +\infty\right)}(\beta)$$

$$= g\left(T(x)\ ;\ \theta\right) \cdot 1, \quad 1 \equiv h(x)$$

όπου $\theta = (a, \beta)$, $T(X) = (X_{(1)}, X_{(n)})$ με $T(x)$ να είναι η αποτίμηση στο δείγμα

$\overset{\text{Π.Κ.Ν}}{\Rightarrow}$ η $T(X)$ είναι επαρκής σ.σ. για το $\theta = (a, \beta)$.

Παρατήρηση 0.0.42. Αν π.χ. εδώ το β είναι γνωστό, τότε

$$f_\alpha(x) = \underbrace{\left(\frac{1}{\beta - \alpha}\right)^n \cdot \mathbb{1}_{\left(-\infty, X_{(1)}\right]}(\alpha)}_{g\left(X_{(1)}\ ;\ \alpha\right)} \cdot \underbrace{\mathbb{1}_{\left[X_{(n)}, +\infty\right)}(\beta)}_{\substack{\| \\ h(x)}}$$

Από τη $g\left(X_{(1)}\ ;\ \alpha\right) \Rightarrow$ το $X_{(1)}$ είναι επαρκής σ.σ. για το α.

Αντίστοιχα

$$f_\beta(x) = \underbrace{\left(\frac{1}{\beta - \alpha}\right)^n \cdot \mathbb{1}_{\left[X_{(n)}, +\infty\right)}(\beta)}_{g\left(X_{(n)}\ ;\ \beta\right)} \cdot \underbrace{\mathbb{1}_{\left(-\infty, X_{(1)}\right]}(\alpha)}_{\substack{\| \\ h(x)}}$$

Από τη $g\left(X_{(n)}\ ;\ \beta\right) \overset{\text{Π.Κ.Ν}}{\Rightarrow}$ το $X_{(n)}$ είναι επαρκής σ.σ. για το β με το α γνωστό.

Πόρισμα 0.0.16. Αν $X_1, X_2, \ldots X_n$ τ.δ. από $\mathcal{U}[0,\theta]$, όπου εδώ $\alpha = 0$ γνωστό και $\beta = \theta$, τότε το $X_{(n)}$ είναι επαρκής σ.σ. για το θ.

Θυμηθείτε επίσης ότι $X_{(n)}$ είναι και η ε.μ.π. $\Rightarrow$ η ε.μ.π. είναι επαρκής για το θ.

Παρατήρηση 0.0.43. Ο δειγματικός μέσος $\overline{X}_n$ μας είχε βγει μέσω της $2\overline{X}_n$ χειρότερη εκτιμήτρια του θ.

Είναι ο δειγματικός μέσος επαρκής σ.σ.;

Έστω $U = \overline{X}_n$. Θα δείξουμε ότι το U δεν είναι επάρκής σ.σ. για το θ.

Γενικά για $x \neq y$ με $f_\theta(y) > 0$ έχουμε

$$\frac{f_\theta(x)}{f_\theta(y)} = \frac{\cancel{\left(\frac{1}{\theta}\right)^n} \cdot \mathbb{1}_{[X(n),+\infty)}(\theta)}{\cancel{\left(\frac{1}{\theta}\right)^n} \cdot \mathbb{1}_{[Y(n),+\infty)}(\theta)}$$

με επιπέον συνθήκη ότι πρέπει $\boxed{Y_{(n)} \leq \theta}$.

Παίρνουμε $x, y : \overline{X} = \overline{Y}$ και $X_{(n)} \neq Y_{(n)}$ π.χ. $x = (0, 0, \ldots, 0, 1), \quad y = \left(\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n}, \frac{1}{n}\right)$

και έχουμε $\overline{X} = \overline{Y} = \dfrac{1}{n}$, ενώ $X_{(n)} = 1$ και $Y_{(n)} = \dfrac{1}{n}$, άρα $X_{(n)} \neq Y_{(n)}$.

Αν πάρουμε το $\theta > X_{(n)} = 1 \Rightarrow \dfrac{f_\theta(x)}{f_\theta(y)} = \dfrac{1}{1} = 1$.

Αν όμως

$$\underset{\substack{\| \\ Y_{(n)}}}{\frac{1}{n}} \leq \theta < \underset{\substack{\| \\ X_{(n)}}}{1}, \quad \text{τότε} \quad \frac{f_\theta(x)}{f_\theta(y)} = \frac{0}{1}$$

άρα η τιμή του λόγου $\dfrac{f_\theta(x)}{f_\theta(y)}$, για τα επιτρεπτά $\theta : \theta \geq \dfrac{1}{n}$, εξαρτάται από την τιμή του $\theta \Rightarrow$ το $\overline{X}_n$ δεν είναι επαρκής σ.σ..

Διάλεξη 31

Πορίσματα του Π.Κ.Ν. [Fisher- Neyman]

Πόρισμα 0.0.17. Έστω T επαρκής σ.σ. για το θ. Τότε

(i) αν $T = h(U)$, τότε η U είναι επαρκής για το θ.

(ii) αν $\theta = \omega(\phi)$, τότε η T είναι επαρκής για το ϕ.

Απόδειξη

(i) Από το Π.Κ.Ν.

$$f_\theta(x) = g(T(x)\ ;\ \theta) \cdot h(x) \overset{\text{υπόθ.}}{=} g(h(U(x))\ ;\ \theta) \cdot h(x)$$

$$= g^*(U(x)\ ;\ \theta) \cdot h(x)$$

$\Longrightarrow$ η $U(X)$ είναι επαρκής σ.σ. για το θ.

(ii)

$$f_\theta(x) = g(T(x)\ ;\ \theta) \cdot h(x) \overset{\text{υπόθ.}}{=} g(T(x)\ ;\ \omega(\phi)) \cdot h(x)$$

$$= g^*(T(x), \phi) \cdot h(x)$$

$\Longrightarrow$ η $T(X)$ είναι επαρκής για το ϕ από το Π.Κ.Ν.

Παράδειγμα 0.0.45. Έστω $X_1, \ldots, X_n$ τ.δ. από $\mathcal{N}(\mu, \sigma^2)$, σ^2 γνωστό.

Η $T = \sum\limits_{i=1}^{n} X_i$ δείξαμε ότι είναι επαρκής για το $\mu \Rightarrow$

$\overset{(i)}{\Longrightarrow} U = (U_1, U_2) = {}^{\text{π.χ.}} \left(X_1, \sum\limits_{i=2}^{n} X_i\right)$ ή $\left(\sum\limits_{i=1}^{\kappa} X_i, \sum\limits_{i=\kappa+1}^{n} X_i\right)$ είναι επαρκής για το μ,
αφού $T = U_1 + U_2 = h(U_1, U_2) = h(U)$

$\overset{(i)}{\Longrightarrow}$ ο δ.μ. $\overline{X}$ είναι επαρκής για το μ, αφού

$$T = \sum_{i=1}^{n} X_i = n \cdot \frac{\sum\limits_{i=1}^{n} X_i}{n} = n \cdot \overline{X} = \underset{\substack{\downarrow \\ n \cdot x}}{h}(\overline{X})$$

$\overset{(ii)}{\Longrightarrow}$ ο δ.μ. $\overline{X}$ είναι επαρκής για το e^{μ}, αφού $\mu = \underbrace{\log}_{\omega} \underbrace{(e^{\mu})}_{\phi}$

Προσοχή! $\overline{X} \not\Longrightarrow$ επαρκής για το μ^2

(δεν μπορούμε να αποφανθούμε, αφού για $\mu \in \mathbb{R}$ το $\mu = \omega(\mu^2)$ δεν μπορούμε να το εκφράσουμε για κάποια συνάρτηση $\omega(\cdot)$).

Άρα δε δουλεύει αυτό το κριτήριο [εδώ πράγματι δείχνεται ότι δεν είναι επαρκής].

Παρατήρηση 0.0.44. Αν μια παράμετρος $\theta > 0$, τότε μπορούμε να συμπεράνουμε την επάρκεια για το θ^2, αφού $\theta = \sqrt{\theta^2}$ για $\theta > 0$.

π.χ. σε τ.δ. Exp$(\theta) > 0$ έχουμε ότι το $\sum\limits_{i=1}^{n} X_i$ είναι επαρκής για το θ

$\Longrightarrow$ το $\sum\limits_{i=1}^{n} X_i$ είναι επαρκής για το θ^2 αφού $\theta = \sqrt{\theta^2} = \omega(\theta^2)$.

Πόρισμα 0.0.18. Έστω T επαρκής σ.σ. για το θ. Αν $h : \mathbb{R}^{\kappa} \to \mathbb{R}^{\kappa}$, $\phi : \mathbb{R}^{s} \to \mathbb{R}^{s}$ 2 συναρτήσεις που είναι 1-1 τουλάχιστον στο S_T και στο Θ αντίστοιχα, όπου $S = \bigcup\limits_{\theta \in \Theta} S_{T,\theta}$ τότε τα εξής είναι ισοδύναμα:

(i) T επαρκής για το θ.

(ii) $U = h(T)$ επαρκής για το θ.

(iii) T επαρκής για το $\Phi = \phi(\theta)$

Παράδειγμα 0.0.46. Όπως πριν, $\overline{X}$ επαρκής για το μ σε τ.δ. $\mathcal{N}(\mu, \sigma^2)$, όπου σ^2 γνωστό $\Longrightarrow \overline{X}$ επαρκής για το e^{μ} (e^{x} είναι "1-1" στο $\mathbb{R}$).

Όμως και πάλι δεν μπορούμε να αποφανθούμε για την επάρκεια για το μ^2 αφού η x^2 δεν είναι "1-1" στο $\mathbb{R}$
$\not\Rightarrow$ επαρκής για το μ^2.

Επίσης,

$\Longrightarrow e^{\overline{X}}$ επαρκής για το μ (διότι e^x, "1-1")
$\not\Rightarrow \left(\overline{X}\right)^2$ είναι επαρκής για το μ (δεν μπορούμε να αποφανθούμε)

Συμπληρωματικές σ.σ. - Ελαχιστκή επάρκεια

Εφ' όσον η επάρκεια ελαττώνει το # δεδομένων χωρίς απώλεια πληροφορίας, μια επαρκής σ.σ. είναι μεγαλύτερου ενδιαφέροντος, αν πετυχαίνει τη μεγαλύτερη δυνατή ελάττωση.

Ορισμός 0.0.26. Μια επαρκής σ.σ. S λέγεται ελαχιστική (minimal) αν $\forall$ T επαρκή σ.σ., $\exists\ g:\quad \underbrace{S = g(T)}_{\text{εξαρτάται απ' την } T}$

Παρατήρηση 0.0.45. Συνήθως στα περισσότερα στατιστικά μοντέλα υπάρχει ελαχιστικά επαρκής σ.σ. Μπορούν βέβαια να βρεθούν εξαιρέσεις.

Θεώρημα 0.0.19. (χαρακτηρισμός ελαχιστικά επαρκών σ.σ.) Έστω X τ.δ. με από κοινού σ.π.π./ σ.κ. $f_\theta(x)$. Η σ.σ. S είναι ελαχιστικά επαρκής για το θ αν και μόνο αν $\dfrac{f_\theta(x)}{f_\theta(y)}$ είναι ανεξ. του θ $\Leftrightarrow$ $S(x) = S(y)$.

Ορισμός 0.0.27. Μια σ.σ. V λέγεται συμπληρωματική (ancillary) αν η κατανομή της είναι ανεξάρτητη του θ.

Παρατήρηση 0.0.46. Μία συμπληρωματική σ.σ. είναι μια ποσότητα οδηγός που είναι ταυτόχρονα στατιστική συνάρτηση, δηλαδή η έκφρασή της δεν εξαρτάται από το θ.

Λογικά σκέφτεται κανείς ότι εφ' όσον η κατανομή της δεν εξαρτάται από το θ, δε μας δίνει πληροφορία για το θ. Παρ' όλα αυτά "περιέργως" σε συνδυασμό με άλλες μας δίνουν πολλές φορές πληροφορία.

Παράδειγμα 0.0.47. Έστω $X_1, \ldots, X_n$ τ.δ. $\mathcal{U}[0, \theta]$, $\theta > 0$. Ας υποθέσουμε ότι δίνεται το $X_{(1)} = \theta \cdot U_{(1)}$, όπου $U_{(1)} = \min\{U_1, \ldots, U_n\}$ όπου U_i α.ι.τ.μ. $\sim \mathrm{Unif}[0, 1]$.

Είδαμε ότι $\mathcal{U}[0,\theta] = \theta \cdot \mathcal{U}[0,1]$.

Ας θέσουμε
$$V = \frac{X_{(n)}}{X_{(1)}} = \frac{\theta \cdot U_{(n)}}{\theta \cdot U_{(1)}} = \frac{U_{(n)}}{U_{(1)}}$$
Τότε η V είναι σ.σ. (συνάρτηση του τ.δ. χωρίς άγνωστη παράμετρο) και έχει κατανομή που δεν εξαρτάται από το θ
$\Longrightarrow$ η V είναι συμπληρωματική σ.σ. (μόνη της δε μου δίνει πληροφορία για το θ).

Λύση

Ας υποθέσουμε ότι μας δίνεται το $X_{(1)}$ και το V.
$$X_{(n)} = X_{(1)} \cdot \frac{X_{(n)}}{X_{(1)}} = X_{(1)} \cdot V$$
$\Longrightarrow$ γνωρίζουμε το δειγματικό μέγιστο.

Συμπεραίνουμε ότι το ζεύγος $\big(X_{(1)}, V\big)$ είναι επαρκής σ.σ. για το θ διότι το $X_{(n)} = h\underbrace{\big(X_{(1)}, V\big)}_{U}$ και το $X_{(n)}$ είναι επαρκής σ.σ. για το θ.

Αρχικά είχαμε το $X_{(1)}$ που δεν είναι επαρκής σ.σ. για το θ (δείξτε το) και προσθέσαμε το V (είναι συμπληρωματική σ.σ.)
$$\Downarrow$$
φτάσαμε σε επαρκή σ.σ. $\big(X_{(1)}, V\big)$ άρα πήραμε πληροφορία για το θ, σε συνδυασμό με το $X_{(1)}$ [και μάλιστα όση πληροφορία χρειαζόμασταν].

Εδώ το $X_{(n)}$ είναι ελαχιστικά επαρκής σ.σ.
Πράγματι, έστω $x, y \in \mathbb{R}^n$.
$$\frac{f_\theta(x)}{f_\theta(y)} = \frac{\frac{1}{\theta^n} \cdot \mathbb{1}_{[X_{(n)},+\infty)}(\theta)}{\frac{1}{\theta^n} \cdot \mathbb{1}_{[Y_{(n)},+\infty)}(\theta)}$$
$$\frac{f_\theta(x)}{f_\theta(y)} \text{ ανεξάρτητο του } \theta \Leftrightarrow x_{(n)} = y_{(n)} \Leftrightarrow S(x) = S(y)$$
Άρα η $S(X) = X_{(n)}$ είναι ελαχιστικά επαρκής σ.σ.

Θα δείξουμε τώρα σε αυτό το παράδειγμα ότι το $X_{(n)}$ με το $\frac{X_{(n)}}{X_{(1)}} = V$ είναι ανεξάρτητες τ.μ., όπου $\frac{X_{(n)}}{X_{(1)}} = V$ συμπληρωματική σ.σ. και $X_{(n)}$ ελαχιστικά επαρκής σ.σ. Θα περιμέναμε ίσως να είναι της ελάχιστης δυνατής διάστασης και να συμπίπτουν με τη διάσταση του παραμετρικού χώρου. Επιπλέον, δε θα θέλαμε να "περιέχουν" συμπληρωματικές σ.σ.

Αρκεί να αποδείξουμε ότι το $\frac{1}{V} = \frac{X_{(1)}}{X_{(n)}}$ με το $X_{(n)}$ είναι ανεξάρτητες τ.μ.
$$\left(X_{(n)} \perp\!\!\!\perp \frac{1}{V} \Longrightarrow X_{(n)} \perp\!\!\!\perp V\right)$$

$\rightarrow$ Ερώτημα: ποια είναι η κατανομή του $\left(X_{(1)}, X_{(n)}\right)$;

$$\left(X_{(1)}, X_{(n)}\right) = \theta \cdot \left(U_{(1)}, U_{(n)}\right)$$

$$\downarrow$$

αρκεί να βρούμε αυτήν την κατανομή

$\rightarrow$ Αρκεί επίσης $U_{(1)} \perp\!\!\!\perp \dfrac{1}{V} = \dfrac{U_{(1)}}{U_{(n)}}$.

$$\left(\text{Υπενθύμιση: } \frac{X_{(1)}}{X_{(n)}} = \frac{\theta \cdot U_{(1)}}{\theta \cdot U_{(n)}} = \frac{U_{(1)}}{U_{(n)}}\right)$$

Έχουμε $\left(\dfrac{U_{(1)}}{U_{(n)}}, U_{(n)}\right) = g\left(U_{(1)}, U_{(n)}\right)$. Θέτουμε $Y = \left(\dfrac{U_{(1)}}{U_{(n)}}, U_{(n)}\right)$ με $Y_1 = \dfrac{U_{(1)}}{U_{(n)}}$, $Y_2 = U_{(n)}$.

Έχουμε αν θέσουμε $X = (X_1, X_2) = \left(U_{(1)}, U_{(n)}\right)$, τότε

$$Y = g(X) = (g_1(X_1, X_2),\ g_2(X_1, X_2)) = \left(\frac{X_1}{X_2}, X_2\right)$$

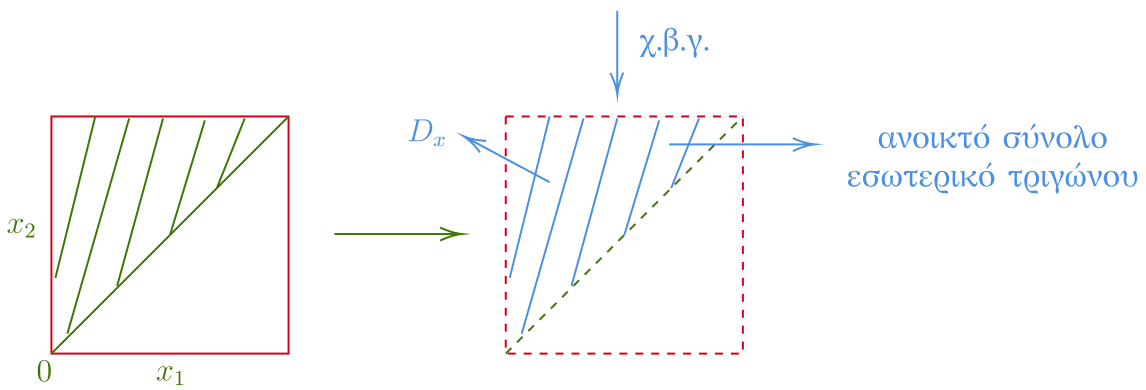
Για το "1-1":

$$\left.\begin{array}{l} Y_1 = \dfrac{X_1}{X_2} \\[2em] Y_2 = X_2 \end{array}\right| \Leftrightarrow \begin{array}{l} X_1 = Y_1 Y_2 \\[2em] X_2 = Y_2 \end{array}$$

Άρα η g αντιστρέφεται και έχουμε

$$g^{-1}(Y_1, Y_2) = (Y_1 Y_2, Y_2)$$

Για ευκολία δουλεύουμε σε ανοικτά υποσύνολα. Αρχικά $0 \le X_1 \le X_2 \le 1$.



$$g : D_x \rightarrow D_y, \quad \underset{(0,1)\times(0,1)}{\overset{}{D_y = g(D_x)}}$$

$$g(X) = \left(\frac{X_1}{X_2}, X_2\right), \quad X_2 \in (0,1) \times (0,1)$$

$$g^{-1} : D_y = (0,1) \times (0,1) \rightarrow D_x, \quad \text{με } g^{-1}(Y_1, Y_2) = (Y_1 Y_2, Y_2)$$

Θεώρημα αλλαγής μεταβλητής

Υπενθύμιση σε διάσταση 1

Αν $Y = g(X),\ g : 1-1$ και διαφορίσιμη με $g^{'}(x) \neq 0$ τότε

$$f_Y(y) = f_X\left(g^{-1}(y)\right) \cdot \left|\frac{dg^{-1}(y)}{dy}\right|$$

Επέκταση για διάσταση > 1

Αν $Y = g(X)$, η g είναι 1-1 και διαφορίσιμη και $|J_g(x)| \neq 0$ τότε

$$f_Y(y) = f_X\left(g^{-1}(y)\right) \cdot |J_{g^{-1}}(x)|, \ \forall y \in D_y$$

$\xrightarrow{B1}$ κατανομή $\left(U_{(1)}, U_{(n)}\right)$

$\xrightarrow{B2}$ πάμε στο μετασχηματισμό

Έχουμε,

$$B1: \quad F_{U_{(1)},U_{(n)}}(x_1,x_2) = \mathbb{P}\left(\underbrace{U_{(1)} \leq x_1}_{A}, \underbrace{U_{(n)} \leq x_2}_{B}\right), \quad 0 < x_1 < x_2 < 1$$

$$= \mathbb{P}\left(\underbrace{U_{(n)} \leq x_2}_{B}\right) - \mathbb{P}\left(\underbrace{U_{(1)} > x_1}_{A^c}\underset{\cap}{,}\underbrace{U_{(n)} \leq x_2}_{B}\right)$$

$$= F_u^n(x_2) - \mathbb{P}\left(x_1 < U_1 \leq x_2, \ldots, x_1 < U_n \leq x_2\right)$$

$$= F_u^n(x_2) - \left(F_u(x_2) - F_u(x_1)\right)^n$$

Για τα διδιάστατα, όταν έχουμε συνεχή τ.μ. (διδιάστατη)

$$\boxed{f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y}}, \quad 0 < x_1 < x_2 < 1$$

$$F_{U_{(1)},U_{(n)}}(x_1,x_2) = x_2^n - (x_2 - x_1)^n \Longrightarrow f_{U_{(1)},U_{(n)}}(x_1,x_2) = n \cdot (n-1) \cdot (x_2 - x_1)^{n-2}$$

$$f_Y(y) = f_X\left(g^{-1}(y)\right) \cdot |J_{g^{-1}}(x)|, \quad \forall y \in (0,1)^2$$

$$F_Y(y) = n(n-1)\left(y_2 - y_1 y_2\right)^{n-2} \cdot y_2 = n(n-1) y_2^{n-1}\left(1 - y_1\right)^{n-2}$$

$$[x_2 = y_2, \ x_1 = y_1 y_2]$$

όπου $g^{-1}(y_1,y_2) = (y_1 y_2, y_2)$.

$$J_{g^{-1}}(y_1,y_2) = \begin{pmatrix} y_2 & y_1 \\ 0 & 1 \end{pmatrix} \Longrightarrow \quad |J_{g^{-1}}(y_1,y_2)| = y_2 > 0$$

Έχουμε

$$f_{Y_1,Y_2}(y_1,y_2) = \underset{\substack{\| \\ h_1(y_1)h_2(y_2)}}{n(n-1) y_2^{n-1}\left(1-y_1\right)^{n-2}}, \qquad 0 < y_1, y_2 < 1$$

$$\Longrightarrow Y_1 \perp\!\!\!\perp Y_2 \Longrightarrow \frac{U_{(1)}}{U_{(n)}} \perp\!\!\!\perp U_{(n)} \quad \to \text{πήραμε την ανεξαρτησία}$$

Τελικά $X_{(n)}$ και $V = \dfrac{X_{(n)}}{X_{(1)}}$ είναι ανεξάρτητες τ.μ.

$\swarrow$ $\searrow$

ελαχιστικά επαρκής σ.σ. συμπληρωματική σ.σ.

Διάλεξη 32

Παράδειγμα 0.0.48. Έστω $X_1, \ldots, X_n$ τ.δ. $\mathcal{N}(\mu,\, \sigma^2)$.

γνωστό

Έχουμε δείξει ότι η $T = \sum\limits_{i=1}^{n} X_i$ είναι επαρκής σ.σ. $\Rightarrow$ $\overline{X}$ επαρκής σ.σ. για το μ.

Δείχνεται εύκολα ότι ο $\overline{X}$ είναι ελαχιστικά επαρκής (άσκηση).

$X_i = \mu + \sigma Z_i$, $Z_i \sim \mathcal{N}(0,1)$ και ανεξ. τ.μ.

Άρα

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 = \frac{\sigma^2}{n-1} \underbrace{\sum_{i=1}^{n}\left(Z_i - \overline{Z}\right)^2}_{\substack{\text{συμπληρωματική} \\ \text{σ.σ.}}}$$

$\overline{X} = \mu + \sigma\overline{Z} \quad \Rightarrow \quad S^2$ είναι σ.σ.σ.

Έχουμε δεχθεί ότι ο $\overline{X} \perp\!\!\!\perp S^2$ [θα το δείξουμε και μέσω θεωρήματος Basu]. Βλέπουμε ότι ο $\overline{X}$, που είναι ελαχιστικά επαρκής, είναι ανεξ. από τη σ.σ.σ. S^2. Άλλη μία φορά η ελαχιστικά επαρκής σ.σ. βγαίνει ανεξάρτητη από συμπληρωματική σ.σ.

Παράδειγμα 0.0.49. Έστω $X_1, \ldots, X_n$ τ.δ. $\mathcal{U}[\theta, \theta+1]$.

$$f_\theta(x) = \prod_{i=1}^{n} \mathbf{1}_{[\theta,\theta+1]}(x_i) = \mathbf{1}_{[\theta,+\infty)}(X_{(1)}) \cdot \mathbf{1}_{(-\infty,\theta+1]}(X_{(n)}) = g\Big(\underset{\substack{\| \\ T(X)}}{(X_{(1)}, X_{(n)})};\theta\Big) \xRightarrow{\text{Π.Κ.Ν.}}$$

η $S(X) = (X_{(1)}, X_{(n)})$ είναι επαρκής σ.σ.

Παρατήρηση 0.0.47. $\dim(\theta) = 1$ ενώ $\dim(S(X)) = 2 \longrightarrow$ μήπως μπορούμε να το ελλατώσουμε κι'άλλο;

Αναρωτιόμαστε αν η ελαχιστικά επαρκής σ.σ. μπορεί να έχει διάσταση 1. Εξετάζουμε ελαχιστική επάρκεια $\theta \leq X_{(1)},\ X_{(n)}\theta + 1 \Leftrightarrow X_{(n)} - 1 \leq \theta \leq X_{(1)}$.

$$\frac{f_\theta(x)}{f_\theta(y)} = \frac{\mathbf{1}_{[X_{(n)}-1, X_{(1)}]}(\theta)}{\mathbf{1}_{[X_{(n)}-1, X_{(1)}]}(\theta)}$$

ανεξ. του $\theta \Leftrightarrow X_{(1)} = Y_{(1)}$ και $X_{(n)} = Y_{(n)} \Leftrightarrow S(x) = S(y)$, όπου $S(x) = (X_{(1)}, X_{(n)})$ και $S(y) = (Y_{(1)}, Y_{(n)})$.
Άρα η $S(X) = (X_{(1)}, X_{(n)})$ είναι ελαχιστικά επαρκής σ.σ.

Ερώτημα: Ποιά είναι η ουσιαστική διαφορά μεταξύ των πρώτων δύο παραδειγμάτων και του τρίτου;

$$\dim(S(X)) = \dim(\theta) \quad \dim(S(X)) > \dim(\theta)$$

Παρατήρηση 0.0.48. Φαίνεται στο τρίτο παράδειγμα η ελαχιστικά επαρκής σ.σ. να «περιέχει» παραπάνω πληροφορία από αυτήν που θα θέλαμε ιδανικά για να φτάσουμε σε διάσταση 1. Παρατηρούμε ότι αν $R = X_{(n)} - X_{(1)}$ είναι το δειγματικό εύρος, τότε $R = h(S)$ και από τη σχέση $X_i = \theta + U_i\ ,\quad U_i \sim \mathcal{U}[0,1]$ (μετασχηματισμός θέσης).

Άρα

$$R = X_{(n)} - X_{(1)} = \theta + U_{(n)} - (\theta + U_{(1)}) = U_{(n)} - U_{(1)} \quad \text{(καταν. ανεξ. του } \theta\text{)}.$$

$$\Rightarrow \quad R \text{ είναι συμπληρωματική σ.σ.}$$

$\Rightarrow$ υπάρχει συνάρτηση της ελαχιστικά επαρκούς σ.σ. που είναι συμπληρωματική σ.σ. και άρα είναι στοχαστικά εξαρτημένη από αυτήν.

Βλέπουμε λοιπόν σε αυτό το παράδειγμα που η ελαχιστικά επαρκής σ.σ. που προέκυψε δεν μας «ικανοποιούσε» ότι περιέχει «μέσα στα σπλάχνα της» [δηλαδή προκύπτει ως συνάρτησή της μία συμπληρωματική σ.σ. (μη τετριμμένη)]. Δεν μπορούμε επομένως να «απαλλαγούμε» από αυτό το μη πληροφοριακό υλικό (το συμπληρωματικό) ελαττώνοντας κι άλλο τη διάσταση των δεδομένων (κρατώντας την επάρκεια). Θα έρθει σε λίγο μία άλλη έννοια, η έννοια της πληρότητας που όταν είναι παρούσα «καθαρίζει» από αυτό το συμπληρωματικό υλικό και έτσι φτάνουμε σε στατιστικά μοντέλα που έχουν πολύ καλύτερες ιδιότητες. Παρατηρείστε ότι στο προηγούμενο παράδειγμα

$$\mathbb{E}(R) = \mathbb{E}(X_{(n)} - X_{(1)}) = \mathbb{E}(U_{(n)} - U_{(1)}) = \frac{n}{n+1} - \frac{1}{n+1} = \frac{n-1}{n+1}.$$

Άρα αν θέσουμε $g(S) = g(X_{(1)}, X_{(n)}) = X_{(n)} - X_{(1)} - \dfrac{n-1}{n+1} \Rightarrow \forall \theta \in \mathbb{R},\ \mathbb{E}_\theta\left[g(S)\right] = 0.$

Παρατήρηση 0.0.49. Αν κάποια σ.σ. T ικανοποιεί $\mathbb{E}_\theta(T) = 0,\ \forall \theta \in \Theta$, τότε αυτή λέγεται αμερόληπτη εκτιμήτρια του μηδενός. Είναι η απουσία αυτής της συμπεριφοράς της ύπαρξης μη μηδενικών αμερόληπτων εκτιμητριών του μηδενός που καθορίζει την έννοια της πληρότητας.

Μία ασθενέστερη έννοια από αυτήν της σ.σ.σ. είναι η συμπληρωματική σ.σ. 1ης τάξης (first-order ancillarity).

Ορισμός 0.0.28. Μία σ.σ. $V = V(x)$ λέγεται 1ης τάξης συμπληρωματική αν $\mathbb{E}_\theta(V)$ είναι ανεξάρτητη του θ, δηλαδή σταθερά.

Παρατήρηση 0.0.50.

(i)

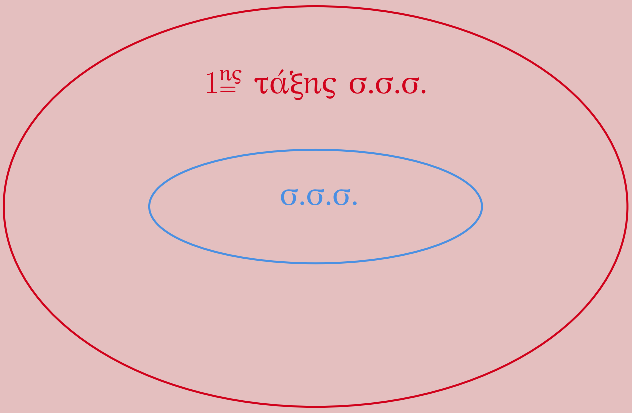
συμπληρωματική $\Rightarrow$ 1ης-τάξης συμπληρωματική
$\not\Leftarrow$ $\downarrow$
(ασθενέστερη έννοια)

π.χ. σε τ.δ. $\mathcal{N}(0, \sigma^2)$

άγνωστο

Έχουμε $\overline{X} \sim \mathcal{N}\left(0, \frac{\sigma^2}{n}\right) \Rightarrow \mathbb{E}_{\sigma^2}\left(\overline{X}\right) = 0, \quad \forall \sigma^2 > 0$

$\Rightarrow$ ο $\overline{X}$ είναι 1ης-τάξης σ.σ.σ. ενώ δεν είναι συμπληρωματική σ.σ. διότι η κατανομή του $\overline{X}$ εξαρτάται από το σ^2.



Αν μία επαρκής σ.σ. είναι «απαλλαγμένη» από 1ης-τάξης σ.σ.σ. (μη τετριμμένες), τότε θα είναι απαλλαγμένη και από σ.σ.σ.

Ορισμός 0.0.29. Έστω $X = (X_1, \ldots, X_n)$ τ.δ. από μία παραμετρική οικογένεια κατανομών $\mathcal{P} = \{P_\theta : \ \theta \in \Theta\}$.

(i) Μία σ.σ. $T = T(X)$ λέγεται πλήρης (complete), αν

$$\mathbb{E}_\theta\left[g(T)\right] = 0\,, \quad \forall \theta \in \Theta \Rightarrow \mathbb{P}_\theta\left[g(T) = 0\right] = 1\,, \quad \forall \theta \in \Theta.$$

(ii) Μία σ.σ. $T = T(X)$ λέγεται φραγμένα πλήρης (boundedly complete), αν η παραπάνω σχέση ισχύει μόνο για φραγμένες g.

(iii) Η οικογένεια των κατανομών που αντισοιχούν στην T λέγεται πλήρης ή φραγμένα πλήρης αντίστοιχα, ανάλογα αν ικανοποιείται η $(i)/(ii)$.

Παρατήρηση 0.0.51.

(i)

$$T \text{ πλήρης} \quad \underset{\not\Leftarrow}{\Rightarrow} \quad T \text{ φραγμένα πλήρης}$$

[Αφού αν ισχύει για όλες τις g, θα ισχύει και για τις φραγμένες.]
Μάλιστα μπορούμε να δώσουμε εύκολα αντιπαράδειγμα για $\not\Leftarrow$ [θα το δείξουμε σε λίγο].

(ii) Αν η T είναι πλήρης (φραγμένα πλήρης) τότε και η $U = h(T)$ είναι πλήρης (αντίστοιχα φραγμένα πλήρης).
Απόδειξη: [για φραγμένα πλήρεις, παρόμοια για πλήρεις]

Έστω g φραγμένη.

$$\mathbb{E}_\theta\left[g(U)\right] = 0\,, \ \ \forall \theta \in \Theta \Rightarrow$$
$$\mathbb{E}_\theta\left[(g \circ h)(T)\right] = 0\,, \ \ \forall \theta \in \Theta \xRightarrow{\text{υπόθ.}}$$
$$(T \text{ φραγμένα πλήρης } + \ g \circ h \text{ φραγμένη}), \text{ έχουμε } \mathbb{P}_\theta\left[(g \circ h)(T) = 0\right] = 1 \Rightarrow$$
$$\mathbb{P}_\theta\left[g(U) = 0\right] = 1\,, \ \ \forall \theta \in \Theta \Rightarrow \ \text{ η } U \text{ είναι φραγμένα πλήρης.}$$

$$\begin{aligned} T \text{ πλήρης } &\Rightarrow \text{ η } T \text{ δεν περιέχει (μη τετριμμένες) 1ης-τάξης σ.σ.σ.} \\ &\Rightarrow \text{ η } T \text{ δεν περιέχει (μη τετριμμένες) σ.σ.σ.} \end{aligned}$$

Απόδειξη:

Πράγματι, έστω $V = h(T)$ 1ης-τάξης σ.σ.σ.
Τότε

$$\begin{aligned} &\mathbb{E}_\theta(V) = \mathbb{E}_\theta\left[h(T)\right] = c\,, \ \ \forall\theta \in \Theta \Rightarrow \\ &\mathbb{E}_\theta\,\underbrace{(h(T) - c)}_{g(T)} = 0\,, \ \ \forall\theta \in \Theta \xRightarrow{T \text{ πλήρης(υπόθ.)}} \\ &\mathbb{P}_\theta[\underbrace{h(T)}_{V} = c] = 1\,, \ \ \forall\theta \in \Theta \quad () \end{aligned}$$

Άρα από T μία πλήρη σ.σ. δε μπορεί να προκύψει μη τετριμμένη 1ης-τάξης σ.σ.σ. της μορφής $h(T)$.

Παράδειγμα 0.0.50. (φραγμένα πλήρους σ.σ. που δεν είναι πλήρης σ.σ.) [*Lehmann-Scheffé ; 1950*]
Έστω X με

$$f_\theta(x) = \begin{cases} \theta\,, & x = -1 \\ (1-\theta)^2\,\theta^x\,, & x = 0, 1, 2, \ldots \end{cases}$$

Εδώ $\theta \in (0,1)$. Θ.δ.ο. η X δεν είναι πλήρης, αλλά είναι φραγμένα πλήρης.

Λύση

Έστω $g:\ \mathbb{E}_\theta\left[g(X)\right] = 0\,,\ \forall\theta \in (0,1)$.
Έχουμε ότι

$$\mathbb{E}_\theta\left[g(X)\right] = g(-1)\theta + \sum_{x=0}^{+\infty} g(x)(1-\theta)^2\theta^x$$

$$\mathbb{E}_\theta\left[g(X)\right] = 0\,, \ \ \forall\theta \in \Theta \Leftrightarrow \sum_{x=0}^{+\infty} g(x)\theta^x = -\frac{g(-1)\theta}{(1-\theta)^2}$$

$$\sum_{x=0}^{+\infty} \theta^x \overset{0<\theta<1}{=} \frac{1}{1-\theta} \Rightarrow \sum_{x=1}^{+\infty} x\theta^{x-1} = \frac{1}{(1-\theta)^2} \Rightarrow \boxed{\sum_{x=1}^{+\infty} x\theta^x = \frac{\theta}{(1-\theta)^2}}$$

Από τα προηγούμενα έχουμε

$$\sum_{x=0}^{+\infty} g(x)\theta^x = -g(-1)\sum_{x=0}^{+\infty} x\theta^x\,, \ \ \forall\theta \in (0,1)$$

Άρα $\forall x = 0, 1, \ldots \quad g(x) = -g(-1)x$.

- Αν $g(-1) = 0 \Rightarrow g(x) = 0\ ,\ \ \forall x = -1, -2, \ldots$ (μηδενική συνάρτηση)
- Αν $g(-1) \neq 0 \xRightarrow{\text{άλλες λύσεις}} g(x) = -g(-1)x\ ,\ \ \forall x = 0, 1, 2, \ldots$
  Για όλα τα $x \neq 0\ ,\ \ g(x) \neq 0$

Στην γενική περίπτωση

$$\exists g \neq 0 \ \text{(άπειρες) με } \mathbb{E}_\theta\left[g(X)\right] = 0\ ,\ \ \forall \theta \in (0,1) \Rightarrow \ \text{η } X \text{ δεν είναι πλήρης!}$$

Οι δυνατές g είναι της μορφής $c{\cdot}x$, $x = 0, 1, \ldots, c \neq 0$, είναι δηλαδή μη φραγμένες. Άρα η X είναι φραγμένα πλήρης [αφού δεν υπάρχει φραγμένη μη μηδενική που να ικανοποιεί τη σχέση $\mathbb{E}_\theta\left[g(X)\right] = 0,\ \forall \theta \in (0,1)$].

Ασκήσεις

Ν.δ.ο. $\sum\limits_{i=1}^{n} X_i$ είναι πλήρης σ.σ. σε τ.δ. $\mathrm{Bin}(\underset{\text{γνωστό}}{N}, \underset{\text{άγνωστο}}{p})$ ή $\mathcal{P}(\underset{\text{άγνωστο}}{\lambda})$.

Διάλεξη 33

Διάλεξη 34

ΜΑΘΗΜΑΤΙΚΗ ΣΤΑΤΙΣΤΙΚΗ - QUIZ 4, 9 Ιανουαρίου 2021

ΕΡΩΤΗΣΕΙΣ:

1. Έστω μία τυχαία μεταβλητή $F \sim F_{n_1,n_2}$, $X_1 \sim \chi^2_{n_1}$, $X_2 \sim \chi^2_{n_2}$ και X_1, X_2 ανεξάρτητες.

   ○ $F \stackrel{d}{=} X_1/X_2$ ○ $F \stackrel{d}{=} X_1^2/X_2^2$ ○ $F \stackrel{d}{=} (n_2X_1)/(n_1X_2)$

2. Αν $\lambda = \sigma_1^2/\sigma_2^2$ είναι ο λόγος διασπορών σε 2 ανεξάρτητα τ.δ. $N(\mu_i, \sigma_i^2)$, $i = 1, 2$, τότε εκτιμώντας με το λόγο δειγματικών διασπορών $\widehat{\lambda} = S_1^2/S_2^2$, έχουμε $\widehat{\lambda}/\lambda$ ακολουθεί

   ○ F_{n_1,n_2} ○ F_{n_1-1,n_2-1} ○ F_{n_1-2,n_2-2}

3. Μία σ.σ. T είναι επαρκής για το θ σε διακριτό μοντέλο, αν για κάθε θ

   ○ $\mathbb{E}_\theta(T) = c$ ○ $\mathbb{P}_\theta(X = x|T = t) = c(x, t)$ ○ $\mathbb{P}_\theta(T = t) = c(t)$

4. Αν εφαρμόζεται το παραγοντικό κριτήριο του Neyman, τότε η σ.π.π. $f_\theta(x)$ γράφεται:

   ○ $g(T(x); \theta)h(x)$ ○ $g(x; \theta)h(x)$ ○ $T(x)g(\theta)$

5. Σε τ.δ. $N(\mu, \sigma^2)$, σ^2 γνωστό, η $T(X) = \sum_{i=1}^n X_i^2$ είναι επαρκής για το άγνωστο $\mu \in \mathbb{R}$.

   ○ Ναι ○ Όχι ○ ακαθόριστο

6. Αν S ελαχιστικά επαρκής και T επαρκής για το θ, τότε για κατάλληλη g

   ○ $S = \min\{T, g(T)\}$ ○ $S = g(T)$ ○ $T = g(S)$

7. Μία διακριτή συμπληρωματική σ.σ. πρώτης-τάξης V ικανοποιεί

   ○ $\mathbb{E}_\theta(V) = c$ ○ $\mathbb{V}_\theta(V) = c$ ○ $\mathbb{P}_\theta(V = v) = c(v)$

8. Μία επαρκής και πλήρης σ.σ. είναι και ελαχιστικά επαρκής σύμφωνα με το θεώρημα

   ○ Neyman ○ Rao-Blackwell ○ Bahadur

9. Για να είναι πλήρης η T για το $\theta \in \Theta$ μέσω Ε.Ο.Κ., πρέπει η Ε.Ο.Κ. να

   ○ είναι καμπυλωμένη ○ είναι μονοπαραμετρική ○ έχει $interior(\Theta) \neq \emptyset$

10. Από το θεώρημα Rao-Blackwell παίρνουμε καλύτερες αμερόληπτες εκτιμήτριες, όταν δεσμεύουμε σε σ.σ. που είναι

    ○ επαρκείς ○ πλήρεις ○ φραγμένα πλήρεις

ΚΑΛΗ ΕΠΙΤΥΧΙΑ!

Απαντήσεις

1	Γ
2	Β
3	Β
4	Α
5	Β
6	Β
7	Α
8	Γ
9	Γ
10	Α

Σχολιασμός του Quiz 4

Διάλεξη 35

Διάλεξη 36

Διάλεξη 37

Διάλεξη 38

Διάλεξη 39

Διάλεξη 40

Ευρετήριο