Λύσεις ασκήσεων 4

Μέθοδος Δέλτα και μέγιστη πιθανοφάνεια

Λύση 1. Εκτίμηση του ρυθμού μιας εκθετικής κατανομής

Η παράμετρος θ είναι ο ρυθμός, άρα

$$\mathbb{E}_\theta(X_1) = \frac{1}{\theta}, \qquad \mathbb{V}_\theta(X_1) = \frac{1}{\theta^2}.$$

(α) Ο ισχυρός νόμος δίνει $\overline{X}_n \xrightarrow{\text{a.s.}} 1/\theta$. Η συνάρτηση $u \mapsto 1/u$ είναι συνεχής στο $1/\theta > 0$, οπότε $T_n \xrightarrow{\text{a.s.}} \theta$.

(β) Από το Κ.Ο.Θ.,

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

Για $g(u) = 1/u$ έχουμε $g'(1/\theta) = -\theta^2$. Η μέθοδος Δέλτα δίνει

$$\sqrt{n}(T_n - \theta) \xrightarrow{d} \mathcal{N}(0, \theta^2).$$

(γ) Η ακριβής ταυτότητα

$$A_n = \sqrt{n}(1 - \theta\overline{X}_n) = -\theta\sqrt{n}\left(\overline{X}_n - \frac{1}{\theta}\right)$$

δίνει απευθείας $A_n \xrightarrow{d} \mathcal{N}(0,1)$. Για την B_n εφαρμόζουμε τη μέθοδο Δέλτα στη συνάρτηση $u \mapsto \log u$, στο σημείο θ, και παίρνουμε επίσης $B_n \xrightarrow{d} \mathcal{N}(0,1)$.
Για τη διαφορά θέτουμε $U_n = \theta\overline{X}_n$ και $h(u) = u - 1 - \log u$. Τότε

$$B_n - A_n = \sqrt{n}\, h(U_n), \qquad h(1) = h'(1) = 0.$$

Εφόσον $\sqrt{n}(U_n - 1) \xrightarrow{d} \mathcal{N}(0,1)$, η μέθοδος Δέλτα με μηδενική παράγωγο δίνει $B_n - A_n \xrightarrow{\mathbb{P}} 0$. Η απόδειξη χρησιμοποιεί τη σχέση των δύο ποσοτήτων, όχι μόνο τις περιθώριες οριακές κατανομές τους.

Για σκέψη. Η ανισότητα $\log u \le u - 1$ δίνει $B_n - A_n \ge 0$. Ο λόγος $h(U_n)/(U_n - 1)^2$, με τιμή $1/2$ στο $U_n = 1$, συγκλίνει κατά πιθανότητα στο $1/2$. Επίσης $n(U_n - 1)^2 \xrightarrow{d} Z^2$, όπου $Z \sim \mathcal{N}(0,1)$. Άρα

$$\sqrt{n}(B_n - A_n) = n\, h(U_n) \xrightarrow{d} \frac{1}{2}Z^2.$$

Η διαφορά χάνεται στην πρώτη κανονική προσέγγιση, αλλά εμφανίζεται σε μικρότερη κλίμακα.

Λύση 2. Η διασπορά στην Bernoulli και η περίπτωση $p = 1/2$

(α) Από το Κ.Ο.Θ.,

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

Για $g(u) = u(1-u)$ έχουμε $g'(p) = 1 - 2p$. Άρα, όταν $p \neq 1/2$,

$$\boxed{\sqrt{n}\big(V_n - p(1-p)\big) \xrightarrow{d} \mathcal{N}\big(0, (1-2p)^2 p(1-p)\big).}$$

(β) Για $p = 1/2$, θέτουμε $Z_n = 2\sqrt{n}(\widehat{p}_n - 1/2)$. Τότε $Z_n \xrightarrow{d} Z \sim \mathcal{N}(0,1)$. Από την ταυτότητα της εκφώνησης,

$$n(V_n - 1/4) = -\frac{1}{4} Z_n^2 \xrightarrow{d} -\frac{1}{4} Z^2.$$

Επομένως

$$\boxed{n(V_n - 1/4) \xrightarrow{d} -\frac{1}{4}\chi^2_1.}$$

Εδώ το $-\frac{1}{4}\chi^2_1$ δηλώνει την κατανομή του $-Z^2/4$. Επίσης,

$$\sqrt{n}(V_n - 1/4) = -\frac{Z_n^2}{4\sqrt{n}} \xrightarrow{\mathbb{P}} 0,$$

από το θεώρημα συνεχούς απεικόνισης και το λήμμα Slutsky. Η κλίμακα $\sqrt{n}$ δίνει εκφυλισμένο όριο· η κλίμακα n δίνει μη εκφυλισμένο όριο.

(γ) Επειδή $S_n^2 = nV_n/(n-1)$,

$$n\left(S_n^2 - \frac{1}{4}\right) = \frac{n}{n-1}\, n\left(V_n - \frac{1}{4}\right) + \frac{n}{4(n-1)}.$$

Συνεπώς

$$\boxed{n(S_n^2 - 1/4) \xrightarrow{d} \frac{1 - Z^2}{4}.}$$

Η διόρθωση της μεροληψίας μετατοπίζει το όριο κατά $1/4$. Στην κλίμακα n η διαφορά των δύο εκτιμητριών δεν είναι αμελητέα.

Για σκέψη. Στο $p = 1/2$ έχουμε $\sigma^2 = 1/4$ και $\mu_4 = 1/16$. Επομένως

$$\mathrm{MSE}(S_n^2) = \frac{1}{8n(n-1)}.$$

Η $V_n = (n-1)S_n^2/n$ έχει μεροληψία $-1/(4n)$. Άρα

$$\mathrm{MSE}(V_n) = \frac{n-1}{8n^3} + \frac{1}{16n^2} = \frac{3n-2}{16n^3}.$$

Η διαφορά των δύο σφαλμάτων είναι

$$\mathrm{MSE}(V_n) - \mathrm{MSE}(S_n^2) = \frac{n^2 - 5n + 2}{16n^3(n-1)}.$$

Για $n = 2, 3, 4$ μικρότερο ΜΤΣ έχει η V_n, ενώ για κάθε $n \geq 5$ μικρότερο ΜΤΣ έχει η S_n^2. Στο συγκεκριμένο σημείο της παραμέτρου η αμερόληπτη διόρθωση δεν είναι αμελητέα στην κλίμακα n.

Λύση 3. Εμπειρική και επαγόμενη εκτιμήτρια μιας πιθανότητας

(α) Η T_n είναι η εμπειρική εκτιμήτρια της πιθανότητας $\mathbb{P}_\lambda(X_1 = 0)$. Η R_n προκύπτει με αντικατάσταση του λ από την εκτιμήτρια ροπών $\overline{X}_n$ στη συνάρτηση $g(\lambda) = e^{-\lambda}$.
Οι $Y_i = \mathbf{1}_{\{X_i=0\}}$ είναι ανεξάρτητες Bernoulli με παράμετρο q. Άρα, από τον ισχυρό νόμο,

$$T_n \xrightarrow{\text{a.s.}} q, \qquad \overline{X}_n \xrightarrow{\text{a.s.}} \lambda.$$

Η συνέχεια της εκθετικής δίνει $R_n \xrightarrow{\text{a.s.}} e^{-\lambda} = q$.

(β) Από το Κ.Ο.Θ. για τις Y_i,

$$\sqrt{n}(T_n - q) \xrightarrow{d} \mathcal{N}(0, q(1-q)).$$

Για τις X_i έχουμε $\mathbb{E}_\lambda(X_i) = \mathbb{V}_\lambda(X_i) = \lambda$, οπότε

$$\sqrt{n}(\overline{X}_n - \lambda) \xrightarrow{d} \mathcal{N}(0, \lambda).$$

Με $g'(\lambda) = -e^{-\lambda}$, η μέθοδος Δέλτα δίνει

$$\sqrt{n}(R_n - q) \xrightarrow{d} \mathcal{N}(0, \lambda e^{-2\lambda}).$$

(γ) Η διαφορά των ασυμπτωτικών διασπορών είναι

$$e^{-\lambda}(1 - e^{-\lambda}) - \lambda e^{-2\lambda} = e^{-2\lambda}(e^{\lambda} - 1 - \lambda) > 0,$$

αφού $e^{\lambda} > 1 + \lambda$ για $\lambda > 0$. Επομένως η R_n έχει μικρότερη διασπορά στην οριακή κανονική κατανομή.

(δ) Η $K_n = \sum_i X_i$ έχει κατανομή Poisson$(n\lambda)$ και $R_n = e^{-K_n/n}$. Άρα

$$\mathbb{E}_\lambda(R_n) = \exp\!\left\{n\lambda(e^{-1/n} - 1)\right\}.$$

Η ανισότητα $e^{-1/n} > 1 - 1/n$ δείχνει ότι αυτή η μέση τιμή είναι μεγαλύτερη από q. Από το ανάπτυγμα της εκθετικής,

$$n\lambda(e^{-1/n} - 1) = -\lambda + \frac{\lambda}{2n} + O(n^{-2}).$$

Επομένως

$$\mathbb{E}_\lambda(R_n) = q\left(1 + \frac{\lambda}{2n} + O(n^{-2})\right), \qquad n\big[\mathbb{E}_\lambda(R_n) - q\big] \longrightarrow \frac{\lambda q}{2}.$$

Η μεροληψία είναι τάξης $1/n$, μικρότερης από την κλίμακα $1/\sqrt{n}$ του σφάλματος. Επίσης η T_n είναι αμερόληπτη, αφού είναι ο μέσος όρος Bernoulli μεταβλητών με παράμετρο q.

Για σκέψη. Θεωρούμε $\mathbb{P}(X_i = 0) = \mathbb{P}(X_i = 2) = 1/2$. Η μέση τιμή είναι 1, αλλά η πιθανότητα του μηδενός είναι $1/2$ και όχι e^{-1}. Ο ισχυρός νόμος δίνει

$$T_n \xrightarrow{\text{a.s.}} \frac{1}{2}, \qquad R_n \xrightarrow{\text{a.s.}} e^{-1}.$$

Η εμπειρική T_n παραμένει συνεπής για την πιθανότητα του μηδενός. Η R_n χρησιμοποιεί επιπλέον τη σχέση $\mathbb{P}(X_i = 0) = e^{-\mathbb{E}(X_i)}$, η οποία ίσχυε στο μοντέλο Poisson και δεν ισχύει εδώ.

Λύση 4. Δειγματική διασπορά χωρίς κανονικότητα

(α) Η ποσότητα είναι μη αρνητική και

$$\mathbb{E}\big[\sqrt{n}(\overline{X}_n - \mu)^2\big] = \frac{\sigma^2}{\sqrt{n}} \longrightarrow 0.$$

Η ανισότητα Markov δίνει το ζητούμενο.

(β) Οι $Y_i = (X_i - \mu)^2$ είναι ανεξάρτητες και ισόνομες, με

$$\mathbb{E}(Y_i) = \sigma^2, \qquad \mathbb{V}(Y_i) = \mu_4 - \sigma^4.$$

Θέτουμε $A_n = n^{-1}\sum_i (X_i - \mu)^2$. Από το Κ.Ο.Θ.,

$$\sqrt{n}(A_n - \sigma^2) \xrightarrow{d} \mathcal{N}(0, \mu_4 - \sigma^4).$$

Η δοσμένη ταυτότητα και το (α) δίνουν

$$\sqrt{n}(M_{2,n} - A_n) = -\sqrt{n}(\overline{X}_n - \mu)^2 \xrightarrow{\mathbb{P}} 0.$$

Επομένως, από το Slutsky,

$$\sqrt{n}(M_{2,n} - \sigma^2) \xrightarrow{d} \mathcal{N}(0, \mu_4 - \sigma^4).$$

Επίσης $M_{2,n} \xrightarrow{\mathbb{P}} \sigma^2$ και

$$\sqrt{n}(S_n^2 - M_{2,n}) = \frac{\sqrt{n}}{n-1} M_{2,n} \xrightarrow{\mathbb{P}} 0.$$

Άρα

$$\sqrt{n}(S_n^2 - \sigma^2) \xrightarrow{d} \mathcal{N}(0, \mu_4 - \sigma^4).$$

(γ) Για $g(u) = \sqrt{u}$ έχουμε $g'(\sigma^2) = 1/(2\sigma)$. Η μέθοδος Δέλτα δίνει

$$\sqrt{n}(S_n - \sigma) \xrightarrow{d} \mathcal{N}\left(0, \frac{\mu_4 - \sigma^4}{4\sigma^2}\right).$$

(δ) Για την $\mathrm{Unif}[-\theta, \theta]$ ισχύουν

$$\mu = 0, \qquad \sigma^2 = \frac{\theta^2}{3}, \qquad \mu_4 = \frac{\theta^4}{5}.$$

Η διαφορά $\mu_4 - \sigma^4$ είναι $4\theta^4/45$. Αντικαθιστώντας στο (γ),

$$\sqrt{n}\left(S_n - \frac{\theta}{\sqrt{3}}\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{\theta^2}{15}\right).$$

Πολλαπλασιάζοντας με $\sqrt{3}$ παίρνουμε

$$\sqrt{n}(\sqrt{3}\, S_n - \theta) \xrightarrow{d} \mathcal{N}\left(0, \frac{\theta^2}{5}\right).$$

Για σκέψη. Η ισότητα $\mu_4 = \sigma^4$ ισοδυναμεί με μηδενική διασπορά της $(X_1 - \mu)^2$. Άρα $(X_1 - \mu)^2 = \sigma^2$ σχεδόν βέβαια. Εφόσον $\mathbb{E}(X_1 - \mu) = 0$, οι δύο δυνατές τιμές $\mu - \sigma$ και $\mu + \sigma$ έχουν πιθανότητα $1/2$ καθεμία. Για ένα τέτοιο δείγμα,

$$M_{2,n} = \sigma^2 - (\overline{X}_n - \mu)^2.$$

Από το Κ.Ο.Θ. και τη συνεχή απεικόνιση,

$$n(M_{2,n} - \sigma^2) \xrightarrow{d} -\sigma^2 Z^2, \qquad Z \sim \mathcal{N}(0,1).$$

Η Bernoulli με $p = 1/2$ είναι η ειδική περίπτωση $\mu = 1/2$ και $\sigma = 1/2$.

Λύση 5. Ακριβής και ασυμπτωτική πιθανότητα σφάλματος

(α) Θέτουμε $Q_n = (n-1)S_n^2/\sigma^2 \sim \chi^2_{n-1}$. Το ενδεχόμενο ενδιαφέροντος γράφεται

$$\{Q_n < (n-1)(1-r)\} \cup \{Q_n > (n-1)(1+r)\}.$$

Επομένως η ακριβής πιθανότητα είναι

$$\boxed{F_{\chi^2_{n-1}}((n-1)(1-r)) + 1 - F_{\chi^2_{n-1}}((n-1)(1+r)).}$$

Τα αυστηρά ή μη αυστηρά όρια δεν αλλάζουν την πιθανότητα, αφού η κατανομή είναι συνεχής.

(β) Από την ασυμπτωτική κατανομή,

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

Η αντίστοιχη κανονική προσέγγιση είναι

$$\mathbb{P}_{\mu,\sigma^2}(|S_n^2 - \sigma^2| > r\sigma^2) \simeq 2\left[1 - \Phi\left(r\sqrt{\frac{n}{2}}\right)\right].$$

Πρόκειται για προσέγγιση πεπερασμένου δείγματος. Το Κ.Ο.Θ. δεν παρέχει από μόνο του φράγμα για το σφάλμα της προσέγγισης, ούτε σχετική ακρίβεια στις ουρές.

(γ) Εδώ $r = 1/4$ και τα δύο όρια της χ^2_{99} είναι 74.25 και 123.75. Άρα η ακριβής πιθανότητα είναι περίπου

$$0.0298512 + 1 - 0.9532350 = 0.0766162.$$

Η κανονική προσέγγιση δίνει

$$2\left[1 - \Phi\left(\frac{5}{2\sqrt{2}}\right)\right] \simeq 2(1 - 0.9614501) = 0.0770998.$$

Οι τιμές είναι περίπου 7.66% και 7.71%, αντίστοιχα.

(δ) Επιλύουμε

$$2\left[1 - \Phi\left(0.2\sqrt{\frac{n}{2}}\right)\right] \le 0.05.$$

Με το δοσμένο ποσοστημόριο αυτό ισοδυναμεί κατά προσέγγιση με

$$0.2\sqrt{\frac{n}{2}} \ge 1.96 \quad \Longleftrightarrow \quad n \ge 2\left(\frac{1.96}{0.2}\right)^2 = 192.08.$$

Συνεπώς επιλέγουμε $n = 193$. Αυτή είναι επιλογή από την κανονική προσέγγιση· για ακριβή έλεγχο της πιθανότητας χρησιμοποιούμε τον τύπο του (α).

Για σκέψη. Η ακριβής κάτω ουρά είναι 0.0298512 και η πάνω ουρά 0.0467650. Η κανονική προσέγγιση δίνει και για τις δύο την τιμή 0.0385499. Άρα η προσέγγιση μείον την ακριβή τιμή είναι 0.0086987 στην κάτω ουρά και -0.0082151 στην πάνω. Τα δύο σφάλματα έχουν αντίθετα πρόσημα και σχεδόν αλληλοαναιρούνται: το συνολικό σφάλμα είναι μόλις 0.0004836. Η καλή προσέγγιση του αθροίσματος δεν συνεπάγεται εξίσου καλή προσέγγιση κάθε ουράς.

Λύση 6. Bernoulli με περιορισμένη παράμετρο

(α) Για $x_i \in \{0,1\}$ και $k = \sum_i x_i$,

$$L_x(p) = p^k(1-p)^{n-k}.$$

Στα άκρα ερμηνεύουμε την έκφραση μέσω των αντίστοιχων πιθανοτήτων, ή με τη σύμβαση για κενά γινόμενα. Αν $k=0$, η $L_x(p) = (1-p)^n$ μεγιστοποιείται μοναδικά στο 0. Αν $k=n$, η $L_x(p) = p^n$ μεγιστοποιείται μοναδικά στο 1.

Για $0 < k < n$,

$$\ell_x'(p) = \frac{k}{p} - \frac{n-k}{1-p} = \frac{k-np}{p(1-p)}.$$

Η λογαριθμοπιθανοφάνεια αυξάνεται μέχρι το k/n και μειώνεται μετά από αυτό. Συνεπώς, σε όλες τις περιπτώσεις,

$$\widehat{p}_n = \frac{K_n}{n} = \overline{X}_n.$$

(β) Η K_n ακολουθεί $\mathrm{Bin}(n,p)$, άρα η πιθανοφάνεια με μόνο δεδομένο το k είναι

$$L_k(p) = \binom{n}{k} p^k(1-p)^{n-k}.$$

Διαφέρει από την L_x κατά θετικό παράγοντα ανεξάρτητο του p και έχει τον ίδιο μεγιστοποιητή.

(γ) Από την ίδια μονοτονία,

$$\boxed{\widehat{p}_n^{\,c} = \begin{cases} 1/4, & \overline{X}_n < 1/4, \\ \overline{X}_n, & 1/4 \le \overline{X}_n \le 3/4, \\ 3/4, & \overline{X}_n > 3/4. \end{cases}}$$

(δ) Στο μοντέλο $p \in [0,1]$, η εμπειρική εκτιμήτρια, η εκτιμήτρια ροπών και η ε.μ.π. συμπίπτουν με την $\overline{X}_n$. Στο περιορισμένο μοντέλο η εμπειρική σχετική συχνότητα παραμένει η $\overline{X}_n$, αλλά μπορεί να βρίσκεται εκτός του παραμετρικού χώρου. Η εξίσωση ροπών $p = \overline{X}_n$ έχει επιτρεπτή λύση μόνο όταν $\overline{X}_n \in [1/4, 3/4]$. Η ε.μ.π. υπάρχει πάντοτε και παίρνει τις τιμές του (γ). Η προβολή της εμπειρικής τιμής στο επιτρεπτό διάστημα είναι εδώ και η ε.μ.π., όχι λύση της αρχικής εξίσωσης ροπών όταν αυτή δεν έχει επιτρεπτή λύση.

Για σκέψη. Στο $p = 1/4$, η συνέπεια της $\overline{X}_n$ συνεπάγεται ότι ο άνω περιορισμός $3/4$ δεν ενεργοποιείται με πιθανότητα που τείνει στο 1. Εκτός ενδεχομένου με πιθανότητα που τείνει στο 0,

$$\sqrt{n}(\widehat{p}_n^{\,c} - 1/4) = \max\{0, \sqrt{n}(\overline{X}_n - 1/4)\}.$$

Το Κ.Ο.Θ. και η συνέχεια της $u \mapsto \max(0,u)$ δίνουν

$$\sqrt{n}(\widehat{p}_n^{\,c} - 1/4) \xrightarrow{d} \max(0, Z), \qquad Z \sim \mathcal{N}(0, 3/16).$$

Το όριο έχει πιθανότητα $1/2$ στο 0 και δεν είναι κανονικό. Για σταθερό p στο εσωτερικό του $[1/4, 3/4]$, κανένας περιορισμός δεν ενεργοποιείται με πιθανότητα που τείνει στο 1, και παραμένει το όριο $\mathcal{N}(0, p(1-p))$.

Λύση 7. Εκθετική κατανομή: ρυθμός και μέση τιμή

(α) Για $s = \sum_i x_i > 0$,

$$\ell_x(\theta) = n \log \theta - \theta s, \qquad \ell_x'(\theta) = \frac{n}{\theta} - s, \qquad \ell_x''(\theta) = -\frac{n}{\theta^2} < 0.$$

Η μοναδική λύση είναι $\theta = n/s$ και, λόγω γνήσιας κοιλότητας, δίνει ολικό μέγιστο. Άρα $\widehat{\theta}_n = 1/\overline{X}_n$.

(β) Έχουμε $m = 1/\theta$ και

$$f(x; m) = m^{-1} e^{-x/m} \mathbf{1}_{(0,\infty)}(x), \qquad \ell_x(m) = -n \log m - \frac{s}{m}.$$

Η παράγωγος είναι $\ell_x'(m) = (s - nm)/m^2$, θετική για $m < \overline{x}$ και αρνητική για $m > \overline{x}$. Επομένως $\widehat{m}_n = \overline{X}_n$. Η μέθοδος ροπών δίνει $m = \overline{X}_n$ ή, στην παραμετροποίηση ρυθμού, $1/\theta = \overline{X}_n$. Οι εκτιμήτριες συμπίπτουν με τις αντίστοιχες ε.μ.π.

(γ) Με $A_n = \sum_i X_i$ και $n \geq 2$,

$$\mathbb{E}_\theta(A_n^{-1}) = \frac{\theta^n}{\Gamma(n)} \int_0^\infty s^{n-2} e^{-\theta s}\, ds.$$

Το ολοκλήρωμα είναι $\Gamma(n-1)/\theta^{n-1}$, οπότε

$$\mathbb{E}_\theta(A_n^{-1}) = \frac{\theta}{n-1}.$$

Άρα

$$\mathbb{E}_\theta(\widehat{\theta}_n) = \frac{n\theta}{n-1}, \qquad \mathrm{Bias}_\theta(\widehat{\theta}_n) = \frac{\theta}{n-1}.$$

Η αμερόληπτη διόρθωση είναι

$$\widetilde{\theta}_n = \frac{n-1}{n}\widehat{\theta}_n = \frac{n-1}{\sum_i X_i}.$$

Δεν είναι ε.μ.π., επειδή διαφέρει από τον μοναδικό μεγιστοποιητή.

(δ) Η διαφορά των κανονικοποιημένων εκτιμητριών είναι

$$\sqrt{n}(\widetilde{\theta}_n - \widehat{\theta}_n) = -\frac{\widehat{\theta}_n}{\sqrt{n}} \xrightarrow{\mathbb{P}} 0.$$

Με την Άσκηση 1 και το λήμμα Slutsky,

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

Για σκέψη. Η ίδια ολοκλήρωση με το (γ), για $n \geq 3$, δίνει

$$\mathbb{E}_\theta(A_n^{-2}) = \frac{\theta^2}{(n-1)(n-2)}.$$

Επομένως

$$\mathrm{MSE}_\theta(c/A_n) = \theta^2 \left\{ \frac{c^2}{(n-1)(n-2)} - \frac{2c}{n-1} + 1 \right\}.$$

Το δεξί μέλος είναι γνησίως κυρτό ως προς c και ελαχιστοποιείται στο $c = n-2$. Το ελάχιστο ΜΤΣ είναι $\theta^2/(n-1)$. Αντίθετα, η ε.μ.π. αντιστοιχεί στο $c = n$ και η αμερόληπτη εκτιμήτρια στο $c = n-1$. Τα αντίστοιχα ΜΤΣ είναι

$$\mathrm{MSE}_\theta(n/A_n) = \frac{(n+2)\theta^2}{(n-1)(n-2)}, \qquad \mathrm{MSE}_\theta((n-1)/A_n) = \frac{\theta^2}{n-2}.$$

Λύση 8. Κανονικό μοντέλο και διαδοχική μεγιστοποίηση

(α) Για γνωστή σ_0^2,

$$\ell_x(\mu) = C - \frac{1}{2\sigma_0^2}\sum_i (x_i - \mu)^2.$$

Η ταυτότητα

$$\sum_i (x_i - \mu)^2 = \sum_i (x_i - \overline{x})^2 + n(\mu - \overline{x})^2$$

δίνει το μοναδικό ολικό μέγιστο στο $\mu = \overline{x}$. Άρα $\widehat{\mu}_n = \overline{X}_n$.

(β) Θέτουμε $Q_0 = \sum_i (x_i - \mu_0)^2 > 0$. Τότε

$$\ell_x(\sigma^2) = C - \frac{n}{2}\log(\sigma^2) - \frac{Q_0}{2\sigma^2}, \qquad \frac{d\ell_x}{d(\sigma^2)} = \frac{Q_0 - n\sigma^2}{2(\sigma^2)^2}.$$

Η παράγωγος αλλάζει πρόσημο από θετικό σε αρνητικό στο Q_0/n. Άρα η ε.μ.π. είναι $n^{-1}\sum_i (X_i - \mu_0)^2$.

(γ) Για κάθε σταθερό $\sigma^2 > 0$, το (α) δίνει $\mu^*(\sigma^2) = \overline{x}$. Στη συνέχεια, το (β) εφαρμοσμένο στη συνάρτηση $\ell_x(\overline{x}, \sigma^2)$ δίνει

$$\boxed{\widehat{\mu}_n = \overline{X}_n, \qquad \widehat{\sigma^2}_n = M_{2,n}.}$$

Κάθε βήμα είναι ολική μεγιστοποίηση, οπότε το ζεύγος είναι κοινό ολικό μέγιστο.
Στην αντίστροφη σειρά, για κάθε σταθερό μ,

$$(\sigma^2)^*(\mu) = A(\mu) := \frac{1}{n}\sum_i (x_i - \mu)^2 > 0.$$

Αντικαθιστώντας,

$$\ell_x\big(\mu, A(\mu)\big) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log A(\mu) - \frac{n}{2}.$$

Η μεγιστοποίηση ως προς μ ισοδυναμεί με την ελαχιστοποίηση του $A(\mu)$ και δίνει $\overline{x}$. Επιστρέφοντας στην πρώτη λύση, παίρνουμε $A(\overline{x}) = n^{-1}\sum_i (x_i - \overline{x})^2$.

(δ) Επειδή $M_{2,n} = (n-1)S_n^2/n$ και $\mathbb{E}(S_n^2) = \sigma^2$,

$$\mathbb{E}(M_{2,n}) = \frac{n-1}{n}\sigma^2, \qquad \mathrm{Bias}(M_{2,n}) = -\frac{\sigma^2}{n}.$$

Η αμερόληπτη διόρθωση είναι η S_n^2. Για το δοσμένο μη σταθερό δείγμα διαφέρει από τη μοναδική τιμή που μεγιστοποιεί την προκύπτουσα πιθανοφάνεια ως προς τη διασπορά, άρα δεν είναι ε.μ.π.

Για σκέψη. Χρησιμοποιούμε $s^2 = nm_2/(n-1)$ και $\sum_i (x_i - \bar{x})^2 = nm_2$. Η διαφορά είναι

$$\ell_x(\bar{x}, m_2) - \ell_x(\bar{x}, s^2) = \frac{n}{2}\log\left(\frac{n}{n-1}\right) - \frac{1}{2}.$$

Είναι θετική και δεν εξαρτάται από το μη σταθερό δείγμα. Τείνει στο 0, αφού $n\log(n/(n-1)) \to 1$. Άρα

$$\frac{L_x(\bar{x}, s^2)}{L_x(\bar{x}, m_2)} \longrightarrow 1.$$

Η αμερόληπτη διόρθωση δεν είναι ε.μ.π. για πεπερασμένο n, παρότι η σχετική πιθανοφάνειά της τείνει στο 1.

Λύση 9. Κανονική κατανομή με μη αρνητική μέση τιμή

(α) Για κάθε σταθερή διασπορά μεγιστοποιούμε τη λογαριθμοπιθανοφάνεια ελαχιστοποιώντας το $\sum_i (x_i - \mu)^2$ υπό τον περιορισμό $\mu \geq 0$. Από τη διάσπαση γύρω από το $\overline{x}$, αρκεί να ελαχιστοποιήσουμε το $(\mu - \overline{x})^2$. Άρα

$$\widehat{\mu}_n = \max(0, \overline{X}_n).$$

(β) Αντικαθιστούμε το μ με τον μεγιστοποιητή του (α). Το δεύτερο μονοδιάστατο πρόβλημα της Άσκησης 8 δίνει

$$\widehat{\sigma^2}_n = \frac{1}{n}\sum_i (X_i - \widehat{\mu}_n)^2 = M_{2,n} + (\overline{X}_n - \widehat{\mu}_n)^2.$$

Επομένως

$$\boxed{(\widehat{\mu}_n, \widehat{\sigma^2}_n) = \begin{cases} (\overline{X}_n, M_{2,n}), & \overline{X}_n \geq 0, \\ (0, M'_{2,n}), & \overline{X}_n < 0. \end{cases}}$$

Το μέγιστο είναι ολικό και μοναδικό, διότι το ίδιο ισχύει σε κάθε ένα από τα δύο βήματα.

(γ) Στο δείγμα $(-2, 0, 1)$,

$$\overline{x} = -\frac{1}{3}, \qquad m'_2 = \frac{5}{3}, \qquad m_2 = \frac{5}{3} - \frac{1}{9} = \frac{14}{9}.$$

Στο περιορισμένο μοντέλο οι εκτιμήσεις είναι $(0, 5/3)$. Χωρίς τον περιορισμό είναι $(-1/3, 14/9)$. Ο περιορισμός στη μέση τιμή επηρεάζει και την εκτίμηση της διασποράς.

Για σκέψη. Στο $\mu = 0$, η $Z_n = \sqrt{n}\,\overline{X}_n/\sigma$ έχει ακριβώς κατανομή $\mathcal{N}(0,1)$. Από το (β),

$$\frac{n(\widehat{\sigma^2}_n - M_{2,n})}{\sigma^2} = \big(\min(Z_n, 0)\big)^2.$$

Η κατανομή έχει μάζα $1/2$ στο 0 και, με το υπόλοιπο βάρος $1/2$, κατανομή χ^2_1. Αυτό ισχύει για κάθε $n \geq 2$. Διαιρώντας με $\sqrt{n}$ προκύπτει

$$\sqrt{n}(\widehat{\sigma^2}_n - M_{2,n}) \xrightarrow{\mathbb{P}} 0.$$

Επομένως, από το κανονικό αποτέλεσμα του Φύλλου 3 και το Slutsky,

$$\sqrt{n}(\widehat{\sigma^2}_n - \sigma^2) \xrightarrow{d} \mathcal{N}(0, 2\sigma^4).$$

Ο περιορισμός αφήνει αμετάβλητο το όριο στην κλίμακα $\sqrt{n}$, όχι όμως τη διαφορά των δύο εκτιμητριών στην κλίμακα n.

Λύση 10. Ομοιόμορφες κατανομές με άγνωστο άκρο

(α) Για μη αρνητικές παρατηρήσεις,

$$L_x(\theta) = \theta^{-n}\mathbf{1}_{[x_{(n)},\infty)}(\theta), \qquad \theta > 0.$$

Η συνάρτηση είναι γνησίως φθίνουσα εκεί όπου είναι θετική. Εφόσον $x_{(n)} > 0$, το μοναδικό μέγιστο επιτυγχάνεται στο $x_{(n)}$:

$$\widehat{\theta}_n = X_{(n)}.$$

(β) Το δείγμα ανήκει στο $[-\theta, \theta]$ αν και μόνο αν

$$\theta \geq c_x := \max\{-x_{(1)}, x_{(n)}\} = \max_i |x_i|.$$

Άρα

$$L_x(\theta) = (2\theta)^{-n}\mathbf{1}_{[c_x,\infty)}(\theta), \qquad \boxed{\widehat{\theta}_n = \max_i |X_i|.}$$

(γ) Οι περιορισμοί είναι

$$-\theta \leq x_{(1)}, \qquad x_{(n)} \leq \theta^2.$$

Θέτοντας $u^+ = \max(u, 0)$, αυτοί ισοδυναμούν, για $\theta > 0$, με

$$\theta \geq d_x := \max\{-x_{(1)}, \sqrt{x_{(n)}^+}\}.$$

Το d_x είναι θετικό επειδή οι παρατηρήσεις δεν είναι όλες μηδέν. Το μήκος του διαστήματος είναι $\theta^2 + \theta$, οπότε

$$L_x(\theta) = (\theta^2 + \theta)^{-n}\mathbf{1}_{[d_x,\infty)}(\theta).$$

Η $\theta^2 + \theta$ αυξάνεται γνησίως για $\theta > 0$. Επομένως

$$\boxed{\widehat{\theta}_n = \max\{-X_{(1)}, \sqrt{X_{(n)}^+}\}.}$$

(δ) Στο (α), $\mathbb{E}_\theta(X_1) = \theta/2$ και η εκτιμήτρια ροπών είναι $2\overline{X}_n$. Στο (β), $\mathbb{E}_\theta(X_1) = 0$ δεν προσδιορίζει το θ, ενώ $\mathbb{E}_\theta(X_1^2) = \theta^2/3$. Άρα η εκτιμήτρια ροπών είναι

$$\sqrt{3M'_{2,n}} = \sqrt{\frac{3}{n}\sum_i X_i^2}.$$

Οι εκτιμήτριες ροπών δεν συμπίπτουν γενικά με τις ε.μ.π. Χρησιμοποιούμε τη δεύτερη μη κεντρική ροπή στο (β), όχι το $M_{2,n}$.

Για σκέψη. Το μέγιστο των παρατηρήσεων δεν αλλάζει, οπότε η ε.μ.π. παραμένει το $x_{(n)}$. Αν L_n και L_{n+m} είναι οι δύο πιθανοφάνειες, οι λόγοι είναι

$$\frac{L_n(\theta_2)}{L_n(\theta_1)} = \left(\frac{\theta_1}{\theta_2}\right)^n, \qquad \frac{L_{n+m}(\theta_2)}{L_{n+m}(\theta_1)} = \left(\frac{\theta_1}{\theta_2}\right)^{n+m}.$$

Ο λόγος πολλαπλασιάζεται με το $(\theta_1/\theta_2)^m < 1$. Οι νέες παρατηρήσεις ενισχύουν τη σχετική πιθανοφάνεια της μικρότερης τιμής, χωρίς να μεταβάλουν τον μεγιστοποιητή. Η ίδια ε.μ.π. δεν σημαίνει ίδια συνάρτηση πιθανοφάνειας.

Λύση 11. Ομοιόμορφη κατανομή θέσης και μη μοναδική ε.μ.π.

(α) Η πιθανοφάνεια ισούται με 1 αν και μόνο αν

$$\theta \le x_{(1)}, \qquad x_{(n)} \le \theta + 1.$$

Άρα

$$L_x(\theta) = \mathbf{1}_{[x_{(n)}-1,\, x_{(1)}]}(\theta), \qquad \arg\max_{\theta\in\mathbb{R}} L_x(\theta) = [x_{(n)} - 1, x_{(1)}].$$

Το διάστημα έχει θετικό μήκος. Τρεις επιλογές ε.μ.π. είναι

$$T_n = X_{(n)} - 1, \qquad U_n = X_{(1)}, \qquad W_n = \frac{X_{(1)} + X_{(n)} - 1}{2}.$$

(β) Για $n = 2$, $\overline{X}_2 = (X_{(1)} + X_{(2)})/2$, οπότε $\overline{X}_2 - 1/2 = W_2$ και είναι ε.μ.π. Για το δείγμα $(0, 0, 0.9)$, το διάστημα των μεγιστοποιητών είναι $[-0.1, 0]$, ενώ

$$\overline{x} - \frac{1}{2} = 0.3 - 0.5 = -0.2.$$

Η εκτίμηση ροπών βρίσκεται εκτός του διαστήματος και έχει μηδενική πιθανοφάνεια. Δεν είναι ε.μ.π.

(γ) Γράφουμε $X_i = \theta + Y_i$, όπου οι Y_i είναι ανεξάρτητες $\mathrm{Unif}[0, 1]$. Από το αποτέλεσμα για το μέγιστο του Φύλλου 3,

$$\max_{i\le n} Y_i \xrightarrow{\text{a.s.}} 1.$$

Εφαρμόζοντας το ίδιο αποτέλεσμα στις $1 - Y_i$, παίρνουμε $\min_{i\le n} Y_i \xrightarrow{\text{a.s.}} 0$. Επομένως

$$X_{(n)} - 1 \xrightarrow{\text{a.s.}} \theta, \qquad X_{(1)} \xrightarrow{\text{a.s.}} \theta.$$

Κάθε εκτιμήτρια A_n με

$$X_{(n)} - 1 \le A_n \le X_{(1)}$$

συγκλίνει σχεδόν βέβαια στο θ, από το θεώρημα παρεμβολής. Η μη μοναδικότητα της ε.μ.π. δεν εμποδίζει εδώ τη συνέπεια κάθε τέτοιας επιλογής.

Για σκέψη. Από τις δοσμένες μέσες τιμές,

$$\mathbb{E}_\theta(A_{n,c}) - \theta = \frac{2c - 1}{n + 1}.$$

Άρα η μοναδική αμερόληπτη επιλογή είναι η $c = 1/2$, δηλαδή η W_n του (α). Με τις δοσμένες διασπορές και τη συνδιακύμανση,

$$\mathbb{V}_\theta(A_{n,c}) = \frac{n\{c^2 + (1 - c)^2\} + 2c(1 - c)}{(n + 1)^2(n + 2)}.$$

Προσθέτοντας το τετράγωνο της μεροληψίας και απλοποιώντας,

$$\mathrm{MSE}_\theta(A_{n,c}) = \frac{1/2 + 6(c - 1/2)^2}{(n + 1)(n + 2)}.$$

Η $c = 1/2$ ελαχιστοποιεί και το ΜΤΣ, για κάθε θ. Τα δύο άκρα, $c = 0$ και $c = 1$, έχουν τετραπλάσιο ΜΤΣ από τη μεσαία επιλογή, παρότι και οι τρεις έχουν την ίδια μέγιστη πιθανοφάνεια.

Λύση 12. Ομοιόμορφη κατανομή με δύο άγνωστα άκρα

(α) Η πιθανοφάνεια είναι

$$L_x(a,b) = (b-a)^{-n}\mathbf{1}_{\{a\le x_{(1)},\ b\ge x_{(n)}\}}, \qquad a<b.$$

Για γνωστό $a \le x_{(1)}$, μειώνεται γνησίως ως προς το επιτρεπτό b, άρα $\widehat{b}_n = X_{(n)}$. Για γνωστό $b \ge x_{(n)}$, αυξάνεται γνησίως ως προς το επιτρεπτό a, άρα $\widehat{a}_n = X_{(1)}$.

(β) Για κάθε $b \ge x_{(n)}$, η πρώτη μεγιστοποίηση δίνει $a^*(b) = x_{(1)}$. Μετά μεγιστοποιούμε τη συνάρτηση $b \mapsto -n\log(b - x_{(1)})$ ως προς $b \ge x_{(n)}$ και βρίσκουμε $b = x_{(n)}$. Το μοναδικό ζεύγος ε.μ.π. είναι

$$(\widehat{a}_n, \widehat{b}_n) = (X_{(1)}, X_{(n)}).$$

Είναι επιτρεπτό επειδή $x_{(1)} < x_{(n)}$.

(γ) Οι εξισώσεις των δύο πρώτων ροπών γράφονται ισοδύναμα

$$\frac{a+b}{2} = \overline{X}_n, \qquad \frac{(b-a)^2}{12} = M_{2,n}.$$

Άρα

$$a_n^{\mathrm{MM}} = \overline{X}_n - \sqrt{3M_{2,n}}, \qquad b_n^{\mathrm{MM}} = \overline{X}_n + \sqrt{3M_{2,n}}.$$

Στο δείγμα $(0,0,0,0,1)$ έχουμε $\bar{x} = 1/5$ και $m_2 = 4/25$. Επομένως

$$b^{\mathrm{MM}} = \frac{1+2\sqrt{3}}{5} < 1.$$

Το εκτιμώμενο διάστημα δεν περιέχει τη μεγαλύτερη παρατήρηση. Η πιθανοφάνεια στο ζεύγος των εκτιμήσεων ροπών είναι 0.

(δ) Οι μεροληψίες των ε.μ.π. είναι $(b-a)/(n+1)$ και $-(b-a)/(n+1)$, αντίστοιχα. Λύνοντας τις δύο γραμμικές σχέσεις μέσων τιμών, βρίσκουμε τις αμερόληπτες εκτιμήτριες

$$\widetilde{a}_n = \frac{nX_{(1)} - X_{(n)}}{n-1}, \qquad \widetilde{b}_n = \frac{nX_{(n)} - X_{(1)}}{n-1}.$$

(ε) Αν $r = x_{(n)} - x_{(1)} > 0$, το διορθωμένο διάστημα περιέχει το αρχικό και έχει μήκος $(n+1)r/(n-1)$. Άρα

$$\frac{L_x(\widetilde{a},\widetilde{b})}{L_x(\widehat{a},\widehat{b})} = \left(\frac{n-1}{n+1}\right)^n < 1.$$

Η αμερόληπτη διόρθωση δεν διατηρεί τη μέγιστη πιθανοφάνεια.

Για σκέψη. Ο λογάριθμος του λόγου τείνει στο -2, άρα ο λόγος τείνει στο e^{-2}, όχι στο 1. Οι απόλυτες διαφορές των διορθωμένων άκρων από τις ε.μ.π. είναι και οι δύο

$$\frac{X_{(n)} - X_{(1)}}{n-1} \le \frac{b-a}{n-1} \longrightarrow 0.$$

Τα άκρα του δείγματος συγκλίνουν σχεδόν βέβαια στα a, b, οπότε και το διορθωμένο ζεύγος είναι ισχυρά συνεπές. Η πιθανοφάνεια μεταβάλλεται με το n: η μικρή απόσταση των εκτιμήσεων δεν συνεπάγεται λόγο που τείνει στο 1. Στο κανονικό παράδειγμα της Άσκησης 8 ο αντίστοιχος λόγος έτεινε στο 1.

Λύση 13. Διαδοχική μεγιστοποίηση σε μη καρτεσιανό σύνολο

(α) Για σταθερό y, επιλέγουμε το $x \in [0, y]$ που βρίσκεται πλησιέστερα στο $3/4$. Άρα

$$x^*(y) = \min(y, 3/4).$$

Για $0 \le y \le 3/4$,

$$f(x^*(y), y) = -(y - 3/4)^2 - (y - 1/4)^2 = -2(y - 1/2)^2 - \frac{1}{8}.$$

Το μέγιστο σε αυτό το διάστημα είναι $-1/8$ και επιτυγχάνεται μοναδικά στο $y = 1/2$. Για $3/4 \le y \le 1$,

$$f(x^*(y), y) = -(y - 1/4)^2 \le -\frac{1}{4}.$$

Επομένως

$$\boxed{(x, y) = (1/2, 1/2), \qquad \max_D f = -1/8.}$$

(β) Για σταθερό x, επιλέγουμε το $y \in [x, 1]$ που βρίσκεται πλησιέστερα στο $1/4$:

$$y^*(x) = \max(x, 1/4).$$

Για $0 \le x \le 1/4$ η τιμή είναι $-(x - 3/4)^2 \le -1/4$. Για $1/4 \le x \le 1$ είναι

$$-(x - 3/4)^2 - (x - 1/4)^2 = -2(x - 1/2)^2 - \frac{1}{8}.$$

Παίρνουμε ξανά το μοναδικό ζεύγος $(1/2, 1/2)$.

(γ) Ένα ζεύγος μένει αμετάβλητο αν και μόνο αν

$$x = \min(y, 3/4), \qquad y = \max(x, 1/4).$$

Οι δύο σχέσεις επιβάλλουν $x = y = t$ με $1/4 \le t \le 3/4$, και κάθε τέτοιο ζεύγος τις ικανοποιεί. Άρα η διαδικασία μπορεί να σταματήσει σε οποιοδήποτε σημείο του διαγώνιου τμήματος

$$\{(t, t) : 1/4 \le t \le 3/4\}.$$

Μόνο το $(1/2, 1/2)$ είναι ολικό μέγιστο. Για παράδειγμα, στο $(1/4, 1/4)$ καμία μεμονωμένη αντικατάσταση δεν αλλάζει το ζεύγος, αλλά η τιμή είναι $-1/4$, αντί της μέγιστης $-1/8$.

Η πλήρης διαδοχική μεγιστοποίηση δεν κρατά μία μόνο τιμή της δεύτερης μεταβλητής: σχηματίζει τη συνάρτηση $y \mapsto f(x^*(y), y)$ σε όλο το $[0, 1]$ και τη μεγιστοποιεί. Αυτό διαφέρει από τις εναλλάξ ενημερώσεις ενός συγκεκριμένου ζεύγους.

Για σκέψη. Στη διαγώνιο έχουμε

$$f(t, t) = -2(t - 1/2)^2 - \frac{1}{8}.$$

Για οποιοδήποτε $0 < \alpha \le 1$ θέτουμε $h = \alpha(1/2 - t)$. Τότε το $t + h$ βρίσκεται μεταξύ των t και $1/2$, άρα το νέο ζεύγος παραμένει στο D, και

$$f(t + h, t + h) - f(t, t) = 2\big[1 - (1 - \alpha)^2\big](t - 1/2)^2 > 0.$$

Η κοινή μεταβολή επιτρέπει βελτίωση που καμία μεταβολή μίας μόνο συντεταγμένης, με την άλλη σταθερή και με τους περιορισμούς του D, δεν επιτυγχάνει.

Λύση 14. Μοναδικό στάσιμο σημείο: τοπικό και ολικό μέγιστο

(α) Οι πρώτες μερικές παράγωγοι είναι

$$f_x(x,y) = -2x + 3(1-x)^2y^2, \qquad f_y(x,y) = -2(1-x)^3y.$$

Από $f_y = 0$ παίρνουμε $y = 0$ ή $x = 1$. Στην περίπτωση $x = 1$ έχουμε $f_x = -2 \neq 0$. Άρα $y = 0$ και η $f_x = 0$ δίνει $x = 0$. Το $(0,0)$ είναι το μοναδικό στάσιμο σημείο.

(β) Έχουμε

$$f_{xx} = -2 - 6(1-x)y^2, \qquad f_{xy} = 6(1-x)^2y, \qquad f_{yy} = -2(1-x)^3.$$

Συνεπώς

$$H_f(0,0) = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}.$$

Ο Εσσιανός είναι αρνητικά ορισμένος, άρα το $(0,0)$ είναι γνήσιο τοπικό μέγιστο με τιμή 0.

(γ) Για $x = 2$,

$$f(2,y) = -4 + y^2 \longrightarrow +\infty \qquad (|y| \to \infty).$$

Άρα το τοπικό μέγιστο δεν είναι ολικό. Η μοναδικότητα του στάσιμου σημείου και το αρνητικά ορισμένο του Εσσιανού σε αυτό δεν αρκούν για ολικό μέγιστο.

(δ) Πρώτα, δεν υπάρχει $z \in I \setminus \{x_0\}$ με $\varphi(z) = \varphi(x_0)$: αν υπήρχε, το Rolle στο διάστημα με άκρα x_0, z θα έδινε δεύτερο στάσιμο σημείο. Άρα το τοπικό μέγιστο είναι γνήσιο.

Αν υπήρχε $b > x_0$ με $\varphi(b) > \varphi(x_0)$, θα επιλέγαμε $a \in (x_0, b)$ αρκετά κοντά στο x_0, ώστε

$$\varphi(a) < \varphi(x_0) < \varphi(b).$$

Η συνέχεια στο $[a,b]$ θα έδινε $z \in (a,b)$ με $\varphi(z) = \varphi(x_0)$, αντίφαση. Για $b < x_0$ το ίδιο επιχείρημα εφαρμόζεται στο αντίστοιχο διάστημα αριστερά του x_0. Συνεπώς $\varphi(x) < \varphi(x_0)$ για κάθε $x \neq x_0$. Χρησιμοποιήθηκε η συνέχεια της φ, που έπεται από την παραγωγισιμότητα, όχι η συνέχεια της φ'.

Για σκέψη. Θέτοντας $x = y = t$, περιορίζουμε την f στη διαγώνιο και παίρνουμε

$$\psi(t) = -2t^2 + 3t^3 - 3t^4 + t^5, \qquad \psi'(t) = -4t + 9t^2 - 12t^3 + 5t^4.$$

Έχουμε $\psi'(1) = -2$ και $\psi'(2) = 12$. Επειδή η ψ' είναι συνεχής, υπάρχει c με $1 < c < 2$ και $\psi'(c) = 0$. Η τιμή c της μονοδιάστατης μεταβλητής αντιστοιχεί στο σημείο (c,c) του επιπέδου. Από τον κανόνα της αλυσίδας,

$$\psi'(c) = f_x(c,c) + f_y(c,c) = 0.$$

Μηδενίζεται λοιπόν η παράγωγος κατά μήκος της διαγωνίου, όχι καθεμία από τις δύο μερικές παραγώγους. Πράγματι, $f_y(c,c) = -2(1-c)^3c > 0$, αφού $c > 1$. Άρα το c είναι στάσιμο σημείο της ψ, αλλά το (c,c) δεν είναι στάσιμο σημείο της f.

Λύση 15. Student: ο παρονομαστής αλλάζει την ακριβή κατανομή

(α) Από τον ορισμό της Student και την ανεξαρτησία των Z_n, Q_n,

$$T_n = \frac{Z_n}{\sqrt{Q_n/(n-1)}} \sim t_{n-1}.$$

Επειδή $M_{2,n} = (n-1)S_n^2/n$,

$$R_n = \sqrt{\frac{n}{n-1}}\, T_n.$$

Η R_n έχει την κατανομή ενός πολλαπλασίου της t_{n-1}, όχι την ίδια t_{n-1}. Ο ζητούμενος συντελεστής είναι $\sqrt{(n-1)/n}$.

(β) Από τη δοσμένη ταυτότητα,

$$\frac{nV_n}{\sigma^2} = Q_n + Z_n^2, \qquad U_n = \frac{\sqrt{n}\, Z_n}{\sqrt{Q_n + Z_n^2}}.$$

Αντικαθιστώντας το Z_n μέσω της T_n, προκύπτει

$$U_n = \frac{\sqrt{n}\, T_n}{\sqrt{n - 1 + T_n^2}}.$$

Η συνάρτηση $t \mapsto \sqrt{n}\, t/\sqrt{n-1+t^2}$ είναι γνησίως αύξουσα και απεικονίζει την $\mathbb{R}$ στο $(-\sqrt{n}, \sqrt{n})$. Άρα η U_n παίρνει τιμές σε αυτό το ανοικτό διάστημα με πιθανότητα 1 και το στήριγμα της κατανομής της είναι το $[-\sqrt{n}, \sqrt{n}]$. Δεν έχει κατανομή Student, αφού κάθε t_ν έχει θετική πυκνότητα σε ολόκληρη την $\mathbb{R}$.

Η πληροφορία ότι $nV_n/\sigma^2 \sim \chi^2_n$ δεν αρκεί για εφαρμογή του ορισμού της Student. Εδώ ο παρονομαστής περιέχει το Z_n^2. Πράγματι, η δεσμευμένη μέση τιμή του $Q_n + Z_n^2$ δοθέντος του $Z_n = z$ είναι $n - 1 + z^2$ και εξαρτάται από το z.

(γ) Η ισχυρή συνέπεια δίνει $S_n \xrightarrow{\text{a.s.}} \sigma$ και $M_{2,n} \xrightarrow{\text{a.s.}} \sigma^2$. Επιπλέον, ο ισχυρός νόμος στις $(X_i - \mu)^2$ δίνει $V_n \xrightarrow{\text{a.s.}} \sigma^2$. Σε κάθε λόγο ο αριθμητής διαιρεμένος με το σ έχει κατανομή $\mathcal{N}(0,1)$. Με το λήμμα Slutsky,

$$T_n \xrightarrow{d} \mathcal{N}(0,1), \qquad R_n \xrightarrow{d} \mathcal{N}(0,1), \qquad U_n \xrightarrow{d} \mathcal{N}(0,1).$$

Οι τρεις ακριβείς κατανομές είναι διαφορετικές, παρότι έχουν το ίδιο όριο.

Για σκέψη. Θέτουμε $Y_i = (X_i - \mu)/\sigma$. Τότε

$$U_n = \frac{\sum_i Y_i}{\sqrt{\sum_i Y_i^2}}.$$

Η U_n είναι φραγμένη και η ταυτόχρονη αλλαγή όλων των προσήμων αλλάζει το πρόσημό της χωρίς να αλλάζει την κατανομή. Άρα $\mathbb{E}(U_n) = 0$. Επιπλέον,

$$U_n^2 = 1 + 2\sum_{i<j} \frac{Y_i Y_j}{\sum_k Y_k^2}.$$

Κάθε όρος του αθροίσματος είναι ολοκληρώσιμος, αφού έχει απόλυτη τιμή το πολύ $1/2$. Η αλλαγή προσήμου μόνο του Y_i δείχνει ότι η μέση τιμή του όρου είναι 0. Επομένως

$$\mathbb{E}(U_n^2) = \mathbb{V}(U_n) = 1.$$

Για $n > 3$, η ανεξαρτησία των Z_n, Q_n δίνει

$$\mathbb{V}(T_n) = \mathbb{E}(T_n^2) = (n-1)\mathbb{E}(Z_n^2)\mathbb{E}(Q_n^{-1}) = \frac{n-1}{n-3} > 1.$$