Ασκήσεις: Πολυμεταβλητή Κανονική Κατανομή

Στατιστική ΙΙ — Τμήμα Οικονομικής Επιστήμης, ΟΠΑ
Διάλεξη 6

Σε αυτό το συμπληρωματικό υλικό παρουσιάζουμε 15 αναλυτικά λυμένα παραδείγματα που καλύπτουν το περιεχόμενο της Διάλεξης 6: τον ορισμό και τις ιδιότητες της πολυμεταβλητής κανονικής κατανομής, τη ροπογεννήτρια συνάρτηση, τις περιθώριες και τις δεσμευμένες κατανομές, καθώς και τους γραμμικούς μετασχηματισμούς. Τα παραδείγματα ακολουθούν αυστηρά τον συμβολισμό της διάλεξης.

1. Ορισμός και ροπές της πολυμεταβλητής κανονικής

Παράδειγμα 1: Διμεταβλητή κανονική πυκνότητα — Σαφής μορφή

Έστω $\boldsymbol{X} = (X_1, X_2)' \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ με

$$\boldsymbol{\mu} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \qquad \boldsymbol{\Sigma} = \begin{pmatrix} 4 & 2 \\ 2 & 9 \end{pmatrix}.$$

(α΄) Υπολογίστε την ορίζουσα $|\boldsymbol{\Sigma}|$ και τον αντίστροφο πίνακα $\boldsymbol{\Sigma}^{-1}$.

(β΄) Υπολογίστε τον συντελεστή συσχέτισης ρ_{12}.

(γ΄) Γράψτε σε σαφή μορφή τη συνάρτηση πυκνότητας $f(x_1, x_2)$.

(δ΄) Υπολογίστε $f(1, 2)$.

Λύση

(α΄) Για πίνακα 2×2: $|\boldsymbol{\Sigma}| = 4 \cdot 9 - 2 \cdot 2 = 32$. Ο αντίστροφος:

$$\boldsymbol{\Sigma}^{-1} = \frac{1}{32} \begin{pmatrix} 9 & -2 \\ -2 & 4 \end{pmatrix}.$$

(β΄) Έχουμε $\sigma_{11} = 4,\ \sigma_{22} = 9,\ \sigma_{12} = 2$. Άρα:

$$\rho_{12} = \frac{\sigma_{12}}{\sqrt{\sigma_{11}\sigma_{22}}} = \frac{2}{\sqrt{4 \cdot 9}} = \frac{2}{6} = \frac{1}{3}.$$

(γ΄) Για $\boldsymbol{x} = (x_1, x_2)'$, θέτουμε $\boldsymbol{x} - \boldsymbol{\mu} = (x_1 - 1,\ x_2 - 2)'$. Η τετραγωνική μορφή είναι:

$$(\boldsymbol{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{\mu}) = \frac{1}{32}\big[9(x_1 - 1)^2 - 4(x_1 - 1)(x_2 - 2) + 4(x_2 - 2)^2\big].$$

Επομένως:

$$f(x_1, x_2) = \frac{1}{2\pi\sqrt{32}} \exp\left\{-\frac{1}{64}\left[9(x_1 - 1)^2 - 4(x_1 - 1)(x_2 - 2) + 4(x_2 - 2)^2\right]\right\}.$$

(δ΄) Στο σημείο $(1, 2)$ η τετραγωνική μορφή μηδενίζεται:

$$f(1, 2) = \frac{1}{2\pi\sqrt{32}} = \frac{1}{8\pi\sqrt{2}} \approx 0{,}02813.$$

Παρατήρηση

Το μέγιστο της πυκνότητας επιτυγχάνεται πάντοτε στο σημείο $\boldsymbol{x} = \boldsymbol{\mu}$. Στο ίδιο σημείο η τετραγωνική μορφή είναι μηδέν και η τιμή της πυκνότητας εξαρτάται αποκλειστικά από τη σταθερά κανονικοποίησης $(2\pi)^{-n/2}|\boldsymbol{\Sigma}|^{-1/2}$.

Παράδειγμα 2: Ροπογεννήτρια — Εύρεση ροπών

Έστω $\boldsymbol{X} = (X_1, X_2, X_3)' \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Χρησιμοποιώντας τη ροπογεννήτρια συνάρτηση

$$M_{\boldsymbol{X}}(\boldsymbol{t}) = \exp\left(\boldsymbol{t}'\boldsymbol{\mu} + \tfrac{1}{2}\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t}\right), \qquad \boldsymbol{t} \in \mathbb{R}^3,$$

δείξτε ότι:

(α΄) $\mathbb{E}[X_i] = \mu_i$, για $i = 1, 2, 3$.

(β΄) $\text{Var}(X_i) = \sigma_{ii}$ και $\text{Cov}(X_i, X_j) = \sigma_{ij}$.

Λύση

(α΄) Από τη θεωρία, $\mathbb{E}[X_i] = \left.\dfrac{\partial M_{\boldsymbol{X}}(\boldsymbol{t})}{\partial t_i}\right|_{\boldsymbol{t}=\boldsymbol{0}}$. Παραγωγίζουμε τον εκθέτη:

$$\frac{\partial}{\partial t_i}\left(\boldsymbol{t}'\boldsymbol{\mu} + \tfrac{1}{2}\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t}\right) = \mu_i + (\boldsymbol{\Sigma}\boldsymbol{t})_i.$$

Άρα

$$\frac{\partial M_{\boldsymbol{X}}}{\partial t_i}(\boldsymbol{t}) = \left[\mu_i + (\boldsymbol{\Sigma}\boldsymbol{t})_i\right] M_{\boldsymbol{X}}(\boldsymbol{t}).$$

Για $\boldsymbol{t} = \boldsymbol{0}$ έχουμε $M_{\boldsymbol{X}}(\boldsymbol{0}) = 1$ και $(\boldsymbol{\Sigma}\boldsymbol{0})_i = 0$, οπότε $\mathbb{E}[X_i] = \mu_i$.

(β΄) Για τη δεύτερη ροπή:

$$\frac{\partial^2 M_{\boldsymbol{X}}}{\partial t_i \partial t_j}(\boldsymbol{t}) = \left[\sigma_{ij} + (\mu_i + (\boldsymbol{\Sigma}\boldsymbol{t})_i)(\mu_j + (\boldsymbol{\Sigma}\boldsymbol{t})_j)\right] M_{\boldsymbol{X}}(\boldsymbol{t}).$$

Στο $\boldsymbol{t} = \boldsymbol{0}$:

$$\mathbb{E}[X_i X_j] = \sigma_{ij} + \mu_i\mu_j \quad \Rightarrow \quad \text{Cov}(X_i, X_j) = \mathbb{E}[X_i X_j] - \mu_i\mu_j = \sigma_{ij}.$$

Ειδικότερα, για $i = j$: $\text{Var}(X_i) = \sigma_{ii}$.

Παρατήρηση

Η ροπογεννήτρια της πολυμεταβλητής κανονικής καθορίζει πλήρως την κατανομή. Συνεπώς δύο πολυμεταβλητά κανονικά διανύσματα με την ίδια μέση τιμή $\boldsymbol{\mu}$ και τον ίδιο πίνακα συνδιακυμάνσεων $\boldsymbol{\Sigma}$ έχουν ταυτόσημη κατανομή.

2. Ανεξαρτησία και από κοινού κανονικότητα

Παράδειγμα 3: Ασυσχέτιστες ⇒ Ανεξάρτητες (από κοινού κανονικότητα)

Έστω $\boldsymbol{X} = (X_1, X_2)' \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ με

$$\boldsymbol{\mu} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}.$$

Δείξτε ότι οι X_1, X_2 είναι ανεξάρτητες παραγοντοποιώντας άμεσα τη συνάρτηση πυκνότητας.

Λύση

Έχουμε $|\boldsymbol{\Sigma}| = \sigma_1^2\sigma_2^2$ και $\boldsymbol{\Sigma}^{-1} = \begin{pmatrix} 1/\sigma_1^2 & 0 \\ 0 & 1/\sigma_2^2 \end{pmatrix}$. Η τετραγωνική μορφή γίνεται:

$$(\boldsymbol{x} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu}) = \frac{x_1^2}{\sigma_1^2} + \frac{x_2^2}{\sigma_2^2}.$$

Άρα:

$$\begin{aligned} f(x_1, x_2) &= \frac{1}{2\pi\sigma_1\sigma_2} \exp\left\{ -\frac{x_1^2}{2\sigma_1^2} - \frac{x_2^2}{2\sigma_2^2} \right\} \\ &= \underbrace{\frac{1}{\sqrt{2\pi}\,\sigma_1} e^{-x_1^2/(2\sigma_1^2)}}_{f_{X_1}(x_1)} \cdot \underbrace{\frac{1}{\sqrt{2\pi}\,\sigma_2} e^{-x_2^2/(2\sigma_2^2)}}_{f_{X_2}(x_2)}. \end{aligned}$$

Η από κοινού πυκνότητα παραγοντοποιείται ως γινόμενο δύο μονοδιάστατων κανονικών πυκνοτήτων. Άρα $X_1 \perp\!\!\!\perp X_2$, με $X_1 \sim \mathcal{N}(0, \sigma_1^2)$ και $X_2 \sim \mathcal{N}(0, \sigma_2^2)$.

Παρατήρηση

Η Πρόταση της διάλεξης: *Αν οι $X_1, \ldots, X_n$ είναι από κοινού κανονικές και ανά δύο ασυσχέτιστες (δηλ. $\rho_{ij} = 0$ για κάθε $i \neq j$), τότε είναι ανεξάρτητες.* Η από κοινού κανονικότητα είναι ουσιαστική υπόθεση· βλ. Παραδείγματα 4 και 13.

Παράδειγμα 4: Κανονικές περιθώριες χωρίς από κοινού κανονικότητα

Έστω $X \sim \mathcal{N}(0, 1)$ και W μια τ.μ. ανεξάρτητη από την X με $P(W = 1) = P(W = -1) = 1/2$. Ορίζουμε $Y = W \cdot X$.

(α΄) Δείξτε ότι $Y \sim \mathcal{N}(0, 1)$.

(β΄) Δείξτε ότι το ζεύγος (X, Y) *δεν* ακολουθεί διμεταβλητή κανονική κατανομή, παρότι κάθε μία από τις X, Y είναι κανονική.

Λύση

(α΄) Υπολογίζουμε τη συνάρτηση κατανομής της Y:

$$\begin{aligned} F_Y(y) &= P(WX \le y) \\ &= P(W=1)P(X \le y) + P(W=-1)P(-X \le y) \\ &= \tfrac{1}{2}\Phi(y) + \tfrac{1}{2}P(X \ge -y) = \tfrac{1}{2}\Phi(y) + \tfrac{1}{2}\Phi(y) = \Phi(y), \end{aligned}$$

όπου χρησιμοποιήσαμε τη συμμετρία της $\mathcal{N}(0,1)$: $P(X \ge -y) = \Phi(y)$. Άρα $Y \sim \mathcal{N}(0,1)$.

(β΄) Παρατηρούμε ότι $X + Y = X(1+W)$. Όμως $1 + W \in \{0, 2\}$ με πιθανότητα $1/2$ η καθεμία. Άρα

$$P(X+Y=0) = P(W=-1) = \tfrac{1}{2}.$$

Εάν το (X,Y) ήταν από κοινού κανονικό, τότε η $X+Y$ θα ήταν (μονοδιάστατη) κανονική, δηλαδή συνεχής τ.μ. — κάθε σημείο θα είχε πιθανότητα 0. Αυτό αντιφάσκει με $P(X+Y=0) = 1/2$. Άρα (X,Y) *δεν* είναι από κοινού κανονικό.

Παρατήρηση

Το παράδειγμα αυτό δείχνει γιατί είναι λάθος να συμπεραίνει κανείς από κοινού κανονικότητα από τις περιθώριες. Η από κοινού κανονική κατανομή απαιτεί πολύ ισχυρότερη συνθήκη: *κάθε* γραμμικός συνδυασμός $a_1X_1 + \cdots + a_nX_n$ να είναι μονοδιάστατη κανονική τ.μ.

3. Περιθώριες κατανομές

Παράδειγμα 5: Περιθώρια από διμεταβλητή κανονική

Έστω $(X_1, X_2)' \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ με $\boldsymbol{\mu} = (\mu_1, \mu_2)'$ και $\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}$. Δείξτε ότι $X_1 \sim \mathcal{N}(\mu_1, \sigma_{11})$ χρησιμοποιώντας τη ροπογεννήτρια.

Λύση

Παρατηρούμε ότι $X_1 = \boldsymbol{t}'\boldsymbol{X}$ με $\boldsymbol{t} = (1,0)'$. Τότε:

$$M_{X_1}(s) = \mathbb{E}[e^{sX_1}] = \mathbb{E}[e^{s\boldsymbol{t}'\boldsymbol{X}}] = M_{\boldsymbol{X}}(s\boldsymbol{t}).$$

Αντικαθιστώντας $s\boldsymbol{t} = (s,0)'$ στη ροπογεννήτρια:

$$\begin{aligned} M_{X_1}(s) &= \exp\big((s\boldsymbol{t})'\boldsymbol{\mu} + \tfrac{1}{2}(s\boldsymbol{t})'\boldsymbol{\Sigma}(s\boldsymbol{t})\big) \\ &= \exp\big(s\mu_1 + \tfrac{1}{2}s^2\sigma_{11}\big) . \end{aligned}$$

Αυτή είναι ακριβώς η ροπογεννήτρια της $\mathcal{N}(\mu_1, \sigma_{11})$, άρα $X_1 \sim \mathcal{N}(\mu_1, \sigma_{11})$. Κατά συμμετρία, $X_2 \sim \mathcal{N}(\mu_2, \sigma_{22})$.

Παρατήρηση

Η τεχνική αυτή γενικεύεται: αν $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ είναι n-διάστατο και $\boldsymbol{a} \in \mathbb{R}^n$, τότε $\boldsymbol{a}'\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{a}'\boldsymbol{\mu},\ \boldsymbol{a}'\boldsymbol{\Sigma}\boldsymbol{a})$. Επιλέγοντας κατάλληλες τιμές του $\boldsymbol{a}$ παίρνουμε τις περιθώριες, αθροίσματα και άλλους γραμμικούς συνδυασμούς.

4. Δεσμευμένες κατανομές

Παράδειγμα 6: Δεσμευμένη κατανομή — Διμεταβλητή περίπτωση

Έστω $(X_1, X_2)' \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ με

$$\boldsymbol{\mu} = \begin{pmatrix} 3 \\ 5 \end{pmatrix}, \qquad \boldsymbol{\Sigma} = \begin{pmatrix} 4 & 3 \\ 3 & 9 \end{pmatrix}.$$

(α´) Βρείτε τη δεσμευμένη κατανομή της X_1 δεδομένου ότι $X_2 = x_2$.

(β´) Υπολογίστε $\mathbb{E}[X_1 \mid X_2 = 8]$ και $\mathrm{Var}(X_1 \mid X_2 = 8)$.

(γ´) Υπολογίστε τον συντελεστή συσχέτισης ρ_{12} και επαληθεύστε ότι $\mathrm{Var}(X_1 \mid X_2 = x_2) = \sigma_{11}(1 - \rho_{12}^2)$.

Λύση

(α´) Από τον τύπο της δεσμευμένης κατανομής στη διμεταβλητή κανονική:

$$\mathbb{E}[X_1 \mid X_2 = x_2] = \mu_1 + \frac{\sigma_{12}}{\sigma_{22}}(x_2 - \mu_2) = 3 + \frac{3}{9}(x_2 - 5) = 3 + \frac{x_2 - 5}{3},$$
$$\mathrm{Var}(X_1 \mid X_2 = x_2) = \sigma_{11} - \frac{\sigma_{12}^2}{\sigma_{22}} = 4 - \frac{9}{9} = 3.$$

Άρα $X_1 \mid X_2 = x_2 \ \sim\ \mathcal{N}\left(3 + \frac{x_2 - 5}{3},\ 3\right)$.

(β´) Για $x_2 = 8$:

$$\mathbb{E}[X_1 \mid X_2 = 8] = 3 + \frac{8 - 5}{3} = 3 + 1 = 4, \qquad \mathrm{Var}(X_1 \mid X_2 = 8) = 3.$$

(γ´) Έχουμε $\rho_{12} = \dfrac{\sigma_{12}}{\sqrt{\sigma_{11}\sigma_{22}}} = \dfrac{3}{\sqrt{36}} = \dfrac{1}{2}$.

Έλεγχος

Ελέγχουμε την ισότητα $\mathrm{Var}(X_1 \mid X_2 = x_2) = \sigma_{11}(1 - \rho_{12}^2)$:

$$\sigma_{11}(1 - \rho_{12}^2) = 4 \cdot \left(1 - \tfrac{1}{4}\right) = 4 \cdot \tfrac{3}{4} = 3. \checkmark$$

Παρατήρηση

Παρατηρούμε δύο θεμελιώδεις ιδιότητες της διμεταβλητής κανονικής:

- Η δεσμευμένη μέση τιμή $\mathbb{E}[X_1 \mid X_2 = x_2]$ είναι *γραμμική συνάρτηση* του x_2.

Αυτό εξηγεί γιατί η γραμμική παλινδρόμηση είναι βέλτιστη όταν οι μεταβλητές είναι από κοινού κανονικές.

- Η δεσμευμένη διακύμανση *δεν εξαρτάται* από το x_2 (ομοσκεδαστικότητα).

Παράδειγμα 7: Αποδόσεις μετοχών — Δεσμευμένη πρόβλεψη

Οι μηνιαίες αποδόσεις (σε %) δύο μετοχών R_A και R_B ακολουθούν διμεταβλητή κανονική κατανομή με

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_A \\ \mu_B \end{pmatrix} = \begin{pmatrix} 0{,}8 \\ 1{,}2 \end{pmatrix}, \qquad \boldsymbol{\Sigma} = \begin{pmatrix} 4 & 2{,}4 \\ 2{,}4 & 9 \end{pmatrix}.$$

(α΄) Βρείτε τη δεσμευμένη κατανομή της R_B δεδομένου ότι $R_A = r_A$.

(β΄) Τον περασμένο μήνα η μετοχή A είχε απόδοση -2%. Ποια είναι η βέλτιστη πρόβλεψη για την απόδοση της B τον ίδιο μήνα; Ποια η τυπική απόκλιση της πρόβλεψης;

Λύση

(α΄) Εδώ $\sigma_{AA} = 4,\ \sigma_{BB} = 9,\ \sigma_{AB} = 2{,}4$. Άρα:

$$\begin{aligned} \mathbb{E}[R_B \mid R_A = r_A] &= \mu_B + \frac{\sigma_{AB}}{\sigma_{AA}}(r_A - \mu_A) = 1{,}2 + \frac{2{,}4}{4}(r_A - 0{,}8) \\ &= 1{,}2 + 0{,}6\,(r_A - 0{,}8), \\ \mathrm{Var}(R_B \mid R_A = r_A) &= \sigma_{BB} - \frac{\sigma_{AB}^2}{\sigma_{AA}} = 9 - \frac{(2{,}4)^2}{4} = 9 - 1{,}44 = 7{,}56. \end{aligned}$$

Άρα $R_B \mid R_A = r_A \ \sim\ \mathcal{N}\big(1{,}2 + 0{,}6(r_A - 0{,}8),\ 7{,}56\big)$.
(β΄) Για $r_A = -2$:

$$\mathbb{E}[R_B \mid R_A = -2] = 1{,}2 + 0{,}6 \cdot (-2 - 0{,}8) = 1{,}2 - 1{,}68 = -0{,}48\%.$$

Η τυπική απόκλιση είναι $\sqrt{7{,}56} \approx 2{,}749\%$.

Παρατήρηση

Αυτό ακριβώς είναι το μοντέλο της απλής γραμμικής παλινδρόμησης όταν (R_A, R_B) διμεταβλητή κανονική: ο *συντελεστής κλίσης* $0{,}6$ είναι σ_{AB}/σ_{AA}, η *σταθερά* είναι $\mu_B - 0{,}6\mu_A = 0{,}72$, και η *διακύμανση των καταλοίπων* είναι $\sigma_{BB}(1 - \rho^2) = 7{,}56$, όπου εδώ $\rho = 2{,}4/\sqrt{36} = 0{,}4$.

Παράδειγμα 8: Τριμεταβλητή κανονική — Περιθώρια υποδιανύσματος

Έστω $\boldsymbol{X} = (X_1, X_2, X_3)' \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ με

$$\boldsymbol{\mu} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \qquad \boldsymbol{\Sigma} = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 4 \end{pmatrix}.$$

Βρείτε την κατανομή του υποδιανύσματος $\boldsymbol{X}_1 = (X_1, X_2)'$.

Λύση

Διαμερίζουμε το $\boldsymbol{X}$ σε $\boldsymbol{X}_1 = (X_1, X_2)'$ (διάστασης 2) και $\boldsymbol{X}_2 = X_3$ (διάστασης 1). Τότε:

$$\boldsymbol{\mu}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \qquad \boldsymbol{\Sigma}_{11} = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}.$$

Γενικά, οποιοδήποτε υποδιάνυσμα μιας πολυμεταβλητής κανονικής είναι και αυτό πολυμεταβλητή κανονική με μέση τιμή το αντίστοιχο τμήμα του $\boldsymbol{\mu}$ και πίνακα συνδιακυμάνσεων το αντίστοιχο διαγώνιο μπλοκ του $\boldsymbol{\Sigma}$. Άρα:

$$\boldsymbol{X}_1 = (X_1, X_2)' \sim \mathcal{N}\left(\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} \right).$$

Παρατήρηση

Για να το επιβεβαιώσουμε αυστηρά, χρησιμοποιούμε τη ροπογεννήτρια: για $\boldsymbol{t} = (t_1, t_2, 0)'$, η $M_{\boldsymbol{X}}(\boldsymbol{t})$ εξαρτάται μόνο από το $\boldsymbol{\mu}_1$ και το $\boldsymbol{\Sigma}_{11}$. Αυτό αποδεικνύει το αποτέλεσμα.

Παράδειγμα 9: Τριμεταβλητή κανονική — Δεσμευμένη κατανομή

Συνεχίζουμε με το διάνυσμα $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ του Παραδείγματος 8. Διαμερίζουμε $\boldsymbol{X}_1 = (X_1, X_2)'$ και $\boldsymbol{X}_2 = X_3$.

(α΄) Βρείτε τη δεσμευμένη κατανομή του $\boldsymbol{X}_1 \mid X_3 = x_3$.

(β΄) Υπολογίστε $\mathbb{E}[\boldsymbol{X}_1 \mid X_3 = 5]$ και $\mathrm{Cov}(\boldsymbol{X}_1 \mid X_3 = 5)$.

Λύση

(α΄) Από τον γενικό τύπο:

$$\mathbb{E}[\boldsymbol{X}_1 \mid \boldsymbol{X}_2 = \boldsymbol{x}_2] = \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{x}_2 - \boldsymbol{\mu}_2),$$
$$\mathrm{Cov}(\boldsymbol{X}_1 \mid \boldsymbol{X}_2 = \boldsymbol{x}_2) = \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}.$$

Εδώ $\boldsymbol{\Sigma}_{22} = 4$, οπότε $\boldsymbol{\Sigma}_{22}^{-1} = 1/4$, και $\boldsymbol{\Sigma}_{12} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $\boldsymbol{\Sigma}_{21} = \boldsymbol{\Sigma}_{12}' = (1,\ 2)$. Υπολογίζουμε:

$$\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1} = \frac{1}{4}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 0{,}25 \\ 0{,}5 \end{pmatrix}.$$

Άρα:

$$\mathbb{E}[\boldsymbol{X}_1 \mid X_3 = x_3] = \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} 0{,}25 \\ 0{,}5 \end{pmatrix}(x_3 - 3).$$

Για τη συνδιακύμανση:

$$\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21} = \frac{1}{4}\begin{pmatrix} 1 \\ 2 \end{pmatrix}(1\ 2) = \frac{1}{4}\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} 0{,}25 & 0{,}5 \\ 0{,}5 & 1 \end{pmatrix}.$$

Άρα:

$$\mathbf{\Sigma}_{11\cdot 2} = \mathbf{\Sigma}_{11} - \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21} = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} - \begin{pmatrix} 0{,}25 & 0{,}5 \\ 0{,}5 & 1 \end{pmatrix} = \begin{pmatrix} 1{,}75 & 0{,}5 \\ 0{,}5 & 2 \end{pmatrix}.$$

(β´) Για $x_3 = 5$:

$$\mathbb{E}[\boldsymbol{X}_1 \mid X_3 = 5] = \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} 0{,}25 \\ 0{,}5 \end{pmatrix}(5-3) = \begin{pmatrix} 1{,}5 \\ 3 \end{pmatrix},$$

και $\mathrm{Cov}(\boldsymbol{X}_1 \mid X_3 = 5) = \begin{pmatrix} 1{,}75 & 0{,}5 \\ 0{,}5 & 2 \end{pmatrix}$.

Παρατήρηση

Παρατηρούμε ότι η δεσμευμένη συνδιακύμανση $\mathbf{\Sigma}_{11\cdot 2}$ *δεν εξαρτάται* από το x_2. Αυτό είναι ιδιαίτερο χαρακτηριστικό της πολυμεταβλητής κανονικής και οδηγεί στη θεωρία της πολυμεταβλητής παλινδρόμησης.

5. Γραμμικοί μετασχηματισμοί

Παράδειγμα 10: Γραμμικός μετασχηματισμός διανύσματος

Έστω $\boldsymbol{X} = (X_1, X_2)' \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right)$ με $|\rho| < 1$. Ορίζουμε

$$\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{b}, \qquad \boldsymbol{A} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \ \boldsymbol{b} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}.$$

(α´) Υπολογίστε $\mathbb{E}[\boldsymbol{Y}]$ και $\mathrm{Cov}(\boldsymbol{Y})$.

(β´) Ποια είναι η κατανομή της $\boldsymbol{Y}$;

(γ´) Δείξτε ότι οι συνιστώσες Y_1 και Y_2 είναι *ανεξάρτητες*.

Λύση

(α´) Χρησιμοποιώντας τις γενικές ιδιότητες $\mathbb{E}[\boldsymbol{A}\boldsymbol{X}+\boldsymbol{b}] = \boldsymbol{A}\mathbb{E}[\boldsymbol{X}]+\boldsymbol{b}$ και $\mathrm{Cov}(\boldsymbol{A}\boldsymbol{X}+\boldsymbol{b}) = \boldsymbol{A}\mathbf{\Sigma}\boldsymbol{A}'$:

$$\mathbb{E}[\boldsymbol{Y}] = \boldsymbol{A}\cdot\mathbf{0} + \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}.$$

Για τη συνδιακύμανση:

$$\boldsymbol{A}\mathbf{\Sigma} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} = \begin{pmatrix} 1+\rho & \rho+1 \\ 1-\rho & \rho-1 \end{pmatrix},$$

και τελικά:

$$\boldsymbol{A}\mathbf{\Sigma}\boldsymbol{A}' = \begin{pmatrix} 1+\rho & 1+\rho \\ 1-\rho & -(1-\rho) \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 2(1+\rho) & 0 \\ 0 & 2(1-\rho) \end{pmatrix}.$$

(β´) Επειδή ο γραμμικός μετασχηματισμός κανονικού διανύσματος παράγει κανονικό διάνυσμα:

$$\boldsymbol{Y} \sim \mathcal{N}\left(\begin{pmatrix}2\\0\end{pmatrix},\ \begin{pmatrix}2(1+\rho) & 0\\ 0 & 2(1-\rho)\end{pmatrix}\right).$$

(γ΄) Ο πίνακας συνδιακυμάνσεων της $\boldsymbol{Y}$ είναι διαγώνιος, άρα $\mathrm{Cov}(Y_1, Y_2) = 0$. Επειδή (Y_1, Y_2) είναι από κοινού κανονικές (ως γραμμικός μετασχηματισμός από κοινού κανονικών), από την πρόταση του μαθήματος: $\rho_{Y_1Y_2} = 0 \;\Rightarrow\; Y_1 \perp\!\!\!\perp Y_2$.

Παρατήρηση

Ο μετασχηματισμός $(X_1 + X_2,\ X_1 - X_2)$ «διαγωνοποιεί» μια διμεταβλητή κανονική με ίσες διακυμάνσεις. Αυτή η τεχνική αποτελεί ιδιαίτερη περίπτωση της φασματικής αποσύνθεσης: πάντοτε μπορούμε να βρούμε γραμμικό μετασχηματισμό $\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X}$ τέτοιο ώστε τα Y_i να είναι ανεξάρτητα.

Παράδειγμα 11: Άθροισμα συσχετισμένων κανονικών μεταβλητών

Έστω $(X_1, X_2)' \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ με $\mu_1 = \mu_2 = 0$, $\sigma_{11} = \sigma_{22} = 1$ και $\mathrm{Cov}(X_1, X_2) = \rho$. Βρείτε την κατανομή της $S = X_1 + X_2$ και της $D = X_1 - X_2$.

Λύση

Γράφουμε $S = \boldsymbol{a}'\boldsymbol{X}$ με $\boldsymbol{a} = (1, 1)'$ και $D = \boldsymbol{b}'\boldsymbol{X}$ με $\boldsymbol{b} = (1, -1)'$. Από τον γενικό τύπο $\boldsymbol{a}'\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{a}'\boldsymbol{\mu}, \boldsymbol{a}'\boldsymbol{\Sigma}\boldsymbol{a})$:

$$\begin{aligned} \mathbb{E}[S] &= 0, \qquad \mathrm{Var}(S) = \boldsymbol{a}'\boldsymbol{\Sigma}\boldsymbol{a} = 1 + 1 + 2\rho = 2(1+\rho),\\ \mathbb{E}[D] &= 0, \qquad \mathrm{Var}(D) = \boldsymbol{b}'\boldsymbol{\Sigma}\boldsymbol{b} = 1 + 1 - 2\rho = 2(1-\rho). \end{aligned}$$

Άρα $S \sim \mathcal{N}(0,\ 2(1+\rho))$ και $D \sim \mathcal{N}(0,\ 2(1-\rho))$.

Έλεγχος

Ελέγχουμε ακραίες τιμές:

- $\rho = 1$ (τέλεια θετική συσχέτιση): $X_1 = X_2$ σ.β. (με τη δοθείσα περιθώρια), άρα $S = 2X_1 \sim \mathcal{N}(0, 4)$. Πράγματι $2(1 + 1) = 4$. ✓ Επίσης $D = 0$ σ.β., δηλαδή $\mathrm{Var}(D) = 0 = 2(1 - 1)$. ✓
- $\rho = -1$ (τέλεια αρνητική): $X_2 = -X_1$, άρα $S = 0$ ($\mathrm{Var} = 0$) και $D = 2X_1 \sim \mathcal{N}(0, 4)$. ✓
- $\rho = 0$: $\mathrm{Var}(S) = \mathrm{Var}(D) = 2$, που είναι το αναμενόμενο για ανεξάρτητες $\mathcal{N}(0, 1)$. ✓

Παράδειγμα 12: Χαρτοφυλάκιο δύο μετοχών — Ελαχιστοποίηση κινδύνου

Οι ετήσιες αποδόσεις δύο μετοχών X_1, X_2 ακολουθούν διμεταβλητή κανονική με

$$\mathbb{E}[X_1] = \mathbb{E}[X_2] = 0{,}10, \qquad \mathrm{Var}(X_1) = 0{,}04,\ \mathrm{Var}(X_2) = 0{,}09,\ \mathrm{Corr}(X_1, X_2) = 0{,}2.$$

Κατασκευάζουμε χαρτοφυλάκιο $P = wX_1 + (1 - w)X_2$, όπου $w \in [0, 1]$.

(α΄) Βρείτε την κατανομή του P ως συνάρτηση του w.

(β΄) Ποια τιμή του w ελαχιστοποιεί τη διακύμανση του P;

(γ΄) Υπολογίστε την ελάχιστη διακύμανση και τη μέση απόδοση σε αυτό το σημείο.

Λύση

(α΄) Έχουμε $\sigma_{12} = \rho\sqrt{\sigma_{11}\sigma_{22}} = 0{,}2 \cdot \sqrt{0{,}04 \cdot 0{,}09} = 0{,}2 \cdot 0{,}06 = 0{,}012$. Άρα:

$$\mathbb{E}[P] = w(0{,}10) + (1-w)(0{,}10) = 0{,}10,$$
$$\text{Var}(P) = w^2(0{,}04) + (1-w)^2(0{,}09) + 2w(1-w)(0{,}012).$$

Εφόσον το P είναι γραμμικός συνδυασμός από κοινού κανονικών, $P \sim \mathcal{N}(0{,}10,\ \text{Var}(P))$.

(β΄) Αναπτύσσοντας: $\text{Var}(P) = 0{,}04w^2 + 0{,}09(1-w)^2 + 0{,}024w(1-w)$. Παραγωγίζοντας ως προς w και θέτοντας ίσο με μηδέν:

$$\frac{d\text{Var}(P)}{dw} = 0{,}08w - 0{,}18(1-w) + 0{,}024(1-2w) = 0.$$

Υπολογίζουμε: $0{,}08w - 0{,}18 + 0{,}18w + 0{,}024 - 0{,}048w = 0 \;\Rightarrow\; (0{,}08 + 0{,}18 - 0{,}048)w = 0{,}18 - 0{,}024$.

$$0{,}212\,w = 0{,}156 \;\Rightarrow\; w^* = \frac{0{,}156}{0{,}212} \approx 0{,}7358.$$

Έλεγχος

Γενικός τύπος για ελάχιστη διακύμανση: $w^* = \dfrac{\sigma_{22} - \sigma_{12}}{\sigma_{11} + \sigma_{22} - 2\sigma_{12}} = \dfrac{0{,}09 - 0{,}012}{0{,}04 + 0{,}09 - 2(0{,}012)} = \dfrac{0{,}078}{0{,}106} \approx 0{,}7358.$ ✓

(γ΄) Αντικαθιστώντας $w^* \approx 0{,}7358$:

$$\begin{aligned}\text{Var}(P^*) &\approx 0{,}04(0{,}7358)^2 + 0{,}09(0{,}2642)^2 + 0{,}024(0{,}7358)(0{,}2642)\\ &\approx 0{,}04(0{,}5414) + 0{,}09(0{,}0698) + 0{,}024(0{,}1944)\\ &\approx 0{,}02166 + 0{,}00628 + 0{,}00467 \approx 0{,}0326.\end{aligned}$$

Τυπική απόκλιση $\approx \sqrt{0{,}0326} \approx 0{,}1806$ (δηλαδή $18{,}06\%$). Η μέση απόδοση παραμένει $\mathbb{E}[P] = 10\%$.

Παρατήρηση

Επειδή οι δύο μετοχές έχουν την ίδια αναμενόμενη απόδοση, η μεταβολή του w αλλάζει μόνο τον κίνδυνο. Το βέλτιστο χαρτοφυλάκιο δίνει σημαντικά μικρότερη διακύμανση ($0{,}0326$) από τη δεύτερη μετοχή μόνη ($0{,}09$) αλλά όχι από την πρώτη μόνη ($0{,}04$). Αυτό συμβαίνει επειδή εδώ η πρώτη μετοχή έχει ήδη τη μικρότερη διακύμανση και $w^* < 1$ σημαίνει ότι το μίγμα μειώνει τον κίνδυνο περαιτέρω, χάρη στη μη τέλεια συσχέτιση.

6. Συνδυαστικές εφαρμογές

Παράδειγμα 13: Ασυσχέτιστες αλλά εξαρτημένες (χωρίς από κοινού κανονικότητα)

Έστω $X \sim \mathcal{N}(0,1)$ και $Y = X^2$.

(α΄) Υπολογίστε $\mathrm{Cov}(X,Y)$.

(β΄) Δείξτε ότι οι X, Y *δεν* είναι ανεξάρτητες.

(γ΄) Συμπεραίνετε ότι το ζεύγος (X,Y) δεν μπορεί να είναι από κοινού κανονικό.

Λύση

(α΄) Υπολογίζουμε:

$$\mathrm{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \mathbb{E}[X^3] - 0 \cdot 1 = \mathbb{E}[X^3].$$

Επειδή $X \sim \mathcal{N}(0,1)$, όλες οι περιττές ροπές είναι μηδέν: $\mathbb{E}[X^3] = 0$. Άρα $\mathrm{Cov}(X,Y) = 0$.

(β΄) Η Y είναι *συνάρτηση* της X: γνωρίζοντας X, γνωρίζουμε ακριβώς την Y. Δηλαδή οι X, Y είναι πολύ έντονα εξαρτημένες. Τυπικός έλεγχος: λ.χ.

$$P(Y \le 0) = P(X^2 \le 0) = P(X = 0) = 0.$$

Αλλά $P(Y \le 0 \mid X = 0) = P(0 \le 0) = 1 \ne 0$. Άρα δεν ισχύει $P(Y \le 0 \mid X) = P(Y \le 0)$, οπότε οι X, Y δεν είναι ανεξάρτητες.

(γ΄) Αν το (X,Y) ήταν από κοινού κανονικό, τότε από την πρόταση της διάλεξης θα είχαμε: $\mathrm{Cov}(X,Y) = 0 \;\Rightarrow\; X \perp\!\!\!\perp Y$. Αυτό αντιφάσκει με το (β΄). Άρα (X,Y) δεν είναι από κοινού κανονικό, παρότι X είναι κανονική.

Παρατήρηση

Αυτό είναι ένα κλασικό αντιπαράδειγμα που δείχνει πόσο ισχυρή είναι η υπόθεση της από κοινού κανονικότητας. **Χωρίς** αυτήν, η ασυσχετισία είναι πολύ ασθενέστερη από την ανεξαρτησία. Στην πράξη, σε χρηματοοικονομικά και οικονομικά δεδομένα, η υπόθεση αυτή πρέπει να ελέγχεται πριν εξαχθούν συμπεράσματα ανεξαρτησίας.

Παράδειγμα 14: Πιθανότητες σε διμεταβλητή κανονική

Οι μηνιαίες δαπάνες X (σε χιλ. ευρώ) και τα έσοδα Y μιας μικρής επιχείρησης ακολουθούν διμεταβλητή κανονική με

$$\boldsymbol{\mu} = \begin{pmatrix} 20 \\ 25 \end{pmatrix}, \qquad \boldsymbol{\Sigma} = \begin{pmatrix} 16 & 8 \\ 8 & 25 \end{pmatrix}.$$

(α΄) Υπολογίστε $P(X > 22)$.

(β΄) Υπολογίστε $P(Y > 30 \mid X = 25)$.

(γ΄) Υπολογίστε $P(Y - X > 10)$ (καθαρό κέρδος μεγαλύτερο από 10 χιλ. ευρώ).

Λύση

(α΄) Έχουμε $X \sim \mathcal{N}(20, 16)$, δηλαδή τυπική απόκλιση 4. Τυποποιούμε:

$$P(X > 22) = P\big(Z > \tfrac{22-20}{4}\big) = P(Z > 0{,}5) = 1 - \Phi(0{,}5) \approx 1 - 0{,}6915 = 0{,}3085.$$

(β΄) Η δεσμευμένη κατανομή $Y \mid X = x$ έχει:

$$\mathbb{E}[Y \mid X = x] = 25 + \frac{8}{16}(x - 20) = 25 + 0{,}5(x - 20),$$
$$\mathrm{Var}(Y \mid X = x) = 25 - \frac{8^2}{16} = 25 - 4 = 21.$$

Για $x = 25$: $\mathbb{E}[Y \mid X = 25] = 25 + 0{,}5 \cdot 5 = 27{,}5$ και $\mathrm{Var}(Y \mid X = 25) = 21$, τυπική απόκλιση $\sqrt{21} \approx 4{,}583$. Άρα:

$$P(Y > 30 \mid X = 25) = P\Big(Z > \tfrac{30-27{,}5}{4{,}583}\Big) = P(Z > 0{,}545) \approx 1 - 0{,}7072 = 0{,}2928.$$

(γ΄) Η $W = Y - X$ είναι γραμμικός συνδυασμός: $W = \boldsymbol{a}'\boldsymbol{X}$ με $\boldsymbol{a} = (-1, 1)'$.

$$\mathbb{E}[W] = -20 + 25 = 5,$$
$$\mathrm{Var}(W) = \sigma_{11} + \sigma_{22} - 2\sigma_{12} = 16 + 25 - 16 = 25.$$

Άρα $W \sim \mathcal{N}(5, 25)$, τυπική απόκλιση 5. Τότε:

$$P(W > 10) = P\big(Z > \tfrac{10-5}{5}\big) = P(Z > 1) \approx 1 - 0{,}8413 = 0{,}1587.$$

Παρατήρηση

Τα ερωτήματα (α΄)–(γ΄) δείχνουν τρία διαφορετικά εργαλεία του μαθήματος: *περιθώρια κατανομή* για το (α΄), *δεσμευμένη κατανομή* για το (β΄), και *γραμμικός μετασχηματισμός* για το (γ΄). Παρατηρήστε ότι στο (γ΄) η θετική συσχέτιση ($\rho > 0$) μειώνει τη διακύμανση της διαφοράς — τα έσοδα και οι δαπάνες κινούνται ενν. προς την ίδια κατεύθυνση, οπότε το κέρδος είναι λιγότερο μεταβλητό.

Παράδειγμα 15: Ολοκληρωμένη εφαρμογή — Δύο ασφαλιστικοί κίνδυνοι

Η ετήσια συνολική ζημία (σε χιλ. ευρώ) από δύο ασφαλιστικά χαρτοφυλάκια μιας εταιρείας, X_1 και X_2, ακολουθεί διμεταβλητή κανονική κατανομή με

$$\boldsymbol{\mu} = \begin{pmatrix} 100 \\ 80 \end{pmatrix}, \qquad \boldsymbol{\Sigma} = \begin{pmatrix} 400 & 120 \\ 120 & 225 \end{pmatrix}.$$

(α΄) Υπολογίστε τον συντελεστή συσχέτισης ρ.

(β΄) Βρείτε την περιθώρια κατανομή της X_1 και την πιθανότητα $P(X_1 > 130)$.

(γ΄) Βρείτε τη δεσμευμένη κατανομή της X_2 δεδομένου $X_1 = 130$.

(δ΄) Υπολογίστε την αναμενόμενη τιμή $\mathbb{E}[X_2 \mid T = 220]$ όταν $T = X_1 + X_2$.

Λύση

(α΄) $\rho = \frac{\sigma_{12}}{\sqrt{\sigma_{11}\sigma_{22}}} = \frac{120}{\sqrt{400 \cdot 225}} = \frac{120}{300} = 0{,}4.$
(β΄) Από τη θεωρία της διάλεξης, η περιθώρια $X_1 \sim \mathcal{N}(100, 400)$, τυπική απόκλιση 20. Άρα:

$$P(X_1 > 130) = P\big(Z > \tfrac{130-100}{20}\big) = P(Z > 1{,}5) \approx 1 - 0{,}9332 = 0{,}0668.$$

(γ΄) Δεσμευμένη κατανομή:

$$\mathbb{E}[X_2 \mid X_1 = x_1] = 80 + \frac{120}{400}(x_1 - 100) = 80 + 0{,}3(x_1 - 100),$$
$$\text{Var}(X_2 \mid X_1 = x_1) = 225 - \frac{120^2}{400} = 225 - 36 = 189.$$

Για $x_1 = 130$: $\mathbb{E}[X_2 \mid X_1 = 130] = 80 + 0{,}3 \cdot 30 = 89$, Var $= 189$. Άρα $X_2 \mid X_1 = 130 \sim \mathcal{N}(89, 189)$.
(δ΄) Χρειαζόμαστε τη δεσμευμένη μέση τιμή $\mathbb{E}[X_2 \mid T = t]$. Ορίζουμε το διάνυσμα $\boldsymbol{Z} = (X_2, T)'$. Αυτό είναι γραμμικός μετασχηματισμός του $\boldsymbol{X}$, άρα από κοινού κανονικό. Υπολογίζουμε:

$$\mathbb{E}[\boldsymbol{Z}] = \begin{pmatrix} 80 \\ 180 \end{pmatrix},$$

$$\text{Var}(X_2) = 225,$$
$$\text{Var}(T) = 865 \quad \text{(ήδη από (δ΄))},$$
$$\text{Cov}(X_2, T) = \text{Cov}(X_2, X_1 + X_2) = \text{Cov}(X_2, X_1) + \text{Var}(X_2) = 120 + 225 = 345.$$

Άρα:

$$\mathbb{E}[X_2 \mid T = t] = 80 + \frac{\text{Cov}(X_2, T)}{\text{Var}(T)}(t - 180) = 80 + \frac{345}{865}(t - 180).$$

Για $t = 220$: $\mathbb{E}[X_2 \mid T = 220] = 80 + \frac{345}{865}(40) = 80 + \frac{13\,800}{865} \approx 80 + 15{,}95 \approx 95{,}95$ χιλ. €.

Έλεγχος

Άμεσος έλεγχος του (ε΄): Η $\mathbb{E}[X_1 \mid T = 220]$ θα πρέπει να ικανοποιεί $\mathbb{E}[X_1 \mid T = 220] + \mathbb{E}[X_2 \mid T = 220] = 220$. Υπολογίζουμε:

$$\text{Cov}(X_1, T) = \text{Var}(X_1) + \text{Cov}(X_1, X_2) = 400 + 120 = 520,$$
$$\mathbb{E}[X_1 \mid T = 220] = 100 + \frac{520}{865}(40) \approx 100 + 24{,}05 \approx 124{,}05.$$

Πράγματι $124{,}05 + 95{,}95 = 220$. ✓

Παρατήρηση

Στο (ε΄) χρησιμοποιήσαμε μια γενική αρχή: αν $\boldsymbol{X}$ είναι πολυμεταβλητό κανονικό, τότε *κάθε* ζεύγος γραμμικών συνδυασμών του είναι διμεταβλητό κανονικό, και μπορούμε να χρησιμοποιήσουμε τους τύπους δεσμευμένης μέσης τιμής απευθείας. Αυτή η ιδιότητα κλειστότητας σε γραμμικούς μετασχηματισμούς είναι ο κύριος λόγος που η πολυμεταβλητή κανονική έχει τόσο ευρεία εφαρμογή στη στατιστική και την οικονομετρία.

Σύνοψη τύπων της Διάλεξης 6

Πολυμεταβλητή κανονική: $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ στον $\mathbb{R}^n$:

$$f(\boldsymbol{x}) = \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\bigl(-\tfrac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\bigr)\,.$$

Ροπογεννήτρια: $M_{\boldsymbol{X}}(\boldsymbol{t}) = \exp\bigl(\boldsymbol{t}'\boldsymbol{\mu} + \tfrac{1}{2}\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t}\bigr)$.
Γραμμικός μετασχηματισμός: Αν $\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{b}$, τότε $\boldsymbol{Y} \sim \mathcal{N}(\boldsymbol{A}\boldsymbol{\mu} + \boldsymbol{b},\ \boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}')$.
Περιθώριες: Κάθε υποδιάνυσμα του $\boldsymbol{X}$ είναι και αυτό κανονικό, με τη μέση τιμή και συνδιακύμανση που αντιστοιχούν στα επιλεγμένα στοιχεία.
Δεσμευμένες (διαμερίζοντας $\boldsymbol{X} = (\boldsymbol{X}_1, \boldsymbol{X}_2)'$):

$$\mathbb{E}[\boldsymbol{X}_1 \mid \boldsymbol{X}_2 = \boldsymbol{x}_2] = \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{x}_2 - \boldsymbol{\mu}_2),$$
$$\mathrm{Cov}(\boldsymbol{X}_1 \mid \boldsymbol{X}_2 = \boldsymbol{x}_2) = \boldsymbol{\Sigma}_{11\cdot 2} = \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}.$$

Διμεταβλητή ειδική περίπτωση:

$$\mathbb{E}[X_1 \mid X_2 = x_2] = \mu_1 + \frac{\sigma_{12}}{\sigma_{22}}(x_2 - \mu_2),$$
$$\mathrm{Var}(X_1 \mid X_2 = x_2) = \sigma_{11} - \frac{\sigma_{12}^2}{\sigma_{22}} = \sigma_{11}(1-\rho^2).$$

Πρόταση: Υπό από κοινού κανονικότητα, $\rho_{ij} = 0 \iff X_i \perp\!\!\!\perp X_j$.