ΠΟΛΥΔΙΑΣΤΑΤΗ ΚΑΝΟΝΙΚΗ ΚΑΤΑΝΟΜΗ

Έστω $\mathbf{X} : \Omega \to \mathbb{R}^d$ τυχαίο διάνυσμα με

$$\mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_d \end{pmatrix}.$$

Έχουμε δει ότι η γνώση της κατανομής καθεμιάς από τις $X_1, X_2, \ldots, X_d$ δεν αρκεί για να προσδιορίσουμε την κατανομή του $\mathbf{X}$, αφού δεν περιέχει πληροφορία για το πώς αυτές σχετίζονται μεταξύ τους. Είναι δυνατόν να περιγράψουμε την κατανομή του $\mathbf{X}$ χρησιμοποιώντας μια μεγαλύτερη συλλογή από μονοδιάστατες κατανομές; Το ακόλουθο λήμμα απαντά θετικά, αν η συλλογή περιλαμβάνει την κατανομή των προβολών του $\mathbf{X}$ σε *οποιαδήποτε* κατεύθυνση του $\mathbb{R}^d$.

Λήμμα 1 Η κατανομή του τυχαίου διανύσματος $\mathbf{X} : \Omega \to \mathbb{R}^d$ προσδιορίζεται πλήρως από την κατανομή των πραγματικών τυχαίων μεταβλητών $\{\ell^T \mathbf{X}\}_{\ell \in \mathbb{R}^d}$, όπου

$$\ell^T \mathbf{X} = \sum_{j=1}^{d} \ell_j X_j.$$

Απόδειξη: Η χαρακτηριστική συνάρτηση του διανύσματος $\mathbf{X}$ είναι η $\phi_{\mathbf{X}} : \mathbb{R}^d \to \mathbb{C}$ με $\phi_{\mathbf{X}}(\ell) = \mathbb{E}\big[e^{i\ell^T \mathbf{X}}\big]$ για $\ell \in \mathbb{R}^d$. Αν η κατανομή της $\ell^T \mathbf{X}$ είναι γνωστή για κάθε $\ell \in \mathbb{R}^d$, είναι γνωστή και η $\phi_{\mathbf{X}}$. Επομένως η κατανομή της $\mathbf{X}$ είναι πλήρως προσδιορισμένη. $\square$

1 Ορισμοί

Ορισμός 1 Θα λέμε ότι το τυχαίο διάνυσμα $\mathbf{X} : \Omega \to \mathbb{R}^d$ ακολουθεί (d-διάστατη) κανονική κατανομή, αν για κάθε $\ell \in \mathbb{R}^d$ η πραγματική τυχαία μεταβλητή $\ell^T \mathbf{X}$ ακολουθεί κανονική κατανομή.

Το ακόλουθο λήμμα εξασφαλίζει ότι αν οι συνιστώσες του $\mathbf{X}$ είναι ανεξάρτητες τυπικές κανονικές τυχαίες μεταβλητές, τότε το $\mathbf{X}$ ακολουθεί d-διάστατη κανονική κατανομή σύμφωνα με τον ορισμό που δώσαμε.

Λήμμα 2 Αν οι $X_1, X_2, \ldots, X_d$ είναι ανεξάρτητες τυπικές κανονικές τυχαίες μεταβλητές, τότε για κάθε $\ell \in \mathbb{R}^d$

$$\ell^T \mathbf{X} \sim \mathcal{N}\big(0, \|\ell\|^2\big).$$

Απόδειξη: Αρκεί να δείξουμε ότι η χαρακτηριστική συνάρτηση της τ.μ. $\ell^T \mathbf{X}$ είναι η $\psi(t) = \exp\left(-\frac{t^2\|\ell\|^2}{2}\right)$.

$$\psi(t) = \mathbb{E}\big[e^{it\ell^T \mathbf{X}}\big] = \mathbb{E}\big[\prod_{j=1}^{d} e^{it\ell_j X_j}\big] = \prod_{j=1}^{d} \mathbb{E}\big[e^{it\ell_j X_j}\big] = \prod_{j=1}^{d} \exp\left(-\frac{(t\ell_j)^2}{2}\right) = \exp\left(-\frac{t^2\|\ell\|^2}{2}\right)$$

Στην τρίτη ισότητα χρησιμοποιήσαμε ότι οι $\{X_j\}_{1 \le j \le d}$ είναι ανεξάρτητες και στην τέταρτη ότι η χαρακτηριστική συνάρτηση της τυπικής κανονικής κατανομής είναι η $\phi(t) = \exp\big(-t^2/2\big)$. $\square$

Η ακόλουθη πρόταση περιγράφει τις παραμέτρους μιας κανονικής κατανομής στον $\mathbb{R}^d$.

Πρόταση 1 Η κατανομή του d-διάστατου τυχαίου διανύσματος $\mathbf{X}$ προσδιορίζεται από το διάνυσμα μέσων τιμών

$$\mathbf{m} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_d \end{pmatrix} \qquad \text{με } \mu_i = \mathbb{E}\big[X_i\big],\ i = 1, 2, \ldots, d. \tag{1}$$

και τον πίνακα συνδιακυμάνσεων

$$\mathbf{\Sigma} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \cdots & \Sigma_{1d} \\ \Sigma_{21} & \Sigma_{22} & \cdots & \Sigma_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{d1} & \Sigma_{d2} & \cdots & \Sigma_{dd} \end{pmatrix} \qquad \text{με } \Sigma_{ij} = \mathrm{Cov}\big(X_i, X_j\big), \qquad i,j = 1,2,\ldots,d. \tag{2}$$

Απόδειξη: Η κανονική κατανομή σε μία διάσταση προσδιορίζεται από δύο παραμέτρους, τη μέση τιμή και τη διασπορά της. Από το Λήμμα 1 συμπεραίνουμε ότι η κατανομή μιας d-διάστατης κανονικής τυχαίας μεταβλητής $\mathbf{X}$ προσδιορίζεται πλήρως, αν γνωρίζουμε για κάθε $\ell \in \mathbb{R}^d$ τη μέση τιμή και τη διασπορά της τυχαίας μεταβλητής $\ell^T\mathbf{X}$. Έχουμε όμως ότι

$$\mathbb{E}\big[\ell^T\mathbf{X}\big] = \mathbb{E}\big[\sum_{i=1}^{d} \ell_i X_i\big] = \sum_{i=1}^{d} \ell_i \mathbb{E}\big[X_i\big] = \ell^T\mathbf{m}.$$

Επίσης,

$$V\big[\ell^T\mathbf{X}\big] = \mathrm{Cov}\big(\ell^T\mathbf{X}, \ell^T\mathbf{X}\big) = \mathrm{Cov}\big(\sum_{i=1}^{d} \ell_i X_i, \sum_{j=1}^{d} \ell_j X_j\big) = \sum_{i,j=1}^{d} \ell_i \ell_j \mathrm{Cov}\big(X_i, X_j\big) = \ell^T\mathbf{\Sigma}\ell. \tag{3}$$

Επόμενως η κατανομή του $\mathbf{X}$ προσδιορίζεται από τα $\mathbf{m}$ και $\mathbf{\Sigma}$. □

Ορισμός 2 Θα συμβολίζουμε με $\mathcal{N}(\mathbf{m}, \mathbf{\Sigma})$ την d-διάστατη κανονική κατανομή που έχει παραμέτρους $\mathbf{m} \in \mathbb{R}^d$ και $\mathbf{\Sigma} \in \Pi_{d\times d}$ και θα γράφουμε $\mathbf{X} \sim \mathcal{N}(\mathbf{m}, \mathbf{\Sigma})$. Επομένως

$$\mathbf{X} \sim \mathcal{N}(\mathbf{m}, \mathbf{\Sigma}) \Leftrightarrow \ell^T\mathbf{X} \sim \mathcal{N}(\ell^T\mathbf{m}, \ell^T\mathbf{\Sigma}\ell), \qquad \text{για κάθε } \ell \in \mathbb{R}^d. \tag{4}$$

Παράδειγμα 1 Στην περίπτωση που εξετάσαμε στο Λήμμα 2, όπου οι $X_1, \ldots, X_d$ είναι ανεξάρτητες, τυπικές κανονικές τυχαίες μεταβλητές, έχουμε ότι $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$.

2 Γραμμικοί μετασχηματισμοί - Περιθώριες κατανομές

Θεώρημα 1 Κάθε γραμμικός μετασχηματισμός μιας κανονικής τυχαίας μεταβλητής ακολουθεί κανονική κατανομή. Συγκεκριμένα, αν $\mathbf{X} \sim \mathcal{N}(\mathbf{m}, \mathbf{\Sigma})$ με $\mathbf{m} \in \mathbb{R}^d$, $\mathbf{\Sigma} \in \Pi_{d\times d}$ και $\mathbf{A} \in \Pi_{n\times d}$, $\mathbf{b} \in \mathbb{R}^n$, τότε για το τυχαίο διάνυσμα $\mathbf{Y} = \mathbf{AX} + \mathbf{b}$ στον $\mathbb{R}^n$ έχουμε

$$\mathbf{Y} \sim \mathcal{N}(\mathbf{Am} + \mathbf{b}, \mathbf{A\Sigma A}^T).$$

Απόδειξη: Έστω $\ell \in \mathbb{R}^n$. Έχουμε ότι

$$\ell^T\mathbf{Y} = \ell^T(\mathbf{AX} + \mathbf{b}) = (\ell^T\mathbf{A})\mathbf{X} + \ell^T\mathbf{b} = (\mathbf{A}^T\ell)^T\mathbf{X} + \ell^T\mathbf{b}. \tag{5}$$

Εφόσον $\mathbf{A}^T\ell \in \mathbb{R}^d$ και $\mathbf{X} \sim \mathcal{N}(\mathbf{m}, \mathbf{\Sigma})$ έχουμε ότι

$$(\mathbf{A}^T\ell)^T\mathbf{X} \sim \mathcal{N}\big((\mathbf{A}^T\ell)^T\mathbf{m}, (\mathbf{A}^T\ell)^T\mathbf{\Sigma}(\mathbf{A}^T\ell)\big) = \mathcal{N}\big(\ell^T(\mathbf{Am}), \ell^T(\mathbf{A\Sigma A}^T)\ell\big)$$

Επομένως, η (5) δίνει ότι

$$\ell^T\mathbf{Y} \sim \mathcal{N}\big(\ell^T(\mathbf{Am} + \mathbf{b}), \ell^T(\mathbf{A\Sigma A}^T)\ell\big) \qquad \text{για κάθε } \ell \in \mathbb{R}^n.$$

Αυτό όμως είναι από τον ορισμό (2) ο ισχυρισμός του Θεωρήματος. □

Πόρισμα 1 Αν το $\mathbf{X}$ είναι τυχαίο διάνυσμα στον $\mathbb{R}^d$ που ακολουθεί κανονική κατανομή και $\mathbf{Y} = P\mathbf{X}$ είναι η προβολή του $\mathbf{X}$ σε κάποιον υπόχωρο του $\mathbb{R}^d$ διάστασης n, τότε το $\mathbf{Y}$ ακολουθεί κανονική κατανομή σε n διαστάσεις. Ειδικότερα, η περιθώρια κατανομή μιας πολυδιάστατης κανονικής κατανομής είναι επίσης κανονική.

Παράδειγμα 2 Έστω $\mathbf{X}$ τυχαίο διάνυσμα στον $\mathbb{R}^3$ που ακολουθεί κανονική κατανομή $\mathcal{N}(\mathbf{m}, \mathbf{\Sigma})$ με

$$\mathbf{m} = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix} \qquad \text{και} \qquad \mathbf{\Sigma} = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 1 & -1 \\ 0 & -1 & 2 \end{pmatrix}$$

Ας υποθέσουμε ότι θέλουμε να βρούμε την (περιθώρια) κατανομή του ζεύγους (X_1, X_3). Από το προηγούμενο πόρισμα, το ζεύγος ακολουθεί κανονική κατανομή. Οι παράμετροι αυτής της κατανομής μπορούν να διαβαστούν από τις παραμέτρους της $\mathbf{X}$. Πράγματι,

$$\mathbb{E}\big[X_1\big] = \mu_1 = 2 \qquad \text{και} \qquad \mathbb{E}\big[X_3\big] = \mu_3 = -1,$$

ενώ

$$\mathrm{Cov}\big(X_1, X_1\big) = \Sigma_{11} = 2, \qquad \mathrm{Cov}\big(X_1, X_3\big) = \mathrm{Cov}\big(X_3, X_1\big) = \Sigma_{13} = \Sigma_{31} = 0, \qquad \mathrm{Cov}\big(X_3, X_3\big) = \Sigma_{33} = 2.$$

Επομένως,

$$\begin{pmatrix} X_1 \\ X_3 \end{pmatrix} \sim \mathcal{N}(\mathbf{m}', \mathbf{\Sigma}'), \qquad \text{όπου } \mathbf{m}' = \begin{pmatrix} 2 \\ -1 \end{pmatrix} \qquad \text{και} \qquad \mathbf{\Sigma}' = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Παρατηρήστε ότι οι παράμετροι της περιθώριας κατανομής προκύπτουν από εκείνες της πλήρους κατανομής, αν διαγράψουμε τις στήλες και τις γραμμές που αντιστοιχούν στις μεταβλητές που δεν μας ενδιαφέρουν (εδώ στη X_2). $\square$

Πόρισμα 2 Σε κάθε διάνυσμα $\mathbf{m} \in \mathbb{R}^d$ και σε κάθε συμμετρικό θετικά ορισμένο πίνακα $\mathbf{\Sigma}$ αντιστοιχεί μία d-διάστατη κανονική κατανομή.

Απόδειξη: Από την (2) βλέπουμε ότι ο πίνακας συνδιακυμάνσεων $\mathbf{\Sigma}$ μιας d-διάστατης κανονικής κατανομής είναι συμμετρικός. Από την (3) βλέπουμε ότι ο $\mathbf{\Sigma}$ είναι θετικά ορισμένος (δηλ. $\ell^T \mathbf{\Sigma} \ell \geq 0$, για κάθε $\ell \in \mathbb{R}^d$) και μάλιστα $\mathbf{\Sigma}\ell = 0$ αν και μόνο αν το τυχαίο διάνυσμα $\mathbf{X}$ παίρνει τιμές στο υπερεπίπεδο $\{x \in \mathbb{R}^d : \ell^T x = c\}$ με πιθανότητα 1, για κάποιο $c \in \mathbb{R}$.

Αντίστροφα, κάθε συμμετρικός θετικά ορισμένος πίνακας $\mathbf{\Sigma}$ μπορεί να γραφεί ως $\mathbf{\Sigma} = \mathbf{L}\mathbf{L}^T$ (π.χ. ανάλυση Cholesky). Μάλιστα, αν $\mathrm{rank}(\mathbf{\Sigma}) = k \leq d$ τότε μπορούμε να επιλέξουμε τον $\mathbf{L}$ ώστε να είναι ένας $d \times k$ πίνακας. Ας θεωρήσουμε τώρα το τυχαίο διάνυσμα $\mathbf{Z}$ στον $\mathbb{R}^k$ με $Z_1, \ldots, Z_k$ ανεξάρτητες τυπικές κανονικές τυχαίες μεταβλητές και ας ορίσουμε $\mathbf{X} = \mathbf{L}\mathbf{Z} + \mathbf{m}$. Από το Λήμμα 2 και το Θεώρημα 1 έχουμε ότι $\mathbf{X} \sim \mathcal{N}(\mathbf{m}, \mathbf{\Sigma})$. $\square$

Έχουμε δει ότι αν μια συλλογή από τυχαίες μεταβλητές $\{X_i\}_{1 \leq i \leq d}$ είναι ανεξάρτητες, τότε είναι και ασυσχέτιστες ενώ το αντίστροφο δεν είναι εν γένει σωστό. Η ανεξαρτησία είναι επομένως πιο ισχυρή συνθήκη από τη μηδενική συσχέτιση. Με τη βοήθεια του Λήμματος 1 θα δείξουμε ότι στο πλαίσιο των από κοινού κανονικών τυχαίων μεταβλητών οι δύο έννοιες ταυτίζονται.

Θεώρημα 2 Αν $\mathbf{X}$ είναι ένα τυχαίο διάνυσμα στον $\mathbb{R}^d$ με $\mathbf{X} \sim \mathcal{N}(\mathbf{m}, \mathbf{\Sigma})$, τότε οι $\{X_i\}_{1 \leq i \leq d}$ είναι ανεξάρτητες αν και μόνο αν είναι ασυσχέτιστες, δηλαδή $\mathrm{Cov}\big(X_i, X_j\big) = 0$ για $i \neq j$.

Απόδειξη: Αρκεί να αποδείξουμε ότι αν $\mathrm{Cov}\big(X_i, X_j\big) = 0$ για $i \neq j$, τότε οι $\{X_i\}_{1 \leq i \leq d}$ είναι ανεξάρτητες. Πράγματι, σ' αυτήν την περίπτωση από την (2) έχουμε ότι ο πίνακας συνδιακυμάνσεων $\mathbf{\Sigma}$ είναι διαγώνιος,

$$\mathbf{\Sigma} = \mathrm{diag}\big[\Sigma_{11}, \ldots, \Sigma_{dd}\big] = \mathrm{diag}\big[V(X_1), \ldots, V(X_d)\big]$$

Χωρίς βλάβη υποθέτουμε ότι $\Sigma_{ii} = V(X_i) > 0$ για κάθε $i = 1, 2, \ldots, d$. Σε διαφορετική περίπτωση η X_i θα ήταν σταθερή με πιθανότητα 1, επομένως ανεξάρτητη από οποιεσδήποτε άλλες τυχαίες μεταβλητές.

Ορίζουμε τώρα τον γραμμικό μετασχηματισμό

$$Y_i = \frac{X_i - \mu_i}{\sqrt{\Sigma_{ii}}}, \qquad i = 1, 2, \ldots, d.$$

Από το Θεώρημα 1 για το τυχαίο διάνυσμα $\mathbf{Y}$ που έχει συντεταγμένες τις Y_i έχουμε $\mathbf{Y} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. Από το Λήμμα 1 έχουμε ότι οι $\{Y_i\}_{1 \leq i \leq d}$ είναι ανεξάρτητες τυπικές κανονικές τυχαίες μεταβλητές. Επομένως και οι $\{X_i\}_{1 \leq i \leq d}$ θα είναι ανεξάρτητες αφού $X_i = \sqrt{\Sigma_{ii}} Y_i + \mu_i$ για $i = 1, \ldots, d$. $\square$

Παράδειγμα 3 Αν οι X, Y είναι ανεξάρτητες τυπικές κανονικές τυχαίες μεταβλητές, θα δείξουμε ότι οι U, W με

$$U = \frac{X+Y}{\sqrt{2}} \qquad \text{και} \qquad W = \frac{X-Y}{\sqrt{2}}$$

είναι επίσης ανεξάρτητες τυπικές κανονικές τυχαίες μεταβλητές.

Εφόσον τον ζεύγος (X, Y) ακολουθεί κανονική κατανομή σε 2 διαστάσεις (Λήμμα 2) και

$$\begin{pmatrix} U \\ W \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}$$

συμπεραίνουμε αμέσως ότι το ζεύγος (U, V) ακολουθεί κανονική κατανομή σε δύο διαστάσεις. Μπορούμε να υπολογίσουμε τις παραμέτρους της κατανομής των U, V είτε από το Θεώρημα 1 είτε παρατηρώντας ότι

$$\mathbb{E}[U] = \frac{1}{\sqrt{2}}\big(\mathbb{E}[X] + \mathbb{E}[Y]\big) = 0 \qquad \text{και ομοίως} \qquad \mathbb{E}[W] = 0,$$

ενώ

$$V(U) = V(W) = \frac{1}{2}\big(V(X) + V(Y)\big) = 1 \qquad \text{και} \qquad \mathrm{Cov}\big(U, W\big) = \frac{1}{2}\mathrm{Cov}\big(X+Y, X-Y\big) = 0.$$

Παράδειγμα 4 Αν οι $X_1, \dots, X_n$ είναι ανεξάρτητες, ισόνομες τυχαίες μεταβλητές με $X_i \sim \mathcal{N}(\mu, \sigma^2)$ και

$$\bar{X} = \frac{X_1 + \cdots + X_n}{n},$$

τότε οι $X_1 - \bar{X}$ και $\bar{X}$ είναι ανεξάρτητες.

Από το Θεώρημα 1 το ζεύγος $(X_1 - \bar{X}, \bar{X})$ ακολουθεί κανονική κατανομή σε δύο διαστάσεις αφού

$$\begin{pmatrix} X_1 - \bar{X} \\ \bar{X} \end{pmatrix} = \begin{pmatrix} 1 - \frac{1}{n} & -\frac{1}{n} & \cdots & -\frac{1}{n} \\ -\frac{1}{n} & -\frac{1}{n} & \cdots & -\frac{1}{n} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}.$$

Από το Θεώρημα 2 αρκεί να ελέγξουμε ότι $\mathrm{Cov}\big(X_1 - \bar{X}, \bar{X}\big) = 0$. Πράγματι,

$$\mathrm{Cov}\big(X_1 - \bar{X}, \bar{X}\big) = \mathrm{Cov}\big(X_1, \bar{X}\big) - \mathrm{Cov}\big(\bar{X}, \bar{X}\big) = \frac{1}{n}\sum_{j=1}^{n} \mathrm{Cov}\big(X_1, X_j\big) - \frac{1}{n^2}\sum_{i,j=1}^{n} \mathrm{Cov}\big(X_i, X_j\big) = \frac{1}{n} - \frac{1}{n^2}n = 0.$$

Φυσικά δεν υπάρχει κάτι ιδιαίτερο για τη μεταβλητή X_1. Θα μπορούσαμε με τον ίδιο τρόπο να δείξουμε ότι οι $X_i - \bar{X}$ και $\bar{X}$ είναι ανεξάρτητες. $\square$

Παράδειγμα 5 Αν το ζεύγος (X, Y) ακολουθεί κανονική κατανομή σε δύο διαστάσεις τότε καθεμιά από τις τυχαίες μεταβλητές X και Y ακολουθεί κανονική κατανομή. Αυτό είναι άμεση συνέπεια του ορισμού, παίρνοντας ως ℓ το μοναδιαίο διάνυσμα σε κάθε κύρια κατεύθυνση. Ισχύει άραγε το αντίστροφο εν γένει; Είναι δηλαδή σωστό ότι αν $X \sim \mathcal{N}(\mu_1, \sigma_1^2)$ και $Y \sim \mathcal{N}(\mu_2, \sigma_2^2)$ τότε το ζεύγος (X, Y) ακολουθεί κανονική κατανομή σε δύο διαστάσεις;

Η απάντηση είναι αρνητική. Θεωρήστε για παράδειγμα $X \sim \mathcal{N}(0, 1)$, μια τυχαία μεταβλητή Z ανεξάρτητη από τη X, η οποία παίρνει τις τιμές $+1$ ή -1 με πιθανότητα $\frac{1}{2}$ την καθεμία και ορίστε

$$Y = XZ.$$

Η Y ακολουθεί επίσης κανονική κατανομή. Πράγματι, αν Φ είναι η συνάρτηση κατανομής της X,

$$\begin{aligned} F_Y(t) &= \mathbb{P}\big[Y \le t\big] = \mathbb{P}\big[Y \le t,\ Z = +1\big] + \mathbb{P}\big[Y \le t,\ Z = -1\big] \\ &= \mathbb{P}\big[X \le t,\ Z = +1\big] + \mathbb{P}\big[X \ge -t,\ Z = -1\big] \\ &= \frac{1}{2}\mathbb{P}\big[X \le t\big] + \frac{1}{2}\mathbb{P}\big[X \ge -t\big] = \frac{\Phi(t) + \big(1 - \Phi(-t)\big)}{2} \\ &= \Phi(t). \end{aligned}$$

Επομένως $Y \sim \mathcal{N}(0,1)$.

Το ζεύγος (X,Y) δεν ακολουθεί κανονική κατανομή σε δύο διαστάσεις. Αν αυτό συνέβαινε, τότε από τον ορισμό η τυχαία μεταβλητή $X+Y$ θα ακολουθούσε κανονική κατανομή. Έχουμε όμως

$$\mathbb{P}\big[X+Y=0\big] = \mathbb{P}\big[X \leq 0,\ Z=+1\big] + \mathbb{P}\big[X \geq 0,\ Z=-1\big] = \frac{1}{2},$$

επομένως η κατανομή της $X+Y$ δεν είναι καν συνεχής. $\square$

3 Δεσμευμένες κατανομές

Μια πολύ χρήσιμη συνέπεια του Θεωρήματος 2 είναι ότι, αν το ζεύγος (X,Y) ακολουθεί κανονική κατανομή, τότε μπορούμε να αναλύσουμε την Y σε δύο συνιστώσες, μία που είναι πολλαπλάσιο της X και μία που είναι ανεξάρτητη από την X.

Λήμμα 3 Έστω ζεύγος (X,Y) που ακολουθεί κανονική κατανομή σε 2 διαστάσεις, με $V(X)>0$. Τότε

$$Y = \lambda X + Z, \qquad \text{όπου } \lambda = \frac{\mathrm{Cov}\big(X,Y\big)}{V(X)}$$

και η τυχαία μεταβλητή Z είναι ανεξάρτητη της X και ακολουθεί κατανομή $\mathcal{N}\big(\mathbb{E}\big[Y\big]-\lambda\mathbb{E}\big[X\big], V(Y)-\lambda^2 V(X)\big)$.

Απόδειξη: Ορίζουμε

$$Z = Y - \frac{\mathrm{Cov}\big(X,Y\big)}{V(X)}X.$$

Το ζεύγος (X,Z) ακολουθεί κανονική κατανομή αφού το ζεύγος (X,Y) ακολουθεί κανονική κατανομή και

$$\begin{pmatrix} X \\ Z \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\lambda & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}.$$

Επιπλέον,

$$\mathrm{Cov}\big(X,Z\big) = \mathrm{Cov}\big(X, Y-\lambda X\big) = \mathrm{Cov}\big(X,Y\big) - \lambda \mathrm{Cov}\big(X,X\big) = 0.$$

Από το Θεώρημα 2 οι X και Z είναι ανεξάρτητες. Από το Πόρισμα 1 η Z ακολουθεί κανονική κατανομή. Για να βρούμε τις παραμέτρους της κατανομής της Z παρατηρούμε ότι

$$\mathbb{E}\big[Y\big] = \lambda\mathbb{E}\big[X\big] + \mathbb{E}\big[Z\big] \Rightarrow \mathbb{E}\big[Z\big] = \mathbb{E}\big[Y\big] - \lambda\mathbb{E}\big[X\big],$$

ενώ εφόσον οι $\lambda X, Z$ είναι ανεξάρτητες έχουμε ότι

$$V(Y) = V(\lambda X) + V(Z) \Rightarrow V(Z) = V(Y) - \lambda^2 V(X).$$

$\square$

Ως τώρα δεν έχουμε δει έναν τύπο για την συνάρτηση πυκνότητας πιθανότητας ενός ζεύγους (X_1,X_2) που ακολουθεί κανονική κατανομή σε δύο διαστάσεις. Θα ασχοληθούμε τώρα με αυτό το πρόβλημα.

Θεώρημα 3 Έστω ζεύγος τυχαίων μεταβλητών (X_1,X_2) που ακολουθεί κανονική κατανομή $\mathcal{N}(\mathbf{m},\mathbf{\Sigma})$ σε 2 διαστάσεις. Αν ο πίνακας συνδιακυμάνσεων $\mathbf{\Sigma}$ είναι αντιστρέψιμος, η συνάρτηση πυκνότητας πιθανότητας του ζευγους δίνεται από τον τύπο

$$f(x_1,x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)} \Big[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2 - 2\rho \left(\frac{x_1-\mu_1}{\sigma_1}\right) \left(\frac{x_2-\mu_2}{\sigma_2}\right) \Big] \right),$$

όπου $\sigma_i = \sqrt{\Sigma_{ii}}$ είναι η τυπική απόκλιση της X_i και $\rho = \frac{\Sigma_{12}}{\sqrt{\Sigma_{11}\Sigma_{22}}}$ είναι ο συντελεστής γραμμικής συσχέτισης των X_1, X_2. Μπορούμε να γράψουμε την πυκνότητα και σε πιο συμπαγή μορφή ως

$$f(\mathbf{x}) = \frac{1}{2\pi\sqrt{|\det(\mathbf{\Sigma})|}} \exp\left(-\frac{(\mathbf{x}-\mathbf{m})^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\mathbf{m})}{2} \right). \tag{6}$$

Απόδειξη: Η ιδέα της απόδειξης είναι να χρησιμοποιήσουμε το Λήμμα 3 ώστε να γράψουμε το ζεύγος (X_1, X_2) ως γραμμικό μετασχηματισμό ενός ζεύγους (U, V) όπου οι U και V είναι ανεξάρτητες κανονικές τυχαίες μεταβλητές. Η πυκνότητα του ζεύγους (U, V) θα είναι τότε γνωστή και μπορούμε να φτάσουμε από αυτήν στην πυκνότητα του ζεύγους (X_1, X_2).

Συγκεκριμένα από το Λήμμα 3 έχουμε ότι

$$X_2 = \frac{\Sigma_{12}}{\Sigma_{11}} X_1 + V$$

όπου η V είναι ανεξάρτητη της X_1 και ακολουθεί κατανομή $\mathcal{N}(\mu_V, \sigma_V^2)$, όπου $\mu_V = \mu_2 - \frac{\Sigma_{12}}{\Sigma_{11}}\mu_1$ και $\sigma_V^2 = \Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}}$. Αν θέσουμε $U = X_1$, εφόσον οι U και V είναι ανεξάρτητες, η από κοινού συνάρτηση πυκνότητας πιθανότητάς τους δίνεται από την

$$f_{U,V}(u, v) = f_{X_1}(u) f_V(v) = \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{(u-\mu_1)^2}{2\sigma_1^2}\right) \frac{1}{\sqrt{2\pi\sigma_V^2}} \exp\left(-\frac{(v-\mu_V)^2}{2\sigma_V^2}\right).$$

Έχουμε όμως $(X_1, X_2) = \Psi(U, V)$, όπου $\Psi(u, v) = (u, \frac{\Sigma_{12}}{\Sigma_{11}} u + v)$. Η Ιακωβιανή του αντίστροφου μετασχηματισμού $(u, v) = (x_1, -\Sigma_{12}/\Sigma_{11}\, x_1 + x_2)$ είναι ίση με

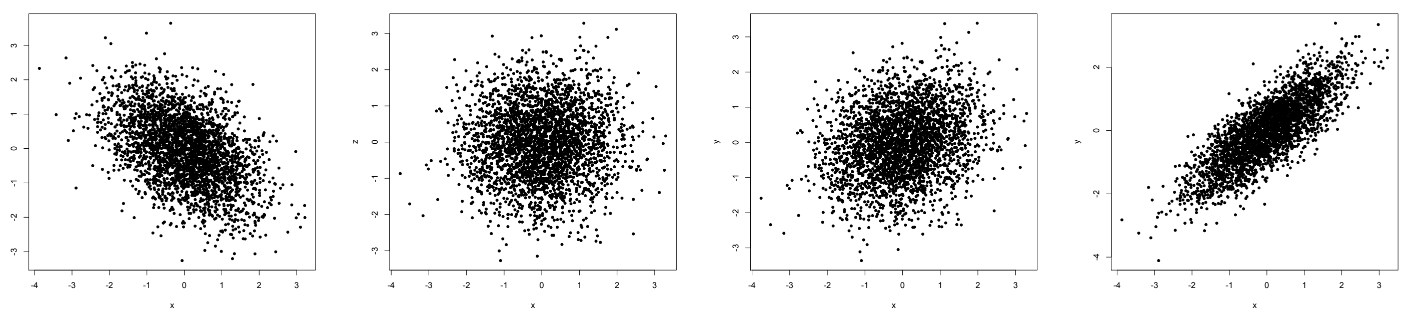
$$J(x_1, x_2) = \begin{vmatrix} 1 & 0 \\ -\frac{\Sigma_{12}}{\Sigma_{11}} & 1 \end{vmatrix} = 1.$$

Επομένως

$$f_{X_1,X_2}(x_1, x_2) = f_{U,V}(x_1, x_2 - \frac{\Sigma_{12}}{\Sigma_{11}}\, x_1) = f_{X_1}(x_1) f_V(x_2 - \frac{\Sigma_{12}}{\Sigma_{11}}\, x_1). \tag{7}$$

Το υπόλοιπο της απόδειξης είναι απλές αλγεβρικές πράξεις που μπορείτε να επιβεβαιώσετε εύκολα. □

Στο ακόλουθο Σχήμα φαίνονται 3.000 δείγματα του ζεύγους (X, Y) για διάφορες τιμές του συντελεστή συσχέτισης ρ. Σε όλες τις περιπτώσεις έχουμε $X \sim \mathcal{N}(0, 1)$ και $Y \sim \mathcal{N}(0, 1)$.



$\rho = -0.5$ $\qquad$ $\rho = 0$ $\qquad$ $\rho = +0.2$ $\qquad$ $\rho = +0.8$

Πόρισμα 3 Αν το ζεύγος τυχαίων μεταβλητών (X_1, X_2) ακολουθεί κανονική κατανομή $\mathcal{N}(\mathbf{m}, \mathbf{\Sigma})$ σε 2 διαστάσεις, τότε η δεσμευμένη συνάρτηση πυκνότητας πιθανότητας της X_2 δεδομένης της X_1 δίνεται από την

$$f_{X_2|X_1}(y|x) = f_V\big(y - \frac{\Sigma_{12}}{\Sigma_{11}}x\big) = \frac{1}{\sqrt{2\pi\Sigma_{22}(1-\rho^2)}} \exp\left(-\frac{1}{2\Sigma_{22}(1-\rho^2)}\Big(y - \mu_2 - \frac{\Sigma_{12}}{\Sigma_{11}}(x-\mu_1)\Big)^2\right).$$

Μπορούμε επομένως να πούμε ότι η δεσμευμένη κατανομή της X_2 δεδομένου ότι $X_1 = x$ είναι κανονική, με μέση τιμή

$$\mathbb{E}\big[X_2\big|\, X_1 = x\big] = \mu_2 + \frac{\Sigma_{12}}{\Sigma_{11}}(x - \mu_1)$$

και διασπορά

$$V\big(X_2\big|\, X_1 = x\big) = \Sigma_{22}(1 - \rho^2).$$

Απόδειξη: Προκύπτει άμεσα από την (7) και τους τύπους για τις μ_V και σ_V^2. □

Αν οι X_1, X_2 έχουν θετική συσχέτιση $(\Sigma_{12} > 0)$ και ξέρουμε ότι η X_1 πήρε μια τιμή x μεγαλύτερη από τη μέση τιμή της, παρατηρήστε ότι η δεσμευμένη μέση τιμή της X_2 είναι κι αυτή μεγαλύτερη από τη μέση τιμή μ_2 της X_2. Αντίθετα, αν $\Sigma_{12} < 0$, η $\mathbb{E}\big[X_2\big|\, X_1 = x\big]$ αποκλίνει από την μ_2 προς την αντίθετη κατεύθυνση από εκείνη που η x αποκλίνει από την μ_1. Παρατηρήστε τέλος ότι η γνώση της τιμής της X_1 μειώνει την αβεβαιότητά μας για την X_2 όταν υπάρχει κάποια μεταξύ τους συσχέτιση ($\rho \neq 0$). Πράγματι, η εκ των προτέρων διασπορά της τυχαίας μεταβλητής X_2 είναι $V(X_2) = \mathrm{Cov}\big(X_2, X_2\big) = \Sigma_{22}$, ενώ η διασπορά της X_2 με δεδομενη την X_1 είναι $V\big(X_2\big|\, X_1 = x\big) = \Sigma_{22}(1-\rho^2) < \Sigma_{22}$.

Παράδειγμα 6 Το ζευγάρι τυχαίων μεταβλητών (X, Y) ακολουθεί κανονική κατανομή σε 2 διαστάσεις με μέση τιμή $\mathbf{m} = \mathbf{0}$ και πίνακα συνδιακυμάνσεων

$$\mathbf{\Sigma} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

Η περιθώρια κατανομή της Y είναι κανονική με μέση τιμή $\mu_2 = 0$ και διασπορά $\Sigma_{22} = 2$. Επομένως η συνάρτηση πυκνότητας πιθανότητας της Y είναι

$$f_Y(y) = \frac{1}{\sqrt{4\pi}} \exp\left(-\frac{y^2}{4}\right), \qquad y \in \mathbb{R}.$$

Με δεδομένο ότι $X = x$, η δεσμευμένη κατανομή της Y είναι κανονική με μέση τιμή

$$\mu_2 + \frac{\Sigma_{12}}{\Sigma_{11}}(x - \mu_1) = 0 + \frac{1}{1}(x - 0) = x$$

και διασπορά

$$\Sigma_{22}(1-\rho^2) = \Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}} = 2 - \frac{1^2}{1} = 1.$$

Επομένως

$$f_{Y|X}(y|x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(y-x)^2}{2}\right), \qquad y \in \mathbb{R}.$$

$\square$

Θα γενικεύσουμε τώρα το Λήμμα 3 και τις συνέπειές του σε περισσότερες διαστάσεις. Ας υποθέσουμε λοιπόν ότι το $\mathbf{X}$ είναι ένα τυχαίο διάνυσμα με τιμές στον $\mathbb{R}^d$ και $\mathbf{X} \sim \mathcal{N}(\mathbf{m}_X, \mathbf{\Sigma}_{XX})$ και αντίστοιχα ότι το $\mathbf{Y}$ είναι ένα τυχαίο διάνυσμα με τιμές στον $\mathbb{R}^n$ και $\mathbf{Y} \sim \mathcal{N}(\mathbf{m}_Y, \mathbf{\Sigma}_{YY})$. Θα υποθέσουμε ακόμα ότι το τυχαίο διάνυσμα $\mathbf{W}$ στον $\mathbb{R}^{d+n}$ με

$$\mathbf{W} = \left(\begin{array}{c} \mathbf{X} \\ \hline \mathbf{Y} \end{array}\right)$$

ακολουθεί κανονική κατανομή $\mathcal{N}(\mathbf{m}, \mathbf{\Sigma})$ στον $\mathbb{R}^{d+n}$. Θα έχουμε τότε

$$\mathbf{m} = \left(\begin{array}{c} \mathbf{m}_X \\ \hline \mathbf{m}_Y \end{array}\right) \qquad \text{και} \qquad \mathbf{\Sigma} = \left(\begin{array}{c|c} \mathbf{\Sigma}_{XX} & \mathbf{\Sigma}_{XY} \\ \hline \mathbf{\Sigma}_{YX} & \mathbf{\Sigma}_{YY} \end{array}\right).$$

Οι πίνακες $\mathbf{\Sigma}_{XY}$ και $\mathbf{\Sigma}_{YX}$ που εμφανίζονται στην παραπάνω σχέση περιέχουν τις συνδιακυμάνσεις συντεταγμένων των $\mathbf{X}$ και $\mathbf{Y}$. Συγκεκριμένα ο $\mathbf{\Sigma}_{XY}$ είναι ένας $d \times n$ πίνακας με

$$(\mathbf{\Sigma}_{XY})_{ij} = \mathrm{Cov}\big(X_i, Y_j\big),\ 1 \le i \le d,\ 1 \le j \le n, \qquad \text{ενώ} \qquad \mathbf{\Sigma}_{YX} = \mathbf{\Sigma}_{XY}^T.$$

Ας προσπαθήσουμε τώρα να βρούμε έναν $n \times d$ πίνακα $\mathbf{\Lambda}$ ώστε

$$\mathbf{Y} = \mathbf{\Lambda}\mathbf{X} + \mathbf{Z} \tag{8}$$

όπου το τυχαίο διάνυσμα $\mathbf{Z}$ ακολουθεί κανονική κατανομή στον $\mathbb{R}^n$ και είναι ανεξάρτητο του $\mathbf{X}$. Η παραπάνω σχέση σημαίνει

$$Y_j = \sum_{i=1}^{d} \Lambda_{ji} X_i + Z_j, \qquad j = 1, 2, \ldots, n.$$

Παίρνοντας τη συνδιακύμανση των δύο μελών με την τυχαία μεταβλητή X_k, για $1 \le k \le d$ έχουμε ότι

$$\mathrm{Cov}\big(Y_j, X_k\big) = \mathrm{Cov}\big(\sum_{i=1}^{d} \Lambda_{ji} X_i, X_k\big) + \mathrm{Cov}\big(Z_j, X_k\big) = \sum_{i=1}^{d} \Lambda_{ji} \mathrm{Cov}\big(X_i, X_k\big),$$

επομένως

$$(\mathbf{\Sigma}_{YX})_{jk} = (\mathbf{\Lambda}\ \mathbf{\Sigma}_{XX})_{jk}, \qquad 1 \le k \le d,\ 1 \le j \le n.$$

Βλέπουμε δηλαδή ότι αν η (8) ισχύει για κάποιο $\mathbf{Z}$ ανεξάρτητο του $\mathbf{X}$ θα πρέπει

$$\mathbf{\Sigma}_{YX} = \mathbf{\Lambda}\ \mathbf{\Sigma}_{XX}.$$

Το ακόλουθο Λήμμα γενικεύει το Λήμμα 3 σε πολλές διαστάσεις.

Λήμμα 4 Έστω ζεύγος τυχαίων διανυσμάτων $(\mathbf{X}, \mathbf{Y}) \sim \mathcal{N}(\mathbf{m}, \mathbf{\Sigma})$ σε $d+n$ διαστάσεις, με

$$\mathbf{m} = \left(\begin{array}{c} \mathbf{m}_X \\ \hline \mathbf{m}_Y \end{array}\right) \qquad \text{και} \qquad \mathbf{\Sigma} = \left(\begin{array}{c|c} \mathbf{\Sigma}_{XX} & \mathbf{\Sigma}_{XY} \\ \hline \mathbf{\Sigma}_{YX} & \mathbf{\Sigma}_{YY} \end{array}\right).$$

και $\mathbf{\Sigma}_{XX}$ αντιστρέψιμο. Τότε

$$\mathbf{Y} = \mathbf{\Lambda}\,\mathbf{X} + \mathbf{Z}, \qquad \text{όπου } \mathbf{\Lambda} = \mathbf{\Sigma}_{YX}\mathbf{\Sigma}_{XX}^{-1} \in \Pi_{n\times d}$$

και το τυχαίο διάνυσμα $\mathbf{Z}$ στον $\mathbb{R}^n$ είναι ανεξάρτητο του $\mathbf{X}$ με

$$\mathbf{Z} \sim \mathcal{N}\big(\mathbf{m}_Y - \mathbf{\Lambda}\,\mathbf{m}_X, \mathbf{\Sigma}_{YY} - \mathbf{\Sigma}_{YX}\mathbf{\Sigma}_{XX}^{-1}\mathbf{\Sigma}_{XY}\big).$$

Απόδειξη: Η απόδειξη ακολουθεί τα επιχειρήματα της μονοδιάστατης περίπτωσης. Ορίζουμε

$$\mathbf{Z} = \mathbf{Y} - \mathbf{\Sigma}_{YX}\mathbf{\Sigma}_{XX}^{-1}\mathbf{X}.$$

Το ζεύγος $(\mathbf{X}, \mathbf{Z})$ ακολουθεί κανονική κατανομή σε $d+n$ διαστάσεις αφού το ζεύγος $(\mathbf{X}, \mathbf{Y})$ ακολουθεί κανονική κατανομή και

$$\left(\begin{array}{c} \mathbf{X} \\ \hline \mathbf{Z} \end{array}\right) = \left(\begin{array}{c|c} \mathbf{I}_d & \mathbf{0} \\ \hline -\mathbf{\Lambda} & \mathbf{I}_n \end{array}\right) \left(\begin{array}{c} \mathbf{X} \\ \hline \mathbf{Y} \end{array}\right).$$

Επιπλέον, για κάθε i, j με $1 \le i \le d$ και $1 \le j \le n$ έχουμε

$$\mathrm{Cov}\big(Z_j, X_i\big) = \mathrm{Cov}\big(Y_j - \sum_{k=1}^{d} \Lambda_{jk} X_k, X_i\big) = \mathrm{Cov}\big(Y_j, X_i\big) - \sum_{k=1}^{d} \lambda_{jk}\mathrm{Cov}\big(X_k, X_i\big) = (\mathbf{\Sigma}_{YX} - \mathbf{\Lambda}\mathbf{\Sigma}_{XX})_{ji} = 0.$$

Από το Θεώρημα 2 οι X_i και Z_j είναι ανεξάρτητες. Από το Πόρισμα 1 η $\mathbf{Z}$ ακολουθεί κανονική κατανομή σε n διαστάσεις. Για να βρούμε τις παραμέτρους της κατανομής της $\mathbf{Z}$ παρατηρούμε ότι για $j \in \{1, \ldots, n\}$

$$\mathbb{E}\big[Y_j\big] = \sum_{k=1}^{d} \Lambda_{jk}\mathbb{E}\big[X_k\big] + \mathbb{E}\big[Z_j\big] \Longleftrightarrow \mathbf{m}_Z = \mathbf{m}_Y - \mathbf{\Lambda}\,\mathbf{m}_X,$$

ενώ εφόσον οι X_k, Z_j είναι ανεξάρτητες, για κάθε $i, j = 1, 2, \ldots, n$ έχουμε

$$\mathrm{Cov}\big(Y_i, Y_j\big) = \mathrm{Cov}\big(\sum_{k=1}^{d} \Lambda_{ik} X_k + Z_i, \sum_{r=1}^{d} \Lambda_{jr} X_r + Z_j\big) = \sum_{k,r=1}^{d} \Lambda_{ik}\mathrm{Cov}\big(X_k, X_r\big)\Lambda_{rj}^T + \mathrm{Cov}\big(Z_i, Z_j\big).$$

Επομένως, αν συμβολίσουμε με $\mathbf{\Sigma}_{ZZ}$ τον πίνακα συνδιακυμάνσεων του $\mathbf{Z}$,

$$\mathbf{\Sigma}_{YY} = \mathbf{\Lambda}\,\mathbf{\Sigma}_{XX}\mathbf{\Lambda}^T + \mathbf{\Sigma}_{ZZ} \Longleftrightarrow \mathbf{\Sigma}_{ZZ} = \mathbf{\Sigma}_{YY} - \mathbf{\Sigma}_{YX}\mathbf{\Sigma}_{XX}^{-1}\mathbf{\Sigma}_{XY}.$$

$\square$

Εντελώς αντίστοιχα μπορούμε να βρούμε τη συνάρτηση πυκνότητας πιθανότητας ενός διανύσματος $\mathbf{X}$ που ακολουθεί κανονική κατανομή σε d διαστάσεις. Ο τύπος (6) παίρνει τη μορφή

$$f(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^d |\det(\mathbf{\Sigma})|}} \exp\left(-\frac{(\mathbf{x}-\mathbf{m})^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\mathbf{m})}{2}\right) \tag{9}$$

ενώ για τη δεσμευμένη κατανομή του $\mathbf{Y}$ δεδομένου του $\mathbf{X}$ έχουμε

$$f_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}) = f_{\mathbf{Z}}\big(\mathbf{y} - \mathbf{\Sigma}_{YX}\mathbf{\Sigma}_{XX}^{-1}\mathbf{x}\big).$$

Μπορούμε επομένως να πούμε ότι η δεσμευμένη κατανομή του $\mathbf{Y}$ δεδομένου ότι $\mathbf{X} = \mathbf{x}$ είναι κανονική, με μέση τιμή

$$\mathbb{E}\big[\mathbf{Y}\big|\,\mathbf{X} = \mathbf{x}\big] = \mathbf{m}_Y + \mathbf{\Sigma}_{YX}\mathbf{\Sigma}_{XX}^{-1}(\mathbf{x} - \mathbf{m}_X) \tag{10}$$

και πίνακα συνδιακυμάνσεων

$$\mathbf{\Sigma}_{YY}\Big|_{\mathbf{X}=\mathbf{x}} = \mathbf{\Sigma}_{ZZ} = \mathbf{\Sigma}_{YY} - \mathbf{\Sigma}_{YX}\mathbf{\Sigma}_{XX}^{-1}\mathbf{\Sigma}_{XY}. \tag{11}$$

Παρατηρήστε και πάλι ότι η γνώση της τιμής του $\mathbf{X}$ μειώνει την αβεβαιότητά μας για το $\mathbf{Y}$ όταν υπάρχει κάποια μεταξύ τους συσχέτιση ($\mathbf{\Sigma}_{XY} \neq \mathbf{0}$). Πράγματι, για κάθε $\ell \in \mathbb{R}^n$ έχουμε

$$\ell^T\mathbf{\Sigma}_{ZZ}\ell = \ell^T\mathbf{\Sigma}_{YY}\ell - \ell^T\mathbf{\Sigma}_{YX}\mathbf{\Sigma}_{XX}^{-1}\mathbf{\Sigma}_{XY}\ell = \ell^T\mathbf{\Sigma}_{YY}\ell - (\mathbf{\Sigma}_{XY}\ell)^T\mathbf{\Sigma}_{XX}^{-1}(\mathbf{\Sigma}_{XY}\ell) \leq \ell^T\mathbf{\Sigma}_{YY}\ell.$$

Παράδειγμα 7 Η τριάδα (X, Y, Z) ακολουθεί κανονική κατανομή σε 3 διαστάσεις με μέση τιμή $\mathbf{m} = \mathbf{0}$ και πίνακα συνδιακυμάνσεων

$$\mathbf{\Sigma} = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 2 & -2 \\ 1 & -2 & 3 \end{pmatrix}.$$

Η κατανομή της Z είναι κανονική με μέση τιμή $\mu_3 = 0$ και διασπορά $\Sigma_{33} = 3$. Επομένως

$$f_Z(z) = \frac{1}{\sqrt{6\pi}}\exp\left(-\frac{z^2}{6}\right), \qquad z \in \mathbb{R}.$$

Για να βρούμε την δεσμευμένη κατανομή της Z δεδομένων των X, Y χωρίζουμε τον πίνακα $\mathbf{\Sigma}$ σε μπλοκ.

$$\mathbf{\Sigma} = \left(\begin{array}{cc|c} 1 & -1 & 1 \\ -1 & 2 & -2 \\ \hline 1 & -2 & 3 \end{array}\right).$$

Αν ξέρουμε ότι $X = x$ και $Y = y$, η δεσμευμένη κατανομή της Z είναι και πάλι κανονική, με μέση τιμή που δίνεται από την (10)

$$\mu_{x,y} = 0 + (1 \quad -2)\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}^{-1}\begin{pmatrix} x - 0 \\ y - 0 \end{pmatrix} = -y.$$

Η διασπορά της δεσμευμένης κατανομής της Z δίνεται από την (11) και είναι

$$\sigma_{x,y}^2 = 3 - (1 \quad -2)\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}^{-1}\begin{pmatrix} 1 \\ -2 \end{pmatrix} = 3 - 2 = 1.$$

Επομένως, η δεσμευμένη συνάρτηση πυκνότητας πιθανότητας της Z είναι

$$f_{Z|X,Y}(z|x,y) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{(z+y)^2}{2}\right), \qquad z \in \mathbb{R}. \tag{12}$$

Εναλλακτικά μπορούμε να επαναλάβουμε το επιχείρημα της απόδειξης του Λήμματος 4. Πρώτα ψάχνουμε να βρούμε σταθερές α, β ώστε η τυχαία μεταβλητή $W = Z - \alpha X - \beta Y$ να είναι ανεξάρτητη από τις X, Y. Από το Θεώρημα 1 η από κοινού κατανομή των X, Y, W είναι κανονική σε 3 διαστάσεις, διότι

$$\begin{pmatrix} X \\ Y \\ W \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\alpha & -\beta & 1 \end{pmatrix}\begin{pmatrix} X \\ Y \\ Z \end{pmatrix}.$$

Από το Θεώρημα 2, προκειμένου η W ανεξάρτητη από τις X, Y αρκεί να έχουμε $\mathrm{Cov}\big(W, X\big) = \mathrm{Cov}\big(W, Y\big) = 0$. Οι εξισώσεις αυτές δίνουν

$$0 = \mathrm{Cov}\big(W, X\big) = \mathrm{Cov}\big(Z - \alpha X - \beta Y, X\big) = \Sigma_{31} - \alpha\Sigma_{11} - \beta\Sigma_{21} = 1 - \alpha + \beta$$

και
$$0=\mathrm{Cov}\big(W,Y\big)=\mathrm{Cov}\big(Z-\alpha X-\beta Y,Y\big)=\Sigma_{32}-\alpha\Sigma_{12}-\beta\Sigma_{22}=-2+\alpha-2\beta.$$
Λύνοντας το σύστημα βρίσκουμε ότι $\alpha=0,\ \beta=-1$, επομένως η $W=Y+Z$ είναι ανεξάρτητη από τις X,Y. Η μέση τιμή της W είναι μηδέν, ενώ η διασπορά της είναι

$$\sigma_W^2=\mathrm{Cov}\big(W,W\big)=\mathrm{Cov}\big(Y+Z,Y+Z\big)=\mathrm{Cov}\big(Y,Z\big)+2\mathrm{Cov}\big(Y,Z\big)+\mathrm{Cov}\big(Z,Z\big)=\Sigma_{22}+2\Sigma_{23}+\Sigma_{33}=2+2(-2)+3=1.$$

Έχουμε επομένως ότι $Z=-Y+W$, όπου η W είναι ανεξάρτητη από τις X,Y με $W\sim\mathcal{N}(0,1)$. Έτσι, με δεδομένο ότι $X=x,Y=y$ η δεσμευμένη κατανομή της Z είναι $\mathcal{N}(-y,1)$ και η δεσμευμένη πυκνότητα δίνεται από την (12).

Μια τρίτη προσέγγιση θα ήταν να βρούμε την από κοινού συνάρτηση πυκνότητας πιθανότητας $f_{X,Y,Z}$ των X,Y,Z, την περιθώρια συνάρτηση πυκνότητας πιθανότητας $f_{X,Y}$ των X,Y και να υπολογίσουμε την δεσμευμένη συνάρτηση πυκνότητας

$$f_{Z|X,Y}(z|x,y)=\frac{f_{X,Y,Z}(x,y,z)}{f_{X,Y}(x,y)}.$$

Αυτός ο τρόπος όμως είναι μάλλον πιο επίπονος και λιγότερο παιδαγωγικός. □

Παράδειγμα 8 Η τριάδα (X,Y,Z) ακολουθεί κανονική κατανομή σε 3 διαστάσεις με μέση τιμή $\mathbf{m}=\mathbf{0}$ και πίνακα συνδιακυμάνσεων

$$\boldsymbol{\Sigma}=\begin{pmatrix}1&1&1\\1&2&2\\1&2&3\end{pmatrix}.$$

Θέλουμε να βρούμε τη δεσμευμένη κατανομή του ζεύγους (X,Y) δεδομένου ότι $Z=z$. Η δεσμευμένη μέση τιμή του ζεύγους δεδομένου ότι $Z=z$ δίνεται από την (10) και είναι

$$\begin{pmatrix}\mu_{x|z}\\\mu_{y|z}\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}+\begin{pmatrix}1\\2\end{pmatrix}(3)^{-1}(z-0)=z\begin{pmatrix}1/3\\2/3\end{pmatrix}.$$

Ο δεσμευμένος πίνακας συνδιακυμάνσεων των X,Y δεδομένου ότι $Z=z$ δίνεται από την (11) και είναι

$$\boldsymbol{\Sigma}_z=\begin{pmatrix}1&1\\1&2\end{pmatrix}-\begin{pmatrix}1\\2\end{pmatrix}(3)^{-1}\,(1\quad 2)=\begin{pmatrix}1&1\\1&2\end{pmatrix}-\begin{pmatrix}1/3&2/3\\2/3&4/3\end{pmatrix}=\begin{pmatrix}2/3&1/3\\1/3&2/3\end{pmatrix}.$$

Μπορούμε τώρα να γράψουμε τη δεσμευμένη συνάρτηση πυκνότητας πιθανότητας του ζεύγους (X,Y)

$$f_{X,Y|Z}(x,y|z)=\frac{\sqrt{3}}{2\pi}\exp\left(-x^2-y^2+xy+zy-\frac{z^2}{3}\right). \tag{13}$$

Εναλλακτικά μπορούμε να επαναλάβουμε το επιχείρημα της απόδειξης. Από το Λήμμα 3 έχουμε

$$X=\frac{\Sigma_{13}}{\Sigma_{33}}Z+W_1=\frac{1}{3}Z+W_1\qquad\text{και}\qquad Y=\frac{\Sigma_{23}}{\Sigma_{33}}Z+W_2=\frac{2}{3}Z+W_2$$

όπου οι W_1 και W_2 είναι ανεξάρτητες της Z. Έχουμε επομένως ότι

$$\begin{pmatrix}X\\Y\end{pmatrix}=Z\begin{pmatrix}1/3\\2/3\end{pmatrix}+\mathbf{W}$$

όπου το διάνυσμα $\mathbf{W}$ είναι ανεξάρτητο της Z. Από το Θεώρημα 1 το $\mathbf{W}$ ακολουθεί κανονική κατανομή σε δύο διαστάσεις αφού

$$\mathbf{W}=\begin{pmatrix}1&0&-1/3\\0&1&-2/3\end{pmatrix}\begin{pmatrix}X\\Y\\Z\end{pmatrix}.$$

Μπορούμε πολύ εύκολα να προσδιορίσουμε τις παραμέτρους της κατανομής του $\mathbf{W}$. Συγκεκριμένα,

$$\mathbb{E}\big[W_1\big]=\mathbb{E}\big[X\big]-\frac{1}{3}\mathbb{E}\big[Z\big]=0\qquad\text{και}\qquad\mathbb{E}\big[W_2\big]=\mathbb{E}\big[Y\big]-\frac{2}{3}\mathbb{E}\big[Z\big]=0.$$

Από το Λήμμα 3 έχουμε ακόμα ότι

$$V(W_1) = V(X) - \left(\frac{1}{3}\right)^2 V(Z) = \Sigma_{11} - \frac{1}{9}\Sigma_{33} = \frac{2}{3} \qquad \text{και} \qquad V(W_2) = V(Y) - \left(\frac{2}{3}\right)^2 V(Z) = \Sigma_{22} - \frac{4}{9}\Sigma_{33} = \frac{2}{3}.$$

Τέλος,

$$\mathrm{Cov}\big(W_1, W_2\big) = \mathrm{Cov}\big(X - \frac{1}{3}Z, Y - \frac{2}{3}Z\big) = \Sigma_{12} - \frac{2}{3}\Sigma_{13} - \frac{1}{3}\Sigma_{32} + \frac{2}{9}\Sigma_{33} = \frac{1}{3}.$$

Έχουμε λοιπόν ότι

$$\begin{pmatrix} X \\ Y \end{pmatrix} = Z \begin{pmatrix} 1/3 \\ 2/3 \end{pmatrix} + \mathbf{W}$$

όπου το $\mathbf{W}$ είναι ανεξάρτητο από την Z με $\mathbf{W} \sim \mathcal{N}\left(\mathbf{0}, \begin{pmatrix} 2/3 & 1/3 \\ 1/3 & 2/3 \end{pmatrix}\right)$. Επομένως

$$f_{X,Y|Z}(x, y|z) = f_{\mathbf{W}}(x - z/3, y - 2z/3)$$

και καταλήγουμε πάλι στην (13).

Μια τρίτη προσέγγιση θα ήταν να βρούμε την από κοινού συνάρτηση πυκνότητας πιθανότητας $f_{X,Y,Z}$ των X, Y, Z, την περιθώρια συνάρτηση πυκνότητας πιθανότητας f_Z της X και να υπολογίσουμε τη δεσμευμένη συνάρτηση πυκνότητας πιθανότητας ως

$$f_{X,Y|Z}(x, y|z) = \frac{f_{X,Y,Z}(x, y, z)}{f_Z(z)}.$$

Για να υπολογίσουμε την $f_{X,Y,Z}$ χρειαζόμαστε την $\det(\mathbf{\Sigma})$=1 και τον αντίστροφο του $\mathbf{\Sigma}$

$$\mathbf{\Sigma}^{-1} = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}.$$

Από τον τύπο (9) έχουμε

$$f_{X,Y,Z} = \frac{1}{(2\pi)^{\frac{3}{2}}} \exp\left(-x^2 - y^2 - \frac{z^2}{2} + xy + yz\right), \qquad (x, y, z) \in \mathbb{R}^3.$$

Από την άλλη, η Z είναι κανονική τυχαία μεταβλητή με μέση τιμή $\mu_3 = 0$ και διασπορά $\Sigma_{33} = 3$. Επομένως

$$f_Z(z) = \frac{1}{\sqrt{6\pi}} \exp\left(-\frac{z^2}{6}\right), \qquad z \in \mathbb{R}.$$

Διαιρώντας τις παραπάνω σχέσεις ξαναβρίσκουμε την (13). $\square$

4 Εφαρμογές στη Στατιστική

Ας θεωρήσουμε μια ακολουθία $\{X_i\}_{1 \le i \le n}$ από ανεξάρτητες, ισόνομες τυχαίες μεταβλητές με κοινή κατανομή $\mathcal{N}(\mu, \sigma^2)$. Αν ορίσουμε

$$\bar{X} = \frac{X_1 + \cdots + X_n}{n},$$

τότε από το Θεώρημα 1 έχουμε ότι

$$\bar{X} \sim \mathcal{N}(\mu, \frac{\sigma^2}{n}) \quad \Longleftrightarrow \quad \sqrt{n}\left(\frac{\bar{X} - \mu}{\sigma}\right) \sim \mathcal{N}(0, 1). \tag{14}$$

Ας ορίσουμε τώρα

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2. \tag{15}$$

Θεώρημα 4 Αν οι $\{X_i\}_{1\le i\le n}$ είναι ανεξάρτητες, ισόνομες τυχαίες μεταβλητές με κοινή κατανομή $\mathcal{N}(\mu,\sigma^2)$, τότε η S^2 που ορίζεται στην (15) είναι ανεξάρτητη της $\bar{X}$ και

$$\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1).$$

Απόδειξη: Από το Παράδειγμα 4 έχουμε ότι οι $X_i-\bar{X}$ είναι ανεξάρτητες από την $\bar{X}$ για $i=1,2,\ldots,n$, επομένως και η S^2 είναι ανεξάρτητη από την $\bar{X}$. Επιπλέον έχουμε ότι

$$\begin{aligned}(n-1)S^2&=\sum_{i=1}^n\left(X_i-\mu-(\bar{X}-\mu)\right)^2=\sum_{i=1}^n\left[(X_i-\mu)^2+(\bar{X}-\mu)^2-2(\bar{X}-\mu)(X_i-\mu)\right]\\&=\sum_{i=1}^n(X_i-\mu)^2+n(\bar{X}-\mu)^2-2(\bar{X}-\mu)\sum_{i=1}^n(X_i-\mu)\\&=\sum_{i=1}^n(X_i-\mu)^2-n(\bar{X}-\mu)^2.\end{aligned}$$

Μπορούμε να ξαναγράψουμε την τελευταία σχέση ως εξής.

$$\sum_{i=1}^n\left(\frac{X_i-\mu}{\sigma}\right)^2=\frac{(n-1)S^2}{\sigma^2}+\left(\sqrt{n}\ \frac{\bar{X}-\mu}{\sigma}\right)^2 \tag{16}$$

Εφόσον οι τυχαίες μεταβλητές $\frac{X_i-\mu}{\sigma}$ είναι ανεξάρτητες τυπικές κανονικές, το αριστερό μέλος της (16) είναι το άθροισμα των τετραγώνων n ανεξάρτητων τυπικών κανονικών τυχαίων μεταβλητών, επομένως ακολουθεί κατανομή $\chi^2(n)$, δηλαδή

$$\sum_{i=1}^n\left(\frac{X_i-\mu}{\sigma}\right)^2\sim\chi^2(n).$$

και έχει χαρακτηριστική συνάρτηση

$$\phi_n(t)=(1-2it)^{-\frac{n}{2}}.$$

Εφόσον η S^2 και η $\bar{X}$ είναι ανεξάρτητες, το δεξί μέλος της (16) είναι το άθροισμα δύο ανεξάρτητων τυχαίων μεταβλητών. Η μία απ' αυτές, λόγω της (14), ακολουθεί κατανομή $\chi^2(1)$

$$\left(\sqrt{n}\ \frac{\bar{X}-\mu}{\sigma}\right)^2\sim\chi^2(1)$$

και έχει χαρακτηριστική συνάρτηση

$$\phi_1(t)=(1-2it)^{-\frac{1}{2}}.$$

Εξισώνοντας τις χαρακτηριστικές συναρτήσεις των δύο μελών της (16) και συμβολίζοντας με ψ τη χαρακτηριστική συνάρτηση της $(n-1)S^2/\sigma^2$ έχουμε

$$\phi_n(t)=(1-2it)^{-\frac{n}{2}}=\psi(t)\times(1-2it)^{-\frac{1}{2}}\Longleftrightarrow\psi(t)=\phi_n(t)=(1-2it)^{-\frac{n-1}{2}}.$$

Εφόσον η χαρακτηριστική συνάρτηση μιας κατανομής χαρακτηρίζει την κατανομή θα πρέπει επομένως

$$\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1),$$

που ολοκληρώνει την απόδειξη του Θεωρήματος. □

Λήμμα 5 Αν $X\sim\mathcal{N}(0,1)$, $Y\sim\chi^2(n)$ και οι X,Y είναι ανεξάρτητες, τότε η συνάρτηση πυκνότητας πιθανότητας της τ.μ.

$$Z=\frac{X}{\sqrt{Y/n}}$$

είναι

$$f_n(z)=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{\pi n}}\left(1+\frac{z^2}{n}\right)^{-\frac{n+1}{2}}. \tag{17}$$

Απόδειξη: Εφόσον οι X και Y είναι ανεξάρτητες, η από κοινού συνάρτηση πυκνότητας πιθανότητας του ζεύγους (X,Y) είναι η

$$f_{X,Y}(x,y) = f(x)g_n(y) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}y^{\frac{n}{2}-1}e^{-\frac{y}{2}}, \qquad (x,y)\in\mathbb{R}\times\mathbb{R}_+.$$

Θα υπολογίσουμε τη συνάρτηση κατανομής πιθανότητας της Z.

$$F_Z(t) = \mathbb{P}\big[Z \le t\big] = \mathbb{P}\big[X \le t\sqrt{Y/n}\big] = \mathbb{P}\big[(X,Y)\in D_t\big] = \iint_{D_t} f(x)g(y)\ dxdy,$$

όπου $D_t = \{(x,y)\in\mathbb{R}\times\mathbb{R}_+ :\ x \le t\sqrt{y/n}\}$. Επομένως, με την αλλαγή μεταβλητής $x = z\sqrt{y/n}$ έχουμε

$$\begin{aligned} F_Z(t) &= \int_0^\infty \int_{-\infty}^{t\sqrt{y/n}} f(x)g(y)dx\ dy = \int_0^\infty \int_{-\infty}^{t} \sqrt{\frac{y}{n}} f\left(\sqrt{\frac{y}{n}}z\right) dz\ g(y)dy \\ &= \int_{-\infty}^{t} \int_0^\infty \sqrt{\frac{y}{n}} f\left(\sqrt{\frac{y}{n}}z\right) g(y)dy\ dz. \end{aligned}$$

Η πυκνότητα της Z είναι λοιπόν ίση με

$$\begin{aligned} f_n(z) &= \int_0^\infty \sqrt{\frac{y}{n}} f\left(\sqrt{\frac{y}{n}}z\right) g(y)dy = \int_0^\infty \sqrt{\frac{y}{2\pi n}} e^{-\frac{yz^2}{2n}} \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} y^{\frac{n}{2}-1} e^{-\frac{y}{2}}\ dy \\ &= \frac{1}{\sqrt{2\pi n}2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} \int_0^\infty y^{\frac{n-1}{2}} e^{-\frac{y}{2}(1+\frac{z^2}{n})}\ dy = \frac{(1+\frac{z^2}{n})^{-\frac{n+1}{2}}}{\sqrt{\pi n}\Gamma\left(\frac{n}{2}\right)} \int_0^\infty t^{\frac{n-1}{2}} e^{-t}\ dt, \end{aligned}$$

όπου στο τελευταίο βήμα κάναμε την αλλαγή μεταβλητής $t = \frac{y}{2}(1+\frac{z^2}{n})$. Επομένως,

$$f_n(z) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{\pi n}} \left(1+\frac{z^2}{n}\right)^{-\frac{n+1}{2}}.$$

□

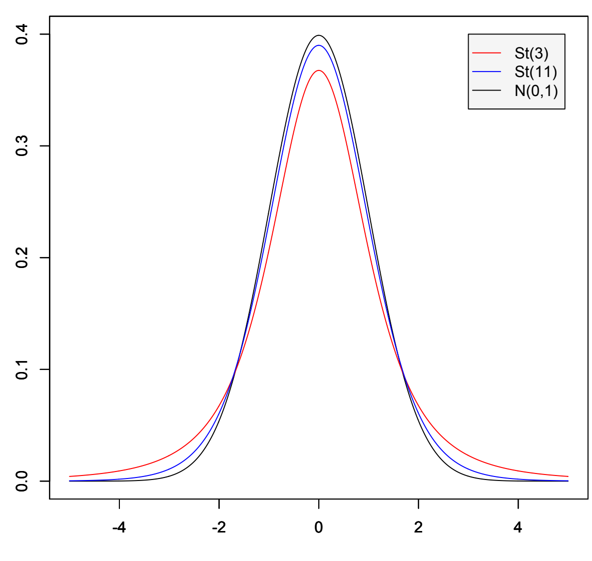
Ορισμός 3 Αν μια τυχαία μεταβλητή T έχει συνάρτηση πυκνότητας πιθανότητας την f_n της (17) λέμε ότι η T ακολουθεί την κατανομή Student με n βαθμούς ελευθερίας και συμβολίζουμε $T \sim St(n)$.

Παρατηρήστε ότι εφόσον $f_n(-z) = f_n(z)$ για κάθε $z\in\mathbb{R}$, η κατανομή Student είναι συμμετρική ως προς το 0. Παρατηρήστε ακόμα ότι καθώς οι βαθμοί ελευθερίας n της κατανομής Student αυξάνονται, αυτή μοιάζει όλο και περισσότερο με την τυπική κανονική κατανομή. Στο διπλανό σχήμα φαίνονται οι πυκνότητες δύο Student με 3 (κόκκινο) και 11 (μπλε) βαθμούς ελευθερίας και μιας τυπικής κανονικής κατανομής (μαύρο). Με τη βοήθεια του τύπου του Stirling μπορείτε να δείτε ότι

$$\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{\pi n}} \overset{n\to\infty}{\longrightarrow} \frac{1}{\sqrt{2\pi}},$$

ενώ

$$\left(1+\frac{z^2}{n}\right)^{-\frac{n+1}{2}} \overset{n\to\infty}{\longrightarrow} e^{-\frac{z^2}{2}}.$$



Πόρισμα 4 Αν οι $\{X_i\}_{1\le i\le n}$ είναι ανεξάρτητες, ισόνομες τυχαίες μεταβλητές με κοινή κατανομή $\mathcal{N}(\mu,\sigma^2)$, τότε

$$T = \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim St(n-1).$$

Απόδειξη: Είδαμε ότι οι τυχαίες μεταβλητές

$$Z_1 = \sqrt{n}\ \frac{\bar{X}-\mu}{\sigma} \sim \mathcal{N}(0,1) \qquad \text{και} \qquad Z_2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

είναι ανεξάρτητες, το Λήμμα 5 δίνει ότι

$$\frac{Z_1}{\sqrt{Z_2/(n-1)}} = \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim St(n-1).$$

□

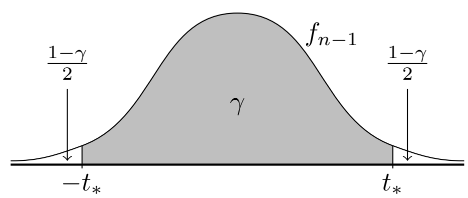
Τα επόμενα δύο θεωρήματα κατασκευάζουν διαστήματα εμπιστοσύνης για τη μέση τιμή μ μιας κανονικής κατανομής χωρίς πληροφορία για τη διασπορά της και για τη διασπορά σ^2 χωρίς πληροφορία για τη μέση τιμή της.

Θεώρημα 5 Αν οι $\{X_i\}_{1\le i\le n}$ είναι ανεξάρτητες, ισόνομες τυχαίες μεταβλητές με κοινή κατανομή $\mathcal{N}(\mu,\sigma^2)$ και $t_*>0$ είναι ένας πραγματικός αριθμός ώστε

$$\int_{-\infty}^{t_*} f_{n-1}(z)\ dz = \frac{1+\gamma}{2}, \tag{18}$$

τότε

$$\mathbb{P}\Big[\bar{X} - t_*\frac{S}{\sqrt{n}} \le \mu \le \bar{X} + t_*\frac{S}{\sqrt{n}}\Big] = \gamma.$$



Απόδειξη: Από το Πόρισμα 4 η τυχαία μεταβλητή $T=\frac{\bar{X}-\mu}{S/\sqrt{n}}$ έχει συνάρτηση πυκνότητα πιθανότητας την f_{n-1}. Από την (18) έχουμε ότι $\mathbb{P}\big[T>t_*\big]=\frac{1-\gamma}{2}$. Από τη συμμετρία της κατανομής έχουμε επίσης $\mathbb{P}\big[T<-t_*\big]=\frac{1-\gamma}{2}$. Επομένως,

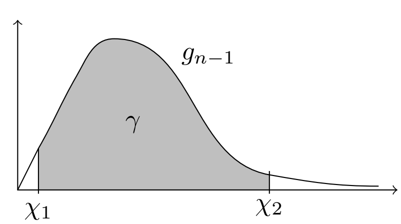
$$\mathbb{P}\big[|T|\le t_*\big]=\gamma \iff \mathbb{P}\Big[-t_* \le \frac{\mu-\bar{X}}{S/\sqrt{n}} \le t_*\Big]=\gamma \iff \mathbb{P}\Big[\bar{X} - t_*\frac{S}{\sqrt{n}} \le \mu \le \bar{X} + t_*\frac{S}{\sqrt{n}}\Big] = \gamma. \qquad \square$$

Θεώρημα 6 Αν οι $\{X_i\}_{1\le i\le n}$ είναι ανεξάρτητες, ισόνομες τυχαίες μεταβλητές με κοινή κατανομή $\mathcal{N}(\mu,\sigma^2)$ και $0<\chi_1<\chi_2$ είναι δύο πραγματικοί αριθμοί ώστε

$$\int_{\chi_1}^{\chi_2} g_{n-1}(z)\ dz = \gamma, \tag{19}$$

όπου g_n είναι η συνάρτηση πυκνότητας πιθανότητας της κατανομής $\chi^2(n)$, τότε

$$\mathbb{P}\Big[\frac{(n-1)S^2}{\chi_2} \le \sigma^2 \le \frac{(n-1)S^2}{\chi_1}\Big] = \gamma.$$



Απόδειξη: Από το Θεώρημα 4 η τυχαία μεταβλητή $Z=\frac{(n-1)S^2}{\sigma^2}$ έχει συνάρτηση πυκνότητα πιθανότητας την g_{n-1}. Από την (19) έχουμε ότι $\mathbb{P}\big[\chi_1\le Z\le\chi_2\big]=\gamma$. Επομένως,

$$\gamma = \mathbb{P}\big[\chi_1 \le \frac{(n-1)S^2}{\sigma^2} \le \chi_2\big] = \mathbb{P}\Big[\frac{(n-1)S^2}{\chi_2} \le \sigma^2 \le \frac{(n-1)S^2}{\chi_1}\Big]. \qquad \square$$

5 Ασκήσεις

Άσκηση 1 Αν οι τ.μ. X, Y είναι ανεξάρτητες, με ομοιόμορφη κατανομή στο διάστημα [0,1] δείξτε ότι οι τ.μ.

$$U = \sqrt{-2\ln X}\cos(2\pi Y) \qquad V = \sqrt{-2\ln X}\sin(2\pi Y)$$

είναι ανεξάρτητες και ακολουθούν την τυπική κανονική κατανομή. Σ' αυτόν τον υπολογισμό βασίζεται η μέθοδος Box-Muller για την προσομοίωση κανονικών τ.μ. στον υπολογιστή. Βλέπετε πώς;

Άσκηση 2 Οι $X_1, X_2, \ldots, X_n$ είναι ανεξάρτητες ισόνομες τ.μ. με κατανομή $\mathcal{N}(0, \sigma^2)$. Ορίζουμε $\bar{X} = \frac{X_1 + \cdots + X_n}{n}$. Ποια είναι η δεσμευμένη σ.π.π. $f_{X_1|\bar{X}}(x \mid \mu)$;

Άσκηση 3 Δείξτε ότι αν το τυχαίο διάνυσμα $(X, Y) \in \mathbb{R}^2$ ακολουθεί τυπική κανονική κατανομή τότε και το διάνυσμα

$$\begin{pmatrix} X' \\ Y' \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}$$

ομοίως ακολουθεί τυπική κανονική κατανομή. Δηλαδή η $\mathcal{N}(\mathbf{0}, \mathbb{I}_2)$ μένει αναλλοίωτη από στροφές.

Άσκηση 4 Η από κοινού κατανομή των τυχαίων μεταβλητών X, Y, Z είναι κανονική σε 3 διαστάσεις με μέση τιμή $\mathbf{m} = \mathbf{0}$ και πίνακα συνδιακυμάνσεων

$$\mathbf{\Sigma} = \begin{pmatrix} 4 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 4 \end{pmatrix}.$$

Βρείτε την δεσμευμένη κατανομή των X, Y δεδομένου ότι $Z = 0$. Βρείτε την απο κοινού κατανομή των τυχαίων μεταβλητών Y και $X - Z$. Ποια είναι η δεσμευμένη κατανομή της Y δεδομένου ότι $X = Z$;

Άσκηση 5 Έστω $X_0 \sim \mathcal{N}(0, 1)$ και

$$X_n = \frac{3}{5}X_{n-1} + \frac{4}{5}W_n$$

για κάθε $n \in \mathbb{N}$, όπου $\{W_n\}_{n\in\mathbb{N}}$ είναι μια ακολουθία από ανεξάρτητες, ισόνομες τυχαίες μεταβλητές με τυπική κανονική κατανομή και ανεξάρτητες από την X_0. Δείξτε ότι η από κοινού κατανομή των X_0, X_1, X_2, X_3 είναι κανονική σε 4 διαστάσεις. Ποιος είναι ο πίνακας συνδιακυμάνσεων; Ποια είναι η περιθώρια κατανομή των X_1, X_2, X_3; Ποια είναι η δεσμευμένη κατανομή της X_2 δεδομένου ότι $X_1 = x$ και $X_3 = y$;

Άσκηση 6 Ας είναι X, Y ανεξάρτητες, ισόνομες τ.μ. με κατανομή $\mathcal{N}(0, 2)$. Ορίζουμε $U = \frac{X+Y}{2}$, $V = \frac{X-Y}{2}$.
α) Ποια είναι η σ.π.π. του τυχαίου διανύσματος (X, Y);
β) Υπολογίστε την σ.π.π. του τυχαίου διανύσματος (U, V). Είναι οι τ.μ. U, V ανεξάρτητες;
γ) Έστω $S \sim \mathcal{N}(0, 1)$ ανεξάρτητη από τις X, Y. Υπολογίστε την σ.π.π. του διανύσματος $(U, \sqrt{V^2 + S^2})$.
δ) Θεωρούμε δύο ακόμα ανεξάρτητες τ.μ. Z, W με κατανομή $\mathcal{N}(0, 2)$ και ανεξάρτητες από τις X, Y.
Αν $A = \begin{pmatrix} X & Z \\ W & Y \end{pmatrix}$, και $\Lambda_1 \geq \Lambda_2$ είναι οι ιδιοτιμές του συμμετρικού πίνακα $\frac{A+A^\top}{2}$ υπολογίστε την από κοινού κατανομή των Λ_1, Λ_2.

Άσκηση 7 Οι στοχαστικές διαδικασίες είναι τυχαίες συναρτήσεις, και ένας τρόπος να τις φανταστεί κανείς είναι ότι η τιμή τους κάθε χρονική στιγμή είναι μια τυχαία μεταβλητή. Η κίνηση Brown είναι μια στοχαστική διαδικασία όπου $B_t \sim \mathcal{N}(0, t)$. Αυτό δεν καθορίζει πλήρως την κίνηση Brown αφού την ίδια ιδιότητα έχει και η διαδικασία $Y_t = \sqrt{t}\chi$ όπου $\chi \sim \mathcal{N}(0, 1)$ και $t \geq 0$. Χρειάζεται να καθορίσουμε και το πώς σχετίζονται μεταξύ τους οι τιμές της διαδικασίας σε διαφορετικούς χρόνους, πώς εξελίσσεται δηλαδή η διαδικασία στο χρόνο. Ο τρόπος που αυτό συμβαίνει για την κίνηση Brown είναι ότι για κάθε $s, t \in \mathbb{R}_+$ με $s < t$ η μεταβολή της $B_t - B_s$ είναι ανεξάρτητη από το παρελθόν της, είναι δηλαδή ανεξάρτητη από τις τ.μ. B_r με $r \leq s$ με κατανομή $\mathcal{N}(0, t - s)$. Αν $r < s < t$ βρείτε την από κοινού κατανομή των B_r, B_s, B_t και την δεσμευμένη σ.π.π. της B_s δεδομένων των B_r, B_t.

Άσκηση 8 Έστω $\{(X_i, Y_i)\}_{1\le i\le n}$ ανεξάρτητα, ισόνομα τυχαία διανύσματα με κατανομή

$$\begin{pmatrix} X_i \\ Y_i \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \mathbf{\Sigma}\right).$$

Ορίζουμε $\bar{X} = \frac{X_1+\cdots+X_n}{n}$, $\bar{Y} = \frac{Y_1+\cdots+Y_n}{n}$ και

$$S_D^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - Y_i + \bar{Y} - \bar{X})^2.$$

Αν το t_* είναι όπως στην (18), δείξτε ότι οποιοσδήποτε κι αν είναι ο πίνακας συνδιακυμάνσεων Σ έχουμε

$$\mathbb{P}\Big[\bar{X} - \bar{Y} - t_*\frac{S_D}{\sqrt{n}} \le \mu_1 - \mu_2 \le \bar{X} - \bar{Y} + t_*\frac{S_D}{\sqrt{n}}\Big] = \gamma. \tag{20}$$

Άσκηση 9 Δίνονται $\{X_i\}_{1\le i\le n}$ ανεξάρτητες, ισόνομες τ.μ. με κατανομή $\mathcal{N}(\mu_1, \sigma^2)$ και $\{Y_j\}_{1\le j\le m}$ ανεξάρτητες, ισόνομες τ.μ. με κατανομή $\mathcal{N}(\mu_2, \sigma^2)$, οι οποίες είναι και ανεξάρτητες από τις $\{X_i\}_{1\le i\le n}$. Ορίζουμε $\nu = \frac{nm}{n+m}$. Ποια κατανομή ακολουθεί η $\bar{X} - \bar{Y}$; Ορίζουμε ακόμα

$$(n+m-2)S^2 = \sum_{i=1}^{n}(X_i - \bar{X})^2 + \sum_{j=1}^{m}(Y_j - \bar{Y})^2 = (n-1)S_X^2 + (m-1)S_Y^2.$$

Ποια κατανομή ακολουθεί η τυχαία μεταβλητή

$$T = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{S/\sqrt{\nu}};$$

Μπορείτε να βρείτε μια σχέση όπως η (20) για τη διαφορά $\mu_1 - \mu_2$ που δεν χρησιμοποιεί την άγνωστη διασπορά σ^2;