Πιθανότητες ΙΙ – 2$^{\text{ο}}$ Μέρος[1]

Μιχάλης Ζαζάνης
Τμήμα Στατιστικής
Οικονομικό Πανεπιστήμιο Αθηνών

[1]Το κεφάλαιο 6 του βιβλίου *Βασικές Αρχές Θεωρίας Πιθανοτήτων* του Sheldon Ross καλύπτει σε μεγάλο βαθμό την ύλη του δευτέρου μέρους του μαθήματος

Κεφάλαιο 1

Συναρτήσεις Συνεχών Τυχαίων Μεταβλητών

1.1 Το πηλίκο δύο συνεχών τυχαίων μεταβλητών

Έστω X, Y, δύο συνεχείς τυχαίες μεταβλητές, ορισμένες στον ίδιο χώρο πιθανοτήτων, με από κοινού πυκνότητα πιθανότητας $f(x, y)$. Αν $Z := Y/X$, θέλουμε να βρούμε την πυκνότητα πιθανότητας της Z.

Αν ορίσουμε $A_z := \{(x, y) : y/x \leq z\}$.

$$F_Z(z) = \int\int_{A_z} f(x, y)dxdy = \int_{-\infty}^{0} \left(\int_{xz}^{\infty} f(x, y)dy\right) dx + \int_{0}^{\infty} \left(\int_{-\infty}^{xz} f(x, y)dy\right) dx$$

Με την αλλαγή μεταβλητής $y = xu$, $dy = xdu$ στο εσωτερικό ολοκλήρωμα,

$$\begin{aligned}
F_Z(z) &= \int_{-\infty}^{0} \left(\int_{z}^{-\infty} xf(x, xu)du\right) dx + \int_{0}^{\infty} \left(\int_{-\infty}^{z} xf(x, xu)du\right) dx \\
&= \int_{-\infty}^{0} \left(\int_{-\infty}^{z} (-x)f(x, xu)du\right) dx + \int_{0}^{\infty} \left(\int_{-\infty}^{z} xf(x, xu)du\right) dx \\
&= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{z} \mid x \mid f(x, xu)du\right) dx \\
&= \int_{-\infty}^{z} \left(\int_{-\infty}^{\infty} \mid x \mid f(x, ux)dx\right) du.
\end{aligned}$$

Από την παραπάνω σχέση συμπεραίνουμε άμεσα ότι ή πυκνότητα της Z δίνεται από τον τύπο

$$f_Z(z) = \int_{-\infty}^{\infty} \mid x \mid f(x, zx)dx. \tag{1.1}$$

Παράδειγμα 1.

Ας υποθέσουμε ότι οι τυχαίες μεταβλητές X και Y είναι ανεξάρτητες τυποποιημένες κανονικές μεταβλητές. Τότε $f(x,y) = \frac{1}{2\pi} e^{-\frac{1}{2}(x^2+y^2)}$ και επομένως η πυκνότητα πιθανότητας του πηλίκου είναι

$$f_Z(z) \;=\; \int_{-\infty}^{\infty} |\, x \,|\, \frac{1}{2\pi} e^{-\frac{1}{2}(x^2+z^2x^2)} dx \;=\; \frac{1}{\pi}\int_0^{\infty} e^{-\frac{x^2}{2}(1+z^2)} d\left(\frac{x^2}{2}\right) \;=\; \frac{1}{\pi}\frac{1}{1+z^2}, \quad z \in \mathbb{R}.$$

Συνάγουμε λοιπόν το συμπέρασμα ότι το πηλίκο δύο ανεξάρτητων, τυποποιημένων κανονικών τυχαίων μεταβλητών έχει κατανομή Cauchy.

Παράδειγμα 2.

Υποθέτουμε τώρα ότι οι τυχαίες μεταβλητές X και Y είναι ανεξάρτητες με κατανομή Γάμμα. Η X έχει παράμετρο σχήματος α_2 ενώ η Y έχει παράμετρο σχήματος α_1. Η παράμετρος κλίμακος είναι ίση με 1 και για τις δύο τυχαίες μεταβλητές. Επομένως, η από κοινού πυκνότητα των X και Y είναι $f(x,y) = \frac{x^{\alpha_2-1}}{\Gamma(\alpha_2)} e^{-x} \frac{y^{\alpha_1-1}}{\Gamma(\alpha_1)} e^{-y}$ και

$$\begin{aligned} f_Z(z) &= \int_0^{\infty} x \frac{x^{\alpha_2-1}}{\Gamma(\alpha_2)} e^{-x} \frac{(zx)^{\alpha_1-1}}{\Gamma(\alpha_1)} e^{-zx} dx \;=\; \frac{z^{\alpha_1-1}}{\Gamma(\alpha_1)\,\Gamma(\alpha_2)} \int_0^{\infty} x^{\alpha_1+\alpha_2-1} e^{-x(1+z)}\, dx \\ &= \frac{1}{\Gamma(\alpha_1)\,\Gamma(\alpha_2)} \frac{z^{\alpha_1-1}}{(1+z)^{\alpha_1+\alpha_2}} \int_0^{\infty} u^{\alpha_1+\alpha_2-1} e^{-u} du \quad \textit{(με την αντικατάσταση } u = x(1+z)\textit{)} \\ &= \frac{1}{B(\alpha_1,\alpha_2)} \frac{z^{\alpha_1-1}}{(1+z)^{\alpha_1+\alpha_2}}, \qquad z \geq 0. \end{aligned} \tag{1.2}$$

Η πυκνότητα πιθανότητας (1.2) ονομάζεται πυκνότητα Βήτα του *δευτέρου είδους* και η σχέση της με την συνηθισμένη πυκνότητα Βήτα είναι η εξής: Αν T είναι μια τυχαία μεταβλητή Βήτα με πυκνότητα πιθανότητας

$$f_T(t) \;=\; \frac{1}{B(\alpha_1,\alpha_2)} t^{\alpha_1-1}(1-t)^{\alpha_2-1}, \quad t \in (0,1),$$

και αν $W := \frac{T}{1-T}$ τότε η τυχαία μεταβλητή W έχει πυκνότητα πιθανότητας που δίνεται από την (1.2). Πράγματι, για $w \geq 0$,

$$F_W(w) = P\left(\frac{T}{1-T} \leq w\right) = P\left(T \leq \frac{w}{1+w}\right) = F_T\left(\frac{w}{1+w}\right).$$

Συνεπώς, η πυκνότητα πιθανότητα της W είναι

$$\begin{aligned} f_W(w) &= F_T'\left(\frac{w}{1+w}\right) \frac{1}{(1+w)^2} = \frac{1}{B(\alpha_1,\alpha_2)} \left(\frac{w}{1+w}\right)^{\alpha_1-1} \left(\frac{1}{1+w}\right)^{\alpha_2-1} \frac{1}{(1+w)^2} \\ &= \frac{1}{B(\alpha_1,\alpha_2)} \frac{w^{\alpha_1-1}}{(1+w)^{\alpha_1+\alpha_2}}, \quad w \geq 0. \end{aligned}$$

Αυτή είναι η πυκνότητα (1.2).

1.2 Η κατανομή χ^2_n

Έστω Z_i, $i = 1, \ldots, n$ τυποποιημένες κανονικές τυχαίες μεταβλητές. Η τυχαία μεταβλητή

$$X_n \;:=\; Z_1^2 + Z_2^2 + \cdots + Z_n^2 \tag{1.3}$$

έχει κατανομή χ^2_n με n βαθμούς ελευθερίας. Η τυχαία μεταβλητή Z^2 έχει συνάρτηση κατανομής

$$F_{Z^2}(x) = P(Z^2 \le x) = P(-\sqrt{x} \le Z \le \sqrt{x}) = \frac{1}{2\pi}\int_{-\sqrt{x}}^{\sqrt{x}} e^{-\frac{1}{2}u^2}du, \quad x \ge 0.$$

Συνεπώς, η πυκνότητα πιθανότητας της τυχαίας μεταβλητής Z^2 είναι

$$f_{Z^2}(x) \;=\; \frac{1}{2x^{1/2}}e^{-\frac{1}{2}(x^{1/2})^2} - \frac{-1}{2x^{1/2}}e^{-\frac{1}{2}(-x^{1/2})^2} \;=\; \frac{1}{2}\,\frac{1}{\Gamma(1/2)}\,\left(\frac{x}{2}\right)^{-1/2} e^{-\frac{x}{2}},\;\; x > 0,$$

λαμβάμνοντας υπ΄ όψιν ότι $\Gamma(1/2) = \sqrt{\pi}$. Επομένως βλέπουμε ότι η τυχαία μεταβλητή Z^2 έχει κατανομή Gamma(1/2,1/2). Κατά συνέπεια η τυχαία μεταβλητή X_n έχει κατανομή Gamma(1/2,n/2) με πυκνότητα πιθανότητας

$$f_{\chi^2_n}(x) \;=\; \frac{1}{2}\,\frac{1}{\Gamma\left(\frac{n}{2}\right)}\,\left(\frac{x}{2}\right)^{\frac{n}{2}-1} e^{-\frac{x}{2}},\;\; x > 0. \tag{1.4}$$

Η πυκνότητα πιθανότητας (1.4) ονομάζεται χ^2 με n βαθμούς ελευθερίας.

1.3 Αλλαγή μεταβλητών

Έστω U μια περιοχή του $\mathbb{R}^2$ και $\mathbf{g} : U \to \mathbb{R}^2$ μια συνάρτηση συνεχώς παραγωγίσιμη και $\mathbf{X} = (X_1, X_2)$ ένα ζεύγος τυχαίων μεταβλητών με από κοινού πυκνότητα πιθανότητας $f_X(x_1, x_2)$ ορισμένες στο U. Αν $g(\mathbf{x}) = g(x_1, x_2) = \begin{pmatrix} g_1(\mathbf{x}) \\ g_2(\mathbf{x}) \end{pmatrix}$ ορίζουμε το διάνυσμα τυχαίων μεταβλητών $\mathbf{Y} = \mathbf{g}(\mathbf{X})$ δηλαδή

$$\begin{aligned} Y_1 &= g_1(X_1, X_2) \\ Y_2 &= g_2(X_1, X_2) \end{aligned}$$

Το ζητούμενο είναι η από κοινού πυκνότητα πιθανότητας $f_Y(y_1, y_2)$ των τυχαίων μεταβλητών (Y_1, Y_2).

Υποθέτουμε αρχικά ότι η συνάρτηση $\mathbf{g}$ είναι 1- προς -1. Εξετάζουμε την καμπύλη $\{(x_1, x_2) : g_1(x_1, x_2) = y_1\}$ και την καμπύλη $\{(x_1, x_2) : g_1(x_1, x_2) = y_1 + \Delta y_1\}$. Παρομοίως τις καμπλυλες $\{(x_1, x_2) : g_2(x_1, x_2) = y_2\}$ και την καμπύλη $\{(x_1, x_2) : g_2(x_1, x_2) = y_2 + \Delta y_2\}$.

Κανόνας για την από κοινού πυκνότητα: Έστω

$$J := \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix}. \tag{1.5}$$

(Συχνά θα συμβολίζουμε συνοπτικά την Ιακωβιανή ορίζουσα J (1.5) ως $\frac{\partial(x_1,x_2)}{\partial(y_1,y_2)}$. Τότε

$$f_Y(y_1,y_2) \;=\; f_X(x_1,x_2)\,|\,J\,|\,. \tag{1.6}$$

Για την εφαρμογή του παραπάνω τύπου, θεωρούμε βεβαίως τα x_1, x_2, στο δεξί μέλος της (1.6) ως συναρτήσεις των y_1, y_2. Τα ακόλουθα παραδείγματα δείχνουν πώς εφαρμόζεται αυτή η μέθοδος.

1.3.1 Παραδείγματα.

Παράδειγμα 1.

Έστω $y_1 = x_1 + x_2$, $y_2 = x_1 - x_2$. Υποθέτουμε ότι οι τυχαίες μεταβλητές X_1, X_2, είναι ανεξάρτητες, τυποποιημένες κανονικές, και συνεπώς $f_X(x_1,x_2) = \frac{1}{2\pi}\, e^{-\frac{1}{2}(x_1^2+x_2^2)}$, $x_1, x_2 \in \mathbb{R}$. Ισχύει ότι

$$\begin{aligned} x_1 + x_2 &= y_1 \\ x_1 - x_2 &= y_2 \end{aligned} \qquad \Longleftrightarrow \qquad \begin{aligned} x_1 &= \tfrac{1}{2}\,(y_1 + y_2) \\ x_2 &= \tfrac{1}{2}\,(y_1 - y_2) \end{aligned}$$

και, επομένως, από την (1.5),

$$J \;=\; \begin{vmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{vmatrix} \;==-\frac{1}{2}.$$

Συνεπώς, από την (1.6)

$$f_Y(y_1,y_2) \;=\; \frac{1}{2\pi}\, e^{-\frac{1}{2}\left(\frac{(y_1+y_2)^2}{4}+\frac{(y_1-y_2)^2}{4}\right)}\,\frac{1}{2} \;=\; \frac{1}{4\pi}\, e^{-\frac{y_1^2}{4}-\frac{y_2^2}{4}}, \qquad y_1, y_2 \in \mathbb{R}.$$

Επομένως οι τυχαίες μεταβλητές $Y_1 = X_1 + X_2$, $Y_2 = X_1 - X_2$, είναι από κοινού κανονικά κατανεμημένες, ανεξάρτητες, με μέσο ο και διασπορά 2.

Παράδειγμα 2.

Έστω ότι οι X_1, X_2, είναι ανεξάρτητες με κατανομή Γάμμα με παραμέτρους (α_1, λ) και (α_2, λ) αντίστοιχα. Συνεπώς

$$f_X(x_1,x_2) \;=\; \lambda^{\alpha_1+\alpha_2}\,\frac{x_1^{\alpha_1-1}x_2^{\alpha_2-1}}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\, e^{-\lambda(x_1+x_2)}\,\mathbf{1}(x_1>0, x_2>0).$$

Ορίζουμε τώρω τις τυχαίες μεταβλητές Y_1, Y_2, ως

$$\begin{array}{rcl} X_1+X_2 & = & Y_1 \\ \frac{X_1}{X_1+X_2} & = & Y_2 \end{array} \quad \Longleftrightarrow \quad \begin{array}{rcl} X_1 & = & Y_1 Y_2 \\ X_2 & = & Y_1(1-Y_2) \end{array}$$

Η Ιακωβιανή είναι

$$J = \begin{vmatrix} y_2 & y_1 \\ 1-y_2 & -y_1 \end{vmatrix} = -y_1.$$

Συνεπώς

$$\begin{aligned} f_Y(y_1,y_2) &= \lambda^{\alpha_1+\alpha_2} \frac{(y_1y_2)^{\alpha_1-1}(y_1-y_1y_2)^{\alpha_2-1}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} e^{-\lambda(y_1y_2+y_1(1-y_2))} y_2 \mathbf{1}(y_1>0, 0<y_2<1) \\ &= \left(\lambda^{\alpha_1+\alpha_2} \frac{y_1^{\alpha_1+\alpha_2-1}}{\Gamma(\alpha_1+\alpha_2)} e^{\lambda y_1}, \mathbf{1}(y_1>0)\right) \left(\frac{y_2^{\alpha_1-1}(1-y_2)^{\alpha_2-1}}{B(\alpha_1,\alpha_2)} \mathbf{1}(0<y_2<1)\right). \end{aligned}$$

Στην τελευταία σχέση πολλαπλασιάσαμε και διαιρέσαμε με $\Gamma(\alpha_1+\alpha_2)$ και χρησιμοποιήσαμε την σχέση $B(\alpha_1,\alpha_2) = \frac{\Gamma(\alpha_1)\Gamma(\alpha_2)}{\Gamma(\alpha_1+\alpha_2)}$. Παρατηρείστε ότι η από κοινού πυκνότητα πιθανότητας για τις τυχαίες μεταβλητές Y_1, Y_2, είναι το γινόμενο των δύο περιθωρίων. Συμπεραίνουμε ότι οι Y_1 και Y_2 είναι ανεξάρτητες. Η Y_1 έχει κατανομή Γάμμα $(\alpha_1+\alpha_2,\lambda)$ και η Y_2 έχει κατανομή Βήτα με παραμέτρους (α_1,α_2).

Παράδειγμα 3.

Έστω ότι οι X, Y, τυχαίες μεταβλητές με από κοινού κατανομή

$$f(x,y) = \frac{1}{y} e^{-x/y} e^{-y}, \quad x,y>0$$

Δείξτε ότι οι τυχαίες μεταβλητές $V = X/Y$ και $W = Y$, είναι ανεξάρτητες και βρείτε τις περιθώριες κατανομές τους.

Ισχύτει ότι

$$\begin{array}{rcl} v & = & x/y \\ w & = & y \end{array} \quad \Longleftrightarrow \quad \begin{array}{rcl} x & = & vw \\ y & = & w \end{array}$$

και

$$J = \frac{\partial(x,y)}{\partial(v,w)} = \begin{vmatrix} w & v \\ 0 & 1 \end{vmatrix} = w.$$

Συνεπώς,

$$f_{VW}(v,w) = f(vw,w)w = \frac{1}{w} e^{-v} e^{-w} w \mathbf{1}(v>0,w>0) = \left(e^{-v}\mathbf{1}(v>0)\right)\left(e^{-w}\mathbf{1}(w>0)\right)$$

και επομένως βλέπουμε ότι οι τυχαίες μεταβλητές V, W είναι ανεξάρτητες, εκθετικές κατανεμημένες με ρυθμό 1.

Παράδειγμα 4.

Έστω ότι οι X, Y, τυχαίες μεταβλητές με από κοινού κατανομή

$$f(x,y) \ = \ \frac{x}{(1+x)^2(1+xy)^2}, \qquad x,y > 0$$

Δείξτε ότι οι τυχαίες μεταβλητές $V = X$ και $W = XY$, είναι ανεξάρτητες και βρείτε τις περιθώριες κατανομές τους.

Ισχύτει ότι

$$\begin{array}{rcl} v & = & x \\ w & = & xy \end{array} \qquad \Longleftrightarrow \qquad \begin{array}{rcl} x & = & v \\ y & = & w/v \end{array}$$

και

$$J \ = \ \frac{\partial(x,y)}{\partial(v,w)} \ = \ \begin{vmatrix} 1 & 0 \\ -w/v^2 & 1/v \end{vmatrix} \ = \ \frac{1}{v}.$$

Συνεπώς,

$$\begin{aligned} f_{VW}(v,w) \ &= \ f(v,w/v)\frac{1}{v} \ = \ \frac{1}{(1+v)^2(1+w)^2}\,\mathbf{1}(w>0,v>0) \\ &= \ \left(\frac{1}{(1+v)^2}\mathbf{1}(v>0)\right)\left(\frac{1}{(1+w)^2}\mathbf{1}(w>0)\right) \end{aligned}$$

και επομένως βλέπουμε ότι οι τυχαίες μεταβλητές V, W είναι ανεξάρτητες. Η πυκνότητα της V π.χ. είναι $\frac{1}{(1+v)^2}$ για $v > 0$. Η πυκνότητα αυτή ανήκει στην οικογένεια των κατανομών Pareto. Είναι ευκολο να διαπιστώσει κανείς ότι $\int_0^\infty (1+v)^{-2}dv = 1$.

Παράδειγμα 5.

Έστω ότι οι X, Y, τυχαίες μεταβλητές με από κοινού κατανομή

$$f(x,y) \ = \ \lambda^2 e^{-\lambda y}\,\mathbf{1}(0 < x < y).$$

Δείξτε ότι οι τυχαίες μεταβλητές $V = \frac{X}{Y-X}$ και $W = Y$, είναι ανεξάρτητες και βρείτε τις περιθώριες κατανομές τους.

Έχουμε τον μετασχηματισμό

$$\begin{array}{rcl} v & = & \frac{x}{y-x} \\ w & = & y \end{array} \qquad \Longleftrightarrow \qquad \begin{array}{rcl} x & = & \frac{vw}{1+v} \\ y & = & w \end{array}$$

και

$$J \ = \ \frac{\partial(x,y)}{\partial(v,w)} \ = \ \begin{vmatrix} \frac{w}{1+v} - \frac{vw}{(1+v)^2} & \frac{v}{1+v} \\ 0 & 1 \end{vmatrix} \ = \ \frac{w}{(1+v)^2}.$$

Συνεπώς,

$$\begin{aligned} f_{VW}(\nu, w) &= f(\nu, w/\nu)\frac{w}{(1+\nu)^2} = \lambda^2 e^{-\lambda w}\frac{w}{(1+\nu)^2}\,\mathbf{1}\left(0 < \frac{\nu w}{1+\nu} < w\right) \\ &= \left(\lambda^2 w e^{-\lambda w}\mathbf{1}(w>0)\right)\left(\frac{1}{(1+\nu)^2}\mathbf{1}(\nu>0)\right). \end{aligned}$$

Στο τελευταίο βήμα χρησιμοποιήσαμε το γεγονός ότι η διπλή ανισότητα $0 < \frac{\nu w}{1+\nu} < w$ είναι ισοδύναμη με τις

$$w > 0 \quad \text{και} \quad 0 < \frac{\nu}{1+\nu} < 1 \Leftrightarrow \nu > 0.$$

Επομένως βλέπουμε ότι οι τυχαίες μεταβλητές V, W είναι ανεξάρτητες. Η πυκνότητα της V είναι Pareto ενώ εκεινη της W (με άλλα λόγια η περιθώρια της πυκνότητα της Y) είναι Γάμμα $(2, \lambda)$.

1.4 Πολικές Συντεταγμένες

Οι σχέση ανάμεσα στις πολικές και καρτεσιανές συντεταγμένες δίνεται από την

$$\begin{aligned} x &= r\cos\theta \\ y &= r\sin\theta \end{aligned}$$

Αν X, Y είναι τυχαίες μεταβλητές με από κοινού πυκνότητα πιθανότητας $f(x, y)$ και αν R, Θ είναι οι τυχαίες μεταβλητές που προκύπτουν από τον μετασχηματισμό (1.7) τότε η από κοινού πυκνότητα των R, Θ, δίνεται από την σχέση

$$f_{R,\Theta}(r, \theta) := f(r\cos\theta, r\sin\theta)r. \tag{1.7}$$

Ο παράγων r στο δεξί μέλος της (1.7) προκύπτει από τον ακόλουθο υπολογισμό:

$$\det(J) = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r\cos^2\theta + r\sin^2\theta = r. \tag{1.8}$$

Αν X, Y είναι ανεξάρτητες τυποποιημένες κανονικές τυχαίες μεταβλητές τότε

$$f(x, y) = \frac{1}{2\pi}\, e^{-\frac{1}{2}\left(x^2+y^2\right)}$$

και οι τυχαίες μεταβλητές που προκύπτουν από τον μετασχηματισμό (1.7) R, Θ έχουν κατανομή

$$f_{R,\Theta}(r, \theta) = \frac{1}{2\pi}\, e^{-\frac{1}{2}\left(r^2\cos^2\theta + r^2\sin^2\theta\right)}\, r = r\, e^{-\frac{1}{2}r^2}\,\frac{1}{2\pi}. \tag{1.9}$$

Η ανωτέρω από κοινού πυκνότητα πιθανότητας είνα το γινόμενο των δύο περιθωρίων, της πυκνότητας της Θ που όπως βλέπουμε είναι ομοιόμορφα κατανεμημένη στο διάστημα $[0, 2\pi)$, και της περιθώριας πυκνότητας της R:

$$f_\Theta(\theta) = \frac{1}{2\pi}\, \mathbf{1}(0 \le \theta < 2\pi), \tag{1.10}$$

$$f_R(r) = re^{-\frac{1}{2}r^2},\ \mathbf{1}(r > 0). \tag{1.11}$$

Η περιθώρια κατανομή της R ονομάζεται κατανομή Rayleigh. Είναι αξιοσημείωτο ότι, όταν οι X, Y είναι ανεξάρτητες, από κοινού κανονικά κατανεμημένες, τότε η απόσταση του τυχαίου σημείου (X, Y) από την αρχή των αξόνων, R, και η γωνία Θ είναι *ανεξάρτητες*.

Κεφάλαιο 2

Το Κεντρικό Οριακό Θεώρημα

Θεώρημα 2.1. *Έστω* $X_1, X_2, X_3, \ldots,$ *μια ακολουθία από ανεξάρτητες, ισόνομες τυχαίες μεταβλητές με μέσο* μ *και διασπορά* $\sigma^2 < \infty$. *Τότε*

$$\lim_{n\to\infty} P\left(\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \leq x\right) \;=\; \Phi(x) \qquad \forall x \in \mathbb{R} \tag{2.1}$$

όπου

$$\Phi(x) \;=\; \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}}\, du\,. \tag{2.2}$$

Ορισμός 2.1 (Σύγκλιση κατά κατανομή). Έστω μια ακολουθία τυχαίων μεταβλητών $\{X_n\}$ ορισμένη πάνω στον χώρο πιθανοτήτων $(\Omega, \mathscr{F}, P)$ και $\{F_n\}$ η αντίστοιχη ακολουθία συναρτήσεων κατανομής $F_n(x) := P(X \leq x)$, $x \in \mathbb{R}$. Έστω επίσης X τυχαία μεταβλητή στον $(\Omega, \mathscr{F}, P)$ με συνάρτηση κατανομής $F(x) = P(X \leq x)$. Η ακολουθία $\{X_n\}$ *συγκλίνει κατά κατανομή στην* X αν,

$$\lim_{n\to\infty} F_n(x) \qquad \text{για κάθε } x \text{ που είναι σημείο συνέχειας της } F. \tag{2.3}$$

Προκειμένου να δηλώσουμε ότι η ακολουθία $\{X_n\}$ συγκλίνει κατά κατανομή γράφουμε $X_n \xrightarrow{d} X$. Επίσης παρατηρούμε ότι αν η κατανομή της οριακής τυχαίας μεταβλητής X, δηλαδή η F είναι συνεχής για κάθε $x \in \mathbb{R}$ (όπως συμβαίνει για παράδειγμα στην περίπτωση της σύγκλισης στην τυποποιημένη κανονική κατανομή, $F(x) = \Phi(x)$)

Αν θέσουμε $S_n := X_1 + X_2 + \cdots + X_n$ όπου $\{X_i\}$, $i = 1, 2, \ldots$ είναι ανεξάρτητες ισόνομες τ.μ. με μέσο μ και διασπορά $\sigma^2 < \infty$, και αν ορίσουμε την ακολουθία τυχαίων μεταβλητών

$$Z_n \;:=\; \frac{S_n - n\mu}{\sigma\sqrt{n}}, \qquad n = 1, 2, \ldots$$

τότε, το κεντρικό οριακό θεώρημα λέει ότι

$$Z_n \xrightarrow{d} Z$$

όπου Z είναι μια τυποποιημένη κανονική τυχαία μεταβλητή. Ισοδύναμα, το κεντρικό οριακό θεώρημα λέει ότι, αν

$$F_n(x) := P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x\right)$$

τότε $F_n(x) \to \Phi(x)$ για κάθε $x \in \mathbb{R}$, όπου Φ η συνάρτηση κατανομής της τυποποιημένης κανονικής τυχαίας μεταβλητής (2.2).

Στην πράξη, το κεντρικό οριακό θεώρημα συνεπάγεται ότι

$$X_1 + X_2 + \cdots + X_n \approx \mathcal{N}(n\mu, n\sigma^2),$$

Παραδείγμα: Έστω $X_n = \xi_1 + \xi_2 + \cdots + \xi_n$ όπου ξ_i έχουν εκθετική κατανομή με ρυθμό λ (ή, ισοδύναμα, $\Gamma(1, \lambda)$). Τότε η τυχαία μεταβλητή X_n έχει κατανομή $\Gamma(n, \lambda)$. Ισχύει ότι, προσεγγιστικά, η κατανομή της X_n είναι $\mathcal{N}\left(\frac{n}{\lambda}, \frac{n}{\lambda^2}\right)$.

Κεφάλαιο 3

Διατεταγμένα Δείγματα

3.1 Η από κοινού πυκνότητα του διατεταγμένου δείγματος n ανεξάρτητων, ισόνομων τυχαίων μεταβλητών

Έστω $X_1, X_2, \ldots, X_n$ ανεξάρτητες, ισόνομες τυχαίες μεταβλητές με συνάρτηση κατανομής $F(x) = P(X_1 \leq x)$, και πυκνότητα πιθανότητας $f(x) = F'(x)$, $x \in \mathbb{R}$. Ονομάζουμε *διατεταγμένο δείγμα* την διάταξη κατ΄ αύξουσα τιμή, $X_{(1)} < X_{(2)}, \ldots, X_{(n)}$. Έτσι, $X_{(1)}$ είναι η μικρότερη από τις $X_1, \ldots, X_n$, $X_{(2)}$ είναι η δεύτερη μικρότερη, κ.ο.κ. $X_{(n)}$ είναι βεβαίως η μεγαλύτερη.

Η από κοινού πυκνότητα των διατεταγμένων τ.μ. $X_{(1)}, \ldots, X_{(n)}$ είναι

$$f_n(x_1, x_2, \ldots, x_n) = n! f(x_1) f(x_2) \cdots f(x_n) \mathbf{1}(x_1 < x_2 < \cdots < x_n). \tag{3.1}$$

Πριν εξετάσουμε τις συνέπειες του αποτελέσματος αυτού, ας δούμε πρώτα συγκεκριμένες περιπτώσεις:

3.2 Το μέγιστο και το ελάχιστο n τυχαίων μεταβλητών

Θα προσδιορίσουμε την κατανομή της τυχαίας μεταβλητής

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

Αν θέσουμε $F_M(x) := P(X_{(n)} \leq x)$ τότε

$$\begin{aligned} F_M(x) &= P(\max(X_1, X_2, \ldots, X_n) \leq x) = P(X_1 \leq x, X_2 \leq x, \cdot, X_n \leq x) = \prod_{i=1}^{n} P(X_i \leq x) \\ &= F(x)^n. \end{aligned} \tag{3.2}$$

Η αντίστοιχη πυκνότητα πιθανότητας είναι

$$f_M(x) = nf(x)F(x)^{n-1}. \tag{3.3}$$

Η κατανομή του ελάχιστου, $X_{(1)}$, $F_m(x) := P(X_{(1)} \leq x)$ μπορεί να βρεθεί με το εξής επιχείρημα:

$$\begin{aligned} 1 - F_m(x) &= P(\min(X_1, X_2, \ldots, X_n) > x) \;=\; P(X_1 > x, X_2 > x, \ldots, X_n > x) \\ &= \prod_{i=1}^{n} P(X_i > x) \;=\; \prod_{i=1}^{n} (1 - F(x)) \;=\; (1 - F(x))^n \end{aligned}$$

συνεπώς

$$F_m(x) \;=\; 1 - (1 - F(x))^n \tag{3.4}$$

και η αντίστοιχη πυκνότητα πιθανότητας είναι

$$f_m(x) \;=\; nf(x)\,(1 - F(x))^{n-1}\;. \tag{3.5}$$

Παράδειγμα 1. Αν οι τυχαίες μεταβλητές X_i, $i = 1, \ldots, n$ είναι ομοιόμορφα κατανεμημένες στο διάστημα $[0, 1]$, τότε

$$F_M(x) = \begin{cases} x^n & \text{αν } 0 \leq x \leq 1 \\ 0 & \text{διαφορετικά.} \end{cases}$$

Αντίστοιχα, για την κατανομή του ελαχίστου, έχουμε

$$F_m(x) = \begin{cases} (1 - x)^n & \text{αν } 0 \leq x \leq 1 \\ 0 & \text{διαφορετικά.} \end{cases}$$

Παράδειγμα 2. Αν οι τυχαίες μεταβλητές X_i είναι εκθετικά κατανεμημένες με ρυθμό λ,

$$F_M(x) \;=\; \left(1 - e^{-\lambda x}\right)^n, \;\; x \geq 0, \qquad f_M(x) \;=\; n\lambda e^{-\lambda x}\left(1 - e^{-\lambda x}\right)^{n-1}, \;\; x \geq 0. \tag{3.6}$$

Η κατανομή του ελαχίστου n ανεξάρτητων τυχαίων μεταβλητών με κοινή παράμετρο λ είναι

$$F_m(x) \;=\; e^{-n\lambda x}, \;\; x \geq 0, \qquad f_m(x) \;=\; n\lambda e^{-n\lambda x}, \;\; x \geq 0. \tag{3.7}$$

Το τελευταίο αυτό αποτέλεσμα είναι πολύ ενδιαφέρον επειδή, όπως φαίνεται από την έκφραση (3.7), η κατανομή του ελαχίστου $X_{(1)}$ είναι *εκθετική* με ρυθμό $n\lambda$.

3.3 Η από κοινού κατανομή του ελαχίστου και του μεγίστου και το εύρος

Έστω $f(x,y)$ η από κοινού πυκνότητα των τυχαίων μεταβλητών $X_{(1)}$, $X_{(n)}$. Για $x<y$ θα ισχύει ότι

$$\begin{aligned} f(x,y)dxdy &= P(X_i \in [x,x+dx] \text{ για κάποιο } i \in \{1,\dots,n\}\ )\, P(X_j \in [y,y+dy] \text{ για κάποιο } j \neq i\ ) \\ &\quad \times P(x < X_k < y \text{ για όλα τα } k \text{ διάφορα των } i \text{ και } j\ ). \end{aligned}$$

Δεδομένου ότι έχουμε n επιλογές για τον δείκτη i (ποια μεταβλητή θα είναι το ελάχιστο δηλαδή) και $n-1$ επιλογές για το j (ποια μεταβλητή θα είναι το μέγιστο) από την προηγούμενη σχέση παίρνουμε

$$f_{m,M}(x,y) \;=\; n(n-1)f(x)f(y)\left(F(y)-F(x)\right)^{n-2}, \quad x<y. \tag{3.8}$$

Από την σχέση αυτή μπορούμε για παράδειγμα να υπολογίσουμε την δεσμευμένη πυκνότητα του μεγίστου δεδομένου του ελαχίστου.

$$\begin{aligned} f_{M|m}(y|x) &= \frac{f_{m,M}(x,y)}{f_m(x)} = \frac{n(n-1)f(x)f(y)\left(F(y)-F(x)\right)^{n-2}}{nf(x)\left(1-F(x)\right)^{n-1}}\mathbf{1}(x<y\leq 1) \\ &= \frac{(n-1)f(y)\left(F(y)-F(x)\right)^{n-2}}{\left(1-F(x)\right)^{n-1}}\mathbf{1}(x<y\leq 1). \end{aligned}$$

Στην περίπτωση που οι τυχαίες μεταβλητές X_i είναι ομοιόμορφα κατανεμημένες στο $[0,1]$,

$$f_{M|m}(y|x) \;=\; \frac{(n-1)\,(y-x)^{n-2}}{(1-x)^{n-1}}.$$

Συγκεκριμένα, η μέση τιμή του μεγίστου, δεδομένου του ελαχίστου, στην περίπτωση αυτή είναι

$$\begin{aligned} E[X_{(n)}|X_{(1)}=x] &= \int_x^1 y f_{M|m}(y|x)dy \;=\; \int_x^1 y\,\frac{(n-1)\,(y-x)^{n-2}}{(1-x)^{n-1}}dy \\ &= \int_x^1 (y-x)\,\frac{(n-1)\,(y-x)^{n-2}}{(1-x)^{n-1}}dy \;+\; x\int_x^1 y\,\frac{(n-1)\,(y-x)^{n-2}}{(1-x)^{n-1}}dy \\ &= \frac{n-1}{(1-x)^{n-1}}\int_x^1 (y-x)^{n-1}dy \;+\; x \\ &= \frac{n-1}{(1-x)^{n-1}}\int_0^{1-x} u^{n-1}du \;+\; x \;=\; \frac{n-1}{n}(1-x)+x \\ &= 1-\frac{1-x}{n}. \end{aligned} \tag{3.9}$$

3.4 Η κατανομή του εύρους

Ώς εύρος ορίζουμε την διαφορά $R := X_{(n-1)} - X_{(1)}$. Η πυκνότητα πιθανότητάς της προσδιορίζεται εύκολα από την από κοινού πυκνότητα πιθανότητας (3.8) ως

$$f_R(r) \;=\; \int_{\mathbb{R}} f_{m,M}(x, x+r)dx \;=\; n(n-1)\int_{\mathbb{R}} f(x)f(x+r)\left(F(x+r)-F(x)\right)^{n-2} dx. \tag{3.10}$$

Η έκφραση αυτή για την πυκνότητα του εύρους παίρνει πολύ απλούστερη μορφή σε ειδικές περιπτώσεις:

Οι τυχαίες μεταβλητές X_i είναι ομοιόμορφα κατανεμημένες Στην περίπτωση αυτή $f(x) = \mathbf{1}(0 < x < 1)$, $f(x+r) = \mathbf{1}(0 < x+r < 1) = \mathbf{1}(0 < x < 1-r)$, και επομένως

$$f_R(r) \;=\; \int_0^{1-r} n(n-1)[(x+r)-x]^{n-2}dx \;=\; n(n-1)(1-r)r^{n-2}, \quad 0 < r < 1. \tag{3.11}$$

Οι τυχαίες μεταβλητές X_i είναι εκθετικά κατανεμημένες Στην περίπτωση αυτή η (3.10) δίνει

$$\begin{aligned} f_R(r) &= n(n-1)\int_0^{\infty} \lambda^2 e^{-\lambda x}e^{-\lambda(x+r)}\left(e^{-\lambda x} - e^{-\lambda(x+r)}\right)^{n-2} dx \\ &= n(n-1)\lambda^2 e^{-\lambda r}\left(1-e^{-\lambda r}\right)^{n-2}\int_0^{\infty} e^{-\lambda n x}dx \\ &= (n-1)\lambda e^{-\lambda r}\left(1-e^{-\lambda r}\right)^{n-2}. \end{aligned} \tag{3.12}$$

3.5 Η κατανομή της υπ΄ αριθμόν k μεταβλητής στην σειρά διάταξης.

Εδώ εξετάζουμε την πυκνότητα πιθανότητας της $X_{(k)}$ δηλαδή της μεταβλητής που σε αύξουσα σειρά έχει την θέση k. Η πυκνότητα πιθανότητας $f_{(k)}(x)dx$ θα είναι η πιθανότητα ένα από τα X να πέσει στο διάστημα $[x, x+dx]$, $k-1$ από τα X είναι μικρότερα από x και τα υπόλοιπα $n-k$ από τα X είναι μεγαλύτερα από το x. Μπορούμε να διαλέξουμε με n τρόπους το X που ανήκει στο $[x, x+dx]$ και, αφού κάνουμε αυτή την επιλογή, να διαλέξουμε με $\binom{n-1}{k-1}$ τρόπους τα X που είναι μικρότερα από x. Μετά από αυτό δεν μένει άλλη επιλογή. Τα υπόλοιπα $n-k$ X που μένουν θα είναι εκείνα που είναι μεγαλύτερα από x. Συνεπώς,

$$f_{(k)}(x) \;=\; n\binom{n-1}{k-1}f(x)F(x)^{k-1}(1-F(x))^{n-k}. \tag{3.13}$$

Όταν τα X_i είναι ομοιόμορφα κατανεμημένα στο $[0,1]$ τότε

$$\begin{aligned} f_{(k)}(x) &= \frac{n!}{(k-1)!(n-k)!} x^{k-1}(1-x)^{n-k}\,\mathbf{1}(0<x<1) \\ &= \frac{\Gamma(n+1)}{\Gamma(k)\Gamma(n-k+1)} x^{k-1}(1-x)^{n-k}\,\mathbf{1}(0<x<1) \\ &= \frac{x^{k-1}(1-x)^{n-k}}{B(k,n-k+1)}\,\mathbf{1}(0<x<1). \end{aligned} \tag{3.14}$$

3.6 Προβλήματα

Πρόβλημα 1. Έστω X_1, X_2, X_3 ανεξάρτητες τυχαίες μεταβλητές με ρυθμό 1. Να βρεθούν α) η από κοινού πυκνότητα των $X_{(1)}, X_{(2)}, X_{(3)}$, οι περιθώριες πυκνότητες των $X_{(k)}$, $k = 1,2,3$, η από κοινού πυκνότητα των $X_{(1)}, X_{(3)}$, και η πυκνότητα του εύρους $R = X_{(3)} - X_{(1)}$.

Λύση: Από την (3.1) έχουμε

$$f(x_1,x_2,x_3) = 6e^{-(x_1+x_2+x_3)}\,\mathbf{1}(0<x_1<x_2<x_3).$$

Η από κοινού πυκνότητα του μεγίστου και του ελαχίστου προκύπτει ολοκληρώνοντας ως προς x_2:

$$\begin{aligned} f_{1,3}(x_1,x_3) &= \int_0^\infty f(x_1,x_2,x_3)dx_2 = 6\int_{x_1}^{x_3} e^{-x_2}dx_2\, e^{-x_1-x_3}\mathbf{1}(0<x_1<x_3) \\ &= 6e^{-x_1-x_3}\left(e^{-x_1}-e^{-x_3}\right)\mathbf{1}(0<x_1<x_3) \end{aligned}$$

Η κατανομή του εύρους είναι

$$f_R(r) = \int_0^\infty 6e^{-2x_1-r}\left(e^{-x_1}-e^{-x_1-r}\right)dx_1 = \int_0^\infty 6e^{-3x_1-r}\left(1-e^{-r}\right)dx_1 = 2e^{-r}\left(1-e^{-r}\right)$$

Η μέση τιμή του εύρους είναι

$$\begin{aligned} \int_0^\infty r f_R(r)dr &= \int_0^r 2re^{-r}\left(1-e^{-r}\right)dr = 2\int_0^\infty re^{-r}dr - 2\int_0^\infty re^{-2r}dr \\ &= 2-\frac{1}{2} = \frac{3}{2}. \end{aligned}$$

Τέλος, οι περιθώριες πυκνότητες των $X_{(1)}, X_{(2)}, X_{(3)}$ είναι

$$f_1(x_1) = 3e^{-3x_1}, \quad f_2(x_2) = 6e^{-2x_2}\left(1-e^{-x_2}\right), \quad f_3(x_3) = 3e^{-x_3}\left(1-e^{-x_2}\right)^2.$$

Πρόβλημα 2. Έστω X_1, X_2, X_3 ανεξάρτητες τυχαίες μεταβλητές, ομοιόμορφα κατανεμημένες στο $[0,1]$. Να βρεθεί η από κοινού κατανομή των $Y_1 = X_{(1)}$, $Y_2 = X_{(2)} - X_{(1)}$, $Y_{(3)} = X_{(3)} - X_{(2)}$.

Λύση: Αν

$$\begin{array}{rcl} y_1 &=& x_1 \\ y_2 &=& x_2 - x_1 \\ y_3 &=& x_3 - x_2 \end{array} \quad \Longleftrightarrow \quad \begin{array}{rcl} x_1 &=& y_1 \\ x_2 &=& y_1 + y_2 \\ x_3 &=& y_1 + y_2 + y_3 \end{array}$$

η Ιακωβιανή του μετασχηματισμού είναι

$$J = \frac{\partial(x_1, x_2, x_3)}{\partial(y_1, y_2, y_3)} = \begin{vmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{vmatrix} = 1,$$

η από κοινού πυκνότητα των $X_{(1)}, X_{(2)}, X_{(3)}$ είναι $f(x_1, x_2, x_3) = 6\mathbf{1}(0 < x_1 < x_2 < x_3 < 1)$ και επομένως

$$f_Y(y_1, y_2, y_3) = 6\,\mathbf{1}(0 < y_1 < y_1 + y_2 < y_1 + y_2 + y_3 < 1) = 6\,\mathbf{1}(0 < y_1 + y_2 + y_3 < 1,\ 0 < y_i < 1, i = 1, 2, 3),$$

Με άλλα λόγια το τυχαίο σημείο (Y_1, Y_2, Y_3) είναι ομοιόμορφα κατανεμημένο στο σύνολο

$$\{(y_1, y_2, y_3);\ y_i \geq 0, i = 1, 2, 3,\ y_1 + y_2 + y_3 \leq 1\}.$$

Κεφάλαιο 4

Η Πολυμεταβλητή Κανονική Κατανομή

4.1 Γραμμικοί συνδιασμοί τυχαίων μεταβλητών

Έστω $X_1, X_2, \ldots, X_n$ τυχαίες μεταβλητές με από κοινού πυκνότητα πιθανότητας $f(x_1, \ldots, x_n)$, μέσους $\mu_i = EX_i$ και συνδιακυμάνσεις $V_{ij} = E[(X_i - \mu_i)(X_j - \mu_j)]$, $i = 1, \ldots, n$, $j = 1, \ldots, n$. Αν

$$Y_i = a_{i1}X_1 + a_{i2}X_2 + \cdots + a_{in}X_n, \quad i = 1, 2, \ldots, m$$

τότε

$$EY_i = E\sum_{j=1}^{n} a_{ij}X_j = \sum_{j=1}^{n} a_{ij}EX_j = \sum_{j=1}^{n} a_{ij}\mu_j \tag{4.1}$$

και

$$E[Y_iY_j] = E[\sum_{k=1}^{n} a_{ik}X_k \sum_{l=1}^{n} a_{jl}X_l] = \sum_{k=1}^{n}\sum_{l=1}^{n} a_{ik}a_{jl}E[X_kX_l] \tag{4.2}$$

Συνεπώς

$$\begin{aligned} \mathrm{Cov}(Y_i, Y_j) &= E[Y_iY_j] - E[Y_i]E[Y_j] = \sum_{k=1}^{n}\sum_{l=1}^{n} a_{ik}a_{jl}E[X_kX_l] - \left(\sum_{k=1}^{n} a_{ik}EX_k\right)\left(\sum_{l=1}^{n} a_{jl}EX_l\right) \\ &= \sum_{k=1}^{n}\sum_{l=1}^{n} a_{ik}a_{jl}\left(E[X_kX_l] - EX_kEX_l\right) = \sum_{k=1}^{n}\sum_{l=1}^{n} a_{ik}V_{ij}a_{jl}. \end{aligned} \tag{4.3}$$

Αν A είναι ο πίνακας $m \times n$

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

τότε η σχέση (4.1) γράφεται και ως $EY = AEX$, ενώ η (4.3) γράφεται και ως

$$[\mathrm{Cov}(Y_i, Y_j)] = AVA^T. \tag{4.4}$$

4.2 Εκτιμήτριες ελαχίστου τετραγωνικού σφάλματος

Έστω Y μια τυχαία μεταβλητή με γνωστή πυκνότητα πιθανότητας $f_Y(y)$. Ζητάμε να βρούμε μια σταθερά, έστω b, η οποία θα αποτελεί εκτιμήτρια για την άγνωστη τυχαία μεταβλητή και θα είναι βέλτιστη υπό την έννοια ότι θα έχει όσο το δυνατόν μικρότερο σφάλμα. Το σφάλμα εν προκειμένω υπολογίζεται ως το *μέσο τετραγωνικό σφάλμα* και δίδεται από τον τύπο

$$e(b) \;=\; \int_{\mathbb{R}} (y-b)^2 f_Y(y)dy. \tag{4.5}$$

Είναι ενδιαφέρον ότι η τιμή της σταθεράς b η οποία ελαχιστοποιεί το μέσο τετραγωνικό σφάλμα είναι η $b = \mu$.

$$\begin{aligned}
e(b) &= \int_{\mathbb{R}} (y-b)^2 f_Y(y)dy \;=\; \int_{\mathbb{R}} (y-\mu+\mu-b)^2 f_Y(y)dy \\
&= \int_{\mathbb{R}} \left[(y-\mu)^2 + 2(y-\mu)(\mu-b) + (\mu-b)^2\right] f_Y(y)dy \\
&= \int_{\mathbb{R}} (y-\mu)^2 f_Y(y)dy \;+\; 2(\mu-b)\int_{\mathbb{R}} (y-\mu)f_Y(y)dy \;+\; (\mu-b)^2 \int_{\mathbb{R}} f_Y(y)dy \\
&= \sigma_Y^2 + (b-\mu)^2.
\end{aligned}$$

(Η τελευταία σχέση, στην οποία σ_Y^2 συμβολίζει την διασπορά της Y προέκυψε λαμβάνοντας υπ΄ όψιν μας ότι $\int_{\mathbb{R}} f_Y(y)dy = 1$ και ότι $\int_{\mathbb{R}}(y-\mu)f_Y(y)dy = 0$.) Από την σχέση αυτή βλέπουμε ότι η ελάχιστη τιμή του $e(b)$ είναι σ_Y^2 και λαμβάνεται για $b = \mu$.

4.2.1 Η δεσμευμένη μέση τιμή ως εκτιμήτρια που ελαχιστοποιεί το μέσο τετραγωνικό σφάλμα

Έστω (X, Y) δύο τυχαίες μεταβλητές με δεδομένη από κοινού πυκνότητα πιθανότητας $f(x, y)$. Αναζητούμε μια εκτιμήτρια για την τιμή της Y αν γνωρίζουμε την τιμή της X. Συνεπώς, η εκτιμήτρια θα πρέπει να έχει την μορφή $g(X)$ όπου g είναι μια συνάρτηση που θέλουμε να προσδιορίσουμε έτσι ώστε να ελαχιστοποιείται το μέσο τετραγωνικό σφάλμα

$$e[g] \;=\; \int_{\mathbb{R}}\int_{\mathbb{R}} (y-g(x))^2 f(x,y)dxdy.$$

Αν χρησιμοποιήσουμε την σχέση $f(x, y) = f_{Y|X}(y|x)f_X(x)$, το μέσο τετραγωνικό σφάλμα γράφεται ως

$$e[g] \;=\; \int_{\mathbb{R}} \left(\int_{\mathbb{R}} (y-g(x))^2 f_{Y|X}(y|x)dy\right) f_X(x)dx.$$

Με βάση την προηγούμενη συζήτηση το μέσο τετραγωνικό σφάλμα ελαχιστοποιείται όταν

$$g(x) \;=\; E[Y|X=x]. \tag{4.6}$$

4.2.2 Η βέλτιστη γραμμική εκτιμήτρια

Ας εξετάσουμε τώρα γραμμικές εκτιμήτριες της μορφής $aX+b$ οι οποίες ελαχιστοποιούν το μέσο τετραγωνικό σφάλμα

$$e(a, b) \ = \ \int_{\mathbb{R}}\int_{\mathbb{R}}(y-ax-b)^2 f(x,y)dxdy$$

Με βάση τα παραπάνω θα ισχύει ότι $b = E[Y - aX] = EY - aEX$. Θέτοντας $\tilde{Y} = Y - EY$, $\tilde{X} = X - EX$ θα έχουμε

$$e(a, b) \ = \ E[(\tilde{Y} - a\tilde{X})^2] = E\tilde{Y}^2 + a^2E\tilde{X}^2 - 2aE[\tilde{X}\tilde{Y}] \ = \ \sigma_Y^2 + a^2\sigma_X^2 - 2ar\sigma_X\sigma_Y.$$

Επομένως, η τιμή του a η οποία ελαχιστοποιεί το μέσο τετραγωνικό σφάλμα είναι

$$a = \frac{r\sigma_Y}{\sigma_X} = \frac{\mathsf{Cov}(X, Y)}{\mathsf{Var}(X)}.$$

Συνεπώς, η βέλτιστη εκτιμήτρια είναι η

$$\mu_y + r\,\frac{\sigma_Y}{\sigma_X}\,(x - \mu_x)$$

και το ελάχιστο σφάλμα είναι

$$\sigma_Y^2(1 - r^2).$$

4.3 Ιδιότητες του πίνακα συνδιακύμανσης

Έστω $(X_1, X_2, \ldots, X_n)$ ένα διάνυσμα τυχαίων μεταβητών οι οποίες θα θεωρήσουμε (χωρίς βλάβη της γενικότητος) ότι έχουν μηδενικούς μέσους. Έστω $V_{ij} = E[X_iX_j] = \mathsf{Cov}(X_i, X_j)$ η συνδιακύμανση των τυχαίων μεταβλητών. Ο πίνακας συνδιακύμανσης V έχει τις εξής δύο ιδιότητες οι οποίες και τον χαρακτηρίζουν: Είναι (προφανώς) *συμμετρικός* δηλαδή $V_{ij} = V_{ji}$ και είναι *θετικά ημιορισμένος*. Υπενθυμίζουμε ότι

Ορισμός 4.1. Ένας τετραγωνικός πίνακας V, $n \times n$, είναι θετικά ορισμένος αν και μόνο αν για κάθε $a \in \mathbb{R}^n$ με $a \neq 0$,

$$a^TVa \ = \ \sum_{i=1}^{n}\sum_{j=1}^{n} a_iV_{ij}a_j > 0.$$

Ο τετραγωνικός πίνακας V ονομάζεται *θετικά ημιορισμένος* αν, για κάθε $a \in \mathbb{R}^n$, $a^TVa \geq 0$.

Για να διαπιστώσουμε ότι ο V είναι θετικά ημιορισμένος αρκεί να παρατηρήσουμε ότι

$$\begin{aligned} a^T V a &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i E[X_i X_j] a_j = E\left[\sum_{i=1}^{n}\sum_{j=1}^{n} a_i X_i X_j a_j\right] = E\left(\sum_{i=1}^{n} a_i X_i\right)\left(\sum_{j=1}^{n} a_j X_j\right) \\ &= E\left(\sum_{i=1}^{n} a_i X_i\right)^2 \geq 0. \end{aligned}$$

Αν οι τυχαίες μεταβλητές είναι γραμμικά ανεξάρτητες, δηλαδή αν δεν υπάρχει μη μηδενικός γραμμικός συνδιασμός $\sum_{i=1}^{n} a_i X_i$ ο οποίος να είναι *ταυτοτικά 0*, ο πίνακας V είναι θετικά ορισμένος.

Από την Γραμμική Άλγεβρα γνωρίζουμε ότι ένας συμμετριτρικός πίνακας V, $n \times n$ έχει n *πραγματικές ιδιοτιμές* $\lambda_1, \ldots, \lambda_n$ (όχι υποχρεωτικά διαφορετικές) και n αντίστοιχα ιδιοδιανύσματα $q_1, q_2, \ldots, q_n$,

$$V q_i = \lambda_i q_i$$

τα οποία είναι ορθογώνια, με μοναδιαίο μήκος, δηλαδή

$$q_i^T q_j = \delta_{ij} := \begin{cases} 1 & \text{αν} \quad i = j, \\ 0 & \text{αν} \quad i \neq j. \end{cases}$$

Αυτό σημαίνει ότι, αν Q είναι ο πίνακας με στήλες τα ιδιοδιανύσματα, $Q = [q_1|q_2|\cdots|q_n]$ τότε $VQ = Q\Lambda$ όπου

$$\Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

είναι ο διαγώνιος πίνακας των ιδιοτιμών. Συνεπώς ο V είναι ορθογώνια διαγωνοποιήσιμος δηλαδή

$$V = Q\Lambda Q^T \quad \text{ή} \quad \Lambda = Q^T V Q \quad \text{όπου} \quad QQ^T = Q^T Q = I. \tag{4.7}$$

Μια εναλλακτική διατύπωση της διαγωνοποίησης ενός συμμετρικού θεωρήματος είναι το λεγόμενο *φασματικό θεώρημα* το οποίο συνοψίζεται στην παρακάτω ισοδύναμη διατύπωση του (4.7

$$V = \sum_{i=1}^{n} \lambda_i U_i\,, \qquad \text{όπου } U_i = q_i q_i^T. \tag{4.8}$$

Στην παραπάνω διατύπωση οι πίνακες $q_i q_i^T$ είναι τετραγωνικοί πίνακες προβολής βαθμού 1. Πράγματι

$$U_i^2 = q_i \left(q_i^T q_i\right) q_i^T = q_i q_i^T = U_i$$

και επομένως ο πίνακας U_i είναι μονοδύναμος και επομένως πίνακας προβολής. Το ότι έχει βαθμό 1 είναι επίσης πολύ εύκολο να αποδειχθεί.

4.4 Η πολυμεταβλητή κανονική κατανομή

Οι τυχαίες μεταβλητές $X_1, X_2, \ldots, X_n$ έχουν από κοινού κανονική κατανομή με μέσο $\mu = (\mu_1, \ldots, \mu_n)^T$ και πίνακα συνδιακυμάνσεων $V = [\mathrm{Cov}(Y_i, Y_j)]$ αν η από κοινού ροπογεννήτριά τους έχει την μορφή

$$M(\theta_1, \ldots, \theta_n) \;=\; E[e^{\theta_1 X_1 + \cdots + \theta_n X_n}] \;=\; e^{\theta^T \mu + \frac{1}{2}\theta^T V \theta} \;=\; e^{\sum_{i=1}^n \theta_i \mu_i + \frac{1}{2}\sum_{i=1}^n \sum_{j=1}^n \theta_i V_{ij} \theta_j}. \qquad (4.9)$$

Ισοδύναμα, (υπό την προϋπόθεση ότι $\det(V) > 0$) οι τυχαίες μεταβλητές $X_1, X_2, \ldots, X_n$ είναι από κοινού κανονικές αν η από κοινού πυκνότητα πιθανότητας δίδεται από την σχέση

$$f(x_1, \ldots, x_n) \;=\; \frac{1}{(2\pi)^{n/2}|V|^{1/2}}\, e^{-\frac{1}{2}(x-\mu)^T V^{-1}(x-\mu)} \qquad x \in \mathbb{R}^n. \qquad (4.10)$$

Στην περίπτωση $n = 2$ με $\sigma_i^2 = \mathrm{Var}(X_i)$, $i = 1, 2$, και $\mathrm{Cov}(X, Y) = \rho\sigma_1\sigma_2$. Συνεπώς ο αντίστροφος του πίνακα συνδιακύμανσης

$$\begin{aligned}
V^{-1} &= \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}^{-1} = \frac{1}{\sigma_1^2\sigma_2^2(1-\rho^2)} \begin{bmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{bmatrix} \\
&= \frac{1}{1-\rho^2} \begin{bmatrix} \sigma_1^{-2} & -\rho(\sigma_1\sigma_2)^{-1} \\ -\rho(\sigma_1\sigma_2)^{-1} & \sigma_2^{-2} \end{bmatrix} \qquad (4.11)
\end{aligned}$$

Επίσης, η ορίζουσα του πίνακα συνδιακύμανσης είναι $| V |= \sigma_1^2\sigma_2^2\left(1-\rho^2\right)$. Συνεπώς, από την σχέση (4.10),

$$\begin{aligned}
f(x_1, x_2) &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\, e^{-\frac{1}{2}(x_1-\mu_1, x_2-\mu_2)V^{-1}\binom{x_1-\mu_1}{x_2-\mu_2}} \\
&= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\, e^{-\frac{1}{2(1-\rho^2)}(x_1-\mu_1, x_2-\mu_2)\begin{bmatrix} \sigma_1^{-2} & -\rho(\sigma_1\sigma_2)^{-1} \\ -\rho(\sigma_1\sigma_2)^{-1} & \sigma_2^{-2} \end{bmatrix}\binom{x_1-\mu_1}{x_2-\mu_2}} \\
&= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\, e^{-\frac{1}{2(1-\rho^2)}\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{x_1-\mu_1}{\sigma_1}\frac{x_2-\mu_2}{\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right)}. \qquad (4.12)
\end{aligned}$$

Το επόμενο κριτήριο κανονικότητας είναι ιδιαίτερα εύχρηστο

Θεώρημα 4.1. *Οι τυχαίες μεταβλητές* $X_1, X_2, \ldots, X_n$ *είναι από κοινού κανονικά κατανεμημένες αν και μόνο αν κάθε γραμμικός συνδιασμός των τυχαίων αυτών μεταβλητών είναι κανονικά κατανεμημένος δηλαδή αν και μόνο αν, για κάθε* $(\theta_1, \ldots, \theta_n)$, *η τυχαία μεταβλητή* $\sum_{i=1}^n \theta_i X_i$ *είναι κανονικά κατανεμημένη.*

Απόδειξη. 1) Έστω $(X_1, \ldots, X_n) \sim \mathcal{N}(\mu, V)$ όπου $\mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}$ και $V = \begin{bmatrix} v_{11} & \cdots & v_{1n} \\ \vdots & & \vdots \\ v_{n1} & \cdots & v_{nn} \end{bmatrix}$.
Έστω $Z := \sum_{i=1}^n \theta_i X_i$. Δεδομένου ότι οι τυχαίες μεταβλητές $(X_1, \ldots, X_n)$ είναι από κοινού

κανονικά κατανεμημένες έχουμε, λόγω της (4.9), ότι

$$\begin{aligned} Ee^{tZ} &= E\left[e^{t(\theta_1 X_1+\cdots+\theta_n X_n)}\right] = e^{\theta^T\mu+\frac{1}{2}\theta^T V\theta} = e^{\sum_{i=1}^n (t\theta_i)\mu_i+\frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n (t\theta_i)V_{ij}(t\theta_j)} \\ &= e^{t\left(\sum_{i=1}^n \theta_i\mu_i\right)+\frac{1}{2}t^2\left(\sum_{i=1}^n\sum_{j=1}^n \theta_i V_{ij}\theta_j\right)}. \end{aligned} \tag{4.13}$$

Συνεπώς, από την μορφή της ροπογεννήτριας, η τυχαία μεταβλητή Z είναι

$$\mathscr{N}\left(\sum_{i=1}^n \theta_i\mu_i\,,\ \sum_{i=1}^n\sum_{j=1}^n \theta_i V_{ij}\theta_j\right). \tag{4.14}$$

Αντίστροφα τώρα, ας υποθέσουμε ότι για κάθε (θ_1, θ_n) η τυχαία μεταβλητή $Z := \sum_{i=1}^n \theta_i X_i$ είναι κανονικά κατανεμημένη. Ο μέσος και η διασπορά είναι αναγκαστικά $\sum_{i=1}^n \theta_i\mu_i$ και $\sum_{i=1}^n\sum_{j=1}^n \theta_i V_{ij}\theta_j$ αντίστοιχα. Αν η Z είναι κανονικά κατανεμημένη τότε, δεδομένης της μορφής της ροπογεννήτριας για την (μονοδιάστατη) κανονική κατανομή

$$E[e^{\sum_{i=1}^n \theta_i X_i}] = Ee^{1\cdot Z} = Ee^{1\cdot\sum_{i=1}^n \theta_i\mu_i+\frac{1}{2}1^2\cdot\sum_{i=1}^n\sum_{j=1}^n \theta_i V_{ij}\theta_j}$$

Συνγκίνοντας με την (4.9) βλέπουμε ότι $(X_1, \ldots, X_n) \sim \mathscr{N}(\mu, V)$. □

Το ακόλουθο θεώρημα είναι ιδιαιτερα χρήσιμο:

Θεώρημα 4.2. *Έστω $X = (X_1, \ldots, X_n)$ κανονικά κατανεμημένο διάνυσμα στον $\mathbb{R}^n$, $\mathscr{N}(\mu, V)$ και A πίνανκας $m \times n$. Τότε το τυχαίο διάνυσμα $Y := AX$ είναι κανονικά κατανεμημένο στον $\mathbb{R}^m$, $\mathscr{N}(A\mu, AVA^T)$.*

Απόδειξη. Ο μέσος και ο πίνακας συνδιακύμανσης του Y έχει ήδη υπολογισθεί στις (4.1) και (4.4). Μένει να δείξουμε ότι η κατανομή του διανύσματος Y είναι πολυμεταβλητή κανονική. Χρησιμοποιώντας το θεώρημα (4.1), έχουμε

$$\sum_{i=1}^m \theta_i Y_i = \sum_{i=1}^m \theta_i \sum_{j=1}^n a_{ij}X_j = \sum_{j=1}^n \left(\sum_{i=1}^m \theta_i a_{ij}\right) X_j = \sum_{j=1}^n t_j X_j.$$

Αφού τα X_j είναι από κοινού κανονικά κατανεμημένα, ο γραμμικός συνδιασμός $\sum_{j=1}^n t_j X_j$ είναι κανονικά κατανεμημένος. Αυτό σημαίνει, από την παραπάνω εξίσωση ότι οποιοσδήποτε γραμμικός συνδιασμός των Y_i είναι κανονικά κατανεμημένος και επομένως από το θεώρημα (4.1) ότι $Y \sim \mathscr{N}(A\mu, AVA^T)$. □

Θεώρημα 4.3. *Έστω X ένα τυχαίο διάνυσμα στον $\mathbb{R}^n$ κανονικά κατανεμημένο με μέσο $\mu \in \mathbb{R}^n$ και πίνακα συνδιακύμανσης V $(n \times n)$. Αν διαμερίσουμε το X σε δύο τμήματα, $X = \begin{pmatrix} Y \\ Z \end{pmatrix}$, όπου*

$$Y_1 = \begin{pmatrix} X_1 \\ \vdots \\ X_m \end{pmatrix}, \quad \textit{και} \quad Y_2 = \begin{pmatrix} X_{m+1} \\ \vdots \\ X_n \end{pmatrix}.$$

Τα τυχαία διανύσματα Y_1 *και* Y_2 *που έχουν διαστάσεις* m *και* $n-m$ *αντίστοιχα, μέσους*

$$\mu_{Y_1} := \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_m \end{pmatrix}, \quad \textit{και} \quad \mu_{Y_2} := \begin{pmatrix} \mu_{m+1} \\ \vdots \\ \mu_n \end{pmatrix},$$

και πίνακα συνδιακύμανσης

$$V = \left[\begin{array}{c|c} \mathsf{E}[(Y_1-\mu_{Y_1})(Y_1-\mu_{Y_1})^T] & \mathsf{E}[(Y_1-\mu_{Y_1})(Y_2-\mu_{Y_2})^T] \\ \hline \mathsf{E}[(Y_2-\mu_{Y_2})(Y_1-\mu_{Y_1})^T] & \mathsf{E}[(Y_2-\mu_{Y_2})(Y_2-\mu_{Y_2})^T] \end{array}\right] = \left[\begin{array}{c|c} V_{11} & V_{12} \\ \hline V_{21} & V_{22} \end{array}\right].$$

Υποθέτουμε ότι

$$\det V_{22} > 0. \tag{4.15}$$

Τότε η δεσμευμένη κατανομή του διανύσματος Y_1 *δεδομένου του* Y_2 *είναι* η πολυμεταβλητή κανονική κατανομή στο $\mathbb{R}^m$ *με διάνυσμα μέσων*

$$\mathsf{E}[Y_1|Y_2] = \mu_{Y_1} + V_{12}V_{22}^{-1}(Y_2-\mu_{Y_2}) \tag{4.16}$$

και πίνακα συνδιακύμανσης

$$\mathsf{Cov}(Y_1|Y_2=y_2) = V_{11} - V_{12}V_{22}^{-1}V_{21}. \tag{4.17}$$

Απόδειξη. Θεωρούμε το διάνυσμα

$$\mathsf{Z} = \left[\begin{array}{c|c} I_m & -V_{12}V_{22}^{-1} \\ \hline 0 & I_{n-m} \end{array}\right] \left[\begin{array}{c} Y_1 \\ \hline Y_2 \end{array}\right].$$

όπου I_m και I_{n-m} ταυτοτικοί πίνακες διαστάσεων m και $n-m$ αντίστοιχα. Από το προηγούμενο θεώρημα γνωρίζουμε ότι το τυχαίο διάνυσμα Z είναι κανονικά κατανεμημένο στον $\mathbb{R}^n$. (Παρατηρείστε επίσης ότι ο V_{22} είναι πίνακας $(n-m)\times(n-m)$ και ο V_{12} είναι $m\times(n-m)$. Το διάνυσμα Z γράφεται και ως $\mathsf{Z} = \begin{pmatrix} Y_1 - V_{12}V_{22}^{-1}Y_2 \\ U_2 \end{pmatrix}$ και επομένως

$$\begin{aligned}
\mathsf{E}[\mathsf{Z}\mathsf{Z}^T] &= \left[\begin{array}{c|c} I_m & -V_{12}V_{22}^{-1} \\ \hline 0 & I_{n-m} \end{array}\right] \mathsf{E}\left[\begin{array}{c|c} Y_1Y_1^T & Y_1Y_2^T \\ \hline Y_2Y_1^T & Y_2Y_2^T \end{array}\right] \left[\begin{array}{c|c} I_m & 0 \\ \hline -V_{22}^{-1}V_{12}^T & I_{n-m} \end{array}\right] \\
&= \left[\begin{array}{c|c} I_m & -V_{12}V_{22}^{-1} \\ \hline 0 & I_{n-m} \end{array}\right] \left[\begin{array}{c|c} V_{11} & V_{12} \\ \hline V_{21} & V_{22} \end{array}\right] \left[\begin{array}{c|c} I_m & 0 \\ \hline -V_{22}^{-1}V_{12}^T & I_{n-m} \end{array}\right] \\
&= \left[\begin{array}{c|c} I_m & -V_{12}V_{22}^{-1} \\ \hline 0 & I_{n-m} \end{array}\right] \left[\begin{array}{c|c} V_{11}-V_{12}V_{22}^{-1}V_{21} & V_{12} \\ \hline 0 & V_{22} \end{array}\right] = \left[\begin{array}{c|c} V_{11}-V_{12}V_{22}^{-1}V_{21} & 0 \\ \hline 0 & V_{22} \end{array}\right].
\end{aligned}$$

Από τον παραπάνω υπολογισμό συμπεραίνουμε ότι τα τυχαία διανύσματα $Y_1 - V_{12}V_{22}^{-1}Y_2$ και Y_2 είναι γραμμικά ανεξάρτητα. □

Για την περίπτωση δυο τυχαίων μεταβλητών,

$$\mathsf{E}[Y|X=x] = \mu_y - \frac{\mathsf{Cov}(X,Y)}{\mathsf{Var}(X)}(x-\mu_x) = \mu_y - \rho\,\frac{\sigma_y}{\sigma_x}(x-\mu_x) \tag{4.18}$$

$$\mathsf{Var}(Y|X=x) = \mathsf{Var}(Y) - \frac{(\mathsf{Cov}(X,Y))^2}{\mathsf{Var}(X)} = \sigma_y^2\left(1-\rho^2\right). \tag{4.19}$$

4.5 Παραδείγματα

Πρόβλημα 1. Έστω ότι η από κοινού πυκνότητα πιθανότητας των τυχαίων μεταβλητών (X, Y) δίδεται από την συνάρτηση

$$f(x,y) = \frac{1}{2\pi} e^{-\frac{1}{2}\left(x^2 - 2xy + 2y^2\right)}$$

Να βρεθούν οι μέσοι, ο πίνακας συνδιακυμάνσεων, ο συντελεστής συσχέτισης, οι περιθώριες πυκνότητες $f_X(x)$, $f_Y(y)$, και η δεσμευμένη πυκνότητα $f_{Y|X}(y|x)$.

Λύση: Ο εκθέτης γράφεται ως $-\frac{1}{2}\left(x^2 - 2xy + 2y^2\right) = -\frac{1}{2}[x, y]V^{-1}\begin{bmatrix} x \\ y \end{bmatrix}$ με $V^{-1} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$. Συνεπώς οι μέσοι είναι 0, ο πίνακας συνδιακύμανσης είναι $V = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$, ο συντελεστής συσχέτισης είναι $\rho = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}} = \frac{1}{\sqrt{1\cdot 2}} = \frac{1}{\sqrt{2}}$. Επίσης

$$f_X(x) = \frac{1}{2\sqrt{\pi}} e^{-\frac{x^2}{4}}, \quad f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}, \quad f_{Y|X}(y|x) = \frac{f(x,y)}{f_Y(y)} = \frac{1}{\sqrt{\pi}} e^{-\left(y - \frac{x}{2}\right)^2}.$$

Βλέπουμε επομένως ότι $Y|X = x \sim \mathscr{N}\left(\frac{x}{2}, \frac{1}{2}\right)$.

Πρόβλημα 2. Έστω ότι η από κοινού πυκνότητα πιθανότητας των τυχαίων μεταβλητών (X, Y) δίδεται από την συνάρτηση

$$f(x,y) = K\, e^{-\frac{1}{2}\left(3x^2 - 6xy + 4y^2 - 6x + 24y + C\right)}$$

Να βρεθούν οι μέσοι, μ_x, μ_y, ο πίνακας συνδιακυμάνσεων, οι τιμές των σταθερών K και C, ο συντελεστής συσχέτισης, οι περιθώριες πυκνότητες $f_X(x)$, $f_Y(y)$, και η δεσμευμένη πυκνότητα $f_{Y|X}(y|x)$.

Λύση: Για τον εκθέτη, τον πίνακα συνδιακύμανσης, V^{-1} και τους μέσους θα πρέπει να ισχύει

$$\begin{aligned}
3x^2 - 6xy + 4y^2 - 6x + 24y + C \;&=\; \begin{bmatrix} x - \mu_x & y - \mu_y \end{bmatrix} V^{-1} \begin{bmatrix} x - \mu_x \\ y - \mu_y \end{bmatrix} \\
&= V_{11}^{-1}(x - \mu_x)^2 + V_{22}^{-1}(y - \mu_y)^2 + 2V_{12}^{-1}(x - \mu_x)(y - \mu_y) \\
&= V_{11}^{-1}x^2 + 2V_{12}^{-1}xy + V_{22}^{-1}y^2 + 2x\left(V_{11}^{-1}\mu_x + V_{12}^{-1}\mu_y\right) + 2y\left(V_{12}^{-1}\mu_x + V_{22}^{-1}\mu_y\right) + \begin{bmatrix} \mu_x & \mu_y \end{bmatrix} V^{-1} \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}
\end{aligned}$$

Στις παραπάνω σχέσεις, V_{ij}^{-1} είναι το στοιχείο i, j του πίνακα V^{-1}. Επίσης, αφού ο πίνακας συνδιακυμάνσεων V είναι συμμετρικός, το ίδιο ισχύει και για τον αντίστροφό του και συνεπώς $V_{12}^{-1} = V_{21}^{-1}$. Από τις σχέσεις αυτές συμπεραίνουμε ότι

$$\begin{aligned}
&V_{11}^{-1} = 3, \quad V_{22}^{-1} = 4, \quad V_{12}^{-1} = -3 \\
&2\left(V_{11}^{-1}\mu_x + V_{12}^{-1}\mu_y\right) = 6, \;\; 2\left(V_{12}^{-1}\mu_x + V_{22}^{-1}\mu_y\right) = 24, \;\; \begin{bmatrix} \mu_x & \mu_y \end{bmatrix} V^{-1} \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}
\end{aligned} \tag{4.20}$$

Η (4.20) δίνει το σύστημα

$$\begin{array}{rcl} 3\mu_x - 3\mu_y & = & 3 \\ -3\mu_x + 4\mu_y & = & 12 \end{array} \iff \begin{array}{rcl} \mu_x & = & -14 \\ \mu_y & = & 15 \end{array}$$

και

$$\begin{bmatrix} \mu_x & \mu_y \end{bmatrix} V^{-1} \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix} = \begin{bmatrix} -14 & 15 \end{bmatrix} \begin{bmatrix} 3 & -3 \\ -3 & 4 \end{bmatrix} \begin{bmatrix} -14 \\ 15 \end{bmatrix} = 228.$$

Ο πίνακας συνδιακύμανσης V θα είναι

$$V = \begin{bmatrix} 3 & -3 \\ -3 & 4 \end{bmatrix}^{-1} = \frac{1}{3} \begin{bmatrix} 4 & 3 \\ 3 & 3 \end{bmatrix} = \begin{bmatrix} 4/3 & 1 \\ 1 & 1 \end{bmatrix}.$$

Συνεπώς η σταθερά $K = \frac{1}{2\pi|V|^{1/2}} = \frac{1}{2\pi\sqrt{1/3}} = \frac{\sqrt{3}}{2\pi}$. Ο συντελεστής συσχέτισης είναι

$$\rho = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}} = \frac{1}{\sqrt{\frac{4}{3}\cdot 1}} = \frac{\sqrt{3}}{2}.$$

Από τον πίνακα συνδιακύμανσης και τους μέσους βλέπουμε ότι $X \sim \mathcal{N}(-14, \frac{4}{3})$ και $Y \sim \mathcal{N}(15, 1)$ συνεπώς

$$f_X(x) = \frac{\sqrt{3}}{2\sqrt{2\pi}} e^{-\frac{3}{8}(x+14)^2}, \qquad f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y-15)^2}.$$

Προκειμένου να υπολογίσουμε την $f_{Y|X}(y|x)$ είναι απλούστερο να επικαλεστούμε το γεγονός ότι

$$\begin{aligned} Y|X = x \quad &\sim \quad \mathcal{N}\left(\mu_y + \rho\frac{\sigma_y}{\sigma_x}(x-\mu_x)\ ,\ \sigma_y^2(1-\rho^2)\right) \\ &= \quad \mathcal{N}\left(15 + \frac{\sqrt{3}}{2}\frac{1}{\frac{2}{\sqrt{3}}}(x+14)\ ,\ 1\cdot(1-\frac{3}{4})\right) = \mathcal{N}\left(\frac{102+3x}{4}\ ,\ \frac{1}{4}\right). \end{aligned}$$

Συνεπώς

$$f_{Y|X}(y|x) \ = \ \sqrt{\frac{2}{\pi}} e^{-2\left(y-\frac{102+3x}{4}\right)^2}.$$

Παράδειγμα 1. Έστω X τυχαίο διάνυσμα στον $\mathbb{R}^3$ με μηδενικούς μέσους και πίνακα συνδιακύμανσης

$$E\left[XX^T\right] \ = \ V \ = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 6 & 0 \\ -1 & 0 & 4 \end{bmatrix}.$$

Ορίζουμε τώρα τις τυχαίες μεταβλητές Y_1, Y_2, Y_3 που δίδονται από τις γραμμικές σχέσεις

$$\begin{aligned} Y_1 &= X_1 + X_3 \\ Y_2 &= 2X_1 - X_2 \\ Y_3 &= -X_2 + 2X_3. \end{aligned}$$

Οι παραπάνω εξισώσεις γράφονται ως $Y = AX$ ή, αναλυτικά,

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 2 & -1 & 0 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}.$$

Συνεπώς, από το προηγούμενο θεώρημα, $Y \sim \mathscr{N}(0, AVA^T)$, όπου

$$AVA^T = \begin{bmatrix} 1 & 0 & 1 \\ 2 & -1 & 0 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 & -1 \\ 2 & 6 & 0 \\ -1 & 0 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 0 & -1 & -1 \\ 1 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 3 & -2 & 4 \\ -2 & 2 & -2 \\ 4 & -2 & 22 \end{bmatrix}.$$

Η ορίζουσα του πίνακα συνδιακύμανσης του Y είναι $\det(AVA^T) = \det(A)\det(V)\det(A^T) = (-4)\cdot 2 \cdot (-4) = 32$ και ο αντίστροφος του πίνακα αυτού είναι

$$\left(AVA^T\right)^{-1} = \frac{1}{16} \begin{bmatrix} 20 & 18 & -2 \\ 18 & 25 & -1 \\ -2 & -1 & 1 \end{bmatrix} = \begin{bmatrix} \frac{5}{4} & \frac{9}{8} & -\frac{1}{8} \\ \frac{9}{8} & \frac{25}{16} & -\frac{1}{16} \\ -\frac{1}{8} & -\frac{1}{16} & \frac{1}{16} \end{bmatrix}.$$

Επομένως η πυκνότητα πιθανότητας του Y είναι

$$f(y_1, y_2, y_3) = \frac{1}{(2\pi)^{\frac{3}{2}}\sqrt{32}} e^{-\frac{1}{32}\left(20y_1^2 + 25y_2^2 + y_3^2 - 36y_1y_2 - 4y_1y_3 - 2y_2y_3\right)}.$$

Παράδειγμα 2. Έστω (X_1, X_2, X_3) τυχαίες μεταβλητές από κοινού κανονικά κατανεμημένες. Αν α) οι X_1 και $X_2 + X_3$ είναι ανεξάρτητες, β) οι X_2 και $X_3 + X_1$ είναι ανεξάρτητες, και γ) οι X_3 και $X_1 + X_2$ είναι ανεξάρτητες, τότε να δείξετε ότι οι X_1, X_2, X_3 είναι ανεξάρτητες.

Λύση: Έστω $Y_1 := X_2 + X_3$, $Y_2 := X_3 + X_1$, και $Y_3 := X_1 + X_2$. Οι (X_1, Y_1) είναι από κοινού κανονικές διότι προκύπτουν από τον γραμμικό συνδιασμό

$$\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix}$$

και επομένως ο πίνακας διακυμάνσεών τους είναι

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} v_{11} & v_{12} + v_{13} \\ v_{21} + v_{31} & v_{22} + v_{23} + v_{32} + v_{33} \end{pmatrix}.$$

Αφού οι X_1, Y_1 είναι ανεξάρτητες,

$$\mathsf{Cov}(X_1, Y_1) = v_{12} + v_{13} = 0. \tag{4.21}$$

Με τον ίδιο τρόπο

$$\begin{pmatrix} X_2 \\ Y_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix}$$

και αντίστοιχα θα προκύψει με τον ίδιο τρόπο ότι

$$\mathsf{Cov}(X_2, Y_2) = \nu_{21} + \nu_{23} = 0. \tag{4.22}$$

και επίσης, επαναλαμβάνοντας το επιχείρημα για τα (X_3, Y_3),

$$\mathsf{Cov}(X_3, Y_3) = \nu_{31} + \nu_{32} = 0. \tag{4.23}$$

Από τις (4.21), (4.22), και (4.23), καθώς και από το γεγονός ότι $\nu_{ij} = \nu_{ji}$ αφού ο πίνακας συνδιακυμάνσεων είναι συμμετρικός, προκύπτει ότι

$$\nu_{12} = \nu_{23} = \nu_{31} = 0$$

και επομένως οι X_1, X_2, X_3 είναι ασυσχέτιστες, και εφ΄ όσον είναι κανονικά κατανεμημένες αυτό σημαίνει ότι είναι και ανεξάρτητες.

Παράδειγμα 3. Έστω X_1, X_2, X_3 ανεξάρτητες τυχαίες μεταβλητές $\mathscr{N}(1,1)$ και $U_1 := 2X_1 - X_2 + X_3$, $U_2 = X_1 + 2X_2 + 3X_3$. Να βρείτε την δεσμευμένη πυκνότητα της U_2 δεδομένου ότι $U_1 = 3$.

Λύση: Οι μέσοι των U_1, U_2, είναι $\mathsf{E}U_1 = 2\mathsf{E}X_1 - \mathsf{E}X_2 + \mathsf{E}X_3 = 2 - 1 + 1 = 2$, $\mathsf{E}U_2 = \mathsf{E}X_1 + 2\mathsf{E}X_2 + 3\mathsf{E}X_3 = 6$. Ο πίνακας συνδιακυμάνσης των U_1, U_2 είναι

$$\begin{pmatrix} 2 & -1 & 1 \\ 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ -1 & 2 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 6 & 3 \\ 3 & 14 \end{pmatrix}.$$

Ισχύει ότι

$$\begin{pmatrix} U_1 \\ U_2 \end{pmatrix} = \mathscr{N}\left(\begin{pmatrix} 2 \\ 6 \end{pmatrix}, \begin{pmatrix} 6 & 3 \\ 3 & 14 \end{pmatrix} \right)$$

και

$$U_1 \,|\, U_2 = 3 \sim \mathscr{N}\left(2 + \frac{3}{14}(3-6)\ ,\ 6 - \frac{3^2}{14}\right) = \mathscr{N}\left(\frac{19}{14}, \frac{75}{14}\right)$$

Παράδειγμα 4. Έστω $X \sim \mathscr{N}(0, V)$ όπου

$$\begin{bmatrix} 1 & -\frac{1}{2} & \frac{3}{2} \\ -\frac{1}{2} & 2 & -1 \\ \frac{3}{2} & -1 & 4 \end{bmatrix}.$$

Να προσδιορίσετε την από κοινού κατανομή των X_1 και X_1+X_2 δεδομένου ότι $X_1+X_2+X_3=0$. Προκειμένου να λύσουμε αυτό το πρόβλημα θέτουμε $Y_1=X_1$, $Y_2=X_1+X_2$, $Y_3=X_1+X_2+X_3$ δηλαδή

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}$$

Επομένως, ο πίνακας συνδιακύμανσης των Y_1, Y_2, Y_3 είναι

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1/2 & 3/2 \\ -1/2 & 2 & -1 \\ 3/2 & -1 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} = \left[\begin{array}{cc|c} 1 & 1/2 & 2 \\ 1/2 & 2 & 5/2 \\ \hline 2 & 5/2 & 7 \end{array}\right]$$

Η από κοινού κατανομή των Y_1, Y_2, δεδομένου $Y_3=0$ είναι κανονική με μέσο $\left(0//0 \right)$ και πίνακα συνδιακυμάνσεων

$$\begin{bmatrix} 1 & 1/2 \\ 1/2 & 2 \end{bmatrix} - \begin{bmatrix} 2 \\ 5/2 \end{bmatrix} \frac{1}{7} \begin{bmatrix} 2 & 5/2 \end{bmatrix} = \begin{bmatrix} 1 & 1/2 \\ 1/2 & 2 \end{bmatrix} - \frac{1}{7} \begin{bmatrix} 4 & 5 \\ 5 & 25/4 \end{bmatrix} = \begin{bmatrix} 3/7 & -3/14 \\ -3/14 & 31/28 \end{bmatrix}$$