Πιθανότητες II – 1$^{\text{ο}}$ Μέρος

Μιχάλης Ζαζάνης
Τμήμα Στατιστικής
Οικονομικό Πανεπιστήμιο Αθηνών

4 Απριλίου 2017

Κεφάλαιο 1

Συνεχείς Τυχαίες Μεταβλητές

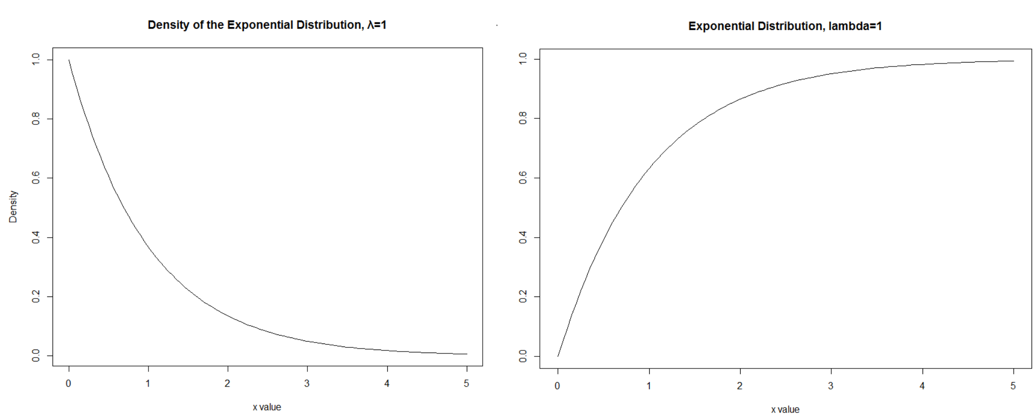
1.1 Η εκθετική κατανομή

Η πυκνότητα πιθανότητας της εκθετικής κατανομής δίδεται από την σχέση

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{αν} \quad x \geq 0, \\ 0 & \text{αν} \quad x < 0. \end{cases} \tag{1.1}$$

Η παράμετρος λ της κατανομής είναι *θετικός* αριθμός. Η συνάρτηση κατανομής δίδεται από την σχέση $F(x) := \int_{-\infty}^{x} f(t)dt$, που δίνει

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{αν} \quad x \geq 0, \\ 0 & \text{αν} \quad x < 0. \end{cases} \tag{1.2}$$



Σχήμα 1.1: Εκθετική κατανομή, $\lambda = 1$

Λήμμα 1.1. *Ισχύει ότι*

$$I_n(x) \ := \ -n! \sum_{k=0}^{n} \frac{x^k}{k!} e^{-x} \ = \ \int x^n e^{-x} dx, \qquad n = 0, 1, 2, \ldots. \tag{1.3}$$

Απόδειξη. Αρκεί να δείξουμε ότι $I_n'(x) = x^n e^{-x}$. Πράγματι,

$$\begin{aligned} I_n'(x) &= \frac{d}{dx}(-n!) \left(1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots + \frac{x^{n-1}}{(n-1)!} + \frac{x^n}{n!}\right) e^{-x} \\ &= -n! \left(1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots + \frac{x^{n-1}}{(n-1)!}\right) e^{-x} \ + \ n! \left(1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots + \frac{x^{n-1}}{(n-1)!} + \frac{x^n}{n!}\right) e^{-x} \\ &= x^n e^{-x}, \end{aligned}$$

και επομένως η (1.3) ισχύει. □

Οι ροπές της κατανομής δίδονται από την σχέση

$$\mathbb{E}X^n \ = \ \int_0^\infty x^n \lambda e^{-\lambda x} dx \ = \ \frac{1}{\lambda^n} \int_0^\infty (\lambda x)^n e^{-\lambda x} \lambda dx \ = \ \frac{1}{\lambda^n} \int_0^\infty u^n e^{-u} du \tag{1.4}$$

όπου $u = \lambda x$. Λαμβάνοντας υπ' όψιν μας το Λήμμα 1.1

$$\int_0^\infty u^n e^{-u} du \ = \ I_n(u) \Big|_0^\infty \ = \ -0 \ - \ (-n!) \ = \ n!. \tag{1.5}$$

(Χρησιμοποιήσαμε το γεγονός ότι $\lim_{u\to\infty} n! \sum_{k=0}^{n} \frac{u^k}{k!} e^{-u} = 0$ καθώς και ότι $I_n(0) = n! \left(1 + \frac{0}{1!} + \frac{0^2}{2!} + \ \cdots + \frac{0^{n-1}}{(n-1)!} + \frac{0^n}{n!}\right) = n!$.) Επομένως, από την (1.4) προκύπτει ότι

$$\mathbb{E}X^n \ = \ \frac{n!}{\lambda^n}, \qquad n = 0, 1, 2, \ldots. \tag{1.6}$$

Συγκεκριμένα, η μέση τιμή της εκθετικής είναι $\mathbb{E}X = \frac{1}{\lambda}$, η δεύτερη ροπή είναι $\mathbb{E}X^2 = \frac{2}{\lambda^2}$, και η διασπορά δίδεται από την σχέση $\mathsf{Var}(X) = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \frac{1}{\lambda^2}$.

Ο *συντελεστής μεταβλητότητας* μιας κατανομής ορίζεται ως

$$c_v = \frac{\sqrt{\mathsf{Var}(X)}}{\mathbb{E}X}$$

Για την εκθετική κατανομή έχουμε $c_v = 1$. Μια από τις σημαντικότερες ιδιότητες της εκθετικής κατανομής, η οποία και την χαρακτηρίζει, είναι η αμνήμων ιδιότητα (ή ιδιότητα έλλειψης μνήμης):

Μια μή αρνητική τ.μ. X έχει την *αμνήμονα ιδιότητα* αν, $\forall s, t > 0$,

$$\mathbb{P}(X > t + s | X > t) = \mathbb{P}(X > s). \tag{1.7}$$

Είναι στοιχειώδες να ελέγξουμε ότι η έκθετική κατανομή ικανοποιεί την (1.7). Πράγματι,

$$\begin{aligned}\mathbb{P}(X > t+s|X > t) &= \frac{\mathbb{P}(X > t+s, X > t)}{\mathbb{P}(X > t)} = \frac{\mathbb{P}(X > t+s)}{\mathbb{P}(X > t)} = \frac{e^{-\lambda(t+s)}}{e^{-\lambda t}} = e^{-\lambda s} \\ &= \mathbb{P}(X > s).\end{aligned}$$

Πολύ πιό ενδιαφέρον όμως είναι το γεγονός ότι η εκθετική κατανομή είναι η μοναδική η οποία έχει αυτή την ιδιότητα. Έστω λοιπόν κάποια τ.μ. X που ικανοποιεί την (1.7) και ας θέσουμε $g(t) = \mathbb{P}(X > t)$. Οπως είδαμε, η (1.7) μπορεί να γραφτεί και ως $\mathbb{P}(X > t+s) = \mathbb{P}(X > t)\mathbb{P}(X > s)$ ή

$$g(t+s) = g(t)g(s). \tag{1.8}$$

Εφαρμόζοντας την παραπάνω σχέση με $s = t = 1$ έχουμε $g(2) = g(1)^2$ και επαγωγικά,

$$g(n) = g(1)^n.$$

Όμοια, $g(1) = g(\frac{1}{n} + \cdots + \frac{1}{n}) = g(\frac{1}{n})^n$ ή

$$g(1/n) = g(1)^{1/n}.$$

Από τις δύο παραπάνω σχέσεις προκύπτει ότι, για κάθε ρητό, έστω $r = p/q$

$$g(p/q) = g(1)^{p/q}$$

ή $g(r) = g(1)^r$. Απομένει να δείξουμε ότι η σχέση αυτή ισχύει για κάθε πραγματικό αριθμό t. Αφού $0 < g(1) < 1$, υπάρχει $\lambda \in (0, \infty)$ τέτοιο ώστε $g(1) = e^{-\lambda}$, Συνεπώς, αν r, s είναι ρητοί τέτοιοι ώστε $r < t < s$, εφόσον η $g(t)$ είναι φθίνουσα συνάρτηση ισχύει ότι

$$e^{-\lambda r} = g(r) \geq g(t) \geq g(s) = e^{-\lambda s}$$

Αφήνοντας $r \uparrow t$, $s \downarrow t$ προκύπτει το ζητούμενο.

1.2 Η συνάρτηση Γάμμα

Η συνάρτηση Γάμμα του Euler αποτελεί γενίκευση του παραγοντικού για μή ακέραια ορίσματα. Για κάθε $\alpha > 0$ ορίζουμε την συνάρτηση

$$\Gamma(\alpha) := \int_0^\infty t^{\alpha-1} e^{-t} dt. \tag{1.9}$$

(Παρατηρείστε ότι το παραπάνω είναι ένα γενικευμένο ολοκλήρωμα.) Ισχύει ότι

$$\Gamma(1) = \int_0^\infty e^{-t} dt = 1. \tag{1.10}$$

Χρησιμοποιώντας παραγοντική ολοκλήρωση βλέπουμε ότι

$$
\begin{aligned}
\Gamma(\alpha+1) &= \int_0^\infty t^\alpha e^{-t} dt = -\int_0^\infty t^\alpha \left(e^{-t}\right)' dt = -t^\alpha e^{-t}\Big|_0^\infty + \int_0^\infty \alpha t^{\alpha-1} e^{-t} dt \\
&= 0 + \alpha \int_0^\infty t^{\alpha-1} e^{-t} dt
\end{aligned}
$$

και συνεπώς έχουμε

$$
\Gamma(\alpha+1) = \alpha\,\Gamma(\alpha), \qquad \alpha > 0. \tag{1.11}
$$

Όταν το α είναι φυσικός αριθμός η παραπάνω σχέση δίνει

$$
\Gamma(n+1) = n\Gamma(n) = n(n-1)\Gamma(n-1) = n(n-1)(n-2)\cdots 2\cdot 1\cdot \Gamma(1)
$$

και λαμβάνοντας υπ΄ όψιν μας την (1.10) έχουμε

$$
\Gamma(n+1) = n!, \qquad n = 0, 1, 2, \ldots. \tag{1.12}
$$

Παρατηρείστε ότι η (1.12) συμφωνεί με την (1.5). Εφαρμόζοντας την σχέση (1.11) επαγωγικά βλέπουμε ότι, για κάθε $n \in \mathbb{N}$ και $\alpha > 0$

$$
\begin{aligned}
\Gamma(\alpha+n) &= (\alpha+n-1)\,\Gamma(\alpha+n-1) = (\alpha+n-1)(\alpha+n-2)\,\Gamma((\alpha+n-2) = \cdots \\
&= (\alpha+n-1)(\alpha+n-2)\,\Gamma(\alpha+n-2)\cdots(\alpha+1)\cdot\alpha\,\Gamma(\alpha).
\end{aligned} \tag{1.13}
$$

1.3 Η κατανομή Γάμμα

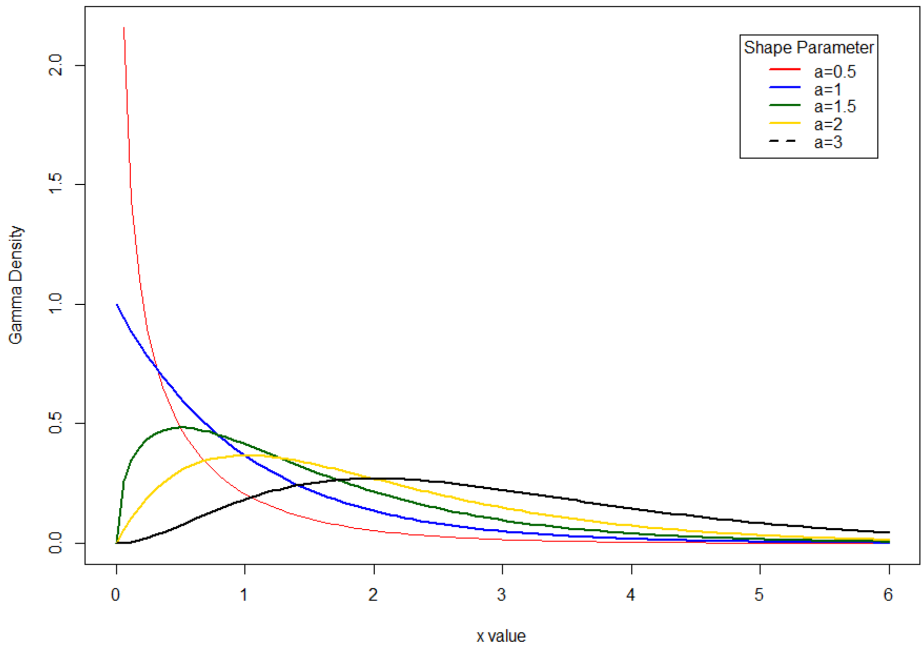
Εμείς θα συναντήσουμε την συνάρτηση Γ στα πλαίσια της κατανομής Γάμμα που έχει πυκνότητα πιθανότητας

$$
f(x) = \begin{cases} \lambda \dfrac{(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} e^{-\lambda x} & \text{αν} \quad x > 0 \\[2ex] 0 & \text{αν} \quad x \le 0. \end{cases} \tag{1.14}
$$

Η παράμετρος λ ονομάζεται παράμετρος κλίμακος ενώ η α παράμετρος σχήματος. Η περίπτωση που η παράμετρος σχήματος είναι ακέραιος αριθμός έχει ιδιαίτερο ενδιαφέρον και θα την εξετάσουμε λεπτομερέστερα στη συνέχεια. (Παρατηρείστε ότι για $\alpha = 1$ παίρνουμε την εκθετική πυκνότητα πιθανότητας.)

Η συνάρτηση κατανομής δεν μπορεί να εκφραστεί μέσω στοιχειωδών συναρτήσεων (εκτός από την περίπτωση που το α είναι φυσικός αριθμός). Εκφράζεται μέσω της μη πλήρους συνάρτησης Γάμμα (incomplete Gamma function) που ορίζεται ως

$$
\Gamma(x;\alpha) := \int_0^x t^{\alpha-1} e^{-t} dt, \qquad x \ge 0.
$$



Σχήμα 1.2: Πυκνότητα Γάμμα για διάφορες τιμές της παραμέτρου σχήματος α και $\lambda = 1$.

Η συνάρτηση αυτή συμπεριλαμβάνεται σε όλα τα προγράμματα στατιστικής ανάλυσης όπως π.χ. η R. Παρατηρείστε ότι $\Gamma(\infty;\alpha) := \lim_{x\to\infty}\Gamma(x;\alpha) = \lim_{\to\infty}\int_0^x t^{\alpha-1}e^{-t}dt = \Gamma(\alpha)$. Χρησιμοποιώντας την μη πλήρη συνάρτηση Γάμμα έχουμε

$$\int_0^x \lambda(\lambda t)^{\alpha-1}e^{-\lambda t}dt = \int_0^{\lambda x} u^{\alpha-1}e^{-u}du = \Gamma(\lambda x;\alpha)$$

και επομένως η συνάρτηση κατανομής εκφράζεται ως

$$F(x) = \begin{cases} \dfrac{\Gamma(\lambda x;\alpha)}{\Gamma(\alpha)} & \text{αν} \quad x > 0 \\ 0 & \text{αν} \quad x \leq 0. \end{cases}$$

Οι ροπές της κατανομής Γάμμα υπολογίζονται εύκολα:

$$\begin{aligned} \mathbb{E}X^n &= \int_{-\infty}^{\infty} x^n f(x)dx = \frac{1}{\Gamma(\alpha)}\int_0^{\infty} x^n\lambda(\lambda x)^{\alpha-1}e^{-\lambda x}dx \\ &= \frac{1}{\lambda^n\Gamma(\alpha)}\int_0^{\infty}\lambda^n x^n\lambda(\lambda x)^{\alpha-1}e^{-\lambda x}dx = \frac{1}{\lambda^n\Gamma(\alpha)}\int_0^{\infty} t^{n+\alpha-1}e^{-t}dt = \frac{1}{\lambda^n}\frac{\Gamma(n+\alpha-1)}{\Gamma(\alpha)}. \end{aligned}$$

Λαμβάνοντας υπ΄ όψιν μας την (1.13) βλέπουμε ότι

$$\mathbb{E}X^n = \frac{1}{\lambda^n}\alpha(\alpha+1)\cdots(\alpha+n-1), \qquad n = 1, 2, \ldots. \tag{1.15}$$

Συγκεκριμένα, οι δύο πρώτες ροπές είναι

$$\mathbb{E}X = \frac{\alpha}{\lambda}, \qquad \mathbb{E}X^2 = \frac{\alpha(\alpha+1)}{\lambda^2}$$

και η διασπορά είναι

$$\mathsf{Var}(X) = \frac{\alpha}{\lambda^2}.$$

1.4 Η κατανομή Erlang

Η κατανομή Erlang είναι μια ειδική περίπτωση της κατανομής Γάμμα όταν η παράμετρος σχήματος είναι ακέραια. Η πυκνότητα πιθανότητας της Erlang δίδεται από τον τύπο

$$f(x) = \begin{cases} \lambda \frac{(\lambda x)^{k-1}}{(k-1)!} e^{-\lambda x} & \text{αν } x \geq 0 \\ 0 & \text{αν } x < 0 \end{cases}.$$

Η παράμετρος k είναι θετικός ακέραιος (όταν $k = 1$ παίρνουμε την εκθετική κατανομή) και η $\lambda > 0$ είναι παράμετρος κλίμακας. Η αντίστοιχη συνάρτηση κατανομής που προκύπτει ολοκληρώνοντας την $f(x)$ δίδεται από τον τύπο

$$F(x) = \begin{cases} 1 - \displaystyle\sum_{m=0}^{k-1} \frac{(\lambda x)^m}{m!} e^{-\lambda x} & \text{αν } x \geq 0 \\ 0 & \text{αν } x < 0 \end{cases}.$$

Παραδείγματος χάριν, όταν $k = 2$, έχουμε $F(x) = 1 - (1 + \lambda x)e^{-\lambda x}$, $(x \geq 0)$.

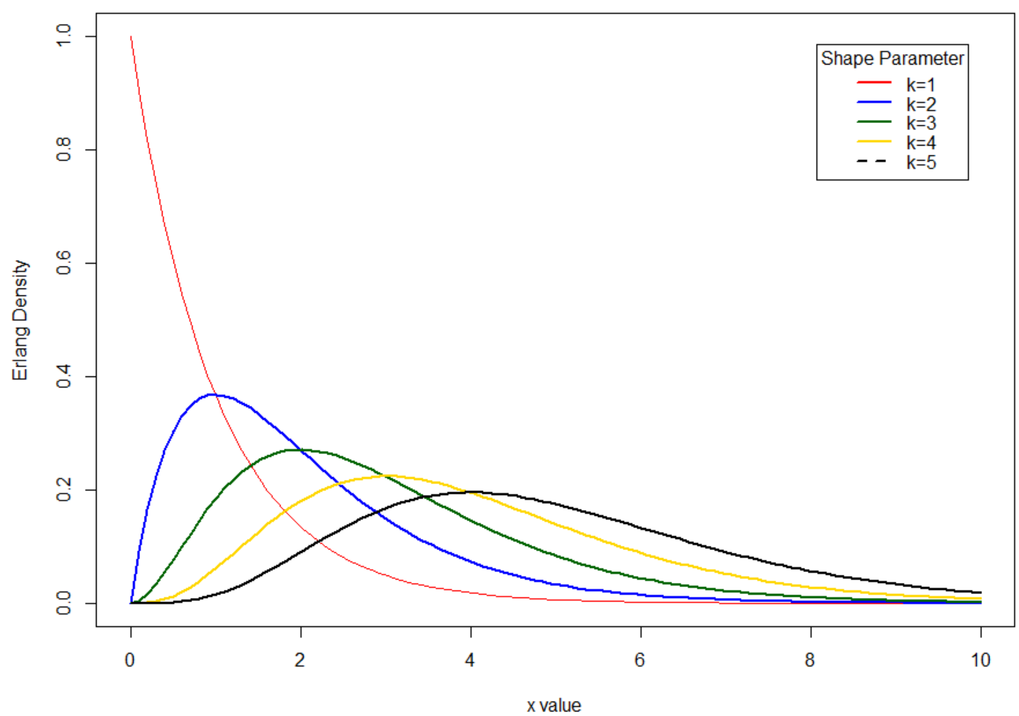
Έστω X τ.μ. με κατανομή Erlang(k, λ). Η ροπή τάξης r θα είναι

$$\mathsf{E}X^r = \int_0^\infty x^r \frac{(\lambda x)^{k-1}}{(k-1)!} e^{-\lambda x} \lambda dx = \frac{1}{(k-1)!\lambda^r} \int_0^\infty y^{r+k-1} e^{-y} dy = \frac{1}{\lambda^r} \frac{(k+r-1)!}{(k-1)!}.$$

Οι δύο πρώτες ροπές της Erlang(k, λ) είναι $\mathbb{E}X = \frac{k}{\lambda}$ και $\mathbb{E}X^2 = \frac{k(k+1)}{\lambda^2}$, και συνεπώς η διασπορά είναι $\mathsf{Var}(X) = \frac{k(k+1)}{\lambda^2} - \frac{k^2}{\lambda^2} = \frac{k}{\lambda^2}$. Ο συντελεστής διακύμανσης που προκύπτει είναι

$$c_V = \frac{\sqrt{k/\lambda^2}}{k/\lambda} = \frac{1}{\sqrt{k}} \leq 1.$$

Βλέπουμε δηλαδή ότι η Erlang έχει πάντα συντελεστή μεταβλητότητας μικρότερο της μονάδας (η ισότητα ισχύει όταν $k = 1$ οπότε παίρνουμε την εκθετική). Όσο μεγαλύτερη είναι η παράμετρος k τόσο μικρότερος γίνεται ο συντελεστής μεταβλητότητας.



Σχήμα 1.3: Πυκνότητα Erlang για διάφορες τιμές της παραμέτρου σχήματος k και $\lambda = 1$.

1.5 Η συνάρτηση Βήτα

Μια άλλη συνάρτηση η οποία σχετίζεται με την συνάρτηση Γάμμα και διαδραματίζει σημαντικό ρόλο είναι η συνάρτηση Βήτα η οποία έχει δύο μη αρνητικά ορίσματα, και η οποία ορίζεται ως

$$B(\alpha,\beta) := \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}dt, \qquad \alpha>0,\ \beta>0. \tag{1.16}$$

Όπως θα δούμε στη συνέχεια η συνάρτηση Βήτα συνδέεται με την Γάμμα μέσω της σχέσης

$$\boxed{B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.} \tag{1.17}$$

Παρατηρούμε ότι η συνάρτηση Βήτα είναι συμμετρική ως προς τα ορίσματά της. Πράγματι

$$B(\alpha,\beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}dt = -\int_1^0 (1-s)^{\alpha}s^{\beta}ds = \int_0^1 s^{\beta}(1-s)^{\alpha}ds = B(\beta,\alpha)$$

όπου στην παραπάνω σχέση κάναμε την αλλαγή μεταβλητής $s = 1-t$. Στην γενική περίπτωση θα δούμε την απόδειξη της (1.17) λίγο αργότερα. Προς το παρόν θα δώσουμε μια απόδειξη στην περίπτωση που ο α είναι φυσικός αριθμός, έστω $\alpha = m \in \mathbb{N}$.

$$\begin{aligned} B(m,\beta) &= \int_0^1 t^{m-1}(1-t)^{\beta-1}dt = -\frac{1}{\beta}\int_0^1 t^{m-1}\left((1-t)^{\beta}\right)'dt \\ &= -\frac{1}{\beta}\left(t^{m-1}(1-t)^{\beta}\Big|_0^1 - (m-1)\int_0^1 t^{m-2}(1-t)^{\beta}dt\right) = \frac{m-1}{\beta}\,B(m-1,\beta+1) \end{aligned}$$

Εφαρμόζοντας την παραπάνω σχέση επαγωγικά έχουμε

$$B(m,\beta) = \frac{m-1}{\beta}\,B(m-1,\beta+1) = \frac{(m-1)(m-2)}{\beta(\beta+1)}\,B(m-2,\beta+2) = \cdots \tag{1.18}$$

$$= \frac{(m-1)(m-2)\cdots 2\cdot 1}{\beta(\beta+1)\cdots(\beta+m-2)}\,B(1,\beta+m-1). \tag{1.19}$$

Όμως

$$B(1,\beta+m-1) = \int_0^1 t^0(1-t)^{\beta+m-2}dt = \frac{1}{\beta+m-1}$$

και επομένως

$$\begin{aligned} B(m,\beta) &= \frac{(m-1)(m-2)\cdots 2\cdot 1}{\beta(\beta+1)\cdots(\beta+m-2)(\beta+m-1)} = \frac{(m-1)!\Gamma(\beta)}{\Gamma(\beta)\,\beta(\beta+1)\cdots(\beta+m-2)(\beta+m-1)} \\ &= \frac{\Gamma(m)\Gamma(\beta)}{\Gamma(\beta+m)} \end{aligned}$$

όπου χρησιμοποιήσαμε το γεγονός ότι $\Gamma(\beta+m) = (\beta+m-1)(\beta+m-2)\cdots(\beta+1)\beta\,\Gamma(\beta)$ (1.13).

Από την (1.17) παίρνουμε και το εξής ενδιαφέρον αποτέλεσμα: Αν $\alpha = \beta = 1/2$ θα έχουμε

$$\mathrm{B}(1/2, 1/2) = \frac{\Gamma(1/2)^2}{\Gamma(1)} \tag{1.20}$$

Όμως $\Gamma(1) = 1$ και $\mathrm{B}(1/2, 1/2) = \int_0^1 t^{-1/2}(1-t)^{-1/2}dt$. Με την αλλαγή μεταβλητής $t = \sin^2 u$, $dt = 2\sin u \cos u\, du$ έχουμε

$$\begin{aligned} \mathrm{B}(1/2, 1/2) &= \int_0^1 (\sin u)^{-1}(1-\sin^2 u)^{-1/2}\, 2\sin u \cos u du \\ &= \int_0^{\pi/2} (\sin u)^{-1}(\cos u)^{-1}\, 2\sin u \cos u du \\ &= \int_0^{\pi/2} 2du = \pi \end{aligned}$$

Από την σχέση αυτή και την (1.20) παίρνουμε

$$\Gamma(1/2)^2 = \pi \qquad \text{ή} \qquad \Gamma(1/2) = \sqrt{\pi}. \tag{1.21}$$

1.6 Η κατανομή Βήτα

Η κατανομή Βήτα έχει πυκνότητα πιθανότητας

$$f(x) = \begin{cases} \dfrac{1}{\mathrm{B}(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1} & \text{αν } \ 0 < x < 1 \\ 0 & \text{διαφορετικά} \end{cases}. \tag{1.22}$$

Όπως βλέπουμε και από το σχήμα (1.4) η μορφή των πυκνοτήτων ποικίλει ανάλογα με τις τιμές των παραμέτρων $\alpha > 0$ και $\beta > 0$. Είναι εύκολο να δούμε ότι αν $\alpha = \beta = 1$ τότε η πυκνότητα πιθανότητας γίνεται

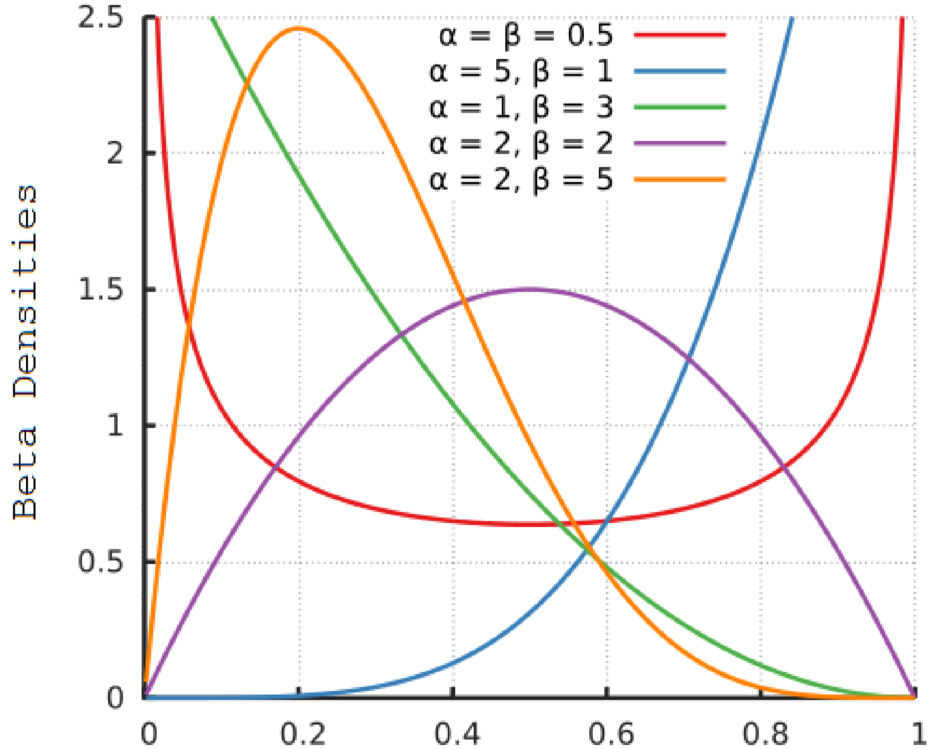
$$f(x) = \begin{cases} 1 & \text{αν } \ 0 < x < 1 \\ 0 & \text{διαφορετικά} \end{cases}$$

δηλαδή ομοιόμορφη στο $[0, 1]$.

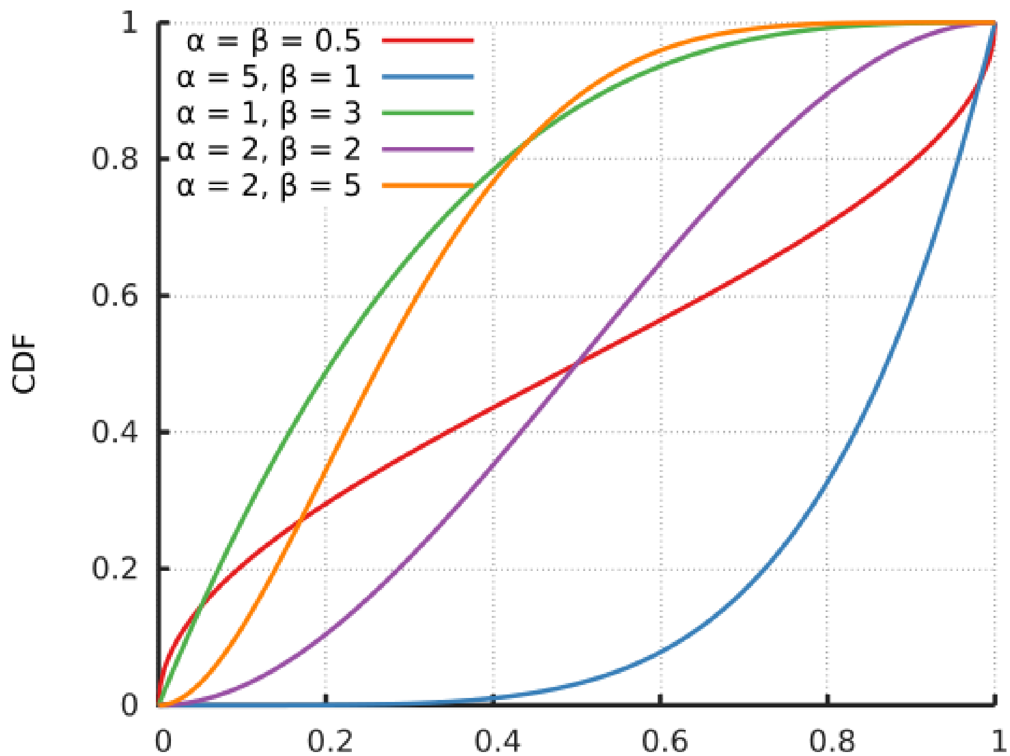
Το σχήμα (1.5) δίνει τις αντίστοιχες συναρτήσεις κατανομής.

Οι ροπές της κατανομής Βήτα υπολογίζονται εύκολα ως εξής:

$$\begin{aligned} \mathbb{E}X^n &= \frac{1}{\mathrm{B}(\alpha,\beta)} \int_0^1 x^n x^{\alpha-1}(1-x)^{\beta-1}dx = \frac{\mathrm{B}(\alpha+n,\beta)}{\mathrm{B}(\alpha,\beta)} \\ &= \frac{\Gamma(\alpha+n)\Gamma(\beta)}{\Gamma(\alpha+\beta+n)} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} = \frac{\Gamma(\alpha+n)}{\Gamma(\alpha)} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha+\beta+n)}. \end{aligned}$$



Σχήμα 1.4: Πυκνότητες Βήτα για διάφορες τιμές των παραμέτρων α και β.



Σχήμα 1.5: Συναρτήσεις κατανομής Βήτα για διάφορες τιμές των παραμέτρων α και β.

Από την τελευταία αυτή έκφραση, χρησιμοποιώντας την (1.13), παίρνουμε

$$\mathbb{E}X^n = \frac{\alpha \cdot (\alpha+1) \cdots (\alpha+n-1)}{(\alpha+\beta)\cdot(\alpha+\beta+1)\cdots(\alpha+\beta+n-1)}, \qquad . \tag{1.23}$$

Συγκεκριμένα οι δύο πρώτες ροπές είναι

$$\mathbb{E}X = \frac{\alpha}{\alpha+\beta}, \qquad \mathbb{E}X^2 = \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)},$$

και η δισπορά είναι

$$\mathsf{Var}(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.$$

1.7 Ροπογεννήτριες συναρτήσεις

Έστω X τ.μ. με συνάρτηση πυκνότητας πιθανότητας f. Η ροπογεννήτρια της X ορίζεται ως η συνάρτηση

$$M_X(\theta) := E[e^{\theta X}] = \int_{-\infty}^{\infty} e^{\theta x} f(x)dx \tag{1.24}$$

για κάθε $\theta \in \mathbb{R}$ για το οποίο το παραπάνω ολοκλήρωμα συγκλίνει. Ακολουθούν παραδείγματα ροπογεννητριών:

Η ροπογεννήτρια συνάρτηση μιας κανονικής τ.μ. Έστω X κανονική τ.μ. με μέση τιμή 0 και διασπορά 1. Η ροπογεννήτρια της X ορίζεται ως

$$M_X(\theta) := Ee^{\theta X} = \int_{-\infty}^{\infty} e^{\theta x} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\theta x - \frac{x^2}{2}} dx.$$

Ο εκθέτης μέσα στο ολοκλήρωμα γράφεται ως $-\frac{1}{2}\left(x^2 - 2\theta x + t^2\right) + \frac{\theta^2}{2}$ και επομένως έχουμε

$$M_X(\theta) = e^{\frac{\theta^2}{2}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x-\theta)^2} dx = e^{\frac{\theta^2}{2}} \tag{1.25}$$

αφού το τελευταίο ολοκλήρωμα ισούται με την μονάδα.

Τώρα ας υποθέσουμε ότι η Y είναι κανονική τ.μ. με μέση τιμή μ και διασπορά σ. Τότε έχουμε $Y = \mu + \sigma X$ και επομένως η ροπογεννήτρια σ′ αυτή την περίπτωση είναι

$$M_Y(\theta) = \mathbb{E}e^{\theta Y} = \mathbb{E}e^{\theta\mu + \theta\sigma X} = e^{t\theta\mu} M_X(\sigma\theta) = e^{\mu\theta + \frac{\sigma^2\theta^2}{2}}.$$

Η ροπογεννήτρια συνάρτηση μιας εκθετικής τ.μ. Έστω X εκθετική τ.μ. με πυκνότητα πιθανότητας $f(x) = \lambda e^{-\lambda x}$ για $x \geq 0$, και $f(x) = 0$ όταν $x < 0$. Η ροπογεννήτριά της δίνεται από την σχέση

$$M_X(\theta) = \int_0^{\infty} e^{\theta x} \lambda e^{-x\lambda} dx = \lambda \int_0^{\infty} e^{-x(\lambda-\theta)} dx = \frac{\lambda}{\lambda - \theta}. \tag{1.26}$$

Παρατηρείστε ότι ο παραπάνω υπολογισμός έχει νόημα μόνον όταν $\theta < \lambda$. Αν το $\theta \geq \lambda$ το ολοκλήρωμα απειρίζεται και η ροπογεννήτρια δεν ορίζεται.

Η ροπογεννήτρια συνάρτηση μιας ομοιόμορφης τ.μ. Ας υποθέσουμε τώρα ότι η X είναι ομοιόμορφα κατανεμημένη στο διάστημα $[a, b]$ με πυκνότητα πιθανότητας $f(x) = \frac{1}{b-a}$ όταν $x \in [a, b]$ και 0 αλλού. Η αντίστοιχη ροπογεννήτρια δίδεται από τη σχέση

$$M_X(\theta) \;=\; \int_a^b e^{\theta x}\frac{1}{b-a}dx \;=\; \frac{e^{\theta b}-e^{\theta a}}{\theta(b-a)}. \tag{1.27}$$

Η ροπογεννήτρια συνάρτηση μιας τ.μ. Γάμμα Έστω X μια τυχαία μεταβλητή με κατανομή Γαμμα(α, λ) δηλαδή με πυκνότητα $f(x) = \lambda\frac{(\lambda x)^{\alpha-1}}{\Gamma(\alpha)}\,e^{-\lambda x}$ όταν $x > 0$ (και 0 διαφορετικά). Η ροπογεννήτρια δίνεται από την έκφραση

$$\begin{aligned}
M(\theta) &= \int_0^\infty e^{\theta x}\lambda\frac{(\lambda x)^{\alpha-1}}{\Gamma(\alpha)}\,e^{-\lambda x}dx \;=\; \frac{\lambda^\alpha}{\Gamma(\alpha)}\int_0^\infty x^{-\alpha-1}\,e^{-x(\lambda-\theta)}dx \\
&= \left(\frac{\lambda}{\lambda-\theta}\right)^\alpha \frac{1}{\Gamma(\alpha)}\int_0^\infty y^{\alpha-1}\,e^{-y}dy \qquad y = x(\lambda-\theta) \\
&= \left(\frac{\lambda}{\lambda-\theta}\right)^\alpha \text{υπό την προϋπόθεση ότι } \theta < \lambda. \qquad (1.28)
\end{aligned}$$

(Αν $\theta \geq \lambda$ το ολοκλήρωμα αποκλίνει και η πιθανογεννήτρια δεν υπάρχει.)

Η ροπογεννήτριες συναρτήσεις παίρνουν το όνομά τους από το την ακόλουθη ιδιότητα: Παραγωγίζοντας την ροπογεννήτρια n φορές έχουμε

$$\frac{d^n}{d\theta^n}M_X(\theta) = \frac{d^n}{d\theta^n}\int_{-\infty}^\infty e^{\theta x}f(x)dx = \int_{-\infty}^\infty \frac{d^n}{d\theta^n}e^{\theta x}f(x)dx = \int_{-\infty}^\infty x^n e^{\theta x}f(x)dx,$$

(Στην παραπάνω εξίσωση υποθέσαμε ότι η εναλλαγή ολοκλήρωσης και παραγώγισης είναι θεμιτή. Δεν θα επεκταθούμε σ΄ αυτό το ζήτημα.) Θέτοντας στην παραπάνω εξίσωση $\theta = 0$, και συμβολίζοντας με $M_X^{(n)}(\theta)$ την n-οστή παράγωγο της M_X έχουμε

$$M_X^{(n)}(0) = \int_{-\infty}^\infty x^n f(x)dx = \mu_n$$

δηλαδή παρατηρούμε ότι η n-οστή παράγωγος της ροπογεννήτριας υπολογισμένη στο 0 συμπίπτει με την n-οστή ροπή της κατανομής. Συνεπώς γνώση της ροπογεννήτριας συνεπάγεται γνώση όλων των ροπών της κατανομής. Ισχύει και το αντίστροφο, δηλαδή ότι οι ροπές προσδιορίζουν την ροπογεννήτρια: Πράγματι, από το θεώρημα του Taylor γνωρίζουμε ότι κάθε συνάρτηση που είναι αναλυτική σε μια περιοχή του μηδενός μπορεί να εκφραστεί ως μια δυναμοσειρά:

$$M_X(\theta) = \sum_{n=0}^\infty \frac{\theta^n}{n!}M_X^{(n)}(0) = \sum_{n=0}^\infty \frac{\theta^n}{n!}\mu_n.$$

Για την τυποποιημένη κανονική κατανομή έχουμε

$$M(\theta) \;=\; e^{-\frac{1}{2}\theta^2} \;=\; 1 + \frac{\theta^2}{2\cdot 1!} + \frac{\theta^4}{2^2\cdot 2!} + \frac{\theta^6}{2^3\cdot 3!} + \cdots + \frac{\theta^{2n}}{2^n\cdot n!} + \cdots$$

Παρατηρούμε ότι οι περιττές δυνάμεις του θ απουσιάζουν πράγμα που σημαίνει ότι οι ροπές περιττής τάξης είναι 0. Για τις ροπές άρτιας τάξης έχουμε

$$\mu_{2n} \;=\; \frac{(2n)!}{2^n\cdot n!} \;=\; \binom{2n}{n}\frac{n!}{2^n} \;=\; 1\cdot 3\cdot 5\cdots(2n-1). \tag{1.29}$$

Για παράδειγμα

$$\mu_2 = 1, \quad \mu_4 = 1\cdot 3, \quad \mu_6 = 1\cdot 3\cdot 5, \quad \mu_8 = 1\cdot 3\cdot 5\cdot 7.$$

Είναι ενδιαφέρον να παρατηρήσουμε ότι η έκφραση για την ροπή τάξης $2n$ που δίδεται στην (1.29) είναι η λύση του ακόλουθου προβλήματος συνδιαστικής ανάλυσης: Με πόσους διαφορετικούς τρόπους μπορούμε να χωρίσουμε $2n$ διαφορετικά αντικείμενα σε n ζευγάρια;

Το πρώτο ζεύγος μπορούμε να το επιλέξουμε με $\binom{2n}{2}$. Το δεύτερο ζευγάρι με $\binom{2n-2}{2}$ (επειδή μετά την επιλογή του πρώτου ζεύγους έχουν μείνει $2n-2$ αντικείμενα). Το τρίτο ζευγάρι με $\binom{2n-4}{2}$ τρόπους και ούτω καθ′ εξής. Το προτελευταίο ($n-1$-οστό) ζευγάρι με $\binom{4}{2}$ τρόπους και το τελευταίο με $\binom{2}{2=1}$ τρόπους. Συνεπώς μπορούμε να φτιάξουμε n αριθμημένα (πρώτο, δεύτερο, τρίτο κ.ο.κ.) ζευγάρια από $2n$ αντικείμενα με

$$\begin{aligned}&\binom{2n}{2}\cdot\binom{2n-2}{2}\cdot\binom{2n-4}{2}\cdots\binom{4}{2}\cdot\binom{2}{2}\\ &\quad= \frac{2n(2n-1)}{2}\cdot\frac{(2n-2)(2n-3)}{2}\cdot\frac{(2n-4)(2n-5)}{2}\cdots\frac{4\cdot 3}{2}\cdot\frac{2\cdot 1}{2} \;=\; \frac{(2n)!}{2^n}.\end{aligned}$$

Δεδομένου όμως ότι δεν μας ενδιαφέρει η σειρά των ζευγαριών και ότι μπορούμε να διατάξουμε n ζευγάρια με $n!$ τρόπους, ο συνολικός αριθμός των τρόπων με τους οποίους μπορούμε να φτιάξουμε n ζευγάρια από $2n$ διαφορετικά αντικείμενα είναι $\frac{(2n)!}{2^n n!}$. Αυτό το αποτέλεσμα συμφωνεί με τον τύπο (1.29).

1.8 Παράμετροι θέσης και κλίμακας

Θεώρημα 1.1. *Έστω X και Y τυχαίες μεταβλητές με συναρτήσεις κατανομής F και G αντίστοιχα. Αν*

$$Y = \mu + \sigma X$$

όπου $\mu \in \mathbb{R}$, $\sigma > 0$ τότε

$$G(x) = F\left(\frac{x-\mu}{\sigma}\right). \tag{1.30}$$

Αν επιπλέον η κατανομή F *είναι συνεχής με πυκνότητα πιθανότητας* f *τότε και η κατανομή* G *είναι συνεχής με πυκνότητα πιθανότητας*

$$g(x) = \frac{1}{\sigma} f\left(\frac{x-\mu}{\sigma}\right). \tag{1.31}$$

Απόδειξη Πράγματι

$$G(x) = \mathbb{P}(Y \le x) = \mathbb{P}\left(\mu + \sigma X \le x\right) = \mathbb{P}\left(X \le \frac{x-\mu}{\sigma}\right) = F\left(\frac{x-\mu}{\sigma}\right)$$

Παραγωγίζοντας την (1.30) ως προς x παίρνουμε την (1.31). ∎

Αν η συνάρτηση κατανομής μιας τ.μ. X μπορεί να γραφτεί υπό την μορφή

$$F(x) = G\left(\frac{x-\mu}{\sigma}\right)$$

όπου $\mu \in \mathbb{R}$, $\sigma > 0$, και G είναι κάποια άλλη συνάρτηση τότε το μ ονομάζεται παράμετρος θέσης και το σ παράμετρος κλίμακας. Η συνάρτηση G είναι τότε $G(x) = F(\mu + \sigma x)$ και κατά συνέπεια είναι επίσης μια συνάρτηση κατανομής.

1.9 Συναρτήσεις τυχαίων μεταβλητών

Έστω X πραγματική τυχαία μεταβλητή με συνάρτηση κατανομής $F_X(x) = \mathbb{P}(X \le x)$ και πυκνότητα πιθανότητας $f_X(x) = F'(x)$, $x \in \mathbb{R}$. Έστω επίσης $g : \mathbb{R} \to \mathbb{R}$ μια πραγματική συνάρτηση. Τότε η $Y := g(X)$ είναι τυχαία μεταβλητή. Η συνάρτηση κατανομής της Y, δίδεται από την σχέση

$$F_Y(y) \;:=\; \mathbb{P}(Y \le y) \;=\; \mathbb{P}(g(X) \le y). \tag{1.32}$$

Δεν θα εξετάσουμε το πρόβλημα στην πλήρη γενικότητά του εδώ. Θα περιοριστούμε στις εξής δύο σημαντικές περιπτώσεις.

1.9.1 Η g είναι γνησίως μονοτονική, παραγωγίσιμη συνάρτηση

Έστω $g : I \to J$ όπου I, J ανοικτά υποσύνολα του $\mathbb{R}$. Εφ' όσον η g είναι γνησίως μονοτονική, η αντίστροφη συνάρτηση $g^{-1} : J \to I$ υπάρχει και είναι παραγωγίσιμη με

$$(g^{-1})'(y) = \frac{1}{g'(g^{-1}(y))} \qquad \text{για κάθε } y \in J.$$

Αν η g είναι γνησίως αύξουσα, τότε και η g^{-1} είναι γνησίως αύξουσα και επομένως η (1.32) δίδει

$$F_Y(y) \;=\; \mathbb{P}(X \le g^{-1}(y)) = F_X(g^{-1}(y)). \qquad (g \text{ γνησίως αύξουσα})$$

Συνεπώς, η πυκνότητα πιθανότητας της Y δίνεται από την σχέση

$$f_Y(y) = \frac{d}{dy} F_Y(y) = f_X(g^{-1}(y)) \, \frac{1}{g'(g^{-1}(y))}. \tag{1.33}$$

Στην περίπτωση που η g είναι γνησίως φθίνουσα το ίδιο ισχύει και για την g^{-1} και επομένως

$$F_Y(y) \;=\; \mathbb{P}(X \geq g^{-1}(y)) = 1 - F_X(g^{-1}(y)). \qquad (g \text{ γνησίως φθίνουσα})$$

Παραγωγίζοντας παίρνουμε σ' αυτή την περίπτωση

$$f_Y(y) = \frac{d}{dy} F_Y(y) = -f_X(g^{-1}(y)) \, \frac{1}{g'(g^{-1}(y))}. \tag{1.34}$$

Παρατηρούμε όμως ότι $g'(g^{-1}(y)) < 0$ στην περίπτωση αυτή. Συνδιάζοντας τις (1.33), (1.34) έχουμε, ανεξαρτήτως αν η g είναι αύξουσα ή φθίνουσα,

$$f_Y(y) \;=\; f_X(g^{-1}(y)) \, \frac{1}{\mid g'(g^{-1}(y)) \mid}. \tag{1.35}$$

Παράδειγμα 1. Έστω ότι η τυχαία μεταβλητή X είναι ομοιόμορφα κατανεμημένη στο διάστημα $(0,1)$, δηλαδή $f_X(x) = \mathbf{1}(0 < x < 1)$ και $g : (0,1) \to (0,\infty)$ είναι η συνάρτηση $g(x) = -\log(1-x)$. (Παρατηρούμε ότι η g είναι αύξουσα.) Αν $y = -\log(1-x)$ τότε $x = 1 - e^{-y}$ δηλαδή $g^{-1}(y) = 1 - e{-}y$. Επίσης $g'(x) = \frac{1}{1-x}$ και $g'(g^{-1}(y)) = \frac{1}{1-(1-e^{-y})} = e^{y}$. Συνεπώς από την (1.35) έχουμε

$$f_y(y) \;=\; f_X(1-e^{-y}) \frac{1}{e^{y}} \;=\; \mathbf{1}(0 < 1 - e^{-y})\, e^{-y} \;=\; e^{-y} \qquad \text{για } y > 0$$

δηλαδή η Y είναι εκθετικά κατανεμημένη τυχαία μεταβλητή με ρυθμό 1.

Παράδειγμα 2. Όπως θα δούμε είναι απλούστερο να ακολουθούμε την διαδικασία από την αρχή αντί να εφαρμόζουμε τον τύπο (1.35). Είναι όμως σημαντικό να κατανούμε ποιοτικά την συμπεριφορά της συνάρτησης g (πεδίο ορισμου, πεδίο τιμών και μονοτονικότητα). Έστω και πάλι ότι η X είναι ομοιόμορφη στο $(0,1)$ και $g(x) = \tan(\pi(x - \frac{1}{2}$.

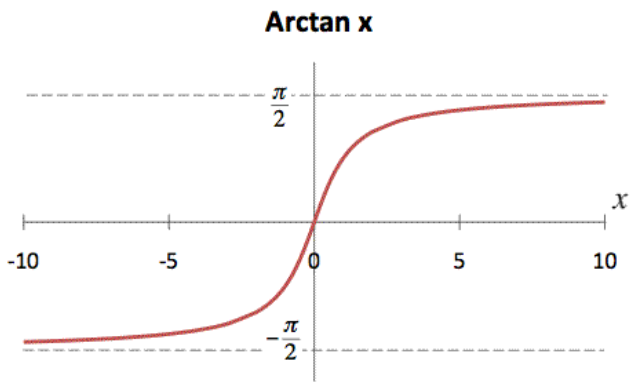
$g(x) = \arctan(\pi(x - \frac{1}{2})$, $x \in \mathbb{R}$. Στο σχήμα 1.6 βλέπουμε το γράφημα της arctan.

Η συνάρτηση είναι γνησίως αύξουσα και το πεδίο τιμών είναι το σύνολο $(-\frac{\pi}{2}, \frac{\pi}{2})$. Επίσης γνωρίζουμε ότι $\frac{d}{dx}\arctan(x) = \frac{1}{1+x^2}$. Αν $Y = \arctan(X)$ τότε

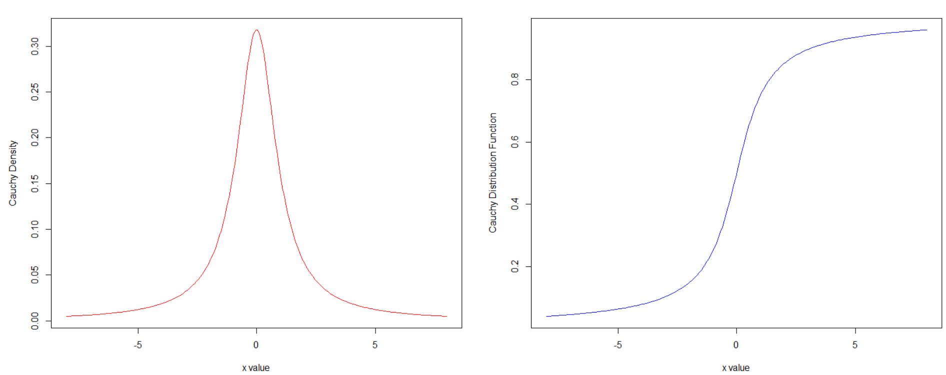
$$\begin{aligned} F_y(y) \;&=\; \mathbb{P}(\tan(\pi(X - \tfrac{1}{2})) \leq y) \;=\; \mathbb{P}(\pi(X - \tfrac{1}{2}) \leq \arctan(y)) \;=\; \mathbb{P}(X \leq \tfrac{1}{2} + \tfrac{1}{\pi}\arctan(y)) \\ &=\; \frac{1}{2} + \frac{1}{\pi}\arctan(y) \end{aligned}$$

(αφού η συνάρτηση κατανομής της X είναι $F_X(x) = x$ όταν $x \in (0,1)$.) Παραγωγίζοντας την παραπάνω σχέση

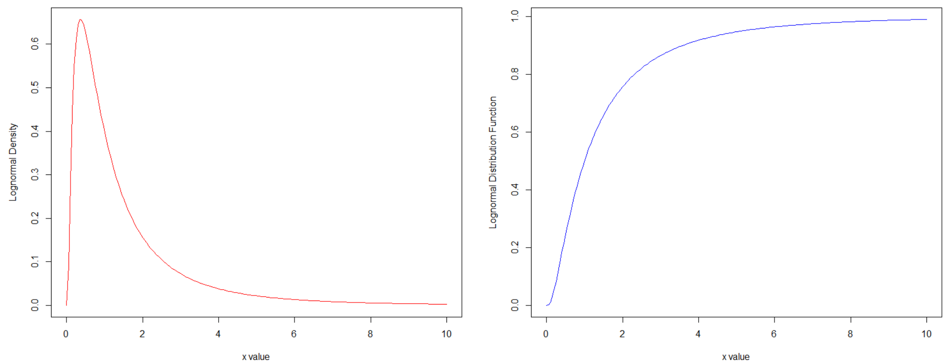
$$f_Y(y) \;=\; \frac{d}{dy}\left(\frac{1}{2} + \frac{1}{\pi}\arctan(y)\right) = \frac{1}{\pi}\frac{1}{1+y^2}, \qquad y \in \mathbb{R}.$$



Σχήμα 1.6: Η συνάρτηση arctan.



Σχήμα 1.7: Πυκνότητα και συνάρτηση κατανομής Cauchy.



Σχήμα 1.8: Λογαριθμοκανονική πυκνότητα και συνάρτηση κατανομής.

Η κατανομή Cauchy δεν έχει ροπές, ούτε καν μέση τιμή. Ο λόγος είναι ότι το ολοκλήρωμα που ορίζει την μέση τιμή δεν συγκλίνει!

Παράδειγμα 3. Η λογαριθμοκανονική κατανομή. Αν η X είναι τυποποιημένη κανονική τυχαία μεταβλητή η $Y := e^X$ ονομάζεται τυποποιημένη λογαριθμοκανονική τυχαία μεταβλητή. Η συνάρτηση κατανομής της X είναι βεβαίως η $\Phi(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2} dt$. Η συνάρτηση κατανομής της Y είναι

$$F_Y(y) = \mathbb{P}(Y \le y) = \mathbb{P}(e^X \le y) = \mathbb{P}(X \le \log(y)) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\log y} e^{-\frac{1}{2}t^2} dt.$$

Παραγωγίζοντας την παραπάνω σχέση παίρνουμε

$$f_Y(y) = \frac{1}{y\sqrt{2\pi}} e^{-\frac{1}{2}(\log y)^2}, \qquad y > 0.$$

Οι ροπές της λογαριμθμοκανονικής κατανομής μπορούν να υπολογισθούν εύκολα από την σχέση που την ορίζει.

$$\mathbb{E}[Y^n] \;=\; \mathbb{E}[(e^X)^n] \;=\; \mathbb{E}[e^{nX}] \;=\; e^{\frac{1}{2}n^2}$$

Στην παραπάνω σχέση χρησιμοποιήσαμε την ροπογεννήτρια της τυποποιημένης κανονικής κατανομής. Συνεπώς

$$\mathbb{E}[Y] = e^{\frac{1}{2}}, \quad \mathbb{E}[Y^2] = e^2, \quad \mathrm{Var}(Y) = e^2 - e = e(e-1).$$

Παράδειγμα 4. Έστω X τυποποιημένη κανονική τυχαία μεταβλητή με κατανομή $\Phi(x) := \mathbb{P}(X \le x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du$ και πυκνότητα πιθανότητας $\varphi(x) := \Phi'(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$. Έστω $Y = X^2$. Θέλουμε να βρούμε την συνάρτηση κατανομής και την πυκνότητα πιθανότητας της Y. Παρατηρείστε ότι $Y = g(X)$ όπου η συνάρτηση $g(x) = x^2$ δεν είναι μονοτονική. Έστω $F(y) := \mathbb{P}(Y \le y) = \mathbb{P}(X^2 \le y)$. Αν $y \le 0$ τότε $F(y) = 0$. Αν $y > 0$ τότε

$$F(y) = \mathbb{P}(-\sqrt{y} \le X \le \sqrt{y}) = \Phi(\sqrt{y}) - \Phi(-\sqrt{y}).$$

Η πυκνότητα της τυχαίας μεταβλητής Y για $y > 0$ θα είναι επομένως

$$\begin{aligned} f(y) &= F'(y) = (\Phi(\sqrt{y}) - \Phi(-\sqrt{y}))' = \frac{1}{2\sqrt{y}}\frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}(\sqrt{y})^2} + \frac{1}{2\sqrt{y}}\frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}(-\sqrt{y})^2} \\ &= \frac{1}{\sqrt{2\pi}} y^{-\frac{1}{2}}\, e^{-\frac{y}{2}} = \frac{1}{\sqrt{\pi}}(y/2)^{-\frac{1}{2}}\, e^{-(y/2)} = \frac{(y/2)^{-\frac{1}{2}}}{\Gamma(1/2)}\, e^{-(y/2)} \end{aligned}$$

Στην παραπάνω σχέση χρησιμοποιήσαμε την (1.21). Παρατηρείστε ότι η Y έχει κατανομή Γαμμα$(1/2, 1/2)$.

Παράδειγμα 5. Έστω X ομοιόμορφη τυχαία μεταβλητή στο $[0, 1]$ και $Y = g(X)$ όπου $g(x) = x(1-x)$. Παρατηρούμε ότι η g δεν είναι μονοτονική και στο διάστημα $[0, 1]$, όπου παίρνει τιμές η X, για κάθε τιμή $y \in [0, \frac{1}{4}]$ η εξίσωση $y = x(1-x)$ έχει δύο λύσεις, $x = \frac{1+\sqrt{1-4y}}{2}$ και $x = \frac{1-\sqrt{1-4y}}{2}$. Παρατηρούμε επίσης ότι το μέγιστο της συνάρτησης g βρίσκεται στο σημείο $x = 1/2$ και είναι $g(1/2) = 1/4$. Συνεπώς $\mathbb{P}(Y \le y) = \mathbb{P}(X(1-X) \le y) = 1$ όταν $y \ge 1/4$. Επίσης, δεδομένου ότι η X παίρνει τιμές στο $[0, 1]$, $\mathbb{P}(Y \le y) = \mathbb{P}(X(1-X) \le y) = 0$ όταν $y \le 0$ αφού η $x(1-x)$ είναι μη αρντητική για $x \in [0, 1]$. Μένει να υπολογίσουμε την πιθανότητα $\mathbb{P}(Y \le y)$ όταν $y \in (0, \frac{1}{4})$.

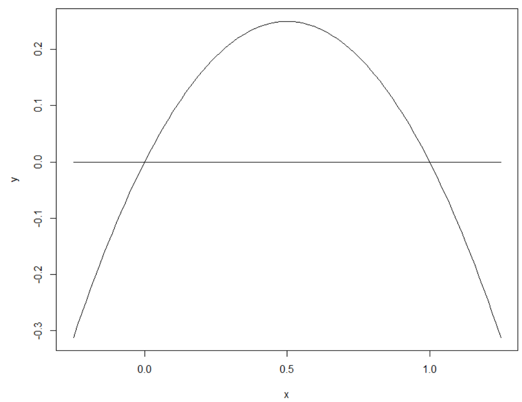
$$\begin{aligned} 1 - F_Y(y) &:= \mathbb{P}(Y > y) = \mathbb{P}(X(1-X) > y) = \mathbb{P}\left(\frac{1-\sqrt{1-4y}}{2} < X < \frac{1+\sqrt{1-4y}}{2}\right) \\ &= \frac{1+\sqrt{1-4y}}{2} - \frac{1-\sqrt{1-4y}}{2} \qquad \text{επειδή η X είναι ομοιόμορφη στο } [0,1] \\ &= \sqrt{1-4y}. \end{aligned}$$

Συνεπώς η συνάρτηση κατανομής της Y είναι

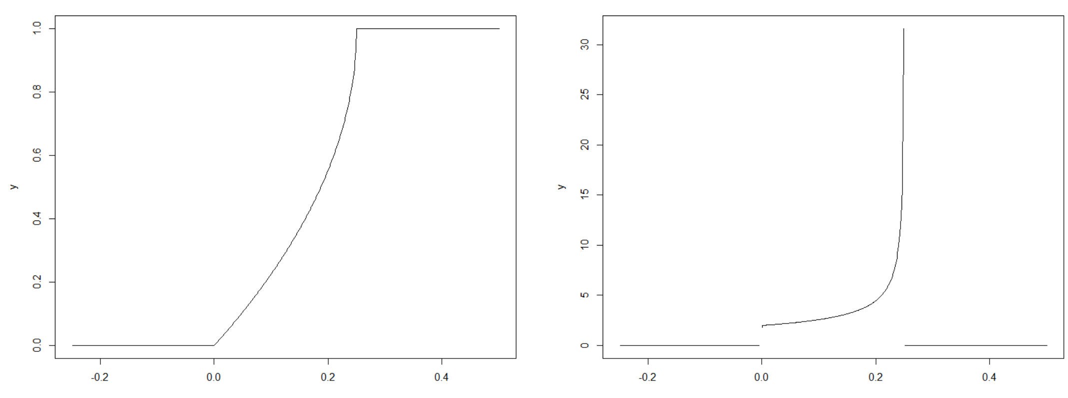
$$F_Y(y) = \begin{cases} 0 & \text{αν} \quad y \le 0 \\ 1 - \sqrt{1-4y} & \text{αν} \quad 0 < y \le \frac{1}{4} \\ 1 & \text{αν} \quad y > \frac{1}{4}. \end{cases}$$

και η πυκνότητα πιθανότητας

$$f_Y(y) = \begin{cases} \frac{2}{\sqrt{1-4y}} & \text{αν} \quad 0 < y \le \frac{1}{4} \\ 0 & \text{διαφορετικά.} \end{cases}$$



Σχήμα 1.9: Η συνάρτηση $g(x) = x(1 - x)$.



Σχήμα 1.10: Η συνάρτηση κατανομής και η πυκνότητα πιθανότητας της Y.

Κεφάλαιο 2

Από κοινού κατανομή τυχαίων μεταβλητών

2.1 Από κοινού κατανομή δύο τυχαίων μεταβλητών

Έστω δυο τυχαίες μεταβλητές X, Y, με τιμές στο $\mathbb{N}_0$ ορισμένες πάνω στο ίδιο χώρο πιθανοτήτων $(\Omega, \mathscr{A}, P)$. Το σύνολο των πιθανοτήτων

$$p_{X,Y}(i,j) := P(X = i, Y = j), \qquad i, j \in \mathbb{N}_0,$$

ονομάζεται *από κοινού κατανομή* των X, Y. Ισχύει ασφαλώς ότι $\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} p_{XY}(i,j) = 1$. Οι

$$\begin{aligned} p_X(i) &:= P(X = i) = \sum_{j=0}^{\infty} p_{X,Y}(i,j) \\ p_Y(j) &:= P(Y = j) = \sum_{i=0}^{\infty} p_{X,Y}(i,j) \end{aligned}$$

ονομάζονται *περιθώριες* κατανομές των X και Y αντίστοιχα. Τα παραπάνω αθροίσματα γράφονται ως άπειρα αθροίσματα μη αρνητικών αριθμών, σε πολλές περιπτώσεις όμως είναι πεπερασμένα αθροίσματα αφού όλοι οι όροι πλην πεπερασμένων μπορεί να είναι μηδέν.

Παράδειγμα 2.1. Θεωρούμε τις τυχαίες μεταβλητές X με τιμές $\{1, 2, 3, 4\}$ και Y με τιμές $\{1, 2\}$. Έστω ότι η από κοινού κατανομή τους δίνεται από τον πίνακα.

Y/X	1	2	3	4
1	1/12	1/12	1/12	1/6
2	1/6	1/12	0	2/6

Οι περιθώριες πιθανότητες δίνονται από τις

$$p_X(1) = 1/4,\ p_X(2) = 1/6,\ p_X(3) = 1/12,\ p_X(4) = 1/2, \qquad p_Y(1) = 5/12,\ p_Y(2) = 7/12.$$

Ορισμός 2.1. Δύο τυχαίες μεταβλητές, X, Y, ονομάζονται *ανεξάρτητες* αν,

$$P(X = i, Y = j) = P(X = i)P(Y = j), \qquad \text{για κάθε } i, j.$$

Ο παραπάνω ορισμός μπορεί επίσης να διατυπωθεί λέγοντας ότι *η από κοινού κατανομή είναι το γινόμενο των περιθωρίων*. Αυτό σημαίνει ότι, για να είναι οι δύο τυχαίες μεταβλητές ανεξάρτητες, θα πρέπει όλα τα ενδεχόμενα της μορφής $\{X = i\}$, $\{Y = j\}$, να είναι ανεξάρτητα.

Όταν οι τυχαίες μεταβλητές X, Y, δεν είναι ανεξάρτητες τότε θα ονομάζονται εξαρτημένες.

Για παράδειγμα οι τυχαίες μεταβλητές με την από κοινού κατανομή που δίνεται στον πίνακα 2.1 δεν είναι ανεξάρτητες αφού $P(X = 1, Y = 1) = 1/6$ ενώ $P(X = 1)P(Y = 1) = 1/3 \times 5/12 = 5/36$.

Ορισμός 2.2. Η δεσμευμένη κατανομή της τυχαίας μεταβλητής X, δεδομένης της Y δίδεται από την σχέση

$$P(X = i|Y = j) \ = \ \frac{P(X = i, Y = j)}{P(Y = j)}.$$

Έστω $f : \mathbb{N}\times\mathbb{N} \longrightarrow \mathbb{R}$, μία συνάρτηση δύο μεταβλητών στους πραγματικούς αριθμούς. Τότε η μέση τιμή $Ef(X, Y)$ δίδεται από την σχέση

$$E[f(X, Y)] = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} f(i, j)p_{X,Y}(i, j). \tag{2.1}$$

Το επόμενο θεώρημα εκφράζει την θεμελιώδη ιδιότητα της γραμμικότητας της μέσης τιμής.

Θεώρημα 2.1. *Αν* X, Y *δύο οποιεσδήποτε τυχαίες μεταβλητές και* a, b, *πραγματικοί αριθμοί, τότε* $E[aX + bY] = aEX + bEY$.

Απόδειξη. Χρησιμοποιώντας τον τύπο (2.1), όπου $f(i, j) = a \cdot i + b \cdot j$,

$$\begin{aligned}
E[aX + bY] &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}(ai + bj)p_{X,Y}(i, j) = a\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} i\, p_{X,Y}(i, j) + b\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} j\, p_{X,Y}(i, j) \\
&= a\sum_{i=0}^{\infty} i\sum_{j=0}^{\infty} p_{X,Y}(i, j) + b\sum_{j=0}^{\infty} j\sum_{i=0}^{\infty} p_{X,Y}(i, j) \\
&= a\sum_{i=0}^{\infty} iP(X = i) + b\sum_{j=0}^{\infty} jP(Y = j) \\
&= aEX + bEY.
\end{aligned}$$

□

Θεώρημα 2.2. *Έστω* f_i, $i = 1, 2$, *δύο συναρτήσεις από τους φυσικούς στους πραγματικούς και* X_1, X_2, *δύο ανεξάρτητες τ.μ. Τότε*

$$E[f_1(X_1)f_2(X_2)] = E[f_1(X_1)]\; E[f_2(X_2)]. \tag{2.2}$$

Το αποτέλεσμα αυτό επεκτείνεται επαγωγικά για οποιοδήποτε αριθμό τ.μ.

Απόδειξη. Χρησιμοποιώντας τον τύπο (2.1) έχουμε

$$\begin{aligned} E[f_1(X_1)f_2(X_2)] &= \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} f_1(n_1)f_2(n_2)P(X_1 = n_1)P(X_2 = n_2) \\ &= \left(\sum_{n_1=0}^{\infty} f_1(n_1)P(X_1 = n_1)\right)\left(\sum_{n_2=0}^{\infty} f_2(n_2)P(X_2 = n_2)\right) \\ &= E[f_1(X_1)]\; E[f_2(X_2)]. \end{aligned}$$

□

Η *συνδιακύμανση* δύο τυχαίων μεταβλητών, X_1, X_2, ορίζεται ως

$$\mathrm{Cov}(X_1, X_2) = E\left[(X_1 - EX_1)(X_2 - EX_2)\right] = E[X_1X_2] - (EX_1)(EX_2).$$

Δυο τ.μ. των οποίων η συνδιακύμανση είναι 0 ονομάζονται *ασυσχέτιστες*. Από την (2.2) προκύπτει ότι όταν δύο τ.μ. είναι ανεξάρτητες τότε είναι και ασυσχέτιστες αφού σ' αυτή την περίπτωση $E[X_1X_2] = (EX_1)(EX_2)$. *Το αντίθετο όμως δεν ισχύει, δηλαδή δύο ασυσχέτιστες τ.μ. δεν είναι υποχρεωτικά ανεξάρτητητες.*

Ας δούμε το ακόλουθο χαρακτηριστικό αντιπαράδειγμα. Έστω X, Y τυχαίες μεταβλητές, ορισμένες στον ίδιο χώρο με τιμές στο σύνολο $\{0, 1, 2\}$ και από κοινού κατανομή που δίνεται από τον ακόλουθο πίνακα.

Y/X	0	1	2
0	1/8	0	1/8
1	0	1/2	0
2	1/8	0	1/8

Η περιθώρια κατανομή της X είναι $p_X(0) = 1/4$, $p_X(1) = 1/2$, $p_X(2) = 1/4$. Η περιθώρια κατανομή της Y είναι ίδια με αυτή της X. Έχουμε συνεπώς ότι $EX = EY = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} = 1$. Η μέση τιμή του γινομένου των X και Y, γράφοντας μόνο τους μη μηδενικούς όρους, είναι $EXY = 1 \cdot 1 \cdot \frac{1}{2} + 2 \cdot 2 \cdot \frac{1}{8} = 1$ και επομένως η συνδιακύμανση είναι $\mathrm{Cov}(X, Y) = E[XY] - EX \cdot EY = 1 - 1 = 0$. Άρα οι X και Y είναι ασυσχέτιστες. Δεν είναι όμως ανεξάρτητες. Για παράδειγμα $P(X = 1, Y = 2) = 0 \neq \frac{1}{2} \cdot \frac{1}{4} = P(X = 1) \cdot P(Y = 2)$, και επομένως η από κοινού κατανομή δεν είναι το γινόμενο των περιθωρίων.

Θεώρημα 2.3. *Για δύο οποιεσδήποτε τ.μ.* X_1, X_2,

$$\mathrm{Var}(X_1 + X_2) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + 2\mathrm{Cov}(X_1, X_2). \tag{2.3}$$

Όταν οι X_1, X_2 *είναι ασυσχέτιστες*, $\mathrm{Var}(X_1 + X_2) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2)$.

Απόδειξη. Χρησιμοποιούμε την ακόλουθη έκφραση για την διασπορά

$$\begin{aligned}
\mathrm{Var}(X_1 + X_2) &= E[(X_1 + X_2)^2] - (E[X_1 + X_2])^2 \\
&= E[X_1^2 + 2X_1X_2 + X_2^2] - \Big((EX_1)^2 + 2EX_1EX_2 + (EX_2)^2\Big) \\
&= E[X_1^2] + 2E[X_1X_2] + E[X_2^2] - (EX_1)^2 - (EX_2)^2 - 2EX_1EX_2
\end{aligned}$$

όπου χρησιμοποιήσαμε κατ΄ επανάληψη τη γραμμικότητα της μέσης τιμής. Συλλέγοντας όρους στο τελευταίο μέλος των παραπάνω εξισώσεων και λαμβάνοντας υπ΄ όψιν ότι $\mathrm{Var}(X_i) = EX_i^2 - (EX_i)^2$, $i = 1, 2$, $\mathrm{Cov}(X_1, X_2) = E[X_1X_2] - EX_1EX_2$, προκύπτει η (2.3). □

2.2 Η ανισότητα του Chebyshev

Θεώρημα 2.4. *Αν* X *είναι τυχαία μεταβλητή με μέσο* μ *και διασπορά* σ^2 *τότε, για κάθε* $\epsilon > 0$,

$$P(|X - \mu| > \epsilon) \le \frac{\sigma^2}{\epsilon^2}. \tag{2.4}$$

Απόδειξη.

$$\begin{aligned}
\sigma^2 &= \sum_k (k-\mu)^2 p_k = \sum_{\{k:|k-\mu|>\epsilon\}} (k-\mu)^2 p_k + \sum_{\{k:|k-\mu|\le\epsilon\}} (k-\mu)^2 p_k \\
&\ge \sum_{\{k:|k-\mu|>\epsilon\}} (k-\mu)^2 p_k \\
&\ge \epsilon^2 \sum_{\{k:|k-\mu|>\epsilon\}} p_k \\
&= \epsilon^2 P(|X-\mu| > \epsilon).
\end{aligned}$$

□

Πόρισμα 2.1. *Αν για μια τυχαία μεταβλητή με μέσο* μ *και διασπορά* σ^2 *ισχύει ότι* $\sigma = 0$ *τότε* $P(X = \mu) = 1$.

Απόδειξη. Από το δεξί μέλος της ανισότητας του Chebyshev ισχύει ότι $P(|X-\mu| > \epsilon) = 0$ για κάθε $\epsilon > 0$. □

Η παραπάνω ανισότητα μας δίνει ένα άνω όριο για την πιθανότητα μια τυχαία μεταβλητή να απέχει περισσότερο από ϵ από τη μέση της τιμή. Το άνω όριο δεν είναι εν γένει ικανοποιητικό για τις περισσότερες πρακτικές εφαρμογές, έχει όμως μεγάλη θεωρητική σημασία. Εδώ θα το χρησιμοποιήσουμε στην απόδειξη του Νόμου των Μεγάλων Αριθμών. Ώς ένα παράδειγμα εφαρμογής της ανισότητας του Chebyshev, ας πάρουμε την X ομοιόμορφη στο σύνολο $\{1, 2, \ldots, 100\}$. Έχουμε $\mathsf{E}X = 50,5$ και $\mathsf{Var}(X) = 833,25$. Με $\epsilon = 40$ η ανισότητα μας λέει ότι

$$\mathsf{P}(|X - 50,5| > 40) \leq \frac{833,5}{1600} = 0,52.$$

Είναι όμως $\mathsf{P}(|X - 50,5| > 40) = \mathsf{P}(X \geq 91) + \mathsf{P}(X \leq 10) = 0,2$. Για μικρές τιμές του ϵ το δεξί μέλος της (2.4) γίνεται μεγαλύτερο από την μονάδα κι΄ έτσι η ανισότητα δεν μας δίνει καμία χρήσιμη πληροφορία. Όπως θα δούμε όμως, σαν θεωρητικό εργαλείο μπορεί να είναι εξαιρετικά χρήσιμη.

2.3 Η ανισότητα Cauchy–Schwarz

Θεώρημα 2.5. *Για δύο οποιεσδήποτε τυχαίες μεταβλητές* X, Y, *με πεπερασμένες διασπορές ισχύει ότι*

$$|\mathsf{Cov}(X, Y)| \leq \sqrt{\mathsf{Var}(X)\mathsf{Var}(Y)}. \tag{2.5}$$

Αν η (2.5) ισχύει ως ισότητα, τότε υπάρχουν $a, b \in \mathbb{R}$ *τέτοια ώστε* $Y = aX + b$.

Απόδειξη. Θεωρούμε την συνάρτηση

$$f(t) := \mathsf{Var}(Y - tX) = \mathsf{Var}(Y) + t^2\mathsf{Var}(X) - 2t\mathsf{Cov}(X, Y) \quad \text{με } t \in \mathbb{R}.$$

Εφ΄ όσον η $f(t)$ είναι διασπορά θα είναι υποχρεωτικά $f(t) \geq 0$ *για κάθε* $t \in \mathbb{R}$. Συνεπώς η διακρίνουσα του τριωνύμου θα είναι αρνητική ή μηδέν, πράγμα που σημαίνει ότι

$$(\mathsf{Cov}(X, Y))^2 \leq \mathsf{Var}(X)\mathsf{Var}(Y).$$

Από την σχέση αυτή προκύπτει άμεσα η (2.5).

Αν η (2.5) ισχύει ως ισότητα τότε η διακρίνουσα του τριωνύμου είναι 0 και υπάρχει ακριβώς μια τιμή του t, έστω $t = a$ τέτοια ώστε $f(a) = 0$. Αυτό όμως σημαίνει ότι $\mathsf{Var}(Y - aX) = 0$ και συνεπώς ότι $\mathsf{P}(Y - aX = b) = 1$ με $b = \mathsf{E}[Y - aX]$ από το πόρισμα της ανισότητας του Chebyshev. Η τιμή του a είναι $\frac{\mathsf{Cov}(X,Y)}{\mathsf{Var}(X)}$. □

Ορισμός 2.3. Η ποσότητα

$$\rho := \frac{\mathsf{Cov}(X, Y)}{\sqrt{\mathsf{Var}(X)\mathsf{Var}(Y)}} \tag{2.6}$$

ονομάζεται συντελεστής συσχέτισης και παίρνει τιμές στο διάστημα $[-1, 1]$.

Το γεγονός ότι ο συντελεστής συσχέτισης ανήκει πάντα στο διάστημα $[-1, 1]$ είναι απόρροια της ανισότητας Cauchy–Schwarz.

2.4 Ο Νόμος των Μεγάλων Αριθμών

Θεώρημα 2.6. *Αν* $X_1, X_2, X_3, \ldots$ *είναι μια ακολουθία από ανεξάρτητες, ισόνομες τυχαίες μεταβλητές με μέσο* μ *και διασπορά* σ^2 *τότε ο αριθμητικός μέσος* $\frac{X_1+X_2+\cdots+X_n}{n}$ *τείνει* κατά πιθανότητα *στην μέση τιμή όταν το* $n \to \infty$ *δηλαδή*

$$\mathsf{P}\left(\left|\frac{X_1+X_2+\cdots+X_n}{n} - \mu\right| > \epsilon\right) \to 0 \quad \forall\, \epsilon > 0 \text{ όταν } n \to \infty. \tag{2.7}$$

Απόδειξη. Ας θέσουμε $Y_n := \frac{1}{n}(X_1 + \cdots + X_n)$. Θα έχουμε $\mathsf{E}Y_n = \frac{1}{n}(\mu + \cdots + \mu) = \mu$ από την γραμμικότητα της μέσης τιμής και αφού οι X_i, $i = 1, 2, \ldots, n$, είναι ανεξάρτητες θα έχουμε επίσης ότι

$$\mathsf{Var}(Y_n) = \frac{1}{n^2}(\sigma^2 + \cdots + \sigma^2) = \frac{\sigma^2}{n}.$$

Το αριστερό μέλος της (2.7) γράφεται ως $\mathsf{P}(|Y_n - \mu| \geq \epsilon)$ και από την ανισότητα του Chebyshev) έχουμε ότι

$$\mathsf{P}(|Y_n - \mu| > \epsilon) \leq \frac{\mathsf{Var}(Y_n)}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2}.$$

Για $\epsilon > 0$ οσοδήποτε μικρό αλλά σταθερό, καθώς το n τείνει στο άπειρο το δεξί μέλος της παραπάνω σχέσης τείνει στο μηδέν! Αυτό αποδεικνύει την (2.7). □

(Ισόνομες τυχαίες μεταβλητές είναι τυχαίες μεταβλητές που έχουν τον ίδιο «νόμο πιθανοτήτων» δηλαδή την ίδια κατανομή. Η υπόθεση αυτή δεν ήταν πραγματικά απαραίτητη εδώ. Αρκεί οι τυχαίες μεταβλητές να είναι ανεξάρτητες (ή ακόμα και απλά ασυσχέτιστες) και να έχουν την ίδια μέση τιμή και διασπορά.)

2.5 Κατανομή του αθροίσματος ανεξαρτήτων τ.μ.

Έστω X, Y, ανεξάρτητες τ.μ. με δεδομένες κατανομές $\mathsf{P}(X = n)$, $\mathsf{P}(Y = m)$, $n, m = 0, 1, 2, \ldots$. Έστω $Z = X + Y$, το άθροισμα των δύο τ.μ. Η κατανομή της Z υπολογίζεται ως εξής

$$\begin{aligned}\mathsf{P}(Z = k) &= \mathsf{P}(X + Y = k) = \sum_{n=0}^{k} \mathsf{P}(X = n, Y = k - n) \\ &= \sum_{n=0}^{k} \mathsf{P}(X = n)\mathsf{P}(Y = k - n).\end{aligned} \tag{2.8}$$

Η τελευταία σχέση οφείλεται στην ανεξαρτησία των X και Y το δε άθροισμα ονομάζεται *συνέλιξη* των δύο κατανομών.

Παράδειγμα 2.2. Έστω X διωνυμική τ.μ. με παραμέτρους n και p δηλαδή $P(X = i) = \binom{n}{i} p^i q^{n-i}$, για $i = 0, 1, 2, \ldots, n$, όπου $q = 1 - p$. Η τ.μ. Y είναι επίσης διωνυμική με παραμέτρους m και p και ανεξάρτητη από την X. Η κατανομή της $Z := X + Y$ δίνεται τότε από την συνέλιξη των δύο διωνυμικών κατανομών δηλαδή

$$\begin{aligned} P(Z = k) &= \sum_{i=0}^{k} P(X = i)P(Y = k - i) = \sum_{i=0}^{k} \binom{n}{i} p^i q^{n-i} \binom{m}{k-i} p^{k-i} q^{m-k+i} \\ &= p^k q^{m+n-k} \sum_{i=0}^{k} \binom{n}{i} \binom{m}{k-i} \\ &= \binom{m+n}{k} p^k q^{m+n-k}. \end{aligned}$$

Στην τελευταία εξίσωση της παραπάνω σχέσης χρησιμοποιήσαμε την συνδυαστική ταυτότητα

$$\sum_{i=0}^{k} \binom{n}{i} \binom{m}{k-i} = \binom{m+n}{k}. \tag{2.9}$$

Παράδειγμα 2.3. Έστω X, Y ανεξάρτητες γεωμετρικές τ.μ. με παράμετρο p. Η κατανομή της $Z := X + Y$ δίνεται από την συνέλιξη

$$\begin{aligned} P(Z = k) &= \sum_{i=0}^{k} P(X = i)P(Y = k - i) = \sum_{i=1}^{k-1} p q^i p q^{k-i} \\ &= (k - 1) q^k p^2. \end{aligned}$$

(Η κατανομή αυτή όπως θα δούμε αργότερα είναι μια ειδική περίπτωση της αρνητικής διωνυμικής κατανομής ή κατανομής Pascal.)

Παράδειγμα 2.4. Έστω X τ.μ. Poisson με παράμετρο α_1, δηλαδή $P(X = k) = \frac{\alpha_1^k}{k!} e^{-\alpha_1}$, $k = 0, 1, 2, \ldots$. Η τ.μ. Y είναι επίσης Poisson με παράμετρο α_2 και ανεξάρτητη της X. Η κατανομή της $Z := X + Y$ δίνεται τότε από την συνέλιξη των δύο κατανομών Poisson δηλαδή

$$\begin{aligned} P(Z = k) &= \sum_{n=0}^{k} P(X = n)P(Y = k - n) = \sum_{n=0}^{k} \frac{\alpha_1^n}{n!} e^{-\alpha_1} \frac{\alpha_2^{k-n}}{(k-n)!} e^{-\alpha_2} \\ &= e^{-(\alpha_1+\alpha_2)} \sum_{n=0}^{k} \frac{\alpha_1^n}{n!} \frac{\alpha_2^{k-n}}{(k-n)!} = \frac{1}{k!} e^{-(\alpha_1+\alpha_2)} \sum_{n=0}^{k} \binom{k}{n} \alpha_1^n \alpha_2^{k-n} \\ &= \frac{(\alpha_1 + \alpha_2)^k}{k!} e^{-(\alpha_1+\alpha_2)}. \end{aligned}$$

Επομένως το άθροισμα δύο ανεξαρτήτων τ.μ. Poisson είναι πάλι Poisson με παράμετρο το άθροισμα των παραμέτρων.

2.6 Παραδείγματα

1. Ας υποθέσουμε ότι ρίχνουμε δύο ζάρια διακριτά μεταξύ τους, π.χ. ένα κόκκινο και ένα πράσινο. Έστω X το αποτέλεσμα της ρίψης του κόκκινου ζαριού και Y το αποτέλεσμα της ρίψης του πράσινου. Η από κοινού κατανομή των X, Y, δίνεται από τον πίνακα

Y/X	1	2	3	4	5	6
1	1/36	1/36	1/36	1/36	1/36	1/36
2	1/36	1/36	1/36	1/36	1/36	1/36
3	1/36	1/36	1/36	1/36	1/36	1/36
4	1/36	1/36	1/36	1/36	1/36	1/36
5	1/36	1/36	1/36	1/36	1/36	1/36
6	1/36	1/36	1/36	1/36	1/36	1/36

Έστω $Z := \min(X, Y)$ και $W := \max(X, Y)$ το μικρότερο και το μεγαλύτερο από τα δύο αποτελέσματα των δύο ζαριών. Οι Z και W είναι βεβαίως τυχαίες μεταβλητές. Η από κοινού κατανομή της X και της Z είναι

Z/X	1	2	3	4	5	6
1	6/36	1/36	1/36	1/36	1/36	1/36
2		5/36	1/36	1/36	1/36	1/36
3			4/36	1/36	1/36	1/36
4				3/36	1/36	1/36
5					2/36	1/36
6						1/36

(Κενές τιμές στον πίνακα είναι μηδέν.) Δεν είναι δύσκολο να καταλάβουμε πώς συμπληρώνεται αυτός ο πίνακας. Για παράδειγμα $P(X = 2, Z = 3) = 0$ αφού είναι αδύνατον το μικρότερο από τα αποτελέσματα να είναι μεγαλύτερο από τό ελάχιστο των δύο. $P(X = 3, Z = 2) = P(X = 3, Y = 2) = 1/36$ αφού σ' αυτή την περίπτωση η τιμή του Z αποκαλύπτει και την τιμή του Y. Τέλος $P(X = 3, Z = 3) = P(X = 3, Y = 1) + P(X = 3, Y = 2) + P(X = 3, Y = 3) = 3/36$. Οι υπόλοιπες τιμές συμπληρώνονται με παρόμοιο τρόπο.

Εντελώς ανάλογα μπορούμε να δούμε ότι

W/X	1	2	3	4	5	6
1	1/36					
2	1/36	2/36				
3	1/36	1/36	3/36			
4	1/36	1/36	1/36	4/36		
5	1/36	1/36	1/36	1/36	5/36	
6	1/36	1/36	1/36	1/36	1/36	6/36

Τέλος η από κοινού κατανομή του maximum και του minimum είναι

W/Z	1	2	3	4	5	6
1	1/36					
2	2/36	1/36				
3	2/36	2/36	1/36			
4	2/36	2/36	2/36	1/36		
5	2/36	2/36	2/36	2/36	1/36	
6	2/36	2/36	2/36	2/36	2/36	1/36

Για να καταλάβουμε πώς συμπληρώνεται ο τελευταίος πίνακας παρατηρούμε ότι $P(W = 2, Z = 2) = P(X = 2, Y = 2) = 1/36$. Τό ίδιο ισχύει βέβαια και για τα άλλα διαγώνια στοιχεία. Παρόμοια, $P(W = 2, Z = 4) = 0$ αφού το maximum δεν μπορεί να είναι μικρότερο από το minimum. Τέλος $P(W = 4, Z = 2) = P(X = 2, Y = 4) + P(X = 4, Y = 2) = 2/36$.

Ας υπολογίσουμε τώρα τις περιθώριες κατανομές των Z και W. Θα χρησιμοποιήσουμε τον τρίτο πίνακα αλλά θα μπορούσαμε εναλλακτικά να χρησιμοποιήσουμε και τους δύο πρώτους. Έχουμε

W	1	2	3	4	5	6
	1/36	3/36	5/36	7/36	9/36	11/36

Z	1	2	3	4	5	6
	11/36	9/36	7/36	5/36	3/36	1/36

Η μέση τιμή της W είναι $EW = \frac{1}{36}(1 \cdot 1 + 2 \cdot 3 + 3 \cdot 5 + 4 \cdot 7 + 5 \cdot 9 + 6 \cdot 11) = 4,4722$. Δεν πρέπει να μας εκπλήσσει το γεγονός ότι η μέση τιμή της W είναι μεγαλύτερη από την $EX = 3,5$. Την μέση τιμή της Z θα μπορούσαμε να την υπολογίσουμε από τον ορισμό. Όμως είναι απλούστερο (και πιο διδακτικό) να παρατηρήσουμε ότι $X + Y = \min(X, Y) + \max(X, Y) = Z + W$. Από την γραμμικότητα της μέσης τιμής θα έχουμε $EX + EY = EZ + EW$ και συνεπώς $EZ = 7 - 4,4722 = 2,5278$.

Οι διασπορές υπολογίζονται επίσης από τον ορισμό: $\mathsf{Var}(X) = 2,9167$, $\mathsf{Var}(W) = \mathsf{Var}(Z) = 1,9715$. Παρατηρείστε ότι λόγω συμμετρίας οι διασπορές των Z και W είναι ίσες.

Οι συνδιακυμάνσεις που υπολογίζονται από τον ορισμό είναι $\mathsf{Cov}(X, Z) = 1,4583$, $\mathsf{Cov}(X, W) = 1,4583$, $\mathsf{Cov}(Z, W) = 0,9452$.

Τέλος οι συντελεστές συσχέτισης είναι $\rho_{XZ} = 0,6082$, $\rho_{XW} = 0,6082$, $\rho_{ZW} = 0,4795$. Παρατηρείστε ότι όλοι οι συντελεστές συσχέτισης είναι θετικοί. Αυτό είναι αναμενόμενο. Αν ξέρουμε ότι το X είναι μεγάλο τότε το minimum των X και Y επίσης θα τείνει να είναι μεγαλύτερο από τη μέση του τιμή, όπως και το maximum. Αν ξέρουμε ότι το minimum είναι μεγαλύτερο από την μέση του τιμή τότε και το maximum θα τείνει να είναι μεγαλύτερο από τη δική του μέση τιμή. Όμως η γνώση του X μας δίνει μεγαλύτερη πληροφορία για το maximum απ΄ ότι η γνώση του minimum και γι΄ αυτό $\rho_{XZ} > \rho_{WZ}$.

2.7 Ασκήσεις

Πρόβλημα 2.1. Έστω X_1, X_2 ανεξάρτητες τυχαίες μεταβλητές Poisson με παραμέτρους α_1, α_2 αντίστοιχα. Αν $S = X_1 + X_2$ είναι το άθροισμά τους να ευρεθεί η πιθανότητα $P(X_1 = k|S = n)$, $k = 0, 1, 2, \ldots, n$.

Πρόβλημα 2.2. X_i, $i = 1, 2, 3$, είναι ανεξάρτητες τυχαίες μεταβλητές με κατανομή Poisson με παραμέτρους α_i αντίστοιχα. Αν $W = X_1 + X_3$ και $V = X_2 + X_3$ να υπολογίσετε τους μέσους των W, V, τις διακυμάνσεις, την συνδιακύμανση και τον συντελεστή συσχέτισης. Να υπολογίσετε επίσης και την από κοινού κατανομή των W και V.

Πρόβλημα 2.3. Η από κοινού κατανομή των τυχαίων μεταβλητών X, Y, που παίρνουν τιμές στο σύνολο $\{0, 1, 2\}$ δίδεται από τον πίνακα

2	1/6	1/6	0
1	1/6	0	1/6
0	0	1/6	1/6
Y/X	0	1	2

Να ευρεθούν οι μέσοι, οι διακυμάνσεις και η συνδιακύμανση των X, Y. Είναι οι X, Y ανεξάρτητες;

2.8 Πιθανογεννήτριες

Έστω X μια διακριτή τυχαία μεταβλητή με κατανομή $\mathbb{P}(X = k) = p_k$, $k = 0, 1, 2, \ldots$. Η πιθανογεννήτρια της X είναι η συνάρτηση

$$G(z) := \sum_{k=0}^{\infty} p_k z^k \tag{2.10}$$

Η μεταβλητή z παίρνει πραγματικές τιμές (στις εφαρμογές που θα συζητήσουμε) και κυμαίνεται στο διάστημα στο οποίο η σειρά (2.10) συγκλίνει. Παρατηρείστε ότι η (2.10) σημαίνει ότι $G(z) := \mathbb{E}[z^X]$.

Η σειρά (2.10) που ορίζει την πιθανογεννήτρια συγκλίνει τουλάχιστον όταν $z \in [-1,1]$. Συμβολίζουμε με $G^{(k)}(z)$ την παράγωγο τάξης k υπολογισμένη στην τιμή z. Τότε

$$p_k = \frac{1}{k!} G^{(k)}(0), \qquad k = 0,1,2,\ldots, \tag{2.11}$$

Παρατηρούμε ότι η κατανομή $\{p_k\}$ προφανώς προσδιορίζει με μοναδικό τρόπο την πιθανογεννήτρια $G(z)$ αλλά και η πιθανογεννήτρια $G(z)$ προσδιορίζει με μοναδικά την κατανομή $\{p_k\}$ μέσω της (2.11).

Οι πιθανογεννήτριες είναι ιδιαίτερα χρήσιμες για τον προσδιορισμό της κατανομής αθροισμάτων ανεξάρτητων διακριτών τυχαίων μεταβλητών. Αν οι X, Y, είναι *ανεξάρτητες* τυχαίες μεταβλητές με πιθανογεννήτριες $G_X(z)$, $G_Y(z)$ αντίστοιχα, η πιθανογεννήτρια του αθροίσματος $Z = X + Y$ είναι $G_Z(z) = G_X(z)G_Y(z)$. Πράγματι, $G_Z(z) = \mathbb{E}[z^{X+Y}] = \mathbb{E}[z^X z^Y] = \mathbb{E}z^X \mathbb{E}z^Y$, όπου, η τελευταία ισότητα ισχύει λόγω της ανεξαρτησίας των X, Y. Η σχέση αυτή επεκτείνεται επαγωγικά για οποιοδήποτε πεπερασμένο πλήθος ανεξάρτητων τ.μ.. Αν X_i, $i = 1,2,\ldots,n$ είναι ανεξάρτητες, ισόνομες τ.μ. με (κοινή) πιθανογεννήτρια $G_X(z)$, τότε το άθροισμά τους $S_n := X_1 + \cdots + X_n$ έχει πιθανογεννήτρια $G_{S_n}(z) = (G_X(z))^n$.

Παρ΄ ότι η πιθανογεννήτρια του αθροίσματος S_n προκύπτει εύκολα ως το γινόμενο των πιθανογεννητριών των X_i, η κατανομή της S_n είναι δυσκολότερο να υπολογιστεί. Βάσει των ανωτέρω,

$$\mathbb{P}(S_n = k) = \frac{1}{k!} \left. \frac{d^k}{dz^k} (G_X(z))^n \right|_{z=0},$$

μια ποσότητα η οποία, εν γένει, δεν είναι εύκολο να υπολογιστεί.

2.8.1 Παραδείγματα πιθανογεννητριών

2.8.2 Η κατανομή Bernoulli

Η τυχαία μεταβλητή

$$X = \begin{cases} 0 & \text{με πιθανότητα } q \\ 1 & \text{με πιθανότητα } p \end{cases}$$

όπου $p \in [0,1]$ και $q = 1 - p$ ονομάζεται Bernoulli. Η κατανομή Bernouli ειναι η πλέον στοιχειώδης και αποτελεί δομικό λίθο για πιο περίπλοκες κατανομές. Η πιθανογεννήτρια της κατανομής αυτής είναι

$$G(z) = z^0 \cdot \mathbb{P}(X=0) + z^1 \cdot \mathbb{P}(X=1) + z^2 \cdot \mathbb{P}(X=2) + \cdots \; = \; q + pz.$$

Η μέση τιμή της X είναι $\mathbb{E}[X] = G'(1) = p$. Η δεύτερη παραγοντική ροπή είναι $\mathbb{E}[X(X-1)] = G''(1) = 0$. Συνεπώς $\mathbb{E}[X^2 - X] = 0$ και $\mathbb{E}[X^2] = \mathbb{E}[X]$, δηλαδή $\mathbb{E}[X^2] = p$. Συνεπώς $\mathsf{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X]) = p - p^2 = p(1-p) = pq$.

2.8.3 Η Διωνυμική κατανομή

Μια τυχαία μεταβλητή X με κατανομή

$$\mathbb{P}(X = k) = \binom{n}{k} p^k q^{n-k}, \qquad \text{αν } k = 0, 1, 2, \ldots, n \text{ και } 0 \text{ διαφορετικά}$$

ονομάζεται διωνυμική. Περιγράφει την πιθανότητα να έχουμε k επιτυχίες σε n ανεξάρτητες δοκιμές, κάθε μια από τις οποίες έχει πιθανότητα επιτυχίας $p \in [0, 1]$ (και πιθανότητα αποτυχίας $q = 1 - p$.) Η πιθανογεννήτρια της διωνυμικής κατανομής είναι

$$G(z) = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} z^k = (q + pz)^n$$

όπου στην τελευταία σχέση χρησιμοποιήσαμε το διωνυμικό θεώρημα. Ο μέσος και η διασπορά μπορούν να υπολογισθούν από την πιθανογεννήτρια ως εξής: $G'(z) = np(q + pz)^{n-1}$ και $G''(z) = n(n-1)p^2(q + pz)^{n-2}$ (υποθέτουμε ότι $n \geq 2$). $\mathbb{E}[X] = G'(1) = np$ (χρησιμοποιήσαμε το γεγονός ότι $p + q = 1$). $\mathbb{E}[X(X-1)] = n(n-1)p^2$. Συνεπώς $\mathbb{E}[X^2] = n(n-1)p^2 + \mathbb{E}[X] = n(n-1)p^2 + np$ και κατά συνέπεια

$$\mathsf{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = n(n-1)p^2 + np - n^2p^2 = np - np^2 = np(1-p) = npq.$$

2.8.4 Η κατανομή Poisson

Η τυχαία μεταβλητή X έχει κατανομή Poisson με παράμετρο $\alpha > 0$ αν

$$\mathbb{P}(X = k) = \frac{1}{k!} \alpha^k e^{-\alpha}, \qquad k = 0, 1, 2, \ldots.$$

Η πιθανογεννήτριά της δίδεται από την

$$G(z) = \sum_{k=0}^{\infty} z^k \frac{1}{k!} \alpha^k e^{-\alpha} = e^{-\alpha} \sum_{k=0}^{\infty} \frac{1}{k!} (\alpha z)^k = e^{-\alpha} e^{z\alpha} = e^{-\alpha(1-z)}.$$

Η μέση τιμή και η διασπορά της κατανομής Poisson υπολογίζεται εύκολα ως $\mathsf{E}X = \mathsf{Var}(X) = \alpha$.

Μια από τις σημαντικότερες ιδιότητες της κατανομής Poisson είναι ότι προκύπτει ως το όριο της διωνυμικής κατανομής $\mathsf{Binom}(n, \alpha/n)$ όταν $n \to \infty$ (δηλαδή στην περίπτωση που έχουμε ένα μεγάλο αριθμό ανεξάρτητων δοκιμών, n, η κάθε μια από τις οποίες έχει μικρή πιθανότητα επιτυχίας, α/n). Αυτό είναι εύκολο να διαπιστωθεί εξετάζοντας την πιθανογεννήτρια της διωνυμικής κατανομής $(n, \alpha/n)$ και παίρνοντας το όριο $n \to \infty$. Πράγματι,

$$\lim_{n\to\infty} \left(1 - \frac{\alpha}{n} + z\frac{\alpha}{n}\right)^n = \lim_{n\to\infty} \left(1 - \frac{\alpha(1-z)}{n}\right)^n = e^{-\alpha(1-z)}$$

το οποίο δείχνει ότι $\mathsf{Binom}(\alpha/n, n) \to \mathsf{Poi}(\alpha)$ όταν $n \to \infty$.

Ισχύει επίσης ότι, αν X_1, X_2 είναι ανεξέαρτητες τυχαίες μεταβλητές Poisson με παραμέτρους α_1, α_2 αντίστοιχα, τότε $X_1 + X_2 \sim \mathsf{Poi}(\alpha_1 + \alpha_2)$. Ο απλούστερος τρόπος να το διαπιστώσουμε είναι να εξετάσουμε την πιθανογεννήτρια $\mathsf{E}z^{X_1+X_2} = \mathsf{E}z^{X_1}\mathsf{E}z^{X_2} = e^{-\alpha_1(1-z)}e^{-\alpha_2(1-z)} = e^{-(\alpha_1+\alpha_2)(1-z)}$.

2.8.5 Η Γεωμετρική Κατανόμή

Η τυχαία μεταβλητή X είναι γεωμετρική με παράμετρο p όταν η κατανομή της δίδεται από την

$$\mathbb{P}(X = k) = q^{k-1}p, \qquad k = 1, 2, 3, \dots, \tag{2.12}$$

όπου $p \in (0, 1)$ και $q = 1 - p$. Η πιθανογεννήτρια συνάρτηση της κατανομής αυτής δίδεται από την σχέση

$$\mathsf{G}(z) = \sum_{k=1}^{\infty} q^{k-1}pz^k = \frac{(1-q)z}{1-qz}. \tag{2.13}$$

Η παράμετρος p ονομάζεται συνήθως 'πιθανότητα επιτυχίας' και η τιμή της X είναι ο αριθμός των ανεξάρτητων δοκιμών που απαιτούνται μέχρι την πρώτη επιτυχία αν κάθε δοκιμή έχει πιθανότητα επιτυχίας p. Εναλλακτικά, μπορούμε να εξετάσουμε τον αριθμό των *αποτυχιών*, Y, μέχρι την πρώτη επιτυχία. Στην περίπτωση αυτή $Y = X - 1$ και

$$\mathbb{P}(Y = k) = q^k p, \qquad k = 0, 1, 2, \dots, \tag{2.14}$$

με πιθανογεννήτρια

$$\mathbb{E}z^Y = \frac{1-q}{1-qz}. \tag{2.15}$$

Μπορούμε να δούμε εύκολα ότι $\mathbb{E}Y = q/p$ και $\mathsf{Var}(Y) = q/p^2$. Επίσης, $\mathbb{E}X = 1 + \mathbb{E}Y = 1/p$ και $\mathsf{Var}(X) = \mathsf{Var}(Y) = q/p^2$.

2.8.6 Η Αρνητική Διωνυμική Κατανομή (Κατανομή Pascal)

Ξεκινάμε με τον ορισμό του διωνυμικού συντελεστή στην περίπτωση που $a \in \mathbb{R}$ και $n \in \mathbb{N}$ ως

$$\binom{a}{n} := \frac{a(a-1)\dots(a-n+1)}{n!}.$$

Αν ο a είναι φυσικός αριθμός τότε $\binom{a}{n} = 0$ για κάθε $n > a$. Αν ο a είναι αρνητικός ακέραιος ή μη ακέραιος πραγματικός αριθμός, τότε $\binom{a}{n} \neq 0$ για κάθε $n \in \mathbb{N}$. Σύμφωνα με το διωνυμικό θεώρημα, για κάθε $|x| < 1$ και $\alpha \in \mathbb{R}$:

$$(1 + x)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k. \tag{2.16}$$

(Αν ο α είναι θετικός ακέραιος τότε $\binom{\alpha}{k} = 0$ για κάθε $k = \alpha+1, \alpha+2, \cdots$ και συνεπώς η άπειρη σειρά (2.16) γίνεται ένα άθροισμα με πεπερασμένο πλήθος όρων: $(1+x)^\alpha = \sum_{k=0}^{\alpha} \binom{\alpha}{k} x^k$.)

Ο διωνυμικός συντελεστής $\binom{-\alpha}{n}$ γράφεται ως

$$\begin{aligned}\binom{-\alpha}{n} &= \frac{(-\alpha)(-\alpha-1)\cdots(-\alpha-n+2)(-\alpha-n+1)}{n!} \\ &= (-1)^n \frac{(\alpha+n-1)(\alpha+n-2)\cdots(\alpha+1)\alpha}{n!} = (-1)^n \binom{\alpha+n-1}{n}.\end{aligned}$$

Συνεπώς, για κάθε $|x| < 1$ ισχύει η ταυτότητα

$$(1-x)^{-\alpha} = \sum_{k=0}^{\infty} \binom{-\alpha}{k} (-x)^k = \sum_{k=0}^{\infty} \binom{\alpha+k-1}{k} x^k. \tag{2.17}$$

Αν $p \in (0,1)$ και $q = 1-p$ τότε *η αρνητική διωνυμική κατανομή* με παραμέτρους p και $\alpha > 0$ ορίζεται ως

$$\mathbb{P}(X = k) = \binom{\alpha+k-1}{k} p^\alpha q^k, \qquad k = 0, 1, 2, \ldots. \tag{2.18}$$

Η παραπάνω έκφραση ορίζει πράγματι μια κατανομή πιθανότητας στους μη αρνητικούς ακεραίους δεδομένου ότι $\binom{\alpha+k-1}{k} > 0$ όταν $\alpha > 0$ για κάθε $k \in \mathbb{N}$ και $\sum_{k=0}^{\infty} \binom{\alpha+k-1}{k} p^\alpha q^k = p^\alpha (1-q)^{-\alpha} = 1$, λόγω της σχέσης (2.17).

Η πιθανογεννήτρια της αρνητικής διωνυμικής κατανομής δίδεται από την

$$G(z) = \sum_{k=0}^{\infty} \binom{\alpha+k-1}{k} p^\alpha q^k z^k = \left(\frac{p}{1-qz}\right)^\alpha.$$

Η μέση τιμή της αρνητικής διωνυμικής τυχαίας μεταβλητής X είναι $\mathbb{E}X = G'(1) = \alpha q \frac{p^\alpha}{(1-q)^{\alpha+1}}$ ή

$$\mathbb{E}X = \alpha \frac{q}{p}.$$

Ομοίως, $\mathbb{E}X(X-1) = G''(1) = \alpha(\alpha+1) q^2 \frac{p^\alpha}{(1-q)^{\alpha+2}} = \alpha(\alpha+1)\left(\frac{q}{p}\right)^2$. Τηυς ωε ηαε $\mathbb{E}X^2 = \alpha(\alpha+1)\left(\frac{q}{p}\right)^2 + \alpha\frac{q}{p}$ ανδ τηυς $\mathsf{Var}(X) = \alpha(\alpha+1)\left(\frac{q}{p}\right)^2 + \alpha\frac{q}{p} - \left(\alpha\frac{q}{p}\right)^2 = \alpha\frac{q}{p}\left(1 + \frac{q}{p}\right)$ ή

$$\mathsf{Var}(X) = \alpha \frac{q}{p^2}.$$

Όταν $\alpha = m \in \mathbb{N}$ η αρνητική διωνυμική κατανομή είναι η κατανομή του αθροίσματος m ανεξάρτητων τυχαίων μεταβλητών γεωμετρικά κατανεμημένων (του τύπου (2.14», η κάθε μια με πιθανότητα επιτυχίας p. Αυτό αποδεικνύεται εύκολα με την βοήθεια των πιθανογεννητριών.

Επίσης

$$\mathbb{E}[X(X-1)\cdots(X-k+1)] = G^{(k)}(1). \tag{2.19}$$

Η έκφραση (2.19) ονομάζεται κατιούσα παραγοντική ροπή τάξης k. Οι συνήθεις ροπές μπορούν να υπολογιστούν από αυτές.

Κεφάλαιο 3

Πολυμεταβλητές Κατανομές

3.1 Πολυωνυμική Κατανομή

Έστω ένα πείραμα που μπορεί να έχει m διαφορετικά αποτελέσματα, τα οποία ονομάζουμε $a_1, a_2, \ldots, a_m$. Η πιθανότητα το αποτέλεσμα του πειράματος να είναι το a_i είναι p_i και $\sum_{i=1}^{m} p_i = 1$. Ένα τέτοιο παράδειγμα είναι η η ρίψη ενός τίμιου ζαριού όπου υπάρχουν έξι διαφορετικά ισοπίθανα αποτελέσματα. Εκτελούμε το πείραμα n φορές και υποθέτουμε ότι κάθε επανάληψη είναι ανεξάρτητη από όλες τις άλλες. Έστω X_1 ο αριθμός των πειραμάτων στα οποία το αποτέλεσμα ήταν a_1, X_2 ο αριθμός των πειραμάτων στα οποία το αποτέλεσμα ήταν a_1 και X_n ο αριθμός των πειραμάτων στα οποία το αποτέλεσμα ήταν a_m. Τα $X_1, X_2, \ldots, X_m$ είναι τυχαίες μεταβλητές οι οποίες βεβαίως δεν είναι ανεξάρτητες μεταξύ τους. Ισχύει μάλιστα, αναγκαστικά, ότι $X_1 + X_2 + \cdots + X_m = n$. Η από κοινού κατανομή των τυχαίων αυτών μεταβλητών ονομάζεται πολυωνυμική. Αν $n_1, n_2, \ldots, n_m$ είναι μη αρνητικοί ακέραιοι, τέτοιοι ώστε $\sum_{i=1}^{m} n_i = n$, τότε

$$P(X_1 = n_1, X_2 = n_2, \ldots, X_m = n_m) = \binom{n}{n_1, n_2, \ldots, n_m} p_1^{n_1} p_2^{n_2} \cdots p_m^{n_m}. \tag{3.1}$$

Ας δούμε πρώτα ένα παράδειγμα. Έστω ένα πείραμα που έχει τρία δυνατά αποτελέσματα, τα a_1, a_2, και a_3 τα οποία συμβαίνουν με πιθανότητες p_1, p_2 και p_3 αντίστοιχα, όπου βεβαίως $p_1 + p_2 + p_3 = 1$ εφόσον ένα από τα τρία αποτελέσματα θα συμβεί κάθε φορά. Ας υποθέσουμε ότι το εκτελούμε 10 φορές και κάθε επανάληψη είναι ανεξάρτητη από όλες τις άλλες. X_i, $i = 1, 2, 3$, είναι ο αριθμός των πειραμάτων των οποίων το αποτέλεσμα ήταν a_i. Ας υπολογίσουμε την πιθανότητα $P(X_1 = 3, X_2 = 5, X_3 = 2)$. Η πιθανότητα να έχουμε τα αποτελέσματα

$$a_1 a_1 a_1 a_2 a_2 a_2 a_2 a_2 a_3 a_3$$

είναι $p_1^3 p_2^5 p_3^2$ λόγω της ανεξαρτησίας των πειραμάτων. Η πιθανότητα μιας οποιασδήποτε άλλης πραγματοποίησης με 3 a_1, 5 a_2 και 2 a_3 είναι η ίδια. Το μόνο θέμα που υπάρχει είναι να δούμε πόσες διαφορετικές πραγματοποιήσεις υπάρχουν με τον συγκεκριμένο αριθμό των τριών αποτελεσμάτων, με πόσους διαφορετικούς τρόπους δηλαδή μπορούμε να βάλουμε στη σειρά 3 a_1, 5 a_2 και 2 a_3. Αυτός ο αριθμός όμως είναι ο αριθμός των επαναληπτικών μεταθέσεων που είναι $\binom{10}{3,5,2} = 2520$. Συνεπώς $P(X_1 = 3, X_2 = 5, X_3 = 2) = \binom{10}{3,5,2} p_1^3 p_2^5 p_3^2$.

Ως ένα δεύτερο παράδειγμα επιβεβαιώστε ότι η πιθανότητα σε 6 ρίψεις ενός τιμίου ζαριού να έχουμε 2 δυάρια, 2 τεσσάρια και 2 εξάρια είναι

$$\binom{6}{0,2,0,2,0,2}\left(\frac{1}{6}\right)^0\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^0\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^0\left(\frac{1}{6}\right)^2 = \frac{6!}{2!0!2!0!2!0!}\frac{1}{6^6} = 0,0019.$$

Η από κοινού κατανομή του αριθμού των διαφορετικών αποτελεσμάτων,

$$(X_1, X_2, \ldots, X_m),$$

ονομάζεται *πολυωνυμική κατανομή*. Το αθροισμα των πολυωνυμικών πιθανοτήτων (3.1) για όλα τα $n_1, n_2, \ldots, n_m$ τέτοια ώστε $\sum_{i=1}^m n_i = n$ είναι μονάδα λόγω του πολυωνυμικού θεωρήματος:

$$\sum_{\{(n_1,\ldots,n_m):\sum_{i=1}^m n_i = n\}} \binom{n}{n_1, n_2, \ldots, n_m} p_1^{n_1} p_2^{n_2} \cdots p_m^{n_m} \;=\; (p_1 + \cdots + p_m)^n = 1^n = 1.$$

Η περιθώρια κατανομή της X_i μπορεί να υπολογισθεί αθροίζοντας τις πολυωνυμικές πιθανότητες.

Ας δούμε ξανά το πείραμα που έχει τρία διαφορετικά αποτελέσματα, a_1, a_2, και a_3 τα οποία συμβαίνουν με πιθανότητες p_1, p_2 και p_3 αντίστοιχα, όπου $p_1 + p_2 + p_3 = 1$, και το οποίο επαναλαμβάνουμε ανεξάρτητα n φορές. Έστω X_1, X_2, X_3, ο αριθμός των αποτελεσμάτων τύπου a_1, a_2, και a_3 στις n επαναλήψεις. Η από κοινού κατανομή των (X_1, X_2, X_3) θα είναι, σύμφωνα με τον τύπο (3.1) για $m = 3$,

$$P(X_1 = i, X_2 = j, X_3 = k) = \binom{n}{i, j, k} p_1^i p_2^j p_3^k. \tag{3.2}$$

Για να βρούμε την περιθώρια κατανομή του X_1 αρκεί να αθροίσουμε ως εξής:

$$P(X_1 = i) \;=\; \sum_{\{(j,k):j+k=n-i,\ j,k\in\mathbb{N}_0\}} \binom{n}{i, j, k} p_1^i p_2^j p_3^k \qquad i = 0, 1, 2, \ldots, n.$$

Λαμβάνοντας υπ΄ όψιν ότι $i + j + k = n$ και

$$\binom{n}{i, j, k} \;=\; \frac{n!}{(n-i)!i!}\frac{(n-i)!}{(n-i-j)!j!} \;=\; \binom{n}{i}\binom{n-i}{j}$$

η παραπάνω έκφραση γράφεται ισοδύναμα ως

$$
\begin{aligned}
P(X_1 = i) &= \sum_{j=0}^{n-i} \binom{n}{i}\binom{n-i}{j} p_1^i p_2^j p_3^{n-i-j} = \binom{n}{i} p_1^i \sum_{j=0}^{n-i} \binom{n-i}{j} p_2^j p_3^{n-i-j} \\
&= \binom{n}{i} p_1^i (p_2 + p_3)^{n-i} = \binom{n}{i} p_1^i (1 - p_1)^{n-i} \qquad (3.3)
\end{aligned}
$$

όπου, στην προτελευταία ισότητα χρησιμοποιήσαμε το διωνυμικό θεώρημα και στην τελευταία το γεγονός ότι $p_2 + p_3 = 1 - p_1$. Από την (3.3) είναι σαφές ότι η X_1 έχει Διωνυμική κατανομή με πιθανότητα επιτυχίας p_1 και αριθμό δοκιμών n. Αυτό μπορεί να το καταλάβει κανείς άμεσα θεωρόντας ότι το πείραμα δεν έχει τρία αποτελέσματα αλλά μόνο δύο: το αποτέλεσμα a_1 που συμβαίνει με πιθανότητα p_1 και το αποτέλεσμα «όχι a_1» (δηλαδή a_2 ή a_3) που συμβαίνει με πιθανότητα $p_2 + p_3 = 1 - p_1$.

Με τον ίδιο ακριβώς τρόπο μπορούμε να αποδείξουμε γενικά (και οι δείκτες δεν πρέπει να μας αποθαρρύνουν) ότι η περιθώριος κατανομή του X_i για την γενική πολυωνυμική κατανομή (3.1) είναι διωνυμική:

$$
P(X_i = n_i) = \binom{n}{n_i} p_i^{n_i} (1 - p_i)^{n-n_i}, \qquad n_i = 0, 1, \ldots, n. \qquad (3.4)
$$

Από τις ιδιότητες της διωνυμικής κατανομής συμπεραίνουμε ότι $\mathsf{E}X_i = np_i$ και $\mathsf{Var}(X_i) = np_i(1 - p_i)$.

Στη συνέχεια μπορούμε να υπολογίσουμε την από κοινού κατανομή των (X_1, X_2) για την πολυωνυμική κατανομή (3.2) με τρία διαφορετικά αποτελέσματα. Αν i, j είναι δύο μη αρνητικοί ακέραιοι τέτοιοι ώστε $i+j \le n$ τότε, λαμβάνοντας υπ' όψιν ότι $i+j+k = n$ έχουμε

$$
\begin{aligned}
P(X_1 = i, X_2 = j) &= \binom{n}{i, j, k} p_1^i p_2^j p_3^k \\
&= \binom{n}{i, j, n-i-j} p_1^i p_2^j (1 - p_1 - p_2)^{n-i-j}, \qquad i + j \le n.
\end{aligned}
$$

Προκειμένου να υπολογίσουμε την συνδιακύμανση των X_1, X_2, υπολογίζουμε πρώτα την μέση τιμή του γινομένου

$$
\begin{aligned}
\mathsf{E}[X_1 X_2] &= \sum_{\{(i,j):\ i+j \le n,\ i \ge 1,\ j \ge 1\}} ij \binom{n}{i, j, n-i-j} p_1^i p_2^j (1 - p_1 - p_2)^{n-i-j} \\
&= \sum_{\{(i,j):\ i+j \le n,\ i \ge 1,\ j \ge 1\}} \frac{n!}{(i-1)!(j-1)!(n-i-j)!} p_1^i p_2^j (1 - p_1 - p_2)^{n-i-j} \\
&= n(n-1)p_1 p_2 \sum_{\{(i,j):\ i-1+j-1 \le n-2,\ i-1 \ge 0,\ j-1 \ge 0\}} \frac{(n-2)!}{(i-1)!(j-1)!(n-2-(i-1)-(j-1))!} \\
&\qquad \times p_1^{i-1} p_2^{j-1} (1 - p_1 - p_2)^{n-2-(i-1)-(j-1)}.
\end{aligned}
$$

Αν θέσουμε $i-1=i'$, $j-1=j$, το τελευταίο σκέλος της παραπάνω εξίσωσης γράφεται ως

$$n(n-1)p_1p_2 \sum_{\{(i',j'):\ i'+j'\le n-2,\ i'\ge 0,\ j'\ge 0\}} \frac{(n-2)!}{i'!j'!(n-2-i'-j')!} \times p_1^{i'} p_2^{j'} (1-p_1-p_2)^{n-2-i'-j'}$$

Ας θέσουμε $k' = n-2-i'-j'$. Παρατηρείστε ότι όταν προσδιορίζονται τα i', j', προσδιορίζεται ταυτόχρονα και το k'. Το παραπάνω άθροισμα μπορεί να γραφτεί ισοδύναμα ως

$$\begin{aligned} &n(n-1)p_1p_2 \sum_{\{(i',j',k'):\ i'+j'+k'=n-2,\ i',j',k'\ge 0\}} \frac{(n-2)!}{i'!j'!k'!} p_1^{i'} p_2^{j'} (1-p_1-p_2)^{k'} \\ &n(n-1)p_1p_2(p_1+p_2+1-p_1-p_2)^{n-2} = n(n-1)p_1p_2 \end{aligned} \tag{3.5}$$

όπου χρησιμοποιήσαμε το πολυωνυμικό θεώρημα. Συνεπώς

$$\mathsf{Cov}(X_1,X_2) = E[X_1X_2] - EX_1EX_2 = n(n-1)p_1p_2 - np_1np_2 = -np_1p_2. \tag{3.6}$$

Ο συντελεστής συσχέτισης ανάμεσα στο X_1 και X_2 είναι

$$\rho_{12} = -\frac{np_1p_2}{\sqrt{np_1(1-p_1)np_2(1-p_2)}} = -\sqrt{\frac{p_1p_2}{(1-p_1)(1-p_2)}}. \tag{3.7}$$

Παρατηρήστε ότι ο συντελεστής συσχέτισης είναι αρνητικός. Αυτό είναι λογικό. Δεδομένου ότι ο αριθμός των πειραμάτων είναι σταθερός, αν έχουμε μεγάλο αριθμό αποτελεσμάτων a_1 τότε αναγκαστικά θα τείνουμε να έχουμε μικρότερο αριθμό αποτελεσμάτων a_2.

Πρόβλημα 3.1. Ρίχνουμε ένα ζάρι 10 φορές. Ποιά είναι η πιθανότητα να πάρουμε 4 εξάρια, 3 πεντάρια, 2 τεσσάρια και ένα τριάρι (με οποιαδήποτε σειρά);
(Απ. $\binom{10}{4,3,2,1}\frac{1}{6^{10}}$.)

3.2 Η πολυμεταβλητή υπεργεωμετρική κατανομή

Ας υποθέσουμε ότι σε μια κάλπη έχουμε K_1 μπάλες χρώματος 1, K_2 μπάλες χρώματος 2, κλπ. και K_m μπάλες χρώματος m. Συνολικά δηλαδή στην κάλπη υπάρχουν $N := K_1 + K_2 + \cdots + K_m$ μπάλες, ασχέτως χρώματος. Από αυτές παίρνουμε, χωρίς επανατοποθέτηση, n μπάλες με $n \le N$. Έστω X_i ο αριθμός των μπαλών χρώματος i, $i = 1,2,\dots,m$ στο δείγμα μεγέθους n που πήραμε. Τότε

$$P(X_1 = n_1, X_2 = n_2, \dots, X_m = n_m) = \frac{\binom{K_1}{n_1}\binom{K_2}{n_2}\cdots\binom{K_m}{n_m}}{\binom{N}{n}}, \qquad n_i \in \mathbb{N}_0,\ \sum_{i=1}^{m} n_i = n. \tag{3.8}$$

Επισημαίνουμε πάλι ότι μπορούμε να χρησιμοποιήσουμε την παραπάνω έκφραση ακόμα και στην περίπτωση που $n_i > K_i$ μια και ο αντίστοιχος διωνυμικός συντελεστής μηδενίζεται. Συνεπώς μπορούμε να χρησιμοποιήσουμε την έκφραση (3.8) για κάθε διάνυσμα μη αρνητικών ακεραίων $(n_1, \ldots, n_m)$.

Ο λόγος που ισχύει η (3.8) είναι ότι μπορούμε να διαλέξουμε n_i μπάλες χρώματος i από K_i που υπάρχουν στην κάλπη με $\binom{K}{n_i}$ τρόπους ενώ μπορούμε να διαλέξουμε n μπάλες ασχέτως χρώματος από τις N που υπάρχουν στην κάλπη με $\binom{N}{n}$ τρόπους.

Ας επαληθεύσουμε τώρα ότι

$$\sum_{\{(n_1,\ldots,n_m)\}:\sum_{i=1}^m n_i = n} P(X_1 = n_1, \ldots, X_m = n_m) \; = \; 1. \tag{3.9}$$

Προκειμένου να το δείξουμε αρκεί να δούμε ότι

$$\sum_{\{(n_1,\ldots,n_m)\}:\sum_{i=1}^m n_i = n} \binom{K_1}{n_1}\binom{K_2}{n_2}\cdots\binom{K_m}{n_m} \;\; = \;\; \binom{N}{n}. \tag{3.10}$$

Ας εξετάσουμε το γινόμενο

$$(1 + x_1)^{K_1}(1 + x_2)^{K_2} \cdots (1 + x_m)^{K_m}.$$

Ο συντελεστής του όρου $x_1^{n_1} x_2^{n_2} \cdots x_m^{n_m}$ θα είναι

$$\binom{K_1}{n_1}\binom{K_2}{n_2}\cdots\binom{K_m}{n_m}.$$

Αν θέσουμε τώρα $x_1 = x_2 = \cdots = x_m = x$ τότε

$$(1 + x)^{K_1}(1 + x)^{K_2} \cdots (1 + x)^{K_m} = (1 + x)^N \tag{3.11}$$

Με βάση τις παραπάνω παρατηρήσεις, ο συντελεστής του x^n ($n \le N = \sum_{i=1}^m K_i$) στο αριστερό μέλος της (3.11) δίνεται από το αριστερό μέλος της (3.10) ενώ ο συντελεστής του x^n στο δεξί μέλος της (3.11) δίνεται από το δεξί μέλος της (3.10).

Η περιθώρια κατανομή του X_i είναι εύκολο να διαπιστώσουμε ότι είναι υπεργεωμετρική. Αρκεί να θεωρήσουμε ότι όλες οι μπάλες που δεν είναι χρώματος i είναι κάποιου άλλου χρώματος που δεν μας ενδιαφέρει και να διαπιστώσουμε ότι

$$P(X_i = n_i) = \frac{\binom{K_i}{n_i}\binom{N-K_i}{n-n_i}}{\binom{N}{n}}, \qquad n_i = 0, 1, \ldots, n. \tag{3.12}$$

Με τον ίδιο τρόπο μπορούμε να δούμε ότι η από κοινού κατανομή των X_i, X_j είναι

$$P(X_i = n_i, X_2 = n_j) = \frac{\binom{K_i}{n_i}\binom{K_j}{n_j}\binom{N-K_i-K_j}{n-n_i-n_j}}{\binom{N}{n}}, \qquad n_i + n_j \le n. \tag{3.13}$$

Η μέση τιμή και η διασπορά των X_i, X_j δίνεται από τις γνωστές εκφράσεις για την υπεργεωμετρική κατανομή.

$$\mathsf{E}X_i = n\frac{K_i}{N}, \qquad \mathsf{Var}(X_i) = n\,\frac{K_i}{N}\,\frac{N-K_i}{N}\,\frac{N-n}{N-1}. \tag{3.14}$$

Για να προσδιορίσουμε την συνδιακύμανση αρκεί να δούμε ότι

$$\begin{aligned}
\mathsf{E}X_1X_2 &= \sum_{(n_1,n_2):\, n_i\geq 0, n_j\geq 0, \{n_1+n_2\leq n\}} n_1 n_2 \frac{\binom{K_i}{n_i}\binom{K_j}{n_j}\binom{N-K_i-K_j}{n-n_i-n_j}}{\binom{N}{n}} \\
&= \frac{K_iK_j}{\binom{N}{n}} \sum_{\{(n_1,n_2):\, n_i\geq 1, n_2\geq 1,\, n_1+n_2\leq n\}} \binom{K_i-1}{n_i-1}\binom{K_j-1}{n_j-1}\binom{N-K_i-K_j}{n-n_i-n_j}
\end{aligned}$$

Αν θέσουμε $l_i = n_i - 1$, $l_j = n_j - 1$ η παραπάνω σχέση μπορεί να γραφεί ως

$$\begin{aligned}
\mathsf{E}X_1X_2 &= \frac{K_iK_j}{\binom{N}{n}} \sum_{\{(l_1,l_2):\, l_i\geq 0, l_2\geq 0,\, l_1+l_2\leq n-2\}} \binom{K_i-1}{l_i}\binom{K_j-1}{l_j}\binom{N-2-(K_i-1)-(K_j-1)}{n-2-l_i-l_j} \\
&= K_1K_2\,\frac{\binom{N-2}{n-2}}{\binom{N}{n}} = n(n-1)\frac{K_1K_2}{N(N-1)}.
\end{aligned}$$

Συνεπώς η συνδιακύμανση των X_1, X_2 είναι

$$\begin{aligned}
\mathsf{Cov}(X_1,X_2) &= \mathsf{E}[X_1X_2] - \mathsf{E}X_1\mathsf{E}X_2 = n(n-1)\frac{K_1K_2}{N(N-1)} - n\frac{K_1}{N}\,n\frac{K_2}{N} \\
&= -\frac{n(N-n)}{N-1}\,\frac{K_1}{N}\,\frac{K_2}{N}.
\end{aligned} \tag{3.15}$$

Τέλος ο συντελεστής συσχέτισης προσδιορίζεται εύκολα από την (3.15) και την (3.14) και μετά από εύκολες απλοποιήσεις προκύπτει ότι

$$\rho = -\sqrt{\frac{K_iK_j}{(N-K_i)(N-K_j)}}\,. \tag{3.16}$$

Πρόβλημα 3.2. Ένα κιβώτιο περιέχει χίλιους λαμπτήρες, 100 από τους οποίους δεν λειτουργούν. Οι χίλιοι αυτοί λαμπτήρες τοποθετούνται τυχαία σε 10 κουτιά. Ποιά είναι η πιθανότητα και οι 100 χαλασμένοι λαμπτήρες να βρεθούν στο ίδιο κουτί; Αν X_i, $i = 1, 2, \ldots, 10$, είναι ο αριθμός των χαλασμένων δοχείων στο κουτί i, ποιά είναι η από κοινού κατανομή των τυχαίων μεταβλητών αυτών;

Πρόβλημα 3.3. Μια κάλπη έχει N_1 σφαιρίδια που γράφουν τον αριθμό 1, N_2 που γράφουν τον αριθμό 2 και N_3 που γράφουν τον αριθμό 3. Διαλέγω (χωρίς επανατοποθέτηση) n σφαιρίδια. Αν X_i ο αριθμός των σφαιριδίων, μεταξύ αυτών που επέλεξα, που γράφουν τον αριθμό i, $i = 1, 2, 3$, να βρείτε την από κοινού κατανομή των X_1, X_2, X_3. Επίσης να ευρεθούν οι μέσοι, διακυμάνσεις και συνδιακυμάνσεις.

Κεφάλαιο 4

Από κοινού κατανομή δύο συνεχών τυχαίων μεταβλητών

4.1 Από κοινού συνάρτηση κατανομής δύο συνεχών τυχαίων μεταβλητών

Έστω X, Y, δύο τυχαίες μεταβλητές. Η από κοινού συνάρτηση κατανομής είναι η συνάρτηση

$$F(x, y) := \mathbb{P}(X \le x, Y \le y), \qquad x, y \in \mathbb{R}. \tag{4.1}$$

Η από κοινού συνάρτηση κατανομής έχει τις εξής ιδιότητες:

1. Είναι αύξουσα ως προς κάθε ένα από τα δύο ορίσματα, δηλαδή αν $x \le x'$ και $y \le y'$ τότε

$$F(x, y) \le F(x', y) \qquad \text{και} \qquad F(x, y) \le F(x, y'). \tag{4.2}$$

2. Ισχύει ότι

$$\begin{aligned}
\lim_{x\to\infty} F(x,y) &= \mathbb{P}(Y \le y) =: F_Y(y) \quad \text{(Η περιθώρια συνάρτηση κατανομής της Y)}\\
\lim_{y\to\infty} F(x,y) &= \mathbb{P}(X \le x) =: F_X(x) \quad \text{(Η περιθώρια συνάρτηση κατανομής της X)}\\
\lim_{\substack{x\to\infty\\ y\to\infty}} F(x,y) &= 1\\
\lim_{x\to-\infty} F(x,y) &= 0 \ \text{ για κάθε } y \text{ και } \lim_{y\to-\infty} F(x,y) = 0 \ \text{ για κάθε } x.
\end{aligned}$$

3. Για κάθε $x_1 < x_2$ και $y_1 < y_2$ ισχύει ότι

$$\mathbb{P}(x_1 < X \le x_2, y_1 < Y \le y_2) = F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) + F(x_1, y_1) \ge 0. \tag{4.3}$$

Η θεμελιώδης σχέση (4.3) αποδεικνύεται ως εξής: Ορίζουμε τα σύνολα

$$\begin{aligned}
&A_{11} = \{X \le x_1, Y \le y_1\}, \quad A_{12} = \{X \le x_1, Y \le y_2\}, \quad A_{21} = \{X \le x_2, Y \le y_1\},\\
&A_{22} = \{X \le x_2, Y \le y_2\}, \quad R = \{x_1 < X \le x_2, y_1 < Y \le y_2\}.
\end{aligned}$$

Παρατηρείστε ότι $R \cap (A_{12} \cup A_{21}) = \emptyset$ και ότι $A_{22} = R \cup (A_{12} \cup A_{21})$. Συνεπώς

$$\mathbb{P}(A_{22}) = \mathbb{P}(R) + \mathbb{P}(A_{12} \cup A_{21})$$

Επίσης $A_{12} \cap A_{21} = A_{11}$ και συνεπώς

$$\mathbb{P}(A_{12} \cap A_{21}) = \mathbb{P}(A_{12}) + \mathbb{P}(A_{21}) - \mathbb{P}(A_{11}).$$

Από τις παραπάνω σχέσεις προκύπτει ότι

$$\mathbb{P}(R) = \mathbb{P}(A_{22}) - \mathbb{P}(A_{12}) - \mathbb{P}(A_{21}) + \mathbb{P}(A_{11}).$$

Η απόδειξη είναι πλήρης δεδομένου ότι $\mathbb{P}(A_{ij}) = F(x_i, y_j)$ για $i, j \in \{1, 2\}$ και ότι $\mathbb{P}(R)$ είναι το αριστερό μέρος της (4.3).

4.2 Μερικές παράγωγοι και διπλά ολοκληρώματα - Μια πρακτική εισαγωγή

Θεωρούμε μια συνάρτηση δύο μεταβλητών $\phi(x, y)$. Η *μερική παράγωγος* της ϕ ως προς την μεταβλητή x συμβολίζεται με $\frac{\partial}{\partial x}\phi$ και ορίζεται ως το όριο

$$\frac{\partial}{\partial x}\phi(x, y) \; = \; \lim_{\delta \to 0} \frac{\phi(x + \delta, y) - \phi(x, y)}{\delta}. \tag{4.4}$$

Αντίστοιχα ορίζουμε την μερική παράγωγο της ϕ ως προς y ως

$$\frac{\partial}{\partial y}\phi(x, y) \; = \; \lim_{\delta \to 0} \frac{\phi(x, y + \delta) - \phi(x, y)}{\delta}.$$

Όπως βλέπουμε από τον ορισμό η μερική παράγωγος ως προς x υπολογίζεται κρατώντας το y σταθερό και παραγωγίζοντας ως προς x. Για παράδειγμα

$$\begin{aligned}
&\frac{\partial}{\partial x}(x^2 y) = 2xy, \qquad \frac{\partial}{\partial x}\left(e^{-xy} + \frac{x}{y} + y\right) = -ye^{-xy} + \frac{1}{y}, \\
&\frac{\partial}{\partial y}\left(e^{-xy} + \frac{x}{y} + y\right) = -xe^{-xy} - \frac{x}{y^2} + 1, \qquad \frac{\partial}{\partial x}\frac{2x + y}{y - x} = \frac{2}{y - x} + \frac{2x + y}{(y - x)^2}.
\end{aligned}$$

Η μερικές παράγωγοι $\frac{\partial \phi}{\partial x}$, $\frac{\partial \phi}{\partial y}$, θα είναι εν γένει συναρτήσεις των x και y και επομένως (αν είναι παραγωγίσιμες) θα μπορούμε να τις παραγωγίσουμε εκ νέου. Αποδεικνύεται ότι, ύπο ορισμένες συνθήκες που πάντοτε ικανοποιούύνται στις περιπτώσεις που θα εξετάσουμε, οι δεύτερες μερικές παράγωγοι $\frac{\partial}{\partial y}\left(\frac{\partial}{\partial x}\phi\right)$ και $\frac{\partial}{\partial x}\left(\frac{\partial}{\partial y}\phi\right)$ είναι πάντα ίσες. Η κοινή τιμή τους ονομάζεται δεύτερη μεικτή μερική παράγωγος ως προς x και y και συμβολίζεται ως $\frac{\partial^2}{\partial x \partial y}\phi$ ή $\frac{\partial^2}{\partial y \partial x}\phi$ αφού η σειρά με την οποία εκτελούνται οι παραγωγήσεις δεν επηρεάζει το τελικό αποτέλεσμα. Δείτε τα ακόλουθα παραδείγματα:

$$\frac{\partial}{\partial y}\left(\frac{\partial}{\partial x}(x^2 y)\right) = \frac{\partial}{\partial y}(2xy) = 2x, \qquad \frac{\partial}{\partial x}\left(\frac{\partial}{\partial y}(x^2 y)\right) = \frac{\partial}{\partial y}\left(x^2\right) = 2x$$

$$\frac{\partial}{\partial y}\left(\frac{\partial}{\partial x}\left(e^{-xy}+\frac{x}{y}+y\right)\right)=\frac{\partial}{\partial y}\left(-ye^{-xy}+\frac{1}{y}\right)=(xy-1)e^{-xy}-\frac{1}{y^2}$$
$$\frac{\partial}{\partial x}\left(\frac{\partial}{\partial y}\left(e^{-xy}+\frac{x}{y}+y\right)\right)=\frac{\partial}{\partial x}\left(-xe^{-xy}-\frac{x}{y^2}+1\right)=(xy-1)e^{-xy}-\frac{1}{y^2}$$

4.2.1 Διπλά ολοκληρώματα

Έστω g μια συνάρτηση δύο μεταβλητών ορισμένη σε ένα παραλλληλόγραμμο χωρίο του $\mathbb{R}^2$, $g:[a,b]\times[c,d]\to\mathbb{R}^+$. Η συνάρτηση αυτή ορίζει μια επιφάνεια αν θεωρήσουμε το σύνολο των σημείων του $\mathbb{R}^3$ $\{(x,y,z): x\in[a,b], y\in[c,d], z=g(x,y)\}$ καθώς επίσης και ένα στερεό, το υποσύνολο του $\mathbb{R}^3$ που απαρτίζεται από τα σημεία

$$\mathscr{S}=\{(x,y,z): x\in[a,b], y\in[c,d], 0\le z\le g(x,y)\}$$

. (Έχουμε υποθέσει ότι η g παίρνει μη αρνητικές τιμές στο πεδίο ορισμού της.) Έστω $\mathscr{P}_x:=\{x_i\}_{i=0,1,\dots,m}$ μια διαμέριση του $[a,b]$ δηλαδή $a=x_0<x_1<x_2<\dots<x_m=b$ και παρομοίως $\mathscr{P}_y:=\{y_j\}_{j=0,1,\dots,n}$ μια διαμέριση του $[c,d]$ με $c=y_0<y_1<\dots<y_n=d$. Κατ΄ αυτό τον τρόπο διαμερίσουμε το ορθογώνιο $[a,b]\times[c,d]$ σε $m\cdot n$ ορθογώνια, $[x_i,x_i+\Delta x_i]\times[y_j,y_j+\Delta y_j]$ όπου $\Delta x_i=x_{i+1}-x_i$ και $\Delta y_j=y_{j+1}-y_j$. Συμβολίζουμε την διαμέριση ως $\mathscr{P}:=\{(x_i,y_j)\}_{\substack{i=0,1,\dots,m\\ j=0,1,\dots,n}}$ και ως μέγεθός της ορίζουμε την ποσότητα $\|\mathscr{P}\|:=\max(\|\mathscr{P}_x\|,\|\mathscr{P}_y)$ όπου $\|\mathscr{P}_x\|=\max_{i=0,1,\dots,m-1}|\Delta x_i|$ και $\|\mathscr{P}_y\|:=\max_{j=0,1,\dots,n-1}|\Delta y_j|$. Αν V είναι ο όγκος του στερεού $\mathscr{S}$ τότε η ποσότητα

$$\sum_{i=0}^{m-1}\sum_{j=0}^{n-1} g(x_i,y_j)\Delta x_i\Delta y_j \tag{4.5}$$

αποτελεί μια προσέγγιση του όγκου από ορθογώνια παραλληλεπίπεδα. Αποδεικνύεται ότι, για ομαλές συναρτήσεις g (πρακτικά μιλώντας, σε κάθε περίπτωση που πρόκειται να μας απασχολήσει) ο όγκος του στερεού $\mathscr{S}$ είναι το όριο του αθροίσματος της (4.5) όταν το μέγεθος της διαμέρισης $\mathscr{P}$ τείνει στο 0 (πράγμα που σημαίνει ότι τόσο το m όσο και το n τείνουν στο άπειρο) δηλαδή

$$V = \lim_{\|\mathscr{P}\|\to 0}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1} g(x_i,y_j)\Delta x_i\Delta y_j. \tag{4.6}$$

Στον μαθηματικό λογισμό αποδεικνύεται ότι, κάτω από συνθήκες που πάντα ισχύουν στην πράξη, μπορούμε να υπολογίσουμε το όριο στην (4.6) ως ένα διπλό όριο

$$V = \lim_{\|\mathscr{P}_x\|\to 0}\lim_{\|\mathscr{P}_y\|\to 0}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1} g(x_i,y_j)\Delta x_i\Delta y_j, \qquad \text{ή ως} \tag{4.7}$$
$$V = \lim_{\|\mathscr{P}_y\|\to 0}\lim_{\|\mathscr{P}_x\|\to 0}\sum_{j=0}^{n-1}\sum_{i=0}^{m-1} g(x_i,y_j)\Delta x_i\Delta y_j. \tag{4.8}$$

(Εν γένει, η σειρά με την οποία υπολογίζει κανείς ένα επαναλαμβανόμενο όριο έχει σημασία και το τελικό αποτέλεσμα μπορεί να εξαρτάται από αυτήν. Εν προκειμένω όμως το θεώρημα του Fubini μας εξασφαλίζει ότι το τελικό αποτέλεσμα είναι ανεξάρτητο από το αν θα πάρουμε πρώτα το όριο $\|\mathscr{P}_x\| \to 0$ ή το $\|\mathscr{P}_y\| \to 0$) Στην περίπτωση που υπολογίζουμε πρώτα το όριο ως προς $\mathscr{P}_y$ (4.7) έχουμε

$$V = \lim_{\|\mathscr{P}_x\|\to 0} \sum_{i=0}^{m-1} \Delta x_i \left(\lim_{\|\mathscr{P}_y\|\to 0} \sum_{j=0}^{n-1} g(x_i, y_j)\Delta y_j \right) \tag{4.9}$$

Για κάθε συγκεκριμένη τιμή x_i, θεωρούμε την συνάρτηση μίας πραγματικής μεταβλητής $g(x_i, y)$ την οποία θεωρούμε ολοκληρώσιμη. Το εσωτερικό όριο είναι ένα άθροισμα Riemann και συνεπώς, αν θέσουμε

$$G(x_i) := \int_c^d g(x_i, y)dy$$

θα έχουμε

$$\lim_{\|\mathscr{P}_y\|\to 0} \sum_{j=0}^{n-1} g(x_i, y_j)\Delta y_j = G(x_i).$$

Αντικαθιστώντας το εσωτερικό όριο με την τιμή αυτή στην (4.9) παίρνουμε

$$V = \lim_{\|\mathscr{P}_x\|\to 0} \sum_{i=0}^{m-1} \Delta x_i G_2(x_i) = \int_a^b G(x)dx = \int_a^b \left(\int_c^d g(x, y)dy \right) dx \tag{4.10}$$

Παίρνοντας τα όρια στην αντίθετη σειρά, δηλαδή πρώτα $\lim_{\|\mathscr{P}_x\|\to 0}$ και μετά $\lim_{\|\mathscr{P}_y\|\to 0}$ καταλήγουμε στο συμπέρασμα ότι ισχύει επίσης

$$V = \int_c^d \left(\int_a^b g(x, y)dx \right) dy. \tag{4.11}$$

Η κοινή τιμή του ορίου είναι το διπλό ολοκλήρωμα

$$\int\int_{[a,b]\times[c,d]} g(x, y)dxdy = \int_a^b \int_c^d g(x, y)dydx = \int_c^d \int_a^b g(x, y)dxdy.$$

4.3 Η από κοινού πυκνότητα πιθανότητας

Παρατηρούμε ότι

$$\mathbb{P}(x < X \le x + \delta, Y \le y) = \mathbb{P}(X \le x + \delta, Y \le y) - \mathbb{P}(X \le x, Y \le y) = F(x + \delta, y) - F(x, y)$$

Επομένως

$$\lim_{\delta\to 0} \frac{1}{\delta} \mathbb{P}(x < X \le x + \delta, Y \le y) = \lim_{\delta\to 0} \frac{F(x + \delta, y) - F(x, y)}{\delta} = \frac{\partial}{\partial x} F(x, y). \tag{4.12}$$

Συμβολίζουμε την συνάρτηση δυο μεταβλητών $\frac{\partial}{\partial x}F(x,y)$ ως $F_x(x,y)$.

Διαισθητικά θέλουμε να ορίσουμε την πυκνότητα πιθανότητας έτσι ώστε, αν Δx και Δy είναι μικρές μεταβολές των τιμών x και y,

$$\mathbb{P}(X \in (x, x+\Delta x], Y \in (y, y+\Delta y]) \approx f(x,y)\,\Delta x \Delta y$$

Η από κοινού πυκνότητα πιθανότητας στο σημείο (x,y) ορίζεται ως

$$f(x,y) \;=\; \lim_{h\to 0}\lim_{\delta\to 0}\frac{1}{h\delta}\,\mathbb{P}(x < X < x+\delta, y < X \le y+h) \tag{4.13}$$

και, λαμβάνοντας υπ΄ όψιν μας την θεμελιώδη σχέση (4.3),

$$\begin{aligned} f(x,y) \;&=\; \lim_{h\to 0}\lim_{\delta\to 0}\frac{1}{h\delta}\,\left(F(x+\delta,y+h) - F(x,y+h) - F(x+\delta,y) + F(x,y)\right) \\ &= \lim_{h\to 0}\frac{1}{h}\left(\lim_{\delta\to 0}\frac{F(x+\delta,y+h)-F(x,y+h)}{\delta} - \lim_{\delta\to 0}\frac{F(x+\delta,y)-F(x,y)}{\delta}\right). \end{aligned}$$

Χρησιμοποιώντας την (4.12) βλέπουμε ότι τα εσωτερικά όρια (όπου $\delta \to 0$) δίδουν $F_x(x,y+h) - F_x(x,y)$. Συνεπώς

$$f(x,y) \;=\; \lim_{h\to 0}\frac{1}{h}\,\left(F_x(x,y+h) - F_x(x,y)\right) \;=\; \frac{\partial}{\partial y}F_x(x,y).$$

Λαμβάνοντας υπ΄ όψιν τον ορισμό της συνάρτησης F_x έχουμε

$$f(x,y) \;=\; \frac{\partial^2}{\partial x \partial y}F(x,y). \tag{4.14}$$

Από το θεμελιώδες θεώρημα του Απειροστικού Λογισμού, για $x_1 < x_2$,

$$F(x_2,y) - F(x_1,y) = \int_{x_1}^{x_2}\frac{\partial}{\partial x}F(x,y)dx = \int_{x_1}^{x_2}F_x(x,y)dx. \tag{4.15}$$

Χρησιμοποιώντας ξανά το θεμελιώδες θεώρημα του Απειροστικού Λογισμού, για $y_1 < y_2$,

$$F_x(x,y_2) - F_x(x,y_1) = \int_{y_1}^{y_2}\frac{\partial}{\partial y}F_x(x,y)dx = \int_{y_1}^{y_2}\frac{\partial^2}{\partial y \partial x}F(x,y)dx = \int_{y_1}^{y_2}f(x,y)dx \tag{4.16}$$

όπου στις παραπάνω εξισώσεις χρησιμοποιήσαμε και την (4.14). Χρησιμοποιώντας τις (4.15) και (4.16) παίρνουμε

$$\begin{aligned} F(x_2,y_2) - F(x_1,y_2) - (F(x_2,y_1) - F(x_1,y_1)) \;&=\; \int_{x_1}^{x_2}\left(F_x(x,y_2) - F_x(x,y_1)\right)dx && (4.17) \\ &=\; \int_{x_1}^{x_2}\int_{y_1}^{y_2} f(x,y)dydx. && (4.18) \end{aligned}$$

Συνεπώς ισχύει ότι

$$\begin{aligned} \mathbb{P}(x_1 < X \le x_2, y_1 < Y \le y_2) \;&=\; F(x_2,y_2) - F(x_1,y_2) - F(x_2,y_1) + F(x_1,y_1) \\ &=\; \int_{x_1}^{x_2}\int_{y_1}^{y_2} f(x,y)dxdy. && (4.19) \end{aligned}$$

Από την σχέση αυτή αφήνοντας $x_1, y_1 \to -\infty$ και $x_2, y_2 \to \infty$ έχουμε

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy \;=\; 1. \tag{4.20}$$

Συναρτήσεις δύο τυχαίων μεταβλητών. Έστω $g : \mathbb{R}^2\mathbb{R}$ μια συνάρτηση δυο μεταβλητών. Η ποσότητα $g(X,Y)$ είναι βεβαίως τυχαία μεταβλητή και η μέση της τιμή υπολογίζεται ως

$$\mathbb{E}[g(X,Y)] \;=\; \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f(x,y)dxdy. \tag{4.21}$$

4.3.1 Οι περιθώριες πυκνότητες πιθανότητας

Από την σχέση (4.19), αφήνοντας το $y_1 \to -\infty$ το $y_2 \to \infty$, και το $x_1 \to -\infty$ παίρνουμε την περιθόρια κατανομή του X.

$$F_X(x) = \mathbb{P}(X \le x) = \mathbb{P}(-\infty < X \le x, -\infty < Y < \infty) = \int_{-\infty}^{x}\left(\int_{-\infty}^{\infty} f(\xi,y)dy\right)d\xi. \tag{4.22}$$

Παραγωγίζουμε την σχέση αυτή παίρνουμε την *περιθώρια πυκνότητα πιθανότητας της* X ως

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy. \tag{4.23}$$

Η περιθώρια πυκνότητα πιθανότητας της Y δίδεται βεβαίως από την σχέση $f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx$.

Η μέση τιμή της τυχαίας μεταβλητής X δίδεται βεβαίως από την σχέση

$$\mathbb{E}X = \int_{-\infty}^{\infty} x f_X(x)dx.$$

Παρατηρείστε ότι τα παραπάνω είναι συνεπή με την εξής έννοια. Αν g είναι η συνάρτηση δυο μεταβλητών που δίνεται από την σχέση $g(x,y) = x$ για κάθε $x, y \in \mathbb{R}$ τότε η (4.21) δίνει

$$\mathbb{E}[g(X,Y)] = \mathbb{E}X = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f(x,y)dydx = \int_{-\infty}^{\infty} x\left(\int_{-\infty}^{\infty} f(x,y)dy\right)dx = \int_{-\infty}^{\infty} x f_X(x)dx.$$

4.3.2 Συνδιακύμανση

Όπως και στην περίπτωση των διακριτών τ.μ. η συνδιακύμανση δύο συνεχών τ.μ. ορίζεται ως $\mathrm{Cov}(X,Y) = \mathbb{E}[XY] - (\mathbb{E}X)(\mathbb{E}Y)$. Λαμβάνοντας υπ΄ όψιν την (4.21) θα έχουμε

$$\mathrm{Cov}(X,Y) \;=\; \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)dxdy \;-\; \int_{-\infty}^{\infty} x f_X(x)dx \int_{-\infty}^{\infty} y f_Y(y)dy.$$

4.3.3 Ανεξαρτησία

Δύο τυχαίες μεταβλητές X, Y ονομάζονται *ανεξάρτητες* αν και μόνο αν η από κοινού πυκνότητα πιθανότητάς τους είναι το γινόμενο των περιθωρίων πυκνοτήτων, αν δηλαδή

$$f(x, y) \; = \; f_X(x) f_Y(y), \qquad \text{για κάθε } x, y \in \mathbb{R}. \tag{4.24}$$

Η σχέση αυτή είναι ισοδύναμη με την συνθήκη να είναι η από κοινού συνάρτηση κατανομής των X, Y το γινόμενο των περιθωρίων συναρτήσεων κατανομής, δηλαδή

$$F(x, y) \; = \; F_X(x) F_Y(y), \qquad \text{για κάθε } x, y \in \mathbb{R}. \tag{4.25}$$

4.3.4 Η δεσμευμένη πυκνότητα πιθανότητας

Όπως είδαμε, αν οι X, Y, είναι *διακριτές τυχαίες μεταβλητές* που παίρνουν τις τιμές x_i, $i = 1, 2, 3, \dots$, και y_j, $j = 1, 2, 3, \dots$, αντίστοιχα με από κοινού πιθανότητα $\mathbb{P}(X = x_i, Y = y_j)$ η δεσμευμένη κατανομή της X δεδομένου ότι $Y = y_j$ υπολογίζεται με βάση τον γνωστό ορισμό για τις δεσμευμένες πιθανότητες, δηλαδή

$$\mathbb{P}(X = x_i | Y = y_j) \; = \; \frac{\mathbb{P}(X = x_i, Y = y_j)}{\mathbb{P}(Y = y_j)}, \qquad \text{για κάθε } i = 1, 2, \dots \tag{4.26}$$

Στην παραπάνω έκφραση η τιμή της y_j είναι σταθερή και εκφράζει το αποτέλεσμα της παρατήρησης της τιμής της Y. Η (4.26) εκφράζει την κατανομή πιθανότητας της X *δεδομένης της πληροφορίας ότι η τιμή της* Y *είναι* y_j.

Θέλουμε να επεκτείνουμε τον παραπάνω ορισμό στην όταν οι τυχαίες μεταβλητές X και Y είναι συνεχείς. Η δυσκολία που εμφανίζεται είναι ότι η πιθανότητα η τυχαία μεταβλητή Y να πάρει μια συγκεκριμένη τιμή στην περίπτωση αυτή είναι 0 και επομένως το αντίστοιχο κλάσμα, $\frac{\mathbb{P}(X=x, Y=y)}{\mathbb{P}(Y=y)}$ δεν έχει νόημα γιατί τόσο ο αριθμητής όσο και ο παρονομαστής είναι 0. Θα προσεγγίσουμε επομένως το πρόβλημα χρησιμοποιώντας την έννοια του ορίου ως εξής. Ορίζουμε πρώτα την δεσμευμένη συνάρτηση κατανομής της X δεδομένου ότι $Y = y$:

$$\begin{aligned}
F_{X|Y}(x|y) \; &= \; \mathbb{P}(X \le x | Y = y) \; := \; \lim_{h \to 0} \mathbb{P}(X \le x | y < Y \le y + h) \\
&= \; \lim_{h \to 0} \frac{\mathbb{P}(X \le x, y < Y \le y + h)}{\mathbb{P}(y < Y \le y + h)} \; = \; \frac{\lim_{h \to 0} \frac{1}{h} \mathbb{P}(X \le x, y < Y \le y + h)}{\lim_{h \to 0} \frac{1}{h} \mathbb{P}(y < Y \le y + h)} \\
&= \; \frac{\lim_{h \to 0} \frac{1}{h} \left(F(x, y + h) - F(x, y) \right)}{\lim_{h \to 0} \frac{1}{h} \left(F_Y(y + h) - F_Y(y) \right)} \; = \; \frac{\frac{\partial}{\partial y} F(x, y)}{\frac{d}{dy} F_Y(y)} \; = \; \frac{\frac{\partial}{\partial y} F(x, y)}{f_Y(y)}.
\end{aligned} \tag{4.27}$$

Συμπεραίνουμε επομένως ότι η *δεσμευμένη συνάρτηση κατανομής* της X δεδομένου ότι η $Y = y$ δίνεται από την σχέση

$$F_{X|Y}(x|y) \; = \; \frac{\frac{\partial}{\partial y} F(x, y)}{f_Y(y)}. \tag{4.28}$$

Η δεσμευμένη πυκνότητα πιθανότητας θα είναι επομένως η μερική παράγωγος ως προς x της (4.28) δηλαδή

$$\frac{\partial}{\partial x}\frac{\frac{\partial}{\partial y}F(x,y)}{f_Y(y)} = \frac{\frac{\partial^2}{\partial x \partial y}F(x,y)}{f_Y(y)}$$

και λαμβάνοντας υπ΄ όψιν μας την (4.14)

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}. \tag{4.29}$$

Έχοντας ορίσει την δεσμευμένη πυκνότητα πιθανότητας μπορούμε εύκολα να ορίσουμε την δεσμευμένη μέση τιμή της X ή οποιασδήποτε συνάρτησης της X ως εξής:

$$\mathbb{E}[X|Y=y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y)dx \tag{4.30}$$

$$\mathbb{E}[X|Y=y] = \int_{-\infty}^{\infty} g(x) f_{X|Y}(x|y)dx \tag{4.31}$$

Παρατηρούμε ότι η $\mathbb{E}[X|Y=y]$ εξαρτάται εν γένει από την τιμή y που παίρνει η Y, και επομένως είναι συνάρτηση της y.

Πρόταση 4.1. *Έστω $\phi(y) := \mathbb{E}[X|Y=y]$. Η μέση τιμή της τυχαίας μεταβλητής $\phi(Y)$ ισούται με την μέση τιμή της X δηλαδή*

$$\mathbb{E}[\phi(Y)] = \int_{-\infty}^{\infty} \phi(y) f_Y(y)dy = \int_{-\infty}^{\infty} \mathbb{E}[X|Y=y] f_Y(y)dy = \mathbb{E}[X] \tag{4.32}$$

Απόδειξη. Λαμβάνοντας υπ΄ όψιν την (4.31) και την (4.29) έχουμε

$$\begin{aligned}\int_{y=-\infty}^{\infty} \mathbb{E}[X|Y=y] f_Y(y)dy &= \int_{y=-\infty}^{\infty}\left(\int_{x=-\infty}^{\infty} x f_{X|Y}(x|y)dx\right) f_Y(y)dy \\ &= \int_{y=-\infty}^{\infty}\left(\int_{x=-\infty}^{\infty} x \frac{f(x,y)}{f_Y(y)}dx\right) f_Y(y)dy.\end{aligned}$$

Το τελευταίο διπλό ολοκλήρωμα γράφεται ως

$$\int_{y=-\infty}^{\infty} \frac{1}{f_Y(y)}\left(\int_{x=-\infty}^{\infty} x\, f(x,y)dx\right) f_Y(y)dy = \int_{y=-\infty}^{\infty}\left(\int_{x=-\infty}^{\infty} x\, f(x,y)dx\right) dy \tag{4.33}$$

επειδή, το εσωτερικό ολοκλήρωμα είναι ως προς x και η $f_Y(y)$ εξαρτάται μόνο από το y και επομένως είναι σταθερά ως προς x. Εναλλάσσοντας την σειρά ολοκλήρωσης στο τελευταίο διπλό ολοκλήρωμα έχουμε

$$\begin{aligned}\int_{x=-\infty}^{\infty}\left(\int_{y=-\infty}^{\infty} x\, f(x,y)dx\right) dy &= \int_{x=-\infty}^{\infty} x\left(\int_{y=-\infty}^{\infty} f(x,y)dx\right) dy \\ &= \int_{x=-\infty}^{\infty} x\, f_Y(y)dy = \mathbb{E}[X].\end{aligned}$$

□

4.3.5 Η διμεταβλητή κανονική κατανομή

Η διμεταβλητή κανονική κατανομή έχει από κοινού πυκνότητα πιθανότητας

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}\, e^{-\frac{1}{2(1-\rho^2)}\left(\frac{(x-\mu_x)^2}{\sigma_x^2} - 2\rho\frac{(x-\mu_x)}{\sigma_x}\frac{(y-\mu_y)}{\sigma_y} + \frac{(y-\mu_y)^2}{\sigma_x^2}\right)} \qquad \forall (x,y) \in \mathbb{R}\times\mathbb{R}. \tag{4.34}$$

Υπολογίζουμε τις περιθώριες πυκνότητες πιθανότητας ως εξής: θέτοντας $u = \frac{x-\mu_x}{\sigma_x}$, $\nu = \frac{y-\mu_y}{\sigma_y}$ (και επομένως $d\nu = \frac{1}{\sigma_y}\, dy$) έχουμε

$$\begin{aligned}
f_X(x) &= \int_{-\infty}^{\infty} f(x,y)dy \;=\; \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}\, e^{-\frac{1}{2(1-\rho^2)}\left(\frac{(x-\mu_x)^2}{\sigma_x^2} - 2\rho\frac{(x-\mu_x)}{\sigma_x}\frac{(y-\mu_y)}{\sigma_y} + \frac{(y-\mu_y)^2}{\sigma_x^2}\right)} dy \\
&= \frac{1}{2\pi\sigma_x\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)}\left(u^2 - 2\rho u\nu + \nu^2\right)} d\nu \\
&= \frac{1}{2\pi\sigma_x\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)}\left(u^2 - \rho^2u^2 + \rho^2u^2 - 2\rho u\nu + \nu^2\right)} d\nu \\
&= \frac{1}{\sigma_x\sqrt{2\pi}}\, e^{-\frac{u^2}{2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}}\, e^{-\frac{(\nu-\rho u)^2}{2(1-\rho^2)}}\, d\nu \\
&= \frac{1}{\sigma_x\sqrt{2\pi}}\, e^{-\frac{u^2}{2}} \;=\; \frac{1}{\sigma_x\sqrt{2\pi}}\, e^{-\frac{(x-\mu_x)^2}{2\sigma_x^2}},
\end{aligned}$$

οπου χρησιμοποιήσαμε το γεγονός ότι η συνάρτηση

$$\phi(\nu) := \frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}}\, e^{-\frac{(\nu-\rho u)^2}{2(1-\rho^2)}}, \qquad \nu \in \mathbb{R}$$

είναι η πυκνότητα πιθανότητας μιας κανονικής τυχαίας μεταβλητής με μέση τιμή ρu και διασπορά $1-\rho^2$ και επομένως το ολοκλήρωμά $\int_{-\infty}^{\infty}\phi(u)du = 1$. Συνεπώς οι περιθώριες πυκνότητες είναι

$$f_X(x) \;=\; \frac{1}{\sigma_x\sqrt{2\pi}}\, e^{-\frac{(x-\mu_x)^2}{2\sigma_x^2}}, \tag{4.35}$$

και με παρόμοιο τρόπο

$$f_Y(y) \;=\; \frac{1}{\sigma_y\sqrt{2\pi}}\, e^{-\frac{(y-\mu_y)^2}{2\sigma_y^2}}. \tag{4.36}$$

Παρατηρείστε ότι, αν $\rho = 0$ η (4.34) γίνεται

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y}\, e^{-\frac{1}{2}\left(\frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma_x^2}\right)} = f_X(x)f_Y(y)$$

και επομένως, στην περίπτωση αυτή οι τυχαίες μεταβλητές X και Y είναι ανεξάρτητες. Από τις (4.35), (4.36), προκύπτει ότι οι τυχαίες μεταβλητές X και Y είναι κανονικά κατανεμημένες: $\mathcal{N}(\mu_x, \sigma_x^2)$ και $\mathcal{N}(\mu_y, \sigma_y^2)$.

Η συνδιακύμανση υπολογίζεται ευκολότερα ως $\mathsf{Cov}(X,Y) = \mathbb{E}[(X-\mu_x)(Y-\mu_y)]$.

$$\mathbb{E}\left[\left(\frac{X-\mu_x}{\sigma_x}\right)\left(\frac{Y-\mu_y}{\sigma_y}\right)\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right)\frac{e^{-\frac{1}{2(1-\rho^2)}\left(\frac{(x-\mu_x)^2}{\sigma_x^2}-2\rho\frac{(x-\mu_x)}{\sigma_x}\frac{(y-\mu_y)}{\sigma_y}+\frac{(y-\mu_y)^2}{\sigma_x^2}\right)}}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}\,dxdy.$$

Θέτοντας $u := \frac{x-\mu_x}{\sigma_x}$, $v := \frac{y-\mu_y}{\sigma_y}$ έχουμε

$$\mathbb{E}\left[\left(\frac{X-\mu_x}{\sigma_x}\right)\left(\frac{Y-\mu_y}{\sigma_y}\right)\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} uv\frac{e^{-\frac{1}{2(1-\rho^2)}\left(u^2-2\rho uv+v^2\right)}}{2\pi\sqrt{1-\rho^2}}\,dudv.$$

Χρησιμοποιώντας το γεγονός ότι $u^2 - 2\rho uv + v^2 = u^2 - 2\rho uv - \rho^2v^2 + v^2 - \rho^2 v^2 = (u-\rho v)^2 + (1-\rho^2)v^2$ γράφουμε το διπλό ολοκλήρωμα ως

$$\int_{v=-\infty}^{\infty}\left(\int_{u=-\infty}^{\infty} uv\frac{e^{-\frac{1}{2(1-\rho^2)}\left((u-\rho v)^2+(1-\rho^2)v^2\right)}}{2\pi\sqrt{1-\rho^2}}\,du\right)dv = \int_{v=-\infty}^{\infty} v\,\frac{e^{-\frac{v^2}{2}}}{\sqrt{2\pi}}\left(\int_{u=-\infty}^{\infty} u\,\frac{e^{-\frac{(u-\rho v)^2}{2(1-\rho^2)}}}{\sqrt{2\pi}\sqrt{1-\rho^2}}\,du\right)dv.$$

Παρατηρείστε ότι η ποσότητα $\frac{e^{-\frac{(u-\rho v)^2}{2(1-\rho^2)}}}{\sqrt{2\pi}\sqrt{1-\rho^2}}$ είναι η πυκνότητα πιθανότητας μιας κανονικής τυχαίας μεταβλητής με μέση τιμή ρv και διασπορά $1-\rho^2$. Συνεπώς το εσωτερικό ολοκλήρωμα είναι η μέση τιμή αυτής της κανονικής τυχαίας μεταβλητής, δηλαδή $\int_{u=-\infty}^{\infty} u\,\frac{e^{-\frac{(u-\rho v)^2}{2(1-\rho^2)}}}{\sqrt{2\pi}\sqrt{1-\rho^2}}\,du = \rho v$ Επομένως

$$\mathbb{E}\left[\left(\frac{X-\mu_x}{\sigma_x}\right)\left(\frac{Y-\mu_y}{\sigma_y}\right)\right] = \int_{v=-\infty}^{\infty}\rho v^2\frac{1}{\sqrt{2\pi}}\,e^{-\frac{1}{2}v^2}dv = \rho.$$

(Χρησιμοποιήσαμε το γεγονός ότι $\int_{v=-\infty}^{\infty} v^2\frac{1}{\sqrt{2\pi}}\,e^{-\frac{1}{2}v^2}dv = 1$ αφού είναι η διασπορά της τυποποιημένης κανονικής τυχαίας μεταβλητής.) Επομένως

$$\mathsf{Cov}(X,Y) = \rho\sigma_x\sigma_y \tag{4.37}$$

και συνεπώς ρ είναι ο συντελεστής συσχέτισης ανάμεσα στην X και στην Y.

Τέλος, υπολογίζουμε την δεσμευμένη πυκνότητα πιθανότητας της X δεδομένου ότι $Y = y$. Για απλούστευση των εκφράσεων χρησιμοποιούμε τις ποσότητες $u := \frac{x-\mu_x}{\sigma_x}$, $v := \frac{y-\mu_y}{\sigma_y}$.

$$\begin{aligned} f_{X|Y}(x|y) &= \frac{f(x,y)}{f_Y(y)} = \frac{e^{-\frac{1}{2(1-\rho^2)}\left(u^2-2\rho uv+v^2\right)}}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}\,\sigma_y\sqrt{2\pi}\,e^{\frac{v^2}{2}} = \frac{e^{-\frac{1}{2(1-\rho^2)}\left(u^2-2\rho uv+v^2\right)}}{\sqrt{2\pi}\,\sigma_x\sqrt{1-\rho^2}} \\ &= \frac{e^{-\frac{1}{2(1-\rho^2)}\left((u-\rho v)^2(1-\rho^2)v^2\right)+\frac{1}{2}v^2}}{\sqrt{2\pi}\,\sigma_x\sqrt{1-\rho^2}} = \frac{1}{\sqrt{2\pi}\,\sigma_x\sqrt{1-\rho^2}}\,e^{-\frac{1}{2(1-\rho^2)}(u-\rho v)^2}. \end{aligned}$$

Αντικαθιστώντας τα u, v με τις τιμές τους παίρνουμε

$$f_{X|Y}(x|y) = \frac{1}{\sqrt{2\pi}\,\sigma_x\sqrt{1-\rho^2}}\,e^{-\frac{1}{2\sigma_x^2(1-\rho^2)}\left(x-\frac{\rho\sigma_x}{\sigma_y}(y-\mu_y)\right)^2}, \quad x\in\mathbb{R}. \tag{4.38}$$

Από την παραπάνω έκφραση διαπιστώνουμε ότι, δεδομένου ότι $Y = y$, η X είναι *κανονικά κατανεμημένη* με μέσο $\frac{\rho\sigma_x}{\sigma_y}(y-\mu_y)$ και διασπορά $\sigma_x^2(1-\rho^2)$.

4.4 Παραδειγματα

Παράδειγμα 1.

Έστω X και Y δύο συνεχείς τυχαίες μεταβλητές με τιμές στο σύνολο $[0,1]$ ορισμένες στον ίδιο χώρο πιθανοτήτων με από κοινού πυκνότητα πιθανότητας που ορίζεται ως

$$f(x,y) = \begin{cases} \dfrac{3}{4} + x(1-y) & \text{αν } 0 \le x \le 1 \text{ και } 0 \le y \le 1, \\ 0 & \text{διαφορετικά.} \end{cases}$$

α) Να υπολογίσετε τις περιθώριες πυκνότητες των τυχαίων μεταβλητών X και Y.
β) Είναι οι X και Y ανεξάρτητες;
γ) Να υπολογίσετε την μέση τιμή των τυχαίων μεταβλητών X και Y
δ) Να υπολογίσετε τις διασπορές $\mathsf{Var}(X)$, $\mathsf{Var}(Y)$ καθώς και την συνδιακύμανση $\mathsf{Cov}(X,Y)$.
ε) Να υπολογίσετε την δεσμευμένη πυκνότητα πιθανότητας $f_{X|Y}(x|y)$ της τυχαίας μεταβλητής X δεδομένου ότι $Y = y$.

$$\begin{aligned} f_X(x) &= \int_0^1 f(x,y)dy = \int_0^1 \left(\frac{3}{4} + x(1-y)\right) dy = \frac{3}{4} + x\int_0^1 (1-y)dy = \frac{3}{4} + x\left(\int_0^1 dy - \int_0^1 y dy\right) \\ &= \frac{3}{4} + x\left(1 - \frac{1}{2}\right) = \frac{3}{4} + \frac{x}{2}. \end{aligned}$$

$$f_Y(y) = \int_0^1 f(x,y)dx = \int_0^1 \left(\frac{3}{4} + x(1-y)\right) dx = \frac{3}{4} + (1-y)\int_0^1 x dx = \frac{3}{4} + (1-y)\frac{1}{2} = \frac{5}{4} - \frac{y}{2}.$$

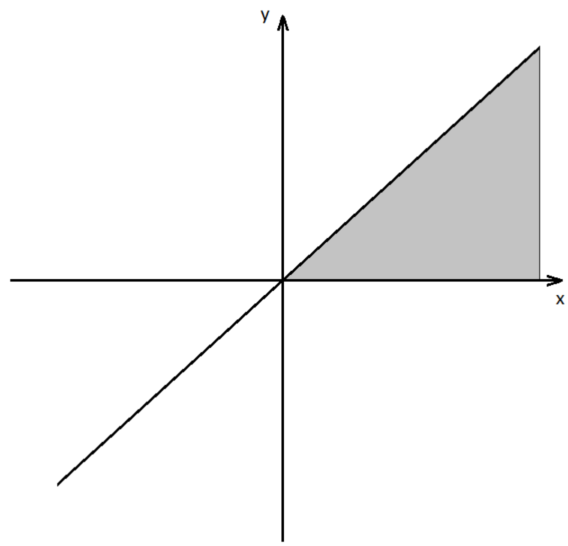
Δεδομένου ότι δεν ισχύει $\frac{3}{4} + x(1-y) = \left(\frac{3}{4} + \frac{x}{2}\right)\left(\frac{5}{4} - \frac{y}{2}\right)$ για κάθε x και y, οι X και Y δεν είναι ανεξάρτητες.

$$\mathbb{E}X = \int_0^1 x f_X(x)dx = \int_0^1 \left(\frac{3}{4}x + \frac{x^2}{2}\right) dx = \frac{3}{4}\int_0^1 x dx + \frac{1}{2}\int_0^1 x^2 dx = \frac{3}{4}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{3} = \frac{13}{24}.$$

$$\mathbb{E}Y = \int_0^1 y f_Y(y)dy = \int_0^1 \left(\frac{5}{4}y - \frac{y^2}{2}\right) dy = \frac{5}{4}\int_0^1 y dy - \frac{1}{2}\int_0^1 y^2 dy = \frac{5}{4}\cdot\frac{1}{2} - \frac{1}{2}\cdot\frac{1}{3} = \frac{11}{24}.$$

$$\mathbb{E}X^2 = \int_0^1 x^2 f_X(x)dx = \int_0^1 \left(\frac{3}{4}x^2 + \frac{x^3}{2}\right) dx = \frac{3}{4}\int_0^1 x^2 dx + \frac{1}{2}\int_0^1 x^3 dx = \frac{3}{4}\cdot\frac{1}{3} + \frac{1}{2}\cdot\frac{1}{4} = \frac{3}{8}.$$

$$\mathbb{E}Y^2 = \int_0^1 y^2 f_Y(y)dy = \int_0^1 \left(\frac{5}{4}y^2 - \frac{y^3}{2}\right) dy = \frac{5}{4}\int_0^1 y^2 dy - \frac{1}{2}\int_0^1 y^3 dy = \frac{5}{4}\cdot\frac{1}{3} - \frac{1}{2}\cdot\frac{1}{4} = \frac{7}{24}.$$



Σχήμα 4.1: Η γκρίζα περιοχή είναι εκείνη όπου η δείκτρια $\mathbf{1}(0 \le y \le x)$ έχει την τιμή 1.

$$\begin{aligned}
\mathbb{E}[XY] &= \int_0^1\int_0^1 \left(\frac{3}{4}xy + x^2(1-y)y\right)dxdy = \frac{3}{4}\int_0^1\int_0^1 xy\,dxdy + \int_0^1\int_0^1 x^2(1-y)y\,dxdy \\
&= \frac{3}{4}\left(\int_0^1 xdx\right)\left(\int_0^1 ydy\right) + \left(\int_0^1 x^2dx\right)\left(\int_0^1 (1-y)ydy\right) \\
&= \frac{3}{4}\cdot\frac{1}{2}\cdot\frac{1}{2} + \frac{1}{3}\left(\int_0^1 ydy - \int_0^1 y^2dy\right) \;=\; \frac{35}{144}.
\end{aligned}$$

Με βάση τους υπολογισμούς αυτούς βρίσκουμε $\mathsf{Var}(X) = \mathsf{Var}(Y) = 0.0816$, $\mathsf{Cov}(X,Y) = -\frac{3}{(24)^2}$, $\rho_{X,Y} = -0.0638$.

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} = \frac{3+4x(1-y)}{5-2y}.$$

$$\mathbb{E}[X|Y=y] = \int_0^1 x\frac{3+4x(1-y)}{5-2y}dx = \frac{3}{5-2y}\int_0^1 xdx + \frac{4}{5-2y}\int_0^1 x^2dx = \frac{17-8y}{6(5-2y)}.$$

Παράδειγμα 2.

Έστω $f(x,y) = e^{-x}\,\mathbf{1}(0 \le y \le x)$ η από κοινού πικνότητα πιθανότητας των τυχαίων μεταβλητών X, Y. Η *δείκτρια συνάρτηση* στην πυκνότητα πιθανότας ορίζεται ως

$$\mathbf{1}(0 \le x \le y) \;=\; \begin{cases} 1 & \text{αν } \; 0 \le x \le y \\ 0 & \text{διαφορετικά} \end{cases}$$

(επομένως εδώ η δείκτρια είναι μια συνάρτηση δύο μεταβλητών που έχει την τιμή 1 στο γραμμοσκιασμένο τμήμα του $\mathbb{R}^2$ στο σχήμα 4.1

Η περιθώρια πυκνότητα της τυχαίας μεταβλητής X θα είναι (αν $x \geq 0$)

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f(x,y)dy = \int_{-\infty}^{\infty} \mathbf{1}(0 \leq y \leq x)e^{-x}dy = e^{-x}\int_{-\infty}^{\infty} \mathbf{1}(0 \leq y \leq x)dy \\ &= e^{-x}\int_0^x dy = x\,e^{-x} \end{aligned}$$

Η πυκνότητα αυτή είναι Gamma(2,1). Παρομοίως, η περιθώρια πυκνότητα της τυχαίας μεταβλητής Y θα είναι

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx = \int_{-\infty}^{\infty} \mathbf{1}(0 \leq y \leq x)e^{-x}dx = \int_y^{\infty} e^{-x}dx = e^{-y}$$

δηλαδή εκθετική (Gamma(1,1)). Η δεσμευμένες πυκνότητες πιθανότητας είναι

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{e^{-x}\,\mathbf{1}(0 \leq y \leq x)}{x\,e^{-x}} = \frac{1}{x}\,\mathbf{1}(0 \leq y \leq x).$$

Η πυκνότητα αυτή είναι *ομοιόμορφη* στο διάστημα $[0, x]$. Επίσης

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} = \frac{e^{-x}\,\mathbf{1}(0 \leq y \leq x)}{e^{-y}} = e^{-(x-y)}\mathbf{1}(0 \leq y \leq x).$$

Η πυκνότητα αυτή είναι μια εκθετική πυκνότητα μετατοπισμένη στα δεξιά κατά y.

Στη συνέχεια θα υπολογίσουμε τις πρώτες ροπές της X και Y. Έχουμε

$$\mathbb{E}X = \int_0^{\infty} f_X(x)x dx = \int_0^{\infty} x^2 e^{-x}dx = \Gamma(3) = 2! = 2$$

και

$$\mathbb{E}[X^2] = \int_0^{\infty} f_X(x)x^2 dx = \int_0^{\infty} x^3 e^{-x}dx = \Gamma(4) = 3! = 6.$$

Συνεπώς $\mathsf{Var}(X) = 6 - 2^2 = 2$. Επίσης

$$\mathbb{E}Y = \int_0^{\infty} f_Y(y)y dy = \int_0^{\infty} y e^{-y}dy = \Gamma(2) = 1$$

και

$$\mathbb{E}Y^2 = \int_0^{\infty} f_Y(y)y^2 dy = \int_0^{\infty} y^2 e^{-y}dy = \Gamma(3) = 2.$$

Η διασπορά είναι $\mathsf{Var}(Y) = 2 - 1^2 = 1$. Τέλος έχουμε

$$\begin{aligned} \mathbb{E}[XY] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)xy\,dxdy = \int_{x=0}^{\infty}\left(\int_{y=0}^{\infty} \mathbf{1}(y \leq x)xye^{-x}dy\right)dx \\ &= \int_{x=0}^{\infty} xe^{-x}\left(\int_{y=0}^{x} y dy\right)dx = \int_{x=0}^{\infty} \frac{x^3}{2}\,e^{-x}dx \\ &= \frac{1}{2}\Gamma(4) = \frac{3!}{2} = 3. \end{aligned}$$

Η συνδιακύμανση είναι $\mathsf{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y] = 3 - 1 \cdot 2 = 1$ και ο συντελεστής συσχέτισης $\rho = \frac{1}{\sqrt{2\cdot 1}} = \frac{\sqrt{2}}{2}$.

Τέλος οι δεσμευμένες μέσες τιμές είναι

$$\begin{aligned}
\mathbb{E}[X|Y=y] &= \int_0^\infty x f_{X|Y}(x|y)dx = \int_0^\infty e^{-(x-y)}\mathbf{1}(0\le y\le x)x dx = \int_y^\infty e^{-(x-y)}x dx \\
&= \int_y^\infty e^{-(x-y)}(x-y+y)dx = \int_y^\infty e^{-(x-y)}(x-y)dx + \int_y^\infty e^{-(x-y)}y dx \\
&= \int_0^\infty e^{-u}u du + y\int_0^\infty e^{-u}dx = \Gamma(2)+y\cdot 1 = y+1
\end{aligned}$$

και

$$\mathbb{E}[Y|X=x] = \int_0^\infty y f_{Y|X}(y|x)dy = \int_0^\infty \frac{1}{x}\mathbf{1}(0\le y\le x)y dy = \int_0^x \frac{y}{x}dy = \frac{x}{2}.$$

Φυσικά,

$$\mathbb{E}X = \int_0^\infty f_Y(y)\mathbb{E}[X|Y=y]dy = \int_0^\infty (y+1)e^{-y}dy = \int_0^\infty y e^{-y}dy + \int_0^\infty e^{-y}dy = 1+1=2,$$

και

$$\mathbb{E}Y = \int_0^\infty f_X(x)\mathbb{E}[Y|X=x]dx = \int_0^\infty \frac{x}{2}x e^{-x}dx = \frac{\Gamma(3)}{2} = 1.$$

Παράδειγμα 3.

Εστω $f(x,y)=\lambda y e^{-yx-\lambda y}$ για $x,y\ge 0$ και 0 διαφορετικά, όπου $\lambda>0$ θετική παράμετρος. Η περιθώρια πυκνότητα της X είναι

$$\begin{aligned}
f_X(x) &= \int_0^\infty \lambda y e^{-\lambda y - xy}dy = \lambda\int_0^\infty y e^{-(x+\lambda)y}dy = \frac{\lambda}{(\lambda+x)^2}\int_0^\infty (\lambda+x)^2 y\, e^{-(\lambda+x)y}\,dy \\
&= \frac{\lambda}{(\lambda+x)^2}.
\end{aligned}$$

Στην τελευταία σχέση χρησιμοποιήσαμε το γεγονός ότι $\alpha^2 y e^{-\alpha y}$ είναι η πυκνότητα της κατανομής Gamma(2,α) (όπου $\alpha=\lambda+x$) και επομένως το ολοκήρωμα ισούται με την μονάδα. Η πυκνότητα $f_x(x)=\frac{\lambda}{(\lambda+x)^2}$, $x\ge 0$ ονομάζεται Pareto και η αντίστοιχη περιθώρια συνάρτηση κατανομής είναι

$$F_X(x) = \int_0^x \frac{\lambda}{(t+\lambda)^2}dt = \lambda\int_\lambda^{(\lambda+x)}\frac{dt}{t^2} = \lambda\left(\frac{1}{\lambda}-\frac{1}{\lambda+x}\right) = 1-\frac{\lambda}{\lambda+x}, \qquad x\ge 0.$$

Η περιθώρια πυκνότητα της Y είναι

$$f_Y(y) = \int_0^\infty \lambda y e^{-\lambda y - xy}dx = \lambda e^{-\lambda y}\int_0^\infty y e^{-xy}dx = \lambda e^{-\lambda y}.$$

Η πυκνότητα αυτή είναι εκθετική (με ρυθμό λ).

Η δεσμευμένες πυκνότητες είναι

$$f_{X|Y}(x|y) \;=\; \frac{\lambda y e^{-yx-\lambda y}}{\lambda e^{-\lambda y}} \;=\; y\, e^{-yx}.$$

Παρατηρείστε ότι η *πυκνότητα αυτή είναι εκθετική με ρυθμό* y. Εδώ y είναι μια θετική παράμετρος (ο ρυθμός) ενώ $x \geq 0$ είναι η μεταβλητή. Αντιστοίχως,

$$f_{Y|X}(y|x) \;=\; \frac{\lambda y e^{-yx-\lambda y}}{\frac{\lambda}{(\lambda+x)^2}} \;=\; (\lambda+x)^2 y\, e^{-(\lambda+x)y}.$$

Παρατηρείστε ότι η πυκνότητα αυτή είναι Gamma($2,\lambda+x$).

Στη συνέχεια υπολογίζουμε τις πρώτες ροπές των X και Y. Παρατηρούμε (ίσως με έκπληξη ...) ότι η μέση τιμή της X είναι άπειρη διότι το γενικευμένο ολοκλήρωμα

$$\int_0^\infty x f_X(x)\,dx \;=\; \int_0^\infty x \frac{\lambda}{(\lambda+x)^2}\,dx$$

αποκλίνει στο $+\infty$. Αντιθέτως $\mathbb{E}[Y] = \int_0^\infty y f_Y(y)\,dy = \int_0^\infty y\lambda e^{-\lambda y}\,dy = \frac{1}{\lambda}$. Επίσης

$$\begin{aligned}
\mathbb{E}[XY] \;&=\; \int_0^\infty \int_0^\infty xy\lambda y e^{-yx-\lambda y} \;=\; \int_{y=0}^\infty \lambda e^{-\lambda y} \left(\int_{x=0}^\infty y^2 x e^{-yx}\,dx \right) dy \\
&=\; \int_0^\infty \lambda e^{-\lambda y} \left(\int_{u=0}^\infty u e^{-u}\,du \right) dy \qquad \text{(αλλαγή μεταβλητής } u = xy) \\
&=\; \int_0^\infty \lambda e^{-\lambda y}\,dy \;=\; 1.
\end{aligned}$$

Όμως η συνδιακύμανση δεν ορίζεται διότι $\mathsf{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$ και η $\mathbb{E}X$ δεν είναι πεπερασμένη. Παρ΄ ότι η μέση τιμή της X είναι άπειρη η δεσμευμένη μέση τιμή (δεδομένου ότι $Y = y$) είναι πεπερασμένη:

$$\begin{aligned}
\mathbb{E}[X|Y=y] \;&=\; \int_0^\infty x f_{X|Y}(x|y)\,dx \;=\; \int_0^\infty xy\, e^{-yx}\,dx \;=\; \frac{1}{y}\int_0^\infty y^2 x\, e^{-yx}\,dx \\
&=\; \frac{1}{y}\int_0^\infty u\, e^{-u}\,du \;=\; \frac{1}{y}.
\end{aligned}$$

Βεβαίως ισχύει ότι

$$\mathbb{E}X \;=\; \int_0^\infty f_Y(y)\mathbb{E}[X|Y=y]\,dy \;=\; \int_0^\infty \frac{1}{y}\lambda e^{-\lambda y}\,dy.$$

Το γενικευμένο ολοκλήρωμα αυτό δεν συγκλίνει (αποκλίνει στο $+\infty$) διότι μπρούμε να το γράψουμε ως $\lim_{\substack{A\to 0\\ B\to\infty}} \int_A^B \frac{1}{y}\lambda e^{-\lambda y}\,dy$. Παρατηρούμε ότι, όταν $A \to 0$ το όριο δεν είναι πεπερασμένο.

Η δεσμευμένη μέση τιμή $\mathbb{E}[Y|X=x]$ υπολογίζεται ως

$$\begin{aligned}
\mathbb{E}[Y|X=x] \;&=\; \int_0^\infty y f_{Y|X}(y|x)\,dy \;=\; \int_0^\infty (\lambda+x)^2 y^2 e^{-y(\lambda+x)}\,dy \\
&=\; \frac{1}{\lambda+x}\int_0^\infty (\lambda+x)^3 y^2 e^{-y(\lambda+x)}\,dy \;=\; \frac{1}{\lambda+x}\int_0^\infty u^2 e^{-u}\,du \;=\; \frac{\Gamma(3)}{x+\lambda} \;=\; \frac{2}{x+\lambda}.
\end{aligned}$$

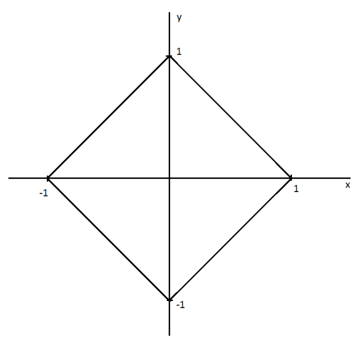
Η μέση τιμή της Y είναι πεπερασμένη και ίση με $\frac{1}{\lambda}$ αφού η περιθώρια πυκνότητα της Y είναι $f_Y(y) = \lambda e^{-\lambda}$, δηλαδή εκθετική με ρυθμό λ. Στο ίδιο αποτέλεσμα καταλήγουμε βεβαίως και αν ολοκληρώσουμε την δεσμευμένη μέση $\mathbb{E}[Y|X = x]$ ως προς την περιθώρια κατανομή της X:

$$\begin{aligned}\mathbb{E}[Y] &= \int_0^\infty \mathbb{E}[Y|X = x] f_X(x)dx = \int_0^\infty \frac{2}{x+\lambda}\frac{\lambda}{(\lambda+x)^2}\,dx = 2\lambda\int_0^\infty \frac{1}{(\lambda+x)^3}\,dx \\ &= 2\lambda\int_\lambda^\infty \frac{dt}{t^3} = 2\lambda\left(-\frac{1}{2t^2}\bigg|_\lambda^\infty\right) = \frac{1}{\lambda}.\end{aligned}$$

Παράδειγμα 4.

Θεωρούμε την από κοινού πυκνότητα πιθανότητας

$$f(x,y) = \begin{cases} \frac{1}{2} & \text{αν } |x|+|y| \le 1, \\ 0 & \text{διαφορετικά.}\end{cases}$$



Σχήμα 4.2: Το χωρίο στο επίπεδο για το οποίο ισχύει $|x| + |y| \le 1$.

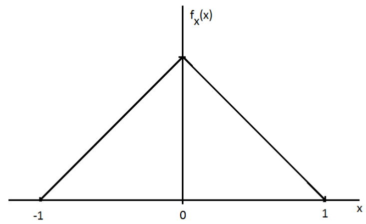
Προκειμένου να υπολογίσουμε την περιθώρια κατανομή της X, παρατηρούμε ότι η από κοινού κατανομή είναι θετική όταν $|y| \le 1 - |x|$ ή ισοδύναμα $-1 + |x| \le y \le 1 - |x|$. Συνεπώς,

$$f_X(x) = \int_{y=-\infty}^{\infty} f(x,y)dy = \int_{y=-1+|x|}^{1-|x|} \frac{1}{2}dy = 1 - |x|, \quad \text{για } |x| \le 1. \tag{4.39}$$

Η περιθώρια κατανομή της X φαίνεται στο σχήμα (4.3).

Η μέση τιμή και η διασπορά της X προκύπτουν ως

$$\mathbb{E}X = \int_{-1}^{1} x(1-|x|)dx = 0$$



Σχήμα 4.3: Η περιθώρια κατανμή $f_X(x) = 1 - |x|$ για $|x| \le 1$.

$$\mathbb{E}X^2 = \int_{-1}^{1} x^2(1-|x|)dx = \int_{-1}^{1} x^2 dx - \int_{-1}^{1} x^2|x|dx = \frac{2}{3} - 2\int_0^1 x^3 dx = \frac{2}{3} - 2\frac{1}{4} = \frac{1}{6}.$$

Προφανώς, $\mathsf{Var}(X) = \frac{1}{6} - 0^2 = \frac{1}{6}$.

Οι υπολογισμοί για την τυχαία μεταβλητή Y είναι ακριβώς ίδιοι λόγω της συμμετρίας ανάμεσα στην X και στην Y.

Προκειμένου να υπολογίσουμε την συνδιακύμανση, υπολογίζουμε την

$$\begin{aligned} \mathbb{E}[XY] &= \int_{x=-1}^{1}\left(\int_{y=-1+|x|}^{1-|x|} \frac{1}{2} xy\,dy\right) dx \\ &= = \int_{x=-1}^{0}\left(\int_{y=-1+|x|}^{1-|x|} \frac{1}{2} xy\,dy\right) dx \; + \; \int_{x=0}^{1}\left(\int_{y=-1+|x|}^{1-|x|} \frac{1}{2} xy\,dy\right) dx. \end{aligned}$$

Το πρώτο ολοκλήρωμα γίνεται

$$\begin{aligned} \int_{x=-1}^{0}\left(\int_{y=-1+|x|}^{1-|x|} \frac{1}{2} xy\,dy\right) dx &= \int_{x=-1}^{0}\left(\int_{y=-1-x}^{1+x} \frac{1}{2} xy\,dy\right) dx \\ &= \frac{1}{2}\int_{x=-1}^{0} x\left(\int_{y=-1-x}^{1+x} y\,dy\right) dx \\ &= \frac{1}{2}\int_{x=-1}^{0} x\left(\frac{(1+x)^2}{2} - \frac{(-1-x)^2}{2}\right) dx \; = \; 0. \end{aligned}$$

Με τον ίδιο τρόπο, το δεύτερο ολοκλρήρωμα είναι

$$\begin{aligned} \int_{0}^{x=1}\left(\int_{y=-1+|x|}^{1-|x|} \frac{1}{2} xy\,dy\right) dx &= \int_{0}^{x=1}\left(\int_{y=-1+x}^{1-x} \frac{1}{2} xy\,dy\right) dx \\ &= \frac{1}{2}\int_{0}^{x=1} x\left(\int_{y=-1+x}^{1-x} y\,dy\right) dx \\ &= \frac{1}{2}\int_{0}^{x=1} x\left(\frac{(1-x)^2}{2} - \frac{(-1+x)^2}{2}\right) dx \; = \; 0. \end{aligned}$$

Συνεπώς, $\mathsf{Cov}(X,Y) = 0 - 0^2 = 0$ και οι τυχαίες μεταβλητές X, Y, είναι ασυσχέτιστες. Δεν είναι όμως ανεξάρτητες.