

Διάλεξη 5: Πιθανότητες & Τυχαία Διανύσματα Σημειώσεις Μαθήματος

Κωνσταντίνος Χατζηλυγερούδης
costashatz@upatras.gr

Δήμητρα Κουμούτσου
dkoumoutsou@upatras.gr

13 Οκτωβρίου, 2025

Contents

1	Πιθανότητα, Δειγματικός Χώρος και Ενδεχόμενα	3
1.1	Δειγματικός Χώρος και Ενδεχόμενα	3
1.2	Μέτρο Πιθανότητας	3
1.3	Διακριτοί και Συνεχείς Δειγματικοί Χώροι	4
1.4	Οπτικοποίηση και Ερμηνεία	4
2	Τυχαίες Μεταβλητές	5
2.1	Ορισμός και Ερμηνεία	5
2.2	Διακριτές και Συνεχείς Τυχαίες Μεταβλητές	5
2.3	Συνήθεις Κατανομές Πιθανότητας	9
2.4	Συμβάσεις Σημειογραφίας	11
3	Διακριτές και Συνεχείς Πιθανότητες	11
3.1	Διακριτές Πιθανότητες	11
3.2	Συνεχείς Πιθανότητες	12
3.3	Διακριτές vs. Συνεχείς Πιθανότητες	13
4	Μέτρα Πιθανότητας και Κατανομές	14
4.1	Μέτρο Πιθανότητας	14
4.2	Συνάρτηση Κατανομής (CDF)	14
4.3	Πυκνότητα και Μάζα Πιθανότητας	14
4.4	Από Κοινού Κατανομές (Joint Distributions)	15
4.5	Περιθωριακές Κατανομές (Marginal Distributions)	15
4.6	Διανυσματικές Τυχαίες Μεταβλητές και Από Κοινού Κατανομές	16
4.7	Περιθωριακές Πυκνότητες και Γεωμετρική Ερμηνεία	16
5	Δεσμευμένες Κατανομές και Θεμελιώδεις Κανόνες Πιθανότητας	17
5.1	Δεσμευμένες Κατανομές	17
5.2	Κανόνες Αθροίσματος και Γινομένου	19

6	Ο Κανόνας του Bayes και η Συμπερασματολογία κατά Bayes	24
6.1	Ο Κανόνας του Bayes: Θεμέλια	24
6.2	Διαισθητική Περίληψη	28
7	Μέση Τιμή, Διασπορά, Συνδιασπορά και Συσχέτιση	28
7.1	Αναμενόμενες Τιμές και Μέση Τιμή	28
7.2	Διασπορά και Τυπική Απόκλιση	29
7.3	Συνδιασπορά	30
7.4	Συσχέτιση	30
7.5	Πολυδιάστατη Συνδιασπορά	31
7.6	Δειγματικοί Μέσοι και Συνδιασπορές	33
7.7	Εσωτερικά Γινόμενα Τυχαίων Μεταβλητών	34
8	Ανεξαρτησία και Υπό Συνθήκη Ανεξαρτησία	35
8.1	Στατιστική Ανεξαρτησία	35
8.2	Υπό Συνθήκη Ανεξαρτησία	36
8.3	Επίδραση της Ανεξαρτησίας στη Συνδιασπορά και στη Διασπορά	36
9	Κανονικές (Gaussian) Κατανομές	37
9.1	Μονοδιάστατη Κανονική Κατανομή	37
9.2	Πολυδιάστατη Κανονική Κατανομή	38
9.3	Από Κοινού, Περιθώριες και Υπό Συνθήκη Κατανομές	38
9.4	Γινόμενο Γκαουσιανών Πυκνοτήτων	42
9.5	Γραμμικοί Μετασχηματισμοί Γκαουσιανών	43
9.6	Δειγματοληψία από Γκαουσιανές Κατανομές	43

1 Πιθανότητα, Δειγματικός Χώρος και Ενδεχόμενα

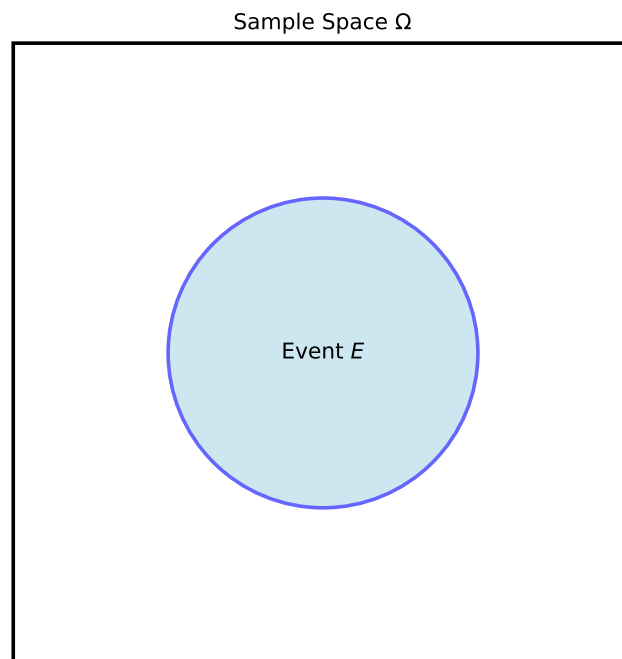
1.1 Δειγματικός Χώρος και Ενδεχόμενα

Στη θεωρία πιθανοτήτων, κάθε τυχαίο πείραμα συνοδεύεται από έναν **δειγματικό χώρο**, που συμβολίζεται με Ω . Ο δειγματικός χώρος είναι το σύνολο όλων των δυνατών εκβάσεων του πειράματος.

- Παράδειγμα: Ρίψη κέρματος $\Omega = \{K, \Gamma\}$.
- Παράδειγμα: Ρίψη ζαριού $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Ένα **ενδεχόμενο** E είναι οποιοδήποτε υποσύνολο του Ω . Αποτελεί μια συλλογή εκβάσεων που ικανοποιούν μια συγκεκριμένη συνθήκη.

- Παράδειγμα: $E = \{\text{ζυγό αποτέλεσμα}\} = \{2, 4, 6\}$.



Σχήμα 1.1: Δειγματικός χώρος Ω και ενδεχόμενο $E \subseteq \Omega$.

1.2 Μέτρο Πιθανότητας

Ένα **μέτρο πιθανότητας** $\mathbb{P}$ αντιστοιχίζει έναν αριθμό μεταξύ 0 και 1 σε κάθε ενδεχόμενο E , εκφράζοντας το πόσο πιθανό είναι να συμβεί.

Τυπικά,

$$\mathbb{P} : \mathcal{F} \subseteq 2^\Omega \rightarrow [0, 1],$$

όπου $\mathcal{F}$ είναι μια συλλογή υποσυνόλων του Ω (μια σ -άλγεβρα) στα οποία μπορεί να οριστεί πιθανότητα.

Το μέτρο ικανοποιεί:

$$\mathbb{P}(\Omega) = 1, \quad \mathbb{P}\left(\bigcup_i E_i\right) = \sum_i \mathbb{P}(E_i) \quad \text{για ανεξάρτητα ενδεχόμενα } E_i.$$

Η πρώτη ιδιότητα εξασφαλίζει ότι το σύνολο των πιθανών εκβάσεων έχει συνολική πιθανότητα 1, ενώ η δεύτερη ότι η πιθανότητα ανεξάρτητων ενδεχομένων είναι αθροιστική.

1.3 Διακριτοί και Συνεχείς Δειγματικοί Χώροι

Ο δειγματικός χώρος μπορεί να είναι **διακριτός** ή **συνεχής**:

Διακριτοί δειγματικοί χώροι. Οι εκβάσεις είναι αριθμήσιμες (πεπερασμένες ή αριθμήσιμα άπειρες). Σε κάθε έκβαση $\omega_i \in \Omega$ αντιστοιχίζεται μια πιθανότητα $\mathbb{P}(\{\omega_i\}) = p_i$, και για κάθε ενδεχόμενο $E \subseteq \Omega$ ισχύει

$$\mathbb{P}(E) = \sum_{\omega_i \in E} p_i, \quad \sum_i p_i = 1.$$

Τυπικά παραδείγματα: κέρματα, ζάρια, κατηγορικές μεταβλητές.

Συνεχείς δειγματικοί χώροι. Οι εκβάσεις είναι μη αριθμήσιμες (π.χ. πραγματικοί αριθμοί). Δεν μπορούμε να αποδώσουμε πιθανότητα σε μεμονωμένη τιμή. Οι πιθανότητες ορίζονται σε διαστήματα μέσω **συνάρτησης πυκνότητας πιθανότητας** (PDF) $p_x(z)$:

$$\mathbb{P}(a \leq x \leq b) = \int_a^b p_x(z) dz, \quad \int_{-\infty}^{\infty} p_x(z) dz = 1.$$

Παραδείγματα: θερμοκρασία, χρόνος, πλάτος θορύβου.

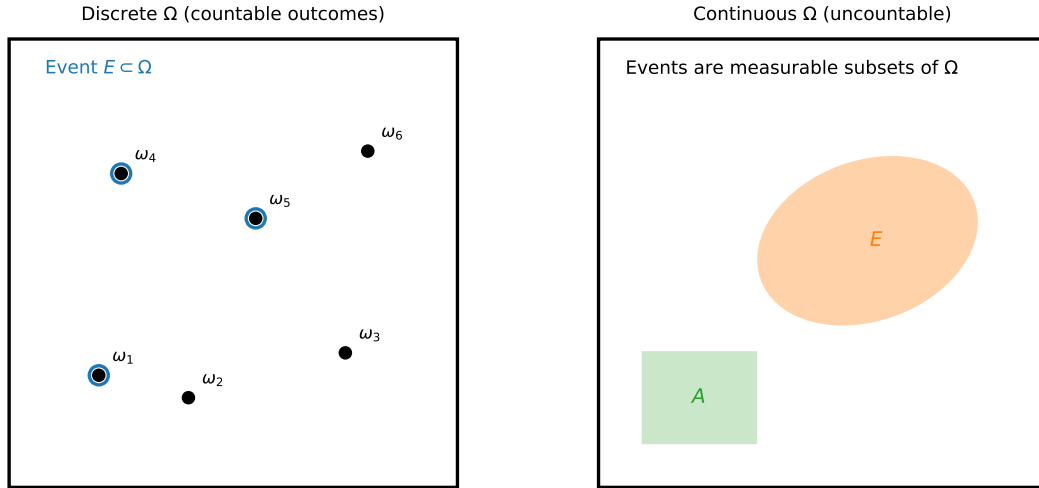
1.4 Οπτικοποίηση και Ερμηνεία

Είναι συχνά χρήσιμο να οπτικοποιούμε τον δειγματικό χώρο Ω ως μια περιοχή που περιέχει όλες τις δυνατές εκβάσεις (Σχ. 1.1). Τα ενδεχόμενα είναι υποσύνολα αυτής της περιοχής, και η πιθανότητα $\mathbb{P}(E)$ αντιστοιχεί στο “μέγεθος” (ή μέτρο) του υποσυνόλου E σε σχέση με το Ω .

Στην **διακριτή** περίπτωση, αυτό αντιστοιχεί στο άθροισμα των πιθανοτήτων των αντίστοιχων εκβάσεων. Στη **συνεχή** περίπτωση, αντιστοιχεί στο ολοκλήρωμα της PDF πάνω στην περιοχή ενδιαφέροντος (Σχ. 1.2).

Σύνοψη:

- Ο **δειγματικός χώρος** Ω περιέχει όλες τις δυνατές εκβάσεις.
- Ένα **ενδεχόμενο** E είναι υποσύνολο του Ω .
- Ένα **μέτρο πιθανότητας** $\mathbb{P}$ δίνει τιμές στο $[0, 1]$, με άθροιση των ανεξάρτητων ενδεχομένων και $\mathbb{P}(\Omega) = 1$.
- Σε **διακριτούς** χώρους, οι πιθανότητες αθροίζονται. Σε **συνεχείς** χώρους, ολοκληρώνονται μέσω της συνάρτησης πυκνότητας πιθανότητας.



Σχήμα 1.2: Διακριτοί vs. συνεχείς δειγματικοί χώροι. Στους διακριτούς αθροίζουμε πιθανότητες· στους συνεχείς ολοκληρώνουμε χρησιμοποιώντας PDF.

2 Τυχαίες Μεταβλητές

2.1 Ορισμός και Ερμηνεία

Μια **τυχαία μεταβλητή** είναι μια μετρήσιμη απεικόνιση που αντιστοιχίζει εκβάσεις ενός τυχαίου πειράματος σε αριθμητικές τιμές:

$$x : \Omega \longrightarrow \mathbb{R}.$$

Πιο γενικά, ένα **τυχαίο διάνυσμα** ορίζεται ως

$$\mathbf{x} : \Omega \longrightarrow \mathbb{R}^M,$$

όπου M είναι η διάσταση του διανύσματος.

Μια τυχαία μεταβλητή δημιουργεί έναν **νόμο πιθανότητας** στο σύνολο τιμών της (τον *φορέα*) $\mathcal{X} \subseteq \mathbb{R}$ μέσω

$$\mathbb{P}(x \in \mathcal{X}) = \mathbb{P}(\{\omega \in \Omega : x(\omega) \in \mathcal{X}\}).$$

Έτσι, το μέτρο πιθανότητας στον δειγματικό χώρο Ω μεταφέρεται σε ένα μέτρο πάνω σε αριθμητικές τιμές στο $\mathbb{R}$.

2.2 Διακριτές και Συνεχείς Τυχαίες Μεταβλητές

Ανάλογα με το αν το στήριγμα $\mathcal{X}$ είναι αριθμήσιμο ή μη-αριθμήσιμο, οι τυχαίες μεταβλητές διακρίνονται σε **διακριτές** και **συνεχείς**.

Διακριτές τυχαίες μεταβλητές. Μια διακριτή τυχαία μεταβλητή x παίρνει τιμές σε ένα αριθμήσιμο σύνολο $\mathcal{X}$ και χαρακτηρίζεται από τη **συνάρτηση μάζας πιθανότητας** (Propability Mass Function - PMF)

$$p_x(z) = \mathbb{P}(x = z), \quad z \in \mathcal{X},$$

με

$$p_x(z) \geq 0, \quad \sum_{z \in \mathcal{X}} p_x(z) = 1.$$

Η πιθανότητα του ενδεχομένου $a \leq x \leq b$ υπολογίζεται αθροίζοντας τις αντίστοιχες μάζες.

Συνεχείς τυχαίες μεταβλητές. Μια συνεχής τυχαία μεταβλητή έχει μη-αριθμήσιμο στήριγμα $\mathcal{X} \subseteq \mathbb{R}$ και περιγράφεται από μια **συνάρτηση πυκνότητας πιθανότητας** (Probability Density Function - PDF)

$$p_x(z) \geq 0, \quad \int_{\mathcal{X}} p_x(z) dz = 1.$$

Οι πιθανότητες ενδεχομένων δίνονται από ολοκληρώματα:

$$\mathbb{P}(a \leq x \leq b) = \int_a^b p_x(z) dz.$$

Αθροιστική Συνάρτηση Κατανομής. Κάθε τυχαία μεταβλητή (διακριτή ή συνεχής) έχει μια **αθροιστική συνάρτηση κατανομής** (Cumulative Distribution Function - CDF)

$$F_x(z) = \mathbb{P}(x \leq z),$$

η οποία δίνει την πιθανότητα η τιμή της τυχαίας μεταβλητής να μην υπερβαίνει το z .

Παραδείγματα. Η ρίψη ζαριού αποτελεί παράδειγμα **διακριτής** τυχαίας μεταβλητής, ενώ ο θόρυβος μέτρησης που μοντελοποιείται ως Γκαουσιανός είναι **συνεχής**.

Παράδειγμα 1: Διακριτή Τυχαία Μεταβλητή (Ρίψη κέρματος)
Δειγματικός χώρος και πιθανότητες:

$$\Omega = \{H, T\}, \quad \mathbb{P}(H) = 0.7, \quad \mathbb{P}(T) = 0.3.$$

Τυχαία μεταβλητή:

$$x(\omega) = \begin{cases} 1, & \omega = H, \\ 0, & \omega = T. \end{cases}$$

Νόμος πιθανότητας:

$$\mathbb{P}(x = 1) = \mathbb{P}(H) = 0.7, \quad \mathbb{P}(x = 0) = \mathbb{P}(T) = 0.3.$$

Άρα η **συνάρτηση μάζας πιθανότητας** (PMF) του x είναι

$$p_x(z) = \begin{cases} 0.7, & z = 1, \\ 0.3, & z = 0, \\ 0, & \text{διαφορετικά.} \end{cases}$$

Το παράδειγμα δείχνει ότι η x απλώς απεικονίζει εκβάσεις σε αριθμητικές τιμές, ενώ οι πιθανότητες προέρχονται από το υποκείμενο μέτρο $\mathbb{P}$.

Παράδειγμα 2: Συνεχής Τυχαία Μεταβλητή (Μονοτονικός Μετασχηματισμός)

Δειγματικός χώρος και μέτρο: Έστω $\Omega = [0, 1]$ με ομοιόμορφο μέτρο πιθανότητας

$$\mathbb{P}(\omega \in [a, b]) = b - a.$$

Ορίζουμε $u(\omega) = \omega$. Τότε $u \sim \text{Uniform}(0, 1)$ με PDF

$$p_u(z_u) = \begin{cases} 1, & 0 \leq z_u \leq 1, \\ 0, & \text{διαφορετικά.} \end{cases}$$

Τυχαία μεταβλητή: Ορίζουμε

$$x(\omega) = \omega^2 = u(\omega)^2.$$

Έτσι η x απεικονίζει το $[0, 1]$ στο $[0, 1]$ μονοτονικά.

Αθροιστική συνάρτηση κατανομής (CDF) της x :

$$F_x(z) = \mathbb{P}(x \leq z) = \mathbb{P}(u^2 \leq z) = \mathbb{P}(u \leq \sqrt{z}) = \sqrt{z}, \quad 0 \leq z \leq 1.$$

Συνάρτηση πυκνότητας πιθανότητας (PDF) της x :

$$p_x(z) = \frac{d}{dz} F_x(z) = \frac{1}{2\sqrt{z}}, \quad 0 < z < 1.$$

Έλεγχος κανονικοποίησης:

$$\int_0^1 p_x(z) dz = \int_0^1 \frac{1}{2\sqrt{z}} dz = [\sqrt{z}]_0^1 = 1.$$

Άρα η PDF της x είναι

$$p_x(z) = \begin{cases} \frac{1}{2\sqrt{z}}, & 0 < z < 1, \\ 0, & \text{διαφορετικά.} \end{cases}$$

Παράδειγμα 3: Συνεχής Τυχαία Μεταβλητή (Μη Μονοτονικός Μετασχηματισμός)

Δειγματικός χώρος και μέτρο: Έστω ξανά $\Omega = [0, 1]$ με ομοιόμορφο μέτρο και

$$u(\omega) = \omega \sim \text{Uniform}(0, 1).$$

Τυχαία μεταβλητή: Ορίζουμε

$$x(\omega) = g(u(\omega)) = \sin(2\pi u).$$

Η συνάρτηση $g(u)$ είναι *μη μονοτονική*, καθώς η $\sin(2\pi u)$ αυξάνεται από 0 σε 1 στο $[0, \frac{1}{4}]$, μειώνεται σε -1 στο $[\frac{1}{4}, \frac{3}{4}]$, και ξανανεβαίνει σε 0 στο $[\frac{3}{4}, 1]$.

Φορέας της x : Αφού $\sin(2\pi u) \in [-1, 1]$, η τυχαία μεταβλητή x λαμβάνει τιμές στο

$$z \in [-1, 1].$$

Γενικός κανόνας μετασχηματισμού: Κατά τον μετασχηματισμό μιας συνεχούς μεταβλητής πρέπει να διατηρείται η πιθανότητα:

$$p_u(u) du = p_x(z) dz.$$

Εφόσον $z = g(u)$, μια μικρή μεταβολή du προκαλεί

$$dz = g'(u) du.$$

Λαμβάνοντας απόλυτες τιμές:

$$p_x(z) |dz| = p_u(u) |du| \implies p_x(z) = \frac{p_u(u)}{|g'(u)|}.$$

Αυτό ισχύει τοπικά όταν ο μετασχηματισμός είναι μονοτονικός.

Όταν όμως το $g(u)$ είναι μη μονοτονικό, η ίδια τιμή z μπορεί να προκύψει από πολλαπλές τιμές u . Τότε πρέπει να αθροίσουμε τις συνεισφορές:

$$p_x(z) = \sum_{u_i: z=g(u_i)} \frac{p_u(u_i)}{|g'(u_i)|}.$$

Αυτός είναι ο γενικός μονοδιάστατος κανόνας μεταβολής μεταβλητής για PDF.

Εφαρμογή στο παρόν παράδειγμα: Με $g(u) = \sin(2\pi u)$ και $p_u(u) = 1$, βρίσκουμε τα u_i από

$$z = \sin(2\pi u) \implies 2\pi u = \arcsin(z) \quad \text{ή} \quad 2\pi u = \pi - \arcsin(z).$$

Έτσι,

$$u_1 = \frac{\arcsin(z)}{2\pi}, \quad u_2 = \frac{1}{2} - \frac{\arcsin(z)}{2\pi}.$$

Η παράγωγος:

$$g'(u) = 2\pi \cos(2\pi u) \implies |g'(u_i)| = 2\pi \sqrt{1 - z^2}.$$

Άρα:

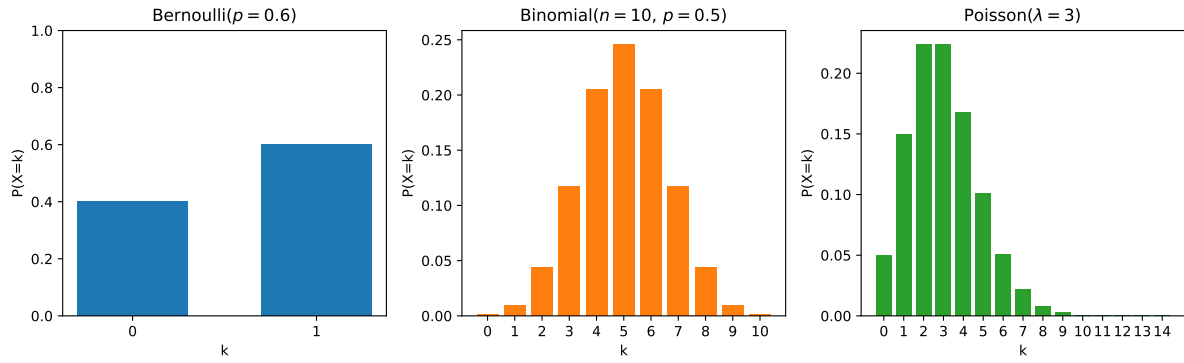
$$p_x(z) = \frac{1}{2\pi \sqrt{1 - z^2}} + \frac{1}{2\pi \sqrt{1 - z^2}} = \frac{1}{\pi \sqrt{1 - z^2}}, \quad -1 < z < 1.$$

Ερμηνεία - Κανονικοποίηση: Αυτή είναι η γνωστή *arcsine κατανομή* στο $[-1, 1]$, με

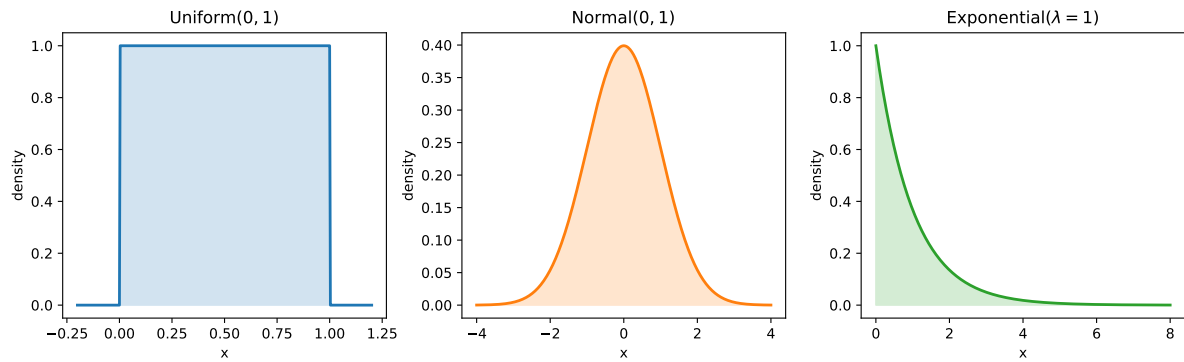
$$\int_{-1}^1 \frac{dz}{\pi \sqrt{1 - z^2}} = 1.$$

Η πυκνότητα είναι μεγαλύτερη κοντά στα άκρα $z = \pm 1$ (όπου το $|g'(u)|$ είναι μικρό, άρα «συμπίεζονται» πιθανότητες) και μικρότερη κοντά στο $z = 0$ (όπου το $|g'(u)|$ είναι μεγάλο). Έτσι ο μη μονοτονικός μετασχηματισμός μετατρέπει την ομοιόμορφη πυκνότητα σε μια έντονα ανομοιόμορφη.

Τελικό σχόλιο: Οι τιμές μιας τυχαίας μεταβλητής δεν καθορίζουν άμεσα την PDF. Η κατανομή προκύπτει από το πώς παράγονται αυτές οι τιμές μέσω του αρχικού πιθανοτικού μοντέλου.



Σχήμα 2.1: Κοινές διακριτές κατανομές πιθανότητας.



Σχήμα 2.2: Κοινές συνεχείς κατανομές πιθανότητας.

2.3 Συνήθεις Κατανομές Πιθανότητας

Πολλά φαινόμενα περιγράφονται με συγκεκριμένες παραμετρικές οικογένειες κατανομών.

Διακριτές κατανομές. Συνηθισμένα παραδείγματα είναι οι κατανομές Bernoulli, Binomial και Poisson (Σχ. 2.1).

Bernoulli(p), $0 < p < 1$. Φορέας: $z \in \{0, 1\}$.

$$p_x(z) = p^z(1-p)^{1-z}, \quad z \in \{0, 1\}.$$

Αθροιστική συνάρτηση κατανομής (CDF):

$$F_x(z) = \mathbb{P}(x \leq z) = \begin{cases} 0, & z < 0, \\ 1-p, & 0 \leq z < 1, \\ 1, & z \geq 1. \end{cases}$$

Binomial(n, p), $n \in \mathbb{N}$, $0 < p < 1$. Φορέας: $z = 0, 1, \dots, n$.

$$p_x(z) = \binom{n}{z} p^z(1-p)^{n-z}, \quad z = 0, \dots, n.$$

Αθροιστική συνάρτηση κατανομής:

$$F_x(z) = \sum_{u=0}^{\lfloor z \rfloor} \binom{n}{u} p^u (1-p)^{n-u}.$$

Συμπαγώς,

$$F_x(z) = I_{1-p}(n - \lfloor z \rfloor, \lfloor z \rfloor + 1),$$

όπου $I_{a,b}(z)$ είναι η κανονικοποιημένη ατελής βήτα.

Poisson(λ), $\lambda > 0$. Φορέας: $z \in \{0, 1, 2, \dots\}$.

$$p_x(z) = e^{-\lambda} \frac{\lambda^z}{z!}, \quad z = 0, 1, 2, \dots$$

Αθροιστική συνάρτηση κατανομής:

$$F_x(z) = e^{-\lambda} \sum_{u=0}^{\lfloor z \rfloor} \frac{\lambda^u}{u!} = P(\lfloor z \rfloor + 1, \lambda),$$

όπου $P(s, u) = \gamma(s, u)/\Gamma(s)$ είναι η κανονικοποιημένη κάτω ατελής γάμμα.

Συνεχείς κατανομές. Κλασικά παραδείγματα είναι η Ομοιόμορφη (Uniform), Εκθετική (Exponential) και Γκαουσιανή (Gaussian) (Σχ. 2.2).

Uniform(a, b), $a < b$. Φορέας: $z \in [a, b]$.

$$p_x(z) = \frac{1}{b-a} \mathbf{1}_{[a,b]}(z),$$

όπου $\mathbf{1}_{[a,b]}(z)$ είναι η ενδεικτική συνάρτηση του διαστήματος $[a, b]$ (1 στο διάστημα μεταξύ a και b , και 0 οπουδήποτε αλλού).

Αθροιστική συνάρτηση κατανομής:

$$F_x(z) = \begin{cases} 0, & z < a, \\ \frac{z-a}{b-a}, & a \leq z \leq b, \\ 1, & z > b. \end{cases}$$

Exponential(λ), $\lambda > 0$. Φορέας: $z \in [0, \infty)$.

$$p_x(z) = \lambda e^{-\lambda z} \mathbf{1}_{[0,\infty)}(z).$$

Αθροιστική συνάρτηση κατανομής:

$$F_x(z) = \int_0^z \lambda e^{-\lambda u} du = 1 - e^{-\lambda z}, \quad z \geq 0,$$

και $F_x(z) = 0$ για $z < 0$.

Gaussian/Normal (μ, σ^2) , $\sigma > 0$. Φορέας: $z \in \mathbb{R}$.

$$p_x(z) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(z-\mu)^2}{2\sigma^2}\right).$$

Αθροιστική συνάρτηση κατανομής:

$$F_x(z) = \mathbb{P}(x \leq z) = \Phi\left(\frac{z-\mu}{\sigma}\right),$$

όπου

$$\Phi(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^u e^{-t^2/2} dt = \frac{1}{2} \left(1 + \operatorname{erf}\left(\frac{u}{\sqrt{2}}\right)\right).$$

Κατανομή	PMF/PDF	CDF $F_x(z)$
Bernoulli(p)	$p^z(1-p)^{1-z}$, $z \in \{0, 1\}$	0 ($z < 0$); $1-p$ ($0 \leq z < 1$); 1 ($z \geq 1$)
Binomial(n, p)	$\binom{n}{z} p^z (1-p)^{n-z}$	$\sum_{u=0}^{\lfloor z \rfloor} \binom{n}{u} p^u (1-p)^{n-u}$
Poisson(λ)	$e^{-\lambda} \frac{\lambda^z}{z!}$	$e^{-\lambda} \sum_{u=0}^{\lfloor z \rfloor} \frac{\lambda^u}{u!}$
Uniform(a, b)	$\frac{1}{b-a} \mathbf{1}_{[a,b]}(z)$	$\frac{z-a}{b-a}$ στο $[a, b]$
Exponential(λ)	$\lambda e^{-\lambda z} \mathbf{1}_{[0,\infty)}(z)$	$1 - e^{-\lambda z}$ ($z \geq 0$)
Normal(μ, σ^2)	$\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(z-\mu)^2}{2\sigma^2}}$	$\Phi\left(\frac{z-\mu}{\sigma}\right)$

Table 1: PMFs/PDFs και CDFs για κοινές κατανομές.

2.4 Συμβάσεις Σημειογραφίας

Για λόγους σαφήνειας και συνέπειας, οι συμβάσεις σημειογραφίας του Πίνακα 2 χρησιμοποιούνται σε όλο το κείμενο. Συνοπτικά: τα πεζά γράμματα δηλώνουν βαθμωτά μεγέθη, τα έντονα πεζά διανύσματα, και τα έντονα κεφαλαία πίνακες. Τα tildes ($\tilde{x}$, $\tilde{\mathbf{x}}$) μπορούν να χρησιμοποιηθούν για να επισημανθεί η τυχαιότητα όταν χρειάζεται, ενώ τα καπέλα ($\hat{x}$, $\hat{\mathbf{x}}$) δηλώνουν εκτιμήσεις.

3 Διακριτές και Συνεχείς Πιθανότητες

3.1 Διακριτές Πιθανότητες

Σε ένα διακριτό δειγματικό χώρο $\Omega = \{\omega_1, \dots, \omega_n\}$, κάθε στοιχειώδες ενδεχόμενο ω_i έχει μια πιθανότητα

$$p_i = \mathbb{P}(\{\omega_i\}), \quad p_i \geq 0, \quad \sum_{i=1}^n p_i = 1.$$

Το σύνολο $\{p_i\}$ ορίζει μια συνάρτηση μάζας πιθανότητας (PMF) πάνω στο Ω .

Τύπος Μεγέθους	Ντετερμινιστικό	Τυχαίο	Σχόλια
Scalar (βαθμωτός)	x	x ή $\tilde{x}$	$x \in \mathbb{R}$: το $\tilde{x}$ χρησιμοποιείται όταν θέλουμε να τονίσουμε ρητά την τυχαιότητα
Vector (διάνυσμα)	$\mathbf{x}$	$\mathbf{x}$ ή $\tilde{\mathbf{x}}$	$\mathbf{x} \in \mathbb{R}^n$: τυχαίο διάνυσμα αν έχει οριστεί πυκνότητα $p(\mathbf{x})$
Matrix (πίνακας)	$\mathbf{A}$	—	$\mathbf{A} \in \mathbb{R}^{m \times n}$: συνήθως θεωρείται ντετερμινιστικός
Estimate (εκτίμηση)	$\hat{x}, \hat{\mathbf{x}}$	—	Χρησιμοποιούνται για εκτιμημένες ή υπολογισμένες τιμές

Table 2: Σημειογραφία για ντετερμινιστικά και τυχαία μεγέθη.

Για οποιοδήποτε **ενδεχόμενο** $E \subseteq \Omega$, η πιθανότητά του είναι

$$\mathbb{P}(E) = \sum_{\omega_i \in E} p_i.$$

Μια **τυχαία μεταβλητή** $x : \Omega \rightarrow \mathbb{R}$ συνεπάγεται πιθανότητες στις τιμές της. Έστω $\mathcal{X} \subseteq \mathbb{R}$ το σύνολο των πιθανών τιμών της. Η PMF δίνεται από:

$$p_x(z) = \mathbb{P}(x = z) = \sum_{\omega_i: x(\omega_i)=z} p_i, \quad z \in \mathcal{X}, \quad \sum_{z \in \mathcal{X}} p_x(z) = 1.$$

Παράδειγμα. Για ένα ζάρι:

$$\Omega = \{1, 2, 3, 4, 5, 6\}, \quad \mathcal{X} = \{1, 2, 3, 4, 5, 6\}, \quad p_i = \frac{1}{6}.$$

Π.χ. η πιθανότητα άρτιου αποτελέσματος:

$$\mathbb{P}(\text{άρτιος}) = \mathbb{P}(\{2, 4, 6\}) = \frac{1}{2}.$$

3.2 Συνεχείς Πιθανότητες

Σε έναν **συνεχή δειγματικό χώρο**, η τυχαία μεταβλητή $x : \Omega \rightarrow \mathbb{R}$ λαμβάνει τιμές σε μη αριθμήσιμο σύνολο $\mathcal{X} \subseteq \mathbb{R}$. Σε αυτή την περίπτωση, μεμονωμένα αποτελέσματα έχουν μηδενική πιθανότητα και οι πιθανότητες περιγράφονται μέσω μιας μη αρνητικής συνάρτησης $p_x(z)$, της **συνάρτησης πυκνότητας πιθανότητας (PDF)**:

$$p_x(z) \geq 0, \quad \int_{\mathcal{X}} p_x(z) dz = 1.$$

Η πιθανότητα ότι x ανήκει σε μια μετρήσιμη περιοχή $A \subseteq \mathcal{X}$ είναι

$$\mathbb{P}(x \in A) = \int_A p_x(z) dz.$$

Η αντίστοιχη **αθροιστική συνάρτηση κατανομής** (CDF) ορίζεται ως

$$F_x(z) = \mathbb{P}(x \leq z) = \int_{-\infty}^z p_x(u) du.$$

Παράδειγμα. Για τυχαία μεταβλητή με ομοιόμορφη τυχειότητα στο $[0, 1]$:

$$\mathcal{X} = [0, 1], \quad p_x(z) = \begin{cases} 1, & 0 \leq z \leq 1, \\ 0, & \text{αλλού,} \end{cases} \quad F_x(z) = \begin{cases} 0, & z < 0, \\ z, & 0 \leq z \leq 1, \\ 1, & z > 1. \end{cases}$$

3.3 Διακριτές vs. Συνεχείς Πιθανότητες

Η βασική διάκριση μεταξύ διακριτών και συνεχών τυχαίων μεταβλητών αφορά τον τρόπο με τον οποίο αποδίδεται η πιθανότητα:

Διακριτή περίπτωση. Οι πιθανότητες αποδίδονται σε συγκεκριμένα σημεία:

$$p_x(z) = \mathbb{P}(x = z), \quad \sum_{z \in \mathcal{X}} p_x(z) = 1.$$

Για οποιοδήποτε υποσύνολο $A \subseteq \mathcal{X}$,

$$\mathbb{P}(x \in A) = \sum_{z \in A} p_x(z).$$

Παράδειγμα: για ένα ζάρι,

$$\mathcal{X} = \{1, 2, 3, 4, 5, 6\}, \quad \mathbb{P}(\text{άρτιος}) = \frac{1}{2}.$$

Συνεχής περίπτωση. Οι πιθανότητες αποδίδονται σε διαστήματα ή περιοχές, όχι σε μεμονωμένα σημεία:

$$p_x(z) \geq 0, \quad \int_{\mathcal{X}} p_x(z) dz = 1, \quad \mathbb{P}(x \in A) = \int_A p_x(z) dz.$$

Παράδειγμα: αν $x \sim \text{Uniform}(0, 1)$,

$$\mathbb{P}(0.2 \leq x \leq 0.5) = 0.3.$$

Κύρια διαφορά.

- **Διακριτές μεταβλητές:** Η πιθανότητα αποδίδεται σε συγκεκριμένες τιμές $z \in \mathcal{X}$.
- **Συνεχείς μεταβλητές:** Η πιθανότητα οποιασδήποτε ακριβούς τιμής είναι μηδέν, $\mathbb{P}(x = z) = 0$: μόνο διαστήματα έχουν μη μηδενική πιθανότητα.

4 Μέτρα Πιθανότητας και Κατανομές

4.1 Μέτρο Πιθανότητας

Το πιο θεμελιώδες αντικείμενο στη θεωρία πιθανοτήτων είναι το **μέτρο πιθανότητας**, συμβολιζόμενο με $\mathbb{P}$. Ορίζεται πάνω σε έναν δειγματικό χώρο Ω και αντιστοιχίζει σε κάθε γεγονός (μετρήσιμο υποσύνολο του Ω) έναν μη αρνητικό αριθμό στο διάστημα $[0, 1]$:

$$\mathbb{P} : \mathcal{F} \subseteq 2^\Omega \rightarrow [0, 1], \quad \mathbb{P}(A) = \text{“η πιθανότητα το γεγονός } A \text{ να συμβεί”}.$$

Το μέτρο $\mathbb{P}$ ικανοποιεί:

$$\mathbb{P}(\Omega) = 1, \quad \mathbb{P}\left(\bigcup_i A_i\right) = \sum_i \mathbb{P}(A_i) \quad \text{για ανεξάρτητα (ασυμβίβαστα) γεγονότα } A_i.$$

Αυτή η διατύπωση μέσω μέτρου αποτελεί τη βάση πάνω στην οποία ορίζονται οι τυχαίες μεταβλητές, οι συναρτήσεις κατανομής και οι πυκνότητες.

4.2 Συνάρτηση Κατανομής (CDF)

Μια τυχαία μεταβλητή $x : \Omega \rightarrow \mathbb{R}$ “δημιουργεί” έναν νόμο πιθανότητας στο $\mathbb{R}$ μέσω της **αθροιστικής συνάρτησης κατανομής (CDF)**:

$$F_x(t) = \mathbb{P}(x \leq t).$$

Η $F_x(t)$ είναι μη φθίνουσα, δεξιά συνεχής και ικανοποιεί:

$$\lim_{t \rightarrow -\infty} F_x(t) = 0, \quad \lim_{t \rightarrow +\infty} F_x(t) = 1.$$

Διαισθητικά, η $F_x(t)$ δίνει την πιθανότητα η τυχαία μεταβλητή x να πάρει τιμή μικρότερη ή ίση από t .

4.3 Πυκνότητα και Μάζα Πιθανότητας

Η **πυκνότητα πιθανότητας (PDF)** ή η **συνάρτηση μάζας πιθανότητας (PMF)** περιγράφει το πώς κατανέμεται η πιθανότητα στις δυνατές τιμές της x .

Διακριτή περίπτωση. Αν η x παίρνει τιμές σε αριθμήσιμο σύνολο $\mathcal{X}$,

$$p_x(z) = \mathbb{P}(x = z), \quad \sum_{z \in \mathcal{X}} p_x(z) = 1.$$

Συνεχής περίπτωση. Όταν η x είναι συνεχής, οι πιθανότητες εκφράζονται μέσω πυκνότητας:

$$p_x(z) = \frac{d}{dz} F_x(z), \quad \int_{-\infty}^{\infty} p_x(z) dz = 1.$$

Η πιθανότητα να βρεθεί η x σε διάστημα είναι

$$\mathbb{P}(a \leq x \leq b) = \int_a^b p_x(z) dz.$$

Η CDF ανακτάται από την PDF μέσω ολοκλήρωσης:

$$F_x(z) = \int_{-\infty}^z p_x(u) du.$$

4.4 Από Κοινού Κατανομές (Joint Distributions)

Όταν εξετάζουμε δύο ή περισσότερες τυχαίες μεταβλητές, η κοινή τους συμπεριφορά περιγράφεται από μια **από κοινού κατανομή** (joint distribution). Για δύο τυχαίες μεταβλητές x, y , η **από κοινού συνάρτηση πυκνότητας πιθανότητας (PDF)** ή **συνάρτηση μάζας πιθανότητας (PMF)** (joint PDF or joint PMF) συμβολίζεται με $p_{x,y}(z_1, z_2)$.

Συνεχής περίπτωση.

$$p_{x,y}(z_1, z_2) \geq 0, \quad \int_{\mathcal{X}} \int_{\mathcal{Y}} p_{x,y}(z_1, z_2) dz_1 dz_2 = 1.$$

Για οποιεσδήποτε μετρήσιμες περιοχές $A \subseteq \mathcal{X}$ και $B \subseteq \mathcal{Y}$,

$$\mathbb{P}(x \in A, y \in B) = \int_A \int_B p_{x,y}(z_1, z_2) dz_1 dz_2.$$

Παράδειγμα. Έστω ότι x και y είναι από κοινού ομοιόμορφες στο μοναδιαίο τετράγωνο:

$$p_{x,y}(z_1, z_2) = \begin{cases} 1, & 0 \leq z_1 \leq 1, 0 \leq z_2 \leq 1, \\ 0, & \text{αλλιώς.} \end{cases}$$

Έλεγχος κανονικοποίησης:

$$\int_0^1 \int_0^1 p_{x,y}(z_1, z_2) dz_1 dz_2 = \int_0^1 \int_0^1 1 dz_1 dz_2 = 1.$$

Διακριτή περίπτωση.

$$p_{x,y}(z_1, z_2) = \mathbb{P}(x = z_1, y = z_2), \quad \sum_{z_1, z_2} p_{x,y}(z_1, z_2) = 1.$$

Παράδειγμα. Για δύο ρίψεις νομίσματος:

$$\Omega = \{HH, HT, TH, TT\}, \quad p_{x,y}(H, H) = \frac{1}{4}, \dots$$

4.5 Περιθωριακές Κατανομές (Marginal Distributions)

Η **περιθωριακή κατανομή** (marginal distribution) μιας μεταβλητής προκύπτει με ολοκλήρωση (ή άθροιση) της από κοινού κατανομής των τιμών όλων των άλλων μεταβλητών.

Συνεχής περίπτωση.

$$p_x(z_1) = \int_{\mathcal{Y}} p_{x,y}(z_1, z_2) dz_2, \quad p_y(z_2) = \int_{\mathcal{X}} p_{x,y}(z_1, z_2) dz_1.$$

Διακριτή περίπτωση. Οι ολοκληρώσεις αντικαθίστανται από αθροίσματα:

$$p_x(z_1) = \sum_{z_2} p_{x,y}(z_1, z_2), \quad p_y(z_2) = \sum_{z_1} p_{x,y}(z_1, z_2).$$

Παράδειγμα. Για το πείραμα δύο ρίψεων νομίσματος:

$$p_x(H) = p_{x,y}(H, H) + p_{x,y}(H, T) = \frac{1}{2}.$$

4.6 Διανυσματικές Τυχαίες Μεταβλητές και Από Κοινού Κατανομές

Μια διανυσματική τυχαία μεταβλητή είναι μια τυχαία μεταβλητή με τιμές στο $\mathbb{R}^M$:

$$\mathbf{x} = [x_1, \dots, x_M]^\top : \Omega \rightarrow \mathbb{R}^M.$$

Η κατανομή της περιγράφεται από μια από κοινού CDF και, όπου υπάρχει, από μια από κοινού PDF.

Από κοινού αθροιστική συνάρτηση κατανομής (CDF).

$$F_{\mathbf{x}}(\mathbf{z}) = \mathbb{P}(x_1 \leq z_1, \dots, x_M \leq z_M), \quad \mathbf{z} \in \mathbb{R}^M.$$

Η συνάρτηση αυτή είναι μη φθίνουσα ως προς κάθε συνιστώσα και ικανοποιεί $F_{\mathbf{x}}(\infty, \dots, \infty) = 1$.

Από κοινού συνάρτηση πυκνότητας πιθανότητας.

$$p_{\mathbf{x}}(\mathbf{z}) = \frac{\partial^M F_{\mathbf{x}}(\mathbf{z})}{\partial z_1 \dots \partial z_M}, \quad \int_{\mathbb{R}^M} p_{\mathbf{x}}(\mathbf{z}) d\mathbf{z} = 1.$$

Στη διακριτή περίπτωση, $p_{\mathbf{x}}(\mathbf{z}) = \mathbb{P}(\mathbf{x} = \mathbf{z})$, ενώ στη συνεχή περίπτωση είναι μια πυκνότητα στο $\mathbb{R}^M$.

4.7 Περιθωριακές Πυκνότητες και Γεωμετρική Ερμηνεία

Για μια διανυσματική τυχαία μεταβλητή $\mathbf{x} = [x_1, \dots, x_M]^\top$, η **περιθωριακή πυκνότητα** (marginal density) μιας συνιστώσας x_i προκύπτει με ολοκλήρωση ως προς όλες τις υπόλοιπες μεταβλητές:

$$p_{x_i}(z_i) = \int p_{\mathbf{x}}(\mathbf{z}) d\mathbf{z}_{\setminus i},$$

όπου $\mathbf{z}_{\setminus i}$ είναι όλα τα στοιχεία του $\mathbf{z}$ εκτός από το z_i .

Γεωμετρική ερμηνεία. Στη συνεχή περίπτωση, η πιθανότητα

$$\mathbb{P}(a_1 \leq x_1 \leq b_1, \dots, a_M \leq x_M \leq b_M)$$

ισούται με τον όγκο κάτω από την από κοινού συνάρτηση πυκνότητας πιθανότητας $p_{\mathbf{x}}(\mathbf{z})$ στην περιοχή που ορίζεται από τα διαστήματα $[a_i, b_i]$.

5 Δεσμευμένες Κατανομές και Θεμελιώδεις Κανόνες Πιθανότητας

5.1 Δεσμευμένες Κατανομές

Μια **δεσμευμένη κατανομή** περιγράφει τον νόμο πιθανότητας μιας τυχαίας μεταβλητής, υπό την προϋπόθεση ότι μια άλλη έχει λάβει συγκεκριμένη τιμή. Για δύο τυχαίες μεταβλητές x και y με από κοινού κατανομή $p_{x,y}(z_1, z_2)$, η δεσμευμένη πυκνότητα του y υπό την προϋπόθεση $x = z_1$ ορίζεται ως

$$p_{y|x}(z_2 | z_1) = \frac{p_{x,y}(z_1, z_2)}{p_x(z_1)}.$$

Ισχύει η κανονικοποίηση:

$$p_{y|x}(z_2 | z_1) \geq 0, \quad \int_{\mathcal{Y}} p_{y|x}(z_2 | z_1) dz_2 = 1,$$

και στη διακριτή περίπτωση το ολοκλήρωμα αντικαθίσταται από άθροισμα.

Η αντίστοιχη **δεσμευμένη πιθανότητα** δίνεται από

$$\mathbb{P}(y \in B | x = z_1) = \int_B p_{y|x}(z_2 | z_1) dz_2.$$

Διαισθητικά, η δέσμευση στο $x = z_1$ περιορίζει την αβεβαιότητα για το y στις τιμές που είναι συμβατές με την παρατήρηση του x .

Η από κοινού πιθανότητα δύο μεταβλητών γράφεται ως

$$\mathbb{P}(y \in B, x \in A) = \int_A \int_B p_{y|x}(z_2 | z_1) p_x(z_1) dz_2 dz_1.$$

Παράδειγμα (Διακριτή περίπτωση). Έστω ότι x και y είναι τα αποτελέσματα δύο ανεξάρτητων ρίψεων ενός νομίσματος:

$$p_{x,y}(H, H) = p_{x,y}(H, T) = p_{x,y}(T, H) = p_{x,y}(T, T) = \frac{1}{4}.$$

Τότε η δεσμευμένη κατανομή του y υπό την προϋπόθεση $x = H$ είναι

$$p_{y|x}(y | H) = \frac{p_{x,y}(H, y)}{p_x(H)} = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2}.$$

Άρα το y παραμένει ομοιόμορφο ακόμη και γνωρίζοντας ότι $x = H$, λόγω ανεξαρτησίας. Σε περίπτωση εξάρτησης, το $p_{y|x}$ θα άλλαζε αναλόγως.

Δεσμευμένες Κατανομές για Πολλαπλές Μεταβλητές

Μια **δεσμευμένη κατανομή** επεκτείνεται φυσικά στην περίπτωση τριών ή περισσότερων τυχαίων μεταβλητών. Για τυχαίες μεταβλητές $x_1, x_2, \dots, x_n$ με από κοινού κατανομή

$$p_{x_1, x_2, \dots, x_n}(z_1, z_2, \dots, z_n),$$

η δεσμευμένη πυκνότητα μιας μεταβλητής x_i , δεδομένων όλων των υπόλοιπων, δίνεται από

$$p_{x_i|x_1,\dots,x_{i-1},x_{i+1},\dots,x_n}(z_i | z_1, \dots, z_{i-1}, z_{i+1}, \dots, z_n) = \frac{p_{x_1,\dots,x_n}(z_1, \dots, z_n)}{p_{x_1,\dots,x_{i-1},x_{i+1},\dots,x_n}(z_1, \dots, z_{i-1}, z_{i+1}, \dots, z_n)}.$$

Ισχύει η κανονικοποίηση:

$$p_{x_i|x_{\setminus i}}(z_i | z_{\setminus i}) \geq 0, \quad \int_{\mathcal{X}_i} p_{x_i|x_{\setminus i}}(z_i | z_{\setminus i}) dz_i = 1,$$

όπου $x_{\setminus i} = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$. Στη διακριτή περίπτωση, το ολοκλήρωμα αντικαθίσταται από άθροισμα.

Η αντίστοιχη δεσμευμένη πιθανότητα γράφεται:

$$\mathbb{P}(x_i \in B | x_{\setminus i} = z_{\setminus i}) = \int_B p_{x_i|x_{\setminus i}}(z_i | z_{\setminus i}) dz_i.$$

Η από κοινού κατανομή μπορεί πάντοτε να παραγοντοποιηθεί με τον **κανόνα αλυσίδας της πιθανότητας**:

$$p(x_1, x_2, \dots, x_n) = p(x_1) p(x_2 | x_1) p(x_3 | x_1, x_2) \cdots p(x_n | x_1, \dots, x_{n-1}).$$

Αυτή η παραγοντοποίηση δείχνει πώς οι εξαρτήσεις μεταξύ των μεταβλητών αποτυπώνονται μέσω δεσμευμένων κατανομών.

Γενικότερα, για οποιαδήποτε ανεξάρτητα μεταξύ τους σύνολα δεικτών $A, B, C \subseteq \{1, \dots, n\}$, η δεσμευμένη πυκνότητα γράφεται ως

$$p(x_A | x_B, x_C) = \frac{p(x_A, x_B | x_C)}{p(x_B | x_C)} = \frac{p(x_A, x_B, x_C)}{p(x_B, x_C)}.$$

Έτσι, η δέσμευση σε πολλές μεταβλητές απλώς περιορίζει την αβεβαιότητα σε έναν χώρο υψηλότερης διάστασης.

Τέλος, η από κοινού πιθανότητα υποσυνόλων μεταβλητών δίνεται από

$$\mathbb{P}(x_A \in A', x_B \in B') = \int_{A'} \int_{B'} p_{x_A|x_B}(z_A | z_B) p_{x_B}(z_B) dz_A dz_B.$$

Δεσμεύσεις σε Διανυσματική Μορφή

Μπορούμε πάντα να θεωρήσουμε τις μεταβλητές $(x_1, x_2, \dots, x_n)$ ως ένα διανυσματικό τυχαίο μέγεθος $\mathbf{x} \in \mathbb{R}^n$. Τότε:

$$p_{x_i|x_{\setminus i}}(z_i | \mathbf{z}_{\setminus i}) = \frac{p_{\mathbf{x}}(\mathbf{z})}{p_{\mathbf{x}_{\setminus i}}(\mathbf{z}_{\setminus i})}.$$

Για συντομία, μπορούμε να γράφουμε ισοδύναμα:

$$p(x_i | \mathbf{x}_{\setminus i}) = \frac{p(\mathbf{x})}{p(\mathbf{x}_{\setminus i})}.$$

Η μορφή αυτή τονίζει ότι οι δεσμευμένες κατανομές είναι απλώς λόγοι κατάλληλων περιθωριακών και από κοινού πυκνοτήτων.

5.2 Κανόνες Αθροίσματος και Γινομένου

Κανόνας Αθροίσματος (Ολοκλήρωση Περιθωριακής Κατανομής)

Ο **κανόνας αθροίσματος** (ή *marginalization*) μάς επιτρέπει να βρούμε την περιθωριακή κατανομή μίας μεταβλητής, αθροίζοντας ή ολοκληρώνοντας την από κοινού κατανομή ως προς την άλλη:

$$p_x(z_1) = \begin{cases} \sum_{z_2 \in \mathcal{Y}} p_{x,y}(z_1, z_2), & \text{διακριτή περίπτωση,} \\ \int_{\mathcal{Y}} p_{x,y}(z_1, z_2) dz_2, & \text{συνεχής περίπτωση.} \end{cases}$$

Αυτό εκφράζει την ιδέα ότι για να βρούμε την περιθωριακή πυκνότητα ως προς x , “ξεχνάμε” ή “αθροίζουμε” τη μεταβλητή y .

Παράδειγμα (Διακριτό). Έστω ότι x και y είναι τα αποτελέσματα δύο ζαριών. Η από κοινού κατανομή $p_{x,y}(i, j)$ αντιστοιχεί σε πιθανότητα $1/36$ για καθεμία από τις 36 ισοπίθανες εκβάσεις. Η περιθωριακή πυκνότητα ως προς x είναι:

$$p_x(i) = \sum_{j=1}^6 p_{x,y}(i, j) = 6 \times \frac{1}{36} = \frac{1}{6}.$$

Άρα το κάθε ζάρι παραμένει ομοιόμορφα κατανεμημένο, ακόμη και όταν τα θεωρούμε από κοινού.

Παράδειγμα (Συνεχές). Αν $x, y \sim \text{Uniform}(0, 1)$ ανεξάρτητα,

$$p_{x,y}(z_1, z_2) = \begin{cases} 1, & 0 \leq z_1, z_2 \leq 1, \\ 0, & \text{αλλιώς,} \end{cases}$$

τότε

$$p_x(z_1) = \int_0^1 p_{x,y}(z_1, z_2) dz_2 = 1,$$

οπότε x είναι περιθωριακά ομοιόμορφη στο $[0, 1]$.

Κανόνας Γινομένου (Παραγοντοποίηση)

Ο **κανόνας γινομένου** γράφει την από κοινού κατανομή ως γινόμενο μιας δεσμευμένης και μίας περιθωριακής πυκνότητας:

$$p_{x,y}(z_1, z_2) = p_{y|x}(z_2 | z_1) p_x(z_1),$$

ή ισοδύναμα,

$$p_{x,y}(z_1, z_2) = p_{x|y}(z_1 | z_2) p_y(z_2).$$

Αυτή η παραγοντοποίηση δείχνει ότι η από κοινού κατανομή μπορεί να κατασκευαστεί από το “τι γνωρίζουμε” (p_x) και το “τι παραμένει αβέβαιο” δεδομένης αυτής της γνώσης ($p_{y|x}$).

Ερμηνεία.

- **Κανόνας αθροίσματος:** Περιθωριοποιείς (αθροίζεις ή ολοκληρώνεις) ως προς τις μεταβλητές που δεν σε ενδιαφέρουν: ουσιαστικά τις «ξεχνάς».
- **Κανόνας γινομένου:** Παραγοντοποιείς την από κοινού κατανομή: την «χτίζεις» από κομμάτια που ήδη γνωρίζεις.

Παράδειγμα: Δύο Ζάρια

Έστω δύο ζάρια με εκβάσεις $x, y \in \{1, 2, 3, 4, 5, 6\}$. Το πρώτο ζάρι είναι δίκαιο (όλες οι πλευρές έχουν ίδια πιθανότητα), ενώ το δεύτερο έχει άνισες πιθανότητες ανά όψη (βλ. Fig. 5.1).

Ζάρι με άνισες πιθανότητες. Οι ακατέργαστες (μη κανονικοποιημένες) «βαρύτητες» των έξι όψεων είναι:

$$\text{loaded_weights} = [0.5, 0.6, 1.5, 1.2, 2.0, 0.8].$$

Κανονικοποιώντας τις, προκύπτει μία έγκυρη PMF:

$$p_y(z) = \frac{\text{loaded_weights}[z]}{\sum_{k=1}^6 \text{loaded_weights}[k]},$$

όπου

$$\sum_{k=1}^6 \text{loaded_weights}[k] = 0.5 + 0.6 + 1.5 + 1.2 + 2.0 + 0.8 = 6.6.$$

Άρα,

$$p_y = \left[\frac{0.5}{6.6}, \frac{0.6}{6.6}, \frac{1.5}{6.6}, \frac{1.2}{6.6}, \frac{2.0}{6.6}, \frac{0.8}{6.6} \right] = [0.0758, 0.0909, 0.2273, 0.1818, 0.3030, 0.1212].$$

Κλασικό Ζάρι. Το πρώτο ζάρι έχει

$$p_x(z) = \frac{1}{6}, \quad z = 1, \dots, 6.$$

Από Κοινού Κατανομή. Υποθέτοντας ότι τα δύο ζάρια είναι ανεξάρτητα, η από κοινού πιθανότητα παραγοντοποιείται ως

$$p_{x,y}(i, j) = p_x(i) p_y(j) = \frac{1}{6} p_y(j).$$

Άρα ο πίνακας της κοινής κατανομής (όλες οι τιμές i, j) έχει σειρές ίδιες μεταξύ τους και αναλογικές προς τις τιμές $p_y(j)$.

Έλεγχος Κανονικοποίησης.

$$\sum_{i=1}^6 \sum_{j=1}^6 p_{x,y}(i, j) = \sum_{i=1}^6 \frac{1}{6} \sum_{j=1}^6 p_y(j) = \frac{1}{6} \times 6 \times 1 = 1.$$

Περιθωριακές πυκνότητες (Marginals). Η περιθωριακή πυκνότητα του x είναι ομοιόμορφη:

$$p_x(i) = \sum_{j=1}^6 p_{x,y}(i, j) = \frac{1}{6} \sum_{j=1}^6 p_y(j) = \frac{1}{6}.$$

Η περιθωριακή πυκνότητα του y είναι εξ ορισμού $p_y(j)$.

Υπό Συνθήκη Κατανομές. Από τον κανόνα του γινομένου:

$$p_{y|x}(j | i) = \frac{p_{x,y}(i, j)}{p_x(i)} = \frac{\frac{1}{6}p_y(j)}{\frac{1}{6}} = p_y(j).$$

Όπως αναμένεται για ανεξάρτητες μεταβλητές, το $p_{y|x}(j | i)$ δεν εξαρτάται από το i .

Πίνακας Περιθωριακών Πυκνοτήτων για το ζάρι με άνισες πιθανότητες.

j	1	2	3	4	5	6
$p_y(j)$	0.0758	0.0909	0.2273	0.1818	0.3030	0.1212

Ερμηνεία.

- Το πρώτο ζάρι παραμένει ομοιόμορφο, $p_x(i) = 1/6$.
- Το δεύτερο ζάρι είναι μεροληπτικό προς τις μεγαλύτερες τιμές (το 5 είναι το πιθανότερο, το 1 το λιγότερο πιθανό).
- Η από κοινού κατανομή $p_{x,y}(i, j) = p_x(i)p_y(j)$ σχηματίζει έναν πίνακα όπου όλες οι σειρές είναι ίδιες (ανεξαρτησία), αλλά οι στήλες διαφέρουν.
- Η υπό συνθήκη κατανομή $p_{y|x}(j | i) = p_y(j)$ επιβεβαιώνει ότι η γνώση του x δεν δίνει καμία πληροφορία για το y .

Σύγκριση με την περίπτωση δύο κλασικών ζαριών. Αν και τα δύο ζάρια είναι κλασικά, τότε $p_y(j) = 1/6$ και

$$p_{x,y}(i, j) = \frac{1}{36},$$

οπότε αναχτούμε την ομοιόμορφη από κοινού κατανομή όπου όλα τα αποτελέσματα είναι ισοπίθانا. Αντίθετα, το «φορτωμένο» ζάρι μετατοπίζει τις κοινές πιθανότητες προς τα αποτελέσματα $(i, 5)$ και $(i, 3)$.

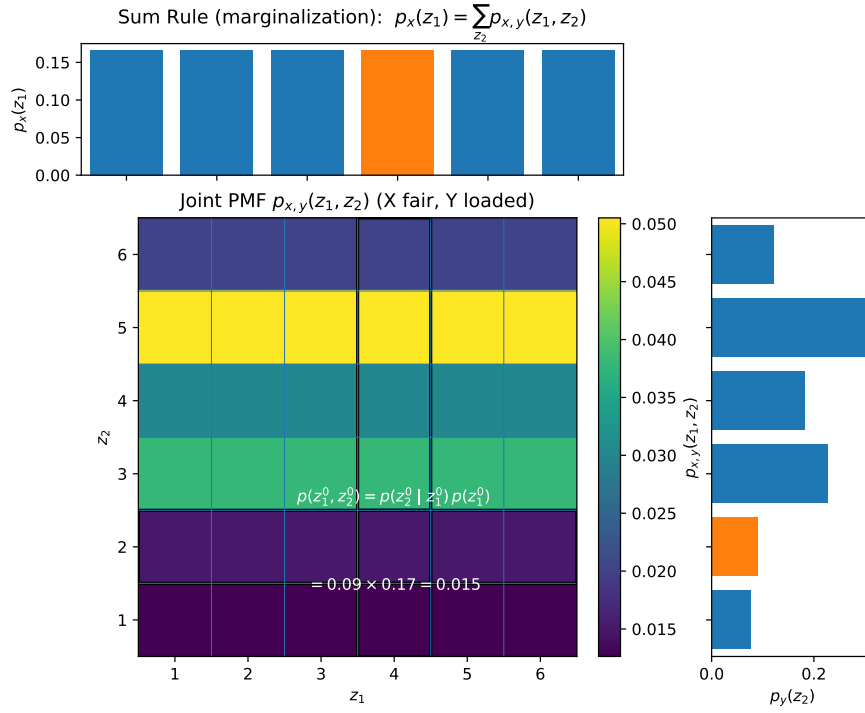
Παράδειγμα: Γκαουσιανή Κατανομή

Έστω ότι (x, y) είναι από κοινού γκαουσιανές τυχαίες μεταβλητές με μηδενική μέση τιμή, μοναδιαίες διασπορές και συσχέτιση $\rho \in (-1, 1)$:

$$p_{x,y}(z_1, z_2) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{z_1^2 - 2\rho z_1 z_2 + z_2^2}{2(1-\rho^2)}\right).$$

Ισοδύναμα, $(x, y) \sim \mathcal{N}(\mathbf{0}, \Sigma)$ με

$$\Sigma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}, \quad \Sigma^{-1} = \frac{1}{1-\rho^2} \begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix}.$$



Σχήμα 5.1: Οπτικοποίηση του κανόνα αθροίσματος και του κανόνα γινομένου για δύο ζάρια. Κάθε κελί αντιστοιχεί σε ένα αποτέλεσμα (z_1, z_2) .

Στόχος. Θέλουμε να δείξουμε ότι οι περιθωριακές πυκνότητες είναι τυπικά κανονικές και να υπολογίσουμε την υπό συνθήκη κατανομή $p_{y|x}(z_2 | z_1)$, επαληθεύοντας στη συνέχεια την παραγοντοποίηση του κανόνα γινομένου

$$p_{x,y}(z_1, z_2) = p_{y|x}(z_2 | z_1) p_x(z_1).$$

1. Περιθωριακή πυκνότητα του x μέσω ολοκλήρωσης ως προς z_2

Ξεκινάμε από την από κοινού κατανομή:

$$p_{x,y}(z_1, z_2) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{z_1^2 - 2\rho z_1 z_2 + z_2^2}{2(1-\rho^2)}\right).$$

Θεωρούμε τον εκθέτη ως τετραγωνική μορφή ως προς το z_2 και συμπληρώνουμε το τετράγωνο:

$$z_1^2 - 2\rho z_1 z_2 + z_2^2 = (z_2 - \rho z_1)^2 + (1 - \rho^2) z_1^2.$$

Άρα

$$p_{x,y}(z_1, z_2) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{(z_2 - \rho z_1)^2}{2(1-\rho^2)}\right) \exp\left(-\frac{z_1^2}{2}\right).$$

Ολοκληρώνουμε ως προς $z_2 \in \mathbb{R}$:

$$p_x(z_1) = \int_{-\infty}^{\infty} p_{x,y}(z_1, z_2) dz_2 = \left[\frac{1}{\sqrt{2\pi}} e^{-\frac{z_1^2}{2}} \right] \underbrace{\left[\frac{1}{\sqrt{2\pi(1-\rho^2)}} \int_{-\infty}^{\infty} e^{-\frac{(z_2 - \rho z_1)^2}{2(1-\rho^2)}} dz_2 \right]}_{=1}.$$

Συνεπώς,

$$p_x(z_1) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z_1^2}{2}\right) = \mathcal{N}(0, 1).$$

Λόγω συμμετρίας το ίδιο ισχύει και για το $p_y(z_2)$.

2. Υπό συνθήκη κατανομή του y δεδομένου $x = z_1$

Από τον ορισμό,

$$p_{y|x}(z_2 | z_1) = \frac{p_{x,y}(z_1, z_2)}{p_x(z_1)}.$$

Χρησιμοποιώντας τις δύο εκφράσεις παραπάνω:

$$p_{y|x}(z_2 | z_1) = \frac{\frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{(z_2-\rho z_1)^2}{2(1-\rho^2)}\right) \exp\left(-\frac{z_1^2}{2}\right)}{\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z_1^2}{2}\right)} = \frac{1}{\sqrt{2\pi(1-\rho^2)}} \exp\left(-\frac{(z_2-\rho z_1)^2}{2(1-\rho^2)}\right).$$

Άρα,

$$p_{y|x}(z_2 | z_1) = \mathcal{N}(\rho z_1, 1 - \rho^2).$$

Αμέσως προκύπτει ότι

$$\mathbb{E}[y | x = z_1] = \rho z_1, \quad \text{Var}[y | x] = 1 - \rho^2.$$

3. Επαλήθευση του κανόνα γινομένου

Πολλαπλασιάζουμε την υπό συνθήκη κατανομή με το περιθώριο:

$$p_{y|x}(z_2 | z_1) p_x(z_1) = \left[\frac{1}{\sqrt{2\pi(1-\rho^2)}} e^{-\frac{(z_2-\rho z_1)^2}{2(1-\rho^2)}} \right] \left[\frac{1}{\sqrt{2\pi}} e^{-\frac{z_1^2}{2}} \right] = \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{(z_2-\rho z_1)^2 + (1-\rho^2)z_1^2}{2(1-\rho^2)}}.$$

Χρησιμοποιώντας την ταυτότητα

$$(z_2 - \rho z_1)^2 + (1 - \rho^2)z_1^2 = z_1^2 - 2\rho z_1 z_2 + z_2^2,$$

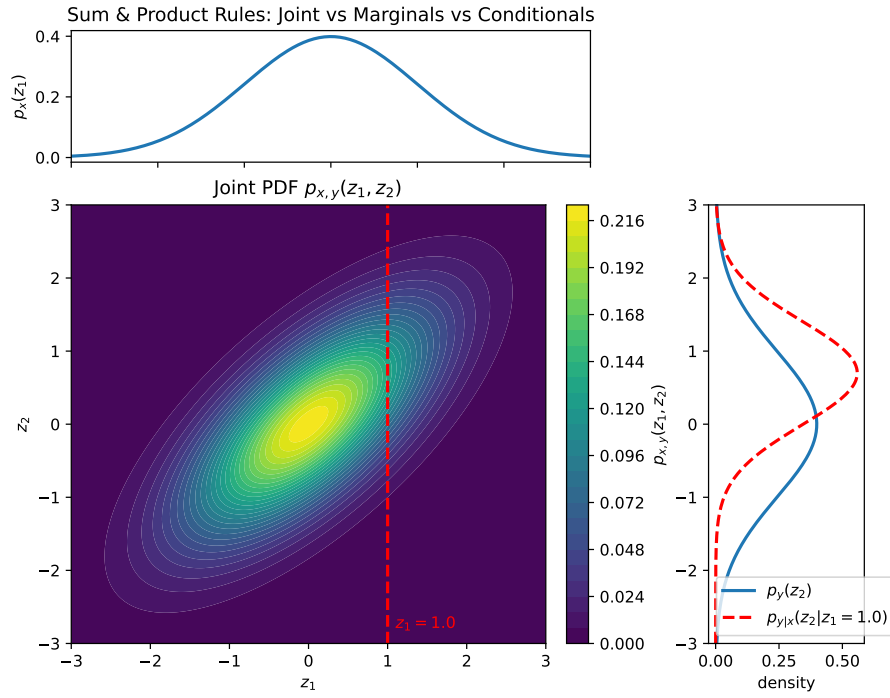
ανακτούμε

$$p_{y|x}(z_2 | z_1) p_x(z_1) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{z_1^2 - 2\rho z_1 z_2 + z_2^2}{2(1-\rho^2)}\right) = p_{x,y}(z_1, z_2).$$

Συμπεράσματα.

$$p_x(z_1) = \mathcal{N}(0, 1), \quad p_{y|x}(z_2 | z_1) = \mathcal{N}(\rho z_1, 1 - \rho^2), \quad p_{x,y}(z_1, z_2) = p_{y|x}(z_2 | z_1) p_x(z_1).$$

Όταν $\rho = 0$, τα x και y είναι ανεξάρτητα, $p_{y|x} = p_y$, και η από κοινού κατανομή παραγοντοποιείται στο γινόμενο δύο τυπικών κανονικών (Γκαουσιανών) κατανομών.



Σχήμα 5.2: Συνεχές παράδειγμα: διδιάστατη Γκαουσιανή κατανομή που απεικονίζει τους κανόνες αθροίσματος και γινομένου. Η υπό συνθήκη κατανομή μειώνει την αβεβαιότητα της μεταβλητής y .

6 Ο Κανόνας του Bayes και η Συμπερασματολογία κατά Bayes

6.1 Ο Κανόνας του Bayes: Θεμέλια

Ο κανόνας του Bayes παρέχει το μαθηματικό πλαίσιο για τον υπολογισμό των “πεποιθήσεών” μας σχετικά με μια άγνωστη ποσότητα μετά την παρατήρηση νέων δεδομένων. Προκύπτει άμεσα από τον κανόνα γινομένου:

$$p_{x,y}(z_1, z_2) = p_{y|x}(z_2 | z_1) p_x(z_1) = p_{x|y}(z_1 | z_2) p_y(z_2).$$

Απλή αναδιάταξη οδηγεί στη γνωστή μορφή του **κανόνα του Bayes**.

Διακριτή περίπτωση.

$$p_{x|y}(z_1 | z_2) = \frac{\underbrace{p_{y|x}(z_2 | z_1)}_{\text{Likelihood}} \underbrace{p_x(z_1)}_{\text{Prior}}}{\underbrace{p_y(z_2)}_{\text{Evidence}}}, \quad p_y(z_2) = \sum_{z_1 \in \mathcal{X}} p_{y|x}(z_2 | z_1) p_x(z_1).$$

Συνεχής περίπτωση.

$$p_{x|y}(z_1 | z_2) = \frac{p_{y|x}(z_2 | z_1) p_x(z_1)}{p_y(z_2)}, \quad p_y(z_2) = \int_{\mathcal{X}} p_{y|x}(z_2 | z_1) p_x(z_1) dz_1.$$

Ερμηνεία.

- **Posterior** (Εκ των υστέρων κατανομή/πιθανότητα/πυκνότητα) $p_{x|y}$: η ενημερωμένη πεποίθηση για το x αφού παρατηρήσουμε το y .
- **Prior** (πρότερη κατανομή/πιθανότητα/πυκνότητα) p_x : η πεποίθησή μας για το x πριν δούμε δεδομένα.
- **Likelihood** (πιθανοφάνεια) $p_{y|x}$: η πιθανότητα των δεδομένων, δεδομένου του x .
- **Evidence** (περιθωριακή πιθανοφάνεια) p_y : η σταθερά κανονικοποίησης που εξασφαλίζει ότι η εκ των υστέρων πιθανότητα “αθροίζει” στην μονάδα.

Συνοπτικά:

$$\text{Posterior} = \frac{\text{Likelihood} \times \text{Prior}}{\text{Evidence}}.$$

Ο κανόνας του Bayes είναι άμεση συνέπεια του **κανόνα γινομένου** και του **κανόνα αθροίσματος**.

Παράδειγμα (Διακριτή περίπτωση): Ιατρικό Τεστ

Εστω ότι εξετάζουμε για μια σπάνια ασθένεια που επηρεάζει το 1% του πληθυσμού:

$$p_x(\text{disease}) = 0.01, \quad p_x(\text{healthy}) = 0.99.$$

Το τεστ έχει:

$$p_{y|x}(\text{positive} \mid \text{disease}) = 0.99, \quad p_{y|x}(\text{positive} \mid \text{healthy}) = 0.05.$$

Θέλουμε να υπολογίσουμε την πιθανότητα να έχει κάποιος την ασθένεια δεδομένου ότι βγήκε θετικός:

$$p_{x|y}(\text{disease} \mid \text{positive}).$$

Βήμα 1: Evidence.

$$\begin{aligned} p_y(\text{positive}) &= p_{y|x}(\text{positive} \mid \text{disease}) p_x(\text{disease}) + p_{y|x}(\text{positive} \mid \text{healthy}) p_x(\text{healthy}) \\ &= 0.99(0.01) + 0.05(0.99) = 0.0594. \end{aligned}$$

Βήμα 2: Posterior.

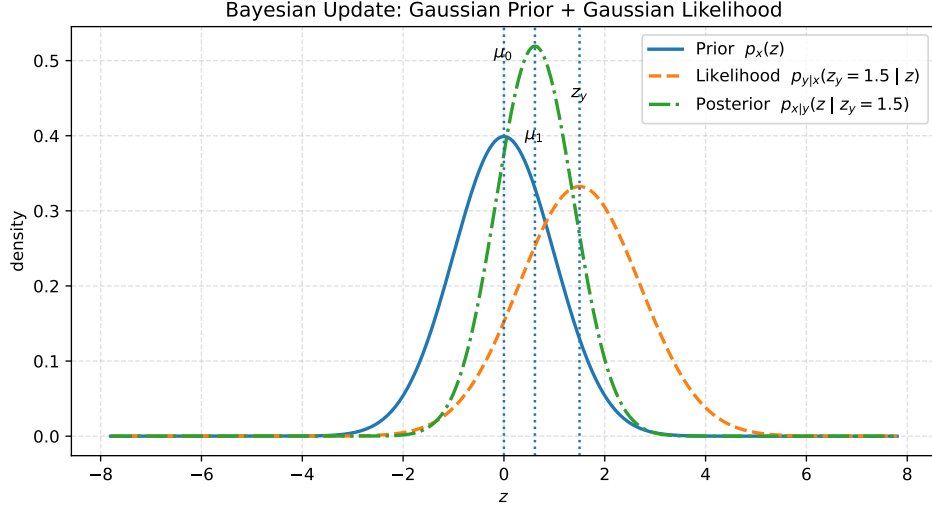
$$p_{x|y}(\text{disease} \mid \text{positive}) = \frac{0.99(0.01)}{0.0594} \approx 0.167.$$

Ερμηνεία: Παρά το πολύ ακριβές τεστ, η πιθανότητα να έχει πράγματι την ασθένεια κάποιος που βρέθηκε θετικός είναι μόλις περίπου 16.7%, λόγω του πολύ χαμηλού *prior* (base rate).

Αυτό δείχνει τη δύναμη του κανόνα του Bayes στη συλλογιστική υπό αβεβαιότητα.

Παράδειγμα (Συνεχή περίπτωση): Γκαουσιανές για prior και πιθανοφάνεια

Ένα κλασικό παράδειγμα συνεχούς συμπερασματολογίας κατά Bayes αφορά συζυγείς (conjugate) Γκαουσιανές κατανομές (Fig. 6.1). Το παράδειγμα δείχνει πώς ενημερώνουμε την πεποίθησή μας για μια άγνωστη συνεχόμενη ποσότητα μετά από μια θορυβώδη μέτρηση.



Σχήμα 6.1: Γκαουσιανό prior, likelihood και posterior. Το posterior βρίσκεται “ενδιάμεσα” και έχει μικρότερη διασπορά.

Πλαίσιο. Θέλουμε να εκτιμήσουμε μια άγνωστη πραγματική ποσότητα (π.χ. πραγματική θερμοκρασία δωματίου), την οποία συμβολίζουμε με θ . Πριν πάρουμε κάποια μέτρηση, έχουμε μια *prior* πεποίθηση, η οποία μοντελοποιείται ως:

$$\theta \sim \mathcal{N}(\mu_{\text{prior}}, \sigma_{\text{prior}}^2),$$

όπου μ_{prior} είναι η αρχική μας εκτίμηση και σ_{prior}^2 η αβεβαιότητά μας. Παίρνουμε μια θορυβώδη μέτρηση:

$$y_{\text{obs}} | \theta \sim \mathcal{N}(\theta, \sigma_{\text{noise}}^2),$$

όπου σ_{noise}^2 είναι η διακύμανση του αισθητήρα. Στόχος: να βρούμε την posterior $p(\theta | y_{\text{obs}})$, δηλαδή την πεποίθησή μας μετά τη μέτρηση.

Βήμα 1: Μορφή του posterior. Από τον κανόνα του Bayes:

$$p(\theta | y_{\text{obs}}) \propto p(y_{\text{obs}} | \theta) p(\theta).$$

Υποκαθιστώντας τις Γκαουσιανές μορφές:

$$p(\theta | y_{\text{obs}}) \propto \exp\left[-\frac{(y_{\text{obs}} - \theta)^2}{2\sigma_{\text{noise}}^2}\right] \exp\left[-\frac{(\theta - \mu_{\text{prior}})^2}{2\sigma_{\text{prior}}^2}\right].$$

Συνδυάζοντας τους εκθέτες:

$$\ln p(\theta | y_{\text{obs}}) = -\frac{1}{2} \left[\frac{(y_{\text{obs}} - \theta)^2}{\sigma_{\text{noise}}^2} + \frac{(\theta - \mu_{\text{prior}})^2}{\sigma_{\text{prior}}^2} \right] + \text{const.}$$

Βήμα 2: Ανάπτυξη και ομαδοποίηση των όρων ως προς θ . Αναπτύσσουμε τα τετράγωνα:

$$\frac{(y_{\text{obs}} - \theta)^2}{\sigma_{\text{noise}}^2} = \frac{\theta^2 - 2\theta y_{\text{obs}} + y_{\text{obs}}^2}{\sigma_{\text{noise}}^2}, \quad \frac{(\theta - \mu_{\text{prior}})^2}{\sigma_{\text{prior}}^2} = \frac{\theta^2 - 2\theta \mu_{\text{prior}} + \mu_{\text{prior}}^2}{\sigma_{\text{prior}}^2}.$$

Άρα:

$$\ln p(\theta | y_{\text{obs}}) = -\frac{1}{2} \left[\theta^2 \left(\frac{1}{\sigma_{\text{noise}}^2} + \frac{1}{\sigma_{\text{prior}}^2} \right) - 2\theta \left(\frac{y_{\text{obs}}}{\sigma_{\text{noise}}^2} + \frac{\mu_{\text{prior}}}{\sigma_{\text{prior}}^2} \right) + \text{const} \right].$$

Βήμα 3: Ολοκλήρωση τετραγώνου. Ορίζουμε

$$A = \frac{1}{\sigma_{\text{noise}}^2} + \frac{1}{\sigma_{\text{prior}}^2}, \quad B = \frac{y_{\text{obs}}}{\sigma_{\text{noise}}^2} + \frac{\mu_{\text{prior}}}{\sigma_{\text{prior}}^2}.$$

Τότε

$$\ln p(\theta | y_{\text{obs}}) = -\frac{1}{2} [A\theta^2 - 2B\theta] + \text{const.}$$

Κάνουμε ολοκλήρωση τετραγώνου:

$$A\theta^2 - 2B\theta = A(\theta - B/A)^2 - \frac{B^2}{A}.$$

Άρα,

$$\ln p(\theta | y_{\text{obs}}) = -\frac{A}{2}(\theta - B/A)^2 + \text{const.}$$

Αναγνωρίζουμε τον εκθέτη της Γκαουσιανής:

$$\theta | y_{\text{obs}} \sim \mathcal{N}(\mu_{\text{post}}, \sigma_{\text{post}}^2),$$

όπου

$$\sigma_{\text{post}}^2 = \frac{1}{A} = \left(\frac{1}{\sigma_{\text{prior}}^2} + \frac{1}{\sigma_{\text{noise}}^2} \right)^{-1}, \quad \mu_{\text{post}} = \frac{B}{A} = \sigma_{\text{post}}^2 \left(\frac{\mu_{\text{prior}}}{\sigma_{\text{prior}}^2} + \frac{y_{\text{obs}}}{\sigma_{\text{noise}}^2} \right).$$

Βήμα 4: Evidence (περιθωριακή πιθανοφάνεια). Το evidence προκύπτει με περιθωριοποίηση (marginalization) ως προς θ :

$$p(y_{\text{obs}}) = \int_{-\infty}^{\infty} p(y_{\text{obs}} | \theta) p(\theta) d\theta.$$

Η συνέλιξη δύο Γκαουσιανών είναι Γκαουσιανή:

$$p(y_{\text{obs}}) = \mathcal{N}(y_{\text{obs}}; \mu_{\text{prior}}, \sigma_{\text{prior}}^2 + \sigma_{\text{noise}}^2).$$

Βήμα 5: Ερμηνεία.

- **Μέση τιμή posterior (ενημερωμένη εκτίμηση):**

$$\mu_{\text{post}} = w_{\text{prior}}\mu_{\text{prior}} + w_{\text{data}}y_{\text{obs}}, \quad w_{\text{prior}} = \frac{\sigma_{\text{noise}}^2}{\sigma_{\text{prior}}^2 + \sigma_{\text{noise}}^2}, \quad w_{\text{data}} = \frac{\sigma_{\text{prior}}^2}{\sigma_{\text{prior}}^2 + \sigma_{\text{noise}}^2}.$$

Αν η μέτρηση είναι πολύ θορυβώδης ($\sigma_{\text{noise}}^2 \gg \sigma_{\text{prior}}^2$), το prior υπερισχύει· αλλιώς, υπερισχύει το δεδομένο.

- **Διακύμανση posterior (ενημερωμένη αβεβαιότητα):**

$$\sigma_{\text{post}}^2 = \left(\frac{1}{\sigma_{\text{prior}}^2} + \frac{1}{\sigma_{\text{noise}}^2} \right)^{-1}.$$

Είναι πάντα μικρότερη από καθεμία από τις δύο αρχικές διακυμάνσεις — συνδυάζοντας δύο ανεξάρτητες πηγές πληροφορίας μειώνουμε την αβεβαιότητα.

- **Evidence (predictive likelihood):** Το $p(y_{\text{obs}})$ μετρά πόσο «πιθανή» είναι η παρατήρηση υπό το μοντέλο: χρήσιμο για σύγκριση μοντέλων ή ρύθμιση υπερπαραμέτρων.

Σύνοψη. Το παράδειγμα δείχνει την ιδιότητα της συζυγίας στα Γκαουσιανά μοντέλα: η posterior παραμένει Γκαουσιανό, με μέση τιμή και διακύμανση που υπολογίζονται αναλυτικά. Η ίδια λογική επεκτείνεται άμεσα και στην πολυδιάστατη περίπτωση, όπου οι πίνακες ακρίβειας (αντίστροφοι πίνακες συνδιακύμανσης) αθροίζονται γραμμικά.

6.2 Διαισθητική Περίληψη

Ο κανόνας του Bayes τυποποιεί τη μάθηση από δεδομένα:

$$\text{Posterior} \propto \text{Likelihood} \times \text{Prior}.$$

Συνδυάζει την προϋπάρχουσα γνώση με την παρατηρούμενη πληροφορία. Στις διακριτές περιπτώσεις αυτό εμφανίζεται ως κανονικοποιημένο σταθμισμένο άθροισμα πιθανών υποθέσεων. Στις συνεχείς περιπτώσεις οδηγεί σε κλειστές μορφές για συζυγείς κατανομές, όπως στο παράδειγμά μας.

Η συμπερασματολογία κατά Bayes είναι πολύ χρήσιμη στη σύγχρονη μηχανική μάθηση, στην επεξεργασία σήματος και στον έλεγχο: παρέχει έναν συστηματικό τρόπο υπολογισμού πεποιθήσεων και ποσοτικοποίησης της αβεβαιότητας.

7 Μέση Τιμή, Διασπορά, Συνδιασπορά και Συσχέτιση

Σε αυτήν την ενότητα εξετάζουμε τις πιο θεμελιώδεις στατιστικές ποσότητες που χρησιμοποιούνται για να περιγράψουν τυχάιες μεταβλητές. Αυτές οι ποσότητες (οι μέσες τιμές, οι διασπορές, οι συνδιασπορές και οι συσχετίσεις) αποτελούν τη ραχοκοκαλιά της θεωρίας πιθανοτήτων, και της στοχαστικής μοντελοποίησης. Περιγράφουν την κεντρική τάση, τη διασπορά και τις σχέσεις ανάμεσα σε τυχάιες ποσότητες.

7.1 Αναμενόμενες Τιμές και Μέση Τιμή

Η **αναμενόμενη τιμή** (ή **μέση τιμή**) μιας τυχάιας μεταβλητής ποσοτικοποιεί την «μέση τιμή» της με βάση τον νόμο πιθανοτήτων της. Διαισθητικά, αντιστοιχεί στην τιμή που θα προέκυπτε κατά μέσο όρο αν το ίδιο τυχαίο πείραμα επαναλαμβανόταν άπειρες φορές.

Έστω ότι x είναι μια τυχαία μεταβλητή με συνάρτηση πυκνότητας πιθανότητας $p_x(z)$ ορισμένη σε έναν δειγματοχώρο $\mathcal{X}$.

Διακριτή περίπτωση. Αν το x λαμβάνει διακριτές τιμές $z \in \mathcal{X}$, η αναμενόμενη τιμή ορίζεται ως το σταθμισμένο άθροισμα:

$$\mathbb{E}[x] = \sum_{z \in \mathcal{X}} z p_x(z).$$

Συνεχής περίπτωση. Αν το x είναι συνεχές, η αναμενόμενη τιμή του δίνεται από ολοκλήρωμα πάνω στη συνάρτηση πυκνότητας πιθανότητας:

$$\mathbb{E}[x] = \int_{\mathcal{X}} z p_x(z) dz.$$

Η αναμενόμενη τιμή είναι μια βαθμωτή σύνοψη της κατανομής της τυχάιας μεταβλητής: δίνει ένα μέτρο της κεντρικής θέσης της. Ωστόσο, δεν παρέχει πληροφορία σχετικά με τη διασπορά ή τη μεταβλητότητά της.

Πολυδιάστατη περίπτωση. Για μία τυχαία διανυσματική μεταβλητή $\mathbf{x} = [x_1, x_2, \dots, x_N]^\top \in \mathbb{R}^N$, η αναμενόμενη τιμή (ή διανυσματική μέση τιμή) ορίζεται ανά συνιστώσα:

$$\mathbb{E}[\mathbf{x}] = \begin{bmatrix} \mathbb{E}[x_1] \\ \vdots \\ \mathbb{E}[x_N] \end{bmatrix} = \begin{bmatrix} \int_{\mathcal{X}_1} z p_{x_1}(z) dz \\ \vdots \\ \int_{\mathcal{X}_N} z p_{x_N}(z) dz \end{bmatrix}.$$

Παραδείγματα.

- Αν $x \sim \text{Bernoulli}(p)$, τότε $\mathbb{E}[x] = p$. Η μέση τιμή αντιστοιχεί στην πιθανότητα επιτυχίας.
- Αν $x \sim \mathcal{N}(\mu, \sigma^2)$, τότε $\mathbb{E}[x] = \mu$.

Γραμμικότητα της αναμενόμενης τιμής. Μία από τις πιο ισχυρές ιδιότητες της αναμενόμενης τιμής είναι η γραμμικότητα:

$$\mathbb{E}[a x + b y] = a \mathbb{E}[x] + b \mathbb{E}[y],$$

για κάθε σταθερές $a, b \in \mathbb{R}$ και για οποιεσδήποτε τυχαίες μεταβλητές x, y . Αυτή η ιδιότητα ισχύει ανεξάρτητα από το αν x και y είναι ανεξάρτητες. Επιτρέπει τον αλγεβρικό χειρισμό αναμενόμενων τιμών σε πιθανολογικές παραθέσεις.

7.2 Διασπορά και Τυπική Απόκλιση

Ενώ η μέση τιμή δείχνει την κεντρική τάση μίας τυχαίας μεταβλητής, δεν παρέχει πληροφορία σχετικά με το πόσο μακριά τείνουν να αποκλίνουν τα δείγματα από αυτήν. Η **διασπορά** μιας τυχαίας μεταβλητής x μετρά τη μέση τετραγωνική απόκλιση από τη μέση τιμή:

$$\text{Var}[x] = \mathbb{E}[(x - \mathbb{E}[x])^2] = \mathbb{E}[x^2] - (\mathbb{E}[x])^2.$$

Η διασπορά είναι πάντα μη αρνητική και έχει τις ίδιες μονάδες με το τετράγωνο του x . Ποσοτικοποιεί το πόσο «απλωμένη» είναι η κατανομή του x .

Η τετραγωνική ρίζα της διασποράς είναι η **τυπική απόκλιση**:

$$\sigma_x = \sqrt{\text{Var}[x]}.$$

Η τυπική απόκλιση συχνά είναι ευκολότερο να ερμηνευτεί, καθώς έχει τις ίδιες μονάδες με τη μεταβλητή.

Παράδειγμα. Αν $x \sim \mathcal{N}(\mu, \sigma^2)$, τότε $\text{Var}[x] = \sigma^2$. Για μια τυχαία μεταβλητή Bernoulli $x \in \{0, 1\}$ με παράμετρο p , η διασπορά είναι

$$\text{Var}[x] = p(1 - p),$$

η οποία μεγιστοποιείται όταν $p = 0.5$.

7.3 Συνδιασπορά

Η **συνδιασπορά** γενικεύει την έννοια της διασποράς ώστε να περιγράψει το πώς δύο τυχαίες μεταβλητές μεταβάλλονται μαζί. Για τυχαίες μεταβλητές x και y , η συνδιασπορά ορίζεται ως:

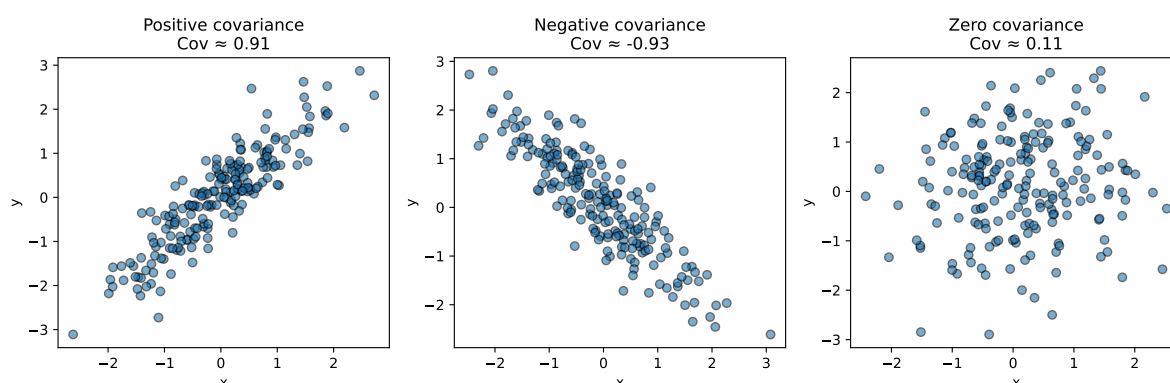
$$\text{Cov}[x, y] = \mathbb{E}[(x - \mathbb{E}[x])(y - \mathbb{E}[y])] = \mathbb{E}[xy] - \mathbb{E}[x]\mathbb{E}[y].$$

Αν $\text{Cov}[x, y] > 0$, τότε x και y τείνουν να αυξάνονται ή να μειώνονται μαζί (θετική γραμμική εξάρτηση). Αν $\text{Cov}[x, y] < 0$, τείνουν να κινούνται σε αντίθετες κατευθύνσεις. Αν $\text{Cov}[x, y] = 0$, είναι γραμμικά μη συσχετισμένες, αν και όχι απαραίτητα ανεξάρτητες.

Διαισθητικά. Η συνδιασπορά μετρά τον ευθυγραμμισμό ανάμεσα στις αποκλίσεις του x και του y από τις αντίστοιχες μέσες τιμές τους. Αν δούμε τα κεντραρισμένα δείγματα $(x - \mathbb{E}[x], y - \mathbb{E}[y])$ ως διανύσματα στο επίπεδο, η συνδιασπορά είναι ανάλογη με το μέσο εσωτερικό τους γινόμενο.

Οπτικό παράδειγμα. Τα διαγράμματα διασποράς μπορούν να αποδώσουν οπτικά τη συνδιασπορά:

- Θετική συνδιασπορά: τα σημεία συγκεντρώνονται κατά μήκος μιας ανοδικής τάσης.
- Αρνητική συνδιασπορά: τα σημεία συγκεντρώνονται κατά μήκος μιας καθοδικής τάσης.
- Μηδενική συνδιασπορά: τα σημεία σχηματίζουν ένα μη συσχετισμένο «σύννεφο» κυκλικού τύπου.



7.4 Συσχέτιση

Ο **συντελεστής συσχέτισης** παρέχει ένα κανονικοποιημένο μέτρο γραμμικής εξάρτησης μεταξύ δύο τυχαίων μεταβλητών:

$$\rho_{x,y} = \frac{\text{Cov}[x, y]}{\sigma_x \sigma_y}, \quad -1 \leq \rho_{x,y} \leq 1.$$

Επανακλιμακώνει τη συνδιασπορά με το γινόμενο των τυπικών αποκλίσεων, παράγοντας έναν αδιάστατο αριθμό. Τιμές της $\rho_{x,y}$ κοντά στο $+1$ υποδεικνύουν ισχυρή θετική γραμμική σχέση, κοντά στο -1 ισχυρή αρνητική σχέση, ενώ κοντά στο 0 ασθενή ή μηδενική γραμμική σχέση.

Σημειώνεται ότι η συσχέτιση μετρά μόνο **γραμμική** συσχέτιση. Δύο τυχαίες μεταβλητές μπορεί να έχουν μηδενική συσχέτιση αλλά να είναι στατιστικά εξαρτημένες μέσω μη γραμμικών σχέσεων.

7.5 Πολυδιάστατη Συνδιασπορά

Πολλαπλές μεταβλητές. Για τυχαία διανύσματα $\mathbf{x} \in \mathbb{R}^n, \mathbf{y} \in \mathbb{R}^m$:

$$\text{Cov}[\mathbf{x}, \mathbf{y}] = \mathbb{E}[(\mathbf{x} - \mathbb{E}[\mathbf{x}])(\mathbf{y} - \mathbb{E}[\mathbf{y}])^\top].$$

$$\text{Cov}[\mathbf{x}, \mathbf{y}] = \begin{bmatrix} \text{Cov}[x_1, y_1] & \text{Cov}[x_1, y_2] & \cdots & \text{Cov}[x_1, y_m] \\ \text{Cov}[x_2, y_1] & \text{Cov}[x_2, y_2] & \cdots & \text{Cov}[x_2, y_m] \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}[x_n, y_1] & \text{Cov}[x_n, y_2] & \cdots & \text{Cov}[x_n, y_m] \end{bmatrix} \in \mathbb{R}^{n \times m}.$$

Μητρώο/Πίνακας συνδιασποράς. Για ένα τυχαίο διάνυσμα $\mathbf{x} = [x_1, x_2, \dots, x_n]^\top$ με μέση τιμή $\boldsymbol{\mu} = \mathbb{E}[\mathbf{x}]$, ο πίνακας συνδιασποράς ορίζεται ως:

$$\boldsymbol{\Sigma} = \text{Cov}[\mathbf{x}, \mathbf{x}] = \mathbb{E}[(\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^\top].$$

Αυτό ο πίνακας περιέχει όλες τις ανά ζεύγη συνδιασπορές μεταξύ των συνιστωσών της $\mathbf{x}$:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \text{Var}[x_1] & \text{Cov}[x_1, x_2] & \cdots & \text{Cov}[x_1, x_n] \\ \text{Cov}[x_2, x_1] & \text{Var}[x_2] & \cdots & \text{Cov}[x_2, x_n] \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}[x_n, x_1] & \text{Cov}[x_n, x_2] & \cdots & \text{Var}[x_n] \end{bmatrix} \in \mathbb{R}^{n \times n}.$$

Ο πίνακας συνδιασποράς είναι συμμετρικός ($\boldsymbol{\Sigma}^\top = \boldsymbol{\Sigma}$) και θετικά ημιορισμένος ($\mathbf{a}^\top \boldsymbol{\Sigma} \mathbf{a} \geq 0$ για κάθε $\mathbf{a} \in \mathbb{R}^n$).

Ερμηνεία. Κάθε διαγώνιο στοιχείο $\Sigma_{ii} = \text{Var}[x_i]$ μετρά τη διασπορά μιας συνιστώσας, ενώ κάθε εκτός-διαγωνίου στοιχείο Σ_{ij} αντιπροσωπεύει τη συνδιασπορά μεταξύ των x_i και x_j . Η δομή του $\boldsymbol{\Sigma}$ αποκαλύπτει τις εξαρτήσεις και τις κλίμακες της πολυμεταβλητής κατανομής.

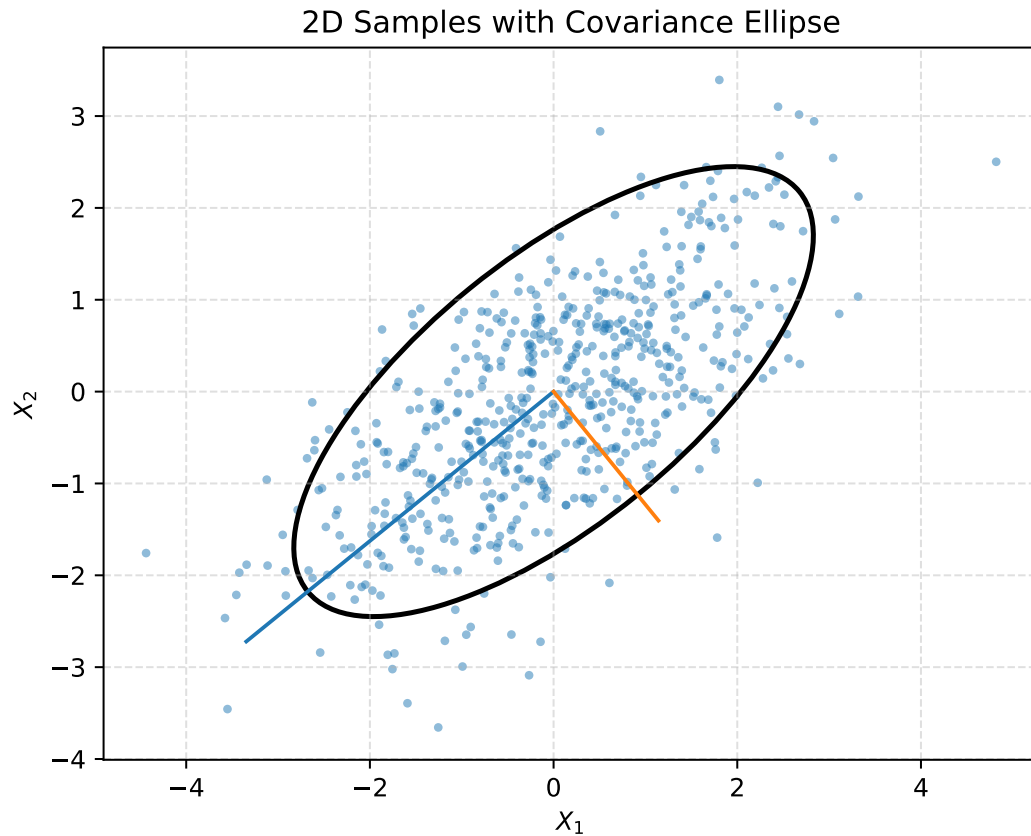
Πίνακας συσχέτισης. Ο πίνακας συσχέτισης αποτελεί μια κανονικοποιημένη, αδιάστατη εκδοχή του πίνακα συνδιασποράς:

$$\mathbf{R} = [R_{ij}] \in \mathbb{R}^{n \times n}, \quad R_{ij} = \frac{\Sigma_{ij}}{\sqrt{\Sigma_{ii}\Sigma_{jj}}}, \quad -1 \leq R_{ij} \leq 1.$$

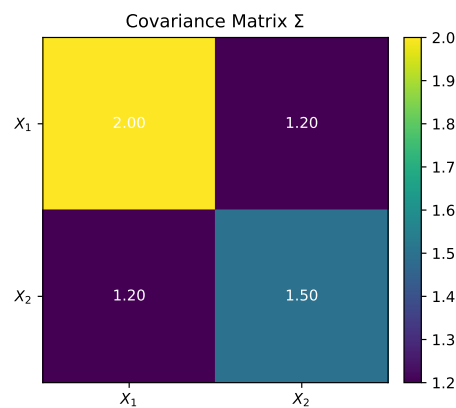
Το $\mathbf{R}$ αποτυπώνει αποκλειστικά τη δύναμη και τη φορά των ανά ζεύγη γραμμικών σχέσεων, αφαιρώντας την επίδραση από διαφορετικές μονάδες και κλίμακες.

Ο προσανατολισμός και η επιμήκυνση της έλλειψης αντανακλούν τη συσχέτιση μεταξύ των συνιστωσών. Ένα τέλεια κυκλικό σχήμα υποδηλώνει ασυσχέτιστες (ανεξάρτητες) συνιστώσες.

Απεικόνιση των $\boldsymbol{\Sigma}$ και $\mathbf{R}$. Το Σχήμα 7.2 δείχνει το μητρώο συνδιασποράς ως heatmap, όπου η ένταση του χρώματος αποτυπώνει μέγεθος και πρόσημο των συνδιασπορών. Το Σχήμα 7.3 παρουσιάζει τον αντίστοιχο πίνακα συσχέτισης, ο οποίος αναδεικνύει τις σχετικές σχέσεις ανεξάρτητα από τις μονάδες μέτρησης.



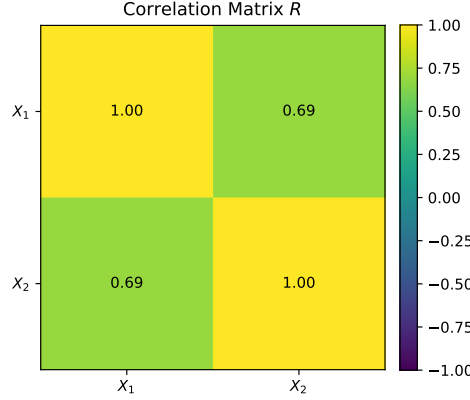
Σχήμα 7.1: Δειγματοληπτικά σημεία 2D Γκαουσιανής με την ελλειψοειδή ισοκαμπύλη που ορίζεται από τη Σ .



Σχήμα 7.2: Heatmap της μήτρας συνδιασποράς Σ .

Περίληψη.

- Η αναμενόμενη τιμή $\mathbb{E}[x]$ περιγράφει την κεντρική τάση.
- Η διασπορά $\text{Var}[x]$ μετρά τη διασπορά γύρω από τον μέσο όρο.
- Η συνδιασπορά $\text{Cov}[x, y]$ ποσοτικοποιεί το πώς δύο μεταβλητές μεταβάλλονται από κοινού.



Σχήμα 7.3: Heatmap της μήτρας συσχέτισης R .

- Η συσχέτιση $\rho_{x,y}$ κανονικοποιεί τη συνδιασπορά και δίνει ένα αδιάστατο μέτρο γραμμικής εξάρτησης.
- Ο πίνακας συνδιασποράς Σ επεκτείνει τις έννοιες αυτές στα τυχαία διανύσματα, κωδικοποιώντας τόσο τις επιμέρους διασπορές όσο και τις διασταυρούμενες συσχετίσεις.

7.6 Δειγματικοί Μέσοι και Συνδιασπορές

Στην πράξη, η πραγματική μέση τιμή και η πραγματική συνδιασπορά μιας τυχαίας μεταβλητής ή ενός τυχαίου διανύσματος σπάνια είναι γνωστές. Αντίθετα, πρέπει να *εκτιμηθούν* από ένα πεπερασμένο σύνολο δεδομένων. Ο δειγματικός μέσος και η δειγματική συνδιασπορά αποτελούν εμπειρικές εκτιμήσεις των θεωρητικών μεγεθών που εισήχθησαν προηγουμένως.

Δειγματικός μέσος. Δοθέντων N ανεξάρτητων πραγματοποιήσεων ενός τυχαίου διανύσματος $\{\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(N)}\}$, με κάθε $\mathbf{x}^{(i)} \in \mathbb{R}^d$, ο **δειγματικός μέσος** ορίζεται ως:

$$\hat{\boldsymbol{\mu}}_{\mathbf{x}} = \frac{1}{N} \sum_{i=1}^N \mathbf{x}^{(i)}.$$

Αποτελεί αμερόληπτη εκτίμηση της πραγματικής μέσης τιμής $\mathbb{E}[\mathbf{x}]$ υπό την υπόθεση i.i.d.:

$$\mathbb{E}[\hat{\boldsymbol{\mu}}_{\mathbf{x}}] = \mathbb{E}[\mathbf{x}].$$

Γεωμετρικά, ο δειγματικός μέσος αντιστοιχεί στο «κέντρο μάζας» του νέφους των δεδομένων στο $\mathbb{R}^d$.

Δειγματικός πίνακας συνδιασποράς. Η εμπειρική συνδιασπορά των δεδομένων εκτιμάται ως εξής:

$$\hat{\Sigma}_{\mathbf{x}} = \frac{1}{N-1} \sum_{i=1}^N (\mathbf{x}^{(i)} - \hat{\boldsymbol{\mu}}_{\mathbf{x}})(\mathbf{x}^{(i)} - \hat{\boldsymbol{\mu}}_{\mathbf{x}})^{\top}.$$

Ο παρονομαστής $N-1$ (αντί του N) εξασφαλίζει αμερόληπτη εκτίμηση του πραγματικού μητρώου συνδιασποράς $\Sigma_{\mathbf{x}}$. Κάθε διαγώνιο στοιχείο $\hat{\Sigma}_{ii}$ δίνει τη δειγματική διασπορά της συνιστώσας x_i , ενώ κάθε εκτός-διαγωνίου στοιχείο $\hat{\Sigma}_{ij}$ μετρά την δειγματική συνδιασπορά μεταξύ x_i και x_j .

Ερμηνεία.

- Το $\hat{\mu}_x$ δίνει το εμπειρικό κέντρο των παρατηρούμενων δεδομένων.
- Το $\hat{\Sigma}_x$ αποτυπώνει τη διασπορά και τον προσανατολισμό του νέφους δεδομένων.
- Τα ιδιοδιανύσματα της $\hat{\Sigma}_x$ αντιστοιχούν στις κύριες διευθύνσεις μεταβολής: χρησιμοποιούνται σε πολλές μεθόδους όπως Principal Component Analysis (PCA).

Python Implementation

```
1 import numpy as np
2
3 # Generate random data: d dimensions, N samples (each sample is a
4   column vector)
5 d, N = 3, 100
6 X = np.random.randn(d, N) #  $X \in \mathbb{R}^{d \times N}$ 
7
8 # Compute sample mean (column vector)
9 mu_hat = np.mean(X, axis=1, keepdims=True) # shape (d, 1)
10
11 # Center the data
12 X_centered = X - mu_hat # subtract mean from each column
13
14 # --- Manual computation of sample covariance ---
15 Sigma_hat_manual = (X_centered @ X_centered.T) / (N - 1)
16
17 # --- Using NumPy's built-in function ---
18 # np.cov expects samples as rows by default, so we transpose X
19 Sigma_hat_np = np.cov(X.T, bias=False)
20
21 print("Sample mean (column vector):\n", mu_hat)
22 print("\nSample covariance (manual):\n", Sigma_hat_manual)
23 print("\nSample covariance (np.cov):\n", Sigma_hat_np)
```

Listing 1: Sample mean and covariance.

7.7 Εσωτερικά Γινόμενα Τυχαίων Μεταβλητών

Οι τυχαίες μεταβλητές μπορούν να θεωρηθούν ως στοιχεία ενός διανυσματικού χώρου, εξοπλισμένου με ένα εσωτερικό γινόμενο το οποίο ορίζεται μέσω της αναμενόμενης τιμής. Αυτή η οπτική συνδέει τη θεωρία πιθανοτήτων με τη γραμμική άλγεβρα και τη λειτουργική ανάλυση, και αποτελεί τη βάση για έννοιες όπως η ορθογωνιότητα, οι προβολές και η εκτίμηση ελαχίστων τετραγώνων.

Ορισμός. Για τυχαία διανύσματα $x, y \in \mathbb{R}^d$, το **εσωτερικό γινόμενο** ορίζεται ως:

$$\langle x, y \rangle = \mathbb{E}[x^\top y].$$

Για βαθμωτές τυχαίες μεταβλητές αυτό απλοποιείται σε

$$\langle x, y \rangle = \mathbb{E}[xy].$$

Η αντίστοιχη νόρμα ορίζεται ως

$$\|x\| = \sqrt{\langle x, x \rangle} = \sqrt{\mathbb{E}[x^\top x]}.$$

Ιδιότητες. Το εσωτερικό γινόμενο μέσω αναμενόμενης τιμής ικανοποιεί τις κλασικές ιδιότητες:

- **Συμμετρία:** $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$.
- **Γραμμικότητα:** $\langle a\mathbf{x} + b\mathbf{z}, \mathbf{y} \rangle = a\langle \mathbf{x}, \mathbf{y} \rangle + b\langle \mathbf{z}, \mathbf{y} \rangle$.
- **Θετικότητα:** $\langle \mathbf{x}, \mathbf{x} \rangle = \mathbb{E}[\mathbf{x}^\top \mathbf{x}] \geq 0$.

Σύνδεση με στατιστικά μεγέθη. Το εσωτερικό γινόμενο ενοποιεί αρκετούς θεμελιώδεις ορισμούς:

$$\text{Μέση τιμή: } \langle \mathbf{x}, \mathbf{1} \rangle = \mathbb{E}[\mathbf{x}],$$

$$\text{Συνδιασπορά: } \text{Cov}[\mathbf{x}, \mathbf{y}] = \langle \mathbf{x} - \mathbb{E}[\mathbf{x}], \mathbf{y} - \mathbb{E}[\mathbf{y}] \rangle,$$

$$\text{Διασπορά: } \text{Var}[\mathbf{x}] = \langle \mathbf{x} - \mathbb{E}[\mathbf{x}], \mathbf{x} - \mathbb{E}[\mathbf{x}] \rangle.$$

Ερμηνεία. Με αυτή την οπτική, οι τυχαίες μεταβλητές λειτουργούν σαν διανύσματα σε έναν χώρο Hilbert, τον $\mathcal{L}^2$ (χώρος των τετραγωνικά ολοκληρώσιμων τυχαίων μεταβλητών), όπου η αναμενόμενη τιμή ορίζει το εσωτερικό γινόμενο και η διασπορά αντιστοιχεί στη νορμά στο τετράγωνο. Η ορθογωνιότητα μεταξύ τυχαίων μεταβλητών, $\langle \mathbf{x}, \mathbf{y} \rangle = 0$, σημαίνει ότι είναι *ανεξάρτητες*. Αυτή η γεωμετρική ερμηνεία θεμελιώνει πολλές μεθόδους στατιστικής εκτίμησης, όπως η γραμμική παλινδρόμηση και οι προβολές ελαχίστων τετραγώνων.

8 Ανεξαρτησία και Υπό Συνθήκη Ανεξαρτησία

Οι τυχαίες μεταβλητές μπορεί να παρουσιάζουν διάφορες μορφές στατιστικής εξάρτησης. Η κατανόηση του πότε δύο (ή περισσότερες) τυχαίες μεταβλητές είναι ανεξάρτητες (ή ανεξάρτητες υπό συνθήκη σε κάποια άλλη μεταβλητή) είναι θεμελιώδης στη θεωρία πιθανοτήτων, στη στατιστική μοντελοποίηση και στην συμπερασματολογία. Η ανεξαρτησία καθορίζει το πώς παραγοντοποιούνται οι αθροιστικές κατανομές και πώς απλοποιούνται οι προσδοκίες, οι διασπορές και οι συνδιασπορές.

8.1 Στατιστική Ανεξαρτησία

Δύο τυχαίες μεταβλητές x και y λέγονται **ανεξάρτητες** όταν η γνώση της μίας δεν παρέχει καμία πληροφορία για την άλλη. Τυπικά,

$$x \perp y \iff p_{x,y}(z_1, z_2) = p_x(z_1) p_y(z_2) \quad \text{για όλα τα } z_1, z_2.$$

Ισοδύναμα, για οποιαδήποτε μετρήσιμα σύνολα A, B ,

$$\mathbb{P}(x \in A, y \in B) = \mathbb{P}(x \in A) \mathbb{P}(y \in B).$$

Με άλλα λόγια, η από κοινού κατανομή των x και y παραγοντοποιείται στο γινόμενο των περιθωριακών τους κατανομών. Διαισθητικά, η παρατήρηση της μίας μεταβλητής δεν μας αποκαλύπτει τίποτα για την άλλη.

Παράδειγμα. Αν ρίξουμε δύο ζάρια, το αποτέλεσμα του ενός είναι ανεξάρτητο από το αποτέλεσμα του άλλου. Η πιθανότητα να δείξει το πρώτο ζάρι 3 και το δεύτερο 5 ισούται με το γινόμενο των ατομικών πιθανοτήτων.

Συνέπειες. Αν x και y είναι ανεξάρτητες:

$$\mathbb{E}[f(x)g(y)] = \mathbb{E}[f(x)]\mathbb{E}[g(y)]$$

για όλες τις μετρήσιμες συναρτήσεις f, g . Αυτή η ιδιότητα απλοποιεί πολλές πράξεις που αφορούν υπολογισμούς αναμενόμενων τιμών και διασπορών.

8.2 Υπό Συνθήκη Ανεξαρτησία

Η υπό συνθήκη ανεξαρτησία γενικεύει την έννοια της ανεξαρτησίας εισάγοντας μια τρίτη μεταβλητή w που «εξηγεί» την εξάρτηση μεταξύ x και y .

Τυπικά, οι τυχαίες μεταβλητές x και y είναι **υπό συνθήκη ανεξάρτητες δεδομένου του w** αν

$$x \perp y \mid w \iff p_{x,y|w}(z_1, z_2 \mid z_3) = p_{x|w}(z_1 \mid z_3) p_{y|w}(z_2 \mid z_3) \quad \text{για όλα τα } z_1, z_2, z_3.$$

Ισοδύναμα,

$$\mathbb{P}(x \in A, y \in B \mid w = z_3) = \mathbb{P}(x \in A \mid w = z_3) \mathbb{P}(y \in B \mid w = z_3).$$

Διαισθητικά. Δεδομένου του w , η γνώση του x δεν παρέχει επιπλέον πληροφορία για το y .

Παράδειγμα. Έστω w ο καιρός, x το αν κάποιος κρατάει ομπρέλα, και y το αν το έδαφος είναι βρεγμένο. Συνολικά, x και y είναι εξαρτημένες μεταβλητές, αλλά γίνονται υπό συνθήκη ανεξάρτητες δεδομένου του w : μόλις γνωρίζουμε τον καιρό, το αν κάποιος κρατάει ομπρέλα δεν προσθέτει πληροφορία για το αν το έδαφος είναι βρεγμένο.

Σημαντικές διακρίσεις.

- Η ανεξαρτησία δεν συνεπάγεται υπό συνθήκη ανεξαρτησία.
- Η υπό συνθήκη ανεξαρτησία δεν συνεπάγεται ανεξαρτησία.

8.3 Επίδραση της Ανεξαρτησίας στη Συνδιασπορά και στη Διασπορά

Η στατιστική ανεξαρτησία οδηγεί σε σημαντικές απλοποιήσεις σε υπολογισμούς αναμενόμενων τιμών και διασπορών, καθώς τα διασταυρούμενα μέλη μηδενίζονται.

Συνδιασπορά υπό ανεξαρτησία. Αν x και y είναι ανεξάρτητες τυχαίες μεταβλητές, τότε:

$$\text{Cov}[x, y] = \mathbb{E}[xy] - \mathbb{E}[x]\mathbb{E}[y] = 0.$$

Άρα, η ανεξαρτησία συνεπάγεται ότι οι μεταβλητές είναι **μη συσχετισμένες**. Ωστόσο, το αντίστροφο δεν ισχύει: $\text{Cov}[x, y] = 0$ δεν σημαίνει απαραίτητα ανεξαρτησία, εκτός από ειδικές περιπτώσεις (π.χ. σε από κοινού Γκαουσιανές μεταβλητές).

Διασπορά αθροίσματος. Για ανεξάρτητες τυχαίες μεταβλητές x και y ,

$$\text{Var}[x + y] = \text{Var}[x] + \text{Var}[y].$$

Πιο γενικά, για ανεξάρτητες μεταβλητές $\{x_i\}_{i=1}^n$:

$$\text{Var}\left(\sum_{i=1}^n x_i\right) = \sum_{i=1}^n \text{Var}[x_i].$$

Το αποτέλεσμα αυτό προκύπτει από τη γραμμικότητα της αναμενόμενης τιμής και τον μη-δενισμό των όρων συνδιασποράς λόγω ανεξαρτησίας.

Συσχέτιση. Η ανεξαρτησία συνεπάγεται επίσης μηδενική συσχέτιση:

$$x \perp y \Rightarrow \rho_{x,y} = 0.$$

Ωστόσο, τυχαίες μεταβλητές με μηδενική συσχέτιση ($\rho_{x,y} = 0$) μπορεί να είναι εξαρτημένες μέσω μη-γραμμικών σχέσεων.

Ερμηνεία.

- Η ανεξαρτησία μηδενίζει διασταυρούμενους όρους σε διασπορές και συνδιασπορές.
- Η μη συσχέτιση είναι ασθενέστερη ιδιότητα: μπορεί να υπάρχουν εξαρτήσεις που δεν είναι γραμμικές.
- Για Γκαουσιανές μεταβλητές, μη συσχέτιση και ανεξαρτησία συμπίπτουν.

Παράδειγμα: Μη συσχετισμένες αλλά εξαρτημένες. Έστω $x \sim \mathcal{U}(-1, 1)$ και $y = x^2$. Τότε $\mathbb{E}[x] = 0$, $\mathbb{E}[y] > 0$ και $\text{Cov}[x, y] = 0$, αλλά το y εξαρτάται σαφώς από το x , άρα οι μεταβλητές δεν είναι ανεξάρτητες.

9 Κανονικές (Gaussian) Κατανομές

Η κανονική (ή *Gaussian*) κατανομή κατέχει κεντρική θέση στην πιθανότητα, τη στατιστική και την επεξεργασία σήματος. Προκύπτει φυσικά ως η οριακή κατανομή αθροισμάτων τυχαίων μεταβλητών (μέσω του Κεντρικού Οριακού Θεωρήματος) και είναι μαθηματικά ιδιαίτερα βολική λόγω της αναλυτικής της μορφής και των ιδιοτήτων κλειστότητας υπό κοινές πράξεις όπως οι γραμμικοί μετασχηματισμοί, η υπό συνθήκη κατανομή και η περιθωριοποίηση.

9.1 Μονοδιάστατη Κανονική Κατανομή

Μια πραγματική τυχαία μεταβλητή x ακολουθεί **κανονική κατανομή** με μέση τιμή μ και διασπορά $\sigma^2 > 0$ αν η συνάρτηση πυκνότητας πιθανότητας (PDF) της είναι

$$p_x(z) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(z - \mu)^2}{2\sigma^2}\right), \quad x \sim \mathcal{N}(\mu, \sigma^2).$$

Παράμετροι.

- Η μέση τιμή μ καθορίζει τη θέση (κέντρο) της κατανομής.
- Η διασπορά σ^2 καθορίζει την εύρος ή διασπορά των τιμών.

Ιδιότητες.

- Η κανονική κατανομή προσδιορίζεται πλήρως από τις δύο πρώτες ροπές (μ, σ^2) .
- Είναι συμμετρική ως προς το μ .
- Περίπου 68.3%, 95.4% και 99.7% των δειγμάτων βρίσκονται εντός 1σ , 2σ και 3σ από τη μέση τιμή, αντίστοιχα.

9.2 Πολυδιάστατη Κανονική Κατανομή

Η πολυδιάστατη κανονική κατανομή γενικεύει την κανονική κατανομή σε τυχαία διανύσματα $\mathbf{x} \in \mathbb{R}^d$.

Ορισμός.

$$p_{\mathbf{x}}(\mathbf{z}) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{z} - \boldsymbol{\mu})^T \Sigma^{-1}(\mathbf{z} - \boldsymbol{\mu})\right), \quad \mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma),$$

όπου

- $\boldsymbol{\mu} \in \mathbb{R}^d$ είναι το διάνυσμα μέσης τιμής,
- $\Sigma \in \mathbb{R}^{d \times d}$ είναι ο πίνακας συνδιασποράς, συμμετρικός και θετικά ορισμένος.

Γεωμετρική ερμηνεία.

- Οι ισοεπίπεδες καμπύλες της πυκνότητας είναι ελλείψεις (ή υπερελλείψεις) με κέντρο το $\boldsymbol{\mu}$.
- Τα ιδιοδιανύσματα του Σ καθορίζουν τις κύριες διευθύνσεις (άξονες) της έλλειψης.
- Οι αντίστοιχες ιδιοτιμές καθορίζουν τα τετράγωνα των μηκών των αξόνων (δηλ. τον βαθμό διασποράς ανά κατεύθυνση).

Βασική ιδιότητα. Η πολυδιάστατη κανονική κατανομή καθορίζεται πλήρως από τις πρώτες δύο ροπές $(\boldsymbol{\mu}, \Sigma)$ και μεγιστοποιεί την εντροπία μεταξύ όλων των κατανομών με δεδομένη μέση τιμή και συνδιασπορά.

9.3 Από Κοινού, Περιθώριες και Υπό Συνθήκη Κατανομές

Έστω

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}\right),$$

όπου $\mathbf{x}_1 \in \mathbb{R}^{d_1}$, $\mathbf{x}_2 \in \mathbb{R}^{d_2}$ και $d_1 + d_2 = d$.

Περιθώριες κατανομές. Κάθε τμήμα (block) είναι επίσης κανονικά κατανομημένο:

$$\mathbf{x}_1 \sim \mathcal{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11}), \quad \mathbf{x}_2 \sim \mathcal{N}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_{22}).$$

Υπό συνθήκη κατανομή. Υπό την συνθήκη $\mathbf{x}_2 = \mathbf{z}_2$:

$$\mathbf{x}_1 \mid \mathbf{x}_2 = \mathbf{z}_2 \sim \mathcal{N}(\boldsymbol{\mu}_{1|2}, \boldsymbol{\Sigma}_{1|2}),$$

όπου

$$\begin{aligned} \boldsymbol{\mu}_{1|2} &= \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12} \boldsymbol{\Sigma}_{22}^{-1} (\mathbf{z}_2 - \boldsymbol{\mu}_2), \\ \boldsymbol{\Sigma}_{1|2} &= \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12} \boldsymbol{\Sigma}_{22}^{-1} \boldsymbol{\Sigma}_{21}. \end{aligned}$$

Ιδιότητα κλειστότητας. Η οικογένεια των κανονικών κατανομών είναι κλειστή ως προς:

- **Περιθωριοποίηση:** κάθε υποσύνολο μεταβλητών είναι κανονικά κατανομημένο.
- **Υπό συνθήκη κατανομή:** οι υπό συνθήκη κατανομές παραμένουν κανονικές.

Ερμηνεία. Η υπό συνθήκη γνώση της $\mathbf{x}_2$ μετατοπίζει γραμμικά τη μέση τιμή της $\mathbf{x}_1$, ενώ η συνδιασπορά της $\mathbf{x}_1$ μειώνεται στις κατευθύνσεις που «εξηγούνται» από την $\mathbf{x}_2$.

Εξαγωγή της υπό συνθήκη κατανομής

Διατύπωση. Διαμερίζουμε

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix}, \quad \boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \quad \boldsymbol{\Sigma}_x = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}, \quad \mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}_x).$$

Από κοινού πυκνότητα. Ο πίνακας ακρίβειας (precision matrix) $\boldsymbol{\Lambda}_x = \boldsymbol{\Sigma}_x^{-1}$,

$$p_x(\mathbf{z}) \propto \exp\left(-\frac{1}{2}(\mathbf{z} - \boldsymbol{\mu})^\top \boldsymbol{\Lambda}_x (\mathbf{z} - \boldsymbol{\mu})\right), \quad \boldsymbol{\Lambda}_x = \begin{bmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{21} & \Lambda_{22} \end{bmatrix}.$$

Υπό συνθήκη $\mathbf{x}_2 = \mathbf{z}_2$. Κρατάμε μόνο τους όρους που εξαρτώνται από $\mathbf{x}_1$:

$$\begin{aligned} & -\frac{1}{2} \begin{bmatrix} \mathbf{x}_1 - \boldsymbol{\mu}_1 \\ \mathbf{z}_2 - \boldsymbol{\mu}_2 \end{bmatrix}^\top \begin{bmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{21} & \Lambda_{22} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 - \boldsymbol{\mu}_1 \\ \mathbf{z}_2 - \boldsymbol{\mu}_2 \end{bmatrix} \\ &= -\frac{1}{2}(\mathbf{x}_1 - \boldsymbol{\mu}_1)^\top \Lambda_{11}(\mathbf{x}_1 - \boldsymbol{\mu}_1) - (\mathbf{x}_1 - \boldsymbol{\mu}_1)^\top \Lambda_{12}(\mathbf{z}_2 - \boldsymbol{\mu}_2) + (\text{όροι ανεξάρτητοι από } \mathbf{x}_1). \end{aligned}$$

Ολοκλήρωση τετραγώνου ως προς $\mathbf{x}_1$. Ορίζουμε

$$\mathbf{h} = \Lambda_{12}(\mathbf{z}_2 - \boldsymbol{\mu}_2).$$

Τότε ο εκθέτης που εξαρτάται από $\mathbf{x}_1$ είναι

$$-\frac{1}{2}(\mathbf{x}_1 - \boldsymbol{\mu}_1)^\top \Lambda_{11}(\mathbf{x}_1 - \boldsymbol{\mu}_1) - (\mathbf{x}_1 - \boldsymbol{\mu}_1)^\top \mathbf{h}.$$

Ολοκληρώνουμε το τετράγωνο:

$$-\frac{1}{2}(\mathbf{x}_1 - \boldsymbol{\mu}_1 + \Lambda_{11}^{-1}\mathbf{h})^\top \Lambda_{11}(\mathbf{x}_1 - \boldsymbol{\mu}_1 + \Lambda_{11}^{-1}\mathbf{h}) + \frac{1}{2}\mathbf{h}^\top \Lambda_{11}^{-1}\mathbf{h}.$$

Επομένως

$$\mathbf{x}_1 \mid \mathbf{x}_2 = \mathbf{z}_2 \sim \mathcal{N}(\boldsymbol{\mu}_1 - \Lambda_{11}^{-1}\Lambda_{12}(\mathbf{z}_2 - \boldsymbol{\mu}_2), \Lambda_{11}^{-1}).$$

Συσχέτιση με τα μπλοκ συνδιασποράς. Χρησιμοποιώντας την ταυτότητα αντιστροφής κατά μπλοκ:

$$\Lambda_{11}^{-1} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}, \quad -\Lambda_{11}^{-1}\Lambda_{12} = \Sigma_{12}\Sigma_{22}^{-1},$$

προκύπτει η τυπική μορφή:

$$\begin{aligned} \mathbf{x}_1 \mid \mathbf{x}_2 = \mathbf{z}_2 &\sim \mathcal{N}(\boldsymbol{\mu}_{1|2}, \Sigma_{1|2}), \\ \boldsymbol{\mu}_{1|2} &= \boldsymbol{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{z}_2 - \boldsymbol{\mu}_2), \\ \Sigma_{1|2} &= \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}. \end{aligned}$$

Python Implementation

```

1 import numpy as np
2
3 # -----
4 # Helper utilities (optional)
5 # -----
6 def is_psd(A, tol=1e-10):
7     """Check positive semidefiniteness."""
8     eigvals = np.linalg.eigvalsh((A + A.T) / 2.0)
9     return np.all(eigvals >= -tol)
10
11 def symmetrize(A):
12     """Numerical symmetrization."""
13     return 0.5 * (A + A.T)
14
15
16 # =====
17 # (A) Build a joint Gaussian from two separate Gaussians
18 # =====
19 # Given:
20 # x1 ~ N(mu1, Sigma11), x2 ~ N(mu2, Sigma22)
21 # We form:
22 # x = [x1; x2] ~ N([mu1; mu2], [[Sigma11, Sigma12], [Sigma21, Sigma22]])
23
24 # Dimensions
25 d1, d2 = 2, 3
26
27 # Means (column vectors)
28 mu1 = np.array([[1.0], [-0.5]]) # shape (d1, 1)
29 mu2 = np.array([[0.2], [1.0], [0.3]]) # shape (d2, 1)
30
31 # Covariances (blocks)
32 Sigma11 = np.array([[1.0, 0.3],
33                     [0.3, 2.0]]) # (d1 x d1)

```



```

34 Sigma22 = np.array([[1.5, 0.1, 0.0],
35                     [0.1, 1.2, 0.2],
36                     [0.0, 0.2, 0.8]])      # (d2 x d2)
37
38 # ---- Case A1: Independent x1 and x2 (no cross-covariance) ----
39 Sigma12_indep = np.zeros((d1, d2))      # (d1 x d2)
40 Sigma21_indep = Sigma12_indep.T         # (d2 x d1)
41
42 mu_joint_indep = np.vstack([mu1, mu2])  # (d1+d2, 1)
43 Sigma_joint_indep = np.block([
44     [Sigma11,      Sigma12_indep],
45     [Sigma21_indep, Sigma22      ]
46 ])
47
48 Sigma_joint_indep = symmetrize(Sigma_joint_indep)
49 assert is_psd(Sigma_joint_indep), "Joint covariance (independent) not
    PSD!"
50
51 # ---- Case A2: Correlated x1 and x2 (non-zero cross-covariance) ----
52 # Provide a valid cross-covariance Sigma12; ensure the final block
    matrix is PSD.
53 Sigma12 = np.array([[0.4, -0.2, 0.1],
54                     [0.6,  0.1, 0.0]])  # (d1 x d2)
55 Sigma21 = Sigma12.T                    # (d2 x d1)
56
57 mu_joint_corr = np.vstack([mu1, mu2])   # same stacked mean
58 Sigma_joint_corr = np.block([
59     [Sigma11, Sigma12],
60     [Sigma21, Sigma22]
61 ])
62
63 Sigma_joint_corr = symmetrize(Sigma_joint_corr)
64 assert is_psd(Sigma_joint_corr), "Joint covariance (correlated) not PSD
    !"
65
66 print("Joint mean (independent):\n", mu_joint_indep)
67 print("Joint covariance (independent):\n", Sigma_joint_indep)
68 print("\nJoint mean (correlated):\n", mu_joint_corr)
69 print("Joint covariance (correlated):\n", Sigma_joint_corr)
70
71
72 # =====
73 # (B) Extract marginals from a larger joint Gaussian
74 # =====
75 # Suppose we are given a big joint:
76 mu_big = mu_joint_corr                # shape (d1+d2, 1)
77 Sigma_big = Sigma_joint_corr           # shape (d1+d2, d1+d2)
78 d = d1 + d2
79
80 # By construction, we assume the variable order is x = [x1; x2]
81 # Index ranges for slicing:
82 idx1 = slice(0, d1)                   # rows/cols for x1-block
83 idx2 = slice(d1, d1 + d2)             # rows/cols for x2-block
84
85 # Extract marginals:
86 mu1_marg = mu_big[idx1, :]            # (d1, 1)
87 mu2_marg = mu_big[idx2, :]            # (d2, 1)
88 Sigma11_marg = Sigma_big[idx1, idx1]  # (d1, d1)

```

```

89 Sigma22_marg = Sigma_big[idx2, idx2]          # (d2, d2)
90 Sigma12_marg = Sigma_big[idx1, idx2]          # (d1, d2)
91 Sigma21_marg = Sigma_big[idx2, idx1]          # (d2, d1)
92
93 print("\n--- Extracted marginals from the big joint ---")
94 print("mu1 (marg):\n", mu1_marg)
95 print("mu2 (marg):\n", mu2_marg)
96 print("Sigma11 (marg):\n", Sigma11_marg)
97 print("Sigma22 (marg):\n", Sigma22_marg)
98 print("Sigma12 (marg):\n", Sigma12_marg)
99 print("Sigma21 (marg):\n", Sigma21_marg)
100
101 # (Optional) Consistency checks against the originals used to build the
    joint
102 assert np.allclose(mu1_marg, mu1)
103 assert np.allclose(mu2_marg, mu2)
104 assert np.allclose(Sigma11_marg, Sigma11)
105 assert np.allclose(Sigma22_marg, Sigma22)
106 assert np.allclose(Sigma12_marg, Sigma12)
107 assert np.allclose(Sigma21_marg, Sigma21)
108 print("\nConsistency checks passed: extracted blocks match originals.")
109
110 # =====
111 # (C) Conditional distribution:  $x_1 \mid x_2 = z_2$ 
112 # =====
113 z2 = np.array([[0.5], [1.2]]) # observed value of  $x_2$ 
114
115 # Conditional mean and covariance
116 mu_cond = mu1_marg + Sigma12_marg @ np.linalg.inv(Sigma22_marg) @ (z2 -
    mu2_marg)
117 Sigma_cond = Sigma11_marg - Sigma12_marg @ np.linalg.inv(Sigma22_marg)
    @ Sigma21_marg
118
119 print("\nConditional mean mu_{1|2}:\n", mu_cond)
120 print("Conditional covariance Sigma_{1|2}:\n", Sigma_cond)

```

Listing 2: Gaussian Distribution: Joint, Marginals and Conditionals

9.4 Γινόμενο Γκαουσιανών Πυκνοτήτων

Το γινόμενο δύο Γκαουσιανών πυκνοτήτων στην ίδια μεταβλητή είναι μια νέα (scaled) Γκαουσιανή.

Αποτέλεσμα. Για $x \in \mathbb{R}^d$,

$$\mathcal{N}(x; \mu_1, \Sigma_1) \mathcal{N}(x; \mu_2, \Sigma_2) = c \mathcal{N}(x; \mu^*, \Sigma^*),$$

όπου

$$\begin{aligned} \Sigma^* &= (\Sigma_1^{-1} + \Sigma_2^{-1})^{-1}, \\ \mu^* &= \Sigma^* (\Sigma_1^{-1} \mu_1 + \Sigma_2^{-1} \mu_2). \end{aligned}$$

Ερμηνεία.

- Οι “ακρίβειες” (αντίστροφες συνδιασπορές) αθροίζονται: η πληροφορία συνδυάζεται γραμμικά.
- Η τελική μέση τιμή είναι ένας σταθμισμένος ως προς την ακρίβεια μέσος των μ_1 και μ_2 .
- Η σταθερά κανονικοποίησης c εξασφαλίζει ότι το γινόμενο ολοκληρώνεται στο ένα.

Εφαρμογές. Η πράξη αυτή εμφανίζεται στην Bayesian ενημέρωση με Γκαουσιανή εκ των προτέρων και πιθανοφάνεια, στα φίλτρα Kalman και στη συνένωση αισθητήρων.

9.5 Γραμμικοί Μετασχηματισμοί Γκαουσιανών

Η οικογένεια των Γκαουσιανών είναι επίσης κλειστή κάτω από αφινικούς μετασχηματισμούς.

Αποτέλεσμα. Αν

$$\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}_x, \boldsymbol{\Sigma}_x) \quad \text{και} \quad \mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b},$$

τότε

$$\mathbf{y} \sim \mathcal{N}(\mathbf{A}\boldsymbol{\mu}_x + \mathbf{b}, \mathbf{A}\boldsymbol{\Sigma}_x\mathbf{A}^\top).$$

Ερμηνεία.

- Η Γκαουσιανή μορφή διατηρείται κάτω από οποιονδήποτε αφινικό (γραμμικό + μετατόπιση) μετασχηματισμό.
- Η μέση τιμή μετασχηματίζεται γραμμικά ενώ η συνδιασπορά μετασχηματίζεται τετραγωνικά.

Εφαρμογές. Οι αφινικοί μετασχηματισμοί Γκαουσιανών είναι θεμελιώδεις σε:

- Γραμμικά δυναμικά συστήματα,
- Μοντέλα χώρου κατάστασης,
- Το βήμα πρόβλεψης του Kalman filter.

9.6 Δειγματοληψία από Γκαουσιανές Κατανομές

Μονοδιάστατη περίπτωση. Αν $x \sim \mathcal{N}(\mu, \sigma^2)$, τότε

$$x = \mu + \sigma z, \quad z \sim \mathcal{N}(0, 1).$$

Πολυδιάστατη περίπτωση. Αν $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, μπορούμε να δειγματοληπτήσουμε μέσω:

$$\mathbf{x} = \boldsymbol{\mu} + \mathbf{L}\mathbf{z}, \quad \mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}),$$

όπου $\mathbf{L}$ είναι οποιοδήποτε μητρώο που ικανοποιεί $\mathbf{L}\mathbf{L}^\top = \boldsymbol{\Sigma}$, π.χ. μέσω ανάλυσης Cholesky.

Μετασχηματισμός Box-Muller (1D). Κλασική μέθοδος παραγωγής ανεξάρτητων κανονικών δειγμάτων από ομοιόμορφα:

$$u_1, u_2 \sim \text{Uniform}(0, 1) \Rightarrow \begin{cases} z_1 = \sqrt{-2 \ln u_1} \cos(2\pi u_2), \\ z_2 = \sqrt{-2 \ln u_1} \sin(2\pi u_2), \end{cases}$$

με $z_1, z_2 \sim \mathcal{N}(0, 1)$.

Ερμηνεία. Η παραγωγή Γκαουσιανών δειγμάτων γίνεται μέσω γραμμικού μετασχηματισμού τυπικά κανονικών δειγμάτων, επιτρέποντας προσομοίωση με οποιαδήποτε μέση τιμή και συνδιασπορά.

Python Implementation

```

1 import numpy as np
2
3 # --- Univariate Gaussian ---
4 mu = 2.0          # mean
5 sigma = 0.5       # standard deviation
6 N = 10           # number of samples
7
8 # Sample from N(mu, sigma^2)
9 z = np.random.randn(N)          # z ~ N(0, 1)
10 x = mu + sigma * z              # x ~ N(mu, sigma^2)
11
12 print("Univariate samples shape:", x.shape)
13
14 # --- Multivariate Gaussian ---
15 d = 3                  # dimensionality
16 mu_vec = np.array([[1.0], [2.0], [-1.0]]) # mean vector (d x 1)
17 Sigma = np.array([[1.0, 0.5, 0.2],
18                  [0.5, 2.0, 0.3],
19                  [0.2, 0.3, 1.5]])         # covariance matrix (d x d)
20
21 # Cholesky decomposition: Sigma = L L^T
22 L = np.linalg.cholesky(Sigma)
23
24 # Sample from N(mu, Sigma)
25 Z = np.random.randn(d, N)          # Z ~ N(0, I)
26 X = mu_vec + L @ Z                 # X ~ N(mu, Sigma)
27
28 print("Multivariate samples shape:", X.shape)

```

Listing 3: Sampling from Gaussian distributions