6

ΚΕΦΑΛΑΙΟ 6

Κανονική Κατανομή

Η κανονική κατανομή, γνωστή και ως κατανομή Gauss, συνιστά αναμφισβήτητα το σημαντικότερο και ευρύτερα εφαρμοζόμενο πιθανοτικό μοντέλο στην Στατιστική Επιστήμη, τη Θεωρία Πιθανοτήτων και την Ανάλυση Δεδομένων. Η κεντρική της θέση στη σύγχρονη στατιστική μεθοδολογία οφείλεται τόσο στις εξαιρετικές μαθηματικές της ιδιότητες όσο και στην εκπληκτική της ικανότητα να περιγράφει με ακρίβεια μια τεράστια ποικιλία φυσικών, βιολογικών, οικονομικών και κοινωνικών φαινομένων. Από τη μέτρηση ανθρωπομετρικών χαρακτηριστικών και την ανάλυση σφαλμάτων στις φυσικές επιστήμες, έως την αποτίμηση χρηματοοικονομικών προϊόντων και τη μοντελοποίηση συμπεριφοράς καταναλωτών, η κανονική κατανομή αποτελεί το αναλυτικό εργαλείο για την κατανόηση και πρόβλεψη της μεταβλητότητας στον πραγματικό κόσμο.

Η θεμελιώδης σημασία της κανονικής κατανομής πηγάζει από το Κεντρικό Οριακό Θεώρημα, ένα από τα σημαντικότερα αποτελέσματα της μαθηματικής στατιστικής, το οποίο διατυπώνει ότι το άθροισμα (ή ο μέσος όρος) μεγάλου αριθμού ανεξάρτητων τυχαίων μεταβλητών προσεγγίζει την κανονική κατανομή, ανεξάρτητα από την κατανομή των επιμέρους μεταβλητών. Αυτή η αξιοσημείωτη ιδιότητα εξηγεί γιατί τόσα πολλά φαινόμενα στη φύση και την κοινωνία ακολουθούν προσεγγιστικά την κανονική κατανομή: οι περισσότερες παρατηρούμενες μεταβλητές αποτελούν αποτέλεσμα της σωρευτικής επίδρασης πολλών μικρών, ανεξάρτητων παραγόντων. Επιπλέον, η κανονική κατανομή διαθέτει μια σειρά ελκυστικών μαθηματικών ιδιοτήτων: είναι πλήρως προσδιοριζόμενη από δύο παραμέτρους (μέση τιμή και διασπορά), διατηρεί το σχήμα της υπό γραμμικούς μετασχηματισμούς και το άθροισμα ανεξάρτητων κανονικών μεταβλητών παράγει πάλι κανονική μεταβλητή.

Το παρόν κεφάλαιο αναπτύσσει συστηματικά τη θεωρία και τις εφαρμογές της κανονικής κατανομής, υιοθετώντας μια προσέγγιση που ισορροπεί μεταξύ μαθηματικής αυστηρότητας και πρακτικής εφαρμοσιμότητας.

6.1 Ιδιότητες και Χαρακτηριστικά

Η κανονική κατανομή είναι η πιο σημαντική και ευρέως χρησιμοποιούμενη συνεχής κατανομή πιθανότητας στη στατιστική και την επιστήμη των δεδομένων. Εμφανίζεται σε πληθώρα φυσικών και κοινωνικών φαινομένων και αποτελεί τη βάση πολλών στατιστικών μεθόδων και τεχνικών ανάλυσης.

Ορισμός 6.1 (*Κανονική Κατανομή*). Μια συνεχής τυχαία μεταβλητή X ακολουθεί *κανονική κατανομή* (normal distribution) με παραμέτρους μ (μέση τιμή) και σ^2 (διασπορά) και συμβολίζεται με

$$X \sim N(\mu, \sigma^2),$$

αν η συνάρτηση πυκνότητας πιθανότητας της είναι:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty$$

Οι παράμετροι της κανονικής κατανομής είναι:

- μ: η μέση τιμή, που καθορίζει το κέντρο της κατανομής
- σ: η τυπική απόκλιση, που καθορίζει το εύρος (spread) της κατανομής
- σ^2: η διασπορά

Η Εικόνα 6.1 απεικονίζει τρεις κανονικές κατανομές με διαφορετικές μέσες τιμές ($\mu = -2, 0, 2$) αλλά την ίδια τυπική απόκλιση ($\sigma = 1$). Παρατηρούμε ότι η μέση τιμή μ καθορίζει το κέντρο της κατανομής, δηλαδή η καμπύλη «*μετακινείται*» οριζόντια χωρίς να αλλάζει το σχήμα της. Οι τρεις καμπύλες έχουν το ίδιο ύψος και το ίδιο εύρος, αλλά είναι κεντραρισμένες σε διαφορετικά σημεία του άξονα x. Αυτό επιβεβαιώνει ότι η παράμετρος μ λειτουργεί ως *παράμετρος θέσης* (location parameter), μετατοπίζοντας την κατανομή χωρίς να επηρεάζει τη διασπορά της.

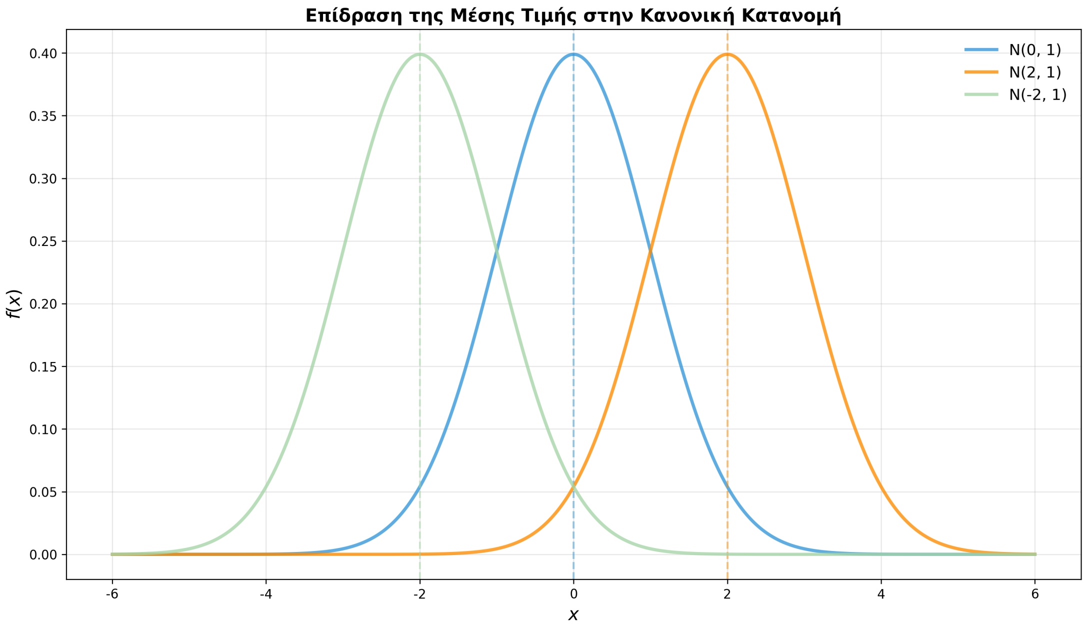
Η Εικόνα 6.2 παρουσιάζει τρεις κανονικές κατανομές με την ίδια μέση τιμή ($\mu = 0$) αλλά διαφορετικές τυπικές αποκλίσεις ($\sigma = 0.5, 1, 2$). Η τυπική απόκλιση σ καθορίζει το εύρος και την «*αιχμηρότητα*» της καμπύλης. Όταν σ είναι μικρό (π.χ. $\sigma = 0.5$), η καμπύλη είναι ψηλή και στενή, υποδεικνύοντας ότι οι τιμές είναι συγκεντρωμένες κοντά στη μέση τιμή. Αντίθετα, όταν σ είναι μεγάλο (π.χ. $\sigma = 2$), η καμπύλη είναι χαμηλή και πλατιά, δείχνοντας μεγαλύτερη διασπορά των τιμών. Το εμβαδόν κάτω από κάθε καμπύλη παραμένει ίσο με 1, ανεξάρτητα από την τιμή του σ, διατηρώντας τη βασική ιδιότητα της κατανομής πιθανότητας.

6.1.1 Βασικές Ιδιότητες

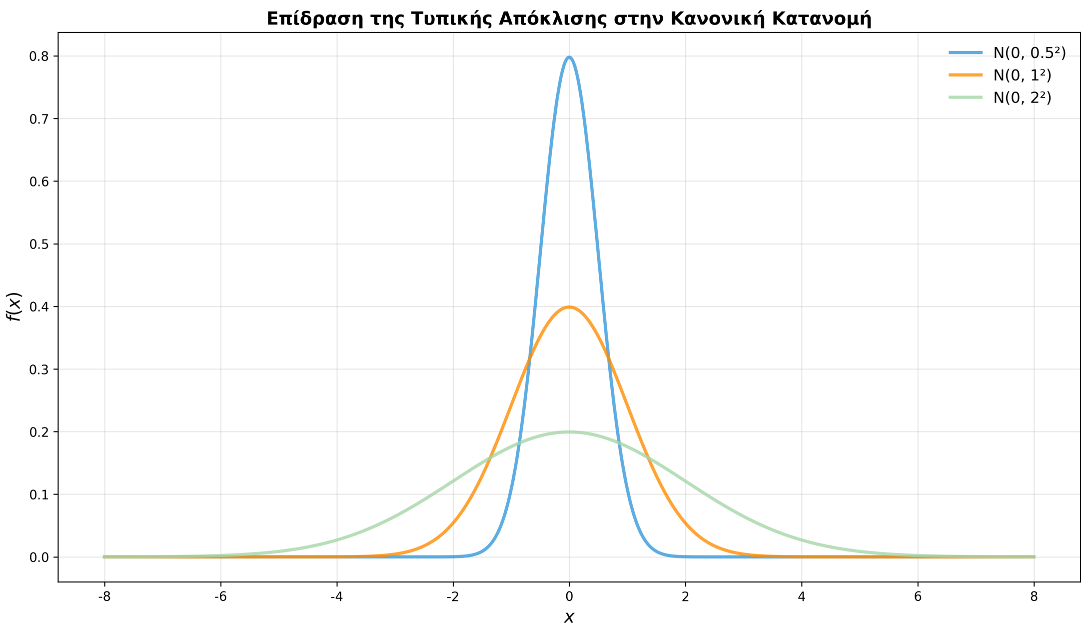
Η κανονική κατανομή διαθέτει μια σειρά χαρακτηριστικών ιδιοτήτων που την καθιστούν ιδιαίτερα χρήσιμη:

1. *Συμμετρία*: Η κατανομή είναι συμμετρική ως προς τη μέση τιμή μ. Αυτό σημαίνει ότι

$$f(\mu + a) = f(\mu - a) \quad \text{για κάθε } a > 0.$$



Εικόνα 6.1: Καμπύλες κανονικής κατανομής με διαφορετικές μέσες τιμές (σταθερή τυπική απόκλιση $\sigma = 1$)



Εικόνα 6.2: Καμπύλες κανονικής κατανομής με διαφορετικές τυπικές αποκλίσεις (σταθερή μέση τιμή $\mu = 0$)

2. *Σχήμα καμπάνας*: Η γραφική παράσταση της συνάρτησης πυκνότητας έχει χαρακτηριστικό καμπανοειδές σχήμα, γνωστό ως καμπύλη Gauss.
3. *Ασύμπτωτη συμπεριφορά*: Η καμπύλη πλησιάζει ασυμπτωτικά τον άξονα x καθώς $x \to \pm\infty$, χωρίς ποτέ να τον αγγίζει.
4. *Μέγιστο*: Η μέγιστη τιμή της συνάρτησης πυκνότητας εμφανίζεται στο $x = \mu$ και είναι:

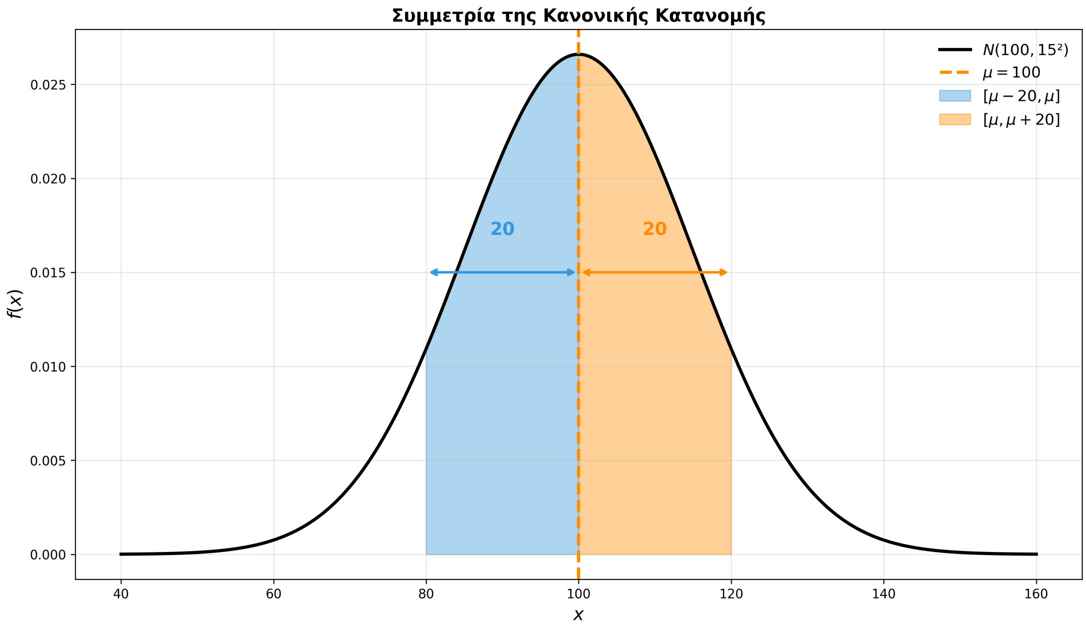
$$f(\mu) = \frac{1}{\sigma\sqrt{2\pi}}$$

5. *Σημεία Καμπής:* Η καμπύλη έχει σημεία καμπής στα $x = \mu \pm \sigma$.
6. *Εμβαδόν:* Το συνολικό εμβαδόν κάτω από την καμπύλη είναι 1:

$$\int_{-\infty}^{\infty} f(x)\,dx = 1$$

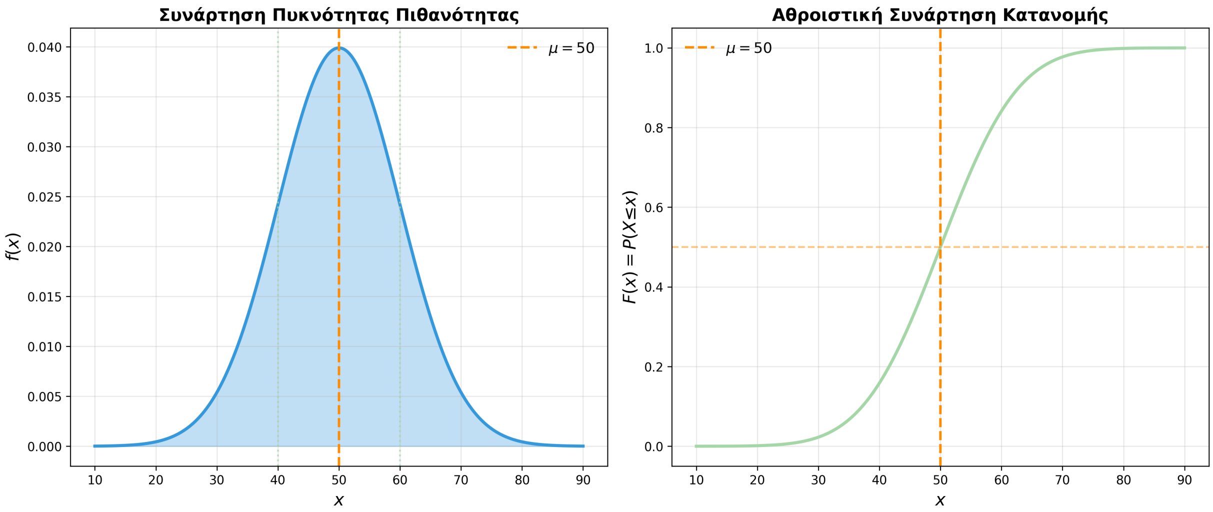
7. *Μέση Τιμή, Διάμεσος, Επικρατούσα τιμή:* Για την κανονική κατανομή ισχύει:

$$\text{Μέση Τιμή} = \text{Διάμεσος} = \text{Επικρατούσα τιμή} = \mu$$



Εικόνα 6.3: Συμμετρία της κανονικής κατανομής ως προς τη μέση τιμή

Η Εικόνα 6.3 επιδεικνύει την ιδιότητα της συμμετρίας της κανονικής κατανομής. Τα δύο χρωματισμένα εμβαδά, ένα αριστερά και ένα δεξιά της μέσης τιμής $\mu = 100$, απέχουν την ίδια απόσταση ($a = 20$) από το κέντρο. Λόγω της συμμετρίας, αυτά τα δύο εμβαδά είναι ίσα, δηλαδή $P(\mu - a \leq X \leq \mu) = P(\mu \leq X \leq \mu + a)$. Αυτή η ιδιότητα σημαίνει ότι η καμπύλη μπορεί να «*διπλωθεί*» κατά μήκος της κατακόρυφης γραμμής στο μ και τα δύο μισά θα ταιριάζουν απόλυτα. Η



Εικόνα 6.4: Συνάρτηση πυκνότητας πιθανότητας (PDF) και αθροιστική συνάρτηση κατανομής (CDF) για $N(50, 10^2)$

συμμετρία είναι μια θεμελιώδης ιδιότητα που διευκολύνει υπολογισμούς πιθανοτήτων και αποτελεί λόγο για τον οποίο η μέση τιμή, η διάμεσος και η επικρατούσα τιμή συμπίπτουν στην κανονική κατανομή.

Η Εικόνα 6.4 παρουσιάζει τόσο τη συνάρτηση πυκνότητας πιθανότητας (PDF) όσο και την αθροιστική συνάρτηση κατανομής (CDF) για την κατανομή $N(50, 10^2)$. Στο αριστερό γράφημα, η καμπανοειδής καμπύλη της PDF δείχνει τη σχετική πιθανότητα διαφορετικών τιμών, με το μέγιστο στη μέση τιμή $\mu = 50$. Το σκιασμένο εμβαδόν κάτω από την καμπύλη αντιπροσωπεύει πιθανότητες. Στο δεξί γράφημα, η CDF δείχνει την πιθανότητα $P(X \leq x)$ για κάθε τιμή x, αυξάνοντας μονοτονικά από 0 έως 1. Παρατηρούμε ότι στο $x = \mu = 50$ η CDF έχει τιμή 0.5, επιβεβαιώνοντας ότι η μέση τιμή διχοτομεί την κατανομή.

6.1.2 Στατιστικά Μέτρα

Για μια τυχαία μεταβλητή $X \sim N(\mu, \sigma^2)$:

- *Μέση τιμή*:
$$E(X) = \mu$$
- *Διασπορά*:
$$\text{Var}(X) = \sigma^2$$
- *Τυπική σπόκλιση*:
$$\text{SD}(X) = \sigma$$
- *Συντελεστής ασυμμετρίας*:
$$\text{Skewness} = 0$$
- *Κύρτωση*:
$$\text{Kurtosis} = 3$$