ΚΕΦΑΛΑΙΟ

8

ΚΑΤΑΝΟΜΕΣ

8.1 Εισαγωγή

Οι κατανομές διακρίνονται σε κατανομές συχνοτήτων, κατανομές πιθανοτήτων και σε δειγματοληπτικές κατανομές. Στη συνέχεια θα γίνει αναλυτική περιγραφή αυτών.

8.2 Κατανομές Συχνοτήτων (Frequency Distributions)

Με τον όρο **Κατανομή Συχνοτήτων** αναφερόμαστε σε έναν πίνακα στον οποίο αναγράφονται όλες οι τιμές της μεταβλητής και οι αντίστοιχες συχνότητες.

Πίνακας 8.1: Παράδειγμα Κατανομής Συχνοτήτων

Αριθμός παιδιών ανά οικογένεια X_i	**Πλήθος οικογενειών** F_i
0	50
1	250
2	360
3	120
4	80
5	40
ΣΥΝΟΛΟ	**900**

8.3 Κατανομές Πιθανοτήτων (Probability Distributions)

Με τον όρο **Κατανομή Πιθανοτήτων** αναφερόμαστε σε έναν πίνακα στον οποίο αναγράφονται όλες οι τιμές της μεταβλητής και οι αντίστοιχες πιθανότητες εμφάνισης αυτών των τιμών. Οι κατανομές πιθανοτήτων σχετίζονται με τις κατανομές συχνοτήτων. Στην πραγματικότητα μπορούμε να δούμε μια κατανομή πιθανοτήτων ως μια θεωρητική κατανομή συχνοτήτων. Τι σημαίνει αυτό; Η θεωρητική κατανομή συχνοτήτων είναι η κατανομή πιθανοτήτων η οποία περιγράφει τα αναμενόμενα αποτελέσματα.

Πίνακας 8.2: Παράδειγμα Κατανομής Πιθανοτήτων

Αριθμός παιδιών ανά οικογένεια X_i	Πιθανότητα P_i
0	0,056
1	0,278
2	0,400
3	0,133
4	0,089
5	0,044
ΣΥΝΟΛΟ	**1,000**

8.3.1 Πείραμα Τύχης (Experiment)

Πείραμα τύχης είναι κάθε διαδικασία η οποία παράγει καλώς καθορισμένα αποτελέσματα. Αυτό σημαίνει ότι σε κάθε απλή επανάληψη ένα και μόνο ένα πιθανό αποτέλεσμα θα εμφανιστεί.

8.3.2 Δειγματικός Χώρος Πειράματος Τύχης (Sample Space)

Το σύνολο των δυνατών αποτελεσμάτων ενός **Πειράματος Τύχης** (Experimental Outcomes) προσδιορίζει τον **Δειγματικό Χώρο του Πειράματος**.

8.3.3 Βασική Αρχή Απαρίθμησης (Counting Roules)

Αν ένα πείραμα τύχης μπορεί να περιγραφεί ως μια διαδικασία k βημάτων, για την οποία υπάρχουν n_1 δυνατά αποτελέσματα στο 1ο βήμα, n_2 δυνατά αποτελέσματα στο 2ο βήμα και n_k στο τελευταίο βήμα, τότε ο συνολικός αριθμός των δυνατών αποτελεσμάτων δίνεται από το γινόμενο $(n_1)\cdot(n_2)\cdot\ldots\ldots\cdot(n_k)$.

8.3.4 Τυχαία Μεταβλητή (Random Variable)

Τυχαία μεταβλητή είναι η αριθμητική περιγραφή των αποτελεσμάτων ενός πειράματος τύχης.

8.3.4.1 Ασυνεχής Τυχαία Μεταβλητή (Discrete Random Variable)

Ασυνεχής τυχαία μεταβλητή είναι αυτή που παίρνει ένα πεπερασμένο πλήθος τιμών ή ένα άπειρο πλήθος τιμών από μια ακολουθία τιμών, όπως 0,1,2,…

8.3.4.2 Συνεχής Τυχαία Μεταβλητή (Continous Random Variable)

Συνεχής τυχαία μεταβλητή είναι αυτή που παίρνει οποιαδήποτε τιμή ενός διαστήματος ή ενός συνόλου διαστημάτων.

8.3.5 Συνάρτηση Πιθανότητας (Probability Function)

Για μια τυχαία μεταβλητή η κατανομή πιθανοτήτων προσδιορίζεται από μια συνάρτηση πιθανότητας P(x), για την οποία ισχύουν:

- ✓ $0 \le P(x) \le 1$
- ✓ $\sum P(x) = 1$

8.3.6 Κατανομές Πιθανοτήτων Ασυνεχούς Τυχαίας Μεταβλητής (Probability Distributions of a Discrete Random Variable)

Η κατανομή πιθανοτήτων μιας διακριτής τυχαίας μεταβλητής είναι ένα γράφημα, ένας πίνακας ή ένας μαθηματικός τύπος που καθορίζει την πιθανότητα την οποία μπορεί να λάβει κάθε δυνατή τιμή της τυχαίας μεταβλητής.

Ορισμένες βασικές παράμετροι ασυνεχών κατανομών δίνονται στη συνέχεια.

Μέση Τιμή (Mean) ή **Αναμενόμενη Τιμή** (Expected Value) ή **Μαθηματική Ελπίδα** δίνεται από τη σχέση:

$E(x) = \mu = \sum x \cdot P(x)$ (1),

όπου P(x) η πιθανότητα εμφάνισης της τιμής χ.

Διακύμανση (Variance) δίνεται από τη σχέση:

$\sigma^2 = \sum \left[x^2 \cdot P(x) \right] - \mu^2$ (2) και

Τυπική Απόκλιση (Std. Deviation) είναι ίση με:

$\sigma = \sqrt{\sigma^2}$ (3).

Μεταξύ των ασυνεχών κατανομών αυτές που παρουσιάζουν μεγαλύτερο ενδιαφέρον είναι η Διωνυμική Κατανομή, η Κατανομή Poisson και η Υπεργεωμετρική Κατανομή.

Στη συνέχεια θα αναφερθούμε αναλυτικά στη Διωνυμική Κατανομή και στην Κατανομή Poisson.

8.3.6.1 Διωνυμική Κατανομή (Binomial Distribution)

Η **Διωνυμική Κατανομή** ή **Διαδικασία Bernoulli** (Bernoulli Prosess) είναι **ασυνεχής κατανομή** πιθανότητας η οποία έχει πολλές εφαρμογές σε πειράματα πολλών βημάτων. Τη συμβολίζουμε με **X~B(n,p)** και διαβάζουμε «***η μεταβλητή Χ ακολουθεί τη διωνυμική κατανομή με συντελεστές n και p***».

Ιδιότητες

- ✓ Το πείραμα αποτελείται από μια σειρά πανομοιότυπων δοκιμών.
- ✓ Σε κάθε δοκιμή δύο μόνο αποτελέσματα μπορεί να υπάρχουν. Το ένα το αναφέρουμε ως **επιτυχία** (success) και το άλλο ως **αποτυχία** (failure).
- ✓ Την πιθανότητα της επιτυχίας τη συμβολίζουμε με p και την πιθανότητα της αποτυχίας με q, είναι δε σταθερές σε όλη τη διάρκεια του πειράματος.
  Ισχύει **p+q=1** (4).
- ✓ Οι δοκιμές είναι ανεξάρτητες.
- ✓ Η τυχαία μεταβλητή Χ μετράει τον αριθμό των επιτυχιών και μπορεί να παίρνει τιμές από 0 έως n.

Συνάρτηση Πιθανότητας

$P(X = x) = \binom{n}{x} \cdot p^x \cdot q^{n-x}$ (5) με χ = 0,1,2,3,…..,n.

Η συνάρτηση αυτή δίνει την **πιθανότητα να έχουμε ακριβώς χ επιτυχίες σε n επαναλήψεις του πειράματος**.

- ✓ Το n εκφράζει τον αριθμό των επαναλήψεων του πειράματος.
- ✓ Το p εκφράζει την πιθανότητα επιτυχίας σε κάθε επανάληψη και το q την πιθανότητα αποτυχίας.
- ✓ Ο διωνυμικός συντελεστής $\binom{n}{x}$ δίνει τους συνδυασμούς των n ανά χ στοιχείων και είναι ίσος με $\binom{n}{x}=\frac{n!}{x!(n-x)!}$ (6).

Σημείωση

- ✓ Αν θέλουμε το πολύ χ επιτυχίες, δηλαδή $P(X \leq x)$, τότε εργαζόμαστε ως εξής:
  $P(X \leq x)=P(X=0)+P(X=1)+.....+P(X=x)$.
- ✓ Αν θέλουμε λιγότερες από χ επιτυχίες, δηλαδή $P(X \prec x)$, τότε εργαζόμαστε ως εξής:
  $P(X \prec x)=P(X=0)+P(X=1)+.....+P(X=x-1)$.
- ✓ Αν θέλουμε τουλάχιστον χ επιτυχίες, δηλαδή $P(X \geq x)$, τότε εργαζόμαστε ως εξής:
  $P(X \geq x)=1\text{-}P(X \prec x)=1\text{-}\left[P(X=0)+P(X=1)+.....+P(X=x-1)\right]$.
- ✓ Αν θέλουμε περισσότερες από χ επιτυχίες, δηλαδή $P(X \succ x)$, τότε εργαζόμαστε ως εξής:
  $P(X \succ x)=1\text{-}P(X \leq x)=1\text{-}\left[P(X=0)+P(X=1)+.....+P(X=x)\right]$.

!!! Όλα τα ανωτέρω ισχύουν με την προϋπόθεση ότι η τιμή του χ είναι πλησιέστερα στο 0 παρά στο n. Σε διαφορετική περίπτωση ακολουθούμε τη διαδικασία που θεωρούμε συντομότερη.

Παράμετροι Διωνυμικής Κατανομής
Μέση τιμή: $E(x)=\mu=np$ (7).
Διακύμανση: $Var(X)=\sigma^2=npq$ (8).
Τυπική απόκλιση: $\sigma=\sqrt{npq}$ (9).
Συντελεστής ασυμμετρίας: $S_k=\frac{(q-p)^2}{npq}$ (10).

- ✓ Αν p=q, τότε Sk=0 (συμμετρική κατανομή).
- ✓ Αν το n είναι πολύ μεγάλο (τείνει στο άπειρο), ο Sk τείνει στο 0 ανεξαρτήτως p και q.
- ✓ Όσο μεγαλύτερη διαφορά υπάρχει μεταξύ p και q τόσο πιο ασύμμετρη είναι η κατανομή.

Συντελεστής Κυρτότητας: $\beta=3+\frac{1-6pq}{npq}$ (11).

Παράδειγμα 1ο: Το πέντε τοις εκατό (5%) των οδηγών φορτηγών αυτοκινήτων στην Αμερική είναι γυναίκες. Υποθέτουμε ότι θα επιλέξουμε τυχαία 10 οδηγούς φορτηγών με σκοπό να πληροφορηθούμε σχετικά με τις συνθήκες εργασίας τους.

α. Είναι η επιλογή των 10 οδηγών ένα διωνυμικό πείραμα;
β. Ποια είναι η πιθανότητα 2 από τους οδηγούς να είναι γυναίκες;
γ. Ποια είναι η πιθανότητα κανένας από τους οδηγούς να μην είναι γυναίκα;
δ. Ποια είναι η πιθανότητα τουλάχιστον ένας από τους οδηγούς να είναι γυναίκα;

Λύση

α. Επειδή έχουμε δύο δυνατά αποτελέσματα (άνδρας-γυναίκα) με σταθερή πιθανότητα (5% και 95%), μπορούμε να πούμε ότι η μεταβλητή X, η οποία μετράει αριθμό γυναικών οδηγών, ακολουθεί τη διωνυμική κατανομή X~B(10, 0,05).

β. Ζητάμε $P(X=2)$. Με δεδομένο ότι p=0,05, q=0,95 και n=10, θα έχουμε από τον τύπο (5):

$$P(X=2)=\binom{10}{2}0,05^2\cdot 0,95^8=\frac{10!}{2!(10-2)!}0,0025\cdot 0,66342=45\cdot 0,0025\cdot 0,66342=7,46\%.$$

γ. Ζητάμε $P(X=0)$.

$$P(X=0)=\binom{10}{0}0,05^0\cdot 0,95^{10}=\frac{10!}{0!(10-0)!}1\cdot 0,5987=1\cdot 1\cdot 0,5987=59,87\%.$$

δ. Ζητάμε $P(X\geq 1)$.

$$P(X\geq 1)=1-P(X\prec 1)=1-\left[P(X=0)\right]=1-0,5987=0,4013=40,13\%.$$

Παράδειγμα 2ο: Ένα τοπικό κρεοπωλείο κατέγραψε το πλήθος των πελατών που πληρώνουν τοις μετρητοίς για μια περίοδο 10 ημερών. Τα αποτελέσματα εμφανίζονται στον επόμενο πίνακα:

Ημέρα	Πλήθος πελατών	Πληρωμή με μετρητά
1	58	28
2	64	38
3	150	82
4	120	75
5	210	84
6	84	51
7	90	63
8	124	92
9	218	165
10	102	54
ΣΥΝΟΛΟ	**1.220**	**732**

Από τα παραπάνω δεδομένα να υπολογίσετε την πιθανότητα ένας τυχαία επιλεγμένος πελάτης να πληρώσει τοις μετρητοίς. Χρησιμοποιώντας την τιμή της πιθανότητας που υπολογίσατε να προσδιορίσετε την πιθανότητα:

α. Τουλάχιστον 5 πελάτες, σε ένα δείγμα 10 πελατών, να πληρώσουν τοις μετρητοίς.
β. Λιγότεροι από 17 πελάτες να πληρώσουν τοις μετρητοίς σε ένα δείγμα 20 πελατών.
γ. Μεταξύ 30% και 60% των πελατών, σε ένα δείγμα 10 πελατών, να μην πληρώσουν τοις μετρητοίς.

8.3.6.2 Προσαρμογή Εμπειρικής σε Διωνυμική Κατανομή

Σε ένα τεστ, 2.000 άτομα κλήθηκαν να απαντήσουν σε ένα σετ 4 ερωτήσεων. Η μεταβλητή χ μετράει επιτυχημένες απαντήσεις και τα αποτελέσματα παρουσιάζονται στον επόμενο πίνακα:

Πίνακας 8.3

x_i	f_i	$f_i x_i$
0	396	0
1	790	790
2	592	1184
3	198	594
4	24	96
ΣΥΝΟΛΟ	**2.000**	**2.664**

Από τον τύπο του αριθμητικού μέσου βρίσκουμε: $\overline{x} = \dfrac{\sum f_i x_i}{\nu} = \dfrac{2.664}{2.000} = 1,332.$

Αυτό σημαίνει ότι κατά μέσο όρο σε κάθε σετ 4 ερωτήσεων απαντήθηκαν σωστά 1,332 ερωτήσεις. Στη συνέχεια από τον τύπο $E(x) = \mu = np$ με αντικατάσταση έχουμε:

$$1,332=4\cdot \mathrm{p} \Leftrightarrow p = \frac{1,332}{4} \Leftrightarrow p = 0,333,$$

που δηλώνει ότι η πιθανότητα να απαντήσει σωστά ένα τυχαία επιλεγμένο άτομο είναι 33,3%. Επομένως η πιθανότητα να απαντήσει λάθος θα είναι q=1-0,333=0,667=66,7%.

Με βάση λοιπόν τα p και q που υπολογίσαμε θα προσδιορίσουμε την πιθανότητα ο αριθμός των σωστών απαντήσεων να είναι 0, 1, 2, 3 και 4 αντίστοιχα, χρησιμοποιώντας τον τύπο της διωνυμικής κατανομής.

$$P(X=0) = \binom{4}{0} 0,333^0 0,667^4 = \frac{4!}{0!(4-0)!} \cdot 1 \cdot 0,19793 = 0,19793 = 19,79\%$$

$$P(X=1) = \binom{4}{1} 0,333^1 0,667^3 = \frac{4!}{1!(4-1)!} \cdot 0,333 \cdot 0,29674 = 4 \cdot 0,333 \cdot 0,29674 = 39,53\%$$

$$P(X=2) = \binom{4}{2} 0,333^2 0,667^2 = \frac{4!}{2!(4-2)!} \cdot 0,110889 \cdot 0,444889 = 6 \cdot 0,110889 \cdot 0,444889 = 29,59\%$$

$$P(X=3) = \binom{4}{3} 0,333^3 0,667^1 = \frac{4!}{3!(4-3)!} \cdot 0,037 \cdot 0,667 = 4 \cdot 0,037 \cdot 0,667 = 9,87\%$$

$$P(X=4) = \binom{4}{4} 0,333^4 0,667^0 = \frac{4!}{4!(4-4)!} \cdot 0,0123 \cdot 1 = 1 \cdot 0,0123 = 1,23\%$$

Στη συνέχεια από τον αρχικό πίνακα 8.3 θα έχουμε τις θεωρητικές συχνότητες.

Πίνακας 8.4

x_i	f_i	P(X=x)	Θεωρητικές συχνότητες $\widehat{f_i} = \left[P(X_i)\right] \cdot \sum f_i$
0	396	0,1979	395,8
1	790	0,3953	790,6
2	592	0,2959	591,8
3	198	0,0986	197,2
4	24	0,0123	24,6
ΣΥΝΟΛΟ	**2.000**	**1,0000**	**2.000,0**

Από τη σύγκριση μεταξύ εμπειρικών και θεωρητικών συχνοτήτων διαπιστώνουμε καλή προσαρμογή της εμπειρικής από τη διωνυμική κατανομή.

8.3.6.3 Κατανομή Poisson

Η κατανομή Poisson είναι μια **ασυνεχής κατανομή** εξίσου σημαντική με τη διωνυμική, η οποία είναι χρήσιμη στο να εκτιμηθούν διάφορα συμβάντα που έχουν λάβει χώρα σε ένα συγκεκριμένο διάστημα χρόνου ή χώρου. Για παράδειγμα η τυχαία μεταβλητή που μας ενδιαφέρει μπορεί να είναι ο αριθμός των αυτοκινήτων που επισκέφθηκαν ένα πλυντήριο αυτοκινήτων μέσα σε ένα διάστημα μιας ώρας ή ο αριθμός των επισκευών που χρειάστηκαν τα αυτοκίνητα σε μια διαδρομή 1.000 χιλιομέτρων. Τη συμβολίζουμε με **Χ~P(λ)** και διαβάζουμε «***η τυχαία μεταβλητή Χ ακολουθεί την κατανομή Poisson με συντελεστή λ***».

Ιδιότητες
- ✓ Η πιθανότητα να συμβεί ένα γεγονός είναι ίδια για κάθε δύο διαστήματα ίδιου πλάτους.
- ✓ Η πραγματοποίηση ή η μη πραγματοποίηση ενός γεγονότος σε ένα διάστημα είναι ανεξάρτητη από την πραγματοποίηση ή τη μη πραγματοποίηση σε κάθε άλλο διάστημα.

Συνάρτηση Πιθανότητας

$$P(X=x) = \lambda^x \cdot \frac{e^{-\lambda}}{x!} \quad (12).$$

Η συνάρτηση αυτή δίνει την πιθανότητα να πραγματοποιηθούν ακριβώς χ γεγονότα σε ένα διάστημα.
- ✓ Το λ εκφράζει τον μέσο όρο των πραγματοποιηθέντων γεγονότων σε ένα διάστημα.
- ✓ e=2,71828

Παράμετροι Κατανομής Poisson

Μέση τιμή: $E(x) = \lambda = np$ (13)

Διακύμανση: $Var(X)=\sigma^2=\lambda$ (14)

Τυπική απόκλιση: $\sigma = \sqrt{\lambda}$ (15)

Συντελεστής ασυμμετρίας: $S_k = \frac{1}{\lambda}$ (16)

Συντελεστής κυρτότητας: $\beta = 3 + \frac{1}{\lambda}$ (17)

Παράδειγμα: Ταξιδιώτες φτάνουν τυχαία και ανεξάρτητα στον σταθμό ελέγχου αποβιβάσεων ενός αεροδρομίου. Ο μέσος αριθμός αφίξεων είναι 10 ταξιδιώτες ανά λεπτό.

α. Ποια η πιθανότητα στη διάρκεια ενός λεπτού να μη φτάσει κανείς ταξιδιώτης;
β. Ποια η πιθανότητα να φτάσουν το πολύ 3 ταξιδιώτες στη διάρκεια ενός λεπτού;
γ. Ποια η πιθανότητα στη διάρκεια 15 δευτερολέπτων να μη φτάσει κανείς ταξιδιώτης;
δ. Ποια η πιθανότητα στη διάρκεια 15 δευτερολέπτων να φτάσει τουλάχιστον 1 ταξιδιώτης;

Λύση

α. Ζητάμε την πιθανότητα P(x=0). Με λ=10 και αντικατάσταση στον τύπο (12) θα έχουμε:

$$P(X=0)=10^0\frac{e^{-10}}{0!}=0,000045399=0,00454\%.$$

β. Ζητάμε την πιθανότητα P(X≤3).

$$P(X\leq 3)=P(X=0)+P(X=1)+P(X=2)+P(X=3)=10^0\frac{e^{-10}}{0!}+10^1\frac{e^{-10}}{1!}+10^2\frac{e^{-10}}{2!}+10^3\frac{e^{-10}}{3!}$$

$$=0,0000454+0,000454+0,00227+0,00757=0,0103394=1,03394\%.$$

γ. Πρέπει πρώτα να βρούμε τη νέα τιμή του λ. Γνωρίζοντας ότι στα 60΄ φτάνουν 10 ταξιδιώτες, είναι εύκολο να υπολογίσουμε ότι στα 15″ αντιστοιχούν (15*10)/60=2,5 ταξιδιώτες. Με λ=2,5 πλέον θέλουμε την πιθανότητα:

$$P(X=0)=2,5^0\frac{e^{-2,5}}{0!}=0,0821 \text{ ή } 8,21\%.$$

δ. Ζητάμε την πιθανότητα P(X≥1) με λ=2,5. Για τον υπολογισμό της πρέπει να κάνουμε τον εξής μετασχηματισμό:

$$\text{P(X}\geq\text{1)=1- P(X<1)=1- P(X=0)=1-}\,2,5^0\frac{e^{-2,5}}{0!}=-1-0,0821=0,9179 \text{ ή } 91,79\%.$$

8.3.6.4 Η Κατανομή Poisson ως Προσέγγιση της Διωνυμικής Κατανομής

Κάποιες φορές, αν επιθυμούμε να αποφύγουμε τη διαδικασία του υπολογισμού της πιθανότητας της διωνυμικής κατανομής, μπορούμε να χρησιμοποιήσουμε αντί αυτής, την κατανομή Poisson. Η κατανομή Poisson μπορεί να είναι λογική προσέγγιση της διωνυμικής κατανομής αλλά μόνο κάτω από συγκεκριμένες συνθήκες, όπως:

- ✓ Το **η** να είναι μεγάλο (η≥20).
- ✓ Το **p** να είναι πολύ μικρό (p≤0,05).

Αντικαθιστώντας στον τύπο της κατανομής Poisson την τιμή του λ με np από τη διωνυμική κατανομή, θα έχουμε την επόμενη σχέση:

$$P(X=x)=(np)^x\cdot\frac{e^{-(np)}}{x!} \quad (18).$$

Παράδειγμα: Μια έρευνα έδειξε ότι μόνο το 2% των επενδυτών θεωρεί το χρηματιστήριο σίγουρη επένδυση. Σε ένα δείγμα 100 επενδυτών ποια είναι η πιθανότητα ώστε:

1. Ακριβώς 2 επενδυτές να θεωρούν ότι το χρηματιστήριο είναι σίγουρη επένδυση;
2. Τουλάχιστον 2 επενδυτές να θεωρούν ότι το χρηματιστήριο είναι σίγουρη επένδυση;

Λύση

Πρόκειται για διωνυμική κατανομή.

1. Ζητάμε την P(X=2). Χρησιμοποιώντας τον τύπο της διωνυμικής κατανομής με p=0,02 q=0,98 και n=100 θα έχουμε:

$$P(X=2)=\binom{100}{2}\cdot 0{,}02^2\cdot 0{,}98^{98}=4950\cdot 0{,}0004\cdot 0{,}138=0{,}2734=27{,}34\%.$$

Επειδή ικανοποιούνται οι απαραίτητες συνθήκες, μπορούμε να χρησιμοποιήσουμε την κατανομή Poisson θέτοντας ως λ την τιμή λ=0,02*100=2.

$$P(X=2)=2^2\cdot\frac{e^{-2}}{2!}=0{,}27067=27{,}067\%.$$

2. Ζητάμε την

$$P(X\geq 2)=1-P(X\prec 2)=1-\left[P(X=0)+P(X=1)\right]=1-\left[\binom{100}{0}\cdot 0{,}02^0\cdot 0{,}98^{100}+\binom{100}{1}\cdot 0{,}02^1\cdot 0{,}98^{99}\right]$$

$$=1-\left[0{,}13262+0{,}27065\right]=1\text{-}0{,}40327=0{,}59673=59{,}673\%.$$

Με την κατανομή Poisson θα έχουμε αντίστοιχα:

$$P(X\geq 2)=1-P(X\prec 2)=1-\left[P(X=0)+P(X=1)\right]=1-\left[2^0\cdot\frac{e^{-2}}{0!}+2^1\cdot\frac{e^{-2}}{1!}\right]$$

$$=1\text{-}[0{,}1353+2\cdot 0{,}1353]=1\text{-}0{,}4059=0{,}5941=59{,}41\%.$$

8.3.7 Κατανομές Πιθανοτήτων Συνεχούς Τυχαίας Μεταβλητής (Probability Distributions of a Continous Random Variable)

Στις συνεχείς τυχαίες μεταβλητές το αντίστοιχο της συνάρτησης πιθανότητας είναι η **Συνάρτηση Πυκνότητας Πιθανότητας** (Probability Density Function) την οποία επίσης συμβολίζουμε με *f(x)*. Η συνάρτηση αυτή δεν παρέχει απευθείας τις ζητούμενες πιθανότητες, ωστόσο η περιοχή κάτω από το γράφημα της *f(x)*, η οποία αντιστοιχεί σε ένα δεδομένο διάστημα, μας παρέχει την πιθανότητα που παίρνει η συνεχής τυχαία μεταβλητή χ. Μια αδυναμία στον καθορισμό της πιθανότητας μιας τυχαίας συνεχούς μεταβλητής είναι ότι η πιθανότητα κάθε μεμονωμένης τιμής είναι μηδέν, γιατί το εμβαδόν της περιοχής που περικλείεται από την *f(x)* και κάθε μεμονωμένο σημείο είναι μηδέν.

Οι πλέον σημαντικές από τις συνεχείς κατανομές είναι η **Ομοιόμορφη** (Uniform), η **Κανονική** (Normal) και η **Εκθετική Κατανομή** (Exponential). Στη συνέχεια θα αναφερθούμε αναλυτικά στη σημαντικότερη όλων, λόγω του πλήθους των εφαρμογών της, την Κανονική Κατανομή.

8.3.7.1 Κανονική Κατανομή (Normal Distribution ή Gauss)

Είναι ίσως η σημαντικότερη κατανομή για να περιγράψει μια τυχαία **συνεχή μεταβλητή**. Χρησιμοποιείται σε μεγάλη ποικιλία πρακτικών εφαρμογών, στις οποίες η τυχαία μεταβλητή μπορεί να είναι το ύψος ή το βάρος των ανθρώπων, το επίπεδο του IQ, το ύψος της βροχόπτωσης και πολλά άλλα. Τη συμβολίζουμε με **X~N(μ, σ²)** και διαβάζουμε «***η τυχαία μεταβλητή X ακολουθεί την κανονική κατανομή με συντελεστές μ και σ²***».

Χαρακτηριστικά της Κανονικής Κατανομής

- ✓ Υπάρχει μια οικογένεια κανονικών κατανομών, αλλά η καθεμία από αυτές διαφοροποιείται από τον αριθμητικό της μέσο μ και την τυπική της απόκλιση σ.
- ✓ Το υψηλότερο σημείο της κανονικής κατανομής είναι ο αριθμητικός μέσος αυτής, που είναι επίσης

και διάμεσος και σημείο μέγιστης συχνότητας της κατανομής. Δηλαδή ισχύει η σχέση $\mu = M_{1/2} = M_0$.

- ✓ Ο αριθμητικός μέσος της κατανομής μπορεί να πάρει κάθε αριθμητική τιμή: θετική, αρνητική ή μηδέν.
- ✓ Η καμπύλη της κανονικής κατανομής είναι μονόκορφη σε μορφή καμπάνας, και συμμετρική, με άξονα συμμετρίας την ευθεία του αριθμητικού μέσου.
- ✓ Η καμπύλη είναι ασυμπτωτική (πλησιάζει τον άξονα XX΄, αλλά δεν τον ακουμπάει παρά μόνο θεωρητικά στο $\pm\infty$).
- ✓ Η τυπική απόκλιση καθορίζει το πλάτος της καμπύλης. Μεγάλη τυπική απόκλιση δημιουργεί καμπύλη πλατιά δείχνοντας και τη διασπορά των τιμών.
- ✓ Το συνολικό εμβαδόν της περιοχής κάτω από την καμπύλη της κανονικής κατανομής είναι 1 και κατανέμεται ισομερώς αριστερά και δεξιά του αριθμητικού μέσου (0,5 αριστερά και 0,5 δεξιά του μέσου).
- ✓ Οι πιθανότητες για την κανονική τυχαία μεταβλητή δίνονται από το εμβαδόν της περιοχής κάτω από την καμπύλη. Έτσι έχουμε:
  α. Το 68,26% του συνολικού εμβαδού βρίσκεται σε απόσταση ±σ από τον αριθμητικό μέσο.
  β. Το 95,44% του συνολικού εμβαδού βρίσκεται σε απόσταση ±2σ από τον αριθμητικό μέσο.
  γ. Το 99,72% του συνολικού εμβαδού βρίσκεται σε απόσταση ±3σ από τον αριθμητικό μέσο.

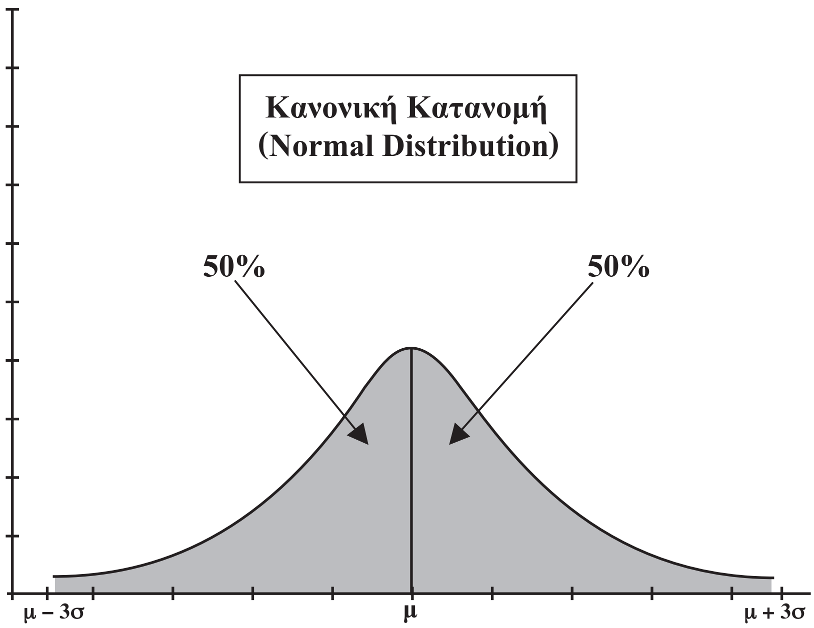
Συνάρτηση Πυκνότητας Πιθανότητας

$$f(x) = \frac{e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}}$$ (19), με π=3,14159 και e=2,71828.

Όταν θέλουμε να υπολογίσουμε την πιθανότητα για ένα διάστημα [a, b], χρησιμοποιούμε τον τύπο:

$$P(a \le x \le b) = \int_a^b f(x)\,dx$$ (20).

Σχήμα 8.1



8.3.7.2 Τυπική Κανονική Κατανομή

Όπως αντιλαμβάνεστε, ο υπολογισμός του ολοκληρώματος είναι συχνά πολύ δύσκολος και χρονοβόρος. Άλλες πάλι φορές δεν έχουμε τις απαιτούμενες μαθηματικές γνώσεις για τον υπολογισμό του. Για τον λόγο αυτό μετατρέπουμε τη **μεταβλητή X** σε **τυπική μεταβλητή Z** μέσω μιας διαδικασίας την οποία ονομάζουμε

Τυπικό Μετασχηματισμό και πλέον έχουμε μια άλλη κατανομή, γνωστή ως **Τυπική Κανονική Κατανομή**.

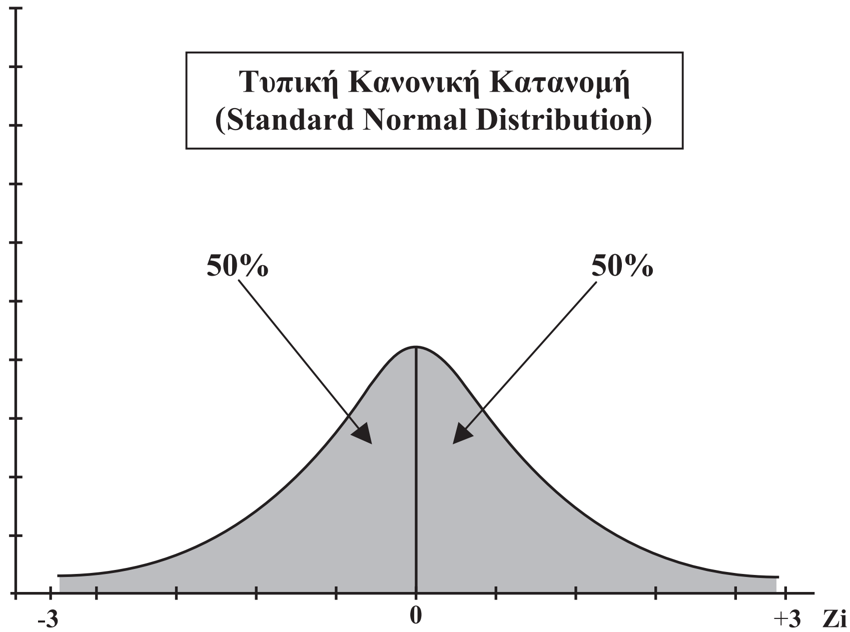
Η τυπική μεταβλητή Ζ δίνεται από τον τύπο $Z = \frac{X-\mu}{\sigma}$ (21), όπου:

- ✓ **Χ**: Η τιμή της οποίας την πιθανότητα θέλουμε να υπολογίσουμε
- ✓ **μ**: Η μέση τιμή
- ✓ **σ**: Η τυπική απόκλιση.

Η πιθανότητα για κάθε συγκεκριμένη τιμή του Ζ δίνεται στον πίνακα Α του παραρτήματος και μας απαλλάσσει από τον υπολογισμό του ολοκληρώματος του τύπου 19.

Η τυπική κατανομή έχει τις ίδιες ιδιότητες με την κανονική κατανομή, αλλά, ανεξαρτήτως αριθμητικού μέσου και τυπικής απόκλισης, αυτή έχει ως αριθμητικό μέσο το 0 και τυπική απόκλιση το 1. Έτσι λοιπόν συμβολίζουμε με **Z~N(0, 1)** και διαβάζουμε ***«η μεταβλητή Ζ ακολουθεί την τυπική κανονική κατανομή με μέση τιμή 0 και διακύμανση 1»***.

Σχήμα 8.2



Παράδειγμα 1ο: Έρευνα σε 10.000 εργαζόμενους του ιδιωτικού τομέα έδειξε ότι ο μέσος μηνιαίος μισθός τους ανέρχεται σε 1.500 €. Αν η τυπική απόκλιση είναι 100€, να υπολογιστεί:

1. Το πλήθος των εργαζομένων με μισθό το πολύ 1.600€.
2. Το πλήθος των εργαζομένων με μισθό μεγαλύτερο των 1.750€.
3. Το πλήθος των εργαζομένων με μισθό το πολύ 1.450€.
4. Το πλήθος των εργαζομένων με μισθό τουλάχιστον 1.400€.
5. Το πλήθος των εργαζομένων με μισθό από 1.450€ έως 1.650€.
6. Το πλήθος των εργαζομένων με μισθό από 1.650€ έως 1.700€.
7. Το πλήθος των εργαζομένων με μισθό από 1.350€ έως 1.400€.

Λύση

1. Για να υπολογίσουμε την πιθανότητα $P(X \le 1.600)$, είναι αναγκαίο να δημιουργήσουμε την τυπική μεταβλητή

$$Z = \frac{x-\mu}{\sigma} = \frac{1.600-1.500}{100} = 1.$$

Επομένως $P(X \le 1.600) = P(Z \le 1) = \Phi(1) = 0{,}5+0{,}3413=0{,}8413=84{,}13\%.$