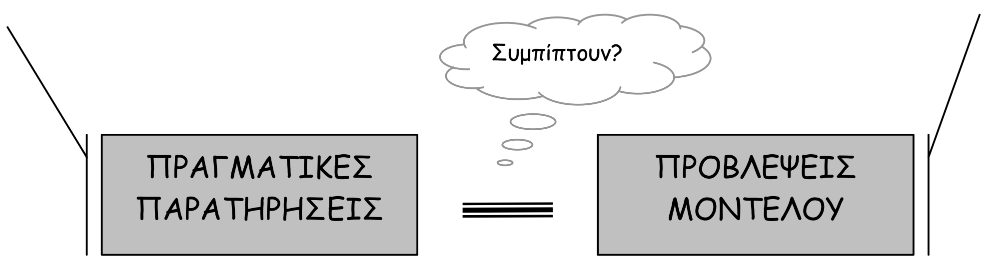
ΕΙΣΑΓΩΓΗ

1. ΜΑΘΗΜΑΤΙΚΑ ΜΟΝΤΕΛΑ

Για να μελετήσουμε ένα φαινόμενο που παρατηρούμε στη φύση συνήθως χτίζουμε ένα μαθηματικό μοντέλο που περιγράφει το φαινόμενο αυτό. Το μοντέλο οφείλει να απλοποιεί τα πράγματα και να αγνοεί τις ασήμαντες λεπτομέρειες.

Για να εξετάσουμε την εγκυρότητα του μοντέλου, μπορούμε να συγκρίνουμε τα αποτελέσματα που προβλέψαμε με βάση το μοντέλο με τις πραγματικές παρατηρήσεις του φαινομένου.



<u>ΠΑΡΑΔΕΙΓΜΑ 1</u>

Φαινόμενο: Στη φύση, αναπτύσσονται δυνάμεις μεταξύ των σωμάτων (βαρύτητα, κλπ)

Μοντέλο: Ο Newton υπολογίζει τη δύναμη μεταξύ δύο σωμάτων ως

$$F = g\frac{m_1 m_2}{r^2}$$

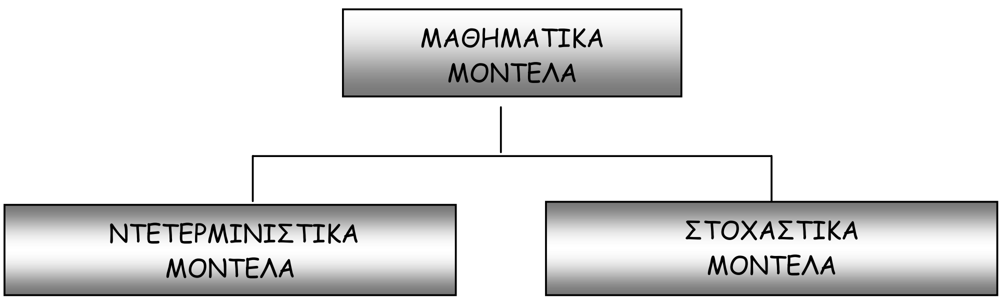
Εγκυρότητα: Μετά από πειράματα φαίνεται ότι ο νόμος του Newton περιγράφει αρκετά καλά την πραγματικότητα. Αργότερα, ο Αϊνστάιν θα δείξει ότι δεν ισχύει πάντοτε ο νόμος αυτός, μπορούμε ωστόσο να τον χρησιμοποιούμε καθώς είναι εύχρηστος και αποτελεί πολύ καλή προσέγγιση των πραγματικών φαινομένων.

ΠΑΡΑΔΕΙΓΜΑ 2

Οι νόμοι του Kepler αποτελούν ένα καλό μοντέλο για την περιγραφή της κίνησης των πλανητών.

---

2. ΣΤΟΧΑΣΤΙΚΑ ΜΟΝΤΕΛΑ



Στα **ντετερμινιστικά μοντέλα** οι συνθήκες ενός πειράματος καθορίζουν πλήρως τα αποτελέσματα.

Π.χ. στο νόμο του Newton που είδαμε προηγουμένως, εάν γνωρίζουμε τα μεγέθη m_1, m_2 και την απόσταση r, μπορούμε να υπολογίσουμε επακριβώς την δύναμη F.

Στα **στοχαστικά μοντέλα** οι συνθήκες ενός πειράματος τύχης καθορίζουν μόνο την "πιθανοτική" συμπεριφορά του αποτελέσματος.
Π.χ. εάν ρίξουμε ένα νόμισμα, δεν γνωρίζουμε το αποτέλεσμα, μπορούμε ωστόσο να περιγράψουμε το αποτέλεσμα με το εξής μοντέλο:
50% πιθανότητα να έρθει ΚΕΦΑΛΗ
50% πιθανότητα να έρθει ΓΡΑΜΜΑ

(Στο δεύτερο πείραμα θα μπορούσαμε να κατασκευάσουμε ένα κάλπικο νόμισμα με δύο κεφαλές για να κερδίζουμε τα στοιχήματα, οπότε το μοντέλο μας θα είναι ντετερμινιστικό!!!).

<u>ΠΑΡΑΔΕΙΓΜΑ 3</u> (στη σεισμολογία)

Εάν καταφέρουμε να ανακαλύψουμε ένα μοντέλο που υπολογίζει την εστία, το χρόνο και το μέγεθος ενός σεισμού θα πρόκειται για ένα ντετερμινιστικό μοντέλο (κάτι σαν και αυτό που ισχυρίζεται η ομάδα VAN).

Προς το παρόν, αρκούμαστε σε στοχαστικά μοντέλα που λαμβάνουν υπόψη στατιστικά στοιχεία του παρελθόντος, μελέτες του υπεδάφους και άλλες μετρήσεις και αποφαίνονται κάπως έτσι

> "υπάρχει μια πιθανότητα 35% να συμβεί ένας σεισμός στη Θεσσαλία
> μέσα στα επόμενα τρία χρόνια".

Μπορεί βέβαια να μη μας ικανοποιεί όσο ένα ντετερμινιστικό μοντέλο, έχει όμως μια αξία καθώς μας προειδοποιεί να πάρουμε προληπτικά μέτρα.

1. *ΒΑΣΙΚΑ ΣΤΟΙΧΕΙΑ ΣΥΝΔΥΑΣΤΙΚΗΣ*

1.1 ΑΠΑΡΙΘΜΗΣΗ

Θα δούμε αργότερα ότι για να βρούμε την πιθανότητα να συμβεί κάποιο γεγονός χρειάζεται πολλές φορές να μετράμε όλες τις δυνατές επιλογές που υπάρχουν για το γεγονός σε σχέση με το σύνολο των επιλογών που έχουμε στη διάθεσή μας. Σαν απλό παράδειγμα αναφέρουμε την ρίψη δύο ζαριών. Ρωτάμε πόσο πιθανό είναι να φέρουμε τουλάχιστον ένα εξάρι. Το σύνολο των επιλογών μας περιέχει 36 δυνατότητες.

1-1	1-2	1-3	1-4	1-5	**1-6**
2-1	2-2	2-3	2-4	2-5	**2-6**
3-1	3-2	3-3	3-4	3-5	**3-6**
4-1	4-2	4-3	4-4	4-5	**4-6**
5-1	5-2	5-3	5-4	5-5	**5-6**
6-1	**6-2**	**6-3**	**6-4**	**6-5**	**6-6**

Από αυτές οι «βολικές» περιπτώσεις που περιέχουν τουλάχιστον ένα εξάρι είναι 11, συγκεκριμένα αυτές που σημειώνονται έντονα στον παραπάνω πίνακα. Άρα λέμε ότι η ζητούμενη πιθανότητα είναι 11 στις 36 ή αλλιώς $^{11}\!/_{36}$.

Η απαρίθμηση των επιλογών σε ένα «πείραμα» δεν είναι πάντοτε εύκολη υπόθεση και χρειάζεται προσοχή. Ξεκινάμε με δύο απλούς κανόνες:

Ας υποθέσουμε ότι μια ενέργεια Α μπορεί να πραγματοποιηθεί με m τρόπους ενώ μια δεύτερη ενέργεια Β με n τρόπους. Με πόσους τρόπους μπορεί να γίνει είτε η μία είτε η άλλη ενέργεια; Προφανώς με

$m+n$ τρόπους

Η αρχή αυτή είναι γνωστή ως ***κανόνας του αθροίσματος*** και θα φανεί χρήσιμη όταν θα «τεμαχίζουμε» ένα πρόβλημα σε μικρότερα «επιμέρους» προβλήματα και θα αθροίζουμε τα αποτελέσματα.

Έστω ότι κάθε επιλογή για την ενέργεια Α μπορεί να συνδυαστεί με οποιαδήποτε επιλογή της ενέργειας Β. Τότε ο συνδυασμός των δύο ενεργειών μπορεί να πραγματοποιηθεί με

$$m \cdot n \quad \text{τρόπους}$$

Η αρχή αυτή είναι γνωστή ως ***κανόνας του γινομένου***. Μπορεί να φαίνεται τετριμμένη αλλά και θα φανεί εξαιρετικά χρήσιμη στη συνέχεια.

<u>ΠΑΡΑΔΕΙΓΜΑ 1</u>

Στις εκλογές ενός συμβουλίου, υπάρχουν 3 υποψήφιοι για τη θέση του προέδρου και 4 για τη θέση του γραμματέα. Με πόσους τρόπους μπορεί να καλυφθεί κάποια θέση (είτε η μία είτε ή άλλη); Με 3+4=7 τρόπους (κανόνας αθροίσματος)
Με πόσους τρόπους μπορούν να καλυφθούν και οι δύο θέσεις; Με 3x4=12 τρόπους (κανόνας γινομένου). Πράγματι,

(Π1,Γ1), (Π1,Γ2), (Π1,Γ3), (Π1,Γ4),
(Π2,Γ1), (Π2,Γ2), (Π2,Γ3), (Π2,Γ4),
(Π3,Γ1), (Π3,Γ2), (Π3,Γ3), (Π3,Γ4)

Πριν προχωρήσουμε στις βασικές περιπτώσεις απαριθμήσεων στη συνδυαστική, ας ξεκαθαρίσουμε κάποιους συμβολισμούς

Το παραγοντικό ορίζεται ως εξής

$$n! = 1 \cdot 2 \cdot 3 \cdots n$$

Δηλαδή 3!=1x2x3=6, καθώς και 4!=24, 1!=1. Συμφωνούμε επίσης ότι 0!=1

Παρατηρούμε ότι όταν έχουμε πηλίκο με παραγοντικά, γίνονται εύκολα απλοποιήσεις, πχ

$$\frac{7!}{5!} = \frac{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5} = \frac{5! \cdot 6 \cdot 7}{5!} = 6 \cdot 7$$

Ένα σύμβολο που θα χρειαστούμε συχνά είναι το $\binom{n}{r}$, το οποίο ορίζεται ως εξής

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}$$

Έτσι

$$\binom{7}{3} = \frac{7!}{3!\cdot 4!}$$

Αν απλοποιήσουμε με τον μεγαλύτερο παράγοντα στον παρονομαστή έχουμε

$$\binom{7}{3} = \frac{7!}{3!\cdot 4!} = \frac{5\cdot 6\cdot 7}{3!} = \frac{5\cdot 6\cdot 7}{1\cdot 2\cdot 3} = 35$$

Ομοίως

$$\binom{100}{99} = \frac{100!}{99!\cdot 1!} = \frac{1\cdot 2\cdots 99\cdot 100}{1\cdot 2\cdots 99} = 100$$

Τώρα είμαστε σε θέση να μελετήσουμε τα βασικά προβλήματα της απαρίθμησης

1.2 ΣΥΝΔΥΑΣΜΟΙ ΚΑΙ ΔΙΑΤΑΞΕΙΣ: ΕΠΙΛΟΓΕΣ r ΑΝΤΙΚΕΙΜΕΝΩΝ ΑΠΟ n

Το πρόβλημα που μελετάμε εδώ είναι

να επιλέξουμε r αντικείμενα από ένα σύνολο n αντικειμένων

Υπάρχουν όμως διάφοροι τρόποι να κάνουμε αυτή την επιλογή.

Αν ενδιαφερόμαστε για τη σειρά με την οποία εμφανίζονται τα επιλεγμένα αντικείμενα μιλάμε για ΔΙΑΤΑΞΕΙΣ. Αν δεν ενδιαφερόμαστε για τη σειρά με την οποία εμφανίζονται τα επιλεγμένα αντικείμενα μιλάμε για ΣΥΝΔΥΑΣΜΟΥΣ.

Υπάρχει και ένας άλλος διαχωρισμός.

Επιλέγουμε αρχικά ένα αντικείμενο. Πριν επιλέξουμε το δεύτερο, το αρχικό θα ξαναμπεί στην «κληρωτίδα» ή όχι; Έτσι μιλάμε για επιλογές ΜΕ ΕΠΑΝΑΛΗΨΗ και ΧΩΡΙΣ ΕΠΑΝΑΛΗΨΗ.

Ας τα δούμε αναλυτικά στη συνέχεια.

Α. <u>ΔΙΑΤΑΞΕΙΣ r αντικειμένων από n (παίζει ρόλο η σειρά)</u>

Έχουμε n αντικείμενα. Πόσοι τρόποι υπάρχουν να επιλέξουμε r αντικείμενα από αυτά και να τα βάλουμε σε μια σειρά; Η απάντηση συμβολίζεται *P(n,r)* και ισούται με

$$\frac{n!}{(n-r)!}$$

δηλαδή, με απλοποίηση

$$n(n-1)(n-2)\cdots(n-r+1)$$

Πράγματι, ας σκεφτούμε για παράδειγμα ότι από 10 άτομα θέλουμε να επιλέξουμε 4 για να μπουν με τη σειρά στις παρακάτω θέσεις

Παρατηρούμε ότι:

Για την 1η θέση έχουμε 10 επιλογές (ένα από τα 10 άτομα)

Για την 2η θέση έχουμε 9 επιλογές (ένα από τα 9 άτομα που περίσσεψαν)

Για την 3η θέση έχουμε 8 επιλογές

Για την 4η θέση έχουμε 7 επιλογές

Συνολικά έχουμε λοιπόν $10\cdot 9\cdot 8\cdot 7$ επιλογές, ή με άλλα λόγια $\frac{10!}{6!}$ όπως λέει και ο παραπάνω τύπος.

Β. <u>ΣΥΝΔΥΑΣΜΟΙ r αντικειμένων από n (δεν παίζει ρόλο η σειρά)</u>

Έχουμε n αντικείμενα. Πόσοι τρόποι υπάρχουν να επιλέξουμε μια ομάδα r αντικειμένων από αυτά; Η απάντηση συμβολίζεται *C(n,r)*, είτε με το σύμβολο που συναντήσαμε πιο πάνω

$$\binom{n}{r}$$

και διαβάζεται «n ανά r». Όπως είδαμε αυτό ισούται με

$$\frac{n!}{r!(n-r)!}$$

ΠΑΡΑΔΕΙΓΜΑ 2 Θέλουμε να επιλέξουμε 2 γράμματα από τα Α,Β,Γ,Δ,Ε. Υπάρχουν

$$\binom{5}{2} = \frac{5!}{2!3!} = \frac{4\cdot 5}{1\cdot 2} = 10 \text{ τρόποι}$$

(Πράγματι, πρόκειται για τα 10 ζευγάρια ΑΒ, ΑΓ, ΑΔ, ΑΕ, ΒΓ, ΒΔ, ΒΕ, ΓΔ, ΓΕ, ΔΕ. Προσέξτε ότι δεν λάβαμε υπόψη τη σειρά, δηλαδή θεωρήσαμε ότι ΑΒ και ΒΑ είναι ίδια)

Αξίζει να σημειωθεί ότι για να επιλέξουμε 3 γράμματα από τα 5 υπάρχουν επίσης

$$\binom{5}{2} = \frac{5!}{3!\cdot 2!} = 10 \text{ τρόποι.}$$

(ήταν αναμενόμενο καθώς, όταν επιλέγουμε 2 αντικείμενα από τα 5, μπορούμε να θεωρήσουμε ότι κάποιος άλλος επιλέγει τα υπόλοιπα 3 αντικείμενα, άρα όσοι τρόποι υπάρχουν για την επιλογή 2 αντικειμένων τόσοι ακριβώς τρόποι υπάρχουν και για την επιλογή 3 αντικειμένων).

Γενικά,

$$\binom{n}{r} = \binom{n}{n-r}$$

διότι και οι δύο αριθμοί ισούνται με

$$\frac{n!}{r!(n-r)!}.$$

Εύκολα επίσης διαπιστώνουμε ότι

$\binom{n}{0} = 1$, (υπάρχει 1 μόνο τρόπος να επιλέξουμε 0 αντικείμενα από τα n: να μην επιλέξουμε κανένα!)

$\binom{n}{n} = 1$, (υπάρχει 1 μόνο τρόπος να διαλέξουμε n αντικείμενα από τα n: να τα επιλέξουμε όλα!)

$\binom{n}{1} = n$, (υπάρχουν n τρόποι να επιλέξουμε 1 αντικείμενο από τα n)

$\binom{n}{n-1} = n$, (υπάρχουν n τρόποι να μην επιλέξουμε ένα αντικείμενο από τα n)

<u>ΠΑΡΑΔΕΙΓΜΑ 3</u> (χαρακτηριστικό) ΛΟΤΤΟ. Πόσοι τρόποι υπάρχουν να επιλέξουμε 6 νούμερα από το 1 ως το 49; (όπως είναι γνωστό, δεν χρειάζεται να τα πετύχουμε με τη σειρά που κληρώνονται). Υπάρχουν λοιπόν

$$\binom{49}{6} = \frac{49!}{6!43!} = \frac{44 \cdot 45 \cdot 46 \cdot 47 \cdot 48 \cdot 49}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6} = 13984816 \text{ δυνατότητες}$$

δηλαδή, περίπου 14 εκατομμύρια συνδυασμοί εξάδων. Εάν παίξουμε 14 στήλες έχουμε πιθανότητα μια στο εκατομμύριο να κερδίσουμε!

<u>ΠΑΡΑΔΕΙΓΜΑ 4</u> Στο Τζόκερ, επιλέγουμε 5 αριθμούς από μια ομάδα 45 αριθμών και ταυτόχρονα 1 αριθμό από μια ομάδα 20 αριθμών. Πόσες δυνατότητες υπάρχουν συνολικά; Εάν παίξουμε 24 στήλες ποια είναι η πιθανότητα να κερδίσουμε;

Υπάρχουν $\binom{45}{5}$ επιλογές για την πρώτη ομάδα και $\binom{20}{1}$ για τη δεύτερη. Άρα συνολικά υπάρχουν

$$\binom{45}{5} \cdot \binom{20}{1} = \frac{45!}{5!(40)!} \cdot 20 = \frac{41 \cdot 42 \cdot 43 \cdot 44 \cdot 45}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5} \cdot 20 = 24.435.180 \text{ δυνατότητες}$$

δηλαδή, περίπου 24,5 εκατομμύρια συνδυασμοί. Εάν παίξουμε 24 στήλες έχουμε πιθανότητα περίπου μια στο εκατομμύριο να κερδίσουμε!

Γ. <u>ΔΙΑΤΑΞΕΙΣ r αντικειμένων από n ΜΕ ΕΠΑΝΑΛΗΨΗ</u>

Δηλαδή κάθε φορά που επιλέγουμε ένα αντικείμενο, το ξαναβάζουμε στην «κληρωτίδα». Η απάντηση είναι

$$n^r$$

Σκεπτόμενοι όπως στην περίπτωση Α, ας πούμε ότι από 10 άτομα έχουμε να επιλέξουμε 4, αλλά αυτή τη φορά κάθε άτομο μπορεί να ξαναεπιλεγεί. Τα ονόματα τους θα τα γράψω σε μια σειρά

Για την 1η θέση έχουμε 10 επιλογές (ένα από τα 10 άτομα)
Για την 2η θέση έχουμε 10 επιλογές (αφού έχουμε ξανά και τα δέκα άτομα)
Για την 3η θέση έχουμε 10 επιλογές
Για την 4η θέση έχουμε 10 επιλογές

Συνολικά έχουμε λοιπόν $10\cdot10\cdot10\cdot10$ επιλογές, ή με άλλα λόγια 10^4 όπως λέει και ο παραπάνω τύπος.

Δ. ΣΥΝΔΥΑΣΜΟΙ r αντικειμένων από n ΜΕ ΕΠΑΝΑΛΗΨΗ

Δίνουμε απευθείας την απάντηση. Είναι

$$\binom{n+r-1}{r}$$

ΠΑΡΑΔΕΙΓΜΑ 5 Ρίχνουμε δύο ζάρια. Πόσες ζαριές υπάρχουν;

Σκεφτόμαστε ότι έχουμε n=6 αριθμούς, τους 1,2,3,4,5,6, και ρωτάμε πόσοι τρόποι υπάρχουν να επιλέξουμε r=2.

Έχουμε ΣΥΝΔΥΑΣΜΟΥΣ, διότι δεν μας ενδιαφέρει η σειρά. Π.χ η ζαριά 3-4 δεν είναι διαφορετική από την 4-3. Έχουμε ΕΠΑΝΑΛΗΨΗ, διότι π,χ η ζαριά 1-1 επιτρέπεται.

Υπάρχουν λοιπόν

$$\binom{6+2-1}{2}=\binom{7}{2}=\frac{6\cdot7}{1\cdot2}=21 \text{ ζαριές}$$

Συνοψίζοντας έχουμε,

	ΧΩΡΙΣ ΕΠΑΝΑΛΗΨΗ	ΜΕ ΕΠΑΝΑΛΗΨΗ
ΔΙΑΤΑΞΕΙΣ	$\frac{n!}{(n-r)!} = n(n-1)\cdots(n-r+1)$	n^r
ΣΥΝΔΥΑΣΜΟΙ	$\binom{n}{r}$	$\binom{n+r-1}{r}$

Παρουσιάζουμε δύο ακόμη ειδικές υποπεριπτώσεις του σκιασμένου κελιού: των διατάξεων χωρίς επανάληψη.

Α1. <u>Διατάξεις (ή μεταθέσεις) *n* αντικειμένων</u>

Με πόσους τρόπους μπορούμε να διατάξουμε στη σειρά *n* διαφορετικά αντικείμενα; (ουσιαστικά διατάσσουμε n αντικείμενα από n)

Για την πρώτη θέση έχουμε *n* επιλογές. Αφού διαλέξουμε το πρώτο αντικείμενο, για τη δεύτερη θέση έχουμε *n-1* επιλογές κ.ο.κ. Συνολικά έχουμε λοιπόν

$$n\cdot(n-1)\cdot(n-2)\cdots 2\cdot 1 = n! \text{ Επιλογές}$$

<u>ΠΑΡΑΔΕΙΓΜΑ 6</u>

Ο **Α**λέξης, ο **Β**ασίλης και ο **Γ**ιώργος μπορούν να μπουν σε μια διάταξη με 3!=6 τρόπους. Πράγματι, οι διατάξεις αυτές είναι

ΑΒΓ ΑΓΒ ΒΑΓ ΒΓΑ ΓΑΒ ΓΒΑ

Α2. <u>Διατάξεις n αντικειμένων όταν υπάρχουν ίδια αντικείμενα.</u>

Έστω ότι έχουμε n αντικείμενα:

- n_1 από τα οποία είναι ίδια, του 1^{ου} είδους
- n_2 από τα οποία είναι ίδια, του 2^{ου} είδους

  ...
- n_k από τα οποία είναι ίδια, του κ$^{\text{-στού}}$ είδους

(οπότε $n_1 + n_2 + \cdots + n_k = n$).

Το πλήθος των διατάξεων των n αντικειμένων δίνεται από τον τύπο

$$\frac{n!}{n_1!\cdot n_2!\cdots n_k!}$$

<u>ΠΑΡΑΔΕΙΓΜΑ 7</u>

α) Με πόσους τρόπους μπορούμε να διατάξουμε τα γράμματα Α,Β,Γ,Δ,Ε;

β) Με πόσους τρόπους μπορούμε να διατάξουμε τα γράμματα Α,Α,Α,Β,Β;

α) Έχουμε διάταξη 5 αντικειμένων, άρα υπάρχουν 5!=120 τρόποι.

β) Σύμφωνα με τον παραπάνω τύπο, έχουμε $n=6$, $n_1=3$, $n_2=2$ και υπάρχουν

$$\frac{5!}{3!\cdot 2!}=10 \text{ τρόποι}$$

Πράγματι, πρόκειται για τις διατάξεις

ΑΑΑΒΒ ΑΑΒΑΒ ΑΒΑΑΒ ΒΑΑΑΒ ΑΑΒΒΑ

ΑΒΑΒΑ ΒΑΑΒΑ ΑΒΒΑΑ ΒΑΒΑΑ ΒΒΑΑΑ

Σημείωση: Οι αριθμοί $\binom{n}{r}$ εμφανίζονται και στην ανάπτυξη του διωνύμου $(a+b)^n$:

$$(a+b)^n=\binom{n}{0}a^n b^0+\binom{n}{1}a^{n-1}b^1+\cdots+\binom{n}{r}a^{n-r}b^r+\cdots+\binom{n}{n}a^0 b^n$$

Π.χ.

$$(a+b)^2=\binom{2}{0}a^2b^0+\binom{2}{1}a^1b^1+\binom{2}{2}a^0b^2=a^2+2ab+b^2$$

$$(a+b)^3=\binom{3}{0}a^3b^0+\binom{3}{1}a^2b^1+\binom{3}{2}a^1b^2+\binom{3}{3}a^0b^3=a^3+3a^2b+3ab^2+b^3$$

1.3 ΠΡΟΒΛΗΜΑΤΑ ΣΥΝΔΥΑΣΤΙΚΗΣ

Σε ένα πρόβλημα συνδυαστικής, αρχικά προσπαθούμε να καταλάβουμε σε ποια περίπτωση εμπίπτει το πρόβλημά μας. Πολλές φορές είναι απαραίτητο να χωρίσουμε το πρόβλημά μας σε περιπτώσεις και να αθροίσουμε (με τον κανόνα του αθροίσματος) τα αποτελέσματα.

Επίσης, πολλές φορές υπάρχουν περιορισμοί που επιβάλουν μικρές τροποποιήσεις στο σκεπτικό της λύσης μας. Ας δούμε ορισμένα προβλήματα.

ΠΑΡΑΔΕΙΓΜΑ 8 Έχουμε τα 24 κεφαλαία γράμματα του ελληνικού αλφαβήτου. Πόσες λέξεις τριών γραμμάτων μπορούμε να σχηματίσουμε; (όχι απαραίτητα με νόημα!)

Έχουμε n=24 γράμματα και επιλέγουμε r=3 . Έχουμε

ΔΙΑΤΑΞΕΙΣ διότι σε μια «λέξη» παίζει ρόλο η σειρά των γραμμάτων
(άλλο ΣΟΙ και άλλο ΙΟΣ)
ΜΕ ΕΠΑΝΑΛΗΨΗ διότι το ίδιο γράμμα μπορεί να επαναληφθεί σε μια λέξη
(πχ στη λέξη ΑΡΑ)

Άρα σύμφωνα με το τυπολόγιο υπάρχουν 24^3 λέξεις.

Μπορούμε βέβαια να σκεφτούμε και με τον τρόπο που δουλέψαμε στην σχετική παράγραφο. Δηλαδή, για το πρώτο γράμμα έχουμε 24 επιλογές, για το δεύτερο 24 επιλογές, για το τρίτο 24 επιλογές, άρα συνολικά 24^3 επιλογές (συνδυασμούς).

ΠΑΡΑΔΕΙΓΜΑ 9 Έχουμε τα 10 ψηφία 0,1,2,3,4,5,6,7,8,9. Πόσους τριψήφιους αριθμούς μπορούμε να δημιουργήσουμε;

Δουλεύοντας όπως πιο πάνω θα λέγαμε 10^3 αριθμοί. Εδώ όμως υπάρχει ένας περιορισμός. Το πρώτο ψηφίο δεν μπορεί να είναι 0 εφόσον μιλάμε για τριψήφιους αριθμούς. Άρα έχουμε 9 επιλογές για τον πρώτο ψηφίο, 10 για το δεύτερο, 10 για το τρίτο, άρα συνολικά, 9x10x10=900 επιλογές.

Πολύ συχνά είναι πιο βολικό να υπολογίζουμε όχι ακριβώς τις περιπτώσεις που ρωτάει η άσκηση αλλά τις υπόλοιπες που εξαιρούνται.

ΠΑΡΑΔΕΙΓΜΑ 10 Στο Παράδειγμα 8 προηγουμένως, σε πόσες λέξεις (τριάδες) υπάρχουν επαναλήψεις γραμμάτων;

1ος τρόπος. Θα τις μετρήσουμε ευθέως αν και είναι πιο περίπλοκο. Αν μπερδευτείτε προχωρήστε απευθείας στον 2ο τρόπο.

Αν έχουμε επανάληψη στην 1η και 2η θέση, δηλαδή έχουμε τη μορφή XXY, υπάρχουν 24 κοινές επιλογές για τις θέσεις αυτές και απομένουν 23 επιλογές για την τρίτη θέση. Άρα υπάρχουν 24x23=552 επιλογές αυτής της μορφής.

Όμοια, υπάρχουν 552 τριάδες της μορφής XYX και 552 της μορφής YXX.

Επίσης υπάρχουν 24 τριάδες της μορφής XXX

Συνολικά λοιπόν υπάρχουν 552+552+552+24=1680 τριάδες με επανάληψη.

2ος τρόπος. Σκεφτόμαστε πιο πονηρά. Συνολικά είδαμε ότι υπάρχουν 24^3 τριάδες. Πόσες τριάδες από αυτές δεν έχουν επαναλήψεις; Δηλαδή πόσες ΔΙΑΤΑΞΕΙΣ, ΧΩΡΙΣ ΕΠΑΝΑΛΗΨΗ, r=3 γραμμάτων από n=24 υπάρχουν;

Ο συνοπτικός πίνακας λέει 24x23x22 = 12144

Άρα οι ζητούμενες περιπτώσεις είναι 24^3 -12144=1680

1.4 ΔΙΑΝΟΜΗ r ΑΝΤΙΚΕΙΜΕΝΩΝ ΣΕ n ΚΟΥΤΙΑ

Εδώ θα εξετάσουμε ένα διαφορετικό πρόβλημα συνδυαστικής. Δεν έχει να κάνει με επιλογή από ένα σύνολο αντικειμένων όπως μέχρι τώρα, αλλά με τοποθέτηση διαφόρων αντικειμένων σε ένα ορισμένο σύνολο από κουτιά.

Θέλουμε να τοποθετήσουμε r αντικείμενα μέσα σε n κουτιά

Υπάρχουν και εδώ περιπτώσεις. Μπορεί να αντικείμενα να είναι ΔΙΑΚΕΚΡΙΜΕΝΑ (πχ γράμματα, αριθμοί, φάκελοι που ξεχωρίζουν μεταξύ τους) ή ΜΗ ΔΙΑΚΕΚΡΙΜΕΝΑ (πχ κόκκινες μπάλες οι οποίες είναι όλες όμοιες). Επίσης μπορεί να ΠΑΙΖΕΙ ΡΟΛΟ Η ΣΕΙΡΑ των αντικειμένων όπως τοποθετούνται στα κουτιά είτε ΝΑ ΜΗΝ ΠΑΙΖΕΙ ΡΟΛΟ Η ΣΕΙΡΑ.

Δίνουμε απευθείας το τυπολόγιο

Διακεκριμένα αντικείμενα όπου δεν παίζει ρόλο η σειρά	n^r
Διακεκριμένα αντικείμενα όπου παίζει ρόλο η σειρά	$\frac{(n+r-1)!}{(n-1)!}$
Μη διακεκριμένα αντικείμενα	$\binom{n+r-1}{r}$

Ας δούμε την εφαρμογή τους σε απλά παραδείγματα.

<u>ΠΑΡΑΔΕΙΓΜΑ 11</u> Έχουμε 3 κουτιά και θέλουμε να τοποθετήσουμε μέσα 2 διαφορετικούς φακέλους (δεν παίζει ρόλο η σειρά με την οποία τοποθετούνται)

Σύμφωνα με τον πρώτο τύπο υπάρχουν $3^2 = 9$ τρόποι. Πράγματι, για να το δούμε στην πράξη, αν ονομάσουμε τους φακέλους Α,Β,Γ οι 9 τρόποι είναι:

Α και Β		
	Α και Β	
		Α και Β
Α	Β	
Α		Β
Β	Α	
Β		Α
	Α	Β
	Β	Β

ΠΑΡΑΔΕΙΓΜΑ 12 Έχουμε 3 κουτιά και θέλουμε να τοποθετήσουμε τα γράμματα Α και Β ενώ παίζει ρόλο η σειρά με την οποία τοποθετούνται σε ένα κουτί.

Σύμφωνα με τον δεύτερο τύπο υπάρχουν

$$\frac{(n+r-1)!}{(n-1)!}=\frac{4!}{2!}=3\cdot 4=12 \text{ τρόποι.}$$

Προσέξτε ότι εδώ είναι διαφορετική η τοποθέτηση Α,Β από την τοποθέτηση Β,Α σε ένα κουτί εφόσον παίζει ρόλο η σειρά. Πέρα λοιπόν από τις 9 παραπάνω περιπτώσεις έχουμε και τις περιπτώσεις

B-A		
	B-A	
		B-A

ΠΑΡΑΔΕΙΓΜΑ 13 Έχουμε 3 κουτιά και θέλουμε να τοποθετήσουμε 2 κόκκινες μπάλες.

Σύμφωνα με τον τρίτο τύπο υπάρχουν

$$\binom{n+r-1}{r}=\binom{4}{2}=\frac{4!}{2!\cdot 2!}=6 \text{ τρόποι.}$$

Πράγματι, οι περιπτώσεις είναι

K-K		
	K-K	
		K-K
K	K	
K		K
	K	K

Εδώ βεβαία είχαμε μικρό αριθμό κουτιών και αντικειμένων και η καταγραφή των περιπτώσεων ήταν εύκολη. Ας δούμε τα ίδια παραδείγματα και με λίγο μεγαλύτερα νούμερα όπου δεν είναι δυνατόν να περιγράψουμε ρητά τις περιπτώσεις και η συνδυαστική μας λύνει τα χέρια.

ΠΑΡΑΔΕΙΓΜΑ 14

α) Έχουμε 10 κουτιά και θέλουμε να τοποθετήσουμε μέσα 4 διαφορετικούς φακέλους:

υπάρχουν $10^4 = 10000$ τρόποι.

β) Έχουμε 10 κουτιά και θέλουμε να τοποθετήσουμε τους αριθμούς 1,2,3,4. Παίζει ρόλο η σειρά με την οποία τοποθετούνται:

υπάρχουν $\dfrac{(n+r-1)!}{(n-1)!} = \dfrac{13!}{9!} = 10 \cdot 11 \cdot 12 \cdot 13 = 17160$ τρόποι.

[Η διαφορά εδώ με το προηγούμενο παράδειγμα είναι ότι για να μπουν πχ οι 3 πρώτοι φάκελοι στο πρώτο κουτί υπάρχει 1 τρόπος, ενώ για να μπουν οι αριθμοί 1,2,3 στο πρώτο κουτί υπάρχουν αρκετοί τρόποι: 1-2-3, 1-3-2, 2-1-3, κλπ. Γι' αυτό έχουμε περισσότερους τρόπους στο δεύτερο παράδειγμα]

γ) Έχουμε 10 κουτιά και θέλουμε να τοποθετήσουμε 4 κόκκινες μπάλες:

υπάρχουν $\dbinom{n+r-1}{r} = \dbinom{13}{4} = \dfrac{13!}{4! \cdot 9!} = 715$ τρόποι.

ΑΣΚΗΣΕΙΣ

1.1 Στο παιχνίδι του ΟΠΑΠ «Extra 5» ο παίκτης επιλέγει 5 αριθμούς από 1 έως 35.
α) Πόσοι συνδυασμοί 5άδων μπορούν να σχηματιστούν; Άρα, αν παίξουμε μόνο μια στήλη (δηλ. μια 5άδα), ποια είναι η πιθανότητα να πετύχουμε 5άρι;
β) Εάν επιλέξουμε 8 αριθμούς, πόσες στήλες (δηλ. 5άδες) παίζουμε ουσιαστικά;

1.2 Μια πιτσαρία χρησιμοποιεί στην κατασκευή της πίτσας της μέχρι 9 διαφορετικά υλικά (μπορεί να μην περιέχει κανένα, να περιέχει μερικά ή ακόμη και τα 9 υλικά)
α) Πόσες πίτσες έχουν ακριβώς τρία είδη;
β) Πόσες πίτσες έχουν το πολύ τρία είδη;
γ) Πόσα είδη πίτσας προσφέρει;

1.3 Θεωρήστε όλες τις ελληνικές λέξεις των 5 γραμμάτων (με κεφαλαία και όχι απαραίτητα με νόημα) που μπορούν να σχηματιστούν. Απαντήστε στα παρακάτω ερωτήματα:
α) Πόσες είναι οι λέξεις αυτές;
β) Πόσες από τις παραπάνω λέξεις ξεκινούν από Α;
γ) Πόσες από τις παραπάνω λέξεις αρχίζουν και καταλήγουν στο ίδιο γράμμα;
δ) Πόσες από τις παραπάνω λέξεις έχουν όλα τα γράμματα διαφορετικά;
ε) Πόσες από τις παραπάνω λέξεις έχουν τουλάχιστον δύο ίδια γράμματα;

1.4 Απαντήστε στα παρακάτω ερωτήματα:
α) Με πόσους τρόπους μπορούμε να διατάξουμε τα γράμματα Α,Β,Γ,Δ,Ε,Ζ;
β) Με πόσους τρόπους μπορούμε να διατάξουμε τα γράμματα Α,Α,Α,Β,Γ,Γ;
γ) Με πόσους τρόπους μπορούμε να διατάξουμε τα γράμματα Α,Α,Α,Α,Α,Β και ποιοι είναι οι τρόποι αυτοί;

1.5 Έχουμε 12 άτομα και θέλουμε να τα χωρίσουμε σε δύο ομάδες των 5 και των 7 ατόμων αντίστοιχα. Με πόσους τρόπους μπορεί να γίνει αυτό
α) αν δεν υπάρχει κανένας άλλος περιορισμός;
β) αν δύο συγκεκριμένα άτομα δεν πρέπει να βρίσκονται στην ίδια ομάδα;

1.6 Ο αριθμός μιας πινακίδας αυτοκινήτου σχηματίζεται από τρία γράμματα μεταξύ των 14 που εμφανίζονται τόσο στο ελληνικό όσο και στο λατινικό αλφάβητο καθώς επίσης και από έναν τετραψήφιο αριθμό. Να υπολογίσετε

α) Πόσες πινακίδες αυτοκινήτων μπορούν να υπάρξουν;

β) Πόσες πινακίδες έχουν τρία κοινά γράμματα

γ) Πόσες πινακίδες έχουν τέσσερις ίδιους αριθμούς

δ) Πόσες πινακίδες έχουν τρία κοινά γράμματα και τέσσερις ίδιους αριθμούς

ε) Πόσες πινακίδες δεν περιέχουν το ψηφίο 0

1.7 Πόσες δυαδικές κωδικές λέξεις των 32 bits μπορούν να σχηματιστούν;
Στα δίκτυα, μια διεύθυνση IP αποτελείται από 32 δυαδικά bits. Διευθύνσεις που έχουν σαν πρώτο bit το 0 χαρακτηρίζονται ως διευθύνσεις κλάσης Α και αφιερώνουν τα 8 πρώτα bit για τη διεύθυνση του δικτύου και τα υπόλοιπα 24 bit για τη διεύθυνση του υπολογιστή. Έχουν δηλαδή τη μορφή

0 - - - - - - -	**- - - - - - - -**	**- - - - - - - -**	**- - - - - - - -**

Έτσι για παράδειγμα η IP διεύθυνση 00000111-0011111-0001000-0001110 (που με μορφή δεκαδικών ψηφίων γράφεται 7.31.8.14) αναφέρεται στο δίκτυο 7 με διεύθυνση υπολογιστή 31.8.14

Δεδομένου ότι τα 8 πρώτα ψηφία δε μπορεί να είναι 0000000 ή 0111111 (το πρώτο χρησιμοποιείται για εσωτερικές λειτουργίες του ίδιου του δικτύου ενώ το δεύτερο για broadcast σε όλα τα δίκτυα)

α) Πόσα δίκτυα μπορούν να εξυπηρετηθούν από διευθύνσεις κλάσης Α;

β) Πόσοι υπολογιστές μπορούν να εξυπηρετηθούν σε κάθε δίκτυο κλάσης Α;

γ) Πόσες είναι συνολικά οι IP διευθύνσεις υπολογιστών κλάσης Α;

1.8 Έχω τρεις φίλους, τον Αγησίλαο, το Βασίλη και το Γιάννη.

α) Διαθέτω 8 διαφορετικά δώρα. Με πόσους τρόπους μπορώ να τους τα μοιράσω; (ένας φίλος μπορεί να πάρει από κανένα μέχρι και τα 8 δώρα!)

β) Διαθέτω 8 καρτέλες με νούμερα: 1,2,3,4,5,6,7,8. Τα μοιράζω στους 3 φίλους μου ώστε να σχηματίσει ο καθένας έναν αριθμό (η να έχει κενό αριθμό)
Πχ δύο από τις δυνατές μοιρασιές είναι οι εξής

Α:352,	Β:71,	Γ: 648
Α:13427	Β:-	Γ:285

Πόσες τέτοιες μοιρασιές υπάρχουν;

γ) Διαθέτω 80 ευρώ σε χαρτονομίσματα των 10 ευρώ. Με πόσους τρόπους μπορώ να μοιράσω το ποσό αυτό στους φίλους μου;

δ) Ποια είναι η απάντηση σε καθεμιά από τις παραπάνω περιπτώσεις αν κάθε φίλος μου πρέπει να πάρει υποχρεωτικά τουλάχιστον ένα αντικείμενο;

ΣΥΝΤΟΜΕΣ ΑΠΑΝΤΗΣΕΙΣ ΤΩΝ ΑΣΚΗΣΕΩΝ

1.1 α) 324632 β) 56

1.2 α) 84 β) 130 γ) 512

1.3 α) $24^5 = 7962624$
β) $24^4 = 331776$
γ) $24^4 = 331776$
δ) P(24,5)=5100480
ε) 2862144

1.4 α) 720 β) 60 γ) 6

1.5 α) 792 β) 540

1.6 α) $14^3 \times 9000 = 24696000$
β) 126000
γ) 24696
δ) 126
ε) $14^3 \times 9^4$

1.7 α) 2^{32}
β) 126
γ) 2^{24}
δ) 126×2^{24}

1.8 α) 6561
β) 1814400
γ) 45
δ) $6561 - 3 \times 2^8 + 3 = 5796$,
P(8,3)x(7!/2!)=846720,
21

2. ΒΑΣΙΚΑ ΣΤΟΙΧΕΙΑ ΠΙΘΑΝΟΤΗΤΩΝ

2.1 ΣΥΝΟΛΑ

Δεχόμαστε σαν σύνολο μια συλλογή από αντικείμενα (δεν μπαίνουμε στη διαδικασία να το ορίσουμε αυστηρά γιατί δε χρειάζεται για το σκοπό μας). Συνήθως συμβολίζεται με ένα κεφαλαίο γράμμα, A,B κλπ. Μπορούμε να το περιγράψουμε με διάφορους τρόπους:

α) καταγράφοντας τα στοιχεία του: $A=\{1,5,8,13\}$
β) με λόγια: Το σύνολο B αποτελείται από όλες τις πόλεις της Ελλάδας
γ) με μια γενική περιγραφή: $C=\{x \mid 0<x<1\}$, δηλαδή το σύνολο όλων των αριθμών x με την ιδιότητα το x να βρίσκεται ανάμεσα στο 0 και το 1.

Υπενθυμίζουμε κάποιους βασικούς συμβολισμούς:

$a \in A$	: το α ανήκει στο σύνολο A
$A \subseteq B$	: το A είναι υποσύνολο του B, δηλαδή κάθε στοιχείο του A ανήκει και στο B
$A = B$	: το A και το B περιέχουν ακριβώς τα ίδια στοιχεία
$A \subset B$	: το A είναι γνήσιο υποσύνολο του B, δηλαδή ισχύει $A \subseteq B$ αλλά όχι $A = B$

Σημειώστε ότι $A = B$ ισοδυναμεί με ($A \subseteq B$ και $B \subseteq A$)

Επίσης μια πλάγια γραμμή πάνω στο σύμβολο της σχέσης ακυρώνει τη σχέση, π.χ. $a \notin A$ σημαίνει ότι το α δεν ανήκει στο σύνολο A. Όμοια και για τα υπόλοιπα σύμβολα.

Ορίζουμε επίσης τα σύνολα

Ø : το κενό σύνολο, το οποίο δεν περιέχει κανένα στοιχείο. Το θεωρούμε υποσύνολο κάθε συνόλου

$A \cup B$: Η **ένωση** των A και B που περιέχει τα στοιχεία που ανήκουν στο A **είτε** στο B (ή και στα δύο)

$A \cap B$: Η **τομή** των A και B που περιέχει τα στοιχεία που ανήκουν στο A **και** στο B **ταυτόχρονα**.

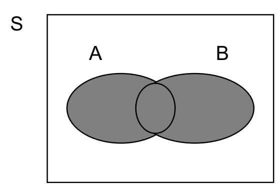
$A \setminus B$: Η **μερική διαφορά** του B από το A που περιέχει τα στοιχεία που ανήκουν στο A **αλλά όχι** στο B

Συνήθως, θεωρούμε ένα αρχικό σύνολο *S* που περιέχει όλα τα στοιχεία που μας ενδιαφέρουν και κατόπιν χρησιμοποιούμε διάφορα υποσύνολά του. Έτσι αν το *A* είναι υποσύνολο του *S*, ορίζουμε

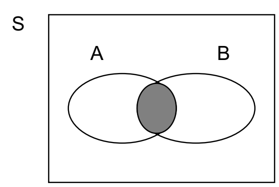
$\overline{A}$: Το **συμπλήρωμα** του *A* που περιέχει τα στοιχεία του *S* που δεν ανήκουν στο A. Δηλαδή

$$\overline{A} = S \setminus A$$

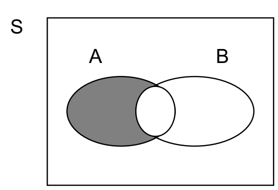
Τα νέα σύνολα που ορίσαμε περιγράφονται πιο παραστατικά (ως σκιασμένες περιοχές) με **διαγράμματα Venn:**



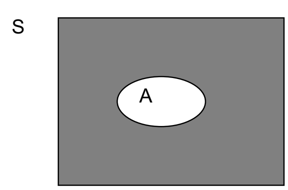
$A \cup B$



$A \cap B$



$A \setminus B$



$\overline{A}$

<u>ΠΑΡΑΔΕΙΓΜΑ 1</u>

Αν $S = \{1,2,3,4,5,6,7,8,9,10\}$ και $A = \{1,2,3,4\}$, $B = \{3,4,5,6,7\}$, τότε

- $A \cup B = \{1,2,3,4,5,6,7\}$
- $A \cap B = \{3,4\}$

- $A \setminus B = \{1,2\}$ και $B \setminus A = \{5,6,7\}$
- $\overline{A} = \{5,6,7,8,9,10\}$ και $\overline{B} = \{1,2,8,9,10\}$

ΠΑΡΑΔΕΙΓΜΑ 2

Χρησιμοποιώντας διαγράμματα Venn μπορείτε να δείξετε ότι

i) $(A \cup B) \cup C = A \cup (B \cup C)$

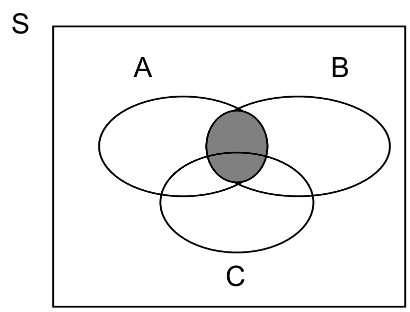
ii) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

iii) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

iv) $\overline{A \cup B} = \overline{A} \cap \overline{B}$

v) $\overline{A \cap B} = \overline{A} \cup \overline{B}$

Π.χ. για το iii), αν παρατηρήσουμε και τα δύο μέλη ξεχωριστά, μας δίνουν το ίδιο αποτέλεσμα:



2.2 ΔΕΙΓΜΑΤΟΧΩΡΟΣ ΚΑΙ ΕΝΔΕΧΟΜΕΝΑ

Όλα τα δυνατά αποτελέσματα ενός πειράματος τύχης θα αποτελούν το **δειγματοχώρο** μας, ενώ με τον όρο **ενδεχόμενο (**ή **γεγονός)** θα εννοούμε κάθε σύνολο που αποτελείται από ορισμένα δυνατά αποτελέσματα. Στη γλώσσα των συνόλων ο δειγματοχώρος θα είναι ένα αρχικό σύνολο S, και κάθε υποσύνολο του S θα ονομάζεται ενδεχόμενο. Αν το υποσύνολο περιέχει μόνο ένα στοιχείο του δειγματοχώρου θα ονομάζεται **απλό ενδεχόμενο** (ή **απλό γεγονός**). Σημειώνουμε ότι τόσο το κενό σύνολο Ø, όσο και το ίδιο το S αποτελούν ενδεχόμενα.

ΠΑΡΑΔΕΙΓΜΑ 3

Ρίχνουμε ένα ζάρι (αυτό είναι το πείραμα τύχης!) και παρατηρούμε τον αριθμό που φέρνουμε. Ο δειγματοχώρος είναι

$$S = \{1,2,3,4,5,6\}$$

Δίνουμε ορισμένα ενδεχόμενα:

- Να φέρουμε άρτιο αριθμό: $A_1 = \{2,4,6\}$.
- Να φέρουμε αριθμό μικρότερο ή ίσο του 2: $A_2 = \{1,2\}$
- $A_3 = \{1,4,5,6\}$
- $A_4 = \{3\}$ (πρόκειται για ένα απλό ενδεχόμενο)
- Ø
- $S = \{1,2,3,4,5,6\}$

ΠΑΡΑΔΕΙΓΜΑ 4

Ρίχνουμε ένα νόμισμα 4 φορές και παρατηρούμε το συνολικό αριθμό ΚΕΦΑΛΩΝ που φέρνουμε.

Εδώ,

$$S = \{0,1,2,3,4\}$$

Δίνουμε και δύο ενδεχόμενα:

- να μη φέρουμε καμία ΚΕΦΑΛΗ: $A_1 = \{0\}$,
- να φέρουμε δύο ή τρεις ΚΕΦΑΛΕΣ: $A_2 = \{2,3\}$

ΠΡΟΣΟΧΗ: Το ενδεχόμενο Ø είναι διαφορετικό από το ενδεχόμενο $\{0\}$. Το τελευταίο περιέχει ένα δυνατό αποτέλεσμα, είναι δηλαδή ένα απλό ενδεχόμενο. Το πρώτο δεν περιέχει κανένα δυνατό αποτέλεσμα.

ΠΑΡΑΔΕΙΓΜΑ 5

[Για να βρούμε το δειγματοχώρο πρέπει να έχουμε μια καθαρή εικόνα για το τι παρατηρούμε. Ας αλλάξουμε π.χ. ελαφρώς το προηγούμενο παράδειγμα]

Ρίχνουμε ένα νόμισμα 4 φορές και παρατηρούμε την σειρά ΚΕΦΑΛΩΝ και ΓΡΑΜΜΑΤΩΝ που φέρνουμε. Τώρα ο δειγματοχώρος είναι

$$\begin{aligned} S = \{ & \text{ΚΚΚΚ}, \text{ΚΚΚΓ}, \text{ΚΚΓΚ}, \text{ΚΚΓΓ}, \text{ΚΓΚΚ}, \text{ΚΓΚΓ}, \text{ΚΓΓΚ}, \text{ΚΓΓΓ}, \\ & \text{ΓΚΚΚ}, \text{ΓΚΚΓ}, \text{ΓΚΓΚ}, \text{ΓΚΓΓ}, \text{ΓΓΚΚ}, \text{ΓΓΚΓ}, \text{ΓΓΓΚ}, \text{ΓΓΓΓ}\}. \end{aligned}$$

Μπορούμε λοιπόν να εκφράσουμε το ενδεχόμενο να φέρουμε περισσότερες ΚΕΦΑΛΕΣ από ΓΡΑΜΜΑΤΑ με το υποσύνολο του δειγματοχώρου:

$$A = \{\text{ΚΚΚΚ}, \text{ΚΚΚΓ}, \text{ΚΚΓΚ}, \text{ΚΓΚΚ}, \text{ΓΚΚΚ}\}$$

<u>ΠΑΡΑΔΕΙΓΜΑ 6</u>

[Ο δειγματοχώρος δεν είναι πάντοτε πεπερασμένος. Μπορεί κάλλιστα να είναι ένα άπειρο σύνολο]

Μια μηχανή κατασκευάζει ένα συγκεκριμένο προϊόν. Κάποια προϊόντα βγαίνουν ελαττωματικά. Η μηχανή συνεχίζει να κατασκευάζει ωσότου συγκεντρωθούν δέκα μη ελαττωματικά προϊόντα. Πόσα προϊόντα είναι δυνατό να κατασκευαστούν συνολικά; Προφανώς πρέπει να κατασκευαστούν τουλάχιστον δέκα προϊόντα. Ο δειγματοχώρος είναι

$$S = \{10, 11, 12, 13, 14, \ldots\}$$

Έστω S ο δειγματοχώρος μας και A, B δύο ενδεχόμενα, δηλαδή $A, B \subseteq S$.

Μπορούμε με τις πράξεις των συνόλων που αναφέραμε νωρίτερα να ορίσουμε τα εξής νέα ενδεχόμενα:

- $A \cup B$: να συμβεί το ενδεχόμενο A **ή** το ενδεχόμενο B
- $A \cap B$: να συμβούν τα ενδεχόμενα Α και Β ταυτόχρονα.
- $A \setminus B$: να συμβεί το ενδεχόμενο Α αλλά όχι το ενδεχόμενο Β.
- $\overline{A}$: να μη συμβεί το ενδεχόμενο Α

Δύο ενδεχόμενα Α και Β θα λέγονται ξένα μεταξύ τους αν δεν μπορούν να συμβούν ταυτόχρονα, δηλαδή αν

$$A \cap B = \emptyset$$

2.3 Η ΠΙΘΑΝΟΤΗΤΑ ΕΝΟΣ ΕΝΔΕΧΟΜΕΝΟΥ

Στην προσπάθειά μας να ορίσουμε την έννοια της πιθανότητας θα χρησιμοποιήσουμε σαν πείραμα το παράδειγμα του ζαριού. Έστω λοιπόν ότι ρίχνουμε ένα ζάρι και με Α συμβολίζουμε το ενδεχόμενο να φέρουμε την ένδειξη 1, δηλαδή

$$S = \{1,2,3,4,5,6\}$$

και

$$A = \{1\}$$

Διαισθητικά καταλαβαίνουμε ότι υπάρχει μια πιθανότητα 1 στις 6 να συμβεί το ενδεχόμενο Α. Αν επαναλάβουμε το πείραμα αρκετές φορές αναμένουμε περίπου στο 1/6 των επαναλήψεων να «πετύχουμε» 1. Όσο αυξάνουμε τον αριθμό των επαναλήψεων τόσο πιο κοντά στο 1/6 θα βρίσκεται η σχετική συχνότητα

$$p_A = \frac{\text{αριθμός επαναλήψεων ενδεχομένου Α}}{\text{συνολικός αριθμός επαναλήψεων}}$$

Για σχετική συχνότητα παρατηρούμε γενικά ότι

1. $0 \leq p_A \leq 1$
2. για το ενδεχόμενο S, το οποίο συμβαίνει πάντοτε, είναι $p_S = 1$.
3. για το ενδεχόμενο $\emptyset$, το οποίο δεν συμβαίνει ποτέ, $p_\emptyset = 0$.
4. αν τα Α και Β είναι ξένα μεταξύ τους, τότε $p_{A \cup B} = p_A + p_B$

Τέλος, όταν ο αριθμός των επαναλήψεων πλησιάζει στο άπειρο, η σχετική συχνότητα p_A πλησιάζει σε έναν συγκεκριμένο αριθμό $P(A)$, ο οποίος θα αποτελεί την πιθανότητα του Α. Ο εμπειρικός αυτός ορισμός της πιθανότητας οδηγεί στον παρακάτω αυστηρότερο ορισμό.

<u>ΟΡΙΣΜΟΣ</u> Έστω S ο δειγματοχώρος ενός πειράματος. Σε κάθε ενδεχόμενο Α αντιστοιχίζουμε έναν αριθμό $P(A)$, που τον ονομάζουμε **πιθανότητα του Α**, με τις εξής ιδιότητες:

1. $0 \le P(A) \le 1$
2. $P(S) = 1$
3. $P(\emptyset) = 0$
4. Αν $A \cap B = \emptyset$, τότε $P(A \cup B) = P(A) + P(B)$

[Κανονικά η ιδιότητα 3 είναι περιττή καθώς προκύπτει από τις ιδιότητες 2 και 4 για $A = S$ και $B = \emptyset$]

Άλλες ιδιότητες που προκύπτουν από τον ορισμό είναι οι εξής:

- $P(\overline{A}) = 1 - P(A)$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, για οποιαδήποτε ενδεχόμενα A και B
- $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(C \cap A)$
  $+ P(A \cap B \cap C)$
- Αν $A \subseteq B$ τότε $P(A) \le P(B)$.

2.4 ΠΕΠΕΡΑΣΜΕΝΟΙ ΔΕΙΓΜΑΤΟΧΩΡΟΙ

Έστω $S = \{a_1, a_2 \ldots, a_n\}$ ένας πεπερασμένος δειγματοχώρος και ότι τα απλά ενδεχόμενα

$$\{a_1\}\ ,\ \{a_2\}\ ,\ \ldots\ ,\ \{a_n\}$$

έχουν αντίστοιχες πιθανότητες

$$p_1\ ,\ p_2\ ,\ \cdots\ ,\ p_n$$

Ισχύουν

α) $p_i \ge 0$ για κάθε $i = 1,2,\ldots,n$

β) $p_1 + p_2 + \cdots + p_n = 1$

Επίσης, για ένα ενδεχόμενο A, η αντίστοιχη πιθανότητα $P(A)$ βρίσκεται εύκολα αθροίζοντας τις επιμέρους πιθανότητες, π.χ. αν $A = \{a_1, a_4, a_5\}$, τότε

$$P(A) = p_1 + p_4 + p_5$$

ΠΑΡΑΔΕΙΓΜΑ 7

Δίνεται ο δειγματοχώρος $S = \{a_1, a_2, a_3\}$ και οι εξής προϋποθέσεις:

Το a_1 έχει διπλάσια πιθανότητα να συμβεί απ' ότι το a_2

Το a_2 έχει διπλάσια πιθανότητα να συμβεί απ' ότι το a_3

Να βρεθεί η πιθανότητα του ενδεχομένου $A = \{a_1, a_2\}$.

Οι προϋποθέσεις μας δίνουν ότι $p_1 = 2p_2$ και $p_2 = 2p_3$ (Άρα, $p_1 = 4p_3$)

Όμως,

$$p_1 + p_2 + p_3 = 1$$

οπότε

$$4p_3 + 2p_3 + p_3 = 1$$

$$7p_3 = 1$$

$$p_3 = 1/7$$

Κατά συνέπεια, $p_1 = 4/7$, $p_2 = 2/7$ και τελικά

$$P(A) = 4/7 + 2/7 = 6/7 .$$

Συνήθως όλα τα δυνατά αποτελέσματα του δειγματοχώρου έχουν την ίδια πιθανότητα να συμβούν. Έτσι αν

$$S = \{a_1, a_2 \ldots, a_n\}$$

έχουμε

$$p_1 = p_2 = \cdots = p_n = \frac{1}{n}$$

Σε μια τέτοια περίπτωση, αν το ενδεχόμενο A που μελετάμε περιέχει r δυνατά αποτελέσματα, ισχύει

$$P(A) = \frac{r}{n} = \frac{\text{πλήθος ζητούμενων αποτελεσμάτων}}{\text{πλήθος δυνατών αποτελεσμάτων}}$$

ΠΑΡΑΔΕΙΓΜΑ 8

Ρίχνουμε ένα ζάρι. Όλα τα δυνατά αποτελέσματα έχουν την ίδια πιθανότητα $1/6$.

Το ενδεχόμενο

A= «να φέρουμε πάνω από 4» = *{5,6}*

έχει πιθανότητα

$$P(A) = \frac{2}{6} = \frac{1}{3}.$$

ΠΑΡΑΔΕΙΓΜΑ 9

Ρίχνουμε ένα νόμισμα δύο φορές και ζητάμε την πιθανότητα να φέρουμε μόνο μια ΚΕΦΑΛΗ.

Χρειάζεται προσοχή στον καθορισμό του δειγματοχώρου και των επιμέρους πιθανοτήτων. Εάν μετράμε τον αριθμό των ΚΕΦΑΛΩΝ που μπορούμε να φέρουμε στις δύο ρίψεις, ο δειγματοχώρος είναι

$$S = \{0,1,2\}$$

Θα ήταν λάθος να θεωρήσουμε ότι και τα τρία δυνατά αποτελέσματα έχουν την ίδια πιθανότητα, δηλαδή $1/3$ καθώς

το 0 λαμβάνεται με έναν τρόπο (ΓΡΑΜΜΑ-ΓΡΑΜΜΑ)

το 1 λαμβάνεται με δύο τρόπους (ΚΕΦΑΛΗ-ΓΡΑΜΜΑ ή ΓΡΑΜΜΑ-ΚΕΦΑΛΗ)

το 2 λαμβάνεται με έναν τρόπο (ΚΕΦΑΛΗ-ΚΕΦΑΛΗ).

Έτσι,

$$P(0) = \frac{1}{4} \qquad P(1) = \frac{2}{4} \qquad P(2) = \frac{1}{4}$$

Η ζητούμενη πιθανότητα είναι λοιπόν $P(1) = \frac{2}{4} = \frac{1}{2}$.

Ίσως θα ήταν καλύτερα να θεωρήσουμε σαν δειγματοχώρο το

$$S' = \{\mathrm{KK},\ \mathrm{K\Gamma},\ \mathrm{\Gamma K},\ \mathrm{\Gamma\Gamma}\}$$

Το ζητούμενο ενδεχόμενο είναι το

$$A = \{\mathrm{K\Gamma},\ \mathrm{\Gamma K}\}$$

με πιθανότητα

$$P(A) = \frac{2}{4} = \frac{1}{2}$$

2.5 ΔΕΣΜΕΥΜΕΝΗ ΠΙΘΑΝΟΤΗΤΑ-ΑΝΕΞΑΡΤΗΤΑ ΕΝΔΕΧΟΜΕΝΑ

Η πιθανότητα ενός ενδεχομένου εξαρτάται πολλές φορές από κάποιο άλλο ενδεχόμενο που επηρεάζει το τυχαίο πείραμά μας. Έτσι για παράδειγμα, η πιθανότητα να βρέξει έτσι απλά, από την πιθανότητα να βρέξει όταν γνωρίζουμε ότι υπάρχει συννεφιά είναι διαφορετική: προφανώς στη δεύτερη περίπτωση η πιθανότητα είναι μεγαλύτερη. Ας δούμε ένα πιο αριθμητικό παράδειγμα.

Έστω ότι από μια τράπουλα με 52 χαρτιά τραβάμε ένα φύλο και κερδίζουμε αν είναι ΚΟΚΚΙΝΟΣ ΑΣΣΟΣ. Η πιθανότητα να κερδίσουμε είναι 2/52 (διότι υπάρχουν δύο κόκκινοι άσσοι: καρό και κούπα!), δηλαδή τελικά 1/26. Εάν όμως κάποιος μας «σφυρίξει» ότι το φύλλο είναι ΚΟΥΠΑ η πιθανότητα να κερδίσουμε αλλάζει. Υπάρχουν 13 ΚΟΥΠΕΣ και από αυτές κερδίζει ο ένας ΑΣΣΟΣ οπότε η πιθανότητα να κερδίσουμε είναι 1/13.

Η δεσμευμένη πιθανότητα έρχεται να εκφράσει αυτή ακριβώς την περίπτωση: την πιθανότητα ενός ενδεχομένου ενώ γνωρίζουμε κάποιο άλλο ενδεχόμενο.

Ορίζουμε

P(B/A) = «η πιθανότητα να συμβεί το ενδεχόμενο Β δεδομένου ότι έχει συμβεί το ενδεχόμενο Α»

Η **δεσμευμένη πιθανότητα** (αλλού θα τη βρείτε ως πιθανότητα υπό συνθήκη) δίνεται από τον τύπο

$$P(B/A) = \frac{P(A \cap B)}{P(A)} \qquad (*)$$

με την προϋπόθεση βέβαια ότι *P(A) > 0.*

<u>ΠΑΡΑΔΕΙΓΜΑ 10</u>

Στο παράδειγμα της τράπουλας που περιγράψαμε πιο πάνω ζητάμε την πιθανότητα να πετύχουμε ΚΟΚΚΙΝΟ ΑΣΣΟ δεδομένου ότι έχει τραβηχτεί ΚΟΥΠΑ.

Θέτουμε

A = «έχει τραβηχτεί ΚΟΥΠΑ»

B = «έχει τραβηχτεί ΚΟΚΚΙΝΟΣ ΑΣΣΟΣ»

και ουσιαστικά ζητάμε την πιθανότητα *P(B/A)*.

Προσέξτε ότι $A \cap B$ = «έχει τραβηχτεί ΑΣΣΟΣ ΚΟΥΠΑ»
Συνεπώς, σύμφωνα με τον τύπο (*),

$$P(B/A) = \frac{P(A \cap B)}{P(A)} = \frac{1/52}{13/52} = \frac{1}{13}$$

Βέβαια, ακολουθώντας τον κλασικό τρόπο, θα μπορούσε να θεωρήσει κανείς σαν δειγματοχώρο μόνο τις ΚΟΥΠΕΣ που είναι 13, ενώ τα φύλλα που κερδίζουν είναι μόνο 1, οπότε

$$P(B/A) = \frac{\text{πλήθος ζητούμενων φύλλων}}{\text{πλήθος από ΚΟΥΠΕΣ}} = \frac{1}{13}$$

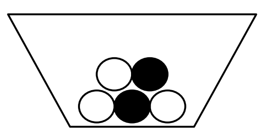
Ο τύπος (*) είναι πιο χρήσιμος στη μορφή

$$P(A \cap B) = P(A)P(B/A) \qquad (**)$$

Μας βοηθάει να βρούμε την πιθανότητα να συμβούν δύο ενδεχόμενα *A* και *B* ταυτόχρονα, όταν το ένα ενδεχόμενο εξαρτάται άμεσα από το άλλο.

ΠΑΡΑΔΕΙΓΜΑ 11

Ένα κουτί περιέχει 3 ΑΣΠΡΕΣ και 2 ΜΑΥΡΕΣ μπάλες. Αν τραβήξουμε δύο μπάλες τη μία μετά την άλλη, ποια είναι η πιθανότητα να είναι και οι δύο ΜΑΥΡΕΣ;



Η πρώτη και σημαντικότερη δουλειά σε τέτοια προβλήματα είναι να καθορίσουμε τα κατάλληλα ενδεχόμενα. Θέτουμε λοιπόν,

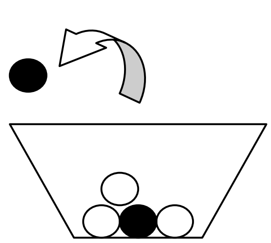
Α= «η πρώτη μπάλα είναι ΜΑΥΡΗ»
Β= «η δεύτερη μπάλα είναι ΜΑΥΡΗ»

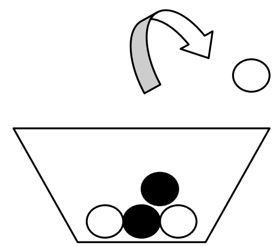
και ουσιαστικά ζητάμε την πιθανότητα $P(A \cap B)$.
Σύμφωνα με τον τύπο (**) έχουμε

$$P(A \cap B) = P(A)P(B/A) = \frac{2}{5} \cdot \frac{1}{4} = \frac{1}{10}$$

<u>ΕΠΕΞΗΓΗΣΗ ΓΙΑ ΤΟ *P(B/A)*</u>*:*

Παρατηρούμε ότι η πιθανότητα *P(A)* είναι $2/5$, καθώς δύο από τις πέντε συνολικά μπάλες είναι ΜΑΥΡΕΣ. Η πιθανότητα *P(B)* είναι πιο δύσκολο να προσδιοριστεί. Η πιθανότητα να τραβήξουμε τη δεύτερη φορά ΜΑΥΡΗ μπάλα εξαρτάται άμεσα από το τι τραβήξαμε την πρώτη φορά.





$A =$ «η πρώτη μπάλα είναι ΜΑΥΡΗ»
οπότε
$$P(B/A) = 1/4$$

$\overline{A} =$ «η πρώτη μπάλα είναι ΑΣΠΡΗ»
οπότε
$$P(B/\overline{A}) = 2/4$$

(Στο παράδειγμά μας χρειαστήκαμε μόνο την δεσμευμένη πιθανότητα $P(B/A) = 1/4$)

Ας υποθέσουμε ότι ρίχνουμε ένα ζάρι δύο φορές και ας θέσουμε

Α = «την πρώτη φορά φέρνουμε άρτιο αριθμό» = {2,4,6}
Β = «την δεύτερη φορά φέρνουμε 6» = {6}

Είναι φανερό ότι τα δύο ενδεχόμενα είναι άσχετα μεταξύ τους, δηλαδή το ενδεχόμενο Α δεν επηρεάζει το ενδεχόμενο Β. Στην περίπτωση αυτή ισχύει

$$P(B/A) = P(B)$$

(που ισούται με $1/6$) καθώς το ενδεχόμενο Α δεν παίζει κανένα ρόλο στη δεύτερη ρίψη. Ο τύπος (**) γίνεται

$$P(A \cap B) = P(A)P(B)$$

ΟΡΙΣΜΟΣ: Όταν ισχύει $P(A\cap B)=P(A)P(B)$ λέμε ότι τα ενδεχόμενα Α και Β είναι ανεξάρτητα.

Συνοψίζοντας, όταν σε ένα πρόβλημα καταλαβαίνουμε ότι τα ενδεχόμενα Α και Β είναι ανεξάρτητα χρησιμοποιούμε τον τύπο

1. $P(A\cap B)=P(A)P(B)$

όταν καταλαβαίνουμε ότι υπάρχει εξάρτηση χρησιμοποιούμε τον τύπο

2. $P(A\cap B)=P(A)P(B/A)$

ΠΑΡΑΔΕΙΓΜΑ 12

Από μία τράπουλα 52 χαρτιών τραβάμε δύο φύλλα, το ένα μετά το άλλο. Να βρεθεί η πιθανότητα να τραβήξουμε δύο ΑΣΣΟΥΣ αν το πρώτο φύλλο

α) ξαναμπεί στην τράπουλα

β) δεν ξαναμπεί στην τράπουλα

πριν προχωρήσουμε στο δεύτερο φύλλο.

Θέτουμε,

Α = «την πρώτη φορά τραβάμε ΑΣΣΟ»

Β = «τη δεύτερη φορά τραβάμε ΑΣΣΟ»

Και στις δύο περιπτώσεις ζητάμε το $P(A\cap B)$.

α) Εφόσον το πρώτο φύλλο ξαναμπαίνει στην τράπουλα, έχουμε ξανά μια τράπουλα 52 φύλλων και η δεύτερη προσπάθεια δεν επηρεάζεται. Συνεπώς, τα δύο ενδεχόμενα είναι ανεξάρτητα και

$$P(A\cap B)=P(A)P(B)=\frac{4}{52}\cdot\frac{4}{52}=\frac{1}{169}$$

β) Εάν την πρώτη φορά τραβήξουμε ΑΣΣΟ και συνεχίσουμε θα μείνει μια τράπουλα με 51 φύλλα και 3 ΑΣΣΟΥΣ. Τα δύο ενδεχόμενα Α και Β είναι εξαρτημένα και έχουμε

$$P(A\cap B)=P(A)P(B/A)=\frac{4}{52}\cdot\frac{3}{51}=\frac{1}{221}$$

Όπως είναι φυσικό, στην περίπτωση β) είναι λιγότερο πιθανό να τραβήξουμε δύο ΑΣΣΟΥΣ καθώς στη δεύτερη προσπάθεια λιγοστεύουν οι διαθέσιμοι ΑΣΣΟΙ.

Στο τελευταίο παράδειγμα ας θέσουμε το ερώτημα «Ποια είναι η πιθανότητα το δεύτερο φύλλο να είναι ΑΣΣΟΣ;», δηλαδή ζητάμε το $P(B)$. Καθώς το B εξαρτάται από το αποτέλεσμα της πρώτης προσπάθειας, θα πρέπει να εξετάσουμε την πιθανότητα του B σε συνδυασμό με όλες τις δυνατές περιπτώσεις της πρώτης προσπάθειας και να αθροίσουμε τα επιμέρους αποτελέσματα. Έτσι,

$$P(B) = P(B/A)P(A) + P(B/\overline{A})P(\overline{A})$$

$$= \frac{3}{51}\cdot\frac{4}{52} + \frac{4}{51}\cdot\frac{48}{52} = \frac{1}{13}$$

Γενικά, αν τα ενδεχόμενα $A_1, A_2, \ldots, A_n$ αποτελούν μια διαμέριση του δειγματοχώρου S, δηλαδή

- είναι ανά δύο ξένα μεταξύ τους και
- $A_1 \cup A_2 \cup \cdots \cup A_n = S$

τότε

$$\boxed{P(B) = P(B/A_1)P(A_1) + P(B/A_2)P(A_2) + \cdots + P(B/A_n)P(A_n)}$$

ΠΑΡΑΔΕΙΓΜΑ 13

50% των ψυγείων της αγοράς κατασκευάζονται από την εταιρεία Α, 25% από την εταιρεία Β και 25% από την εταιρεία Γ. Ένα ποσοστό 2% τόσο της εταιρείας Α όσο και της Β είναι ελαττωματικά ενώ το ποσοστό αυτό για την εταιρεία Γ είναι 4%. Εάν διαλέξουμε ένα ψυγείο στην τύχη, ποια είναι η πιθανότητα να είναι ελαττωματικό;

Τα πάντα ξεκαθαρίζονται αν καθορίσουμε κατάλληλα ενδεχόμενα. Θέτουμε λοιπόν

- A_1 = «το ψυγείο είναι της εταιρείας Α» με $P(A_1) = 50/100$

- A_2 = «το ψυγείο είναι της εταιρείας Β» με $P(A_2) = {}^{25}\!/_{100}$

- A_3 = «το ψυγείο είναι της εταιρείας Γ» με $P(A_3) = {}^{25}\!/_{100}$

ενώ

- B = «το ψυγείο είναι ελαττωματικό»

και ουσιαστικά ζητάμε την πιθανότητα $P(B)$. Τα δεδομένα μας λένε

$$P(B/A_1) = {}^{2}\!/_{100} \qquad P(B/A_2) = {}^{2}\!/_{100} \qquad P(B/A_3) = {}^{4}\!/_{100}$$

Έχουμε

$$P(B) = P(B/A_1)P(A_1) + P(B/A_2)P(A_2) + P(B/A_3)P(A_3)$$

$$= \frac{2}{100}\cdot\frac{50}{100} + \frac{2}{100}\cdot\frac{25}{100} + \frac{4}{100}\cdot\frac{25}{100}$$

$$= \frac{1}{100} + \frac{1}{200} + \frac{1}{100} = \frac{1}{40} = 2.5\%$$

Η ζητούμενη πιθανότητα είναι 2.5%

ΑΣΚΗΣΕΙΣ

2.1 Εκτελούμε τα παρακάτω πειράματα τύχης:

α) Ρίχνουμε ένα ζάρι και μελετάμε τις ενδείξεις.

β) Ρίχνουμε δύο ζάρια και μετράμε το άθροισμα των ενδείξεων.

γ) Ρίχνουμε ένα ζάρι 12 φορές. Μετράμε το πλήθος των 6αριών που μπορούμε να φέρουμε.

δ) Ρίχνουμε 12 ζάρια. Μετράμε το πλήθος των 6αριών που μπορούμε να φέρουμε.

ε) Ρίχνουμε ένα ζάρι μέχρι να φέρουμε δύο 6άρια. Μετράμε το πλήθος των ρίψεων.

Για κάθε πείραμα απαντήστε στα παρακάτω ερωτήματα:

Ποιος είναι ο δειγματοχώρος; Εκτιμάτε ότι έχουν όλα τα απλά ενδεχόμενα την ίδια πιθανότητα; Ποιο απλό ενδεχόμενο νομίζετε ότι είναι πιο πιθανό;

2.2 Με βάση τα αποτελέσματα που βρήκατε στην άσκηση 1.1, αν επιλέξουμε 8 αριθμούς στο «Extra 5» ποια είναι η πιθανότητα να πετύχουμε 5άρι;

2.3. Θεωρήστε όλες τις ελληνικές λέξεις των 5 γραμμάτων (όχι απαραίτητα με νόημα) που μπορούν να σχηματιστούν. Με βάση τα αποτελέσματα που βρήκατε στην άσκηση 1.3, να υπολογίσετε τις παρακάτω πιθανότητες:

α) μια λέξη να αρχίζει από Α;

β) μια λέξη να αρχίζει και να καταλήγει στο ίδιο γράμμα;

γ) μια λέξη να έχει όλα τα γράμματα διαφορετικά;

δ) μια λέξη να έχει τουλάχιστον δύο ίδια γράμματα;

2.4 Με βάση τα αποτελέσματα που βρήκατε στην άσκηση 1.6, ποια είναι η πιθανότητα μια πινακίδα αυτοκινήτου να έχει

α) τρία κοινά γράμματα

β) τέσσερις ίδιους αριθμούς

γ) τρία κοινά γράμματα και τέσσερις ίδιους αριθμούς

δ) αριθμό που να μην περιέχει το ψηφίο 0

2.5 Να υπολογίσετε τα εξής

α) Ποια είναι η πιθανότητα ένας φυσικός αριθμός από το 0 μέχρι το 999 να περιέχει το ψηφίο 7 (τουλάχιστον μια φορά);

β) Ποια είναι η πιθανότητα ένας τετραψήφιος αριθμός να περιέχει το ψηφίο 7 ; [υπόδειξη: προτιμότερο να μετρήσετε πόσοι αριθμοί δεν περιέχουν το ψηφίο 7]

2.6 Με βάση τα αποτελέσματα που βρήκατε στην άσκηση 1.8, αν μοιράσω 8 διαφορετικά δώρα σε 3 φίλους μου με τυχαίο τρόπο, ποια είναι η πιθανότητα να πάρουν όλοι από ένα δώρο τουλάχιστον;

2.7 Δύο δοχεία, Α και Β, περιέχουν από 5 άσπρες και 3 μαύρες μπάλες το καθένα. Διαλέγουμε τυχαία μια μπάλα από το δοχείο Α και την τοποθετούμε στο δοχείο Β. Στη συνέχεια διαλέγουμε τυχαία μια μπάλα από το δοχείο Β.

α) Ποια είναι η πιθανότητα η πρώτη μπάλα να είναι άσπρη και η δεύτερη μαύρη;

β) Ποια είναι η πιθανότητα η δεύτερη μπάλα να είναι μαύρη;

2.8 Ποιες είναι οι απαντήσεις στην προηγούμενη άσκηση εάν αντί για 5 άσπρες και 3 μαύρες μπάλες είχαμε x και y αντίστοιχα;

2.9 Τα ποσοστά 3 κομμάτων σε κάποιες εκλογές είναι

Κόμμα Α: 40%

Κόμμα Β: 35%

Κόμμα Γ: 25%

Ωστόσο, μετά από ένα χρόνο οι συσπειρώσεις των κομμάτων (ποσοστά που παραμένουν στο κόμμα τους) έχουν ως εξής:

Στο Κόμμα Α: 50%

Στο Κόμμα Β: 60%

Στο Κόμμα Γ: 80%

α) Ποιο είναι γενικά το ποσοστό των συσπειρωμένων ψηφοφόρων;

β) Αν ένας ψηφοφόρος δηλώνει συσπειρωμένος στο κόμμα του ποια είναι η πιθανότητα να είχε ψηφίσει το κόμμα Γ;

ΣΥΝΤΟΜΕΣ ΑΠΑΝΤΗΣΕΙΣ ΤΩΝ ΑΣΚΗΣΕΩΝ

2.1 α) S={1,2,3,4,5,6}

β) S={2,3,4,5,6,7,8,9,10,11,12}

γ) και δ) S={0,1,2,3,4,5,6,7,8,9,10,11,12}

ε) S={2,3,4,5,6,…}

2.2 1/5797

2.3 α) 1/24

β) 1/24

γ) 5100480/7962624

δ) 2862144/7962624

2.4 α) 1/196

β) 1/1000

γ) 1/19600

δ) 729/1000

2.5 α) 271/1000 β) 3168/9000

2.6 5796/6561=88.34%

2.7 α) 5/24 β) 3/8

2.8 α) $\frac{xy}{(x+y)(x+y+1)}$ β) $\frac{xy+y(y+1)}{(x+y)(x+y+1)}$

2.9 α) 61% β) 32.79%

3. ΚΑΤΑΝΟΜΕΣ ΤΥΧΑΙΩΝ ΜΕΤΑΒΛΗΤΩΝ

3.1 ΤΥΧΑΙΕΣ ΜΕΤΑΒΛΗΤΕΣ

Είδαμε ότι τα αποτελέσματα ενός πειράματος τύχης, τα οποία αποτελούν το δειγματοχώρο μας, είναι συνήθως αριθμοί. Ακόμη όμως και στην περίπτωση όπου δεν έχουμε αριθμούς αλλά περιγραφικά αποτελέσματα, π.χ. ΚΕΦΑΛΗ-ΓΡΑΜΜΑ, θα μπορούσαμε να αντιστοιχίσουμε σε κάθε δυνατό αποτέλεσμα έναν πραγματικό αριθμό (π.χ. ΚΕΦΑΛΗ=0, ΓΡΑΜΜΑ=1), έτσι ώστε να μελετήσουμε πιο εύκολα το πείραμα και να συναγάγουμε χρήσιμα συμπεράσματα. Μπορούμε π.χ. να βρούμε έναν μέσο όρο αυτών των τιμών για να δούμε που περίπου κυμαίνονται τα δυνατά αποτελέσματα. Μπορούμε επίσης να αποφανθούμε αν οι τιμές αυτές είναι αρκετά συγκεντρωμένες γύρω από τον μέσο όρο ή πιο διάσπαρτες!

Με τον τρόπο αυτό καθορίζουμε μια **Τυχαία Μεταβλητή** που παίρνει τιμές στο σύνολο των πραγματικών αριθμών. Τις τυχαίες μεταβλητές θα τις συμβολίζουμε με κεφάλαια γράμματα, ενώ τις τιμές που παίρνουν με μικρά. Έτσι λέμε ότι η τυχαία μεταβλητή X παίρνει την τιμή a. Την αντίστοιχη πιθανότητα, να είναι δηλαδή $X = a$, τη συμβολίζουμε $P(X = a)$. Στο παράδειγμα του ζαριού, αν X είναι η τυχαία μεταβλητή που δείχνει την ένδειξη μετά από μια ρίψη, η X παίρνει τις τιμές 1,2,3,4,5,6. Είναι

$$P(X=1) \;=\; P(X=2) \;=\; \cdots \;=\; P(X=6) \;=\; \frac{1}{6}$$

Μία τυχαία μεταβλητή X μπορεί να είναι **διακριτή** ή **συνεχής.**

Η X είναι **διακριτή** εάν παίρνει τιμές

- από ένα πεπερασμένο σύνολο διακεκριμένων τιμών,
  π.χ. στο παράδειγμα του ζαριού από το σύνολο {1,2,3,4,5,6}, είτε
- από ένα σύνολο άπειρο μεν αλλά αριθμήσιμο (το λέμε απειραριθμήσιμο)
  π.χ. από το σύνολο των θετικών φυσικών {1,2,3,4,5,...}

Η X είναι **συνεχής** εάν παίρνει (άπειρες) τιμές

- σε ένα διάστημα πραγματικών αριθμών (a,b).

  π.χ. αν X εκφράζει το ύψος σε μέτρα των δέντρων ενός δάσους που παίρνει τιμές στο διάστημα (1,30)
  ή αν X εκφράζει το χρόνο που μεσολαβεί μεταξύ δύο τηλεφωνικών κλήσεων σε ένα τηλεφωνικό κέντρο, που παίρνει τιμές στο διάστημα $(0,+\infty)$ (θεωρητικά μετά από μία κλήση μπορεί να μην ξαναγίνει κλήση ποτέ)
  Τα άκρα a,b του διαστήματος λοιπόν μπορεί να είναι $-\infty$, $+\infty$

3.2 ΚΑΤΑΝΟΜΗ ΜΙΑΣ ΔΙΑΚΡΙΤΗΣ ΤΥΧΑΙΑΣ ΜΕΤΑΒΛΗΤΗΣ

Έστω ότι η τυχαία μεταβλητή X μπορεί να πάρει τις διακριτές τιμές

$$x_1,\ x_2,\ x_{3,}\ x_{4,}\ \ldots$$

(το πλήθος τους μπορεί να είναι είτε πεπερασμένο είτε απειραριθμήσιμο)
με αντίστοιχες πιθανότητες

$$p_1,\ p_2,\ p_{3,}\ p_{4,}\ \cdots$$

Οι πιθανότητες δεν μπορεί να είναι αρνητικοί αριθμοί, ενώ το άθροισμά τους πρέπει να είναι 1. Δηλαδή, ισχύουν οι προϋποθέσεις

a) $p_i \geq 0$, για $i = 1,2,3,\ldots$

b) $\sum_i p_i = p_1 + p_2 + p_3 + \cdots = 1$

Τότε λέμε ότι έχουμε ορίσει μια διακριτή κατανομή με **συνάρτηση πιθανότητας** *p*(*x*), όπου

$$p(x_1) = p_1, \quad p(x_2) = p_2, \quad \text{κλπ}$$

<u>Σημείωση</u>: Στην ιδιότητα b), αν έχουμε πεπερασμένο άθροισμα γράφουμε $\sum_{i=1}^{n} p_i = 1$, ενώ αν έχουμε απειραριθμήσιμο γράφουμε $\sum_{i=1}^{+\infty} p_i = 1$.

ΠΑΡΑΔΕΙΓΜΑ 1

Ρίχνουμε δύο ζάρια και συμβολίζουμε με Χ την τυχαία μεταβλητή που δηλώνει το άθροισμα των ενδείξεων. Πρόκειται για διακριτή κατανομή με συνάρτηση πιθανότητας

x	2	3	4	5	6	7	8	9	10	11	12
p(x)	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Π.χ. είναι $p(5)=\frac{4}{36}$, καθώς από τους 36 συνολικά συνδυασμούς, οι συνδυασμοί που δίνουν άθροισμα πέντε είναι 4, συγκεκριμένα

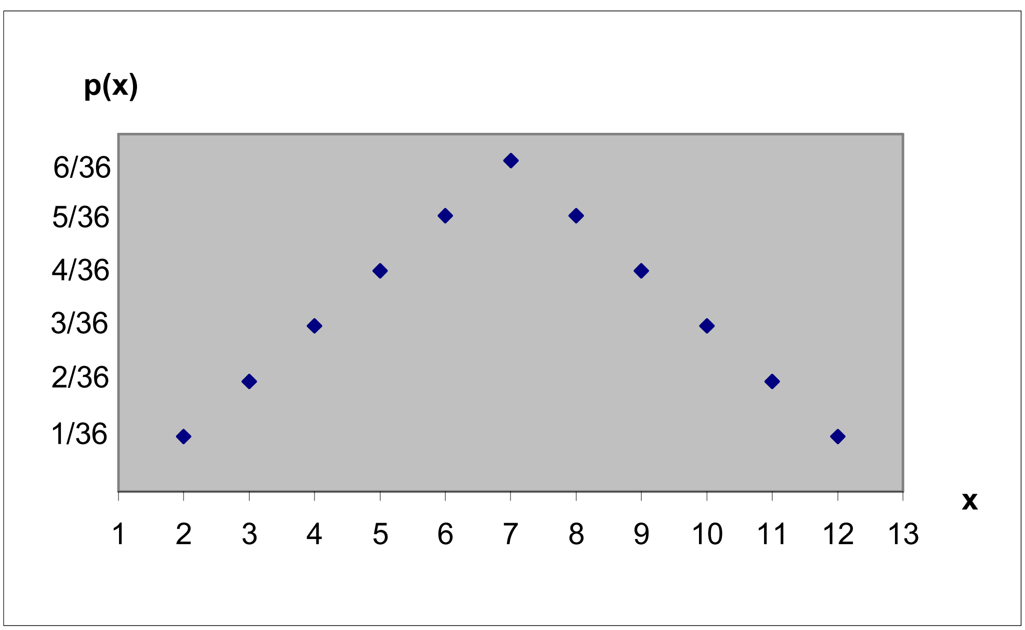
1-4, 2-3. 3-2 και 4-1.

Οι δύο προϋποθέσεις της συνάρτησης πιθανότητας ισχύουν:

a) $p(x)\geq 0$, για $x=2,3,\ldots,12$

b) $p(2)+p(3)+\cdots+p(12)=\frac{36}{36}=1$

Η γραφική παράσταση της συνάρτησης πιθανότητας είναι



Παρατηρούμε ότι η συνάρτηση $p(x)$ χαρακτηρίζει πλήρως την κατανομή, δηλαδή όταν την γνωρίζουμε, γνωρίζουμε επακριβώς την κατανομή.

Στο τελευταίο παράδειγμα θέτουμε το ερώτημα:

Ποια είναι η πιθανότητα να έχουμε άθροισμα από 5 μέχρι και 8;

Δεν έχουμε παρά να αθροίσουμε τις αντίστοιχες πιθανότητες:

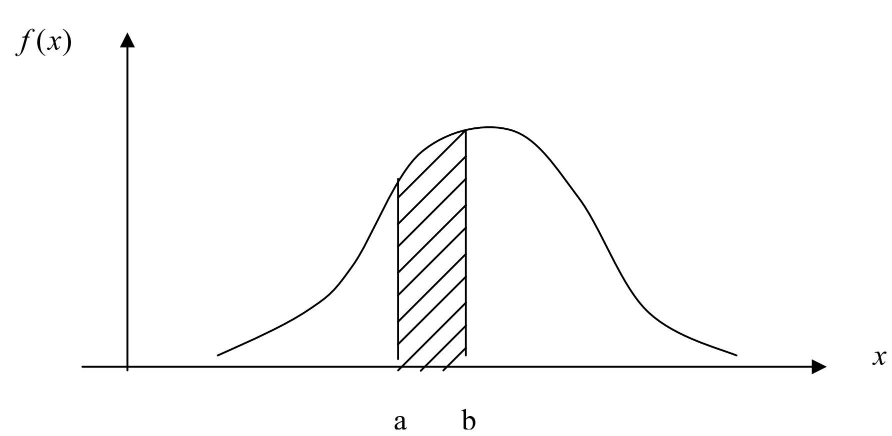
$$P(5 \le X \le 8) = \sum_{x=5}^{8} p(x) = p(5) + p(6) + p(7) + p(8)$$

$$= \frac{4}{36} + \frac{5}{36} + \frac{6}{36} + \frac{5}{36} = \frac{20}{36}$$

Θυμηθείτε το ερώτημα αυτό όταν θα επεκταθούμε στη συνεχή κατανομή και θα αναζητάμε την πιθανότητα σε ένα διάστημα.

3.3 ΚΑΤΑΝΟΜΗ ΜΙΑΣ ΣΥΝΕΧΟΥΣ ΤΥΧΑΙΑΣ ΜΕΤΑΒΛΗΤΗΣ

Οι ιδέες της προηγούμενης παραγράφου μπορούν να επεκταθούν και στην περίπτωση όπου η μεταβλητή X είναι συνεχής. Όπως εκεί είχαμε μια χαρακτηριστική συνάρτηση p(x) με διακριτές τιμές, εδώ θα περιγράψουμε μια αντίστοιχη συνεχή συνάρτηση f(x). Η γραφική παράσταση της f(x) δεν είναι ένα σύνολο σημείων όπως στην προηγούμενη παράγραφο, αλλά μια συνεχής γραμμή.



ΠΑΡΑΤΗΡΗΣΕΙΣ

- Το εμβαδόν της περιοχής ανάμεσα στην f(x) και τον άξονα Ox είναι 1 [συγκρίνετε με το άθροισμα όλων των τιμών της p(x)]

- Το εμβαδόν της σκιασμένης περιοχής μας δίνει την πιθανότητα να βρίσκεται η X μεταξύ των τιμών a και b, δηλαδή την

$$P(a \le X \le b)$$

  Στη συνεχή κατανομή δεν έχει νόημα να ζητάμε την πιθανότητα σε κάποιο σημείο

$$P(X = a)$$

  καθώς αυτή είναι απειροελάχιστη.

  [συγκρίνετε με την τελευταία παρατήρηση της προηγούμενης παραγράφου]

- Η συνάρτηση f(x) ονομάζεται ***pdf*** (probability density function – συνάρτηση πυκνότητας πιθανότητας)

Αν θυμηθούμε ότι το εμβαδόν κάτω από μια καμπύλη αντιστοιχεί σε ολοκλήρωμα, είμαστε έτοιμοι να δώσουμε τον αυστηρό ορισμό της συνεχούς κατανομής.

ΟΡΙΣΜΟΣ Η μεταβλητή X ακολουθεί συνεχή κατανομή (είναι δηλαδή συνεχής τυχαία μεταβλητή) εάν υπάρχει συνάρτηση f, την οποία θα καλούμε pdf, που μας δίνει την πιθανότητα για οποιοδήποτε διάστημα (a,b) σύμφωνα με τον τύπο

$$P(a < X < b) = \int_a^b f(x)dx$$

Η f(x) πληροί τις προϋποθέσεις

a) $f(x) \ge 0$ για κάθε *x*

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

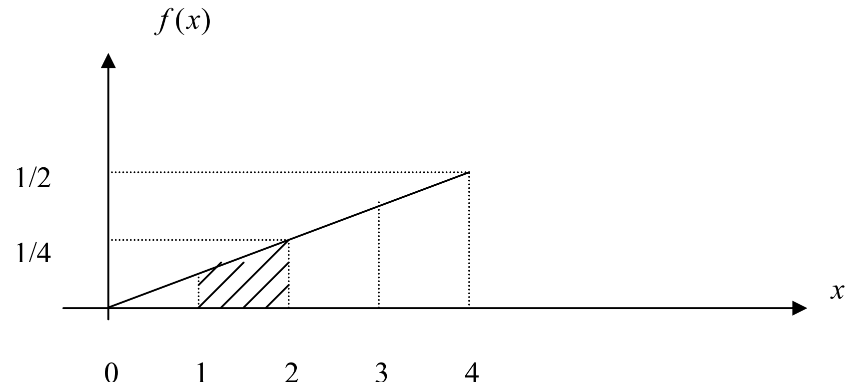
Σημείωση: καθώς η πιθανότητα σε σημείο δεν έχει νόημα θεωρούμε ότι τα άκρα του διαστήματος δεν επηρεάζουν την πιθανότητα του διαστήματος, με άλλα λόγια

$$P(a < X < b) = P(a \le X \le b) = P(a \le X < b) = P(a < X \le b)$$

ΠΑΡΑΔΕΙΓΜΑ 2

Έστω X συνεχής τυχαία μεταβλητή που παίρνει τιμές στο διάστημα [0,4], με pdf

$$f(x) = \frac{1}{8}x$$



Η f(x) είναι πράγματι pdf καθώς ικανοποιεί τις προϋποθέσεις:

a) $f(x) = \frac{1}{8}x \geq 0$ για κάθε $x \in [0,4]$

b) $\int_{-\infty}^{+\infty} f(x)dx = \int_0^4 \frac{1}{8}xdx = \left[\frac{x^2}{16}\right]_0^4 = 1 - 0 = 1$

Επίσης,

$$P(1 < X < 2) = \int_1^2 f(x)dx = \int_1^2 \frac{1}{8}xdx = \left[\frac{x^2}{16}\right]_1^2 = \frac{4}{16} - \frac{1}{16} = \frac{3}{16}$$

3.4 (ΑΘΡΟΙΣΤΙΚΗ) ΣΥΝΑΡΤΗΣΗ ΚΑΤΑΝΟΜΗΣ - cdf

Έστω X τυχαία μεταβλητή, διακριτή ή συνεχής. Ορίζουμε την cdf (cumulative distribution function – συνάρτηση κατανομής) F της X ως

$$F(y) = P(X \leq y)$$

δηλαδή η *F(y)* είναι η πιθανότητα η μεταβλητή να παίρνει τιμή μέχρι και y.

- Αν X διακριτή, αθροίζουμε όλες τις πιθανότητες των τιμών μέχρι και y

$$F(y) = \sum_{x_j \le y} p(x_j)$$

- Αν X συνεχής, βρίσκουμε την πιθανότητα στο διάστημα [-∞,y], δηλαδή

$$F(y) = \int_{-\infty}^{y} f(x)dx$$

<u>ΠΑΡΑΤΗΡΗΣΕΙΣ</u>: 1.Αν η f είναι συνεχής σε ένα διάστημα (a,b), (ενώ εκτός του (a,b) είναι 0), η τιμή της cdf είναι ξεκάθαρη για y εκτός του διαστήματος:

Αν $y < a$, τότε $F(y) = 0$, διότι $F(y) = \int_{-\infty}^{y} f(x)dx = \int_{-\infty}^{y} 0dx = 0$.

Αν $y \ge b$, τότε $F(y) = 1$, διότι $F(y) = \int_{-\infty}^{y} f(x)dx = \int_{a}^{b} f(x)dx = 1$

Μένει λοιπόν κάθε φορά να εξετάζουμε την $F(y)$ για τιμές $a \le y < b$.

2. Στην περίπτωση της συνεχούς μεταβλητής $F'(x) = f(x)$ σχεδόν παντού.

ΠΑΡΑΔΕΙΓΜΑ 3

Έστω X συνεχής τυχαία μεταβλητή με pdf

$$f(x) = \begin{cases} 2x, & 0 < x < 1 \\ 0, & \text{αλλού} \end{cases}$$

Να διαπιστώσετε ότι η f είναι πράγματι pdf. Να βρεθεί η cdf. Τέλος να βρεθεί η πιθανότητα να είναι $X \le 0.7$. Ποιο είναι πιθανότερο, να είναι $X \le 0.7$ ή $X \ge 0.7$;

Η f είναι πράγματι pdf διότι,

a) $f(x) \ge 0$

b) $\int_{-\infty}^{+\infty} f(x)dx = \int_{0}^{1} 2xdx = \left[x^2\right]_0^1 = 1 - 0 = 1$

Για την cdf, αν $0 \le y < 1$ τότε

$$F(y) = \int_{-\infty}^{y} f(x)dx = \int_{0}^{y} 2xdx = \left[x^2\right]_0^y = y^2.$$

Άρα,

$$F(y) = \begin{cases} 0, & y < 0 \\ y^2, & 0 \le y < 1 \\ 1, & y \ge 1 \end{cases}$$

Προσέξτε ότι $F'(x) = f(x)$ (εκτός φυσικά των σημείων $x = 0$ και $x = 1$).

Η ζητούμενη πιθανότητα είναι

$$P(X \le 0.7) = \int_{-\infty}^{0.7} f(x)dx = \int_0^{0.7} 2xdx = \left[x^2\right]_0^{0.7} = (0.7)^2 - 0 = 0.49$$

Ας προσέξουμε όμως ότι την πιθανότητα αυτή τη δίνει και η cdf:

$$P(X \le 0.7) = F(0.7) = (0.7)^2 = 0.49$$

Τέλος είναι $P(X \ge 0.7) = 1 - 0.49 = 0.51$, άρα το δεύτερο ενδεχόμενο είναι πιθανότερο.

ΠΑΡΑΔΕΙΓΜΑ 4

Έστω ότι η Χ είναι διακριτή μεταβλητή που παίρνει τις τιμές 0,1,2 με αντίστοιχες πιθανότητες

$$p(0) = \frac{1}{3} \qquad p(1) = \frac{1}{6} \qquad p(2) = \frac{1}{2}$$

Να επιβεβαιώσετε ότι η p(x) είναι πράγματι συνάρτηση πιθανότητας. Να βρεθεί η cdf. Ποια είναι η πιθανότητα να έχουμε $X \le 1.5$;

Η p(x) είναι πράγματι συνάρτηση πιθανότητας διότι πληροί τις προϋποθέσεις:

a) $p(x) \ge 0$ για $x = 0,1,2$

b) $p(0) + p(1) + p(2) = \frac{1}{3} + \frac{1}{6} + \frac{1}{2} = 1$

Για την cdf παρατηρούμε ότι

$$F(0) = p(0) = \frac{1}{3}$$

$$F(1) = p(0) + p(1) = \frac{1}{3} + \frac{1}{6} = \frac{1}{2}$$

$$F(2) = p(0) + p(1) + p(2) = 1$$

ενώ ανάμεσα στις διακριτές τιμές η F(y) δεν αλλάζει. Άρα,

$$F(y) = \begin{cases} 0, & y < 0 \\ 1/3, & 0 \le y < 1 \\ 1/2, & 1 \le y < 2 \\ 1, & y \ge 2 \end{cases}$$

Η πιθανότητα $P(X \le 1.5)$ μπορεί να βρεθεί με δύο τρόπους:

- $P(X \le 1.5) = p(0) + p(1) = \frac{1}{3} + \frac{1}{6} = \frac{1}{2}$

ή

- $P(X \le 1.5) = F(1.5) = \frac{1}{2}$, διότι η $F(y)$ μας δίνει ακριβώς την $P(X \le y)$

3.5 Η ΜΕΣΗ ΤΙΜΗ *E(X)=μ*

Πριν δώσουμε τον ορισμό της μέσης τιμής μιας μεταβλητής X ας ανατρέξουμε στην εμπειρία μας. Γνωρίζουμε ότι η μέση τιμή *μ* (ή αλλιώς μέσος όρος) των τιμών

$$x_1, x_2, \ldots, x_n$$

βρίσκεται ως εξής

$$\mu = \frac{x_1 + x_2 + \cdots x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Π.χ. ο μέσος όρος των τιμών 5,5,5,8,8,6,6,6,6,2 είναι

$$\mu = \frac{5+5+5+8+8+6+6+6+6+2}{10} = 5.7$$

Αν ομαδοποιήσουμε βέβαια τα δεδομένα έχουμε

$$\mu = \frac{3*5+2*8+4*6+1*2}{10} = \frac{3}{10}*5 + \frac{2}{10}*8 + \frac{4}{10}*6 + \frac{1}{10}*2 = 5.7$$

Παρατηρούμε λοιπόν ότι όταν οι τιμές

$$x_1, x_2, \ldots, x_k$$

εμφανίζονται αντίστοιχα

$$\lambda_1, \lambda_2, \ldots, \lambda_k$$

φορές, (οπότε το συνολικό πλήθος των τιμών είναι $n = \lambda_1 + \lambda_2 + \cdots + \lambda_\kappa$) και η μέση τιμή είναι

$$\mu = \frac{\lambda_1 x_1 + \lambda_2 x_2 + \cdots \lambda_k x_k}{n} = x_1 \frac{\lambda_1}{n} + x_2 \frac{\lambda_2}{n} + \cdots + x_k \frac{\lambda_k}{n} = \sum_{i=1}^{k} x_i p(x_i)$$

όπου $p(x_i) = \dfrac{\lambda_i}{n}$ είναι η πιθανότητα εμφάνισης της τιμής x_i ανάμεσα στις n τιμές.

Τώρα είμαστε έτοιμοι να δώσουμε και επίσημα τον ορισμό της μέσης τιμής.

<u>ΟΡΙΣΜΟΣ</u> Διακρίνουμε δύο περιπτώσεις για τη μέση τιμή *E(X)* (είτε *μ*.)

α) Έστω X διακριτή μεταβλητή που παίρνει τις τιμές

$$x_1, x_2, \ldots$$

με αντίστοιχες πιθανότητες

$$p(x_1), p(x_2), \ldots$$

Η μέση τιμή της μεταβλητής X είναι

$$\boxed{E(X) = \sum_i x_i p(x_i)}$$

β) Έστω X συνεχής μεταβλητή με pfd τη συνάρτηση f(x).

Η μέση τιμή της μεταβλητής X είναι

$$\boxed{E(X) = \int_{-\infty}^{+\infty} x f(x) dx}$$

<u>Σημείωση</u>: Για να είμαστε μαθηματικά νόμιμοι στους παραπάνω ορισμούς θα πρέπει να προσθέσουμε μια προϋπόθεση:

α) $\sum_i |x_i| p(x_i) < +\infty$ β) $\int_{-\infty}^{+\infty} |x| f(x) dx < +\infty$

Το επόμενο θεώρημα είναι αρκετά χρήσιμο.

<u>ΘΕΩΡΗΜΑ</u> (χωρίς απόδειξη) Έστω *X* τυχαία μεταβλητή και *A(X)* μια συνάρτηση του *X*. Η μέση τιμή της *Y=A(X)* είναι

Α) αν *X* διακριτή

$$E(Y) = \sum_i A(x_i)p(x_i)$$

β) αν *X* συνεχής

$$E(Y) = \int_{-\infty}^{+\infty} A(x)f(x)dx$$

Προσοχή! Σύμφωνα με το Θεώρημα, ενώ η μέση τιμή μιας συνεχούς π.χ. μεταβλητής *X* δίνεται από τον τύπο

$$E(X) = \int_{-\infty}^{+\infty} xf(x)dx$$

η μέση τιμή της $A(X)=3X^2+1$ δίνεται από τον τύπο

$$E(3X^2+1) = \int_{-\infty}^{+\infty} (3x^2+1)f(x)dx$$

και όχι από τον τύπο

$$E(3X^2+1) = \int_{-\infty}^{+\infty} (3x^2+1)f(3x^2+1)dx$$

όπως λανθασμένα θα μπορούσε να περιμένει κανείς.

3.6 Η ΔΙΑΣΠΟΡΑ *V(X)*

Ας ξεκινήσουμε με μια παρατήρηση. Έστω ότι

η μεταβλητή Χ λαμβάνει τις τιμές 30,40,50,60,70
η μεταβλητή Υ λαμβάνει τις τιμές 10,30,50,70,90

με ίσες πιθανότητες - 1/5 την κάθε τιμή. Η μέση τιμή και στις δύο περιπτώσεις είναι 50. Π.χ. για τη πρώτη μεταβλητή

$$E(X) = \sum_{i=1}^{5} x_i p(x_i) = 30\cdot\frac{1}{5} + 40\cdot\frac{1}{5} + 50\cdot\frac{1}{5} + 60\cdot\frac{1}{5} + 70\cdot\frac{1}{5} = 50$$

Η διαφορά ανάμεσα στις δύο περιπτώσεις είναι ότι οι τιμές της Υ είναι πιο «απλωμένες» από τις τιμές της Χ. Άρα, εκτός από την μέση τιμή που εκφράζει μια κεντρική τιμή των τιμών χρειαζόμαστε και ένα μέτρο για το πόσο «διάσπαρτες» είναι οι τιμές γύρω από την κεντρική τιμή. Το ρόλο αυτό θα τον παίξει η λεγόμενη διασπορά. Εάν *μ* είναι η μέση τιμή των τιμών

$$x_1, x_2, \ldots, x_n$$

ορίζουμε

$$\textit{διασπορά} = \frac{(x_1-\mu)^2 + (x_2-\mu)^2 + \cdots + (x_n-\mu)^2}{n}$$

Στο παράδειγμά μας

$$\textit{διασπορά(X)} = \frac{(30-50)^2 + (40-50)^2 + (50-50)^2 + (60-50)^2 + (70-50)^2}{5} = 200$$

ενώ

$$\textit{διασπορά(Y)} = \frac{(10-50)^2 + (30-50)^2 + (50-50)^2 + (70-50)^2 + (90-50)^2}{5} = 800$$

που υποδηλώνει ότι οι τιμές τις Υ είναι πιο «διάσπαρτες» από αυτές της Χ.

Αν προσέξουμε τον ορισμό της διασποράς θα δούμε ότι στην ουσία παίρνουμε τη μέση τιμή των $(x_i-\mu)^2$. Είμαστε λοιπόν και πάλι έτοιμοι να δώσουμε τον επίσημο ορισμό της διασποράς.

ΟΡΙΣΜΟΣ Έστω *X* τυχαία μεταβλητή (διακριτή ή συνεχής). Η **διασπορά** της *X* συμβολίζεται με *V(X)* (είτε με σ^2) και ορίζεται ως η μέση τιμή της $(X\text{-}\mu)^2$, δηλαδή

$$\boxed{V(X) = E\,[X-\mu]^2}$$

Η τιμή $\sigma = \sqrt{V(X)}$ είναι γνωστή ως **τυπική απόκλιση** της *X.*

Ένας ισοδύναμος αλλά αρκετά πιο εύκολος τύπος της διασποράς για τις εφαρμογές είναι

$$\boxed{V(X) = E(X^2) - \mu^2}$$

Πράγματι,

$$V(X) = E\,[X-\mu]^2 = E\,[X^2 - 2X\mu + \mu^2] = E\,[X^2] - E[2X\mu] + E[\mu^2]$$

$$= E\,[X^2] - 2\mu E(X) + \mu^2 = E\,[X^2] - 2\mu\cdot\mu + \mu^2 = E\,[X^2] - \mu^2$$

<u>ΠΑΡΑΔΕΙΓΜΑ 5</u>

Έστω Χ διακριτή τυχαία μεταβλητή με συνάρτηση πιθανότητας που δίνεται από τον πίνακα

X	1	2	3	4	5
p(x)	0.1	0.2	0.4	0.2	0.1

Η μέση τιμή της *X* είναι

$$E(X) = \sum_{i=1}^{5} x_i\, p(x_i) = 1\cdot 0.1 + 2\cdot 0.2 + 3\cdot 0.4 + 4\cdot 0.2 + 5\cdot 0.1 = 3$$

Η διασπορά της *X* είναι

$$V(X) = E\,[X-\mu]^2 = \sum_{i=1}^{5} (x_i - 3)^2\, p(x_i) = 4\cdot 0.1 + 1\cdot 0.2 + 0\cdot 0.4 + 1\cdot 0.2 + 4\cdot 0.1 = 1.2$$

Εναλλακτικά, με τον δεύτερο τύπο, βρίσκουμε πρώτα

$$E(X^2) = \sum_{i=1}^{5} x_i^{\,2}\, p(x_i) = 1^2\cdot 0.1 + 2^2\cdot 0.2 + 3^2\cdot 0.4 + 4^2\cdot 0.2 + 5^2\cdot 0.1 = 10.2$$

οπότε

$$V(X) = E(X^2) - \mu^2 = 10.2 - 3^2 = 1.2$$

Τέλος, η τυπική απόκλιση είναι

$$\sigma = \sqrt{1.2} \cong 1.095$$

<u>ΠΑΡΑΔΕΙΓΜΑ 6</u>

Έστω Χ συνεχής τυχαία μεταβλητή με pdf

$$f(x) = \begin{cases} 1+x, & -1 \le x \le 0 \\ 1-x, & 0 < x \le 1 \end{cases}$$

Η μέση τιμή της Χ είναι

$$E(X) = \int_{-\infty}^{+\infty} xf(x)dx = \int_{-1}^{1} xf(x)dx = \int_{-1}^{0} x(1+x)dx + \int_{0}^{1} x(1-x)dx$$

$$= \int_{-1}^{0} (x+x^2)dx + \int_{0}^{1} (x-x^2)dx = \left[\frac{x^2}{2} + \frac{x^3}{3}\right]_{-1}^{0} + \left[\frac{x^2}{2} - \frac{x^3}{3}\right]_{0}^{1} = 0$$

Για τη διασπορά, βρίσκουμε πρώτα

$$E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x)dx = \int_{-1}^{0} x^2(1+x)dx + \int_{0}^{1} x^2(1-x)dx$$

$$= \int_{-1}^{0} (x^2+x^3)dx + \int_{0}^{1} (x^2-x^3)dx = \left[\frac{x^3}{3} + \frac{x^4}{4}\right]_{-1}^{0} + \left[\frac{x^3}{3} - \frac{x^4}{4}\right]_{0}^{1} = \frac{1}{6}$$

οπότε,

$$V(X) = E(X^2) - \mu^2 = \frac{1}{6} - 0^2 = \frac{1}{6}$$

<u>ΠΑΡΑΔΕΙΓΜΑ 7</u>

Δίνεται η συνάρτηση

$$f(x) = \begin{cases} 3(x^2 - 2x + 1) & 1 < x < 2 \\ 0 & \text{αλλού} \end{cases}$$

α) Να δειχθεί ότι η f είναι pdf μιας τυχαίας μεταβλητής Χ.

β) Να βρεθεί η πιθανότητα του ενδεχομένου "$X < 3/2$". Ποια η πιθανότητα να είναι θετική η τιμή της Υ =10Χ-15 ;

γ) Να βρεθεί η μέση τιμή της Χ.

δ) Να βρεθεί η μέση τιμή της $\dfrac{1}{(X-1)^2}$

ε) Ποια η πιθανότητα των ενδεχομένων

Α = "$0 < X < 2$", Β = "$X = 2$" και C = "$X = 3/2$"

Έχουμε λοιπόν

α) Ελέγχουμε τις δύο ιδιότητες της pdf.

- $f(x) = 3(x^2 - 2x + 1) = 3(x-1)^2 \geq 0$ για κάθε $x \in R$.
- $\int_{-\infty}^{+\infty} f(x)dx = \int_1^2 (3x^2 - 6x + 3)dx = \left[x^3 - 3x^2 + 3x\right]_1^2 = (8 - 12 + 6) - (1 - 3 + 3) = 1$

Άρα η f είναι πράγματι pdf.

β) Έχουμε

$$P(X < \frac{3}{2}) = \int_{-\infty}^{3/2} f(x)dx = \int_1^{3/2} (3x^2 - 6x + 3)dx =$$

$$= \left[x^3 - 3x^2 + 3x\right]_1^{3/2} = (\frac{27}{8} - \frac{27}{4} + \frac{9}{2}) - (1 - 3 + 3) = \frac{9}{8} - 1 = \frac{1}{8}$$

Επίσης, $Y > 0 \Leftrightarrow 10X - 15 > 0 \Leftrightarrow X > \frac{3}{2}$. Δηλαδή το να είναι η Y θετική σημαίνει να είναι $X > \frac{3}{2}$ και άρα, από το προηγούμενο αποτέλεσμα

$$P(Y > 0) = P(X > \frac{3}{2}) = 1 - \frac{1}{8} = \frac{7}{8}$$

γ) Για τη μέση τιμή έχουμε

$$E(X) = \int_{-\infty}^{+\infty} xf(x)dx = \int_1^2 3x(x^2 - 2x + 1)dx = \int_1^2 (3x^3 - 6x^2 + 3x)dx =$$

$$= \left[3\frac{x^4}{4} - 2x^3 + 3\frac{x^2}{2}\right]_1^2 = (12 - 16 + 6) - (\frac{3}{4} - 2 + \frac{3}{2}) = 2 - \frac{1}{4} = \frac{7}{4}$$

δ) $E(\frac{1}{(X-1)^2}) = \int_{-\infty}^{+\infty} \frac{1}{(x-1)^2} f(x)dx = \int_1^2 \frac{3(x^2 - 2x + 1)}{(x-1)} dx = \int_1^2 3dx = \left[3x\right]_1^2 = 3$

ε) Το ενδεχόμενο A καλύπτει όλο το δειγματοχώρο, άρα $P(A) = 1$

Τα ενδεχόμενα B και C αφορούν σημεία σε μια συνεχή κατανομή, άρα $P(B) = 0$ και $P(C) = 0$

ΑΣΚΗΣΕΙΣ

3.1 Έστω ότι η Χ είναι διακριτή μεταβλητή που παίρνει τις τιμές 2,5,10 με αντίστοιχες πιθανότητες

$$p(2)=\frac{1}{2} \qquad p(5)=\frac{1}{4} \qquad p(10)=\frac{1}{4}$$

α) Να επιβεβαιώσετε ότι η p(x) είναι πράγματι συνάρτηση πιθανότητας.

β) Να βρεθούν η μέση τιμή E(X) και η διασπορά V(X) της Χ.

γ) Να βρεθεί η cdf.

δ) Ποια είναι η πιθανότητα να έχουμε $X \le 7$;

3.2 Τα αποτελέσματα σε μια εξεταστική του μαθήματος «Στοχαστικά Μοντέλα Ουρών Αναμονής» είχαν ως εξής

Βαθμός	**Αριθμός Γραπτών**	**Βαθμός**	**Αριθμός Γραπτών**
10	6		
9	27	4	15
8	16	3	16
7	22	2	6
6	22	1	10
5	23	0	3
Σύνολο επιτυχόντων	116	Σύνολο μη επιτυχόντων	50
Σύνολο Γραπτών 166			

α) Διαλέγουμε ένα γραπτό στην τύχη και παριστάνουμε με Χ το βαθμό του. Να περιγράψετε την τυχαία μεταβλητή Χ: (είναι διακριτή ή συνεχής; υπάρχει pdf; υπάρχει συνάρτηση πιθανότητας; Ποια είναι η cdf;)

β) Ποια είναι η μέση τιμή της βαθμολογίας; Βρείτε τη μέση τιμή E(X)

γ) Πόσο διάσπαρτη είναι η βαθμολογία; Βρείτε τη διασπορά V(X).

δ) Διαλέγουμε ένα γραπτό στην τύχη και στη συνέχεια ένα δεύτερο γραπτό. Ποια είναι η πιθανότητα να έχουν βαθμολογηθεί και τα δύο επιτυχώς;

3.3 Η τυχαία μεταβλητή Χ παίρνει τιμές στο διάστημα [0,4] με pdf

$$f(x)=\frac{1}{16}x+\frac{1}{8}, \quad x\in[0,4]$$

α) Επιβεβαιώστε ότι η f είναι pdf.

β) Να βρεθεί η πιθανότητα η τιμή X να παίρνει τιμές από 3 μέχρι 4.

γ) Να βρεθεί η πιθανότητα η τιμή X να παίρνει τιμές από 3 μέχρι 5.

δ) Ποιο είναι πιο πιθανό, η X πάρει τιμή στο διάστημα $[0,2]$ ή στο διάστημα $[2,4]$;

ε) Να βρεθεί η μέση τιμή και η διασπορά της X

στ) Να βρεθεί η μέση τιμή της X+1 και της 3X

3.4 Δίνεται η συνάρτηση

$$f(x)=\begin{cases}\dfrac{x^3+2x}{8}, & 0<x<2\\ 0, & \text{αλλού}\end{cases}$$

α) Να δειχθεί ότι η f είναι pdf μιας συνεχούς τυχαίας μεταβλητής X.

β) Να βρεθεί η μέση τιμή και η διασπορά της X.

γ) Σε ένα πρόγραμμα προσομοίωσης ενός δικτύου, μια δυαδική πηγή εκπέμπει τα σύμβολα

0, αν η X πάρει τιμή στο διάστημα $[0,1]$

1, αν η X πάρει τιμή στο διάστημα $[1,2]$

Ποια είναι η πιθανότητα του κάθε συμβόλου της πηγής;

3.5 Δίνεται η συνάρτηση

$$f(x)=\begin{cases}1/x, & 1<x<e\\ 0, & \text{αλλού}\end{cases}$$

α) Να δειχθεί ότι η f είναι pdf μιας συνεχούς τυχαίας μεταβλητής X.

β) Να βρεθεί η πιθανότητα η τιμή X να παίρνει τιμές από 1 μέχρι 2.

γ) Να βρεθούν η μέση τιμή E(X) και η διασπορά V(X) της X.

δ) Σε ποιο σημείο πρέπει να χωρίσουμε το διάστημα $(1,e)$ έτσι ώστε τα δύο υποδιαστήματα που προκύπτουν να είναι ισοπίθανα;

3.6 Δίνεται η συνάρτηση

$$f(x)=\begin{cases}x, & 0\le x<1\\ 2-x, & 1\le x\le 2\end{cases}$$

α) Να δειχθεί ότι η f(x) είναι pdf μιας συνεχούς τυχαίας μεταβλητής X.

β) Να βρεθούν η μέση τιμή E(X) και η διασπορά V(X) της X.

γ) Να βρεθούν οι πιθανότητες

η X πάρει τιμή στο διάστημα $[0,1/2]$, η X πάρει τιμή στο διάστημα $[1/2,2]$

3.7 Δίνεται η συνάρτηση

$$f(x)=\begin{cases} ax-ax^2, & 0<x<1 \\ 0, & \text{αλλού} \end{cases}$$

α) Να προσδιοριστεί η σταθερά α ώστε η f(x) να είναι pdf μιας συνεχούς τυχαίας μεταβλητής X.

β) Να βρεθεί η πιθανότητα η τιμή X να παίρνει τιμές από 0 μέχρι $1/2$.

γ) Να βρεθούν η μέση τιμή E(X) και η διασπορά V(X) της X.

3.8 Δίνεται η συνάρτηση $f(x)=\frac{1}{6}x$ στο διάστημα (2,α)

α) Να προσδιοριστεί η σταθερά α ώστε η f(x) να είναι pdf μιας συνεχούς τυχαίας μεταβλητής X.

β) Να βρεθεί η μέση τιμή E(X) και η διασπορά V(X) της X.

γ) Να βρεθεί η cdf της X.

δ) Σε ένα πρόγραμμα προσομοίωσης ενός δικτύου, μια δυαδική πηγή εκπέμπει τα σύμβολα

0, αν η X πάρει τιμή στο διάστημα $[2,3]$

1, αν η X πάρει τιμή στο διάστημα $[3,a]$

Ποια είναι η πιθανότητα του κάθε συμβόλου της πηγής;

ε) Ποια είναι η πιθανότητα να σταλεί η ακολουθία **001**

3.9 Δίνεται η διακριτή τυχαία μεταβλητή X που παίρνει τις άπειρες τιμές

$$x_1,\ x_2,\ x_{3,}\ \ x_{4,}\ \ldots$$

με συνάρτηση πιθανότητας $p(x_i)=\frac{1}{2^i}$, δηλαδή με αντίστοιχες πιθανότητες

$$\frac{1}{2},\ \frac{1}{4},\ \frac{1}{8},\ \frac{1}{16},\ \ldots$$

α) να επιβεβαιώσετε ότι πρόκειται για συνάρτηση πιθανότητας

β) υπάρχει μέση τιμή της X αν $x_i=2^i$ για κάθε i; Γιατί;

γ) Ποια είναι μέση τιμή της X αν $x_i=\frac{1}{2^i}$ για κάθε i; Ποια είναι η διασπορά;

(υπενθυμίζουμε ότι το άθροισμα απείρων όρων γεωμετρικής σειράς με πρώτο όρο a και λόγο λ είναι $\frac{\alpha}{1-\lambda}$)

ΣΥΝΤΟΜΕΣ ΑΠΑΝΤΗΣΕΙΣ ΤΩΝ ΑΣΚΗΣΕΩΝ

3.1 β) E(X)=4.75 και V(X)=171/16=10.68

δ) 0.75

3.2 β) E(X)=5.79 δ) $\frac{115}{166}\cdot\frac{116}{166}$

3.3 β) και γ) 11/32

δ) στο [2,4]

ε) E(X)=7/3 και V(X)=11/9

στ) E(X+1)=10/3 και E(3X)=7

3.4 β) E(X)=22/15 και V(X)=7/3-(22/15)^2

γ) P(**0**)=5/32 και P(**1**)=27/32

3.5 α) Παρατηρήστε ότι $\int_1^e \frac{1}{x}dx = [\ln x]_1^e = 1$

β) ln2

γ) E(X)$= e-1$ και V(X)$= \frac{4e-e^2-3}{2}$

δ) στο $\sqrt{e}$ =1.648

3.6 β) E(X)=1 και V(X)=1/6

γ) 1/8 και 7/8

3.7 α) α=6 β) 1/2 γ) E(X)=1/2 και V(X)=1/20

3.8 α) α=4 β) E(X)=28/9 και V(X)=26/81

δ) P(**0**)=5/12 και p(**1**)=7/12

ε) p(**001**)=175/1728

3.9 β) Όχι, δεν ικανοποιείται η προϋπόθεση $\sum_i |x_i| p(x_i) < +\infty$

γ) E(X)=1/3 και V(X)=2/63

4. *ΟΡΙΣΜΕΝΑ ΠΑΡΑΔΕΙΓΜΑΤΑ ΚΑΤΑΝΟΜΩΝ*

Υπάρχει μια πληθώρα εφαρμογών στην καθημερινή μας ζωή όπου κάποιο βασικό μέγεθος ακολουθεί τυχαία κατανομή και η μελέτη της κατανομής αυτής οδηγεί σε χρήσιμα συμπεράσματα. Για καλή μας τύχη, οι εφαρμογές αυτές μπορούν να ομαδοποιηθούν, καθώς πολλά από τα χαρακτηριστικά που επηρεάζουν την κατανομή είναι παρόμοια. Έτσι, για μια ομάδα προβλημάτων που έχουν αρκετά κοινά χαρακτηριστικά μελετάμε ένα γενικό μοντέλο κατανομής και βγάζουμε συμπεράσματα για όλα τα προβλήματα που ακολουθούν αυτό το μοντέλο, αγνοώντας επιμέρους ιδιαιτερότητες που μπορεί να παρουσιάζει η κάθε εφαρμογή.

Στο κεφάλαιο αυτό θα παρουσιάσουμε ενδεικτικά ορισμένες κατανομές ώστε να εξοικειωθούμε με το αντικείμενο. Μέσα σ’ αυτές είναι και δύο κατανομές που παρουσιάζονται συχνά σε εφαρμογές τηλεπικοινωνιακών δικτύων και εξυπηρετούν το σκοπό μας. Από κει και πέρα ο ενδιαφερόμενος αναγνώστης μπορεί να βρει αρκετές άλλες κατανομές στη σχετική βιβλιογραφία.

Οι κατανομές που θα δούμε είναι οι εξής

ΔΙΑΚΡΙΤΕΣ	ΣΥΝΕΧΕΙΣ
Διωνυμική	Ομοιόμορφη
Poisson	Κανονική
	Εκθετική

Θα ξεκινήσουμε με τη Διωνυμική και την Ομοιόμορφη κατανομή για να δούμε από ένα εύκολο παράδειγμα κάθε κατηγορίας. Θα συνεχίσουμε με την Κανονική κατανομή που παρουσιάζεται πολύ συχνά στη φύση αλλά και στις τηλεπικοινωνίες ειδικότερα (πχ περιγραφή του τυχαίου θορύβου σε ένα κανάλι). Και θα κλείσουμε με την κατανομή Poisson και την Εκθετική κατανομή οι οποίες εμφανίζονται στα κυριότερα μοντέλα ουρών αναμονής και θα τις χρειαστούμε ιδιαίτερα στο επόμενο Κεφάλαιο.

4.1 Η ΔΙΩΝΥΜΙΚΗ ΚΑΤΑΝΟΜΗ

Ας ξεκινήσουμε με ένα παράδειγμα λύνοντάς το με τον κλασικό τρόπο που γνωρίζαμε ως τώρα και στη συνέχεια θα δούμε πως τα ίδιο αποτέλεσμα θα το παίρνουμε πολύ πιο εύκολα με έναν απλό τύπο. Ρίχνουμε ένα ζάρι 5 φορές και μετράμε τις εξάρες. Προφανώς μπορεί να πετύχουμε 0,1,2,3,4 ή και 5 εξάρες. Θέτουμε το ερώτημα: Ποια είναι η πιθανότητα να πετύχουμε 2 εξάρες;

Η πιθανότητα σε μια ρίψη να φέρουμε εξάρα είναι $p=\frac{1}{6}$ ενώ να μην φέρουμε εξάρα είναι $1-p=\frac{5}{6}$. Η πιθανότητα για παράδειγμα να φέρουμε εξάρα τις δύο πρώτες φορές και κάτι άλλο τις επόμενες τρεις είναι

$$\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^3$$

Ναι αλλά δεν είναι απαραίτητο να έρθουν οι δύο εξάρες στην αρχή, υπάρχουν και άλλοι τρόποι να φέρουμε δύο εξάρες μέσα στις 5 ρίψεις. Πόσοι είναι αυτοί οι τρόποι; Πρόκειται για τον συνδυασμό «5 ανά 2». Άρα υπάρχουν

$$\binom{5}{2}=\frac{5!}{2!\cdot 3!}=10 \text{ τρόποι}$$

Πράγματι, οι τρόποι είναι:

66*, 6*6**, 6**6*, 6***6, *66**, *6*6*, *6**6, **66*, **6*6, ***66**

Άρα η πιθανότητα να πετύχουμε 2 εξάρες στις 5 ρίψεις είναι

$$\binom{5}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^3$$

Είμαστε έτοιμοι τώρα να παρουσιάσουμε τη διωνυμική κατανομή.

Η Διωνυμική κατανομή είναι διακριτή με συνάρτηση πιθανότητας

$$p(x)=\binom{N}{x}p^x(1-p)^{N-x}, \qquad x=0,1,2,\ldots,N \qquad (*)$$

όπου τα *p* και *N* είναι δεδομένα. Δείχνεται ότι

$$E(X)=Np \qquad V(X)=Np(1-p)$$

ΠΕΡΙΓΡΑΦΗ ΠΡΟΒΛΗΜΑΤΟΣ: Ας υποθέσουμε ότι ένα πείραμα έχει δύο αποτελέσματα:

«επιτυχία» με πιθανότητα p

«αποτυχία» με πιθανότητα $1-p$

π.χ. Σε μία απλή ρίψη ενός ζαριού μπορούμε να ονομάσουμε την ΕΞΑΡΑ ως «επιτυχία» και κάθε άλλο αποτέλεσμα ως «αποτυχία». Τότε

$$p=\frac{1}{6} \qquad \text{και} \qquad 1-p=\frac{5}{6}$$

Θέτουμε το ερώτημα: Αν επαναλάβουμε το πείραμα *N* φορές, πόσες φορές αναμένουμε να έχουμε «επιτυχία»; Συμβολίζουμε με *X* τη μεταβλητή που εκφράζει τον αριθμό των «επιτυχιών». Λαμβάνει τις τιμές

0,1,2,3,...,N

Η πιθανότητα να έχουμε *x* επιτυχίες, δηλαδή η *P(X=x)* δίνεται από τον τύπο (*) πιο πάνω.

ΠΑΡΑΔΕΙΓΜΑ 1

Ρίχνουμε ένα ζάρι 5 φορές.

α) Ποια είναι η πιθανότητα να φέρουμε 2 ΕΞΑΡΕΣ;

β) Ποια είναι η πιθανότητα να φέρουμε το πολύ 2 ΕΞΑΡΕΣ;

γ) Ποια είναι η πιθανότητα να φέρουμε τουλάχιστον 3 ΕΞΑΡΕΣ;

Θεωρούμε ως «επιτυχία» την ΕΞΑΡΑ, οπότε

$$p=\frac{1}{6} \qquad 1-p=\frac{5}{6} \qquad N=5$$

α) Προφανώς εδώ ζητάμε την πιθανότητα $P(X=2)$. Είναι

$$P(X=2)=p(2)=\binom{N}{k}p^k(1-p)^{N-k}=\binom{5}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^3=\frac{5!}{2!3!}\frac{1}{6^2}\frac{5^3}{6^3}=\frac{1250}{7776}.$$

Η πιθανότητα είναι λοιπόν περίπου 0.16=16%

β) Εδώ ζητάμε την πιθανότητα $P(X\leq 2)$. Υπολογίζουμε με τον ίδιο τρόπο τις πιθανότητες να φέρουμε καμία, μία ή δύο ΕΞΑΡΕΣ και έχουμε

$$P(X \le 2) = p(0)+p(1)+p(2) = \frac{3125}{7776}+\frac{3125}{7776}+\frac{1250}{7776}=\frac{7500}{7776}.$$

γ) Όμοια

$$P(X \ge 3) = p(3)+p(4)+p(5) = \frac{250}{7776}+\frac{25}{7776}+\frac{1}{7776}=\frac{276}{7776}.$$

Σκεπτόμενοι όμως πιο πονηρά θα μπορούσαμε να πούμε

$$P(X \ge 3) = 1 - P(X \le 2) = 1 - \frac{7500}{7776} = \frac{276}{7776}$$

<u>ΠΑΡΑΔΕΙΓΜΑ 2</u>

Ρίχνουμε ένα νόμισμα έξι φορές.

α) Ποια είναι η πιθανότητα να φέρουμε 2 ΚΕΦΑΛΕΣ;

β) Ποια είναι η πιθανότητα να φέρουμε τουλάχιστον 4 ΚΕΦΑΛΕΣ;

γ) Ποια είναι η πιθανότητα να φέρουμε το πολύ 3 ΚΕΦΑΛΕΣ;

Θεωρούμε ως «επιτυχία» την ΚΕΦΑΛΗ, οπότε

$$p=\frac{1}{2} \qquad 1-p=\frac{1}{2} \qquad N=6$$

α) Προφανώς εδώ ζητάμε την πιθανότητα $P(X=2)$. Είναι

$$P(X=2)=p(2)=\binom{N}{k}p^k(1-p)^{N-k}=\binom{6}{2}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^4=\frac{6!}{2!\cdot 4!}\frac{1}{4}\frac{1}{16}=\frac{15}{64}.$$

β) Υπολογίζουμε τις τιμές $p(4), p(5), p(6)$ με τον ίδιο τρόπο (κάντε το για άσκηση). Οπότε

$$P(X \ge 4) = p(4)+p(5)+p(6)=\frac{15}{64}+\frac{6}{64}+\frac{1}{64}=\frac{22}{64}=\frac{11}{32}.$$

γ) Θα μπορούσαμε να υπολογίσουμε

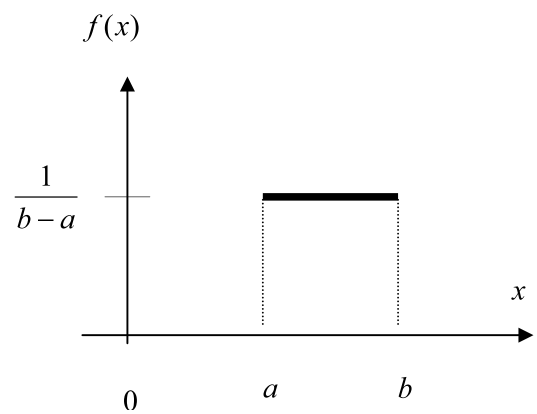
$$P(X \le 3) = p(0)+p(1)+p(2)+p(3)$$

αλλά εφόσον γνωρίζουμε ότι $P(X \ge 4)=\frac{11}{32}$, βρίσκουμε

$$P(X \le 3) = 1 - P(X \ge 4) = 1 - \frac{11}{32} = \frac{21}{32}.$$

4.2 Η ΟΜΟΙΟΜΟΡΦΗ ΚΑΤΑΝΟΜΗ

Πρόκειται για συνεχή κατανομή με pdf μια σταθερή συνάρτηση *f(x)=c* σε κάποιο διάστημα *[a,b]*. Η σταθερή τιμή πρέπει να είναι ίση με c=*1/(b-a)* ώστε να εξασφαλίζονται οι ιδιότητες της pdf (βλέπε ιδιότητα β. πιο κάτω). Δηλαδή,



$$f(x) = \begin{cases} \dfrac{1}{b-a}, & a \le x \le b \\ 0, & \alpha\lambda\lambda\acute{\omega}\varsigma \end{cases}$$

Η *f* είναι πράγματι pdf διότι,

a) $f(x) \ge 0$

b) $\int_{-\infty}^{+\infty} f(x)dx = \int_a^b \frac{1}{b-a}dx = \frac{1}{b-a}\left[x\right]_a^b = \frac{1}{b-a}(b-a) = 1$

Για τη μέση τιμή και τη διασπορά έχουμε

$$E(X) = \frac{a+b}{2} \qquad V(X) = \frac{(b-a)^2}{12}$$

Πράγματι,

$$E(X) = \int_{-\infty}^{+\infty} xf(x)dx = \int_a^b x\frac{1}{b-a}dx = \frac{1}{b-a}\left[\frac{x^2}{2}\right]_a^b = \frac{1}{b-a}(\frac{b^2-a^2}{2}) = \frac{b+a}{2}$$

Επίσης,

$$E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x)dx = \int_a^b x^2\frac{1}{b-a}dx = \frac{1}{b-a}\left[\frac{x^3}{3}\right]_a^b = \frac{1}{b-a}(\frac{b^3-a^3}{3}) = \frac{b^2+ba+a^2}{3}$$

οπότε

$$V(X) = E(X^2) - \mu^2 = \frac{b^2+ba+a^2}{3} - \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)^2}{12}$$

<u>ΠΑΡΑΔΕΙΓΜΑ 3</u>

Διαλέγουμε τυχαία ένα σημείο επάνω στο ευθύγραμμο τμήμα *[0,2].* Ποια είναι η πιθανότητα να βρίσκεται ανάμεσα στο 1 και το 1.5;

Η τυχαία μεταβλητή Χ που εκφράζει το επιλεγόμενο σημείο είναι ομοιόμορφα κατανεμημένη με

$$f(x) = \frac{1}{b-a}, \qquad \text{όπου } a = 0 \text{ και } b = 2.$$

δηλαδή $f(x) = \frac{1}{2}$.

Οπότε,

$$P(1 \le X \le 1.5) = \int_1^{1.5} \frac{1}{2} dx = \frac{1}{2}\left[x\right]_1^{1.5} = \frac{1}{2}(1.5-1) = 0.25.$$

<u>Παρατήρηση</u>: Εδώ βέβαια ο υπολογισμός μπορεί να γίνει και γεωμετρικά. Εφόσον το Χ κατανέμεται ομοιόμορφα, θα μπορούσαμε να σκεφτούμε πιο απλά

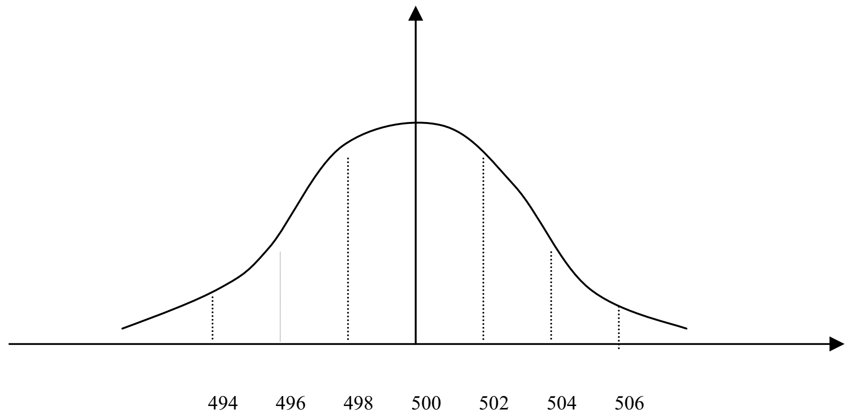
$$P(1 \le X \le 1.5) = \frac{\text{μήκος τμήματος } [1,1.5]}{\text{μήκος τμήματος } [0,2]} = \frac{1.5-1}{2-0} = 0.25$$

4.3 Η ΚΑΝΟΝΙΚΗ ΚΑΤΑΝΟΜΗ (Gauss)

Μία από τις πιο σημαντικές συνεχείς κατανομές είναι η κανονική (είναι γνωστή και ως *Gaussian*). Πολλά μεγέθη στη φύση ακολουθούν αυτήν την κατανομή, όπως το ύψος των ανθρώπων, το βάρος των ανθρώπων, ο θόρυβος που επηρεάζει τη μετάδοση ενός σήματος, το πραγματικό βάρος του καφέ σε μια συσκευασία των 500γρ, κλπ. Σε όλα αυτά τα παραδείγματα υπάρχει μια «κυρίαρχη» μέση τιμή, ενώ τα αντίστοιχα μεγέθη «απλώνονται» συμμετρικά γύρω από αυτήν την μέση τιμή.

Ας μείνουμε για παράδειγμα στην τελευταία περίπτωση του καφέ. Μια συσκευασία των 500γρ περιέχει κατά μέσο όρο 500γρ καφέ (υποθέτουμε ότι γενικά δεν μας κλέβει ο προμηθευτής!). Στην πραγματικότητα όμως, αν εξετάσουμε διάφορα δείγματα συσκευασιών θα διαπιστώσουμε ότι άλλες είναι λίγο βαρύτερες κι άλλες λίγο ελαφρύτερες.

Αν αποτυπώσουμε τις παρατηρήσεις μας σε ένα διάγραμμα θα δούμε



Δηλαδή αρκετές συσκευασίες ζυγίζουν γύρω στα 500γρ ενώ όσο απομακρυνόμαστε από τα 500γρ (είτε προς τα πάνω, είτε προς τα κάτω) τόσο πιο απίθανο είναι να βρούμε μια συσκευασία με αυτό το βάρος. Πρόκειται για ένα παράδειγμα όπου το παρατηρούμενο βάρος ακολουθεί όπως λέμε «κανονική κατανομή»

Συνήθως, όταν ένα μέγεθος επηρεάζεται από πολλούς τυχαίους παράγοντες ακολουθεί κανονική κατανομή. Η τυχαία μεταβλητή X της κανονικής κατανομής έχει pdf

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

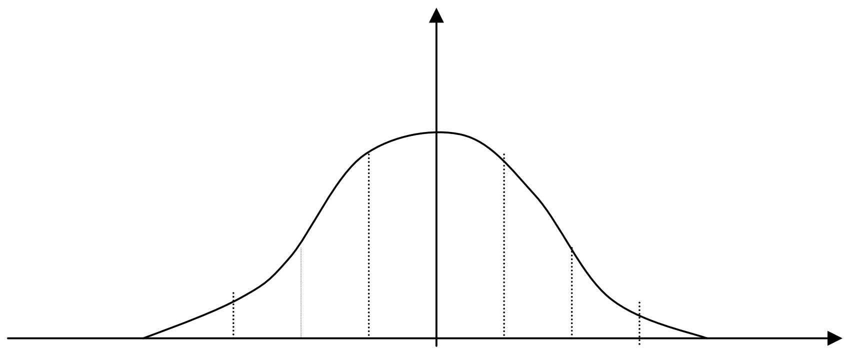
όπου $\mu=E(X)$ είναι η μέση τιμή και σ η τυπική απόκλιση.

Συμβολίζεται με $N(\mu,\sigma^2)$ και για να δηλώσουμε ότι η X ακολουθεί την κανονική κατανομή γράφουμε

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

ΠΑΡΑΤΗΡΗΣΗ: Για να προλάβουμε οποιαδήποτε δυσάρεστη έκπληξη λόγω της πολυπλοκότητας της pdf, να πούμε ότι στην πράξη δε θα χρειαστεί να υπολογίζουμε ολοκληρώματα της $f(x)$ όπως κάναμε ως τώρα. Υπάρχει ένας έτοιμος πίνακας που μας δίνει κάθε φορά τις ζητούμενες τιμές. Πιο κάτω θα πούμε πως χρησιμοποιούμε τον πίνακα αυτόν.

Ξεχωρίζουμε την κανονική κατανομή $N(0,1)$. Στην κατανομή αυτή η τυχαία μεταβλητή X λέγεται τυποποιημένη, παίρνει θετικές και αρνητικές τιμές γύρω από τη μέση τιμή *μ=0,* ενώ η τυπική απόκλιση που δείχνει πόσο «απλώνονται» οι τιμές της γύρω από τη μέση τιμή είναι *σ=1.* Η κατανομή πιθανοτήτων για την τυποποιημένη μεταβλητή *X* φαίνεται στο παρακάτω γράφημα



Στις εφαρμογές είναι βολικό να μετασχηματίζουμε την οποιαδήποτε τυχαία μεταβλητή *X* που ακολουθεί την κανονική κατανομή $N(\mu,\sigma^2)$ σε μια μεταβλητή *Z* σύμφωνα με τον τύπο

$$Z = \frac{X - \mu}{\sigma}$$

Τότε η *Z* ακολουθεί την κανονική κατανομή *N(0,1)*. Η νέα μεταβλητή *Z* λέγεται **τυποποιημένη**. Ας δείξουμε με ένα παράδειγμα πως εφαρμόζεται ο μετασχηματισμός αυτός.

ΠΑΡΑΔΕΙΓΜΑ 4

Έστω ότι η μεταβλητή X που εκφράζει το βάρος του καφέ στο αρχικό παράδειγμα ακολουθεί κανονική κατανομή με μ=500 και σ=4, δηλαδή

$$X \sim N(500,16)$$

Ας υποθέσουμε ότι ζητάμε τις πιθανότητες

- μια συσκευασία να ζυγίζει κάτω από 508γρ, δηλαδή την πιθανότητα

$$P(X < 508)$$

και

- μια συσκευασία να ζυγίζει από 492γρ μέχρι 504γρ, δηλαδή την πιθανότητα

$$P(492 < X < 504).$$

Χρησιμοποιούμε το μετασχηματισμό

$$Z = \frac{X - 500}{4}$$

και τον εφαρμόζουμε σε όλα τα μέλη της εκάστοτε ανισότητας.
Έτσι

$$X < 508 \quad \Leftrightarrow \quad \frac{X - 500}{4} < \frac{508 - 500}{4} \quad \Leftrightarrow \quad Z < 2$$

και για τη ζητούμενη πιθανότητα ισχύει

$$P(X < 508) = P(Z < 2).$$

Επίσης

$$492 < X < 504 \quad \Leftrightarrow \quad \frac{492 - 500}{4} < \frac{X - 500}{4} < \frac{504 - 500}{4} \quad \Leftrightarrow \quad -2 < Z < 1$$

και για τη ζητούμενη πιθανότητα ισχύει

$$P(492 < X < 504) = P(-2 < Z < 1).$$

Τις πιθανότητες $P(Z < 2)$ και $P(-2 < Z < 1)$ θα τις υπολογίσουμε παρακάτω.

Στο παραπάνω διάγραμμα της τυποποιημένης μεταβλητής *Z* ισχύουν τα εξής:

- Η καμπύλη είναι συμμετρική ως προς τον κάθετο άξονα, δηλαδή οι τιμές της κατανέμονται συμμετρικά ως προς τη μέση τιμή μ=0.
- Η συνολική επιφάνεια ανάμεσα στην καμπύλη και τον οριζόντιο άξονα έχει εμβαδόν ίσο με 1:

$$P(-\infty < Z < +\infty) = \int_{-\infty}^{+\infty} f(x)dx = 1$$

  (0.5 στο θετικό κομμάτι και 0.5 στο αρνητικό εφόσον η καμπύλη είναι συμμετρική)

Με αριθμητικές μεθόδους υπολογίζεται ότι

- Στο διάστημα (-1,1) βρίσκεται το 68.27% των τιμών, δηλαδή
$$P(-1 < Z < 1) = 0.6827$$
- Στο διάστημα (-2,2) βρίσκεται το 95.45% των τιμών, δηλαδή
$$P(-2 < Z < 2) = 0.9545$$
- Στο διάστημα (-3,3) βρίσκεται το 99.73% των τιμών -σχεδόν όλες, δηλαδή
$$P(-3 < Z < 3) = 0.9973$$

Θυμηθείτε όμως ότι
$$Z = \frac{X-\mu}{\sigma}$$
οπότε
$$-1 < Z < 1 \Leftrightarrow \quad -1 < \frac{X-\mu}{\sigma} < 1 \Leftrightarrow -\sigma < \mathrm{X} - \mu < \sigma \Leftrightarrow \quad \mu - \sigma < \mathrm{X} < \mu + \sigma$$
Με τον ίδιο τρόπο,

$-2 < Z < 2$ ισοδυναμεί με $\mu - 2\sigma < \mathrm{X} < \mu + 2\sigma$

$-3 < Z < 3$ ισοδυναμεί με $\mu - 3\sigma < \mathrm{X} < \mu + 3\sigma$

Πρακτικά αυτό σημαίνει ότι, αν $X \sim N(\mu, \sigma^2)$, τότε

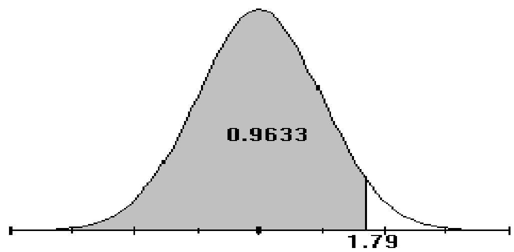
- είμαστε 68.27% σίγουροι ότι η *X* βρίσκεται στο διάστημα *(μ-σ,μ+σ)*.
- είμαστε 95.45% σίγουροι ότι η *X* βρίσκεται στο διάστημα *(μ-2σ,μ+2σ)*.
- είμαστε 99.73% σίγουροι ότι η *X* βρίσκεται στο διάστημα *(μ-3σ,μ+3σ)*.

Στο Παράδειγμα 4 συγκεκριμένα

- Η πιθανότητα να ζυγίζει μια συσκευασία 496γρ - 504γρ είναι 68.27%
- Η πιθανότητα να ζυγίζει μια συσκευασία 492γρ - 508γρ είναι 95.45%
- Η πιθανότητα να ζυγίζει μια συσκευασία 488γρ - 512γρ είναι 99.73%

Για να υπολογίσουμε την πιθανότητα για οποιοδήποτε άλλο διάστημα χρησιμοποιούμε τον Πίνακα του παραρτήματος που βρίσκεται στο τέλος του συγγράμματος, ο οποίος δίνει τις τιμές της cdf της τυποποιημένης κανονικής κατανομής
$$F(a) = P(Z < a)$$

Για παράδειγμα, ο πίνακας δίνει F(1.79) = 0.9633, που σημαίνει ότι η πιθανότητα να ισχύει $Z < 1.79$ είναι 96,33%



Πρέπει να σημειωθεί ότι ο πίνακας δίνει τις τιμές της F για $a \geq 0$. Λόγω της συμμετρίας της γραφικής παράστασης της κανονικής κατανομής γύρω από τη μέση τιμή 0, μπορούμε εύκολα να υπολογίσουμε την F και για αρνητικές τιμές

$$F(-a) = P(Z < -a) = P(Z > a) = 1 - P(Z < a) = 1 - F(a)$$

Έτσι πχ F(-1.79) = 1 - F(1.79) = 1 - 0.9633 = 0.0367

Συνοψίζοντας, για να υπολογίσουμε την πιθανότητα σε ένα διάστημα της τυχαίας μεταβλητής *X*, μετασχηματίζουμε πρώτα την μεταβλητή *X* στην τυποποιημένη μεταβλητή *Z* και κατόπιν να χρησιμοποιούμε το σχετικό πίνακα.

ΠΑΡΑΔΕΙΓΜΑ 5

α) $P(Z < 1.2) = F(1.2) = 0.8849$

(στην πρώτη στήλη του πίνακα αναζητούμε την τιμή 1.2. Δίπλα βρίσκουμε την ζητούμενη πιθανότητα)

β) $P(Z < 1.23) = F(1.23) = 0.8907$

(εφόσον έχουμε και δεύτερο δεκαδικό ψηφίο, το 3, θα αναζητήσουμε και πάλι την γραμμή της τιμής 1.2, αλλά αυτή τη φορά θα κοιτάξουμε τη στήλη .03)

γ) $P(Z > 1.2) = 1 - P(Z < 1.2) = 1 - 0.8849 = 0.1151$

δ) $P(-1.23 < Z) = P(Z < 1.23) = 0.8907$

(λόγω συμμετρίας της γραφικής παράστασης γύρω από το 0)

ε) $P(0.81 < Z < 1.94) = F(1.94) - F(0.81) = 0.9738 - 0.7910 = 0.1828$

στ) $P(-0.46 < Z < 2.21) = F(2.21) - F(-0.46) = F(2.21) - 1 + F(0.46) =$

$$= 0.6772 - 1 + 0.9864 = 0.6636$$

ζ) στο Παράδειγμα 4, αφήσαμε μια εκκρεμότητα. Συμπληρώνουμε τώρα:

$P(Z < 0) = F(0) = 0.5$

$P(-2 < Z < 1) = F(1) - F(-2) = F(1) - 1 + F(2) = 0.8413 + 0.9772 - 1 = 0.8115$

ΠΑΡΑΔΕΙΓΜΑ 6

Το μέσο βάρος 500 σπουδαστών του ΤΕΙ Λάρισας είναι 75.75 κιλά (προφανώς το δείγμα δεν είχε σπουδάστριες!) ενώ η τυπική απόκλιση 7.5. Με την προϋπόθεση ότι το βάρος ακολουθεί κανονική κατανομή να υπολογίσετε πόσοι σπουδαστές ζυγίζουν

α) μεταξύ 60 και 78 κιλών β) πάνω από 92 κιλά.

Έστω ότι η μεταβλητή *X* παριστάνει το βάρος ενός φοιτητή. Έχουμε ότι

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

με

$$\mu = 75.75 \qquad \sigma = 7.5$$

Χρησιμοποιούμε το μετασχηματισμό

$$Z = \frac{X - 75.75}{7.5}$$

οπότε

α) Έχουμε

$$P(60 < X < 78) = P(\frac{60 - 75.75}{7.5} < \frac{X - 75.75}{7.5} < \frac{78 - 75.75}{7.5})$$

$$= P(-2.1 < Z < 0.3) = F(0.3) - F(-2.1) = F(0.3) + F(2.1) - 1 = 0.6$$

Συνεπώς, το 60% των σπουδαστών, δηλαδή 300 σπουδαστές (από τους 500) ζυγίζουν μεταξύ 60 και 78 κιλών.

β) Έχουμε

$$P(X > 92) = P(\frac{X - 75.75}{7.5} > \frac{92 - 75.75}{7.5}) = P(Z > 2.17) = 1 - F(2.17)$$
$$= 1 - 0.9850 = 0.015 = 1.5\%$$

Συνεπώς, το 1.5% των σπουδαστών, δηλαδή 7.5! (ας πούμε 7) σπουδαστές ζυγίζουν πάνω από 92 κιλά.

4.4 Η ΚΑΤΑΝΟΜΗ POISSON

Πρόκειται για διακριτή τυχαία μεταβλητή με συνάρτηση πιθανότητας

$$p(x) = \frac{e^{-\lambda}\lambda^x}{x!} \qquad x = 0,1,2,3,\ldots$$

όπου λ είναι μια παράμετρος.

<u>Σημείωση</u>: Πριν πούμε οτιδήποτε να υπενθυμίσουμε ότι η εκθετική συνάρτηση e^x αναλύεται σε σειρά ως εξής

$$e^x = \frac{x^0}{0!} + \frac{x^1}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

Η *p(x)* είναι όντως συνάρτηση πιθανότητας καθώς

α) $p(x) \geq 0$

β) $\sum_{x=0}^{+\infty} p(x) = p(0) + p(1) + p(2) + \cdots = e^{-\lambda}\frac{\lambda^0}{0!} + e^{-\lambda}\frac{\lambda^1}{1!} + e^{-\lambda}\frac{\lambda^2}{2!} + \cdots$

$$= e^{-\lambda}\left(1 + \frac{\lambda^1}{1!} + \frac{\lambda^2}{2!} + \cdots\right) = e^{-\lambda}e^{\lambda} = 1$$

Για τη μέση τιμή και τη διασπορά της X έχουμε

$$E(X) = \lambda \qquad\qquad V(X) = \lambda$$

Πράγματι,

$$E(X) = \sum_{x=0}^{+\infty} xp(x) = 0 \cdot e^{-\lambda} \frac{\lambda^0}{0!} + 1 \cdot e^{-\lambda} \frac{\lambda^1}{1!} + 2 \cdot e^{-\lambda} \frac{\lambda^2}{2!} + \cdots$$

$$= e^{-\lambda} \lambda \left(1 + \frac{\lambda^1}{1!} + \frac{\lambda^2}{2!} + \cdots \right) = e^{-\lambda} \lambda \, e^{\lambda} = \lambda$$

Με τον ίδιο τρόπο βρίσκουμε (ο υπολογισμός αφήνεται για άσκηση)

$$E(X^2) = \lambda + \lambda^2$$

οπότε,

$$V(X) = E(X^2) - \mu^2 = \lambda + \lambda^2 - \lambda^2 = \lambda$$

ΠΕΡΙΓΡΑΦΗ: Η κατανομή Poisson βρίσκει εφαρμογή σε αρκετά προβλήματα. Αναφέρουμε μερικά παραδείγματα όπου η μεταβλητή X εκφράζει

1. τον αριθμό των τηλεφωνικών κλήσεων σε ένα τηλεφωνικό κέντρο
2. τον αριθμό των ατυχημάτων σε ορισμένο χρονικό διάστημα
3. τον αριθμό των τυπογραφικών λαθών σε ένα βιβλίο
4. τον αριθμό των βακτηρίων σε ένα cm^3.

Γενικά εμφανίζεται όταν έχουμε έναν αριθμό παρατηρήσεων σε κάποιο διάστημα (το οποίο μπορεί να είναι χρονικό διάστημα, απόσταση, όγκος κλπ). Οι παρατηρήσεις αυτές πρέπει να ικανοποιούν ορισμένες προϋποθέσεις. Τις αναφέρουμε για το ενδεικτικό παράδειγμα των τηλεφωνικών κλήσεων σε ένα χρονικό διάστημα:

- η πιθανότητα να υπάρχει κλήση σε ένα χρονικό διάστημα (t_1,t_2) είναι ανάλογη (προσεγγιστικά) με το μήκος του διαστήματος $\delta=t_2-t_1$.
- Η πιθανότητα να έχουμε δύο κλήσεις σε ένα μικρό χρονικό διάστημα είναι αμελητέα.

- Ο αριθμός των κλήσεων που πραγματοποιούνται σε δύο ξένα μεταξύ τους χρονικά διαστήματα είναι ανεξάρτητα γεγονότα.

Αν η συχνότητα των κλήσεων είναι v (κλήσεις ανά μονάδα χρόνου), ο αριθμός X_t των κλήσεων στο διάστημα *[0,t]* ακολουθεί κατανομή Poisson με συνάρτηση πιθανότητας

$$p_t(x) = \frac{e^{-\lambda}\lambda^x}{x!} \qquad\qquad \lambda = vt$$

Προσοχή! Για κάθε τιμή *t* έχουμε μια διαφορετική κατανομή X_t.

ΠΑΡΑΔΕΙΓΜΑ 7

Σε ένα τηλεφωνικό κέντρο γίνονται κατά μέσο όρο 300 κλήσεις την ώρα. Να βρεθεί η πιθανότητα

α) να πραγματοποιηθούν 3 κλήσεις κατά τη διάρκεια 2 λεπτών

β) να πραγματοποιηθούν 6 κλήσεις κατά τη διάρκεια 4 λεπτών

Έστω X_t ο αριθμός των κλήσεων σε χρονικό διάστημα *t* (λεπτών)

Η συχνότητα των κλήσεων είναι

$v =$ 300 κλήσεις/ώρα = 5 κλήσεις/λεπτό

Η παράμετρος της κατανομής είναι $\lambda = 5t$ (εξαρτάται κάθε φορά από το χρόνο *t*).

α) Έχουμε την κατανομή X_t με $t = 2$, οπότε $\lambda = 10$. Ζητάμε την πιθανότητα να είναι X_t=3, δηλαδή την

$$p_2(3) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-10}10^3}{3!} = \frac{1000}{6}e^{-10} \approx 0.0075$$

β) Εδώ έχουμε την κατανομή X_t με $t = 4$, οπότε $\lambda = 20$. Ζητάμε την πιθανότητα να είναι X_t=6, δηλαδή την

$$p_4(6) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-20}20^6}{6!} \approx 0.00018$$

ΠΑΡΑΔΕΙΓΜΑ 8

Σε ένα βιβλίο υπάρχει κατά μέσο όρο 1 ορθογραφικό λάθος ανά 2 σελίδες. Ποια είναι η πιθανότητα να μην υπάρχουν λάθη σε 5 σελίδες του βιβλίου;

Έστω X_t ο αριθμός των λαθών σε διάστημα *t* σελίδων.

Η συχνότητα των λαθών είναι

$$v = 0.5 \text{ λάθη / σελίδα}$$

οπότε για *t=5*, έχουμε $\lambda = vt = 0.5 \cdot 5 = 2.5$. Ζητάμε την πιθανότητα να είναι X_t=0, δηλαδή την

$$p_5(0) = \frac{e^{-2.5} 2.5^0}{0!} = e^{-2.5} = 0.082$$

δηλαδή η πιθανότητα να μην υπάρχουν λάθη σε 5 σελίδες του βιβλίου είναι 8.2%.

4.5 Η ΕΚΘΕΤΙΚΗ ΚΑΤΑΝΟΜΗ

Πρόκειται για συνεχή κατανομή X με pdf

$$f(x) = \begin{cases} ve^{-vx}, & x \geq 0 \\ 0, & \alpha\lambda\lambda o \acute{\upsilon} \end{cases}$$

όπου $v \geq 0$ είναι μια παράμετρος.

Η f είναι πράγματι pdf καθώς

α) $f(x) \geq 0$

β) $\int_{-\infty}^{+\infty} f(x)dx = \int_0^{+\infty} ve^{-vx}dx = \int_0^{+\infty} e^{-vx}dvx = \lim_{t \to +\infty} \left[-e^{-vx}\right]_0^t = \lim_{t \to +\infty} \left[-e^{-vx} + 1\right] = 1$

Δείχνεται ότι

$$E(X) = \frac{1}{v} \qquad V(X) = \frac{1}{v^2}$$

(οι υπολογισμοί αφήνονται για άσκηση).

Το παρακάτω αποτέλεσμα συνδέει την εκθετική κατανομή με την κατανομή Poisson.

ΠΡΟΤΑΣΗ: Έστω X_t ο αριθμός των κλήσεων σε ένα τηλεφωνικό κέντρο στο χρονικό διάστημα *t.* Όπως είναι γνωστό η μεταβλητή X_t ακολουθεί κατανομή Poisson. Ονομάζουμε

T = το χρονικό διάστημα μεταξύ δύο διαδοχικών κλήσεων

Τότε η T είναι συνεχής μεταβλητή και ακολουθεί εκθετική κατανομή.

Απόδειξη: Έστω *ν* η συχνότητα των κλήσεων. Παρατηρούμε ότι
είτε πούμε

«ο αριθμός των κλήσεων σε χρονικό διάστημα t είναι 0»

είτε

«το χρονικό διάστημα μεταξύ 2 διαδοχικών κλήσεων είναι μεγαλύτερο από t»

είναι ακριβώς το ίδιο πράγμα. Βλέποντας τα ως γεγονότα στη γλώσσα των πιθανοτήτων αυτό σημαίνει ότι

$$\{X_t = 0\} \qquad = \qquad \{T > t\}$$

Συνεπώς

$$P(T > t) = P(X_t = 0) = p_t(0) = \frac{e^{-\nu t}(\nu t)^0}{0!} = e^{-\nu t}$$

Αν *F(t)* είναι η cdf της μεταβλητής *T*, τότε από τον ορισμό της cdf

$$F(t) = P(-\infty < T \le t) = 1 - P(T > t) = 1 - e^{-\nu t}$$

οπότε η pdf της *T* είναι

$$f(t) = F'(t) = \nu e^{-\nu t} \qquad\qquad t > 0$$

δηλαδή η *T* ακολουθεί εκθετική κατανομή.

ΠΑΡΑΔΕΙΓΜΑ 9

Η τυχαία μεταβλητή X ακολουθεί εκθετική κατανομή (με παράμετρο ν)
Έστω Α το ενδεχόμενο "X>1" και Β το ενδεχόμενο "X>2".

α) Να υπολογιστούν οι πιθανότητες P(A) και P(B).

β) Να υπολογιστεί η P(B/A) [βρείτε πρώτα το ενδεχόμενο $A \cap B$].

γ) Να βρεθεί η παράμετρος ν όταν $P(A) = P(\overline{A})$.

α) Ας υπολογίσουμε γενικά την $P(X > a)$ για $a > 0$.

$$P(X > a) = 1 - P(X < a) = 1 - \int_{-\infty}^{a} f(x)dx = 1 - \int_{0}^{a} ve^{-vx}dx =$$

$$= 1 - \left[-e^{-vx}\right]_0^a = 1 + \left[e^{-vx}\right]_0^a = 1 + (e^{-av} - 1) = e^{-av}$$

Άρα, P(A)$= P(X > 1) = e^{-v}$ και P(B)$= P(X > 2) = e^{-2v}$.

β) Είναι $A \cap B$ = "X>2" = B. Άρα,

$$\text{P(B/A)} = \frac{P(A \cap B)}{P(A)} = \frac{P(B)}{P(A)} = \frac{e^{-2v}}{e^{-v}} = e^{-2v+v} = e^{-v} = \frac{1}{e^{v}}.$$

γ) $P(A) = P(\overline{A}) \Leftrightarrow P(A) = 1 - P(A) \Leftrightarrow 2P(A) = 1 \Leftrightarrow P(A) = \frac{1}{2}$

$$\Leftrightarrow e^{-v} = \frac{1}{2} \Leftrightarrow \frac{1}{e^{v}} = \frac{1}{2} \Leftrightarrow e^{v} = 2 \Leftrightarrow v = \ln 2$$

ΑΣΚΗΣΕΙΣ

4.1 Έστω ότι μια τυχαία μεταβλητή Χ ακολουθεί κανονική κατανομή με μέση τιμή μ=50 και διασπορά σ^2 =16. Να υπολογιστούν οι πιθανότητες

$$P(X > 50), \quad P(55 < X < 60), \quad P(48 < X < 53), \quad P(X < 48)$$

4.2 Έστω ότι η μεταβλητή Χ ακολουθεί κανονική κατανομή με μ=8 και σ=5. Να υπολογιστούν οι πιθανότητες

$$P(0 < X < 8), \quad P(1.2 < X < 5.3), \quad P(5.3 < X < 10).$$

4.3 Το ποσό των χρημάτων που ξοδεύει για τσιγάρα ένας σπουδαστής ανά μήνα είναι μια τυχαία μεταβλητή Χ με μέση τιμή μ=100 ευρώ και τυπική απόκλιση σ=15 ευρώ. (θεωρούμε ότι η Χ ακολουθεί κανονική κατανομή). Ποια είναι η πιθανότητα να ξοδέψει σε ένα μήνα

α) από 95 μέχρι 120 ευρώ;

β) τουλάχιστον 120 ευρώ;

4.4 Ένα πακέτο καφέ 500gr στην πραγματικότητα μπορεί να ζυγίζει λίγο λιγότερο ή λίγο περισσότερο από 500gr. Ας θεωρήσουμε ότι η τυχαία μεταβλητή Χ που εκφράζει το βάρος ενός πακέτου καφέ ακολουθεί κανονική κατανομή με μέση τιμή 500gr και διασπορά $\sigma^2 = 4$. Σε ένα δείγμα 1000 πακέτων καφέ, πόσα αναμένεται να ζυγίζουν

α) τουλάχιστον 497.5 gr.

β) από 498gr μέχρι 502.5 gr.

4.5 Σε ένα τηλεφωνικό κέντρο γίνονται κατά μέσο όρο 300 κλήσεις την ώρα.

α) ποιος πιστεύετε ότι είναι ο πιθανότερος αριθμός κλήσεων κατά τη διάρκεια 2 λεπτών; Ποια είναι η πιθανότητα να συμβούν τόσες κλήσεις;

β) ποιο είναι πιο πιθανό: να πραγματοποιηθεί μόνο μια κλήση ή να μην πραγματοποιηθούν καθόλου κλήσεις κατά τη διάρκεια 2 λεπτών; Συγκρίνετε τις δύο πιθανότητες.

γ) ποιο είναι πιο πιθανό: να πραγματοποιηθεί 1 κλήση κατά τη διάρκεια ενός λεπτού οι 2 κλήσεις κατά τη διάρκεια δύο λεπτών; Συγκρίνετε τις δύο πιθανότητες.

4.6 Στο τηλεφωνικό κέντρο της εταιρείας PANIC πραγματοποιούνται κατά μέσο όρο 240 κλήσεις την ώρα! Για την επέκταση του τηλεφωνικού δικτύου απαιτούνται κάποιες εκτιμήσεις όσον αφορά την πιθανότητα κλήσεων σε μικρά χρονικά διαστήματα, οπότε καλείστε να απαντήσετε στα παρακάτω ερωτήματα.

α) Έστω Α η πιθανότητα να μην πραγματοποιηθεί καμία κλήση κατά τη διάρκεια 2 λεπτών. Να υπολογιστεί το Α.

β) Δείξτε ότι η πιθανότητα να πραγματοποιηθεί μόνο μια κλήση κατά τη διάρκεια 2 λεπτών είναι οκταπλάσια, δηλαδή 8Α.

γ) Ποια είναι η πιθανότητα να πραγματοποιηθούν τουλάχιστον δύο κλήσεις κατά τη διάρκεια 2 λεπτών; (σε συνάρτηση του Α)

δ) Δείξτε ότι η πιθανότητα να μην πραγματοποιηθεί καμία κλήση κατά τη διάρκεια 4 λεπτών είναι A^2.

4.7 Η διασπορά της τυχαίας μεταβλητής Χ της εκθετικής κατανομής

$$f(x) = ve^{-vx}, \quad x \geq 0 \qquad (\text{με } v > 0)$$

είναι V(X)= $1/v^2$. Να αποδειχθεί για v=2.

4.8 Η τυχαία μεταβλητή Χ παίρνει τιμές στο διάστημα [0,1] με pdf

$$f(x) = 2x^3 + \frac{1}{2}, \quad x \in [0,1]$$

(μπορείτε να επιβεβαιώσετε για εξάσκηση ότι η f είναι pdf και να υπολογίσετε τη μέση τιμή και τη διασπορά της Χ)

Σε ένα πρόγραμμα προσομοίωσης ενός δικτύου, μια δυαδική πηγή εκπέμπει τα σύμβολα

0, αν η Χ πάρει τιμή στο διάστημα $[0, 1/2]$

1, αν η Χ πάρει τιμή στο διάστημα $[1/2, 1]$

α) Ποια είναι η πιθανότητα του κάθε συμβόλου της πηγής;

β) Ποια είναι η πιθανότητα να γίνει εκπομπή της ακολουθίας **11000**;

γ) Ποια είναι η πιθανότητα να γίνει εκπομπή της ακολουθίας **01001**;

δ) Σε μια ακολουθία 5 bits, ποια είναι ή πιθανότητα να υπάρχουν δύο 1;

[Υπόδειξη: στο δ) μπορείτε να χρησιμοποιήσετε τη διωνυμική κατανομή]

ΣΥΝΤΟΜΕΣ ΑΠΑΝΤΗΣΕΙΣ ΤΩΝ ΑΣΚΗΣΕΩΝ

4.1	0.5	0.0994	0.4649	0.3085
4.2	0.4452	0.2077	0.3608	

4.3 α) 0.5375=53.75% β) 0.0918=9.18%

4.4 α) 894 περίπου β) 736 περίπου

4.5 α) 10 κλήσεις με πιθανότητα $\frac{e^{-10}10^{10}}{10!} \cong 0.125$

β) Είναι 10 φορές πιο πιθανό να πραγματοποιηθεί μια κλήση

γ) το πρώτο γεγονός είναι περίπου 15 φορές πιο πιθανό από το δεύτερο

4.6 α) A= e^{-8}

γ) 1-9A

4.7 θεωρητική

4.8 α) P(**0**)=9/32 και P(**1**)=23/32

β) και γ) P(**11000**)=P(**01001**)= $\frac{23^2 9^3}{32^5}$

δ) $10 \cdot \frac{23^2 9^3}{32^5}$