



Περιεχόμενα

1. ΘΕΩΡΙΑ ΚΕΦΑΛΑΙΟΥ 1 - ΣΥΝΑΡΤΗΣΕΙΣ	3
1. Ορισμός συνάρτησης	3
2. Πράξεις με συναρτήσεις.....	3
3. Γραφική παράσταση συνάρτησης	3
4. Γραφικές παραστάσεις βασικών συναρτήσεων	3
5. Μονοτονία – Ακρότατα.....	4
6. Ιδιότητες ορίων	4
7. Συνέχεια συνάρτησης.....	4
8. Παράγωγος της f στο $x = x_0$	5
9. Παράγωγος συνάρτησης	5
10. Παραγωγή βασικών συναρτήσεων (Αποδείξεις)	6
• Η παράγωγος της σταθερής συνάρτησης $f(x) = c$	6
• Η παράγωγος της ταυτοτικής συνάρτησης $f(x) = x$	6
• Η παράγωγος της συνάρτησης $f(x) = x^p$	6
11. Κανόνες Παραγωγίσης (Αποδείξεις).....	7
• Η παράγωγος της συνάρτησης $cf(x)$	7
• Η παράγωγος της συνάρτησης $f(x) + g(x)$	7
12. Βασικοί τύποι και κανόνες παραγωγίσης.....	7
13. Κριτήρια πρώτης παραγώγου	8
Γενικές οδηγίες εν όψει εξέτασης.....	8
2. ΘΕΩΡΙΑ ΚΕΦΑΛΑΙΟΥ 2 - Στατιστική.....	11
Πληθυσμός – Μεταβλητές.....	11
Πίνακες Κατανομής Συχνοτήτων	11
Γραφική Παράσταση Κατανομής Συχνοτήτων	12
α) Ραβδόγραμμα.....	12
β) Διάγραμμα Συχνοτήτων.....	12
γ) Κυκλικό Διάγραμμα.....	13
δ) Σημειόγραμμα.....	13
ε) Χρονόγραμμα.....	13
Ομαδοποίηση των Παρατηρήσεων.....	14
Ιστόγραμμα Συχνοτήτων.....	14
α) Κλάσεις Ίσου Πλάτους.....	14
Καμπύλες Συχνοτήτων.....	15
ΜΕΤΡΑ ΘΕΣΗΣ ΚΑΙ ΔΙΑΣΠΟΡΑΣ.....	16
Μέτρα Θέσης	16
α) Μέση Τιμή.....	16

β) Σταθμικός Μέσος	17
γ) Διάμεσος	17
Διάμεσος σε Ομαδοποιημένα Δεδομένα.....	17
Μέτρα Διασποράς	18
α) Εύρος	18
β) Διακύμανση	18
γ) Τυπική Απόκλιση.....	19
δ) Συντελεστής Μεταβολής	19
.....	20
.....	20
2. Γενικές οδηγίες εν όψει εξέτασης.....	21
Ιδιότητες.....	25
Ιδιότητες μέσης τιμής και διαμέσου	25
Ιδιότητες μέτρων διασποράς και σχετικής διασποράς.....	26
3. Κανονική Κατανομή (Κατανομή Gauss).....	27
1. ΘΕΩΡΙΑ ΚΕΦΑΛΑΙΟΥ 3 - Πιθανότητες.....	28
Πείραμα Τύχης.....	28
Δειγματικός Χώρος.....	28
Ενδεχόμενα	28
Πράξεις με Ενδεχόμενα	28
Ασυμβίβαστα Ενδεχόμενα	29
Έννοια και Ιδιότητες Σχετικής Συχνότητας.....	29
Κλασικός Ορισμός Πιθανότητας	30
Αξιοματικός Ορισμός Πιθανότητας	30
Κανόνες Λογισμού των Πιθανοτήτων	31
2. Γενικές οδηγίες εν όψει εξέτασης.....	32

⁽¹⁾ Επιμέλεια οδηγιών: Γαμβρέλλης Αργύρης – Μέρος των οδηγιών προέρχεται από τις αντίστοιχες οδηγίες του Π.Ι.



Αυτή η εργασία χορηγείται με άδεια [Creative Commons Αναφορά προέλευσης - Μη Εμπορική Χρήση - Παρόμοια Διανομή 3.0 Ελλάδα](https://creativecommons.org/licenses/by-nc-sa/3.0/deed.el).

1. ΘΕΩΡΙΑ ΚΕΦΑΛΑΙΟΥ 1 - ΣΥΝΑΡΤΗΣΕΙΣ

1. Ορισμός συνάρτησης

Συνάρτηση (function) είναι μια διαδικασία με την οποία κάθε στοιχείο ενός συνόλου A αντιστοιχίζεται σε ένα ακριβώς στοιχείο κάποιου άλλου συνόλου B . Το σύνολο A λέγεται **πεδίο ορισμού** της συνάρτησης, ενώ το B λέγεται **πεδίο τιμών**.

Το $f(x)$ λέγεται **τιμή της f στο x** . Το γράμμα x , που συμβολίζει οποιοδήποτε στοιχείο του A , ονομάζεται **ανεξάρτητη μεταβλητή**, ενώ το y , που παριστάνει την τιμή της συνάρτησης στο x και εξαρτάται από την τιμή του x , λέγεται **εξαρτημένη μεταβλητή**.

2. Πράξεις με συναρτήσεις

Αν δύο συναρτήσεις f, g ορίζονται και οι δύο σε ένα σύνολο A , τότε ορίζονται και οι συναρτήσεις:

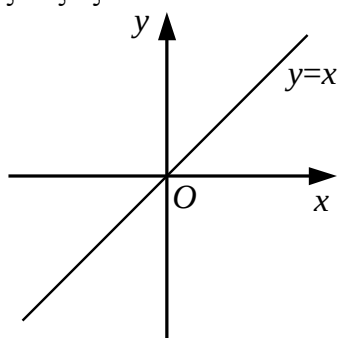
- Το άθροισμα $S = f + g$, με $S(x) = f(x) + g(x)$, $x \in A$
- Η διαφορά $D = f - g$, με $D(x) = f(x) - g(x)$, $x \in A$
- Το γινόμενο $P = f \cdot g$, με $P(x) = f(x) \cdot g(x)$, $x \in A$ και
- Το πηλίκο $R = \frac{f}{g}$, με $R(x) = \frac{f(x)}{g(x)}$, όπου $x \in A$ και $g(x) \neq 0$.

3. Γραφική παράσταση συνάρτησης

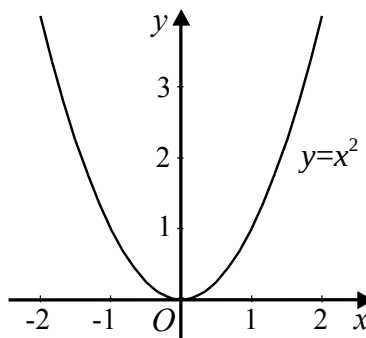
Αν f μια συνάρτηση με πεδίο ορισμού ένα σύνολο A , τότε **γραφική παράσταση** ή **καμπύλη της f** σε ένα καρτεσιανό σύστημα συντεταγμένων Oxy λέγεται το σύνολο των σημείων $M(x, (f(x)))$ για όλα τα $x \in A$.

4. Γραφικές παραστάσεις βασικών συναρτήσεων

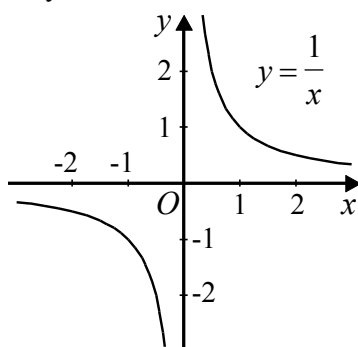
Στα παρακάτω σχήματα φαίνονται οι γραφικές παραστάσεις ορισμένων συναρτήσεων που γνωρίσαμε σε προηγούμενες τάξεις.



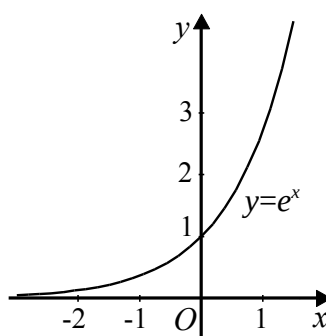
(α) Η καμπύλη της συνάρτησης $f(x) = x$ είναι η διχοτόμος της 1ης και 3ης γωνίας των αξόνων.



(β) Η καμπύλη της συνάρτησης $f(x) = x^2$ είναι μια παραβολή.

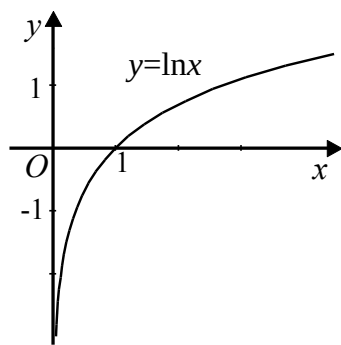


(γ) Η καμπύλη της συνάρτησης



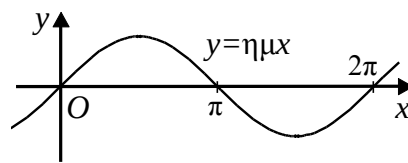
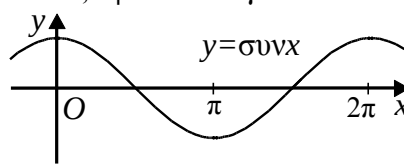
(δ) Η καμπύλη της εκθετικής συνάρτησης

$f(x) = \frac{1}{x}$ είναι μια υπερβολή.



(ε) Η καμπύλη της λογαριθμικής συνάρτησης $f(x) = \ln x$ είναι “δεξιά” του άξονα yy' , αφού ο λογάριθμος ορίζεται μόνο για $x > 0$.

$f(x) = e^x$ είναι “πάνω” από τον άξονα $x'x$, αφού $e^x > 0$ για κάθε $x \in \mathbb{R}$.



(στ) Οι συναρτήσεις $f(x) = \sin x$ και $g(x) = \cos x$ είναι περιοδικές με περίοδο 2π .

5. Μονοτονία – Ακρότατα

Μια συνάρτηση f λέγεται:

- **γνησίως αύξουσα** σε ένα διάστημα Δ του πεδίου ορισμού της, όταν για οποιαδήποτε σημεία $x_1, x_2 \in \Delta$ με $x_1 < x_2$, ισχύει $f(x_1) < f(x_2)$, και
- **γνησίως φθίνουσα** στο Δ , όταν για οποιαδήποτε σημεία $x_1, x_2 \in \Delta$ με $x_1 < x_2$, ισχύει $f(x_1) > f(x_2)$.
- Μια συνάρτηση που είναι γνησίως αύξουσα ή γνησίως φθίνουσα λέγεται **γνησίως μονότονη**.

Μια συνάρτηση f με πεδίο ορισμού το A λέμε ότι παρουσιάζει:

- **τοπικό μέγιστο** στο $x_1 \in A$, όταν $f(x) \leq f(x_1)$, για κάθε x σε μια περιοχή του x_1 , και
- **τοπικό ελάχιστο** στο $x_2 \in A$, όταν $f(x) \geq f(x_2)$, για κάθε x σε μια περιοχή του x_2 .

6. Ιδιότητες ορίων

Αν οι συναρτήσεις f και g έχουν στο x_0 όρια πραγματικών αριθμών, δηλαδή αν $\lim_{x \rightarrow x_0} f(x) = \ell_1$ και

$\lim_{x \rightarrow x_0} g(x) = \ell_2$ όπου ℓ_1 και ℓ_2 πραγματικοί αριθμοί, τότε αποδεικνύεται ότι:

- $\lim_{x \rightarrow x_0} (f(x) + g(x)) = \ell_1 + \ell_2$
- $\lim_{x \rightarrow x_0} (kf(x)) = k\ell_1$
- $\lim_{x \rightarrow x_0} (f(x)g(x)) = \ell_1\ell_2$
- $\lim_{x \rightarrow x_0} \left(\frac{f(x)}{g(x)} \right) = \frac{\ell_1}{\ell_2}$
- $\lim_{x \rightarrow x_0} (f(x))^v = \ell_1^v$
- $\lim_{x \rightarrow x_0} \sqrt[v]{f(x)} = \sqrt[v]{\ell_1}$.

7. Συνέχεια συνάρτησης

Μια συνάρτηση f με πεδίο ορισμού A λέγεται συνεχής, αν για κάθε $x_0 \in A$ ισχύει $\lim_{x \rightarrow x_0} f(x) = f(x_0)$.

- Χαρακτηριστικό γνώρισμα μιας συνεχούς συνάρτησης σε ένα κλειστό διάστημα είναι ότι η γραφική της παράσταση είναι μια συνεχής καμπύλη, δηλαδή για το σχεδιασμό της δε χρειάζεται να σηκώσουμε το μολύβι από το χαρτί.
- Αποδεικνύεται ότι οι γνωστές μας συναρτήσεις, πολυωνυμικές, τριγωνομετρικές, εκθετικές, λογαριθμικές, αλλά και όσες προκύπτουν από πράξεις μεταξύ αυτών είναι συνεχείς συναρτήσεις.

8. Παράγωγος της f στο $x = x_0$

Αν το όριο $\lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h}$ υπάρχει και είναι πραγματικός αριθμός, τότε λέμε ότι η f είναι

παραγωγίσιμη στο σημείο x_0 του πεδίου ορισμού της. Το όριο αυτό ονομάζεται **παράγωγος της f στο x_0** , συμβολίζεται με $f'(x_0)$ και διαβάζεται “ f τονούμενο του x_0 ”. Έχουμε λοιπόν:

$$f'(x_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

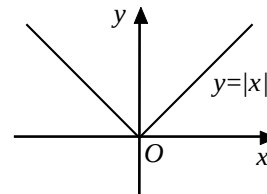
- Η παράγωγος της f στο x_0 εκφράζει:
 - το **ρυθμό μεταβολής** (rate of change) του $y = f(x)$ ως προς το x , όταν $x = x_0$.
 - τον **συντελεστή διεύθυνσης της εφαπτομένης** της καμπύλης που είναι η γραφική παράσταση μιας συνάρτησης f στο σημείο $(x_0, f(x_0))$
- Η ταχύτητα ενός κινητού που κινείται ευθύγραμμα και η θέση του στον άξονα κίνησής του εκφράζεται από τη συνάρτηση $x = f(t)$ θα είναι τη χρονική στιγμή t_0
 $v(t_0) = f'(t_0)$, δηλαδή ο ρυθμός μεταβολής της $f(t)$ ως προς t όταν $t = t_0$.

ΣΧΟΛΙΟ

Υπάρχουν και συναρτήσεις οι οποίες δεν έχουν παράγωγο σε ένα σημείο. Όπως είναι, για παράδειγμα, η συνάρτηση $f(x) = |x|$ στο $x_0 = 0$. Διότι όταν $h < 0$, έχουμε

$$\lim_{h \rightarrow 0} \frac{f(0 + h) - f(0)}{h} = \lim_{h \rightarrow 0} \frac{-h}{h} = -1,$$

ενώ όταν $h > 0$, έχουμε $\lim_{h \rightarrow 0} \frac{f(0 + h) - f(0)}{h} = \lim_{h \rightarrow 0} \frac{h}{h} = 1$, που σημαίνει ότι δεν υπάρχει το $\lim_{h \rightarrow 0} \frac{f(0 + h) - f(0)}{h}$.



9. Παράγωγος συνάρτησης

Έστω μια συνάρτηση f με πεδίο ορισμού το A , και B το σύνολο των $x \in A$ στα οποία η f είναι παραγωγίσιμη. Τότε ορίζεται μια νέα συνάρτηση, με την οποία κάθε $x \in B$ αντιστοιχίζεται στο $f'(x) = \lim_{h \rightarrow 0} \frac{f(x + h) - f(x)}{h}$. Η συνάρτηση αυτή λέγεται (πρώτη) **παράγωγος (derivative) της f και συμβολίζεται με f'** .

- Η παράγωγος της συνάρτησης f' λέγεται **δεύτερη παράγωγος** της f και συμβολίζεται με f'' .
- αν η τετμημένη ενός κινητού που κινείται ευθυγράμμως είναι $x(t)$ τη χρονική στιγμή t , τότε η ταχύτητά του θα είναι $v(t) = x'(t)$. Αν η συνάρτηση v είναι παραγωγίσιμη, τότε η επιτάχυνση του κινητού τη χρονική στιγμή t θα είναι η παράγωγος της ταχύτητας, δηλαδή θα ισχύει $a(t) = v'(t)$ ή ισοδύναμα $a(t) = x''(t)$.

10. Παραγωγή βασικών συνάρτησεων (Αποδείξεις)

- Η παράγωγος της σταθερής συνάρτησης $f(x) = c$

Έχουμε

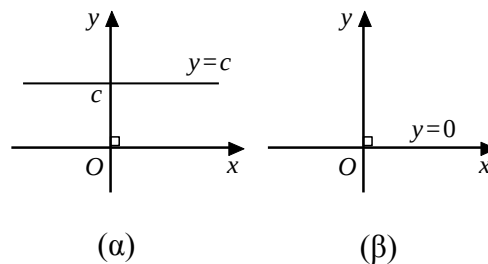
$$f(x+h) - f(x) = c - c = 0$$

και για $h \neq 0$,

$$\frac{f(x+h) - f(x)}{h} = 0,$$

$$\text{οπότε } \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} = 0.$$

Άρα $(c)' = 0$.



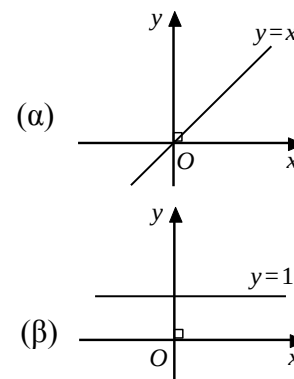
- Η παράγωγος της ταυτοτικής συνάρτησης $f(x) = x$

Έχουμε $f(x+h) - f(x) = (x+h) - x = h$, και για $h \neq 0$,

$$\frac{f(x+h) - f(x)}{h} = \frac{h}{h} = 1.$$

$$\text{Επομένως } \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \rightarrow 0} 1 = 1.$$

Άρα $(x)' = 1$.



- Η παράγωγος της συνάρτησης $f(x) = x^p$

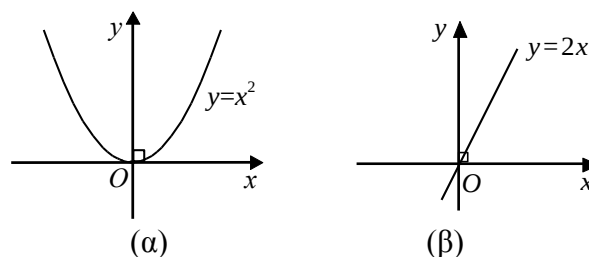
Έστω η συνάρτηση $f(x) = x^2$. Έχουμε

$$f(x+h) - f(x) = (x+h)^2 - x^2 = x^2 + 2xh + h^2 - x^2 = (2x+h)h,$$

$$\text{και για } h \neq 0, \frac{f(x+h) - f(x)}{h} = \frac{(2x+h)h}{h} = 2x+h.$$

$$\text{Επομένως, } \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \rightarrow 0} (2x+h) = 2x.$$

Άρα $(x^2)' = 2x$



Αποδεικνύεται ότι

$$(x^v)' = vx^{v-1}, \text{ όπου } v \text{ φυσικός.}$$

Ο τύπος αυτός ισχύει και στην περίπτωση που ο εκθέτης είναι ρητός αριθμός. Άρα $(x^p)' = px^{p-1}$, όπου p ρητός αριθμός. Για παράδειγμα:

$$\diamond \left(\frac{1}{x}\right)' = (x^{-1})' = -1 \cdot x^{-1-1} = -x^{-2} = \frac{-1}{x^2},$$

$$\diamond (\sqrt{x})' = \left(x^{\frac{1}{2}}\right)' = \frac{1}{2} \cdot x^{\frac{1}{2}-1} = \frac{1}{2} x^{-\frac{1}{2}} = \frac{1}{2\sqrt{x}}.$$

11. Κανόνες Παραγώγισης (Αποδείξεις)

• Η παράγωγος της συνάρτησης $cf(x)$

Έστω η συνάρτηση $F(x) = cf(x)$. Έχουμε

$$F(x+h) - F(x) = cf(x+h) - cf(x) = c(f(x+h) - f(x)), \text{ και για } h \neq 0$$

$$\frac{F(x+h) - F(x)}{h} = \frac{c(f(x+h) - f(x))}{h} = c \frac{f(x+h) - f(x)}{h}.$$

$$\text{Επομένως: } \lim_{h \rightarrow 0} \frac{F(x+h) - F(x)}{h} = \lim_{h \rightarrow 0} \left[c \frac{f(x+h) - f(x)}{h} \right] = cf'(x).$$

$$\text{Άρα } (c \cdot f(x))' = c \cdot f'(x).$$

• Η παράγωγος της συνάρτησης $f(x) + g(x)$

Έστω η συνάρτηση $F(x) = f(x) + g(x)$. Έχουμε

$$\begin{aligned} F(x+h) - F(x) &= (f(x+h) + g(x+h)) - (f(x) + g(x)) \\ &= (f(x+h) - f(x)) + (g(x+h) - g(x)), \end{aligned}$$

$$\text{και για } h \neq 0, \quad \frac{F(x+h) - F(x)}{h} = \frac{f(x+h) - f(x)}{h} + \frac{g(x+h) - g(x)}{h}.$$

Επομένως

$$\lim_{h \rightarrow 0} \frac{F(x+h) - F(x)}{h} = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} + \lim_{h \rightarrow 0} \frac{g(x+h) - g(x)}{h} = f'(x) + g'(x). \text{ Άρα}$$

$$(f(x) + g(x))' = f'(x) + g'(x)$$

12. Βασικοί τύποι και κανόνες παραγώγισης

$(c)' = 0$	$(cf(x))' = cf'(x)$
$(x)' = 1$	$(f(x) + g(x))' = f'(x) + g'(x)$
$(x^\rho)' = \rho x^{\rho-1}$	$(f(x)g(x))' = f'(x)g(x) + f(x)g'(x)$
$(\sqrt{x})' = \frac{1}{2\sqrt{x}}$	$\left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x)g(x) - f(x)g'(x)}{(g(x))^2}$
$(\eta\mu x)' = \sigma\upsilon\nu x$	
$(\sigma\upsilon\nu x)' = -\eta\mu x$	
$(e^x)' = e^x$	

$(\ln x)' = \frac{1}{x}$	$(f(g(x)))' = f'(g(x)) \cdot g'(x)$
--------------------------	-------------------------------------

13. Κριτήρια πρώτης παραγώγου

- Αν μια συνάρτηση f είναι παραγωγίσιμη σε ένα διάστημα Δ και ισχύει $f'(x) > 0$, για κάθε εσωτερικό σημείο του Δ , τότε η f είναι γνησίως αύξουσα στο Δ .
- Αν μια συνάρτηση f είναι παραγωγίσιμη σε ένα διάστημα Δ και ισχύει $f'(x) < 0$, για κάθε εσωτερικό σημείο του Δ , τότε η f είναι γνησίως φθίνουσα στο Δ .
- Αν για μια συνάρτηση f ισχύουν $f'(x_0) = 0$ για $x_0 \in (a, \beta)$, $f'(x) > 0$ στο (a, x_0) και $f'(x) < 0$ στο (x_0, β) , τότε η f παρουσιάζει στο διάστημα (a, β) για $x = x_0$ μέγιστο.
- Αν για μια συνάρτηση f ισχύουν $f'(x_0) = 0$ για $x_0 \in (a, \beta)$, $f'(x) < 0$ στο (a, x_0) και $f'(x) > 0$ στο (x_0, β) , τότε η f παρουσιάζει στο διάστημα (a, β) για $x = x_0$ ελάχιστο.

Γενικές οδηγίες εν όψει εξέτασης

1. Κάθε φορά που εργαζόμαστε με μια συνάρτηση f , η πρώτη μας δουλειά είναι να βρούμε το πεδίο ορισμού της D_f . Θέτουμε, αν χρειάζεται, τους εξής περιορισμούς:
 - Οι **παρονομαστές** πρέπει να είναι **διαφορετικοί από το 0**,
 - Οι **υπόρριζες ποσότητες (οποιασδήποτε τάξης ρίζας)** πρέπει να είναι **μεγαλύτερες ή ίσες του 0**,
 - Αν ο τύπος της συνάρτησης περιέχει παραστάσεις της μορφής **$\ln f(x)$ ή $\log f(x)$** , απαιτούμε: **$f(x) > 0$** ,
 - Αν ο τύπος της συνάρτησης περιέχει παραστάσεις της μορφής **$\exp f(x)$ ή $\sin f(x)$** , απαιτούμε: **$f(x) \neq k\pi + \pi/2$ ή $f(x) \neq k\pi$, $k \in \mathbb{Z}$** .

Ενδεχομένως στις ασκήσεις να χρειαστεί να εφαρμόσουμε συνδυασμό των παραπάνω περιορισμών.

2. Αν f, g συναρτήσεις τότε θυμόμαστε ότι:
 - $D_{f+g} = D_f \cap D_g$, $D_{f-g} = D_f \cap D_g$
 - $D_{f \cdot g} = D_f \cap D_g$
 - $D_{f/g} = D_f \cap D_g - \{x \in \mathbb{R} \mid g(x) = 0\}$
3. Αν C_f, C_g οι γραφικές παραστάσεις δυο συναρτήσεων f, g , πρέπει να θυμόμαστε ότι:
 - η φράση «...το σημείο $A(x_0, y_0)$ ανήκει στη C_f ...» σημαίνει: **$f(x_0) = y_0$**
 - Τα σημεία στα οποία η C_f τέμνει τον $x'x$ έχουν τετμημένες τις **ρίζες της εξίσωσης $f(x) = 0$** .
 - Το σημείο τομής (αν υπάρχει) της C_f με τον $y'y$ είναι το **$O(0, f(0))$** (αν βέβαια $0 \in D_f$).
 - Τα κοινά σημεία των C_f, C_g έχουν τετμημένες τις **ρίζες της εξίσωσης $f(x) = g(x)$** .
 - Τα διαστήματα στα οποία η C_f βρίσκεται **πάνω (αντ. κάτω) από την C_g** , βρίσκονται από τη λύση της ανίσωσης **$f(x) > g(x)$ (αντ. $f(x) < g(x)$)**.
 - Γενικά για να μελετήσουμε τη σχετική θέση των C_f, C_g , χρειάζεται να μελετήσουμε το πρόσημο της διαφοράς **$\delta(x) = f(x) - g(x)$** . Ισχύει:
 - ✓ **$f(x) - g(x) > 0 \Leftrightarrow$ Η C_f βρίσκεται πάνω από την C_g**
 - ✓ **$f(x) - g(x) < 0 \Leftrightarrow$ Η C_f βρίσκεται κάτω από την C_g**
 - ✓ **$f(x) - g(x) = 0 \Leftrightarrow$ Οι C_f, C_g τέμνονται.**
4. **Υπολογισμός ορίου:** Για να υπολογίσουμε το όριο $A = \lim_{x \rightarrow x_0} f(x)$ ακολουθούμε πάντοτε τα εξής βήματα:

- Αντικαθιστούμε στον τύπο της f όπου x το x_0 και αν δεν δημιουργείται απροσδιοριστία, τότε $A = \lim_{x \rightarrow x_0} f(x) = f(x_0)$. (αυτή η ενέργεια προϋποθέτει βέβαια τη συνέχεια της f στο x_0)
- Αν η αντικατάσταση μας οδηγεί σε κλάσμα της μορφής $0/0$ και ο τύπος της f είναι πηλίκο πολυωνύμων (ρητή συνάρτηση), τότε παραγοντοποιούμε αριθμητή και παρονομαστή με σκοπό να απλοποιήσουμε τον παράγοντα $(x - x_0)$. Αφού απλοποιήσουμε, προχωράμε πάλι σε αντικατάσταση του x με το x_0 .
- Αν η αντικατάσταση μας οδηγεί σε κλάσμα της μορφής $0/0$ και ο τύπος της f έχει στον αριθμητή ή/και στον παρονομαστή ριζικό (π.χ. $\sqrt{P(x)}$), τότε πολλαπλασιάζουμε αριθμητή και παρονομαστή με κατάλληλη συζυγή παράσταση. Αφού εκτελέσουμε τη διαφορά τετραγώνων, προχωράμε πάλι σε παραγοντοποίηση αριθμητή και παρονομαστή, απλοποίηση του παράγοντα $(x - x_0)$ και τελικά σε αντικατάσταση του x με το x_0 .
- Αν δεν μπορούμε να κάνουμε κάτι από τα προηγούμενα, εξετάζουμε αν το όριο που αντιμετωπίζουμε είναι όριο κάποιας γνωστής παραγώγου:

Παράδειγμα : Να βρεθεί το $A = \lim_{x \rightarrow 0} \frac{e^x - 1}{x}$.

Λύση: Φανερά δεν μπορούμε να εφαρμόσουμε καμία από τις παραπάνω οδηγίες. Αν $f(x) = e^x$, $x \in \mathbb{R}$,

τότε $f(0) = 1$ και $f'(0) = \lim_{h \rightarrow 0} \frac{f(0+h) - f(0)}{h} = \lim_{h \rightarrow 0} \frac{e^h - 1}{h}$ άρα $A = f'(0) = 1$.

5. Να θυμόμαστε πάντοτε τις δύο ερμηνείες της παραγώγου $f(x_0)$:

- Είναι ο συντελεστής διεύθυνσης της εφαπτομένης της γραφικής παράστασης της f στο σημείο $A(x_0, f(x_0))$
- Είναι ο ρυθμός μεταβολής της f , ως προς x , όταν $x = x_0$.

6. Προσοχή στο πεδίο ορισμού της f' : Δεν προκύπτει από περιορισμούς στον τύπο της f' αλλά καθορίζεται από τα $x \in D_f$ στα οποία η f είναι παραγωγίσιμη.

Παράδειγμα: Αν $f(x) = \ln x$, $x > 0$, τότε $f'(x) = \frac{1}{x}$ με $x > 0$ (κι όχι $x \neq 0$).

7. Οι παρακάτω πίνακες είναι « εκ των ων ουκ άνευ... »

✓ **Παράγωγοι βασικών συναρτήσεων**

f	c	x	x^p	$\sqrt{x}$	$\eta \mu x$	$\sigma \upsilon \nu x$	e^x	$\ln x$
f'	0	1	$\rho x^{\rho-1}$	$\frac{1}{2\sqrt{x}}$	$\sigma \upsilon \nu x$	$-\eta \mu x$	e^x	$\frac{1}{x}$

✓ **Κανόνες παραγώγισης:**

- $(c \cdot f(x))' = c \cdot f'(x)$
- $(f(x) + g(x))' = f'(x) + g'(x)$
- $(f(x) \cdot g(x))' = f'(x) \cdot g(x) + f(x) \cdot g'(x)$

$$\text{iv. } \left(\frac{f(x)}{g(x)} \right)' = \frac{f'(x) \cdot g(x) - f(x) \cdot g'(x)}{g^2(x)}$$

- $(f(g(x)))' = f'(g(x)) \cdot g'(x)$ (Κανόνας της αλυσίδας)

Ειδικά για τον κανόνα της αλυσίδας καλό είναι να θυμόμαστε τα παρακάτω παραδείγματα εφαρμογής του:

f	$f^v(x)$	$\sqrt{f(x)}$	$\eta \mu f(x)$	$\sigma \upsilon \nu f(x)$	$e^{f(x)}$	$\ln f(x)$
f'	$v \cdot f^{v-1}(x) \cdot f'(x)$	$\frac{f'(x)}{2\sqrt{f(x)}}$	$\sigma \upsilon \nu f(x) \cdot f'(x)$	$-\eta \mu f(x) \cdot f'(x)$	$e^{f(x)} \cdot f'(x)$	$\frac{f'(x)}{f(x)}$

Αλλά και αυτά : $(a^x)' = (e^{x \cdot \ln a})' = e^{x \cdot \ln a} \cdot (x \cdot \ln a)' = a^x \cdot \ln a$

$$(\log x)' = \left(\frac{\ln x}{\ln 10} \right)' = \frac{1}{x \cdot \ln 10}$$

8. Για τις ασκήσεις που αναφέρονται στην εφαπτομένη (ε) της γραφικής παράστασης συνάρτησης f, πρέπει να θυμόμαστε τα εξής:

- Πρώτη μας δουλειά είναι να βρούμε το σημείο επαφής. Αν δεν μας δίνεται, τότε θεωρούμε ότι είναι το $A(x_0, f(x_0))$ με $x_0 \in D_f$.
- Οποιαδήποτε «γεωμετρική» πληροφορία σχετικά με τη διεύθυνση της εφαπτομένης (ε) (π.χ. η εφαπτομένη (ε) είναι παράλληλη/κάθετη σε γνωστή ευθεία, η εφαπτομένη (ε) σχηματίζει με τον x'x γωνία με γνωστή εφαπτομένη κ.α.) πρέπει να την αξιοποιούμε βάσει της θεμελιώδους ισότητας: $\lambda_\varepsilon = f'(x_0) = \varepsilon\varphi\omega$, όπου :
 - λ_ε ο συντελεστής διεύθυνσης της εφαπτομένης (ε) ,
 - $f'(x_0)$ η παράγωγος της f στο σημείο επαφής και
 - $\varepsilon\varphi\omega$ η εφαπτομένη της γωνίας που σχηματίζει η εφαπτομένη (ε) με τον x'x.

Θυμόμαστε πάντα και τα κριτήρια: $\varepsilon_1 \parallel \varepsilon_2 \Leftrightarrow \lambda_1 = \lambda_2$ και $\varepsilon_1 \perp \varepsilon_2 \Leftrightarrow \lambda_1 \cdot \lambda_2 = -1$.

(Οι ευθείες $\varepsilon_1, \varepsilon_2$ θεωρούνται μη κατακόρυφες).

- Η εξίσωση της εφαπτομένης (ε) της γρ. παράστασης της f στο $(x_0, f(x_0))$ έχει τη μορφή: $y = f'(x_0) \cdot x + \beta$ (1) ή $y - f(x_0) = f'(x_0)(x - x_0)$ (2).
- Αν μας δίνεται ότι η εφαπτομένη (ε) διέρχεται από κάποιο σημείο B(κ, λ), τότε οι συντεταγμένες του B ικανοποιούν την εξίσωση της (ε).

9. **Μονοτονία – Ακρότατα Συναρτήσεων:** Για να μελετήσουμε μια συνάρτηση f ως προς τη μονοτονία και τα ακρότατα, πρέπει:

- να βρούμε το πεδίο ορισμού A της f και να εξετάσουμε αν είναι παραγωγίσιμη,
- να βρούμε την παράγωγο $f'(x)$
- να βρούμε (αν υπάρχουν) τις ρίζες της $f'(x) = 0$ και να λύσουμε μια από τις ανισώσεις $f'(x) > 0$ ή $f'(x) < 0$,
- να κατασκευάσουμε τον πίνακα προσήμου της f' , στον οποίο πάντοτε βάζουμε τις ρίζες της $f'(x)$, τις τιμές για τις οποίες η f δεν ορίζεται και τα άκρα του πεδίου ορισμού της f.
- τέλος, με βάση το πρόσημο της f' , να βρούμε τη μονοτονία της f στο αντίστοιχο διάστημα.

Προσοχή: Μας αρκεί η συνέχεια στα άκρα των διαστημάτων για να χαρακτηρίσουμε μονότονη την f σε κλειστό διάστημα. Επίσης, ένα άκρο διαστήματος μπορεί να συμπεριλαμβάνεται σε δυο διαστήματα μονοτονίας (π.χ. η $f(x) = x^2$ είναι γν. φθίνουσα στο $(-\infty, 0]$ και γν. αύξουσα στο $[0, +\infty)$).

10. **Προβλήματα:** Μερικές γενικές οδηγίες για την επίλυση προβλημάτων είναι:

- Διαβάζουμε προσεκτικά το πρόβλημα και σίγουρα όχι μόνο μια φορά... Διαβάζουμε την εκφώνηση μέχρι να κατανοήσουμε πλήρως τα δεδομένα και τα ζητούμενα. Αν χρειαστεί, κατασκευάζουμε γεωμετρικό σχήμα με προσοχή και ακρίβεια.
- Βρίσκουμε τύπο ο οποίος δίνει τις τιμές του μεγέθους που πρέπει να μεγιστοποιηθεί/ελαχιστοποιηθεί. Συνήθως αυτός ο τύπος περιέχει περισσότερες από μια μεταβλητές. Από τα δεδομένα του προβλήματος ή/και γενικές γνώσεις, επιλέγουμε κατάλληλη μεταβλητή x ώστε ο τύπος που βρήκαμε να εκφράζεται ως συνάρτηση f(x). (**Προσοχή:** το υπό μεγιστοποίηση/ελαχιστοποίηση μέγεθος πρέπει να είναι το f(x) κι όχι το x.)
- Μελετάμε τη συνάρτηση f(x) ως προς τα ακρότατα (**προσοχή στο πεδίο ορισμού!**) και προσδιορίζουμε την ακρότατη τιμή.
- **Προβλήματα κίνησης:** Αν ένα σώμα κινείται πάνω σε άξονα x'x και η θέση του δίνεται από μια συνάρτηση x(t), όπου t ο χρόνος, τότε πρέπει να θυμόμαστε ότι:
 - ✓ Η ταχύτητα v(t) είναι ο ρυθμός μεταβολής της συνάρτησης θέσης, δηλαδή ισχύει: $v(t) = x'(t)$,
 - ✓ Η επιτάχυνση a(t) είναι ο ρυθμός μεταβολής της ταχύτητας, δηλαδή ισχύει: $a(t) = v'(t)$,
 - ✓ Το σώμα είναι ακίνητο αν και μόνο αν $v(t) = 0$,

- ✓ Το σώμα κινείται προς τη θετική κατεύθυνση αν και μόνο αν $v(t) > 0$,
- ✓ Το σώμα κινείται προς την αρνητική κατεύθυνση αν και μόνο αν $v(t) < 0$,

✂ Προβλήματα ρυθμού μεταβολής: (αχρείαστες να'ναι οι οδηγίες . . .)

Εφαρμόζουμε τις οδηγίες αντιμετώπισης των προβλημάτων, μόνο που εδώ δεν ενδιαφερόμαστε για μέγιστη/ελάχιστη τιμή συνάρτησης, αλλά για τιμή της παραγώγου συνάρτησης $f'(x_0)$.

2. ΘΕΩΡΙΑ ΚΕΦΑΛΑΙΟΥ 2 - Στατιστική

Στατιστική είναι ένα σύνολο αρχών και μεθοδολογιών για:

- το σχεδιασμό της διαδικασίας συλλογής δεδομένων
- τη συνοπτική και αποτελεσματική παρουσίασή τους
- την ανάλυση και εξαγωγή αντίστοιχων συμπερασμάτων.

Ο κλάδος της Στατιστικής που ασχολείται με τον πρώτο στόχο λέγεται **σχεδιασμός πειραμάτων** (experimental design) ενώ, με τον δεύτερο ασχολείται η **περιγραφική στατιστική** (descriptive statistics), που αποτελεί και το αντικείμενο μελέτης μας στη συνέχεια. Τέλος, η **επαγωγική στατιστική** ή **στατιστική συμπερασματολογία** (inferential statistics) περιλαμβάνει τις μεθόδους με τις οποίες γίνεται η προσέγγιση των χαρακτηριστικών ενός μεγάλου συνόλου δεδομένων, με τη μελέτη των χαρακτηριστικών ενός μικρού υποσυνόλου των δεδομένων.

Πληθυσμός – Μεταβλητές

Σε ένα σύνολο που θέλουμε να εξετάσουμε τα στοιχεία του ως προς ένα ή περισσότερα χαρακτηριστικά του λέγεται **πληθυσμός** (population). Τα στοιχεία του πληθυσμού συχνά αναφέρονται και ως μονάδες ή άτομα του πληθυσμού. Τα χαρακτηριστικά ως προς τα οποία εξετάζουμε έναν πληθυσμό λέγονται **μεταβλητές** (variables) και τις συμβολίζουμε συνήθως με τα κεφαλαία γράμματα $X, Y, Z, B, \dots$. Οι δυνατές τιμές που μπορεί να πάρει μια μεταβλητή λέγονται **τιμές της μεταβλητής**.

Τις μεταβλητές τις διακρίνουμε:

1. Σε **ποιοτικές** ή **κατηγορικές** μεταβλητές, των οποίων οι τιμές τους δεν είναι αριθμοί. Τέτοιες είναι, για παράδειγμα, η ομάδα αίματος (με τιμές A, B, AB, O), το φύλο (με τιμές αγόρι, κορίτσι), οι συνέπειες του καπνίσματος (με τιμές καρδιακά νοσήματα, καρκίνος κτλ), όπως επίσης και η οικονομική κατάσταση και η υγεία των ανθρώπων (που μπορεί να χαρακτηριστεί ως κακή, μέτρια, καλή ή πολύ καλή), καθώς και το ενδιαφέρον των μαθητών για τη Στατιστική, που μπορεί να χαρακτηριστεί ως υψηλό, μέτριο, χαμηλό ή μηδαμινό.
2. Σε **ποσοτικές** μεταβλητές, των οποίων οι τιμές είναι αριθμοί και διακρίνονται:
 - i) Σε **διακριτές** μεταβλητές, που παίρνουν μόνο “μεμονωμένες” τιμές. Τέτοιες μεταβλητές είναι, για παράδειγμα, ο αριθμός των υπαλλήλων μιας επιχείρησης (με τιμές $1, 2, \dots$), το αποτέλεσμα της ρίψης ενός ζαριού (με τιμές $1, 2, \dots, 6$) κτλ.
 - ii) Σε **συνεχείς** μεταβλητές, που μπορούν να πάρουν οποιαδήποτε τιμή ενός διαστήματος πραγματικών αριθμών (α, β) . Τέτοιες μεταβλητές είναι το ύψος και το βάρος των μαθητών της Γ' Λυκείου, ο χρόνος που χρειάζονται οι μαθητές να απαντήσουν στα θέματα μιας εξέτασης, η διάρκεια μιας τηλεφωνικής συνδιάλεξης κτλ.

Πίνακες Κατανομής Συχνοτήτων

- Ας υποθέσουμε ότι $x_1, x_2, \dots, x_k$ είναι οι τιμές μιας μεταβλητής X , που αφορά τα άτομα ενός δείγματος μεγέθους n , $k \leq n$. Στην τιμή x_i αντιστοιχίζεται η (απόλυτη) **συχνότητα** (frequency) v_i , δηλαδή ο φυσικός αριθμός που δείχνει πόσες φορές εμφανίζεται η τιμή x_i της εξεταζόμενης μεταβλητής X στο σύνολο των παρατηρήσεων. Ισχύει: $v_1 + v_2 + \dots + v_k = n$
- Αν διαιρέσουμε τη συχνότητα v_i με το μέγεθος n του δείγματος, προκύπτει η **σχετική συχνότητα** (relative frequency) f_i της τιμής x_i , δηλαδή $f_i = \frac{v_i}{n}$, $i = 1, 2, \dots, k$.

Για τη σχετική συχνότητα ισχύουν οι ιδιότητες:

(i) $0 \leq f_i \leq 1$ για $i=1,2,...,κ$ αφού $0 \leq v_i \leq v$.

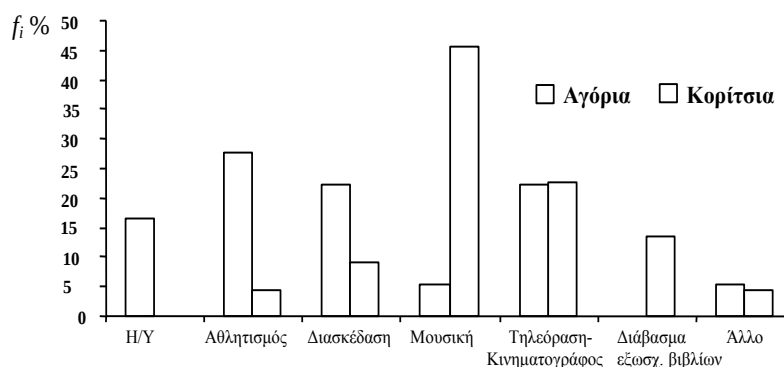
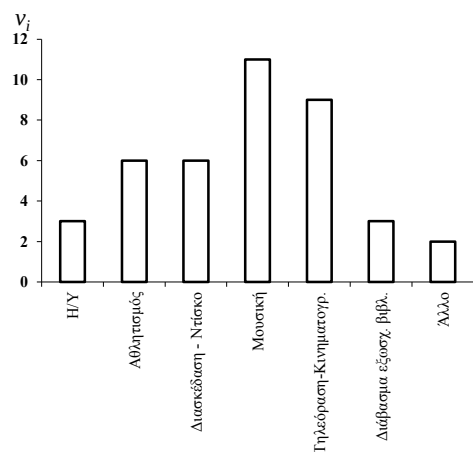
(ii) $f_1 + f_2 + ... + f_κ = 1$, αφού $f_1 + f_2 + ... + f_κ = \frac{v_1}{v} + \frac{v_2}{v} + ... + \frac{v_κ}{v} = \frac{v_1 + v_2 + ... + v_κ}{v} = \frac{v}{v} = 1$.

- **Σχετικές συχνότητες** f_i **επί τοις εκατό** $f_i \% = 100f_i$. Ισχύει $f_1 \% + f_2 \% + f_3 \% + f_4 \% = 100\%$.
- Οι ποσότητες x_i, v_i, f_i για ένα δείγμα συγκεντρώνονται σε ένα συνοπτικό πίνακα, που ονομάζεται **πίνακας κατανομής συχνοτήτων** ή απλά **πίνακας συχνοτήτων**.
- Για μια μεταβλητή, το σύνολο των ζευγών (x_i, v_i) λέμε ότι αποτελεί την **κατανομή συχνοτήτων** και το σύνολο των ζευγών (x_i, f_i) , ή των ζευγών $(x_i, f_i \%)$, την **κατανομή των σχετικών συχνοτήτων**.
- Στην περίπτωση των **ποσοτικών μεταβλητών** χρησιμοποιούνται συνήθως και οι λεγόμενες **αθροιστικές συχνότητες** (cumulative frequencies) N_i και οι **αθροιστικές σχετικές συχνότητες** (cumulative relative frequencies) F_i , οι οποίες εκφράζουν το πλήθος και το ποσοστό αντίστοιχα των παρατηρήσεων που είναι μικρότερες ή ίσες της τιμής x_i . Συχνά οι F_i πολλαπλασιάζονται επί 100 εκφραζόμενες έτσι επί τοις εκατό, δηλαδή $F_i \% = 100F_i$. Ισχύει: Αν οι τιμές $x_1, x_2, ..., x_κ$ μιας ποσοτικής μεταβλητής X είναι σε **αύξουσα διάταξη**, τότε η αθροιστική συχνότητα της τιμής x_i είναι $N_i = v_1 + v_2 + ... + v_i$. Όμοια, η αθροιστική σχετική συχνότητα είναι $F_i = f_1 + f_2 + ... + f_i$, για $i=1,2,...,κ$. Είναι φανερό ότι ισχύουν οι σχέσεις: 1) $v_1 = N_1$, $v_2 = N_2 - N_1, ..., v_κ = N_κ - N_{κ-1}$ και 2) $f_1 = F_1$, $f_2 = F_2 - F_1, ..., f_κ = F_κ - F_{κ-1}$.

Γραφική Παράσταση Κατανομής Συχνοτήτων

α) Ραβδόγραμμα

Το **ραβδόγραμμα** (barchart) χρησιμοποιείται για τη γραφική παράσταση των τιμών μιας ποιοτικής μεταβλητής. Το ραβδόγραμμα αποτελείται από ορθογώνιες στήλες που οι βάσεις τους βρίσκονται πάνω στον οριζόντιο ή τον κατακόρυφο άξονα. Σε κάθε τιμή της μεταβλητής X αντιστοιχεί μια ορθογώνια στήλη της οποίας το ύψος είναι ίσο με την αντίστοιχη συχνότητα ή σχετική συχνότητα. Έτσι έχουμε αντίστοιχα το **ραβδόγραμμα συχνοτήτων** και το **ραβδόγραμμα σχετικών συχνοτήτων**. Τόσο η απόσταση μεταξύ των στηλών όσο και το μήκος των βάσεων τους καθορίζονται αυθαίρετα.



(γ)

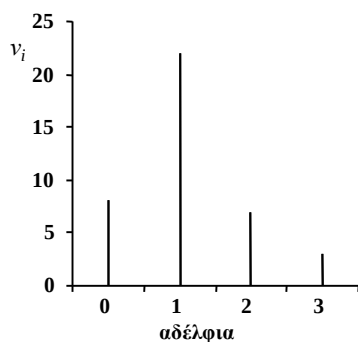
Ραβδόγραμμα σχετικών συχνοτήτων για την απασχόληση των μαθητών του πίνακα 4 ανάλογα με το φύλο.

β) Διάγραμμα Συχνοτήτων

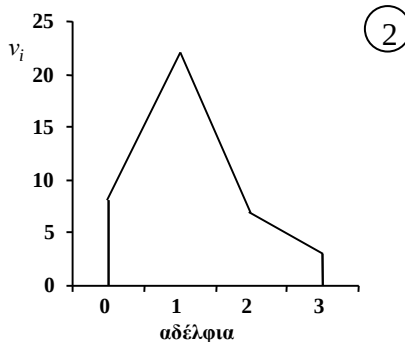
Στην περίπτωση που έχουμε μια ποσοτική μεταβλητή αντί του ραβδογράμματος χρησιμοποιείται το **διάγραμμα συχνοτήτων** (line diagram). Αυτό μοιάζει με το ραβδόγραμμα με μόνη διαφορά ότι αντί να χρησιμοποιούμε συμπαγή ορθογώνια υψώνουμε σε κάθε x_i (υποθέτοντας ότι $x_1 < x_2 < ... < x_κ$) μία κάθετη γραμμή με μήκος ίσο προς την αντίστοιχη συχνότητα, όπως φαίνεται στο σχήμα 2(α). Μπορούμε επίσης αντί των συχνοτήτων v_i στον

κάθετο άξονα να βάλουμε τις σχετικές συχνότητες f_i , οπότε έχουμε το **διάγραμμα σχετικών συχνοτήτων**.

Ενώνοντας τα σημεία (x_i, v_i) ή (x_i, f_i) έχουμε το λεγόμενο **πολύγωνο συχνοτήτων** ή **πολύγωνο σχετικών συχνοτήτων**, αντίστοιχα.



(α)



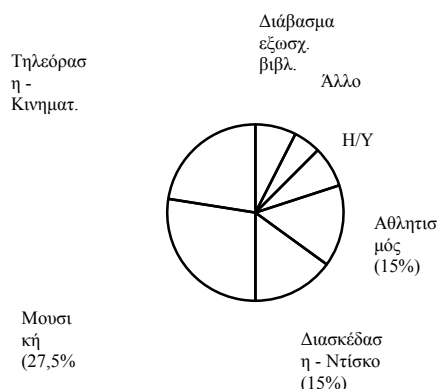
(β)

Διάγραμμα συχνοτήτων (α) και πολύγωνο συχνοτήτων (β) για τη μεταβλητή “αριθμός αδελφών”

γ) Κυκλικό Διάγραμμα

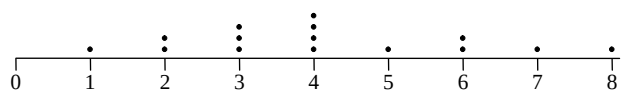
Το **κυκλικό διάγραμμα** (piechart) χρησιμοποιείται για τη γραφική παράσταση **τόσο των ποιοτικών όσο και των ποσοτικών δεδομένων**, όταν οι διαφορετικές **τιμές της μεταβλητής** είναι **σχετικά λίγες**. Το κυκλικό διάγραμμα είναι ένας κυκλικός δίσκος **χωρισμένος σε κυκλικούς τομείς, τα εμβαδά ή τα τόξα** των οποίων **ανάλογα** προς τις **αντίστοιχες συχνότητες v_i ή τις σχετικές συχνότητες f_i** των τιμών x_i της μεταβλητής. συμβολίσουμε με α_i το αντίστοιχο τόξο ενός κυκλικού τμήματος στο κυκλικό διάγραμμα συχνοτήτων, τότε

$$\alpha_i = v_i \frac{360^\circ}{n} = 360^\circ f_i \text{ για } i=1,2,\dots,k.$$



δ) Σημειόγραμμα

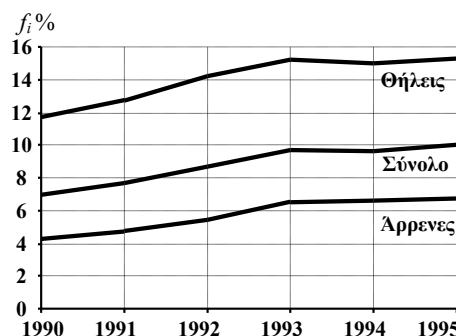
Όταν έχουμε λίγες παρατηρήσεις, η κατανομή τους μπορεί να περιγραφεί με το **σημειόγραμμα** (dot diagram), στο οποίο οι τιμές παριστάνονται γραφικά σαν σημεία υπεράνω ενός οριζώντιου άξονα. Στο σχήμα έχουμε το σημειόγραμμα των χρόνων (σε λεπτά) 4,2,3,1,5,6,4,2,3,4,7,4,8,6,3 που χρειάστηκαν δεκαπέντε μαθητές, για να λύσουν ένα πρόβλημα.



χρόνος (σε λεπτά)

ε) Χρονόγραμμα

Το **χρονόγραμμα** ή **χρονολογικό διάγραμμα** χρησιμοποιείται για τη γραφική απεικόνιση της διαχρονικής εξέλιξης ενός οικονομικού, δημογραφικού ή άλλου μεγέθους. Ο οριζόντιος άξονας χρησιμοποιείται συνήθως ως άξονας μέτρησης του χρόνου και ο κάθετος ως άξονας μέτρησης της εξεταζόμενης μεταβλητής. Στο σχήμα έχουμε το χρονόγραμμα του ποσοστού ανεργίας στη χώρα μας από το 1990 έως το 1995. (Πηγή ΕΣΥΕ).



Ομαδοποίηση των Παρατηρήσεων

Οι πίνακες συχνοτήτων και κατ'αναλογία τα αντίστοιχα διαγράμματα είναι δύσκολο να κατασκευαστούν, όταν το πλήθος των τιμών μιας μεταβλητής είναι αρκετά μεγάλο. Αυτό μπορεί να συμβεί είτε στην περίπτωση μιας διακριτής μεταβλητής είτε, πολύ περισσότερο, στην περίπτωση μιας συνεχούς μεταβλητής, όπου αυτή μπορεί να πάρει οποιαδήποτε τιμή στο διάστημα ορισμού της. Σ'αυτές τις περιπτώσεις είναι απαραίτητο να ταξινομηθούν (ομαδοποιηθούν) τα δεδομένα σε μικρό πλήθος ομάδων, που ονομάζονται και **κλάσεις** (class intervals), έτσι ώστε κάθε τιμή να ανήκει μόνο σε μία κλάση. Τα άκρα των κλάσεων καλούνται **όρια των κλάσεων** (class boundaries). Συνήθως υιοθετούμε την περίπτωση που μια κλάση περιέχει το κάτω άκρο της (κλειστή αριστερά) αλλά όχι το άνω άκρο της (ανοικτή δεξιά), δηλαδή που οι κλάσεις είναι της μορφής $[,)$. Οι παρατηρήσεις κάθε κλάσης θεωρούνται όμοιες, οπότε μπορούν να "αντιπροσωπευθούν" από τις **κεντρικές τιμές**, τα κέντρα δηλαδή κάθε κλάσης.

• Το πρώτο βήμα στην ομαδοποίηση των δεδομένων είναι η εκλογή του αριθμού k των ομάδων ή κλάσεων. Ο αριθμός αυτός συνήθως ορίζεται αυθαίρετα από τον ερευνητή σύμφωνα με την πείρα του. Γενικά όμως μπορεί να χρησιμοποιηθεί ως οδηγός ο παρακάτω πίνακας:

Μέγεθος δείγματος N	Αριθμός κλάσεων k	Μέγεθος δείγματος N	Αριθμός κλάσεων k
< 20	5	200 – 400	9
20 – 50	6	400 – 700	10
50 – 100	7	700 – 1000	11
100 – 200	8	≥ 1000	12

• Το δεύτερο βήμα είναι ο προσδιορισμός του πλάτους των κλάσεων. **Πλάτος μιας κλάσης** ονομάζεται η διαφορά του κατωτέρου από το ανώτερο όριο της κλάσης. Στην πλειονότητα των πρακτικών εφαρμογών οι κλάσεις έχουν το ίδιο πλάτος. Φυσικά υπάρχουν και περιπτώσεις όπου επιβάλλεται οι κλάσεις να έχουν άνισο πλάτος, όπως, για παράδειγμα, στις κατανομές εισοδήματος, ημερών απεργίας κτλ. Για να κατασκευάσουμε ισοπλάτεις κλάσεις, χρησιμοποιούμε το **εύρος** (range) R του δείγματος, δηλαδή τη διαφορά της μικρότερης παρατήρησης από τη μεγαλύτερη παρατήρηση του συνολικού δείγματος. Τότε υπολογίζουμε το πλάτος c των κλάσεων διαιρώντας το εύρος R διά του αριθμού των κλάσεων k , στρογγυλεύοντας, αν χρειαστεί για λόγους διευκόλυνσης, πάντα προς τα πάνω.

• Το επόμενο βήμα είναι η κατασκευή των κλάσεων. Ξεκινώντας από την μικρότερη παρατήρηση, ή για πρακτικούς λόγους λίγο πιο κάτω από την μικρότερη παρατήρηση, και προσθέτοντας κάθε φορά το πλάτος c δημιουργούμε τις k κλάσεις. Αυτονόητο είναι ότι η μεγαλύτερη τιμή του δείγματος θα (πρέπει να) ανήκει οπωσδήποτε στην τελευταία κλάση.

• Το πλήθος των παρατηρήσεων n_i που προκύπτουν από τη διαλογή για την κλάση i καλείται **συχνότητα της κλάσης** αυτής ή **συχνότητα της κεντρικής τιμής** x_i , $i=1,2,...,k$.

Πρέπει να προσεχτεί ότι:

- Καμία παρατήρηση δεν μπορεί να μείνει έξω από κάποια κλάση.
- Οι κεντρικές τιμές διαφέρουν μεταξύ τους όσο και το πλάτος των κλάσεων, που εδώ είναι ίσο με 6.
- Μία παρατήρηση που συμπίπτει με το άνω άκρο μιας κλάσης θα τοποθετηθεί κατά τη διαλογή στην αμέσως επόμενη κλάση. Για παράδειγμα, ο μαθητής με ύψος 180 θα τοποθετηθεί στην πέμπτη κλάση [180, 186).

Ιστογράμματα Συχνοτήτων

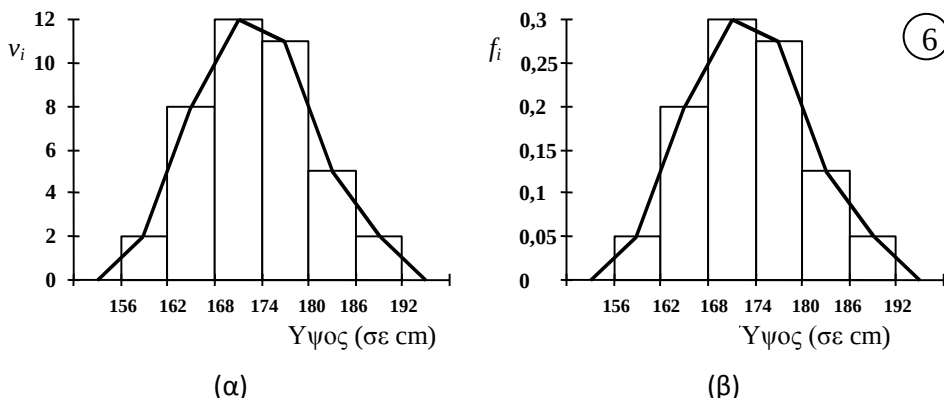
Η αντίστοιχη γραφική παράσταση ενός πίνακα συχνοτήτων με ομαδοποιημένα δεδομένα γίνεται με το λεγόμενο **ιστόγραμμα** (histogram) συχνοτήτων. Στον οριζόντιο άξονα ενός συστήματος ορθογωνίων αξόνων σημειώνουμε, με κατάλληλη κλίμακα, τα όρια των κλάσεων. Στη συνέχεια, κατασκευάζουμε διαδοχικά ορθογώνια (ιστούς), από καθένα από τα οποία έχει βάση ίση με το πλάτος της κλάσης και ύψος τέτοιο, ώστε το **εμβαδόν του ορθογωνίου να ισούται με τη συχνότητα της κλάσης αυτής**.

α) Κλάσεις Ίσου Πλάτους

Θεωρώντας το πλάτος c ως μονάδα μέτρησης του χαρακτηριστικού στον οριζόντιο άξονα, το ύψος κάθε ορθογωνίου είναι ίσο προς τη συχνότητα της αντίστοιχης κλάσης, έτσι ώστε να ισχύει πάλι ότι το εμβαδόν των

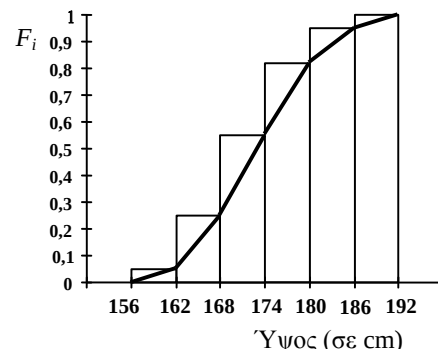
ορθογωνίων είναι ίσο με τις αντίστοιχες συχνότητες. Επομένως, στον κατακόρυφο άξονα σε ένα ιστόγραμμα συχνотήτων βάζουμε τις συχνότητες. Με ανάλογο τρόπο κατασκευάζεται και το **ιστόγραμμα σχετικών συχνотήτων**, οπότε στον κάθετο άξονα βάζουμε τις σχετικές συχνотήτες.

Αν στα ιστογράμματα συχνотήτων θεωρήσουμε δύο ακόμη υποθετικές κλάσεις, στην αρχή και στο τέλος, με συχνотήτα μηδέν και στη συνέχεια ενώσουμε τα μέσα των άνω βάσεων των ορθογωνίων, σχηματίζεται το λεγόμενο **πολύγωνο συχνотήτων** (frequency polygon). Το εμβαδόν του χωρίου που ορίζεται από το πολύγωνο συχνотήτων και τον οριζόντιο άξονα είναι ίσο με το άθροισμα των συχνотήτων, δηλαδή με το μέγεθος του δείγματος n . Όμοια κατασκευάζεται από το ιστόγραμμα σχετικών συχνотήτων και το **πολύγωνο σχετικών συχνотήτων** με εμβαδόν ίσο με 1, (βλέπε σχήμα 6).



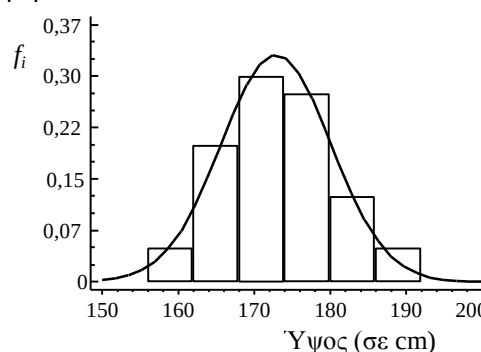
Ιστόγραμμα και πολύγωνο (α) συχνотήτων και (β) σχετικών συχνотήτων για τα δεδομένα του πίνακα 9.

Με τον ίδιο τρόπο κατασκευάζονται και τα **ιστογράμματα αθροιστικών συχνотήτων** και **αθροιστικών σχετικών συχνотήτων**. Αν ενώσουμε σε ένα ιστόγραμμα αθροιστικών συχνотήτων τα **δεξιά άκρα** (όχι μέσα) των άνω βάσεων των ορθογωνίων με ευθύγραμμα τμήματα βρίσκουμε το **πολύγωνο αθροιστικών συχνотήτων** (ogive) της κατανομής. Στο διπλανό σχήμα παριστάνεται το ιστόγραμμα και το πολύγωνο αθροιστικών σχετικών συχνотήτων για το ύψος των μαθητών του πίνακα 9.



Καμπύλες Συχνотήτων

Εάν υποθέσουμε ότι ο αριθμός των κλάσεων για μια συνεχή μεταβλητή είναι αρκετά μεγάλος (τείνει στο άπειρο) και ότι το πλάτος των κλάσεων είναι αρκετά μικρό (τείνει στο μηδέν), τότε η πολυγωνική γραμμή συχνотήτων τείνει να πάρει τη μορφή μιας ομαλής καμπύλης, η οποία ονομάζεται **καμπύλη συχνотήτων** (frequency curve), όπως δείχνει το διπλανό σχήμα. Οι καμπύλες συχνотήτων έχουν μεγάλη εφαρμογή στη Στατιστική, όπου οι ιδιότητες τους μπορούν να χρησιμοποιηθούν για την εξαγωγή χρήσιμων συμπερασμάτων.



Η κατανομή (β), με “κωδωνοειδή” μορφή λέγεται **κανονική κατανομή** (normal distribution) και παίζει σπουδαίο ρόλο στη Στατιστική. Όταν οι παρατηρήσεις “κατανέμονται” ομοιόμορφα σε ένα διάστημα $[a, b]$, όπως στην κατανομή (α), η κατανομή λέγεται **ομοιόμορφη**. Όταν οι παρατηρήσεις δεν είναι συμμετρικά κατανεμημένες, η κατανομή λέγεται ασύμμετρη με θετική ασυμμετρία όπως στην κατανομή (γ) ή αρνητική ασυμμετρία όπως στην κατανομή (δ).



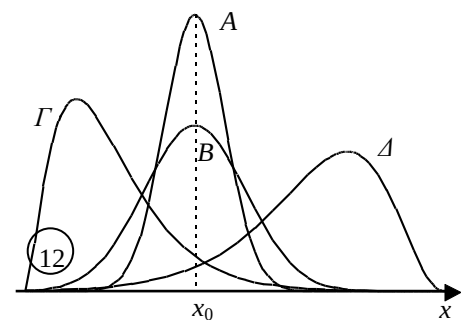
Μερικές χαρακτηριστικές κατανομές συχνοτήτων

ΜΕΤΡΑ ΘΕΣΗΣ ΚΑΙ ΔΙΑΣΠΟΡΑΣ

Ορίζουμε κάποια **μέτρα** (αριθμητικά μεγέθη), που να μας δίνουν α) τη θέση του “κέντρου” των παρατηρήσεων στον οριζόντιο άξονα και β) τη διασπορά των παρατηρήσεων, δηλαδή πόσο αυτές εκτείνονται γύρω από το “κέντρο” τους. Τα πρώτα τα καλούμε **μέτρα θέσης** της κατανομής (location measures), ενώ τα δεύτερα **μέτρα διασποράς** ή **μέτρα μεταβλητότητας** (measures of variability).

Εκτός από τα μέτρα θέσης και διασποράς μιας κατανομής πολλές φορές είναι απαραίτητος και ο προσδιορισμός κάποιων άλλων μέτρων, που καθορίζουν τη **μορφή** της κατανομής. Κατά πόσο δηλαδή η αντίστοιχη καμπύλη συχνοτήτων είναι συμμετρική ή όχι ως προς την ευθεία $x = x_0$, για δεδομένο σημείο x_0 του άξονα $0x$. Τα μέτρα αυτά, που συνήθως εκφράζονται σε συνάρτηση με τα μέτρα θέσης και διασποράς, καλούνται **μέτρα ασυμμετρίας** (measures of skewness).

Υπολογίζοντας από ένα σύνολο δεδομένων κάποια από τα ανωτέρω μέτρα, μπορούμε να έχουμε μια σύντομη περιγραφή της μορφής της καμπύλης συχνοτήτων. Στο σχήμα 12 οι καμπύλες συχνοτήτων A και B είναι συμμετρικές με το ίδιο “κέντρο” x_0 , αλλά η B έχει μεγαλύτερη μεταβλητότητα από την A . Οι καμπύλες Γ και Δ είναι ασύμμετρες, με τη Γ όπως λέμε να παρουσιάζει θετική ασυμμετρία και τη Δ αρνητική ασυμμετρία. Το “κέντρο” της Γ είναι αριστερότερα του x_0 , ενώ της Δ είναι δεξιότερα του x_0 . Η Δ παρουσιάζει μεγαλύτερη μεταβλητότητα από τη Γ .



Μέτρα Θέσης

Τα πιο συνηθισμένα μέτρα που χρησιμοποιούνται για την περιγραφή της θέσης ενός συνόλου δεδομένων πάνω στον οριζόντιο άξονα $0x$, εκφράζοντας την “κατά μέσο όρο” απόστασή τους από την αρχή των αξόνων, είναι ο αριθμητικός μέσος ή μέση τιμή (arithmetic mean or average), η διάμεσος (median)

α) Μέση Τιμή

Η μέση τιμή ενός συνόλου n παρατηρήσεων αποτελεί το σπουδαιότερο και χρησιμότερο μέτρο της Στατιστικής και ορίζεται ως το άθροισμα των παρατηρήσεων διά του πλήθους των παρατηρήσεων.

Όταν σε ένα δείγμα μεγέθους n οι παρατηρήσεις μιας μεταβλητής X είναι $t_1, t_2, \dots, t_n$, τότε η μέση τιμή συμβολίζεται

με $\bar{x}$ και δίνεται από τη σχέση:
$$\bar{x} = \frac{t_1 + t_2 + \dots + t_n}{n} = \frac{\sum_{i=1}^n t_i}{n} = \frac{1}{n} \sum_{i=1}^n t_i$$
 όπου το σύμβολο $\sum_{i=1}^n t_i$ παριστάνει μια συντομογραφία του αθροίσματος $t_1 + t_2 + \dots + t_n$ και διαβάζεται “άθροισμα των t_i από $i=1$ έως n ”. Συχνά, όταν δεν υπάρχει πρόβλημα σύγχυσης, συμβολίζεται και ως $\sum t_i$ ή ακόμα πιο απλά με $\sum t$.

Σε μια κατανομή συχνοτήτων, αν $x_1, x_2, \dots, x_k$ είναι οι τιμές της μεταβλητής X ή τα κέντρα των κλάσεων, εφ' όσον έχουμε κλάσεις, με συχνότητες $v_1, v_2, \dots, v_k$ αντίστοιχα, η μέση τιμή ορίζεται ισοδύναμα από τη σχέση:

$$\bar{x} = \frac{x_1 v_1 + x_2 v_2 + \dots + x_k v_k}{v_1 + v_2 + \dots + v_k} = \frac{\sum_{i=1}^k x_i v_i}{\sum_{i=1}^k v_i} = \frac{1}{v} \sum_{i=1}^k x_i v_i$$

Η παραπάνω σχέση ισοδύναμα γράφεται: $\bar{x} = \sum_{i=1}^k x_i \frac{v_i}{v} = \sum_{i=1}^k x_i f_i$ όπου f_i οι σχετικές συχνότητες.

β) Σταθμικός Μέσος

Στις περιπτώσεις που δίνεται διαφορετική βαρύτητα (έμφαση) στις τιμές $x_1, x_2, \dots, x_v$ ενός συνόλου δεδομένων, τότε αντί του αριθμητικού μέσου χρησιμοποιούμε τον **σταθμισμένο αριθμητικό μέσο** ή **σταθμικό μέσο** (weighted mean). Εάν σε κάθε τιμή $x_1, x_2, \dots, x_v$ δώσουμε διαφορετική βαρύτητα, που εκφράζεται με τους λεγόμενους συντελεστές στάθμισης (βαρύτητας) $w_1, w_2, \dots, w_v$, τότε ο σταθμικός μέσος βρίσκεται από τον τύπο:

$$\bar{x} = \frac{x_1 w_1 + x_2 w_2 + \dots + x_v w_v}{w_1 + w_2 + \dots + w_v} = \frac{\sum_{i=1}^v x_i w_i}{\sum_{i=1}^v w_i}$$

γ) Διάμεσος

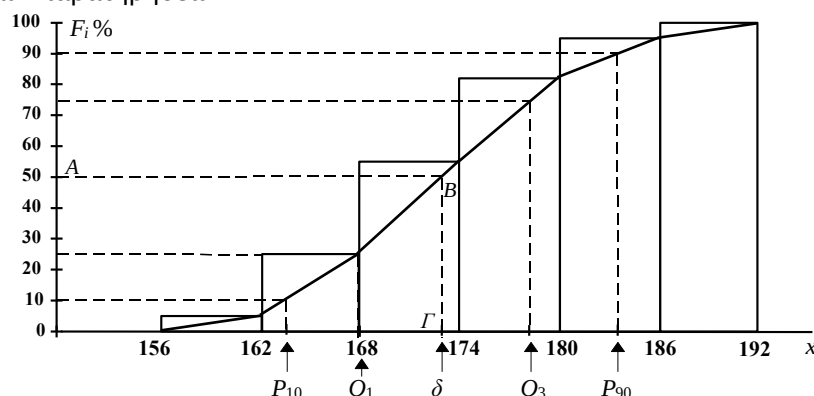
Η μέση τιμή δεν ενδείκνυται ως μέτρο θέσης ("κέντρο") των παρατηρήσεων που έχουν ακραίες τιμές. Αντίθετα, ένα άλλο μέτρο θέσης που δεν επηρεάζεται από ακραίες παρατηρήσεις είναι η **διάμεσος** (median), η οποία ορίζεται ως εξής:

Διάμεσος (δ) ενός δείγματος n παρατηρήσεων οι οποίες έχουν διαταχθεί σε αύξουσα σειρά ορίζεται ως η μεσαία παρατήρηση, όταν το n είναι περιττός αριθμός, ή ο μέσος όρος (ημιάθροισμα) των δύο μεσαίων παρατηρήσεων όταν το n είναι άρτιος αριθμός.

Ή
η διάμεσος είναι η τιμή για την οποία το πολύ 50% των παρατηρήσεων είναι μικρότερες από αυτήν και το πολύ 50% των παρατηρήσεων είναι μεγαλύτερες από την τιμή αυτήν

Διάμεσος σε Ομαδοποιημένα Δεδομένα

Θεωρούμε ιστόγραμμα αθροιστικών σχετικών συχνοτήτων με την πολυγωνική γραμμή, σχήμα 13. Η διάμεσος, όπως ορίστηκε, αντιστοιχεί στην τιμή $x = \delta$ της μεταβλητής X (στον οριζόντιο άξονα), έτσι ώστε το 50% των παρατηρήσεων να είναι μικρότερες ή ίσες του δ . Δηλαδή, η διάμεσος θα έχει αθροιστική σχετική συχνότητα $F_i = 50\%$. Εφόσον στον κάθετο άξονα έχουμε τις αθροιστικές σχετικές συχνότητες, από το σημείο Α (50% των παρατηρήσεων) φέρουμε την $AB \parallel Ox$ και στη συνέχεια τη $B\Gamma \perp Ox$. Τότε, στο σημείο Γ αντιστοιχεί η διάμεσος δ των παρατηρήσεων.



σχήμα 13

Μέτρα Διασποράς

α) Εύρος

Το απλούστερο από τα μέτρα διασποράς είναι το **εύρος** ή **κύμανση** (range) (R), που ορίζεται ως η διαφορά της ελάχιστης παρατήρησης από τη μέγιστη παρατήρηση, δηλαδή:

Εύρος $R = \text{Μεγαλύτερη παρατήρηση} - \text{Μικρότερη παρατήρηση}$

Όταν έχουμε ομαδοποιημένα δεδομένα, το εύρος δίνεται από τη διαφορά του κατώτερου ορίου της πρώτης κλάσης από το ανώτερο όριο της τελευταίας κλάσης.

Το εύρος είναι ένα αρκετά απλό μέτρο, που υπολογίζεται εύκολα δε θεωρείται όμως αξιόπιστο μέτρο διασποράς, γιατί βασίζεται μόνο στις δυο ακραίες παρατηρήσεις.

β) Διακύμανση

Ως ένα μέτρο διασποράς παίρνουμε τον μέσο όρο των τετραγώνων των αποκλίσεων των t_i από τη μέση τιμή τους $\bar{x}$. Το μέτρο αυτό καλείται **διακύμανση** ή **διασπορά** (variance) και ορίζεται από τη σχέση $s^2 = \frac{1}{v} \sum_{i=1}^v (t_i - \bar{x})^2$ Ο

τύπος αυτός αποδεικνύεται ότι μπορεί να πάρει την ισοδύναμη μορφή: $s^2 = \frac{1}{v} \left\{ \sum_{i=1}^v t_i^2 - \frac{\left(\sum_{i=1}^v t_i \right)^2}{v} \right\}$ η οποία

διευκολύνει σημαντικά τους υπολογισμούς κυρίως όταν η μέση τιμή $\bar{x}$ δεν είναι ακέραιος αριθμός.

Όταν έχουμε πίνακα συχνοτήτων ή ομαδοποιημένα δεδομένα, η διακύμανση ορίζεται από τη σχέση:

$$s^2 = \frac{1}{v} \sum_{i=1}^k (x_i - \bar{x})^2 v_i \text{ ή την ισοδύναμη μορφή: } s^2 = \frac{1}{v} \left\{ \sum_{i=1}^k x_i^2 v_i - \frac{\left(\sum_{i=1}^k x_i v_i \right)^2}{v} \right\}.$$

όπου $x_1, x_2, \dots, x_k$ οι τιμές της μεταβλητής (ή τα κέντρα των κλάσεων) με αντίστοιχες συχνότητες $v_1, v_2, \dots, v_k$.

Κλάσεις [-)	Κεντρικές τιμές x_i	Συχνότητα v_i	x_i^2	$x_i v_i$	$x_i^2 v_i$
156-162	159	2	25281	318	50562
162-168	165	8	27225	1320	217800
168-174	171	12	29241	2052	350892
174-180	177	11	31329	1942	344619
180-186	183	5	33489	915	167445
186-192	189	2	35721	378	71442
Σύνολο:		$v = 40$	—	$\sum x_i v_i =$ 6930	$\sum x_i^2 v_i =$ 1202760

$$\text{Επομένως: } s^2 = \frac{1}{v} \left\{ \sum_{i=1}^k x_i^2 v_i - \frac{\left(\sum_{i=1}^k x_i v_i \right)^2}{v} \right\} = \frac{1}{40} \left\{ 1202760 - \frac{6930^2}{40} \right\} = 53,4$$

γ) Τυπική Απόκλιση

Η διακύμανση είναι μια αξιόπιστη παράμετρος διασποράς, αλλά έχει ένα μειονέκτημα. Δεν εκφράζεται με τις μονάδες με τις οποίες εκφράζονται οι παρατηρήσεις. Για παράδειγμα, αν οι παρατηρήσεις εκφράζονται σε cm, η διακύμανση εκφράζεται σε cm^2 . Αν όμως πάρουμε τη θετική τετραγωνική ρίζα της διακύμανσης, θα έχουμε ένα μέτρο διασποράς που θα εκφράζεται με την ίδια μονάδα μέτρησης του χαρακτηριστικού, όπως ακριβώς είναι και όλα τα άλλα μέτρα θέσης, που εξετάσαμε έως τώρα. Η ποσότητα αυτή λέγεται **τυπική απόκλιση** (standard deviation), συμβολίζεται με s και δίνεται από τη σχέση: $s = \sqrt{s^2}$

Αξίζει να σημειωθεί ότι αν η καμπύλη συχνοτήτων για το χαρακτηριστικό που εξετάζουμε είναι **κανονική ή περίπου κανονική**, τότε η τυπική απόκλιση s έχει τις παρακάτω ιδιότητες:

i) το **68%** περίπου των παρατηρήσεων βρίσκεται στο διάστημα

$$(\bar{x} - s, \bar{x} + s)$$

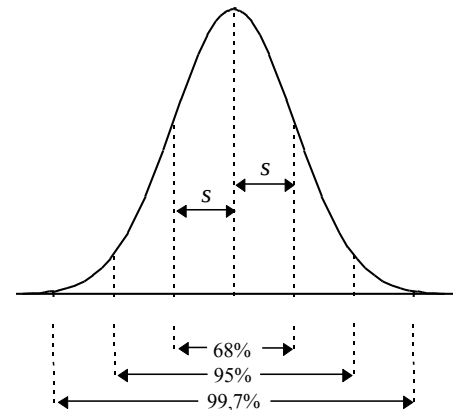
ii) το **95%** περίπου των παρατηρήσεων βρίσκεται στο διάστημα

$$(\bar{x} - 2s, \bar{x} + 2s)$$

iii) το **99,7%** περίπου των παρατηρήσεων βρίσκεται στο διάστημα

$$(\bar{x} - 3s, \bar{x} + 3s)$$

iv) το εύρος ισούται περίπου με έξι τυπικές αποκλίσεις, δηλαδή **$R \approx 6s$** .



δ) Συντελεστής Μεταβολής

Ένα μέτρο με το οποίο μπορούμε να ξεπεράσουμε τις δυσκολίες των διαφορετικών μονάδων μέτρησης των μεταβλητών και το οποίο μας βοηθά στη σύγκριση ομάδων τιμών, που είτε εκφράζονται σε διαφορετικές μονάδες μέτρησης είτε εκφράζονται στην ίδια μονάδα μέτρησης, αλλά έχουν σημαντικά διαφορετικές μέσες τιμές, είναι ο **συντελεστής μεταβολής ή συντελεστής μεταβλητότητας** (coefficient of variation), ο οποίος ορίζεται από το λόγο:

$$CV = \frac{\text{τυπική απόκλιση}}{|\text{μέση τιμή}|} = \frac{s}{|\bar{x}|}$$

Ο συντελεστής μεταβολής επί τοις εκατό, είναι $CV\% = \frac{\text{τυπική απόκλιση}}{|\text{μέση τιμή}|} \cdot 100\% = \frac{s}{|\bar{x}|} \cdot 100\%$ και συνεπώς είναι

ανεξάρτητος από τις μονάδες μέτρησης και παριστάνει ένα μέτρο **σχετικής διασποράς** των τιμών και όχι της απόλυτης διασποράς, όπως έχουμε δει έως τώρα.

Αν $CV_A\% = 12\%$ και $CV_B\% = 25\%$ λέμε ότι έχουμε μεγαλύτερη **ομοιογένεια** στα δεδομένα Α παρά στα Β αφού $12 < 25$.

Γενικά δεχόμαστε ότι ένα δείγμα τιμών μιας μεταβλητής θα είναι ομοιογενές, εάν ο συντελεστής μεταβολής δεν ξεπερνά το 10%. Δηλ. αν $CV\% \leq 10$

Εφαρμογή 3, σελ. 99

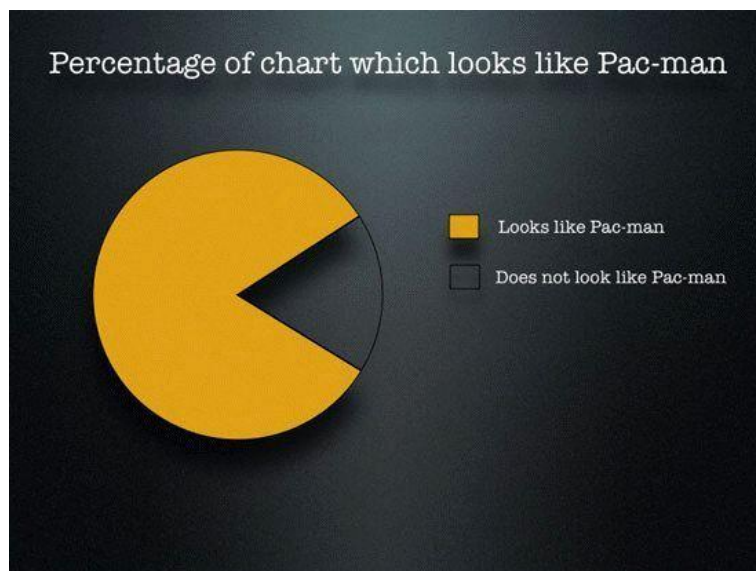
Έστω $x_1, x_2, \dots, x_n$ n παρατηρήσεις με μέση τιμή $\bar{x}$ και τυπική απόκλιση s_x .

α) Αν $y_1, y_2, \dots, y_n$ είναι οι παρατηρήσεις που προκύπτουν αν προσθέσουμε σε καθεμιά από τις $x_1, x_2, \dots, x_n$ μια σταθερά c , τότε:

$$\text{i) } \bar{y} = \bar{x} + c, \quad \text{ii) } s_y = s_x$$

β) Αν $y_1, y_2, \dots, y_n$ είναι οι παρατηρήσεις που προκύπτουν αν πολλαπλασιάσουμε τις $x_1, x_2, \dots, x_n$ επί μια σταθερά c , τότε:

$$\text{i) } \bar{y} = c \bar{x}, \quad \text{ii) } s_y = |c| s_x$$



2. Γενικές οδηγίες εν όψει εξέτασης

1. Παρουσίαση στατιστικών δεδομένων

Για τη συμπλήρωση πινάκων συχνοτήτων θυμόμαστε πάντοτε:

- Τις βασικές ιδιότητες των συχνοτήτων, δηλαδή:
 - ✓ $0 \leq v_i \leq v$ και $v_1 + v_2 + \dots + v_k = v$, όπου v το μέγεθος το δείγματος και k το πλήθος των πιθανών τιμών.
 - ✓ $0 \leq f_i \leq 1$ και $f_1 + f_2 + \dots + f_k = 1$
(αντ. $0 \leq f_i\% \leq 100$ και $f_1\% + f_2\% + \dots + f_k\% = 100$),
 - ✓ $N_1 = v_1, N_{i+1} = N_i + v_i$ για $i = 1, \dots, k-1$ και $N_k = v$,
 - ✓ $F_1 = f_1, F_{i+1} = F_i + f_i$ για $i = 1, \dots, k-1$ και $F_k = 1$
(αντ. $F_1\% = f_1\%, F_{i+1}\% = F_i\% + f_i\%$ για $i = 1, \dots, k-1$ και $F_k\% = 100$)
 - ✓ $\alpha_i = 360^\circ \cdot f_i$ και $\alpha_1 + \alpha_2 + \dots + \alpha_k = 360^\circ$
- Αν το μέγεθος v του δείγματος διαιρεί το 100, τότε οι $f_i\%$ είναι ακέραια πολλαπλάσια των v_i (αφού το πηλίκο $100/v$ είναι ακέραιος). Γενικά, οι συχνότητες v_i, f_i ή $v_i, f_i\%$ είναι πάντοτε ανάλογες.
- Το πλήθος ή το ποσοστό παρατηρήσεων που έχουν τιμή:
 - ✓ το πολύ x_i είναι ίσες με N_i ή $F_i\%$
 - ✓ τουλάχιστον x_i είναι $v - N_{i-1}$ ή $(100 - F_{i-1})\%$
- Για τα ιστογράμματα προσέχουμε:
 - ✓ Το πολύγωνο συχνοτήτων ($v_i, f_i, f_i\%$) έχει κορυφές τα μέσα των άνω βάσεων των ορθογωνίων του ιστογράμματος. Η αρχή και το τέλος είναι τα μέσα δυο υποθετικών κλάσεων συχνότητας 0.
 - ✓ Το πολύγωνο αθροιστικών συχνοτήτων ($v_i, f_i, f_i\%$) έχει κορυφές τα δεξιά άκρα των άνω βάσεων των ορθογωνίων του ιστογράμματος. Η αρχή είναι το αριστερό άκρο της κάτω βάσης του ορθογωνίου της πρώτης κλάσης. Δεν θεωρούμε κλάσεις συχνότητας μηδέν.
 - ✓ Χαρακτηριστικά εμβαδά:
 - Μεταξύ πολυγώνου συχνοτήτων v_i και οριζόντιου άξονα: v
 - Μεταξύ πολυγώνου σχετ. συχν. f_i και οριζόντιου άξονα: 1
 - Μεταξύ πολυγώνου σχετ. συχν. $f_i\%$ και οριζόντιου άξονα: 100
- Στα ομαδοποιημένα δεδομένα (κλάσεις) θεωρούμε πάντοτε ότι οι παρατηρήσεις είναι ομοιόμορφα κατανεμημένες. Αυτό σημαίνει ότι για να βρούμε το πλήθος ή το ποσοστό παρατηρήσεων που ανήκουν σε ένα τμήμα κλάσης, παίρνουμε το αντίστοιχο τμήμα της συχνότητας v_i ή $f_i\%$ της κλάσης.

Παραπομπές/Σημειώσεις

Δες και τα λυμένα παρ. 4 και 5 στις σελ. 130- 131

2. Εύρεση μέτρων θέσης και διασποράς

Τρία από τα 6 μέτρα που έχουμε συναντήσει, το εύρος R , η τυπική απόκλιση S και ο συντελεστής μεταβολής CV υπολογίζονται πάντοτε με τον ίδιο τρόπο :

- ❖ $R = \text{Μέγιστη παρατήρηση} - \text{Ελάχιστη παρατήρηση}$
- ❖ $S = \sqrt{S^2}$

$$\diamond CV = \frac{S}{\bar{X}}$$

Ο ενδεξιγμένος τρόπος εύρεσης των υπόλοιπων μέτρων θέσης και διασποράς είναι:

A. Δείγμα λίγων παρατηρήσεων $t_1, t_2, \dots, t_v$ ($v \leq 15$)

- Η μέση τιμή υπολογίζεται από τον τύπο: $\bar{X} = \frac{1}{v} \cdot \sum_{i=1}^v t_i$
- Για να βρούμε τη διάμεσο δ , διατάσσουμε αρχικά τις παρατηρήσεις σε αύξουσα σειρά και :
 - Αν το πλήθος v των παρατηρήσεων είναι **περιττός** αριθμός, η διάμεσος δ είναι η **μεσαία*** παρατήρηση. [*θέση $(v+1)/2$]
 - Αν το πλήθος v των παρατηρήσεων είναι **άρτιος** αριθμός, η διάμεσος δ είναι το **ημιάθροισμα των δύο μεσαίων**** παρατηρήσεων [**θέσεις $v/2$ και $(v/2)+1$]
- Για να βρούμε τη διακύμανση S^2 χρησιμοποιούμε :
 - τον τύπο $S^2 = \frac{1}{v} \cdot \sum_{i=1}^v (t_i - \bar{X})^2$, όταν η μέση τιμή $\bar{X}$ είναι ακέραιος αριθμός, ή
 - τον τύπο $S^2 = \frac{1}{v} \left\{ \sum_{i=1}^v t_i^2 - \frac{\left(\sum_{i=1}^v t_i \right)^2}{v} \right\}$, όταν η μέση τιμή $\bar{X}$ είναι δεκαδικός αριθμός.

➤ **Παρατήρηση :** Αν στις παρατηρήσεις μας εμφανίζονται συχνές επαναλήψεις των ίδιων αριθμών, είναι προτιμότερο να εργαστούμε με την κατανομή (x_i, v_i) (περίπτωση Β), έστω κι αν το μέγεθος v είναι μικρό.

B. Δείγμα με γνωστή κατανομή συχνοτήτων (x_i, v_i) ή σχετικών συχνοτήτων (x_i, f_i) , όπου x_i οι πιθανές τιμές της μεταβλητής.

- Η μέση τιμή υπολογίζεται από τον τύπο: $\bar{X} = \frac{1}{v} \cdot \sum_{i=1}^k x_i v_i$ ή $\bar{X} = \sum_{i=1}^k x_i f_i$
- Για να βρούμε τη διάμεσο δ , χρησιμοποιούμε την κατανομή των αθροιστικών συχνοτήτων N_i . Συγκεκριμένα:
 - Αν v περιττός, τότε:
 - στην περίπτωση που υπάρχει τιμή $N_i = \frac{v+1}{2}$ (δηλαδή ίση με την θέση της μεσαίας παρατήρησης), τότε $\delta = x_i$.
 - στην περίπτωση που δεν υπάρχει τιμή $N_i = \frac{v+1}{2}$, τότε βρίσκουμε τις τιμές N_i, N_{i+1} για τις οποίες ισχύει: $N_i < \frac{v+1}{2} < N_{i+1}$. Τότε ισχύει: $\delta = x_{i+1}$.

Δες το λυμένο παρ. 7
στη σελ. 156

➤ Αν n άρτιος, τότε:

- στην περίπτωση που υπάρχει τιμή $N_i = \frac{n}{2}$ (δηλαδή ίση με την θέση της πρώτης από τις δύο μεσαίες παρατηρήσεις), τότε $\delta = \frac{x_i + x_{i+1}}{2}$.
- στην περίπτωση που δεν υπάρχει τιμή $N_i = \frac{n}{2}$ (δηλαδή ίση με την θέση της πρώτης από τις δύο μεσαίες παρατηρήσεις), τότε βρίσκουμε τις τιμές N_i, N_{i+1} για τις οποίες ισχύει: $N_i < \frac{n}{2} + 1 \leq N_{i+1}$.
Τότε ισχύει: $\delta = x_{i+1}$

➤ **Παρατήρηση:** Μια σύνθετη περίπτωση είναι να δοθεί η κατανομή σχετικών συχνοτήτων (x_i, f_i) , χωρίς να γνωρίζουμε το μέγεθος n του δείγματος. Τότε χρησιμοποιούμε την αθροιστική σχετική συχνότητα $F_i\%$ ως εξής:

- Αν σε κάποια πιθανή τιμή x_i συμπληρώνεται το 50% των παρατηρήσεων, τότε η διάμεσος δ είναι ίση με: $\delta = \frac{x_i + x_{i+1}}{2}$
- Αν δεν συναντάμε το 50% στις τιμές της $F_i\%$, τότε η διάμεσος ισούται με την πιθανή τιμή x_i , της οποίας η αντίστοιχη τιμή $F_i\%$ ξεπερνάει πρώτη το 50%.

□ Για να βρούμε τη διακύμανση S^2 χρησιμοποιούμε:

- τους τύπους: $S^2 = \frac{1}{n} \cdot \sum_{i=1}^k (x_i - \bar{X})^2 \cdot v_i$ ή $S^2 = \sum_{i=1}^k (x_i - \bar{X})^2 \cdot f_i$, όταν η μέση τιμή $\bar{X}$ είναι ακέραιος αριθμός, ή

- τους τύπους $S^2 = \frac{1}{n} \left\{ \sum_{i=1}^k x_i^2 v_i - \frac{\left(\sum_{i=1}^k x_i v_i \right)^2}{n} \right\}^{\#}$ ή $S^2 = \left\{ \sum_{i=1}^k x_i^2 f_i - \frac{\left(\sum_{i=1}^k x_i f_i \right)^2}{n} \right\}$

όταν η μέση τιμή $\bar{X}$ είναι δεκαδικός αριθμός.

$$(\# \text{ ο τύπος παίρνει και τη μορφή } S^2 = \frac{1}{n} \left\{ \sum_{i=1}^k x_i^2 v_i - n \bar{X}^2 \right\})$$

Γ. Ομαδοποιημένα δεδομένα (x_i οι κεντρικές τιμές των κλάσεων)

□ Η μέση τιμή υπολογίζεται από τον τύπο: $\bar{X} = \frac{1}{n} \cdot \sum_{i=1}^k x_i v_i$ ή $\bar{X} = \sum_{i=1}^k x_i f_i$

□ Για να βρούμε τη διάμεσο δ , κατασκευάζουμε πίνακα αθροιστικών σχετικών συχνοτήτων $F_i\%$ και ενεργούμε ως εξής:

- Αν υπάρχει τιμή $F_i\% = 50\%$, τότε η διάμεσος ισούται με το άνω άκρο της αντίστοιχης κλάσης.
- Αν ισχύει $F_i\% < 50 < F_{i+1}\%$, τότε εφαρμόζουμε την αναλογία:

$$\frac{\delta - (\text{κάτω άκρο της } (i+1) \text{ κλάσης})}{\text{πλάτος κλάσεων}} = \frac{50 - F_i}{f_{i+1}}$$

- Αν δοθεί το πολύγωνο των αθροιστικών σχετικών συχνοτήτων $F_i\%$ τότε από τον κατακόρυφο άξονα, στο ύψος του 50% φέρνουμε οριζόντια γραμμή μέχρι το πολύγωνο. Κατόπιν, φέρνουμε κατακόρυφη από το

Δες το λυμένο παρ. 7
στη σελ. 156

Δες το λυμένο παρ. 8
στη σελ. 157

Δες το λυμένο παρ. 9
στη σελ. 157

σημείο που συναντήσαμε το πολύγωνο προς τον οριζόντιο άξονα. Το σημείο του οριζόντιου άξονα που συναντάμε μας δίνει την τιμή της διαμέσου δ.

Προσοχή: Επειδή με τον παραπάνω τρόπο δεν είναι δυνατό να προσδιοριστεί με ακρίβεια η διάμεσος, πιθανότατα θα μας δίνεται δεδομένο που οδηγεί σε απ'ευθείας εύρεση της διαμέσου.(π.χ. «...η ευθεία $x = 10$ διαιρεί το χωρίο μεταξύ του πολυγώνου συχνοτήτων και του οριζόντιου άξονα σε δυο ισεμβαδικά χωρία...»).

□ Για να βρούμε τη διακύμανση S^2 χρησιμοποιούμε :

➤ τους τύπους : $S^2 = \frac{1}{v} \cdot \sum_{i=1}^k (x_i - \bar{X})^2 \cdot v_i$ ή $S^2 = \sum_{i=1}^k (x_i - \bar{X})^2 \cdot f_i$, όταν η μέση

τιμή $\bar{X}$ είναι ακέραιος αριθμός, ή

➤ τους τύπους $S^2 = \frac{1}{v} \left\{ \sum_{i=1}^k x_i^2 v_i - \frac{\left(\sum_{i=1}^k x_i v_i \right)^2}{v} \right\}^{\#}$ ή $S^2 = \left\{ \sum_{i=1}^k x_i^2 f_i - \frac{\left(\sum_{i=1}^k x_i f_i \right)^2}{n} \right\}$

όταν η μέση τιμή $\bar{X}$ είναι δεκαδικός αριθμός.

([#] ο τύπος παίρνει και τη μορφή $S^2 = \frac{1}{v} \left\{ \sum_{i=1}^k x_i^2 v_i - v \bar{X}^2 \right\}$)

!!! Αν μεταβληθούν ομοιόμορφα οι παρατηρήσεις ενός δείγματος , τότε τα μέτρα δεν υπολογίζονται ξανά από την αρχή. Γνωρίζουμε (εφ. 3 σελ 99) ότι :

1. Αν σε κάθε παρατήρηση προστεθεί ο ίδιος αριθμός c , τότε :

$$\bar{X}' = \bar{X} + c, \delta' = \delta + c, R' = R, S'^2 = S^2 \text{ και } S' = S.$$

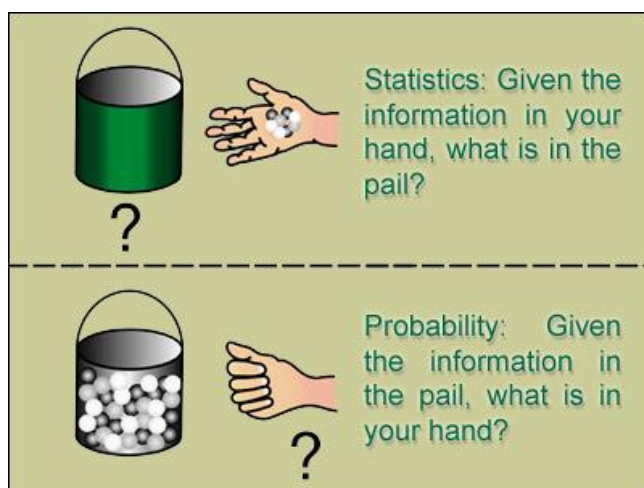
2. Αν κάθε παρατήρηση πολλαπλασιαστεί με τον ίδιο αριθμό $c \neq 0$, τότε :

$$\bar{X}' = c \bar{X}, \delta' = c \delta, R' = |c| R, S'^2 = c^2 \cdot S^2 \text{ και } S' = |c| S.$$

3. Αν κάθε παρατήρηση μεταβληθεί κατά ποσοστό $t\%$, τότε εφαρμόζουμε τα

συμπεράσματα της περίπτωσης (2) με $c = 1 + \frac{t}{100}$ ή $c = 1 - \frac{t}{100}$ για την

αύξηση ή τη μείωση αντίστοιχα .



Δες το λυμένο παρ. 21
στη σελ. 166

Ιδιότητες

Πέρα απ'το να υπολογίζουμε τα διάφορα μέτρα θέσης και διασποράς, θα πρέπει να γνωρίζουμε και τις στατιστικές τους ιδιότητες:

Ιδιότητες μέσης τιμής και διαμέσου

Πλεονεκτήματα

Μειονεκτήματα

Μέση τιμή

- Για τον υπολογισμό της χρησιμοποιούνται όλες οι τιμές.
- Είναι μοναδική για κάθε σύνολο δεδομένων.
- Είναι εύκολα κατανοητή.
- Ο υπολογισμός της είναι σχετικά εύκολος.
- Έχει μεγάλη εφαρμογή για περαιτέρω στατιστική ανάλυση.
- Επηρεάζεται πολύ από ακραίες τιμές.
- Μπορεί να μην αντιστοιχεί σε δυνατή τιμή της μεταβλητής. Όταν η X είναι διακριτή, με ακέραιες τιμές, τότε η μέση τιμή μπορεί να μην είναι ακέραιος.
- Δεν υπολογίζεται για ποιοτικά δεδομένα.

Διάμεσος

- Είναι εύκολα κατανοητή.
- Δεν επηρεάζεται από ακραίες τιμές.
- Υπολογίζεται και στην περίπτωση που οι ακραίες κλάσεις είναι ανοικτές.
- Ο υπολογισμός της είναι απλός.
- Είναι μοναδική σε κάθε σύνολο δεδομένων.
- Δεν χρησιμοποιούνται όλες οι τιμές για τον υπολογισμό της.
- Είναι δύσκολη η εφαρμογή της για περαιτέρω στατιστική ανάλυση.
- Δεν υπολογίζεται για ποιοτικά δεδομένα.
- Για τον υπολογισμό της μπορεί να χρειαστεί προσέγγιση.

ΠΡΟΣΟΧΗ !!!

1. Αν ισχύει η σχέση $\bar{x} < \delta$ έχουμε **αρνητική ασυμμετρία**

Αν ισχύει η σχέση $\bar{x} = \delta$ έχουμε **συμμετρική κατανομή** (όχι κατ'ανάγκη κανονική).

Αν ισχύει η σχέση $\bar{x} > \delta$ έχουμε **θετική ασυμμετρία**.

2. Ισχύει $\sum_{i=1}^v t_i = n \bar{x}$ (πρέπει να θυμάστε αυτή την ιδιότητα, χρησιμεύει στις

περιπτώσεις που συγχωνεύουμε δείγματα ή χωρίζουμε ένα δείγμα σε μικρότερα. Μας οδηγεί στον σταθμικό μέσο, όπως έχουμε δει σε ασκήσεις.)

3. Η μέση τιμή είναι πάντοτε αριθμός ανάμεσα στην μεγαλύτερη και στη μικρότερη παρατήρηση. Η διάμεσος είναι η τιμή για την οποία το πολύ 50% των παρατηρήσεων είναι μικρότερες από αυτήν και το πολύ 50% των παρατηρήσεων είναι μεγαλύτερες από την τιμή αυτή.

Ποιο είναι όμως το καλύτερο μέτρο θέσης μιας κατανομής; Η επιλογή του καταλληλότερου μέτρου θέσης εξαρτάται από το σκοπό για τον οποίο θα χρησιμοποιηθεί. Αν επιθυμούμε περαιτέρω στατιστική επεξεργασία, τότε η

μέση τιμή προσφέρεται περισσότερο. Αν όμως ο σκοπός είναι βασικά περιγραφικός, τότε πρέπει να χρησιμοποιείται το μέτρο που περιγράφει καλύτερα τα δεδομένα. Η παρουσία ακραίων παρατηρήσεων (πολύ μικρών ή πολύ μεγάλων αναφορικά με τις άλλες παρατηρήσεις) είναι συχνά ένα από τα βασικότερα κριτήρια για την επιλογή κατάλληλου μέτρου θέσης. Η διάμεσος μένει γενικά ανεπηρέαστη από τις ακραίες τιμές του δείγματος. Η μέση τιμή όμως επηρεάζεται σημαντικά από τις τιμές αυτές, επομένως δεν ενδείκνυται σε τέτοιες περιπτώσεις.

Ιδιότητες μέτρων διασποράς και σχετικής διασποράς

Πλεονεκτήματα

Μειονεκτήματα

Εύρος

- Είναι πολύ απλό στον υπολογισμό.
- Χρησιμοποιείται αρκετά στον έλεγχο ποιότητας.
- Μπορεί να χρησιμοποιηθεί για την εκτίμηση της τυπικής απόκλισης.
- Δε θεωρείται αξιόπιστο μέτρο διασποράς, επειδή βασίζεται μόνο στις δυο ακραίες παρατηρήσεις.
- Δε χρησιμοποιείται για περαιτέρω στατιστική ανάλυση.

Διασπορά και τυπική απόκλιση

- Λαμβάνονται υπόψη για τον υπολογισμό τους όλες οι παρατηρήσεις.
- Έχουν μεγάλη εφαρμογή στη στατιστική συμπερασματολογία.
- Σε κανονικούς πληθυσμούς το 68%, 95%, 99,7% των παρατηρήσεων βρίσκονται στα διαστήματα $\bar{x} \pm s$, $\bar{x} \pm 2s$ και $\bar{x} \pm 3s$ αντίστοιχα.
- Το κυριότερο μειονέκτημα της διασποράς είναι ότι δεν εκφράζεται στις ίδιες μονάδες με το χαρακτηριστικό. Το μειονέκτημα αυτό παύει να υπάρχει με τη χρησιμοποίηση της τυπικής απόκλισης
- Απαιτούνται περισσότερες αλγεβρικές πράξεις για τον υπολογισμό τους παρά στα άλλα μέτρα.

Συντελεστής μεταβολής

- Είναι καθαρός αριθμός.
- Χρησιμοποιείται ως μέτρο σύγκρισης της μεταβλητότητας, όταν έχουμε ίδιες ή και διαφορετικές μονάδες μέτρησης.
- Χρησιμοποιείται ως μέτρο ομοιογένειας ενός πληθυσμού.
- Δεν ενδείκνυται στην περίπτωση που η μέση τιμή είναι κοντά στο μηδέν.

ΠΡΟΣΟΧΗ !!!

1. Η τυπική απόκλιση είναι πάντοτε μη αρνητικός αριθμός .
2. Ο συντελεστής μεταβολής είναι πάντοτε καθαρός αριθμός (χωρίς μονάδες μέτρησης) και το όριο ομοιογένειας είναι το 10% ή 0,1 (μέχρι και αυτή την τιμή του CV το δείγμα θεωρείται ομοιογενές).
3. Στις περιπτώσεις που οι παρατηρήσεις μεταβάλλονται ομοιόμορφα , για τον συντελεστή μεταβολής CV ισχύουν:

- Στον πολλαπλασιασμό των παρατηρήσεων επί c (ή στη μεταβολή κατά ποσοστό) ο CV μένει ο ίδιος.
- Στην πρόσθεση ενός αριθμού c σε κάθε παρατήρηση ο CV αλλάζει (αφού μένει ίδια η τυπική απόκλιση ενώ η μέση τιμή αλλάζει).

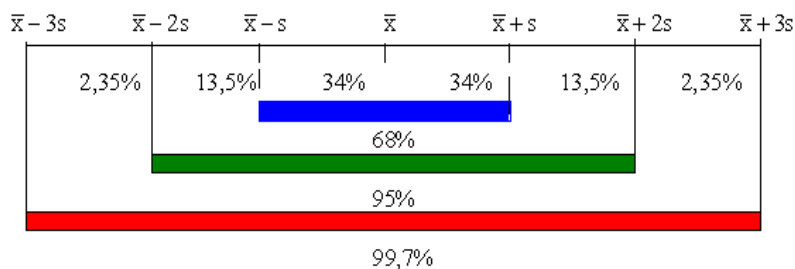
3.Κανονική Κατανομή (Κατανομή Gauss)

Αν έχουμε κανονική ή περίπου κανονική κατανομή παρατηρήσεων, τότε οπωσδήποτε θυμόμαστε τα παρακάτω:

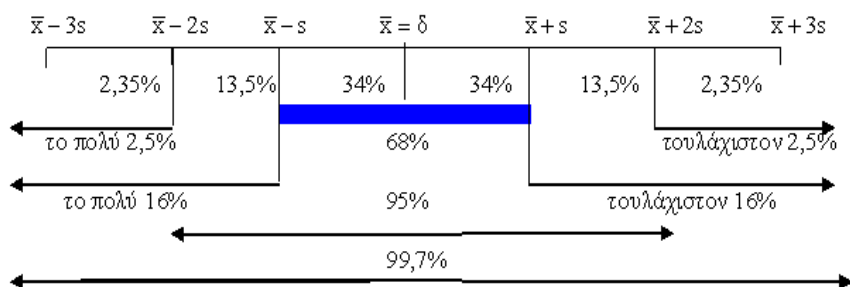
- ◆ Ισχύουν : $\bar{x} = \delta$ και $R \approx 6S$.

Πίνακας κατανομής των παρατηρήσεων της μεταβλητής X

Υπολογισμός του ποσοστού των παρατηρήσεων που βρίσκονται στα αντίστοιχα διαστήματα



Ανάλυση εκφράσεων για τις ασκήσεις - προβλήματα



- ◆ Συνήθως μας δίνονται δύο πληροφορίες σχετικά με τις επτά χαρακτηριστικές τιμές $\bar{x} - 3s$, $\bar{x} - 2s$, $\bar{x} - s$, $\bar{x}$, $\bar{x} + s$, $\bar{x} + 2s$, $\bar{x} + 3s$, οπότε γράφοντας τις αντίστοιχες σχέσεις προκύπτει ένα σύστημα με αγνώστους $\bar{x}$, s το οποίο και λύνουμε.
- ◆ Τα ποσοστά των παρατηρήσεων σε κάθε διάστημα δηλώνουν ταυτόχρονα και την πιθανότητα να επιλέξουμε τυχαία ένα στοιχείο του δείγματος και να ανήκει στο αντίστοιχο διάστημα.

Δες τα λυμένα παρ. 25, 26, 27 στις σελ. 170 – 171

1. ΘΕΩΡΙΑ ΚΕΦΑΛΑΙΟΥ 3 - Πιθανότητες

Πείραμα Τύχης

Κάθε πείραμα κατά το οποίο η γνώση των συνθηκών κάτω από τις οποίες εκτελείται καθορίζει πλήρως το αποτέλεσμα λέγεται **αιτιοκρατικό** (deterministic) πείραμα.

Υπάρχουν όμως και πειράματα των οποίων δεν μπορούμε εκ των προτέρων να προβλέψουμε το αποτέλεσμα, μολονότι επαναλαμβάνονται (φαινομενικά τουλάχιστον) κάτω από τις ίδιες συνθήκες. Ένα τέτοιο πείραμα ονομάζεται **πείραμα τύχης** (random experiment).

Δειγματικός Χώρος

Όλα τα αποτελέσματα που μπορούν να εμφανιστούν σε ένα πείραμα τύχης λέγονται **δυνατά αποτελέσματα** ή **δυνατές περιπτώσεις** του πειράματος. Το σύνολο των δυνατών αποτελεσμάτων λέγεται **δειγματικός χώρος** (sample space) και συμβολίζεται συνήθως με το γράμμα Ω . Αν δηλαδή $\omega_1, \omega_2, \dots, \omega_k$ είναι τα δυνατά αποτελέσματα ενός πειράματος τύχης, τότε ο δειγματικός χώρος του πειράματος θα είναι το σύνολο:

$$\Omega = \{\omega_1, \omega_2, \dots, \omega_k\}.$$

Ενδεχόμενα

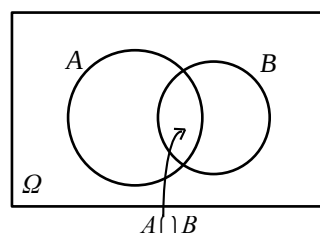
Το σύνολο που έχει ως στοιχεία ένα ή περισσότερα αποτελέσματα ενός πειράματος τύχης λέγεται **ενδεχόμενο** (event) ή γεγονός. Είναι φανερό ότι ένα ενδεχόμενο είναι υποσύνολο του δειγματικού χώρου. Ένα ενδεχόμενο λέγεται **απλό** όταν έχει ένα μόνο στοιχείο και **σύνθετο** αν έχει περισσότερα στοιχεία. Όταν το αποτέλεσμα ενός πειράματος, σε μια συγκεκριμένη εκτέλεσή του είναι στοιχείο ενός ενδεχομένου, τότε λέμε ότι το ενδεχόμενο αυτό **πραγματοποιείται** ή **συμβαίνει**. Γι'αυτό τα στοιχεία ενός ενδεχομένου λέγονται και **ευνοϊκές περιπτώσεις** για την πραγματοποίησή του. Ο ίδιος ο δειγματικός χώρος Ω ενός πειράματος θεωρείται ότι είναι ενδεχόμενο, το οποίο μάλιστα πραγματοποιείται πάντοτε, αφού όποιο και αν είναι το αποτέλεσμα του πειράματος θα ανήκει στο Ω . Γι'αυτό το Ω λέγεται **βέβαιο ενδεχόμενο**. Δεχόμαστε ακόμα ως ενδεχόμενο και το κενό σύνολο $\emptyset$ που δεν πραγματοποιείται σε καμιά εκτέλεση του πειράματος τύχης. Γι'αυτό λέμε ότι το $\emptyset$ είναι **το αδύνατο ενδεχόμενο**.

Το πλήθος των στοιχείων ενός ενδεχομένου A θα το συμβολίζουμε με $N(A)$.

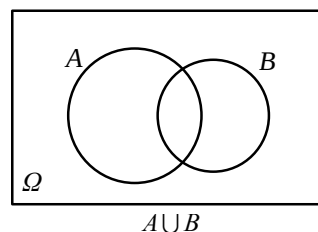
Πράξεις με Ενδεχόμενα

Τα ενδεχόμενα είναι υποσύνολα του δειγματικού χώρου Ω . Επομένως, μεταξύ των ενδεχομένων ενός πειράματος μπορούν να οριστούν οι γνωστές πράξεις μεταξύ των συνόλων, από τις οποίες προκύπτουν νέα ενδεχόμενα. Έτσι, αν A και B είναι δύο ενδεχόμενα, έχουμε:

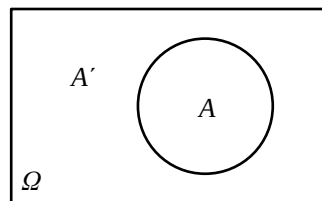
- Το ενδεχόμενο $A \cap B$, που διαβάζεται “ A τομή B ” ή “ A και B ” και πραγματοποιείται, όταν πραγματοποιούνται συγχρόνως τα A και B .



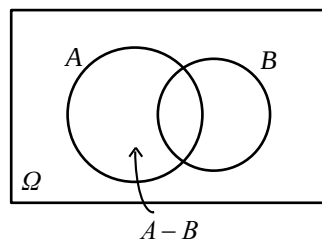
- Το ενδεχόμενο $A \cup B$, που διαβάζεται “ A ένωση B ” ή “ A ή B ” και πραγματοποιείται, όταν πραγματοποιείται ένα τουλάχιστον από τα A , B .



- Το ενδεχόμενο A' , που διαβάζεται “όχι A ” ή “συμπληρωματικό του A ” και πραγματοποιείται, όταν δεν πραγματοποιείται το A . Το A' λέγεται και “αντίθετο του A ”.



- Το ενδεχόμενο $A - B$, που διαβάζεται “διαφορά του B από το A ” και πραγματοποιείται, όταν πραγματοποιείται το A αλλά όχι το B . Είναι εύκολο να δούμε ότι $A - B = A \cap B'$.



Στον παρακάτω πίνακα τα A και B συμβολίζουν ενδεχόμενα ενός πειράματος και το ω ένα αποτέλεσμα του πειράματος αυτού. Στην αριστερή στήλη του πίνακα αναγράφονται διάφορες σχέσεις για τα A και B διατυπωμένες στην κοινή γλώσσα, και στη δεξιά στήλη αναγράφονται οι ίδιες σχέσεις αλλά διατυπωμένες στη γλώσσα των συνόλων.

Το ενδεχόμενο A πραγματοποιείται	$\omega \in A$
Το ενδεχόμενο A δεν πραγματοποιείται	$\omega \in A'$ (ή $\omega \notin A$)
Ένα τουλάχιστον από τα A και B πραγματοποιείται	$\omega \in A \cup B$
Πραγματοποιούνται αμφότερα τα A και B	$\omega \in A \cap B$
Δεν πραγματοποιείται κανένα από τα A και B	$\omega \in (A \cup B)'$
Πραγματοποιείται μόνο το A	$\omega \in A - B$ (ή $\omega \in A \cap B'$)
Η πραγματοποίηση του A συνεπάγεται την πραγματοποίηση του B	$A \subseteq B$

Ασυμβίβαστα Ενδεχόμενα

Δύο ενδεχόμενα A και B λέγονται **ασυμβίβαστα**, όταν $A \cap B = \emptyset$.

Δύο ασυμβίβαστα ενδεχόμενα λέγονται επίσης **ξένα μεταξύ τους** ή **αμοιβαίως αποκλειόμενα**.

Έννοια και Ιδιότητες Σχετικής Συχνότητας

Αν σε n εκτελέσεις ενός πειράματος ένα ενδεχόμενο A πραγματοποιείται k φορές, τότε ο λόγος $\frac{k}{n}$ ονομάζεται **σχετική συχνότητα του A** και συμβολίζεται με f_A . Ιδιαίτερα αν ο δειγματικός χώρος ενός πειράματος είναι το πεπερασμένο σύνολο $\Omega = \{\omega_1, \omega_2, \dots, \omega_\lambda\}$ και σε n εκτελέσεις του πειράματος αυτού τα απλά ενδεχόμενα $\{\omega_1\}, \{\omega_2\}, \dots, \{\omega_\lambda\}$ πραγματοποιούνται $\kappa_1, \kappa_2, \dots, \kappa_\lambda$ φορές αντιστοίχως, τότε για τις σχετικές συχνότητες $f_1 = \frac{\kappa_1}{n}, f_2 = \frac{\kappa_2}{n}, \dots, f_\lambda = \frac{\kappa_\lambda}{n}$ των απλών ενδεχομένων θα έχουμε:

$$1. \quad 0 \leq f_i \leq 1, \quad i=1,2,\dots,\lambda \quad (\text{αφού } 0 \leq \kappa_i \leq n)$$

$$2. f_1 + f_2 + \dots + f_\lambda = \frac{\kappa_1 + \kappa_2 + \dots + \kappa_\lambda}{\nu} = \frac{\nu}{\nu} = 1.$$

Το συμπέρασμα ότι οι σχετικές συχνότητες πραγματοποίησης των ενδεχομένων ενός πειράματος σταθεροποιούνται γύρω από κάποιους αριθμούς (όχι πάντοτε ίδιους), καθώς ο αριθμός των δοκιμών του πειράματος επαναλαμβάνεται απεριόριστα ονομάζεται **στατιστική ομαλότητα** ή **νόμος των μεγάλων αριθμών**. Το εμπειρικό αυτό εξαγόμενο επιβεβαιώνεται και θεωρητικά.

Κλασικός Ορισμός Πιθανότητας

Γενικά, σε ένα πείραμα με ν ισοπίθانا αποτελέσματα η σχετική συχνότητα ενός ενδεχομένου με κ στοιχεία θα τείνει στον αριθμό $\frac{\kappa}{\nu}$. Γι' αυτό είναι εύλογο σε ένα πείραμα με ισοπίθانا αποτελέσματα να ορίσουμε ως πιθανότητα του ενδεχομένου A τον αριθμό:

$$P(A) = \frac{\text{Πλήθος Ευνοϊκών Περιπτώσεων}}{\text{Πλήθος Δυνατών Περιπτώσεων}} = \frac{N(A)}{N(\Omega)}.$$

Έτσι, έχουμε τον **κλασικό** ορισμό της πιθανότητας, που διατυπώθηκε από τον Laplace το 1812.

Από τον προηγούμενο ορισμό προκύπτει άμεσα ότι:

1. $P(\Omega) = \frac{N(\Omega)}{N(\Omega)} = 1$
2. $P(\emptyset) = \frac{0}{N(\Omega)} = 0$
3. Για κάθε ενδεχόμενο A ισχύει $0 \leq P(A) \leq 1$, αφού το πλήθος των στοιχείων ενός ενδεχομένου είναι ίσο ή μικρότερο από το πλήθος των στοιχείων του δειγματικού χώρου.

Αξιωματικός Ορισμός Πιθανότητας

Έστω $\Omega = \{\omega_1, \omega_2, \dots, \omega_\nu\}$ ένας δειγματικός χώρος με πεπερασμένο πλήθος στοιχείων. Σε κάθε απλό ενδεχόμενο $\{\omega_i\}$ αντιστοιχίζουμε έναν πραγματικό αριθμό, που τον συμβολίζουμε με $P(\omega_i)$, έτσι ώστε να ισχύουν:

- $0 \leq P(\omega_i) \leq 1$
- $P(\omega_1) + P(\omega_2) + \dots + P(\omega_\nu) = 1.$

Τον αριθμό $P(\omega_i)$ ονομάζουμε πιθανότητα του ενδεχομένου $\{\omega_i\}$.

Ως πιθανότητα $P(A)$ ενός ενδεχομένου $A = \{\alpha_1, \alpha_2, \dots, \alpha_\kappa\} \neq \emptyset$ ορίζουμε το άθροισμα $P(\alpha_1) + P(\alpha_2) + \dots + P(\alpha_\kappa)$, ενώ ως πιθανότητα του αδύνατου ενδεχομένου $\emptyset$ ορίζουμε τον αριθμό $P(\emptyset) = 0$.

Αν $P(\omega_i) = \frac{1}{\nu}$, $i = 1, 2, \dots, \nu$, τότε έχουμε τον κλασικό ορισμό της πιθανότητας ενός ενδεχομένου. Στην πράξη, ιδιαίτερα στην περίπτωση που δεν ισχύει ο κλασικός ορισμός της πιθανότητας, ως πιθανότητα ενός ενδεχομένου A λαμβάνεται το όριο της σχετικής του συχνότητας.

ΣΧΟΛΙΟ

Όταν έχουμε ένα δειγματικό χώρο $\Omega = \{\omega_1, \omega_2, \dots, \omega_\nu\}$ και χρησιμοποιούμε τη φράση “παίρνουμε τυχαία ένα στοιχείο του Ω ”, εννοούμε ότι όλα τα δυνατά αποτελέσματα είναι ισοπίθانا με πιθανότητα $P(\omega_i) = \frac{1}{\nu}$, $i = 1, 2, \dots, \nu$.

Κανόνες Λογισμού των Πιθανοτήτων

Για τις πιθανότητες των ενδεχομένων ενός δειγματικού χώρου Ω ισχύουν οι παρακάτω ιδιότητες, γνωστές ως “κανόνες λογισμού των πιθανοτήτων”. Οι κανόνες αυτοί θα αποδειχθούν στην περίπτωση που τα απλά ενδεχόμενα είναι ισοπίθανα. Αποδεικνύεται όμως ότι ισχύουν και στην περίπτωση που τα απλά ενδεχόμενα δεν είναι ισοπίθανα.

1. Για οποιαδήποτε **ασυμβίβαστα** μεταξύ τους ενδεχόμενα A και B ισχύει:

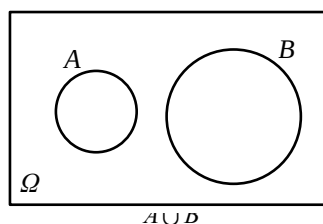
$$P(A \cup B) = P(A) + P(B)$$

ΑΠΟΔΕΙΞΗ

Αν $N(A) = \kappa$ και $N(B) = \lambda$, τότε το $A \cup B$ έχει $\kappa + \lambda$ στοιχεία, γιατί αλλιώς τα A και B δε θα ήταν ασυμβίβαστα. Δηλαδή, έχουμε $N(A \cup B) = \kappa + \lambda = N(A) + N(B)$.

Επομένως:

$$\begin{aligned} P(A \cup B) &= \frac{N(A \cup B)}{N(\Omega)} \\ &= \frac{N(A) + N(B)}{N(\Omega)} \\ &= \frac{N(A)}{N(\Omega)} + \frac{N(B)}{N(\Omega)} \\ &= P(A) + P(B). \end{aligned}$$



Η ιδιότητα αυτή είναι γνωστή ως **απλός προσθετικός νόμος** (simply additive law) και ισχύει και για περισσότερα από δύο ενδεχόμενα. Έτσι, αν τα ενδεχόμενα A , B και Γ είναι ανά δύο ασυμβίβαστα θα έχουμε $P(A \cup B \cup \Gamma) = P(A) + P(B) + P(\Gamma)$.

2. Για δύο συμπληρωματικά ενδεχόμενα A και A' ισχύει:

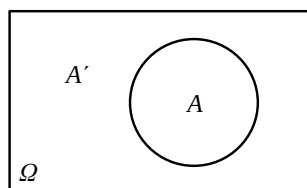
$$P(A') = 1 - P(A)$$

ΑΠΟΔΕΙΞΗ

Επειδή $A \cap A' = \emptyset$, δηλαδή τα A και A' είναι ασυμβίβαστα, έχουμε διαδοχικά, σύμφωνα με τον απλό προσθετικό νόμο:

$$\begin{aligned} P(A \cup A') &= P(A) + P(A') \\ P(\Omega) &= P(A) + P(A') \\ 1 &= P(A) + P(A'). \end{aligned}$$

Οπότε $P(A') = 1 - P(A)$.



3. Για δύο ενδεχόμενα A και B ενός δειγματικού χώρου Ω ισχύει:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

ΑΠΟΔΕΙΞΗ

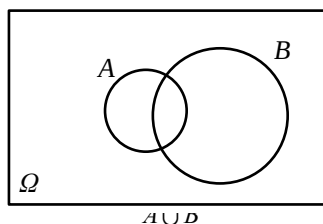
Για δυο ενδεχόμενα A και B έχουμε

$$N(A \cup B) = N(A) + N(B) - N(A \cap B), \quad (1)$$

αφού στο άθροισμα $N(A) + N(B)$ το πλήθος των στοιχείων του $A \cap B$ υπολογίζεται δυο φορές.

Αν διαιρέσουμε τα μέλη της (1) με $N(\Omega)$ έχουμε:

$$\frac{N(A \cup B)}{N(\Omega)} = \frac{N(A)}{N(\Omega)} + \frac{N(B)}{N(\Omega)} - \frac{N(A \cap B)}{N(\Omega)}$$



και επομένως

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Η ιδιότητα αυτή είναι γνωστή ως **προσθετικός νόμος** (additive law).

4. $\boxed{\text{Αν } A \subseteq B, \text{ τότε } P(A) \leq P(B)}$

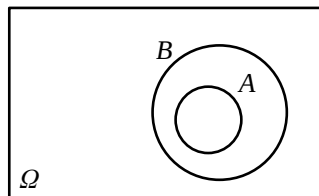
ΑΠΟΔΕΙΞΗ

Επειδή $A \subseteq B$ έχουμε διαδοχικά:

$$N(A) \leq N(B)$$

$$\frac{N(A)}{N(\Omega)} \leq \frac{N(B)}{N(\Omega)}$$

$$P(A) \leq P(B).$$



5. Για δύο ενδεχόμενα A και B ενός δειγματικού χώρου Ω ισχύει

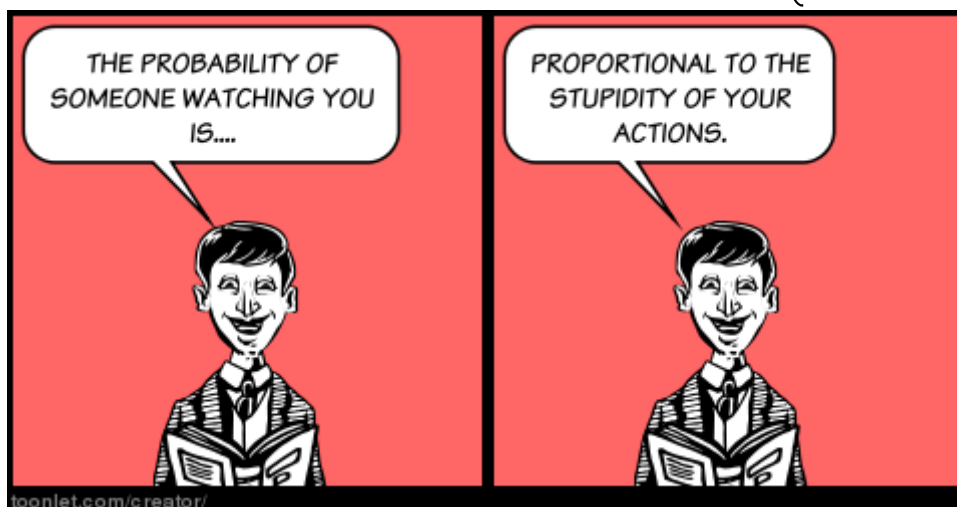
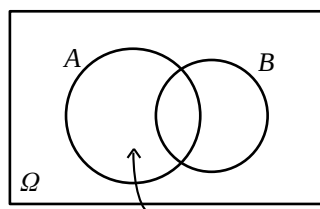
$$P(A - B) = P(A) - P(A \cap B).$$

ΑΠΟΔΕΙΞΗ

Επειδή τα ενδεχόμενα $A - B$ και $A \cap B$ είναι ασυμβίβαστα και $(A - B) \cup (A \cap B) = A$, έχουμε:

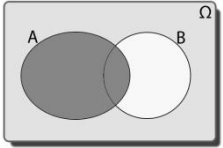
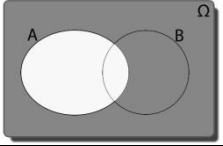
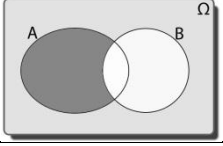
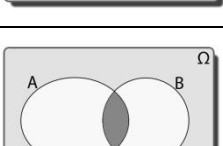
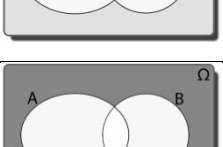
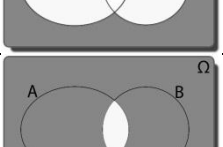
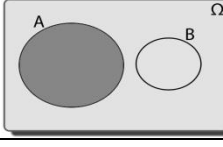
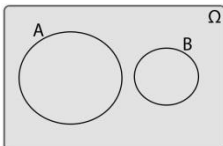
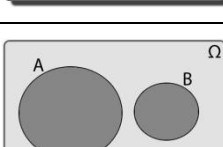
$$P(A) = P(A - B) + P(A \cap B).$$

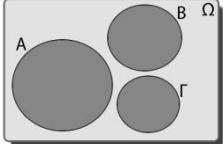
Άρα $P(A - B) = P(A) - P(A \cap B).$



2. Γενικές οδηγίες εν όψει εξέτασης

1. Όταν ένα πείραμα τύχης αποτελείται από τουλάχιστον δύο φάσεις (π.χ. ρίψη ενός ζαριού και ενός κέρματος), τότε ο δειγματικός χώρος Ω προσδιορίζεται ευκολότερα με την βοήθεια ενός δένδροδιαγράμματος, στο οποίο περιγράφουμε διαδοχικά κάθε φάση και τα αποτελέσματά της.
2. **Ενδεχόμενα – Διαγράμματα Venn – Πιθανότητες:**
Οι παρακάτω πίνακες είναι απολύτως απαραίτητοι στην επίλυση ασκήσεων:

	Ενδεχόμενο	Συμβολισμός	Διάγραμμα	Πιθανότητα
1	Πραγματοποιείται το A	A		$P(A)$
2	Δεν πραγματοποιείται το A	A'		$P(A') = 1 - P(A)$
3	Πραγματοποιείται μόνο το A	$A - B = A \cap B'$		$P(A-B) = P(A) - P(A \cap B)$
4	Πραγματοποιείται το A ή το B (Πραγματοποιείται ένα τουλάχιστον από τα A και B)	$A \cup B$		$P(A \cup B) = P(A) + P(B) - P(A \cap B)$
5	Πραγματοποιείται το A και το B (Πραγματοποιούνται τα A και B συγχρόνως)	$A \cap B$		$P(A \cap B)$
6	Δεν πραγματοποιείται κανένα από τα A και B	$(A \cup B)' = A' \cap B'$		$P((A \cup B)') = 1 - P(A \cup B)$
7	Πραγματοποιείται το πολύ ένα από τα A και B	$(A \cap B)' = A' \cup B'$		$P((A \cap B)') = 1 - P(A \cap B)$
8	Πραγματοποιείται μόνο ένα από τα A και B	$(A - B) \cup (B - A)$		$P((A - B) \cup (B - A)) = P(A - B) + P(B - A)$
Ασυμβίβαστα ενδεχόμενα				
1	Πραγματοποιείται μόνο το A	$A - B = A$		$P(A - B) = P(A)$
2	Πραγματοποιείται το A και το B (Πραγματοποιούνται τα A και B συγχρόνως)	$A \cap B = \emptyset$		$P(A \cap B) = 0$
3	Πραγματοποιείται το A ή το B (Πραγματοποιείται ένα τουλάχιστον από τα A και B)	$A \cup B$		$P(A \cup B) = P(A) + P(B)$

4	Πραγματοποιείται ένα τουλάχιστον από τα A, B και Γ	$A \cup B \cup \Gamma$		$P(A \cup B \cup \Gamma) = P(A) + P(B) + P(\Gamma)$
---	---	------------------------	--	---

3. Όταν για κάποιο δειγματικό χώρο $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ χρησιμοποιούνται φράσεις όπως «...επιλέγουμε τυχαία ένα στοιχείο του Ω ...», εννοείται ότι όλα τα δυνατά αποτελέσματα είναι ισοπίθανα με πιθανότητα: $P(\omega_i) = \frac{1}{n}, i = 1, \dots, n$.
4. Αν ισχύει: $P(A) + P(B) > 1$, τότε τα ενδεχόμενα A, B δεν είναι ασυμβίβαστα. Το αντίστροφο δεν ισχύει γενικά.
5. Σε δειγματικό χώρο Ω με ισοπίθανα ενδεχόμενα ισχύουν οι ισοδυναμίες:
 - $P(A) = 1 \Leftrightarrow A = \Omega$
 - $P(A) = 0 \Leftrightarrow A = \emptyset$
 Οι παραπάνω ισοδυναμίες δεν ισχύουν γενικά σε δειγματικό χώρο Ω με μη ισοπίθανα ενδεχόμενα.
6. Αν έχουμε δειγματικό χώρο $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ με **μη ισοπίθανα** ενδεχόμενα, τότε πρέπει αρχικά είναι να βρούμε όλες τις πιθανότητες $P(\omega_i), i = 1, 2, \dots, n$. Για να το πετύχουμε αυτό συνδυάζουμε κατάλληλα τα δεδομένα με την βασική ιδιότητα: $P(\omega_1) + P(\omega_2) + \dots + P(\omega_n) = 1$.
7. **Συγκεντρωτικός πίνακας τύπων:**
 - i. $P(A \cup B) = P(A) + P(B) - P(A \cap B) \Leftrightarrow P(A) + P(B) = P(A \cup B) + P(A \cap B)$
 - ii. $P(A') = 1 - P(A)$
 - iii. Αν $A \subseteq B$, τότε $P(A) \leq P(B)$.
 - iv. $P(A - B) = P(A \cap B') = P(A) - P(A \cap B)$
 - v. $P(B - A) = P(B \cap A') = P(B) - P(A \cap B)$
8. **Χρήσιμες σχέσεις για τον υπολογισμό πιθανοτήτων:**
 - Δεν ξεχνάμε ποτέ ότι: $0 \leq P(A) \leq 1$, για οποιοδήποτε ενδεχόμενο $A \subseteq \Omega$.
 - Με τη βοήθεια του διπλανού διαγράμματος Venn προκύπτουν οι σχέσεις:
 - i. $P(A \cap B) \leq P(A) \leq P(A \cup B)$
 - ii. $P(A \cap B) \leq P(B) \leq P(A \cup B)$
 - iii. $P(A - B) \leq P(A) \leq P(A \cup B)$
 - iv. $P(B - A) \leq P(B) \leq P(A \cup B)$
 - v. $P(A' \cap B') \leq P(A' \cup B')$
 - Αν A, B ενδεχόμενα ενός δειγματικού χώρου Ω , τότε ισχύουν οι ισότητες:
 - i. $(A \cap B)' = A' \cup B'$
 - ii. $(A \cup B)' = A' \cap B'$ **(Κανόνες De Morgan)**
 - Αν μας ζητείται να αποδείξουμε σχέση που περιέχει τις πιθανότητες $P(A)$ και $P(A')$, τότε χρησιμοποιούμε την βασική ιδιότητα $P(A') = 1 - P(A)$ ώστε να απαλείψουμε την $P(A')$ (ή την $P(A)$). Ανάλογα εργαζόμαστε και για να αξιοποιήσουμε δεδομένη σχέση που περιέχει τις $P(A)$ και $P(A')$.