Στατιστική Ι

Γιώργος Τσιρογιάννης

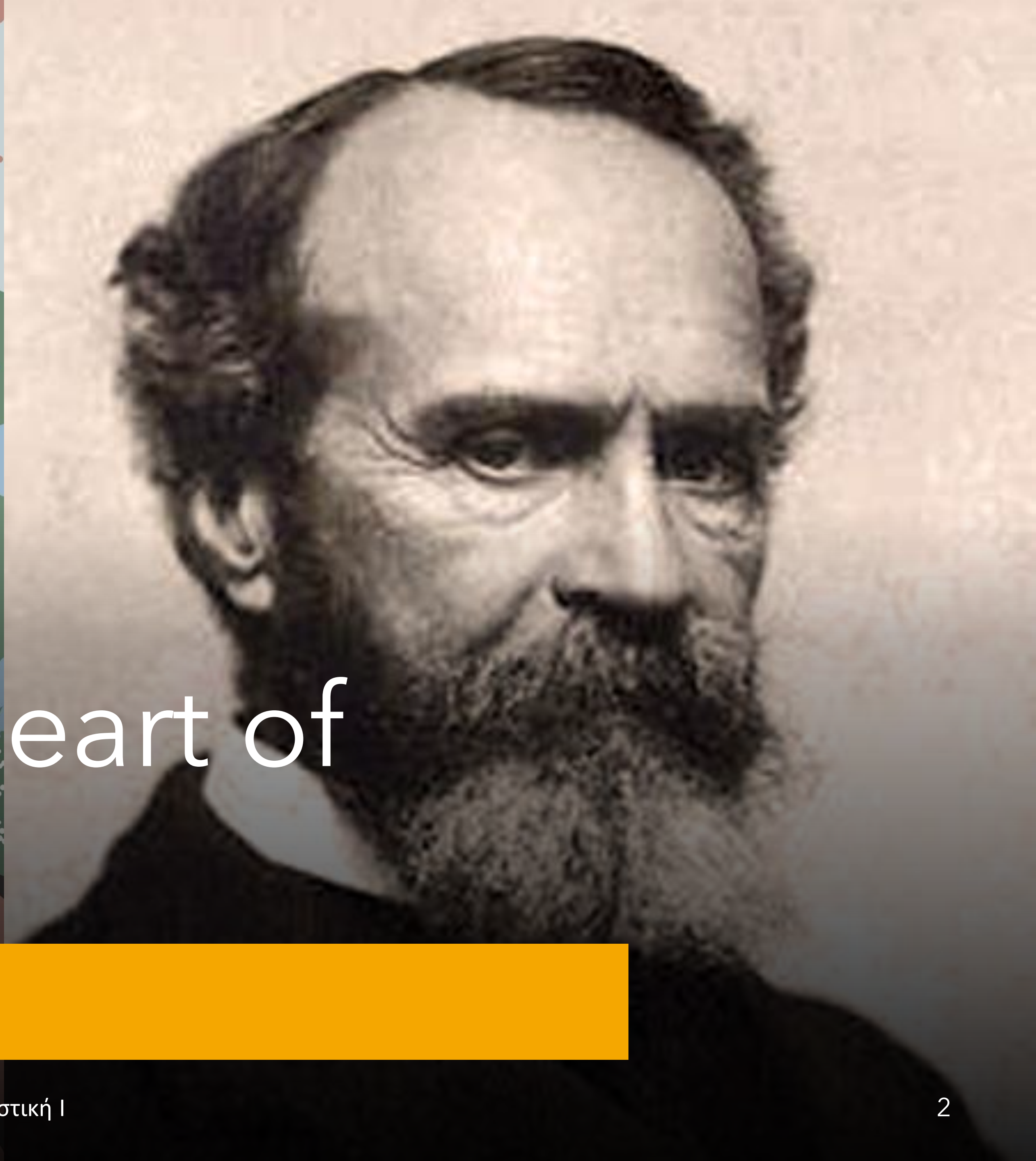
Τμήμα Διοίκησης Επιχειρήσεων Αγροτικών Προϊόντων και Τροφίμων,

Πανεπιστήμιο Πατρών

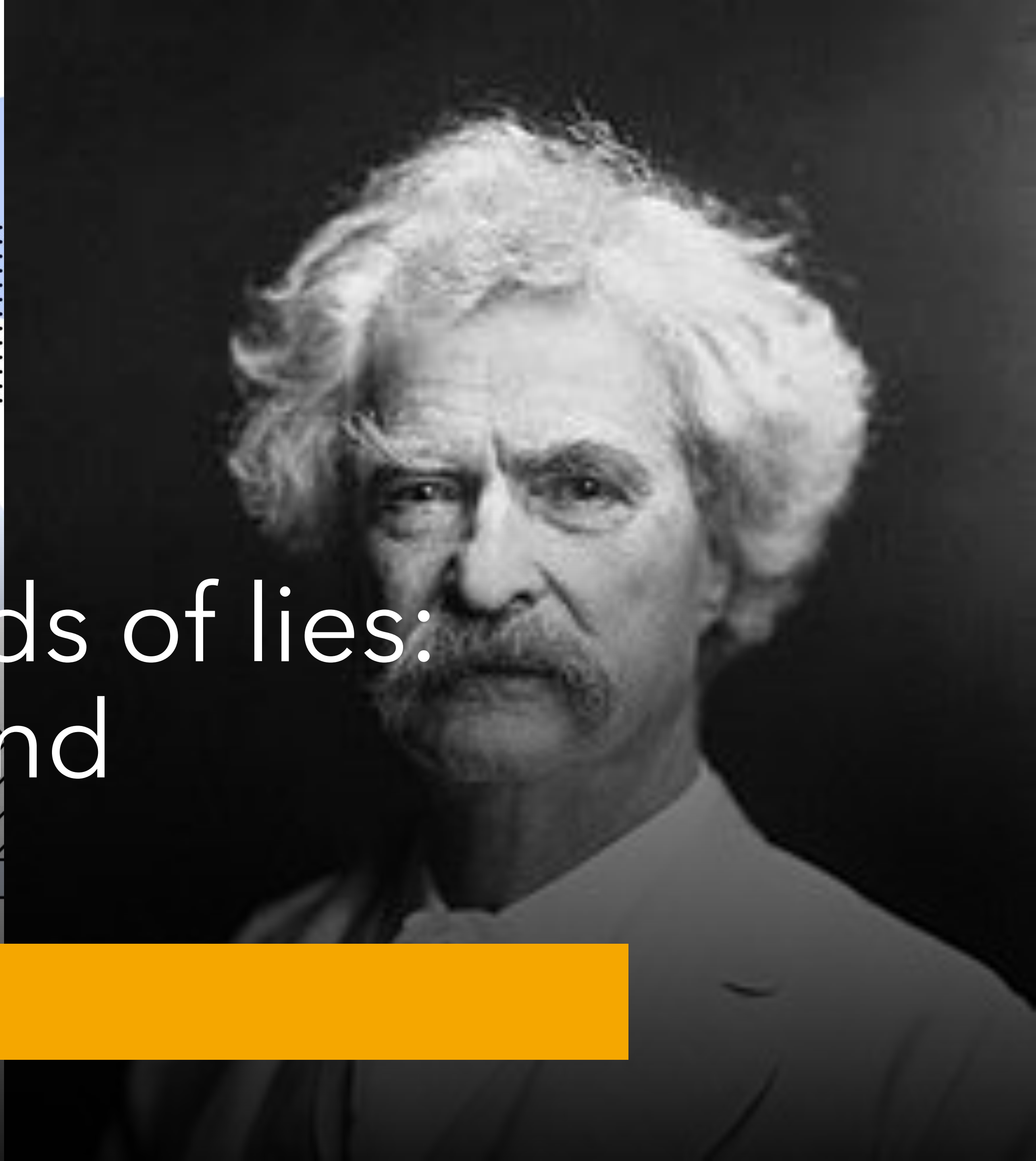
Ακαδημαϊκό έτος: 2020-2021

Statistics are the heart of democracy.

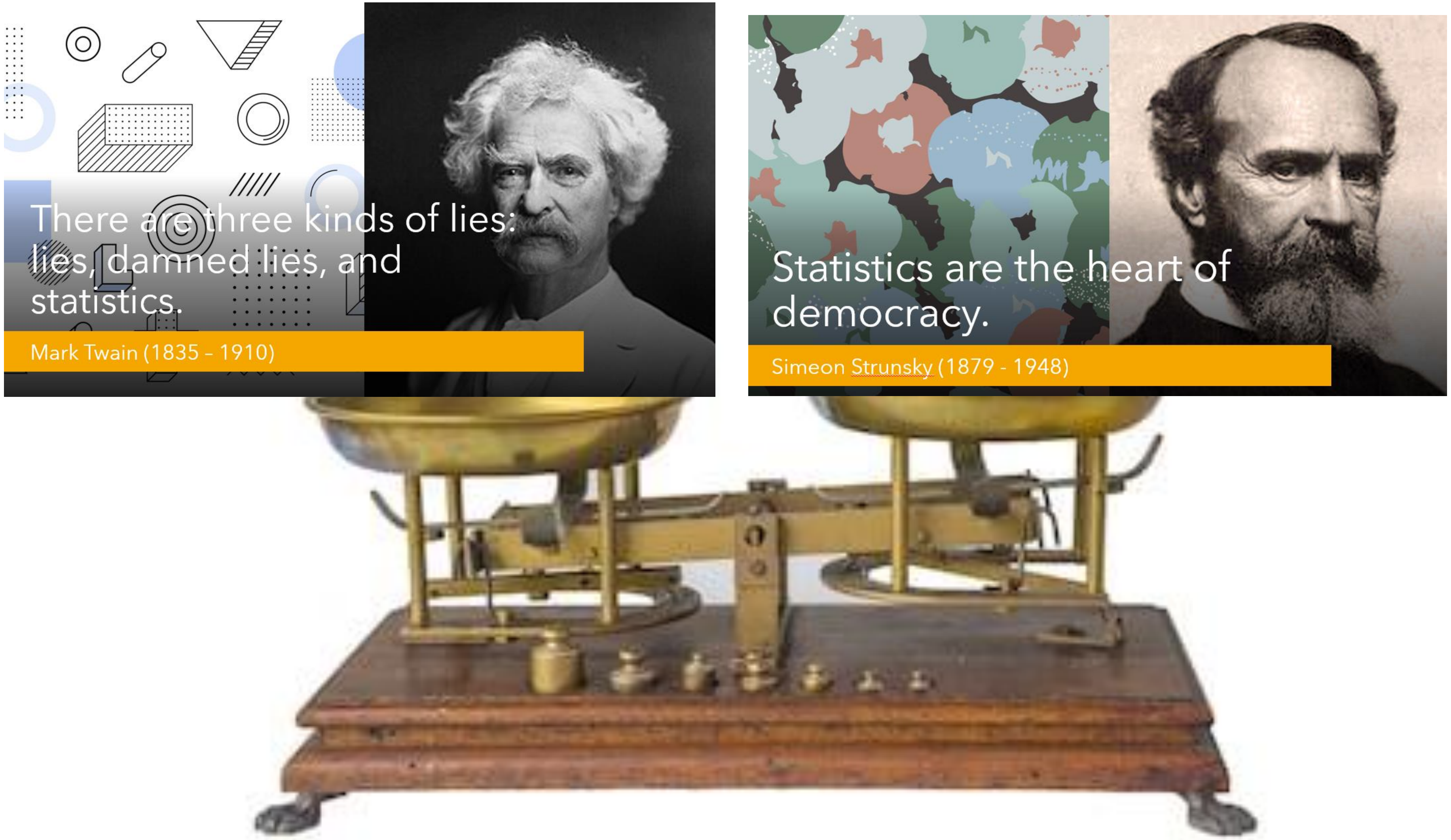
Simeon Strunsky (1879 - 1948)

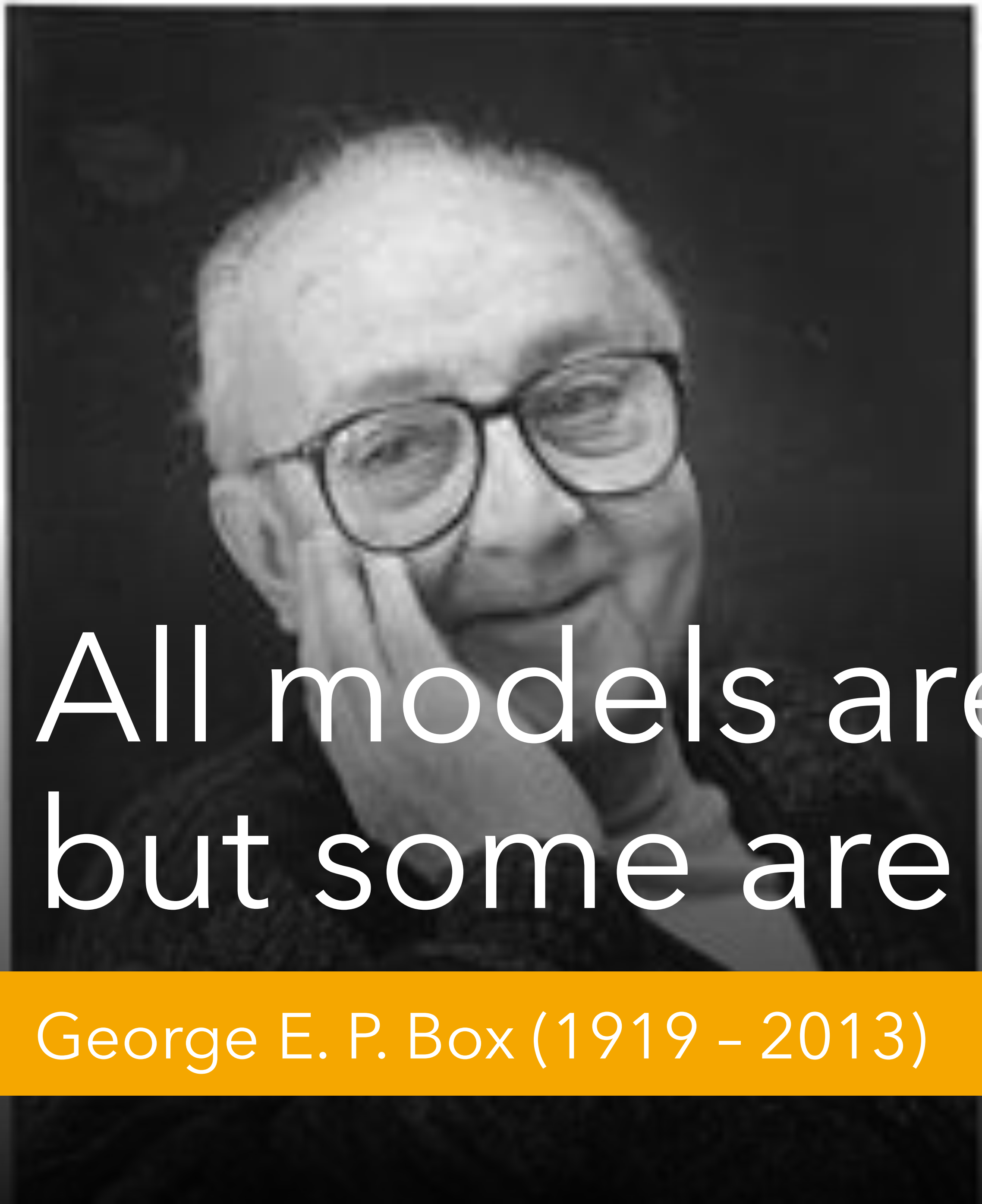


There are three kinds of lies: lies, damned lies, and statistics.



Mark Twain (1835 – 1910)


There are three kinds of lies: lies, damned lies, and statistics.
Mark Twain (1835 – 1910)
Statistics are the heart of democracy.
Simeon Strunsky (1879 - 1948)



All models are wrong, but some are useful.

George E. P. Box (1919 – 2013)

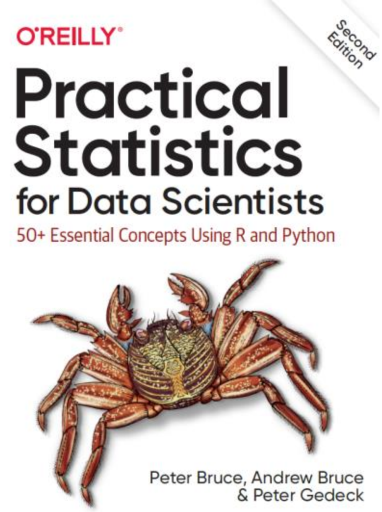
Υλικό μαθήματος

Κύριο σύγγραμμα:

- Εισαγωγή στις Πιθανότητες και τη Στατιστική, Γ. Παπαδόπουλος, 9789600117004

Παραδείγματα σε python πχ:

- Practical Statistics for Data Scientists, 2nd edition, 9781492072942



Διάλεξη 1η

Απαρίθμηση

Πολλαπλασιαστική αρχή

Διατάξεις και Μεταθέσεις

Επαναληπτικές διατάξεις

Μεταθέσεις κ ειδών στοιχείων

Συνδυασμοί

Συνδυασμοί με επανάληψη

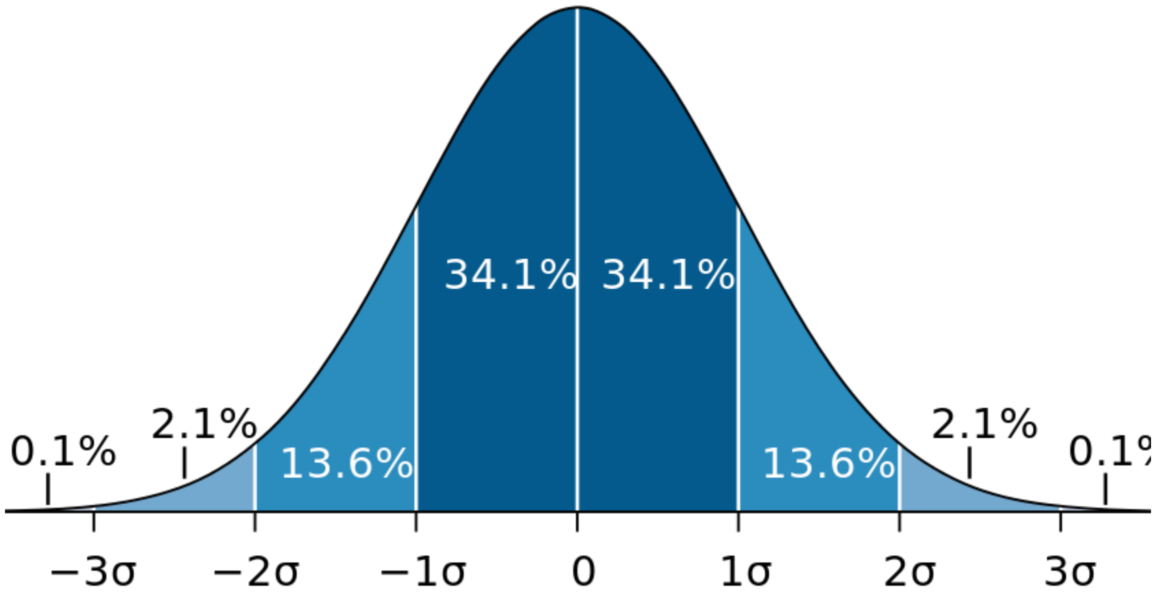
2.1 έως 2.7

Θεωρία Πιθανοτήτων

Κλάδος των μαθηματικών

Πιθανότητα είναι το **μέτρο προσδοκίας** ότι ένα γεγονός (ή ενδεχόμενο) θα συμβεί. Η πιθανότητα είναι ποσοτικά προσδιορισμένη ως νούμερο ανάμεσα στο 0 και το 1 (όπου το 0 υποδεικνύει το αδύνατο και το 1 το βέβαιο). Όσο μεγαλύτερη η πιθανότητα για ένα γεγονός, με τόσο μεγαλύτερη προσδοκία αναμένουμε ότι το γεγονός αυτό θα συμβεί.

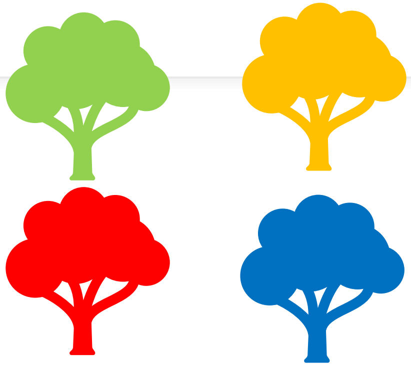
ΒΙΚΙΠΑΙΔΕΙΑ
Η ελεύθερη εγκυκλοπαίδεια



Απαρίθμηση

Πολλαπλασιαστική αρχή

- Παράδειγμα: 4 δένδρα



Πόσοι τρόποι να φυτέψουμε μια δενδροστοιχία με όλα τα δένδρα;

4 επιλογές δένδρων

3 επιλογές δένδρων

2 επιλογές δένδρων

1 επιλογή δένδρου

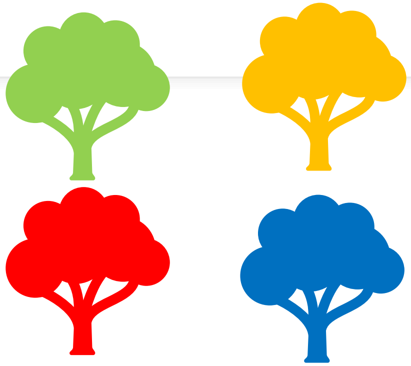
4 * 3 * 2 * 1 = 24 τρόποι

Πολλαπλασιαστική αρχή

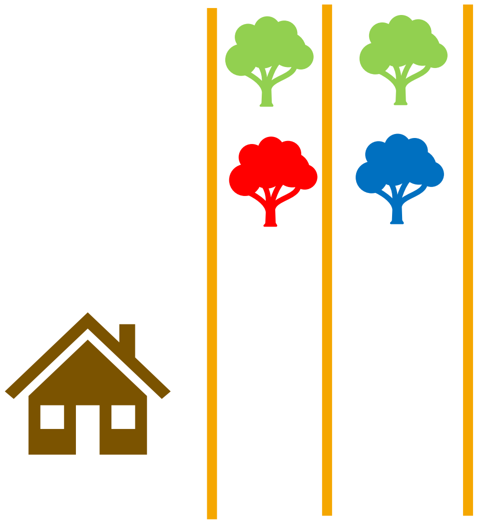
- Αν η απαρίθμηση αποτελείται από διαδοχικά ανεξάρτητα βήματα, και το πλήθος των επιλογών ενός βήματος είναι πλήρως καθορισμένο όταν είναι γνωστά τα προηγούμενα βήματα, τότε μπορούμε να πολλαπλασιάσουμε τους τρόπους των επιμέρους βημάτων.
- $\alpha_1 = \nu_1, \alpha_2 = \nu_2 \ldots \alpha_\kappa = \nu_\kappa \rightarrow \nu_1 * \nu_2 * \cdots * \nu_\kappa$

Διατάξεις και Μεταθέσεις

- Παράδειγμα: 4 δένδρα



Πόσοι τρόποι να φυτέψουμε μια δενδροστοιχία με 2 δένδρα;



4 επιλογές δένδρων

3 επιλογές δένδρων

4 * 3 = 12 τρόποι

Διατάξεις και Μεταθέσεις (Ορισμός)

- Έστω X σύνολο με ν στοιχεία και ακέραιος $\kappa \leq \nu$, κάθε διατεταγμένη κ-αδα $(\alpha_1, \alpha_2 \ldots \alpha_\kappa)$ ονομάζεται διάταξη των ν ανά κ (κ −permutation).
- Αν $\kappa = \nu$, τότε συμπεριλαμβάνει όλα τα στοιχεία και καλείται διάταξη.
- Το πλήθος των διατάξεων είναι: $\nu(\nu-1)(\nu-2)\ldots(\nu-\kappa+1)$ ή $\frac{\nu!}{(\nu-\kappa)!}$ όπου $p! = p(p-1)(p-2)\ldots 1$
- Συμβολισμός του πλήθους διατάξεων $(\nu)_\kappa$ ή P_κ^ν

Επαναληπτικές διατάξεις (γενικά)

- Αν τα στοιχεία όπου κάνουμε τις διατάξεις μπορούν να επαχρησιμοποιηθούν, τότε μιλάμε για διάταξη των ν ανά κ με επανάληψη (κ −permutation with repetition)
- Το πλήθος των διαφορετικών κ-αδων είναι ν^{κ}

Μεταθέσεις κ ειδών στοιχείων

- Παράδειγμα: 3 διαφορετικά γράμματα 

της λέξης ΑΛΛΟ

Πόσες διαφορετικές λέξεις με 4 γράμματα [XXXX];

Προφανής αλλά λάθος απάντηση (όπως δενδροστοιχία): 4!

12 διαφορετικές λέξεις:
ΆΛΛΟ, ΑΛΟΛ, ΛΑΛΟ, ΛΑΟΛ, ΟΑΛΛ, ΛΛΑΟ, ΛΛΟΑ, ΑΟΛΛ, ΛΟΑΛ, ΛΟΛΑ, ΟΛΛΑ, ΟΛΑΛ

Παρατήρηση: από τα 4! «αφαιρούμε» τις επαναλήψεις διαιρώντας με το πλήθος των διατάξεων που θα είχε η κάθε «τσάντα»: $\frac{4!}{1!*2!*1!}$

Μεταθέσεις κ ειδών στοιχείων (γενικά)

- Έστω ένα σύνολο $\mathrm{X} = \{x_1, x_2, \dots, x_k\}$, k διαφορετικών στοιχειών με ($k \geq 2$). Κάθε διατεταγμένη ν-αδα που δημιουργείται αν χρησιμοποιηθεί το x_1 συνολικά ν_1,… το x_k συνολικά ν_k φορές, ονομάζεται μετάθεση των k ειδών στοιχειών από ν (όπου $\nu = \nu_1 + \nu_2 + \cdots + \nu_\kappa$)
- Το πλήθος τον διαφορετικών μεταθέσεων είναι

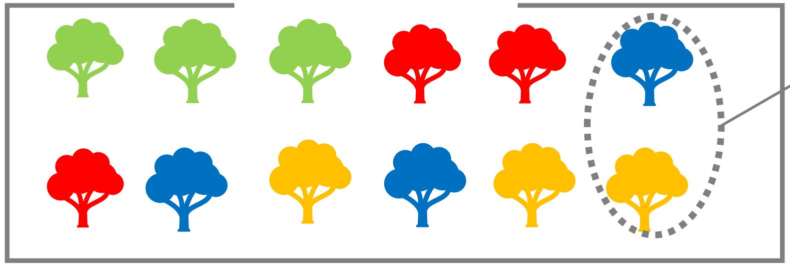
$$\binom{\nu}{\nu_1, \nu_2, \dots, \nu_\kappa} = \frac{\nu!}{\nu_1!\, \nu_2! \dots \nu_\kappa!}$$

Συνδυασμοί

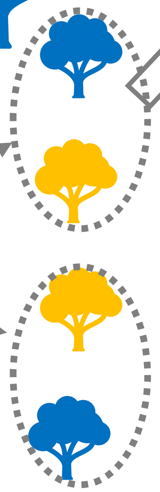
- Παράδειγμα: 4 δένδρα

Πόσοι τρόποι να επιλέξουμε 2 δένδρα;

συνδυασμοί



Για κάθε 2-αδα έχουμε 2 πιθανές διατάξεις



Οπότε αν ξεκινήσουμε από τις διατάξεις $(\nu)_\kappa = \frac{\nu!}{(\nu-\kappa)!}$ ή

$$(4)_2 = \frac{4!}{(4-2)!} = \frac{4*3*2*1}{2*1} = 12$$

και «βγάλουμε» τις διπλοεγγραφές μας μένουν οι συνδυασμοί

$$\frac{\frac{\nu!}{(\nu-\kappa)!}}{\kappa!} = \frac{12}{2*1} = 6$$

Συνδυασμοί (γενικά)

- Έστω X σύνολο με ν στοιχεία και ακέραιος $\kappa \leq \nu$. Συνδυασμός των ν στοιχείων ανά κ ονομάζεται κάθε μη διατεταγμένο υποσύνολο του X με κ στοιχεία.
- Το πλήθος των διαφορετικών συνδυασμών είναι

$$\binom{\nu}{\kappa} = \frac{\frac{\nu!}{(\nu-\kappa)!}}{\kappa!} = \frac{\nu!}{\kappa!\,(\nu-\kappa)!}$$

Συνδυασμοί με επανάληψη (γενικά)

- Έστω X σύνολο με ν στοιχεία και ακέραιος $\kappa \leq \nu$. Επαναληπτικός συνδυασμός των ν στοιχείων ανά κ ονομάζεται κάθε μη διατεταγμένο υποσύνολο του X με κ στοιχεία όπου επιτρέπεται κάποιο στοιχείο να ξαναχρησιμοποιηθεί.
- Το πλήθος των διαφορετικών συνδυασμών είναι

$$\begin{bmatrix}\nu\\ \kappa\end{bmatrix} = \binom{\nu+\kappa-1}{\kappa} = \frac{(\nu+\kappa-1)!}{\kappa!\,(\nu-1)!}$$

Συνδυασμοί με επανάληψη (παράδειγμα)

- Παράδειγμα: 3 είδη από σοκολάτες

Πόσες διαφορετικές επιλογές για συσκευασίες των 15;

Πρόκειται για επιλογή 15 στοιχείων από 3.

Δεν τίθεται θέμα διάταξης στην κάθε συσκευασία.

⇨ Επαναληπτικές διατάξεις 15 στοιχείων από 3

⇨ $$\begin{bmatrix}\nu\\ \kappa\end{bmatrix} = \frac{(\nu+\kappa-1)!}{\kappa!\,(\nu-1)!} = \frac{(3+15-1)!}{15!\,(3-1)!} = 136$$

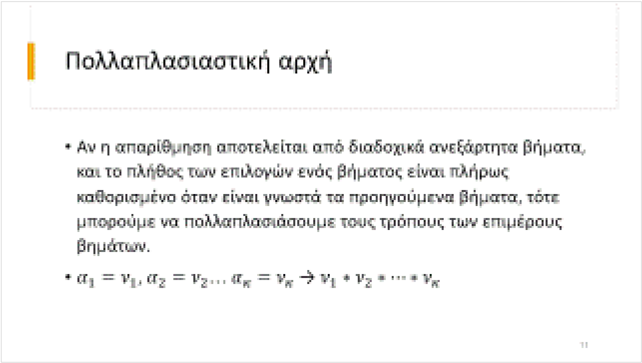
Παραδείγματα

Παράδειγμα 1

- Ζητούμενο: πόσα διαφορετικά αλφαριθμητικών μήκους 4 με περιορισμούς α)1^{η} και 4^{η} θέση σύμφωνο, β) 2^{η} θέση ψηφίο, γ)3^{η} φωνήεν.

✓ Ανεξάρτητα βήματα

✓ Πλήρως καθορισμένα από τα προηγούμενα


Πολλαπλασιαστική αρχή

- Αν η απαρίθμηση αποτελείται από διαδοχικά ανεξάρτητα βήματα, και το πλήθος των επιλογών ενός βήματος είναι πλήρως καθορισμένο όταν είναι γνωστά τα προηγούμενα βήματα, τότε μπορούμε να πολλαπλασιάσουμε τους τρόπους των επιμέρους βημάτων.
- $\alpha_1 = \nu_1, \alpha_2 = \nu_2 \dots \alpha_\kappa = \nu_\kappa \rightarrow \nu_1 * \nu_2 * \dots * \nu_\kappa$

[X X X X]

17 10 7 17

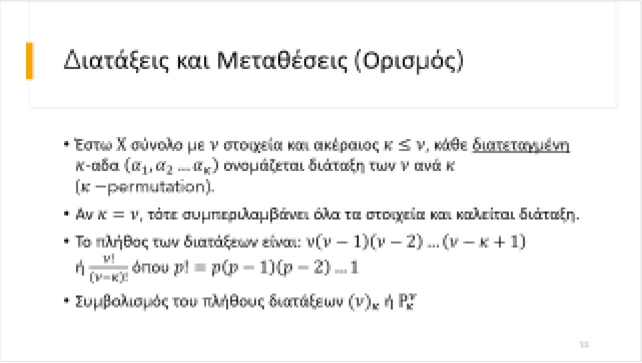
Πολλαπλασιαστική αρχή ⇒ 17*10*7*17 = 20230

Παράδειγμα 2

- Ζητούμενο: πόσες διαφορετικές πολυπεπτιδικές αλυσίδες μήκους 5 (διαφορετικών μεταξύ τους αμινοξέων) μπορούν αν προκύψουν από 12 διαφορετικά αμινοξέα.

Η σειρά αμινοξέων στα πολυπεπτιδια είναι σημαντική (αλυσίδα)

Το κάθε αμινοξύ συμμετέχει μόνο μια φορά το πολύ σε κάθε πολυπεπτίδιο



Διατάξεις και Μεταθέσεις (Ορισμός)

- Έστω X σύνολο με ν στοιχεία και ακέραιος $\kappa \leq \nu$, κάθε διατεταγμένη κ-αδα $(\alpha_1, \alpha_2 \ldots \alpha_\kappa)$ ονομάζεται διάταξη των ν ανά κ (κ −permutation).
- Αν $\kappa = \nu$, τότε συμπεριλαμβάνει όλα τα στοιχεία και καλείται διάταξη.
- Το πλήθος των διατάξεων είναι: $\nu(\nu-1)(\nu-2)\ldots(\nu-\kappa+1)$ ή $\frac{\nu!}{(\nu-\kappa)!}$ όπου $p! = p(p-1)(p-2)\ldots 1$
- Συμβολισμός του πλήθους διατάξεων $(\nu)_\kappa$ ή P_κ^ν

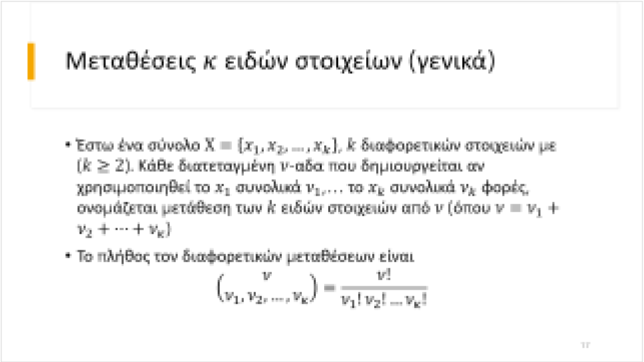
$$(\nu)_\kappa = \frac{\nu!}{(\nu-\kappa)!} = \frac{12!}{(12-5)!} = 95040$$

Παράδειγμα 3

- Ζητούμενο: πόσες διαφορετικές δενδροστοιχίες που αποτελούνται από 5 ελιές, 3 νεραντζιές, 5 πορτοκαλιές και 7 λεμονιές (όλα τα δένδρα αν είδος είναι όμοια).

Δενδροστοιχίες = σειρά/διάταξη είναι σημαντική

Μήκος δενδροστοιχίας = 5+3+5+7 = 20



Μεταθέσεις κ ειδών στοιχείων (γενικά)

- Έστω ένα σύνολο $X = \{x_1, x_2, \dots, x_k\}$, k διαφορετικών στοιχείων με $(k \geq 2)$. Κάθε διατεταγμένη ν-άδα που δημιουργείται αν χρησιμοποιηθεί το x_1 συνολικά $\nu_1, \dots$ το x_k συνολικά ν_k φορές, ονομάζεται μετάθεση των k ειδών στοιχείων από ν (όπου $\nu = \nu_1 + \nu_2 + \dots + \nu_\kappa$)
- Το πλήθος τον διαφορετικών μεταθέσεων είναι

$$\binom{\nu}{\nu_1, \nu_2, \dots, \nu_\kappa} = \frac{\nu!}{\nu_1!\,\nu_2!\dots\nu_\kappa!}$$

$$\binom{\nu}{\nu_1, \nu_2, \dots, \nu_\kappa} = \frac{\nu!}{\nu_1!\,\nu_2!\dots\nu_\kappa!}$$

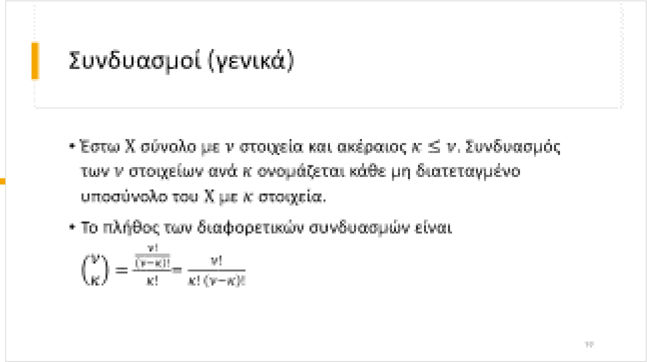
$$\binom{20}{5,3,5,7} = \frac{20!}{5! * 3! * 5! * 7!} = 5587021440$$

Παράδειγμα 4

- Ζητούμενο: πόσες διαφορετικές συνθέσεις τριμελούς επιτροπής μπορούν να προκύψουν από 8 άτομα.

Τριμελής επιτροπή χωρίς σύνθεση = αδιάφορη διάταξη

Αδιάφορη διάταξη σε επιλογή από σύνολο = συνδυασμοί



Συνδυασμοί (γενικά)

- Έστω X σύνολο με ν στοιχεία και ακέραιος $\kappa \leq \nu$. Συνδυασμός των ν στοιχείων ανά κ ονομάζεται κάθε μη διατεταγμένο υποσύνολο του X με κ στοιχεία.
- Το πλήθος των διαφορετικών συνδυασμών είναι

$$\binom{\nu}{\kappa} = \frac{\frac{\nu!}{(\nu-\kappa)!}}{\kappa!} = \frac{\nu!}{\kappa!\,(\nu-\kappa)!}$$

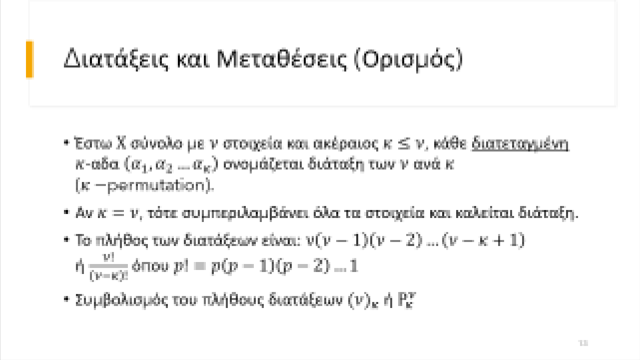
$$\binom{\nu}{\kappa} = \frac{\nu!}{\kappa!\,(\nu-\kappa)!} \text{ ή}$$

$$\binom{8}{3} = \frac{8!}{3!\,(8-3)!} = 56$$

Παράδειγμα 5

- Ζητούμενο: πόσες διαφορετικές συνθέσεις τριμελούς επιτροπής μπορούν να προκύψουν από 8 άτομα. Σύνθεση: πρόεδρος, αντιπρόεδρος και γραμματέας

Τριμελής επιτροπή με σύνθεση με επιλογή από μεγαλύτερο σύνολο = κ - διάταξη



Διατάξεις και Μεταθέσεις (Ορισμός)

- Έστω X σύνολο με ν στοιχεία και ακέραιος $\kappa \leq \nu$, κάθε <u>διατεταγμένη</u> κ-αδα $(\alpha_1, \alpha_2 \ldots \alpha_\kappa)$ ονομάζεται διάταξη των ν ανά κ (κ −permutation).
- Αν $\kappa = \nu$, τότε συμπεριλαμβάνει όλα τα στοιχεία και καλείται διάταξη.
- Το πλήθος των διατάξεων είναι: $\nu(\nu-1)(\nu-2)\ldots(\nu-\kappa+1)$ ή $\frac{\nu!}{(\nu-\kappa)!}$ όπου $p! = p(p-1)(p-2)\ldots 1$
- Συμβολισμός του πλήθους διατάξεων $(\nu)_\kappa$ ή P_κ^ν

$$(\nu)_\kappa = \frac{\nu!}{(\nu-\kappa)!} \text{ ή}$$

$$(8)_3 = \frac{8!}{(8-3)!} = 336$$