# Στατιστική Ι

**Γιώργος Τσιρογιάννης**

Τμήμα Διοίκησης Επιχειρήσεων Αγροτικών Προϊόντων και Τροφίμων,

Πανεπιστήμιο Πατρών

Ακαδημαϊκό έτος: 2022-2023

# Διάλεξη 5η

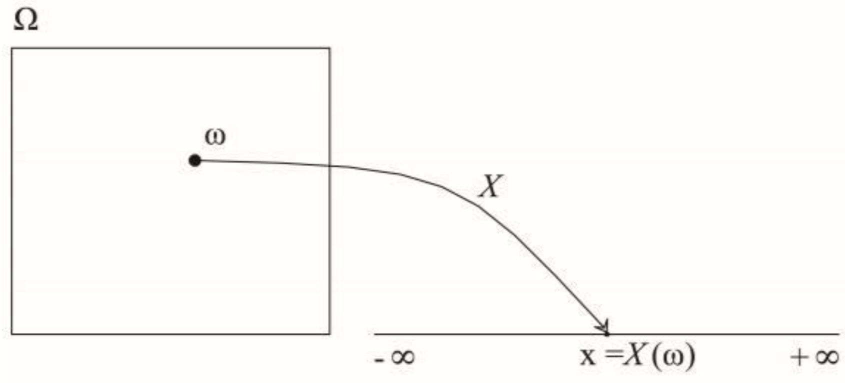
Τυχαίες μεταβλητές

Συνάρτηση κατανομής

Παράμετροι διακριτών τμ


5.1 – 5.3

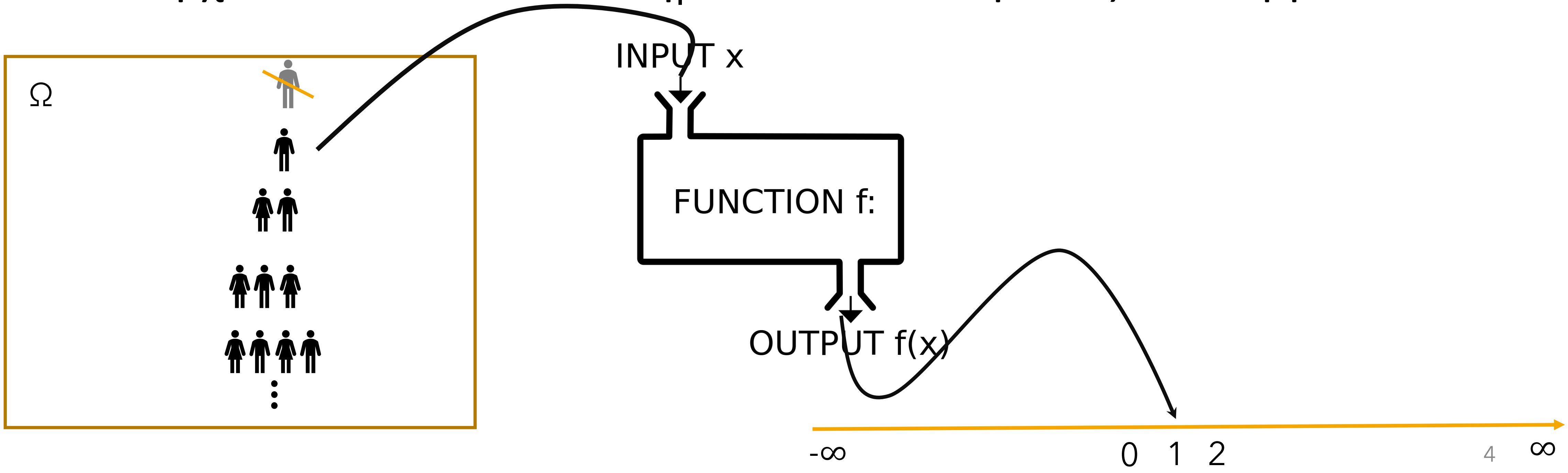
# Τυχαία μεταβλητή




- Μια **συνάρτηση** που αντιστοιχίζει το αποτέλεσμα που εμφανίζεται κατά την εκτέλεση ενός πειράματος τύχης σε έναν **πραγματικό αριθμό**

# Τυχαίες μεταβλητές (συν.)

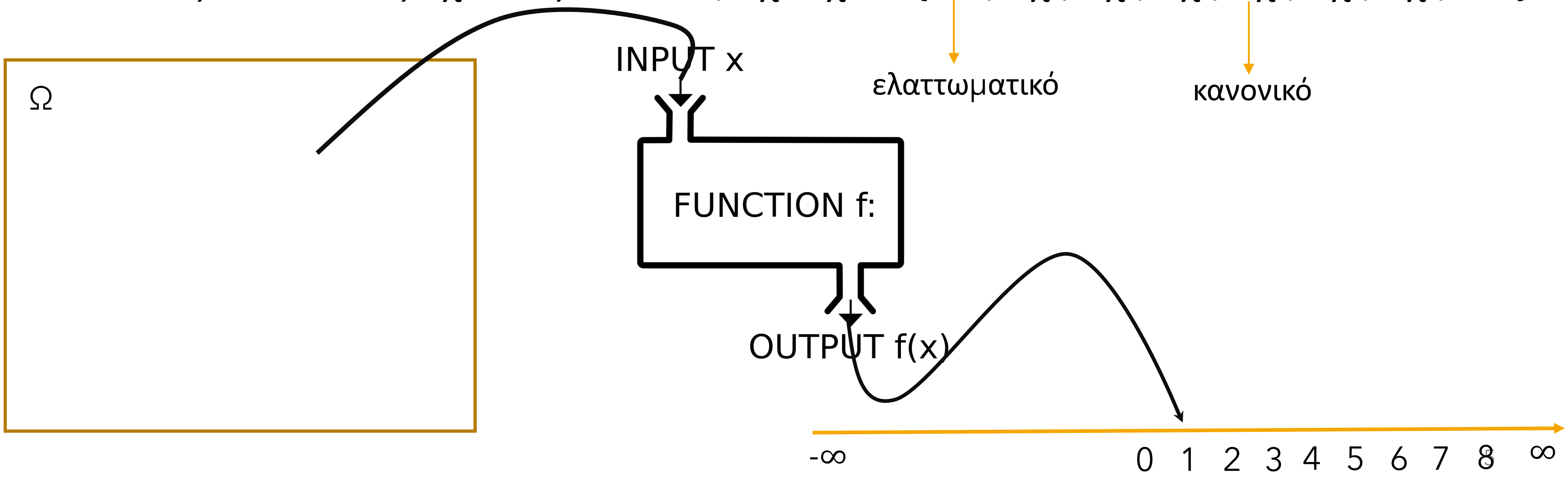
- Μια συνάρτηση που καταμετρά τον αριθμό των πελατών που εισέρχονται σε ένα κατάστημα 9-11 Δευτέρα ως και Σάββατο



# Τυχαίες μεταβλητές (συν.)

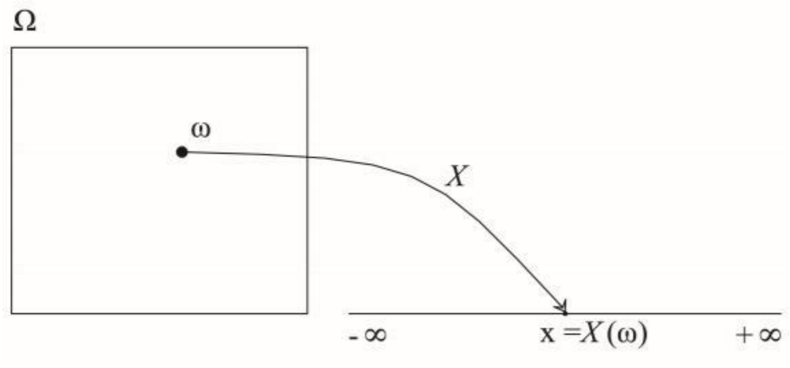
- Ελέγχεται κάθε προϊόν από μια παρτίδα οκτώ προϊόντων για να διαπιστωθεί αν είναι ή όχι ελαττωματικό

Ω=όλες οι δυνατές οχτάδες από ναι, όχι πχ ω={ναι, όχι, όχι, όχι, όχι, όχι, όχι, ναι}



# Τυχαίες μεταβλητές (συν.)

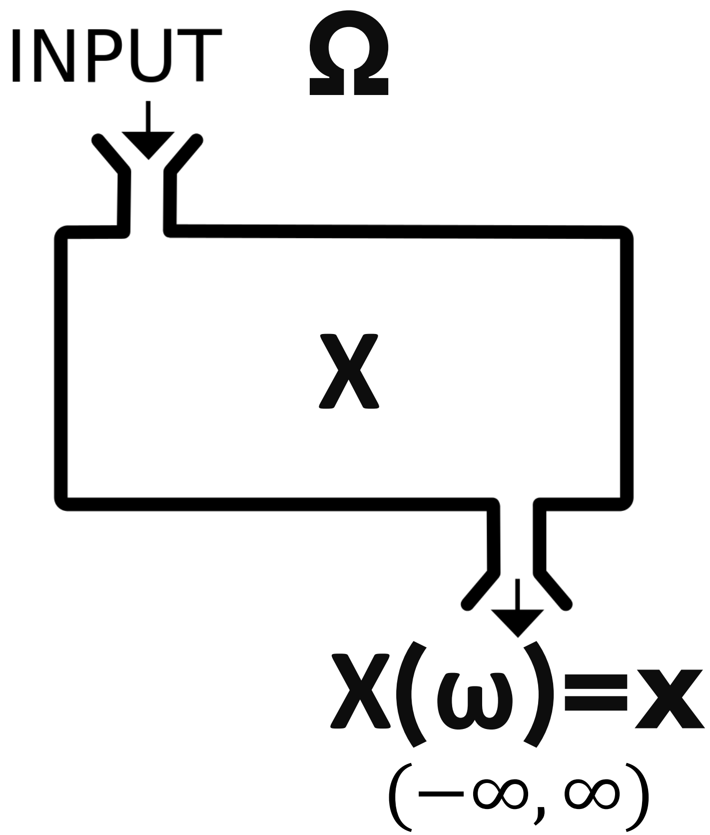
- Μια τυχαία μεταβλητή συνήθως συμβολίζεται με ένα κεφαλαίο γράμμα X, Y, T, Z…
- Με τα αντίστοιχα μικρά γράμματα x, y, t, z,… συμβολίζουμε τις τιμές των τυχαίων μεταβλητών
- Για να δηλώσουμε ότι όταν το αποτέλεσμα του πειράματος τύχης είναι το $\omega \in \Omega$ η τμ λαμβάνει την τιμή x, τότε γράφουμε X(ω)=x ή X=x.



# Τυχαίες μεταβλητές (συν.)

- Το πεδίο τιμών είναι το $X: \Omega \to (-\infty, \infty)$

$$X(\Omega) = R_X = \{x \in (-\infty, +\infty) : X(\omega) = x,\ \omega \in \Omega\} \subseteq (-\infty, +\infty)$$

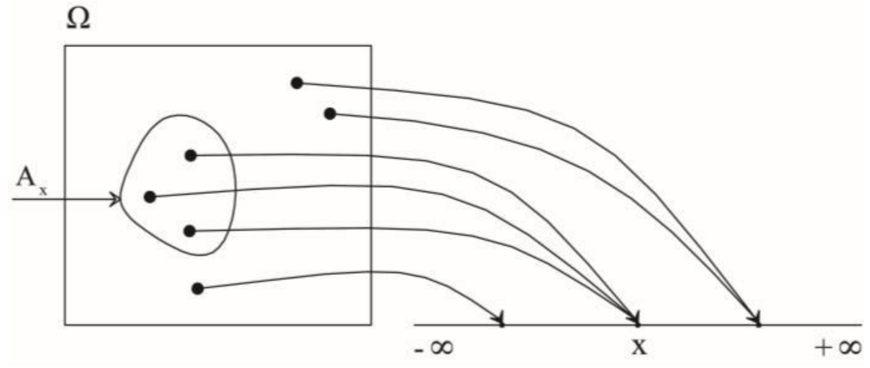


# Τυχαίες μεταβλητές (συν.)

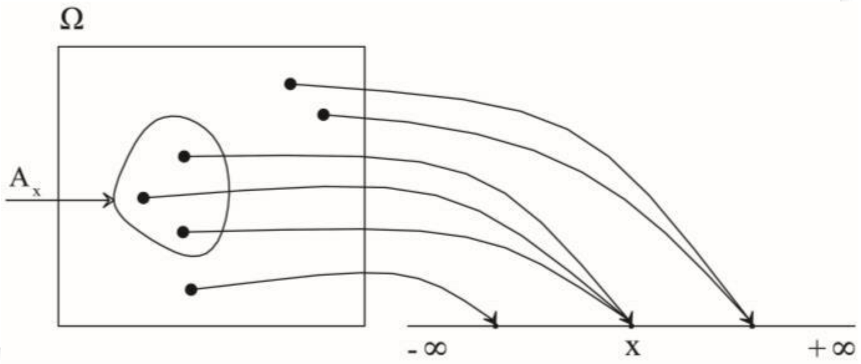
- Δεν είναι απαραίτητο μια τυχαία μεταβλητή να είναι ένα προς ένα.
- Διαφορετικά δειγματικά σημεία του δειγματικού χώρου Ω μπορεί να αντιστοιχίζονται, μέσω μιας τυχαίας μεταβλητής, σε ίσες τιμές της
- Για κάθε πραγματικό αριθμό x ορίζεται το σύνολο

$$A_x = \{\omega \in \Omega : X(\omega) = x\}$$

δηλαδή, ορίζεται το σύνολο των στοιχείων του Ω που μέσω της τ.μ. X αντιστοιχίζονται στον πραγματικό αριθμό x

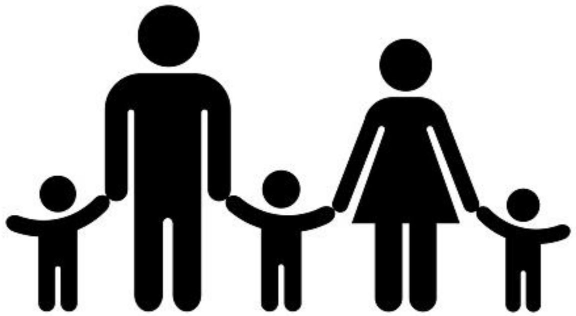


# Πιθανότητα τμ



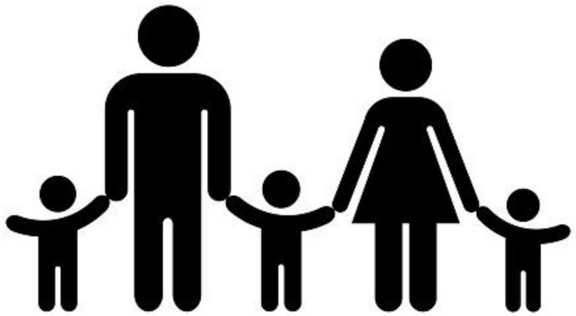

- Για κάθε ενδεχόμενο $A_x$ ( $A_x = \{\omega \in \Omega : X(\omega) = x\}$ ) η πιθανότητα να υλοποιηθεί (δηλαδή $P(A_x)$) ισούται με την πιθανότητα η τμ X να πάρει την τιμή x, και συμβολίζεται με P(X(ω)=x) ή απλά P(X=x)

# Παράδειγμα



- Ποια είναι η πιθανότητα μια τρίτεκνη οικογένεια, που επιλέγεται τυχαία από τα αρχεία του Δήμου Αθηναίων, να έχει ακριβώς ένα αγόρι.

# Παράδειγμα (συν)



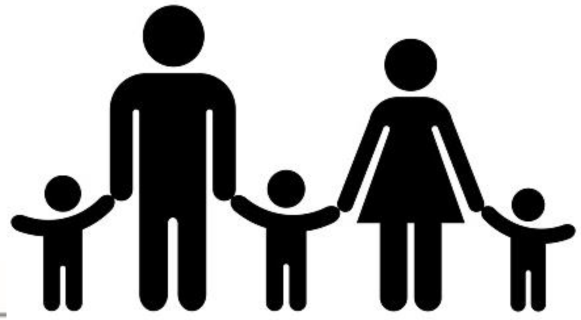
- ΔΧ Ω={ααα, κκκ, ακα, ακκ, κκα, αακ, καα, κακ}
- Κωδικοποίηση κάθε απλού ενδεχομένου με μια τριάδα:
  $1^{ο}$ παιδί $2^{ο}$ παιδί $3^{ο}$ παιδί
- «ακριβώς ένα αγόρι» : {ακκ, κκα, κακ}
- Ορίζουμε την τμ Χ που δέχεται σαν είσοδο (πεδίο ορισμού) όλο τον Ω και υπολογίζει ένα πραγματικό αριθμό

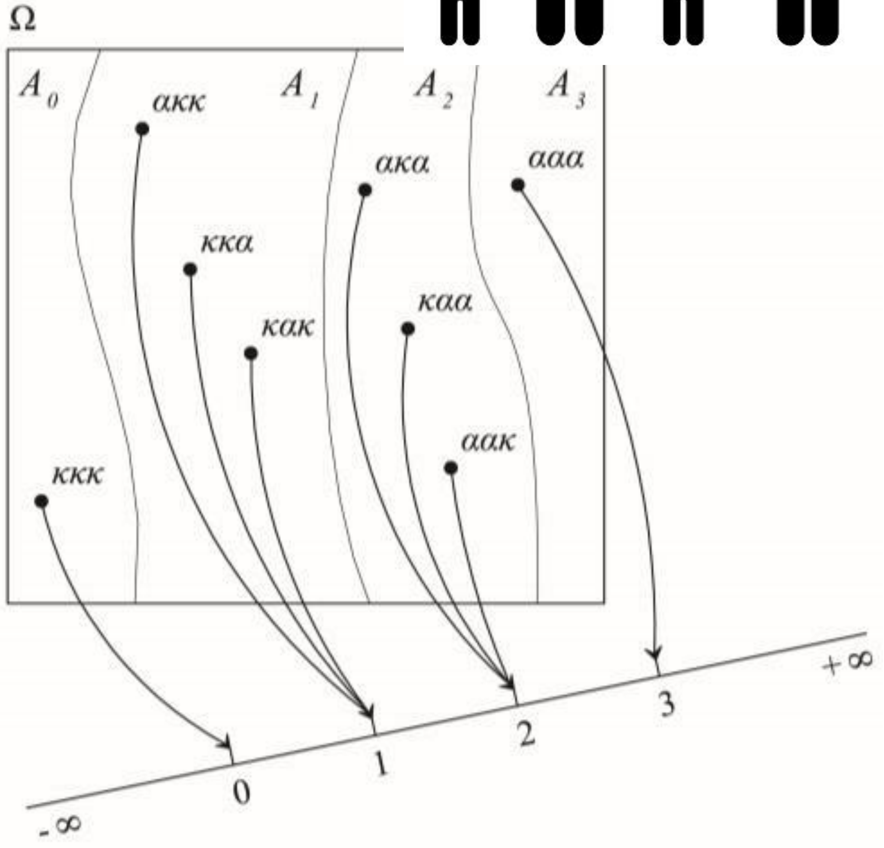
$$X: \{ααα, κκκ, ακα, ακκ, κκα, αακ, καα, κακ\} \rightarrow (-\infty, \infty)$$

# Παράδειγμα (συν)

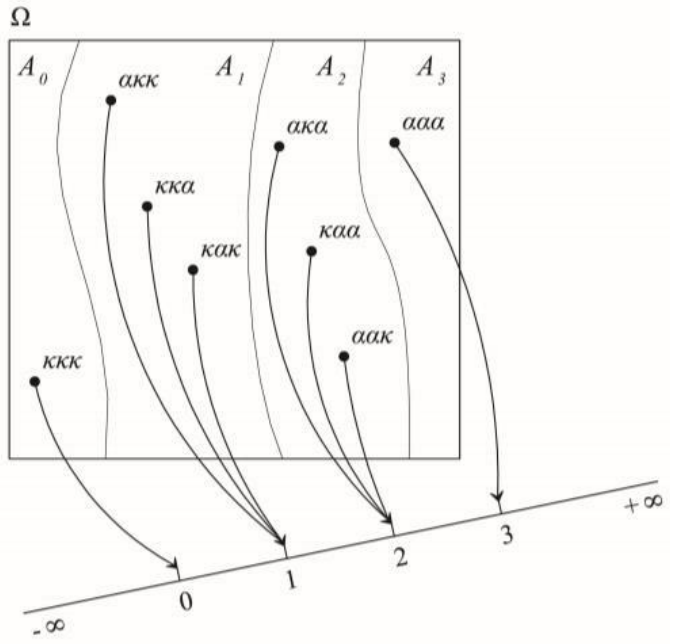
- Πως η X αντιστοιχίζει σε κάθε ενδεχόμενο έναν αριθμό;
- Μετρώντας τον αριθμό των αγοριών ανά τριαδα:

X(ααα)=3, X(κκκ)=0, X(ακα)=2, X(ακκ)=1, X(κκα)=1, X(αακ)=2, X(καα)=2, X(κακ)=1





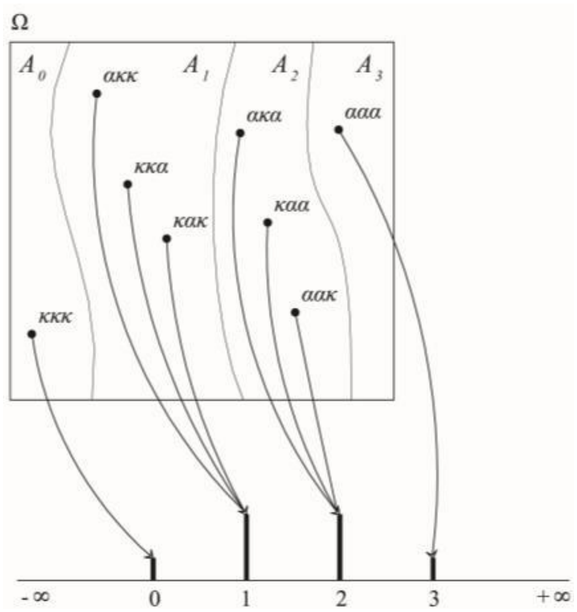
# Παράδειγμα (συν)

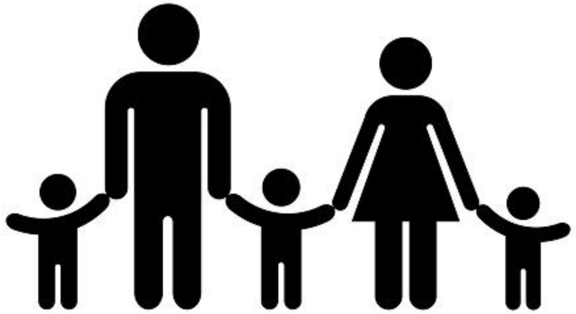


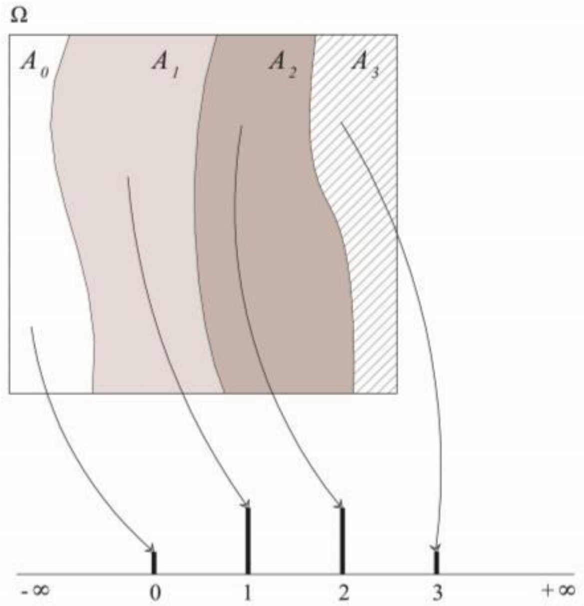

- Το πεδίο τιμών της τμ X είναι $\mathrm{R}_X = \{0,1,2,3\}$
- Το γεγονός «ακριβώς ένα αγόρι», αντιστοιχεί στο ενδεχόμενο {ακκ, κκα, κακ}, που εκφράζεται από την τμ X όταν πάρει την τιμή 1
- Δηλαδή: $\mathrm{A}_1 = \{\omega \in \Omega,\ \mathrm{X}(\omega) = 1\} =$ {ακκ, κκα, κακ}

# Παράδειγμα (συν)

- Μπορούμε να «ξεχάσουμε» το ενδεχόμενο {ακκ, κκα, κακ} και να υπολογίσουμε την πιθανότητα «ακριβώς ένα αγόρι» μέσω της τμ X και της τιμής X=1

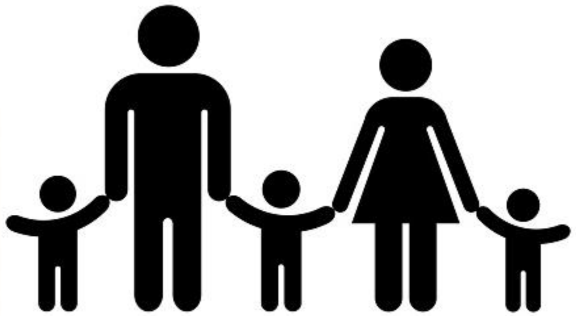


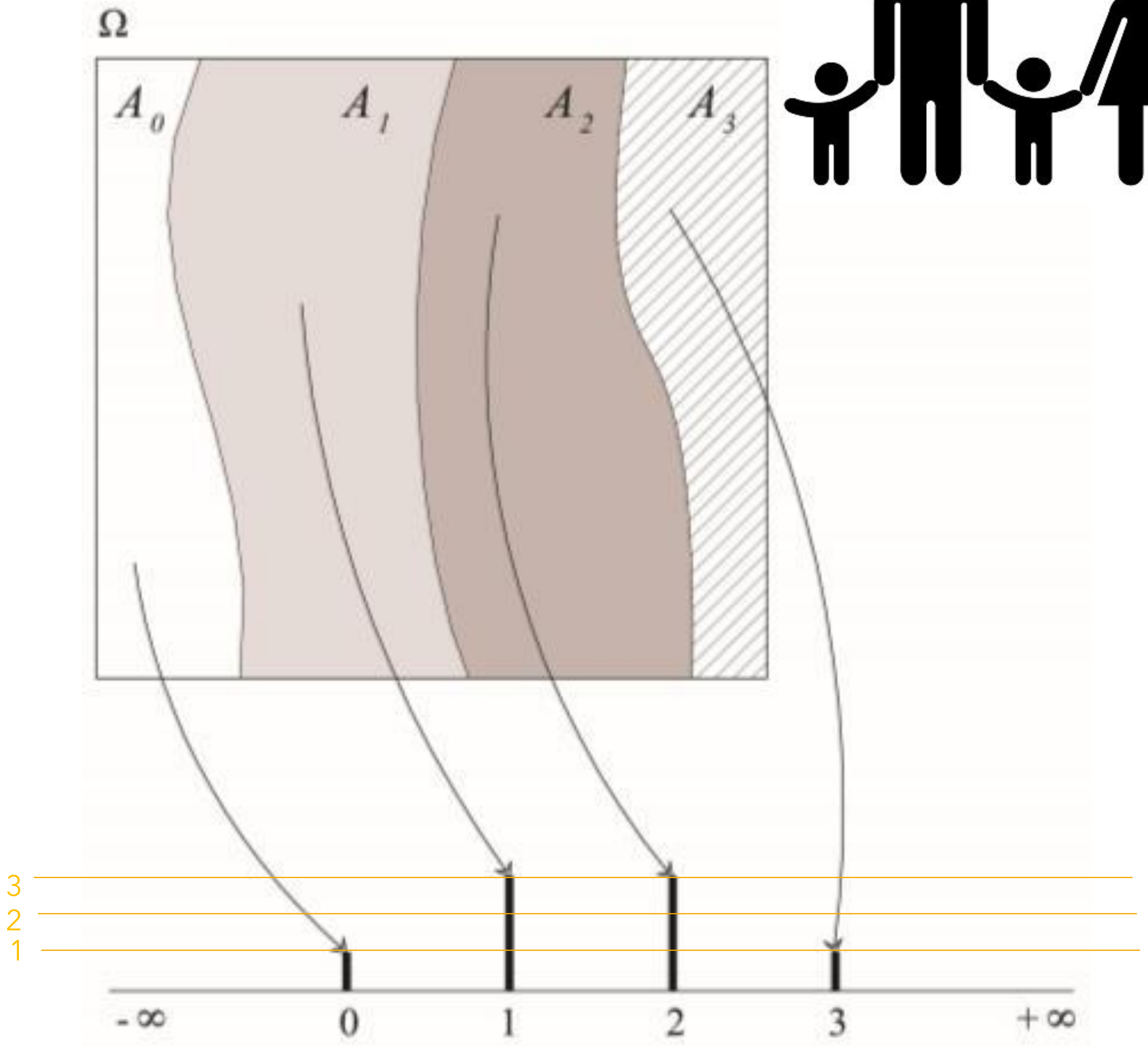



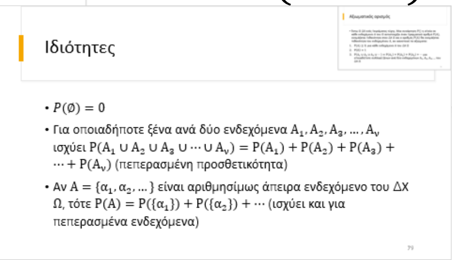


# Παράδειγμα (συν)

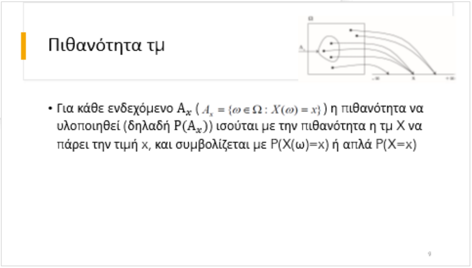
- Πιθανότητα «ακριβώς ένα αγόρι» $P(A_1) = P(X = 1) = \frac{3}{8}$
- Πιθανότητα «ακριβώς δύο αγόρια» $P(A_2) = P(X = 2) = \frac{3}{8}$
- Πιθανότητα «ακριβώς τρία αγόρια» $P(A_3) = P(X = 3) = \frac{1}{8}$
- Πιθανότητα «το πολύ ένα αγόρι» $P(X \leq 1) = P(X = 0) + P(X = 1) = \frac{4}{8} = \frac{1}{2}$







Ιδιότητες

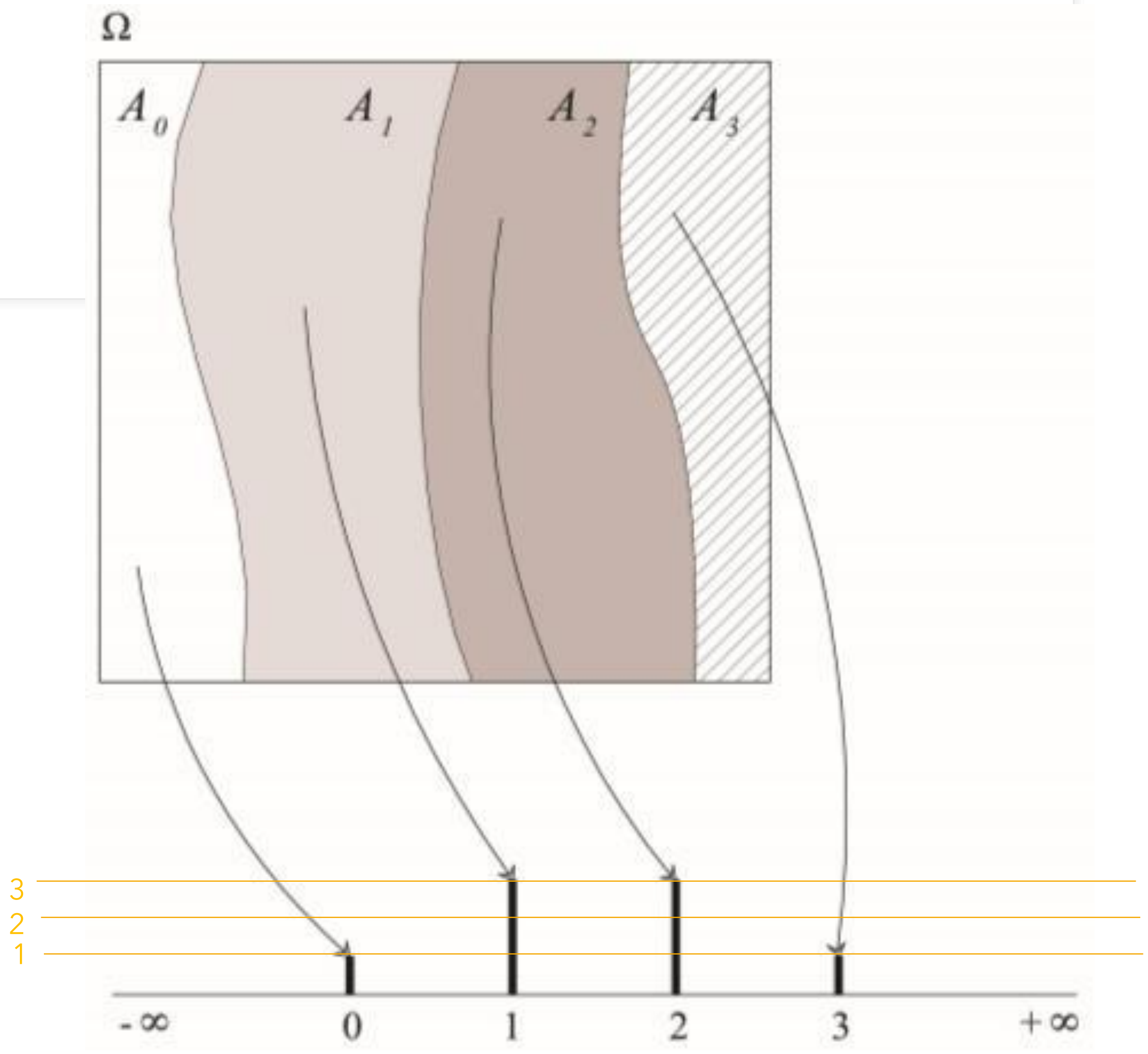
- $P(\emptyset) = 0$
- Για οποιαδήποτε ξένα ανά δύο ενδεχόμενα $A_1, A_2, A_3, \dots, A_\nu$ ισχύει $P(A_1 \cup A_2 \cup A_3 \cup \dots \cup A_\nu) = P(A_1) + P(A_2) + P(A_3) + \dots + P(A_\nu)$ (πεπερασμένη προσθετικότητα)
- Αν $A = \{\alpha_1, \alpha_2, \dots\}$ είναι αριθμησίμως άπειρα ενδεχόμενο του ΔΧ Ω, τότε $P(A) = P(\{\alpha_1\}) + P(\{\alpha_2\}) + \cdots$ (ισχύει και για πεπερασμένα ενδεχόμενα)


Πιθανότητα τμ

- Για κάθε ενδεχόμενο $A_x$ ($A_x = \{\omega \in \Omega : X(\omega) = x\}$) η πιθανότητα να υλοποιηθεί (δηλαδή $P(A_x)$) ισούται με την πιθανότητα η τμ X να πάρει την τιμή x, και συμβολίζεται με P(X(ω)=x) ή απλά P(X=x)

# Συνάρτηση κατανομής τυχαίας μεταβλητής

- Όταν γνωρίζουμε πως κατανέμονται οι πιθανότητες για τις τιμές μιας τμ γνωρίζουμε τα πάντα για αυτή και τα ενδεχόμενα που εκφράζει.




P(X=x) 1/8 3/8 3/8 1/8

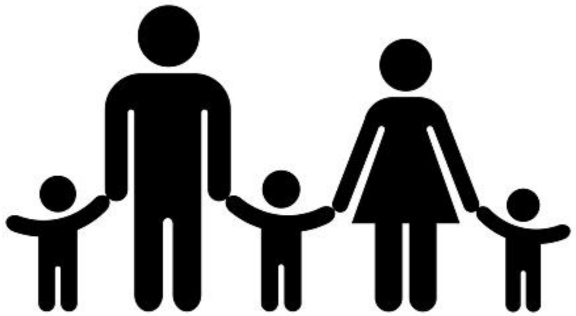
# Ορισμός συνάρτησης κατανομής τυχαίας μεταβλητής

Έστω X μια τυχαία μεταβλητή η οποία ορίζεται σε ένα δειγματικό χώρο Ω. Η πραγματική συνάρτηση F με τύπο

$$F(x) = P(X \leq x) = P(\{\omega \in \Omega : X(\omega) \leq x\}), \quad -\infty < x < +\infty$$

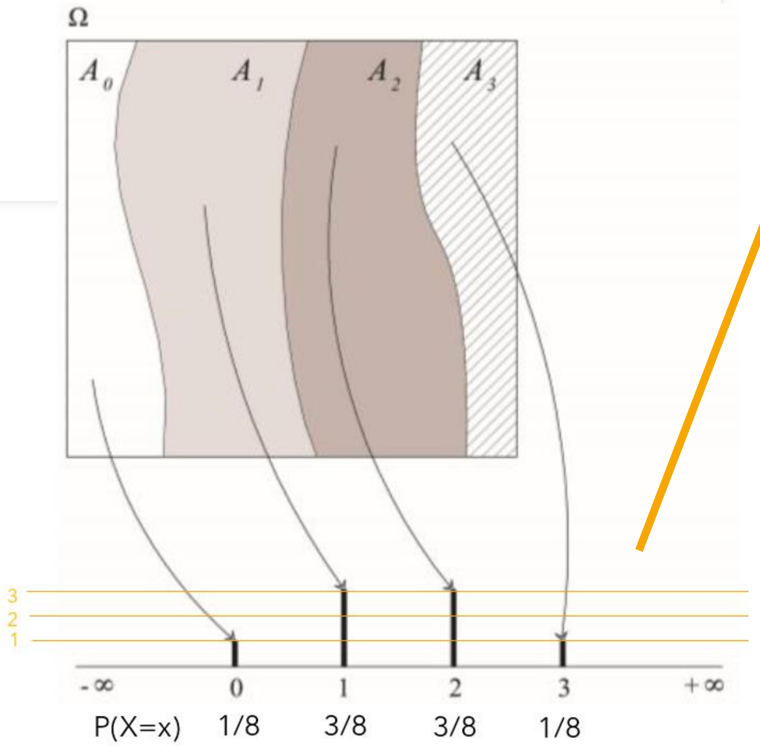
ονομάζεται συνάρτηση κατανομής (σ.κ.) ή αθροιστική συνάρτηση κατανομής της τυχαίας μεταβλητής X (cumulative distribution function/cdf).

# Παράδειγμα

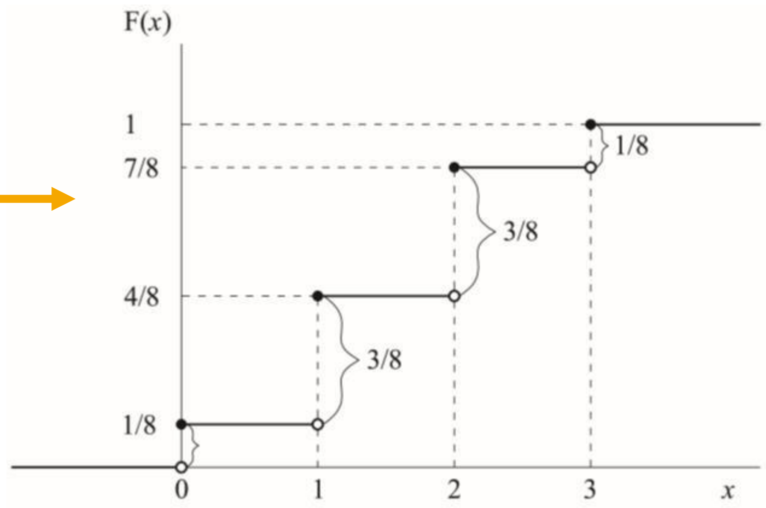


- Ποια είναι η πιθανότητα μια τρίτεκνη οικογένεια, που επιλεγεται τυχαία από τα αρχεία του Δήμου Αθηναίων

$$F(x)=P(X\le x)=P(\{\omega\in\Omega : X(\omega)\le x\}),\ -\infty<x<+\infty$$




$$F(x)=P(X\le x)=\begin{cases}0, & -\infty<x<0\\ 1/8, & 0\le x<1\\ 4/8, & 1\le x<2\\ 7/8, & 2\le x<3\\ 1, & 3\le x<+\infty\end{cases}$$

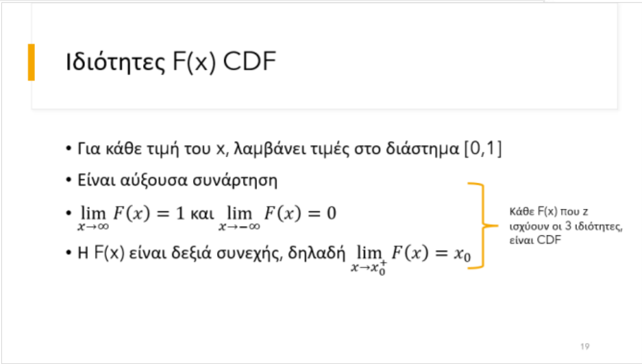


# Ιδιότητες F(x) CDF

- Για κάθε τιμή του x, λαμβάνει τιμές στο διάστημα [0,1]
- Είναι αύξουσα συνάρτηση
- $\lim_{x\to\infty} F(x) = 1$ και $\lim_{x\to-\infty} F(x) = 0$
- Η F(x) είναι δεξιά συνεχής, δηλαδή $\lim_{x\to x_0^+} F(x) = x_0$

Κάθε F(x) που z ισχύουν οι 3 ιδιότητες, είναι CDF

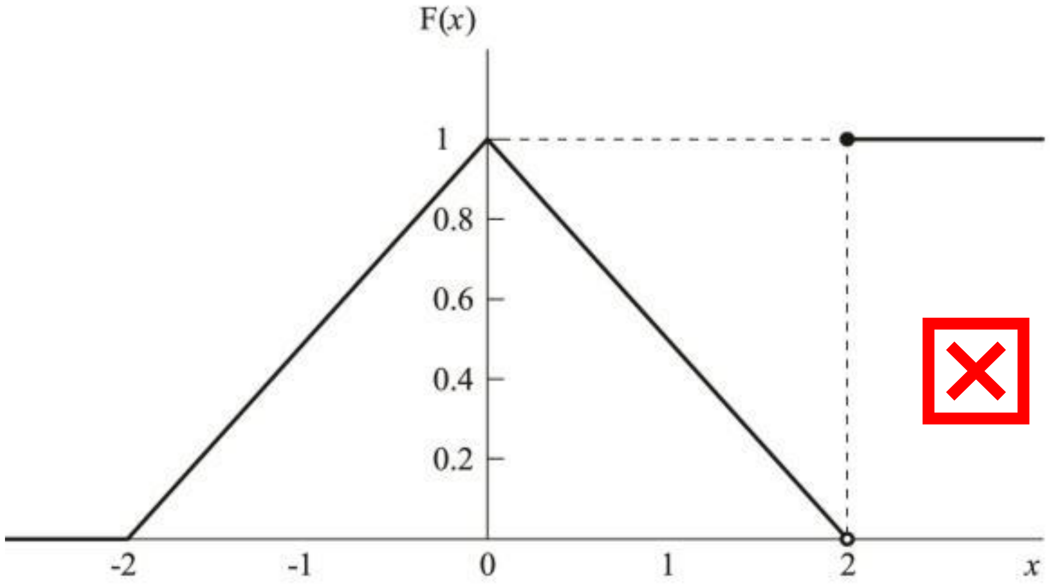
# Κατάλληλες για CDF?

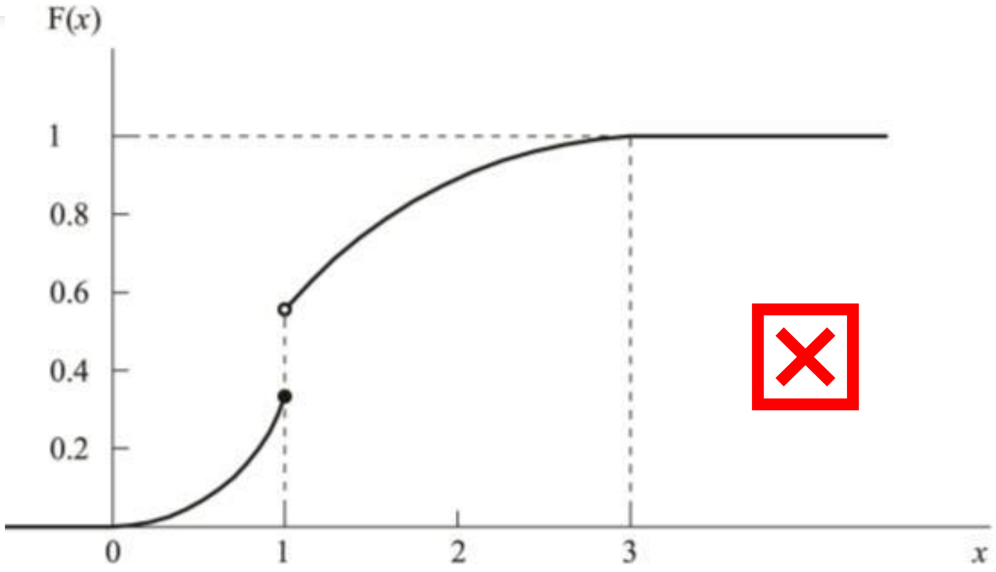

Ιδιότητες F(x) CDF

- Για κάθε τιμή του x, λαμβάνει τιμές στο διάστημα [0,1]
- Είναι αύξουσα συνάρτηση
- $\lim_{x\to\infty} F(x) = 1$ και $\lim_{x\to-\infty} F(x) = 0$
- Η F(x) είναι δεξιά συνεχής, δηλαδή $\lim_{x\to x_0^+} F(x) = x_0$

Κάθε F(x) που z ισχύουν οι 3 ιδιότητες, είναι CDF

19

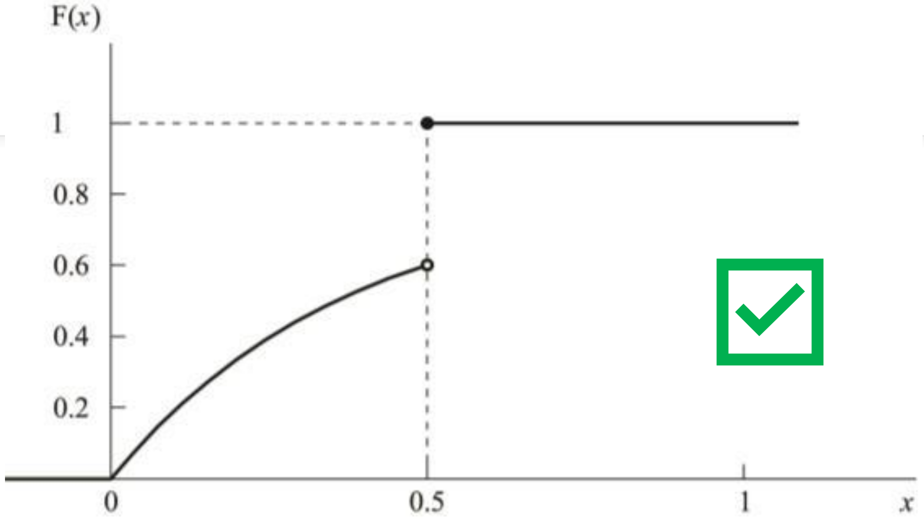


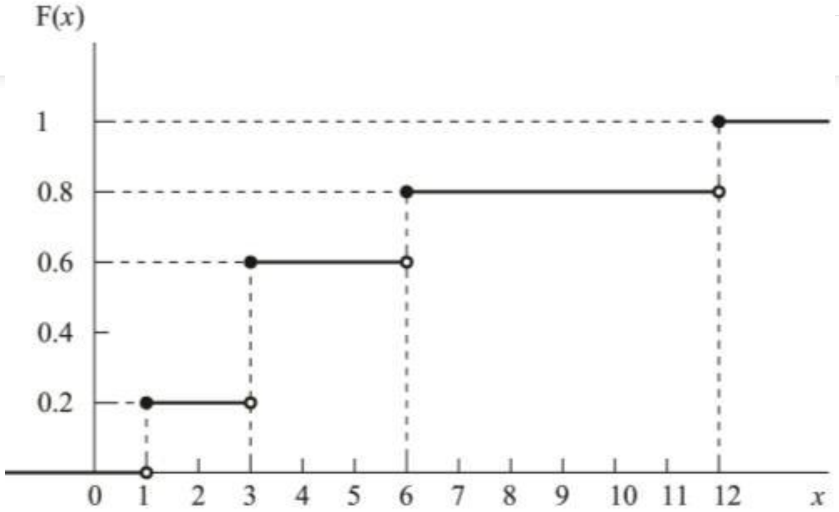



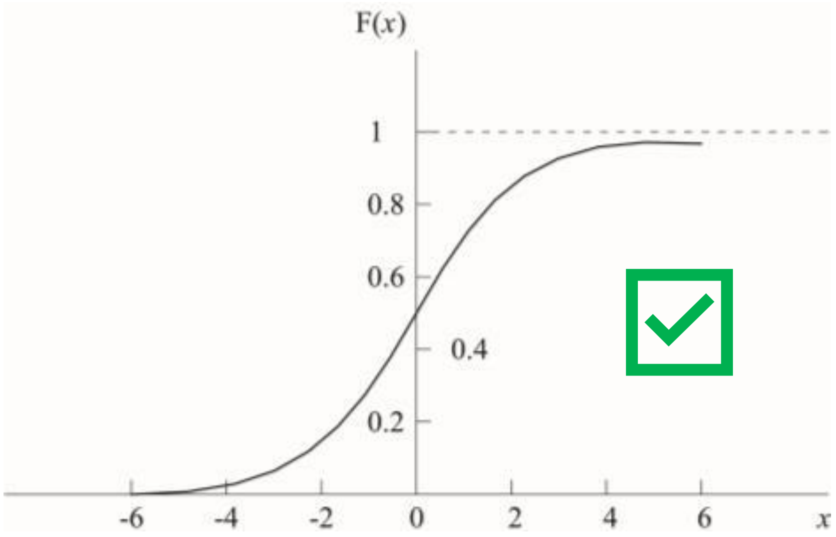

Μη αύξουσα

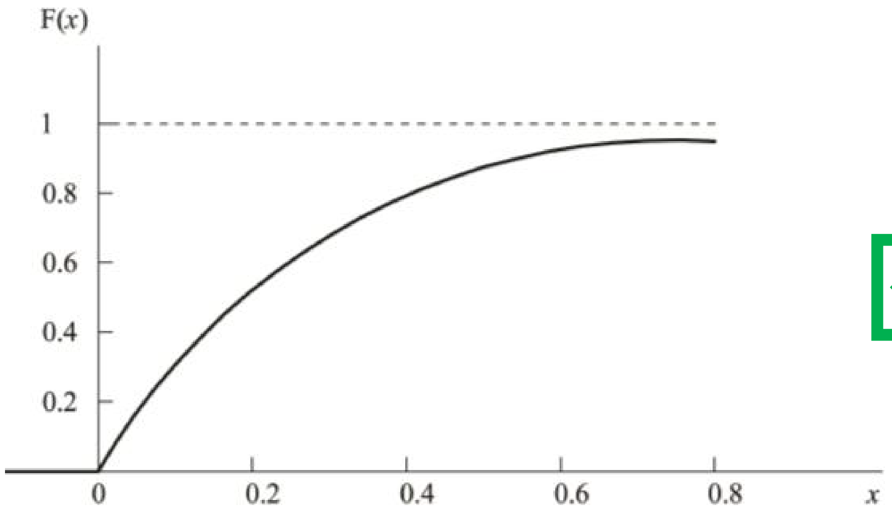
Για x = 1 δεν είναι δεξιά συνεχής

# Κατάλληλες για CDF?




## Ιδιότητες F(x) CDF

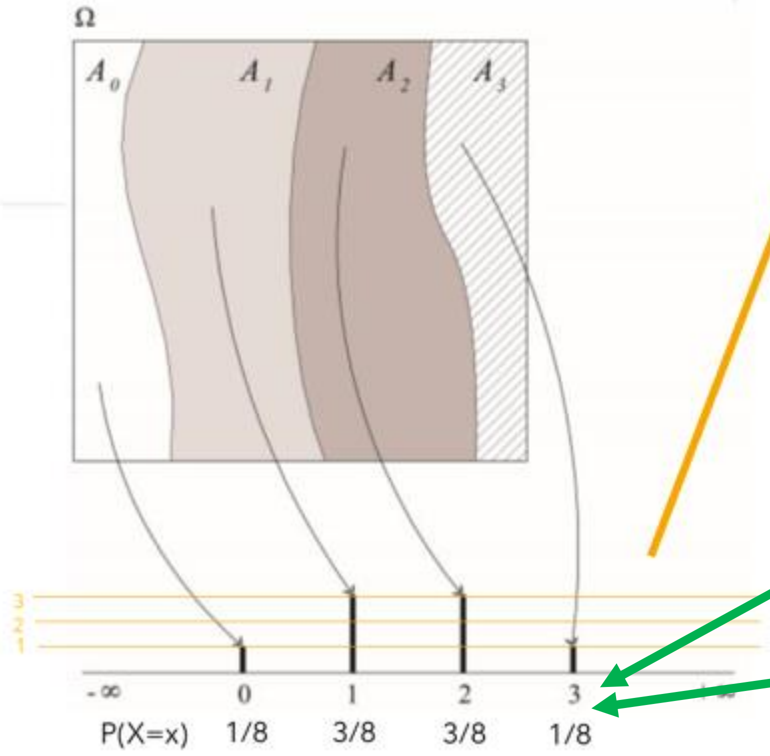
- Για κάθε τιμή του x, λαμβάνει τιμές στο διάστημα [0,1]
- Είναι αύξουσα συνάρτηση
- $\lim_{x\to\infty} F(x) = 1$ και $\lim_{x\to-\infty} F(x) = 0$
- Η F(x) είναι δεξιά συνεχής, δηλαδή $\lim_{x\to x_0^+} F(x) = x_0$

$\left.\begin{matrix} \\ \\ \\ \end{matrix}\right\}$ Κάθε F(x) που z ισχύουν οι 3 ιδιότητες, είναι CDF

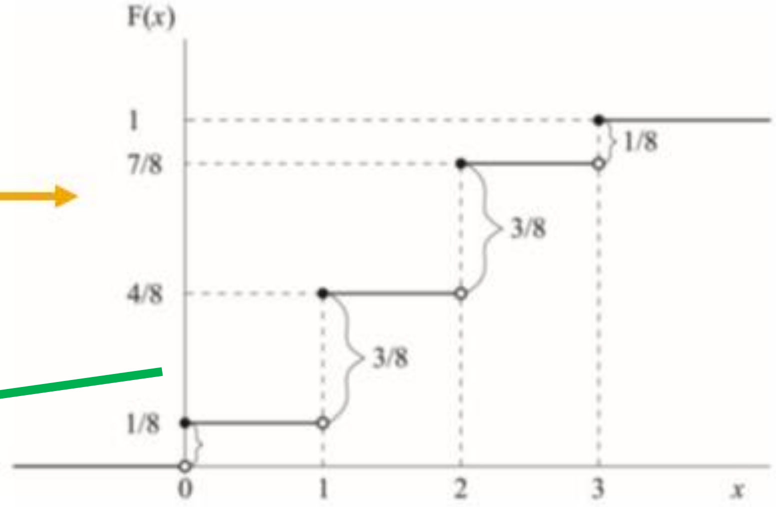
19

# Υπολογισμός πιθανότητας από CDF

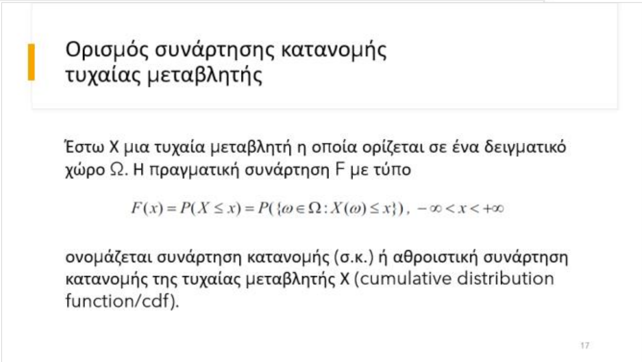
$$F(x) = P(X \leq x) = P(\{\omega \in \Omega : X(\omega) \leq x\}), \quad -\infty < x < +\infty$$




$$F(x) = P(X \leq x) = \begin{cases} 0, & -\infty < x < 0 \\ 1/8, & 0 \leq x < 1 \\ 4/8, & 1 \leq x < 2 \\ 7/8, & 2 \leq x < 3 \\ 1, & 3 \leq x < +\infty \end{cases}$$



# Υπολογισμός πιθανότητας από CDF



### Ορισμός συνάρτησης κατανομής τυχαίας μεταβλητής

Έστω X μια τυχαία μεταβλητή η οποία ορίζεται σε ένα δειγματικό χώρο Ω. Η πραγματική συνάρτηση F με τύπο

$$F(x) = P(X \leq x) = P(\{\omega \in \Omega : X(\omega) \leq x\}),\ -\infty < x < +\infty$$

ονομάζεται συνάρτηση κατανομής (σ.κ.) ή αθροιστική συνάρτηση κατανομής της τυχαίας μεταβλητής X (cumulative distribution function/cdf).

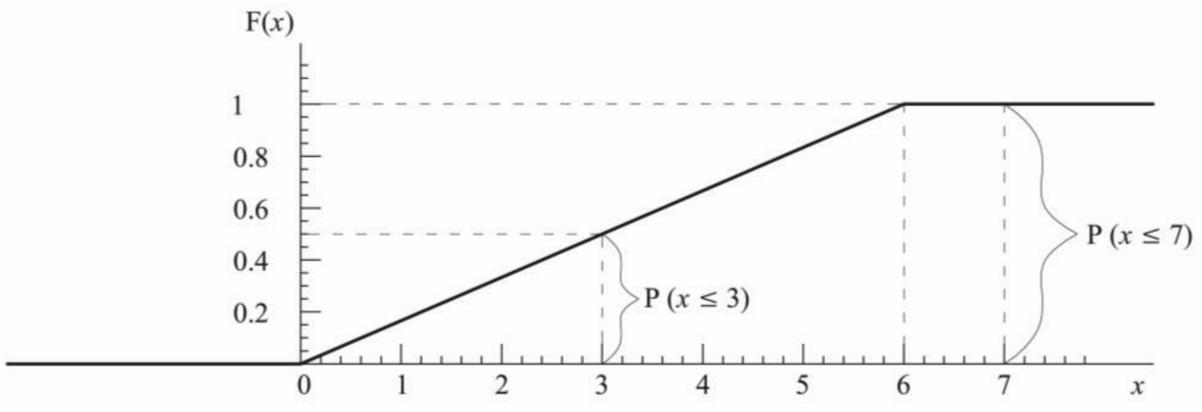
17

- Για πραγματικούς $a, b$ $(\mathrm{a} < \mathrm{b})$ και $\mathrm{F(x)}$ αθροιστική συνάρτηση κατανομής, ισχύουν
  - $\mathrm{P(X \leq b) = F(b)}$
  - $\mathrm{P(X > a) = 1 - P(x \leq a) = 1 - F(a)}$
  - $P(X < b) = F(b-)$ (αριστερό όριο της F στο x)
  - $P(X \geq a) = 1 - P(X < a) = 1 - F(a-)$
  - $P(X = a) = P(X \leq a) - P(X < a) = F(a) - F(a-)$
  - $P(a < X \leq b) = F(b) - F(a)$
  - $P(a \leq X \leq b) = F(b) - F(a-)$
  - $P(a < X < b) = F(b-) - F(a)$
  - $P(a \leq X < b) = F(b-) - F(a-)$

# Παράδειγμα

- Η ποσότητα ελιών (σε τόνους) που επεξεργάζεται σε μια βάρδια λειτουργίας ένα ελαιοτριβείο περιγράφεται από μια τυχαία μεταβλητή Χ με συνάρτηση κατανομής

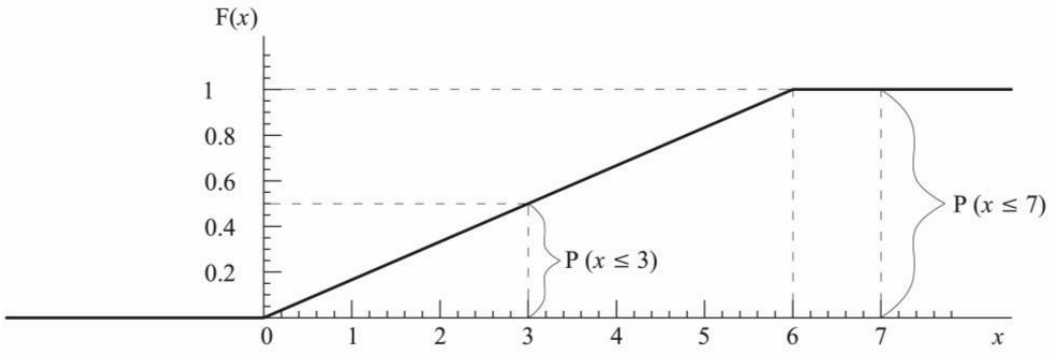
$$F(x)=\begin{cases}0, & -\infty < x < 0\\ x/6, & 0 \le x < 6\\ 1, & 6 \le x < +\infty\end{cases}$$

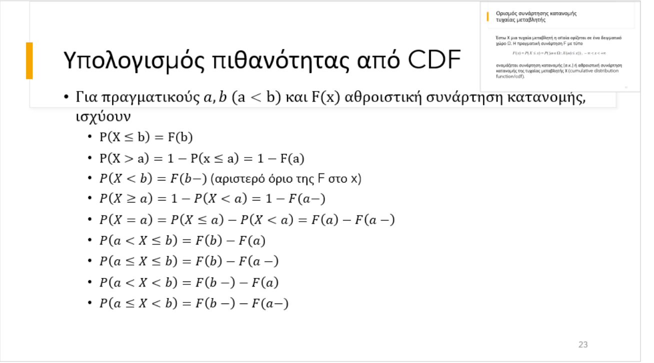


# Παράδειγμα (συν)

$$F(x)=\begin{cases}0, & -\infty < x < 0\\ x/6, & 0 \le x < 6\\ 1, & 6 \le x < +\infty\end{cases}$$

- Πιθανότητα το ελαιοτριβείο σε μια βάρδια να επεξεργαστεί το πολύ 3 τόνους: $P(X \le 3) = F(3) = 3/6 = 1/2$
- Πιθανότητα να επεξεργασθεί σε μια βάρδια το πολύ 7 τόνους:
  $$P(X \le 7) = F(7) = 1$$
- Πιθανότητα να επεξεργασθεί σε μια βάρδια λιγότερο από 3 τόνους $P(X < 3) = F(3-) = 3/6 = 1/2$



## Υπολογισμός πιθανότητας από CDF

- Για πραγματικούς $a, b$ $(a < b)$ και $F(x)$ αθροιστική συνάρτηση κατανομής, ισχύουν
  - $P(X \le b) = F(b)$
  - $P(X > a) = 1 - P(x \le a) = 1 - F(a)$
  - $P(X < b) = F(b-)$ (αριστερό όριο της F στο x)
  - $P(X \ge a) = 1 - P(X < a) = 1 - F(a-)$
  - $P(X = a) = P(X \le a) - P(X < a) = F(a) - F(a-)$
  - $P(a < X \le b) = F(b) - F(a)$
  - $P(a \le X \le b) = F(b) - F(a-)$
  - $P(a < X < b) = F(b-) - F(a)$
  - $P(a \le X < b) = F(b-) - F(a-)$

23

# Παράδειγμα (συν)

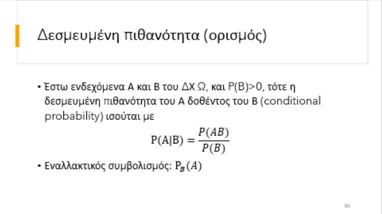
$$F(x)=\begin{cases}0, & -\infty < x < 0\\ x/6, & 0 \le x < 6\\ 1, & 6 \le x < +\infty\end{cases}$$

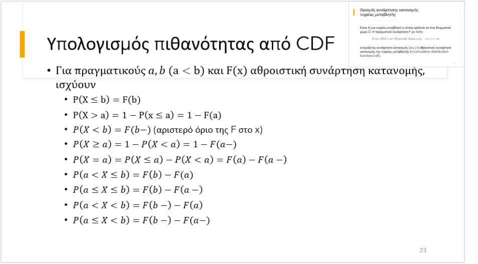
- Πιθανότητα να επεξεργασθεί σε μια βάρδια λιγότερο από 7 τόνους $P(X<7)=F(7-)=1$
- Πιθανότητα το ελαιοτριβείο σε μια βάρδια να επεξεργαστεί περισσότερους από 4.5 τόνους

$$P(X>4.5)=1-P(X\le 4.5)=1-F(4.5)=1-\frac{4.5}{6}=0.25$$

- πιθανότητα η ποσότητα που θα επεξεργαστεί το ελαιοτριβείο σε μια βάρδια να μην ξεπεράσει τους 5.5 τόνους δεδομένου ότι ήδη ξεπέρασε τους 2 τόνους

$$P(X\le 5.5 \mid X>2)=\frac{P(2<X\le 5.5)}{P(X>2)}=\frac{F(5.5)-F(2)}{1-F(2)}=\frac{3.5/6}{4/6}=0.875$$

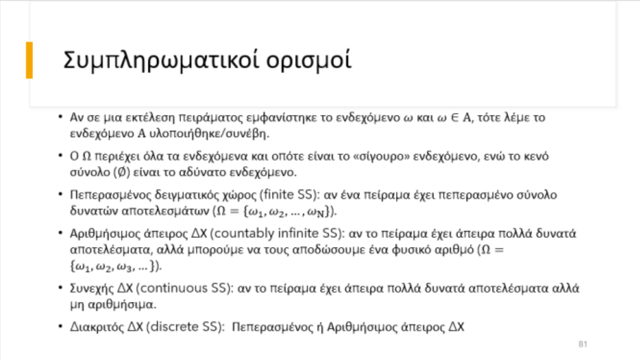


# Διακριτές τυχαίες μεταβλητές

# Διακριτές τυχαίες μεταβλητές

- Μια τυχαία μεταβλητή X ονομάζεται διακριτή ή απαριθμητή τυχαία μεταβλητή (discrete random variable) αν το σύνολο τιμών της, $R_X$, είναι πεπερασμένο ή αριθμησίμως άπειρο σύνολο.



## Συμπληρωματικοί ορισμοί

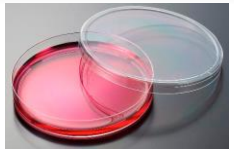
- Αν σε μια εκτέλεση πειράματος εμφανίστηκε το ενδεχόμενο $\omega$ και $\omega \in A$, τότε λέμε το ενδεχόμενο A υλοποιήθηκε/συνέβη.
- Ο Ω περιέχει όλα τα ενδεχόμενα και οπότε είναι το «σίγουρο» ενδεχόμενο, ενώ το κενό σύνολο ($\emptyset$) είναι το αδύνατο ενδεχόμενο.
- Πεπερασμένος δειγματικός χώρος (finite SS): αν ένα πείραμα έχει πεπερασμένο σύνολο δυνατών αποτελεσμάτων ($\Omega = \{\omega_1, \omega_2, \dots, \omega_N\}$).
- Αριθμήσιμος άπειρος ΔΧ (countably infinite SS): αν το πείραμα έχει άπειρα πολλά δυνατά αποτελέσματα, αλλά μπορούμε να τους αποδώσουμε ένα φυσικό αριθμό ($\Omega = \{\omega_1, \omega_2, \omega_3, \dots\}$).
- Συνεχής ΔΧ (continuous SS): αν το πείραμα έχει άπειρα πολλά δυνατά αποτελέσματα αλλά μη αριθμήσιμα.
- Διακριτός ΔΧ (discrete SS): Πεπερασμένος ή Αριθμήσιμος άπειρος ΔΧ

81

# Παραδείγματα διακριτών τυχαίων μεταβλητών

τυχαία μεταβλητή που εκφράζει

- τον αριθμό των βακτηριδίων σε 1cm2 μιας πλάκας Petri



- τον αριθμό των αυτοκινήτων που εισέρχονται στην αττική οδό από τον κόμβο Κύμης με κατεύθυνση την Ελευσίνα 6-9 το πρωί μια εργάσιμη ημέρα
- τον αριθμό των κλήσεων σε ένα τηλεφωνικό κέντρο 1-2 το μεσημέρι μια εργάσιμη ημέρα
- το ύψος της απόδοσης μιας επένδυσης
- τον αριθμό των σεισμών μεγέθους τουλάχιστον 4 Richter που συμβαίνουν σε ένα έτος στο Ιόνιο
- των θανάτων σε ένα μήνα από μια συγκεκριμένη ασθένεια σε μια πόλη
- το ύψος της αποζημίωσης ενός αγρότη για τις ζημιές από τον παγετό
- τον αριθμό των σωματιδίων που εκπέμπονται από ραδιενεργό υλικό σε ορισμένο χρονικό διάστημα
- τον αριθμό των μηχανικών βλαβών που συμβαίνουν σε ένα εργοτάξιο στη διάρκεια της απογευματινής βάρδιας
- τον αριθμό των φυτών μιας καλλιέργειας που έδωσαν περισσότερους από 5καρπούς.

# Συνάρτηση (μάζας) πιθανότητας

- Έστω X διακριτή τυχαία μεταβλητή με σύνολο τιμών $\mathrm{R}_X$. Η πραγματική συνάρτηση $f$ με

$$f(x) = \begin{cases} P(X = x), & \text{για κάθε } x \in R_X \\ 0, & \text{για κάθε } x \notin R_X \end{cases}$$

ονομάζεται συνάρτηση πιθανότητας ή συνάρτηση μάζας πιθανότητας της X (probability mass function).

# Ιδιότητες συνάρτησης (μάζας) πιθανότητας

- Ως συνάρτηση πιθανότητας (ΣΠ) μιας διακριτής τυχαίας μεταβλητής X που έχει σύνολο τιμών $R_X = \{x_1, x_2, \ldots, x_\nu, \ldots\}$ ορίζεται μια πραγματική συνάρτηση $f$ η οποία ικανοποιεί τις ακόλουθες τρεις ιδιότητες:

$$\alpha)\ \ f(x) = 0\,,\ \text{για κάθε}\ \ x \neq x_1, x_2, \ldots, x_\nu, \ldots$$

$$\beta)\ \ f(x_i) \geq 0\ \ \text{για κάθε}\ \ i = 1, 2, \ldots$$

$$\gamma)\ \ f(x_1) + f(x_2) + \ldots + f(x_\nu) + \ldots = \sum_{i=1}^{\infty} f(x_i) = 1$$

# Συνάρτηση κατανομής διακριτής τμ

- Έστω $f$ η συνάρτηση πιθανότητας και $F$ η συνάρτηση κατανομής μιας διακριτής τ.μ. Χ με σύνολο τιμών $R_X = \{x_1, x_2, \ldots, x_\nu, \ldots\}$

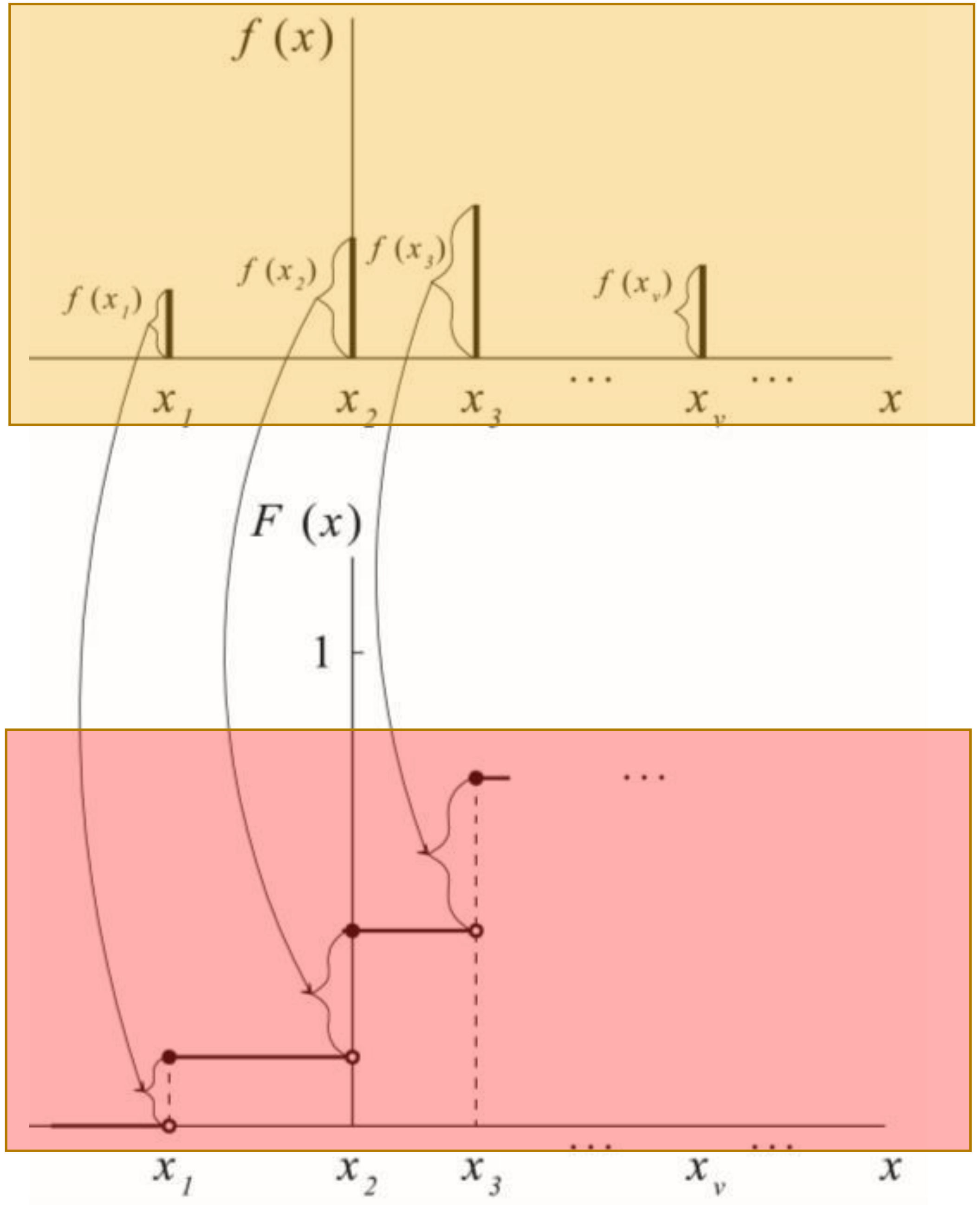
*α) Η F μπορεί να υπολογισθεί μέσω της f από τον τύπο*

$$F(x) = P(X \le x) = \sum_{i: x_i \le x} f(x_i), \; x \in R.$$

*β) Η f μπορεί να υπολογισθεί μέσω της F από τους τύπους*

$$f(x_1) = F(x_1) \;\; και \;\; f(x_k) = F(x_k) - F(x_{k-1}), \; k = 2, 3, \ldots$$

# Γραφική συσχέτιση συνάρτησης πιθανότητας και κατανομής



# Παράδειγμα

- Έστω Y η τυχαία μεταβλητή που εκφράζει τον αριθμό εργατικών ατυχημάτων που συμβαίνουν σε μια ημέρα στη βιομηχανική ζώνη Α. Στον πίνακα αντιστοιχίας που ακολουθεί φαίνεται η συνάρτηση πιθανότητας f της Y.

| $y$ | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| $f(y)=P(Y=y)$ | 0.366 | 0.370 | 0.182 | 0.062 | 0.016 | 0.004 |

| $y$ | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| $f(y)=P(Y=y)$ | 0.366 | 0.370 | 0.182 | 0.062 | 0.016 | 0.004 |
| $F(y)=P(Y\le y)$ | 0.366 | 0.736 | 0.918 | 0.980 | 0.996 | 1.000 |

## Συνάρτηση κατανομής διακριτής τμ

- Έστω $f$ η συνάρτηση πιθανότητας και $F$ η συνάρτηση κατανομής μιας διακριτής τ.μ. X με σύνολο τιμών $R_X = \{x_1, x_2, \dots, x_\nu, \dots\}$

*α) Η F μπορεί να υπολογισθεί μέσω της f από τον τύπο*
$$F(x) = P(X \le x) = \sum_{i: x_i \le x} f(x_i),\ x \in R.$$

*β) Η f μπορεί να υπολογισθεί μέσω της F από τους τύπους*
$$f(x_1) = F(x_1) \text{ και } f(x_k) = F(x_k) - F(x_{k-1}),\ k = 2,3,\dots$$

32

# Ομοιόμορφη διακριτή τμ

Η συνάρτηση πιθανότητας f μιας ομοιόμορφης διακριτής τ.μ. X με $R_X = \{x_1, x_2, \dots, x_\nu\}$ είναι:

$$f(x) = P(X = x) = \begin{cases} \dfrac{1}{\nu}, & x = x_1, x_2, \dots, x_\nu \\ 0, & x \neq x_1, x_2, \dots, x_\nu \end{cases}$$

Η συνάρτηση κατανομής είναι $F(x_k) = P(X \leq x_k) = \dfrac{k}{\nu}, \quad k = 1, 2, \dots, \nu$



Υποθέτουμε $x_1 < x_2 < \cdots < x_\nu$

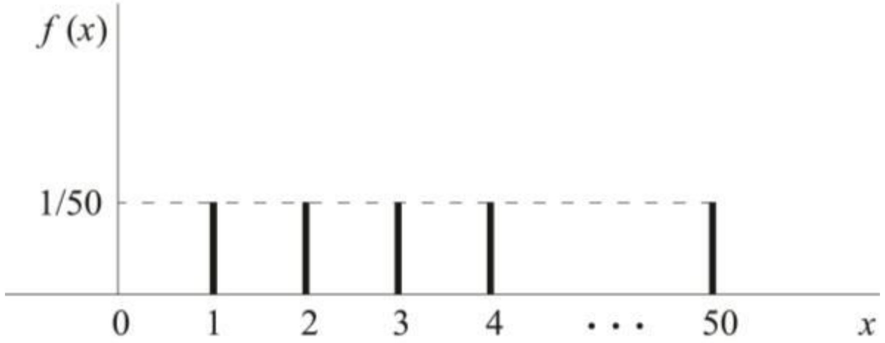
# Παράδειγμα

- Σε μια παρτίδα 50 προϊόντων υπάρχει ένα ελαττωματικό. Τα προϊόντα ελέγχονται το ένα μετά το άλλο, χωρίς επανάθεση, μέχρι να βρεθεί το ελαττωματικό. Ποια είναι η πιθανότητα να χρειαστούν:
  a) ακριβώς 12 έλεγχοι
  b) το πολύ 12 έλεγχοι
  c) περισσότεροι από 30 αλλά όχι περισσότεροι από 40 έλεγχοι

# Παράδειγμα (συν)

$$f(x) = P(X = x) = \begin{cases} \dfrac{1}{50}, & x = 1, 2, \ldots, 50 \\ 0, & x \neq 1, 2, \ldots, 50 \end{cases}$$

$$F(x_k) = P(X \le x_k) = \frac{k}{\nu}, \quad k = 1, 2, \ldots, 50$$



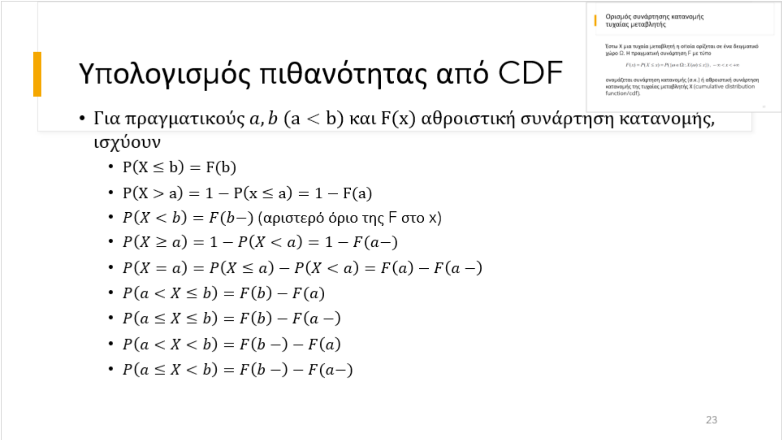
# Παράδειγμα (συν)

a) ακριβώς 12 έλεγχοι $P(X = 12) = f(12) = 1/50$

b) το πολύ 12 έλεγχοι $P(X \le 12) = F(12) = 12/50$

c) περισσότεροι από 30 αλλά όχι περισσότεροι από 40 έλεγχοι

$$P(30 < X \le 40) = F(40) - F(30) = 40/50 - 30/50 = 10/50$$



## Υπολογισμός πιθανότητας από CDF

- Για πραγματικούς $a, b$ ($a < b$) και $F(x)$ αθροιστική συνάρτηση κατανομής, ισχύουν
  - $P(X \le b) = F(b)$
  - $P(X > a) = 1 - P(x \le a) = 1 - F(a)$
  - $P(X < b) = F(b-)$ (αριστερό όριο της F στο x)
  - $P(X \ge a) = 1 - P(X < a) = 1 - F(a-)$
  - $P(X = a) = P(X \le a) - P(X < a) = F(a) - F(a-)$
  - $P(a < X \le b) = F(b) - F(a)$
  - $P(a \le X \le b) = F(b) - F(a-)$
  - $P(a < X < b) = F(b-) - F(a)$
  - $P(a \le X < b) = F(b-) - F(a-)$

### Ορισμός συνάρτησης κατανομής τυχαίας μεταβλητής

Έστω X μια τυχαία μεταβλητή η οποία ορίζεται σε ένα δειγματικό χώρο Ω. Η πραγματική συνάρτηση F με τύπο

$$F(x) = P(X \le x) = P(\{\omega \in \Omega : X(\omega) \le x\}), \ -\infty < x < +\infty$$

ονομάζεται συνάρτηση κατανομής (σ.κ.) ή αθροιστική συνάρτηση κατανομής της τυχαίας μεταβλητής X (cumulative distribution function/cdf).

23

# Παράμετροι διακριτής τυχαίας μεταβλητής

# Παράδειγμα

- Το κέρδος (ή ζημιά) ενός παραγωγού ανά τεμάχιο προϊόντος που παράγει είναι τυχαία μεταβλητή με τις πιθανότητες του παρακάτω πίνακα:

| $x$ (σε €) | -2 | 1 | 3 | 5 |
|---|---|---|---|---|
| $f(x) = P(X = x)$ | 0.02 | 0.10 | 0.80 | 0.08 |

- Πόσο είναι το αναμενόμενο κέρδος;

# Παράδειγμα (συν)

$$\frac{-2+1+3+5}{4} = 1.75€$$

$$(-2)\cdot 0.02 + 1\cdot 0.10 + 3\cdot 0.80 + 5\cdot 0.08 = 2.86€$$

$$E(X) = (-2)\cdot P(X=-2) + 1\cdot P(X=1) + 3\cdot P(X=3) + 5\cdot P(X=5)$$

| $x$ (σε €) | -2 | 1 | 3 | 5 |
|---|---|---|---|---|
| $f(x) = P(X=x)$ | 0.02 | 0.10 | 0.80 | 0.08 |

# Μέση τιμή / αναμενόμενη τιμή (expected value)

Έστω Χ μια διακριτή τυχαία μεταβλητή με σύνολο τιμών $R_X = \{x_1, x_2, \ldots, x_\nu, \ldots\}$ και συνάρτηση πιθανότητας $f$. Η μέση τιμή της Χ (mean value) συμβολίζεται $E(X)$ και δίνεται από τον τύπο

$$E(X) = \sum_{i=1}^{\infty} x_i f(x_i) = x_1 f(x_1) + x_2 f(x_2) + \ldots + x_\nu f(x_\nu) + \ldots$$

$$\acute{\eta}$$

$$E(X) = x_1 P(X = x_1) + x_2 P(X = x_2) + \ldots + x_\nu P(X = x_\nu) + \ldots$$

$$\acute{\eta}$$

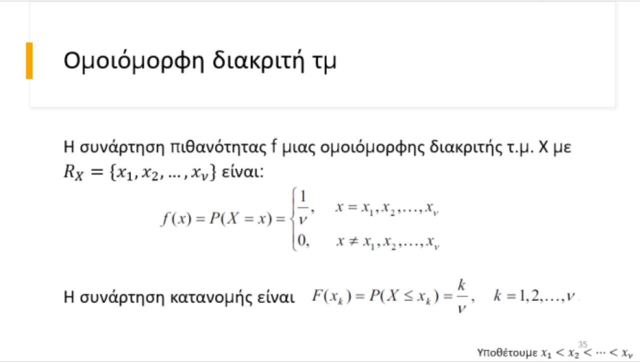
$$E(X) = \sum_{x \in R_X} x f(x) = \sum_{x \in R_X} x P(X = x)$$

*εφόσον η σειρά* $\sum_{i=1}^{\infty} x_i f(x_i)$ *συγκλίνει* απόλυτα, δηλαδή, αν $\sum_{i=1}^{\infty} |x_i| f(x_i) < \infty$

# Σχόλια

- Η μέση τιμή μιας τυχαίας μεταβλητής Χ είναι μια ένδειξη για τη θέση γύρω από την οποία παίρνει τιμές η τ.μ.

# Ομοιόμορφη διακριτή τμ



## Ομοιόμορφη διακριτή τμ

Η συνάρτηση πιθανότητας f μιας ομοιόμορφης διακριτής τ.μ. X με $R_X = \{x_1, x_2, \dots, x_\nu\}$ είναι:

$$f(x) = P(X = x) = \begin{cases} \frac{1}{\nu}, & x = x_1, x_2, \dots, x_\nu \\ 0, & x \neq x_1, x_2, \dots, x_\nu \end{cases}$$

Η συνάρτηση κατανομής είναι $F(x_k) = P(X \leq x_k) = \frac{k}{\nu}, \quad k = 1, 2, \dots, \nu$

35

Υποθέτουμε $x_1 < x_2 < \dots < x_\nu$

- Η μέση τιμή της είναι

$$\mu = E(X) = \sum_{i=1}^{\nu} x_i f(x_i) = x_1 f(x_1) + x_2 f(x_2) + \ldots + x_\nu f(x_\nu) =$$

$$= x_1 \frac{1}{\nu} + x_2 \frac{1}{\nu} + \ldots + x_\nu \frac{1}{\nu} = \frac{x_1 + x_2 + \ldots + x_\nu}{\nu}$$

# Μέση τιμή συνάρτησης τμ

- Έστω X μια διακριτή τυχαία μεταβλητή με σύνολο τιμών $R_X = \{x_1, x_2, \dots., x_\nu, \dots\}$ και συνάρτηση πιθανότητας $f$. Για οποιαδήποτε πραγματική συνάρτηση $g$ της X, η μέση τιμή της τυχαίας μεταβλητής $g(X)$ δίνεται από τον τύπο

$$E[g(X)] = \sum_{x \in R_X} g(x) f(x)$$

$$\text{εφόσον η σειρά } \sum_{x \in R_X} g(x) f(x) \text{ συγκλίνει απόλυτα.}$$

# Γραμμικότητα της μέσης τιμής

- Αν X διακριτή τυχαία μεταβλητή και $\alpha, \beta$, πραγματικοί αριθμοί, τότε

$$E(\alpha X + \beta) = \alpha E(X) + \beta$$

# Παράδειγμα

- Ο αριθμός των προϊόντων που πουλάει μια επιχείρηση σε διάστημα μιας εβδομάδας είναι διακριτή τυχαία μεταβλητή, έστω Χ, με συνάρτηση πιθανότητας

$$f(x) = P(X = x) = \begin{cases} 0.1, & x = 0 \\ 0.2, & x = 1 \\ 0.6, & x = 2 \\ 0.1, & x = 3. \end{cases}$$

- Έστω ότι το κέρδος από την πώληση ενός προϊόντος είναι 200€ και ότι τα πάγια εβδομαδιαία έξοδα της επιχείρησης είναι 100€. Ενδιαφερόμαστε για το αναμενόμενο εβδομαδιαίο κέρδος της επιχείρησης

# Παράδειγμα (συν)

- Πρέπει να υπολογίσουμε τη μέση τιμή της τυχαίας μεταβλητής

$$Y = g(X) = 200X - 100$$

- Η μέση τιμή της X είναι

$$E(X) = 0 \cdot 0.1 + 1 \cdot 0.2 + 2 \cdot 0.6 + 3 \cdot 0.1 = 1.7$$

- Οπότε η μτ της X ειναι

$$E(200X - 100) = 200E(X) - 100 = 200 \cdot 1.7 - 100 = 240\,€.$$

# Μέση τιμή γραμμικού συνδυασμού συναρτήσεων διακριτής τυχαίας μεταβλητής

- Έστω X μια διακριτή τυχαία μεταβλητή και $k$ πραγματικές συναρτήσεις $g_1, g_2, g_3, \dots, g_k$ της X. Αν $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_k$ είναι πραγματικοί αριθμοί, τότε

$$E[\alpha_1 g_1(X) + \alpha_2 g_2(X) + \ldots + \alpha_k g_k(X)] =$$
$$= \alpha_1 E[g_1(X)] + \alpha_2 E[g_2(X)] + \ldots + \alpha_k E[g_k(X)]$$

# Παράδειγμα

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

$$E(\alpha \ln X + \beta \eta\mu X + \gamma) = \alpha E(\ln X) + \beta E(\eta\mu X) + \gamma$$

$$E[(X-\alpha)^2] = E(X^2 - 2\alpha X + \alpha^2) = E(X^2) - 2\alpha E(X) + \alpha^2$$

# Μέση τιμή γραμμικού συνδυασμού διακριτών τυχαίων μεταβλητών

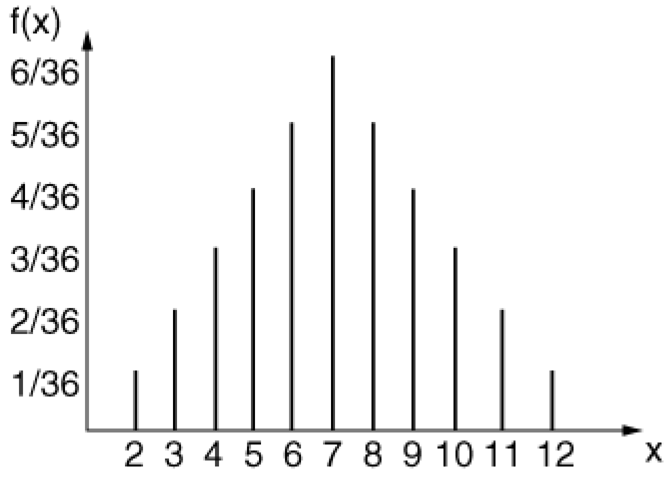
- Για οποιεσδήποτε διακριτές τυχαίες μεταβλητές $\mathrm{X}_1, \mathrm{X}_2, \mathrm{X}_3, \ldots, \mathrm{X}_k$ των οποίων οι μέσες τιμές υπάρχουν και για οποιουσδήποτε πραγματικούς αριθμούς $\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_k$ ισχύει:

$$E(\alpha_1 X_1 + \alpha_2 X_2 + \ldots + \alpha_k X_k) = \alpha_1 E(X_1) + \alpha_2 E(X_2) + \ldots + \alpha_k E(X_k)$$

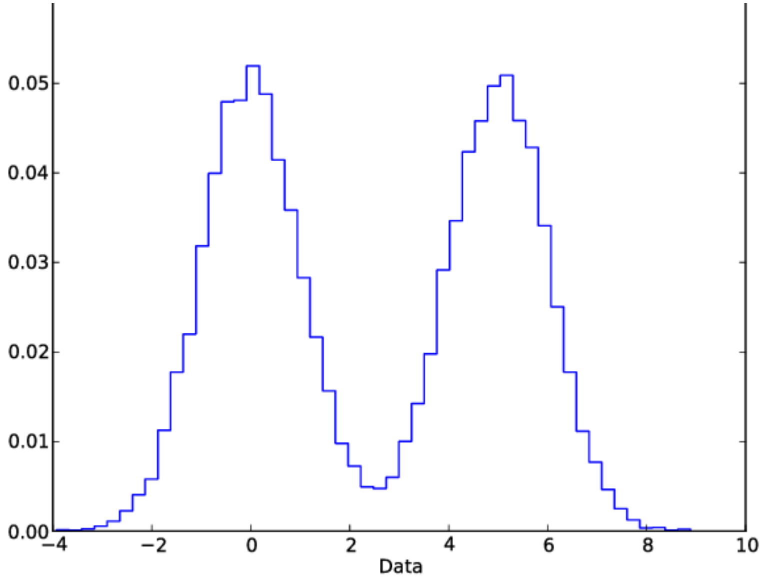
# Κορυφή (mode)

Έστω $x_1, x_2, x_3, \ldots, x_k$ οι δυνατές τιμές μιας διακριτής τυχαίας μεταβλητής $X$ και έστω επίσης, $x_1 < x_2 < x_3 < \cdots < x_k$. Μια τιμή $x_\nu$ της X λέγεται κορυφή (mode) της κατανομής της αν

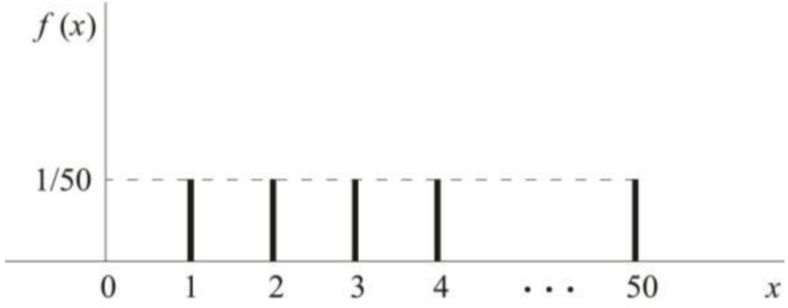
$$P(X = x_{\nu-1}) < P(X = x_\nu) \;\; \text{και} \;\; P(X = x_{\nu+1}) < P(X = x_\nu)$$




μονοκόρυφη




δικόρυφη



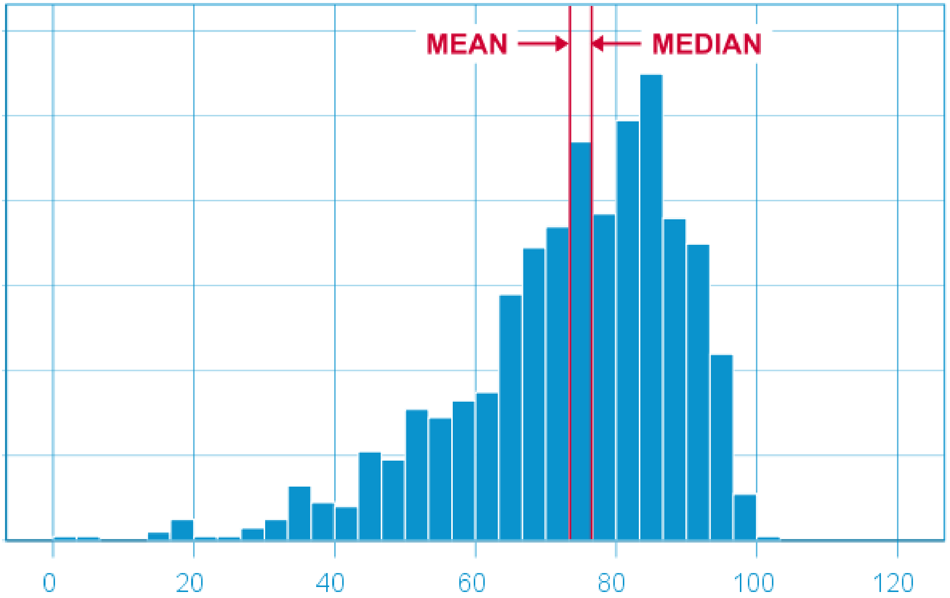

καμία κορυφή

# Διάμεσος

Ένας πραγματικός αριθμός $\delta$ ονομάζεται διάμεσος (median) μιας τ.μ. ή της κατανομής μιας τ.μ. X, αν

$$P(X < \delta) \leq \frac{1}{2} \text{ και } P(X > \delta) \leq \frac{1}{2}$$

Η *διάμεσος* χωρίζει την κατανομή σε δύο ίσα (με όρους πιθανότητας) μέρη.



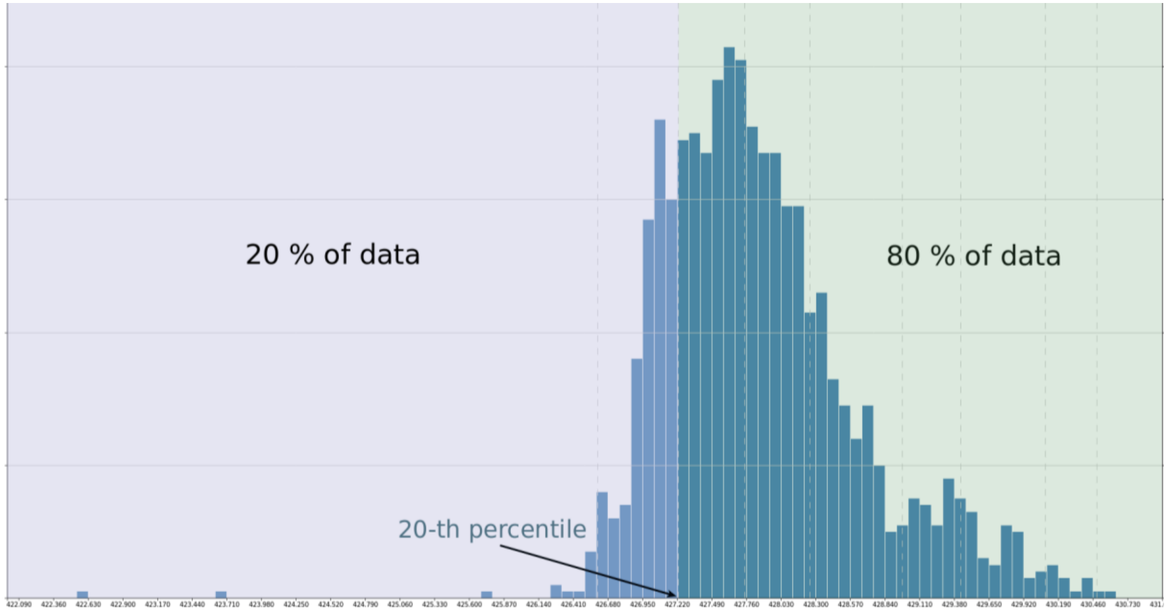
# p-ποσοστιαίο σημείο

Ένας πραγματικός αριθμός $x_p, 0 < p < 1$ p-ποσοστιαίο σημείο (quantile) ή 100p-οστό εκατοστημόριο (percentile) ή της κατανομής μιας τ.μ. Χ, αν

$$P(X < x_p) \leq p \text{ και } P(X > x_p) \leq 1 - p$$

το p-ποσοστιαίο σημείο χωρίζει την κατανομή σε δύο μέρη υπό την αναλογία $p$: $(1 - p)$ με 0<$p$<1

το 0.5-ποσοστιαίο σημείο ταυτίζεται με την διαμεσο



# Παράμετροι διασποράς: διακύμανση και τυπική απόκλιση =>

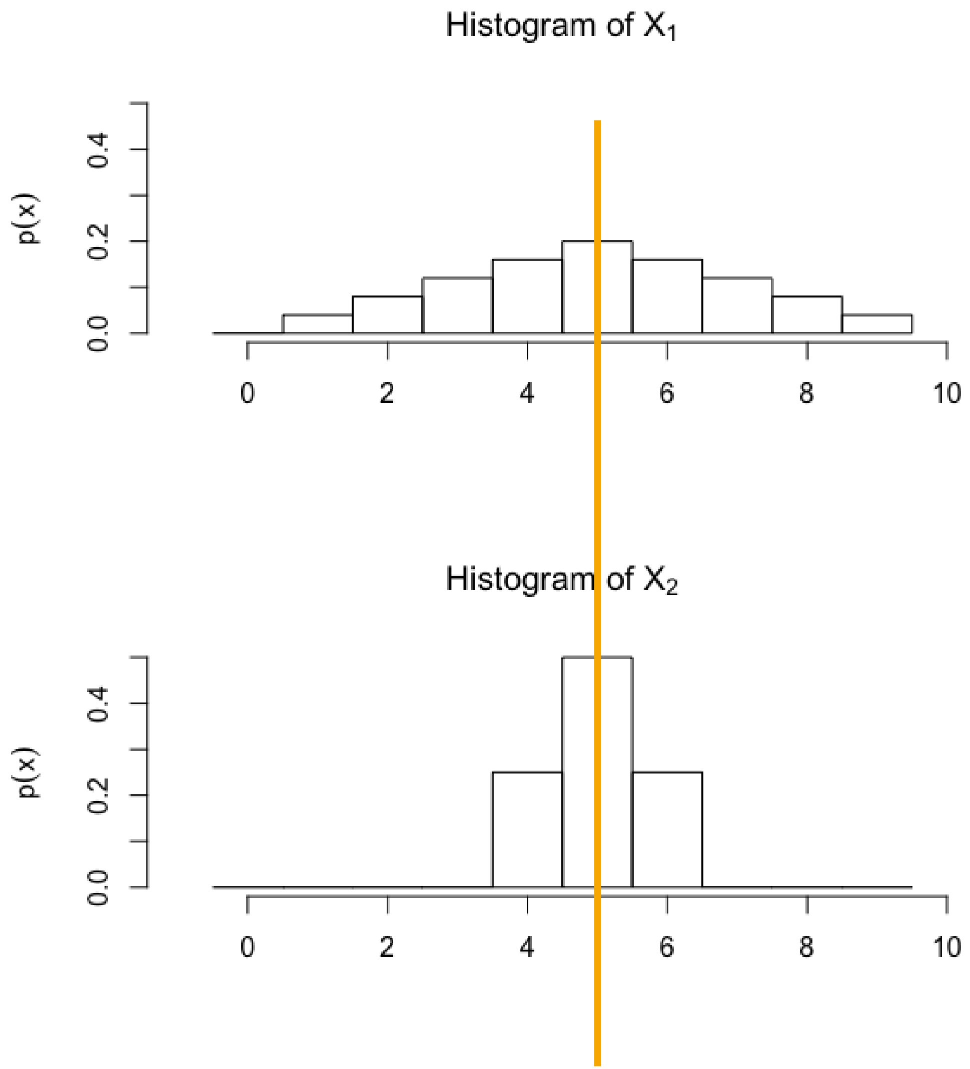
# Κίνητρο

| $x$ | -1 | 1 |
|---|---|---|
| $P(X = x)$ | 0.5 | 0.5 |

| $y$ | -1000 | 1000 |
|---|---|---|
| $P(Y = y)$ | 0.5 | 0.5 |

# Κίνητρο (συν)

- Η μέση τιμή μας δίνει όπως είδαμε μια ένδειξη για τη θέση γύρω από την οποία παίρνει τιμές η τ.μ. (σταθμισμένος μέσος)
- Δεν μας δίνει απολύτως καμία πληροφορία για το πόσο αυτές είναι διασκορπισμένες (ή συγκεντρωμένες)



# Παράδειγμα


- Η απόδοση X μιας μετοχής, έστω A, μπορεί να είναι 20%, 1% ή -20 % , ανάλογα με το αν η οικονομία το ερχόμενο έτος βρεθεί σε κατάσταση ανάπτυξης, στασιμότητας ή ύφεσης. Η απόδοση Y μιας άλλης μετοχής, έστω B, μπορεί αντίστοιχα να είναι 3%, 1%, -3%. Για την οικονομία, η πιθανότητα να βρεθεί σε κατάσταση ανάπτυξης, στασιμότητας, ύφεσης, είναι αντίστοιχα 0.10, 0.80, 0.10.

# Παράδειγμα (συν)


- Αναμενόμενη απόδοση της μετοχής Α:


$$\mu_X = E(X) = \sum_x x f(x) = 0.2 \cdot 0.1 + 0.01 \cdot 0.8 + (-0.2) \cdot 0.1 = 0.008$$

- Αναμενόμενη απόδοση της μετοχής Β:

$$\mu_Y = E(Y) = \sum_y y f(y) = 0.03 \cdot 0.1 + 0.01 \cdot 0.8 + (-0.03) \cdot 0.1 = 0.008$$


Μέση τιμή / αναμενόμενη τιμή (expected value)

Έστω X μια διακριτή τυχαία μεταβλητή με σύνολο τιμών $R_X = \{x_1, x_2, \dots, x_\nu, \dots\}$ και συνάρτηση πιθανότητας $f$. Η μέση τιμή της X (mean value) συμβολίζεται $E(X)$ και δίνεται από τον τύπο

$$E(X) = \sum_{i=1}^{\infty} x_i f(x_i) = x_1 f(x_1) + x_2 f(x_2) + \dots + x_\nu f(x_\nu) + \dots$$

ή

$$E(X) = x_1 P(X = x_1) + x_2 P(X = x_2) + \dots + x_\nu P(X = x_\nu) + \dots$$

ή

$$E(X) = \sum_{x \in R_X} x f(x) = \sum_{x \in R_X} x P(X = x)$$

*εφόσον η σειρά* $\sum_{i=1}^{\infty} x_i f(x_i)$ *συγκλίνει* απόλυτα, δηλαδή, αν $\sum_{i=1}^{\infty} |x_i| f(x_i) < \infty$

42

# Διακύμανση διακριτής τυχαίας μεταβλητής

- Έστω Χ μια διακριτή τυχαία μεταβλητή με σύνολο τιμών $R_X = \{x_1, x_2, \ldots., x_\nu, \ldots\}$ και συνάρτηση πιθανότητας $f$. Έστω επίσης ότι υπάρχει η μέση τιμή $\mathrm{E(X)} = \mu$. Διακύμανση (variance) της Χ ονομάζεται η μέση τιμή της τυχαίας μεταβλητής $(\mathrm{X} - \mu)^2$ και συμβολίζεται με $Var(X)$ η $\sigma^2$.

$$\sigma^2 = Var(X) = E[(X-\mu)^2] = \sum_{i=1}^{\infty}(x_i - \mu)^2 f(x_i)$$

ή

$$\sigma^2 = Var(X) = \sum_{x \in R_X}(x-\mu)^2 f(x).$$

# Τυπική απόκλιση

- Η θετική τετραγωνική ρίζα της διακύμανσης της X, ονομάζεται τυπική απόκλιση της X (standard deviation).

$$\sigma = \sqrt{Var(X)} = \sqrt{E[(X-\mu)^2]}$$

# Σχόλια

- Η τυπική απόκλιση μιας τυχαίας μεταβλητής Χ είναι ένα μέτρο της συγκέντρωσης (ή της διασποράς) των τιμών της γύρω από τη μέση τιμή της.
- Μια ανεπιθύμητη συνέπεια της εμφάνισης τετραγωνικής δύναμης στον ορισμό της διακύμανσης είναι ότι η μονάδα μέτρησής της δεν είναι η ίδια με τη μονάδα μέτρησης της τυχαίας μεταβλητής αλλά με το τετράγωνό της.
- Στην τυπική απόκλιση δεν εμφανίζεται το πρόβλημα αυτό .

# Παράδειγμα (συν)


- Η απόδοση Χ μιας μετοχής, έστω Α, μπορεί να είναι 20%, 1% ή -20 % , ανάλογα με το αν η οικονομία το ερχόμενο έτος βρεθεί σε κατάσταση ανάπτυξης, στασιμότητας ή ύφεσης. Η απόδοση Υ μιας άλλης μετοχής, έστω Β, μπορεί αντίστοιχα να είναι 3%, 1%, -3%. Για την οικονομία, η πιθανότητα να βρεθεί σε κατάσταση ανάπτυξης, στασιμότητας, ύφεσης, είναι αντίστοιχα 0.10, 0.80, 0.10.

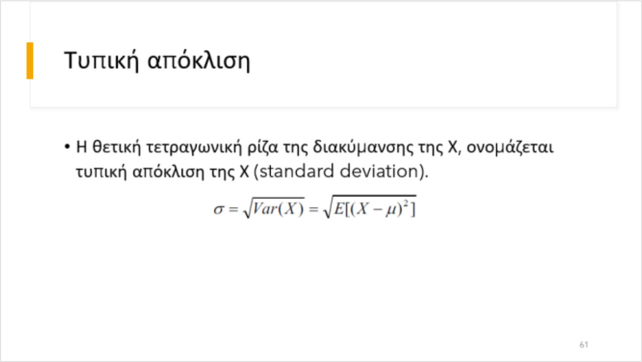
# Παράδειγμα (συν)

- Διακύμανση και την τυπική απόκλιση της απόδοσης κάθε μετοχής

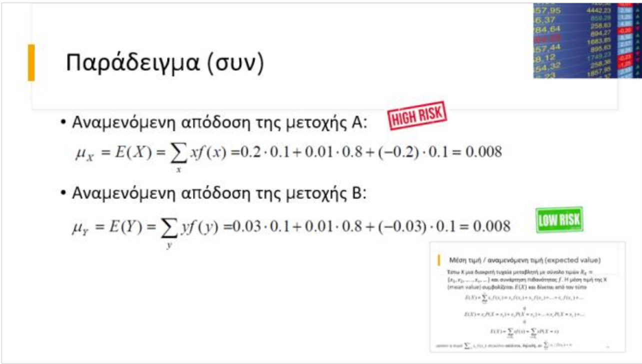
$$\sigma_X^2 = \sum_{x \in R_X} (x-\mu_X)^2 f(x) = (0.2-0.008)^2 \cdot 0.1 + (0.01-0.008)^2 \cdot 0.8 + (-0.2-0.008)^2 \cdot 0.1 = 0.008$$

$$\sigma_Y^2 = \sum_{y \in R_Y} (y-\mu_Y)^2 f(y) = (0.03-0.008)^2 \cdot 0.1 + (0.01-0.008)^2 \cdot 0.8 + (-0.03-0.008)^2 \cdot 0.1 = 0.0002$$

$$\sigma_X = \sqrt{0.008} = 0.089 \qquad \sigma_Y = \sqrt{0.0002} = 0.014$$




0.089/0.014=6.357



# Χρήσιμες ιδιότητες

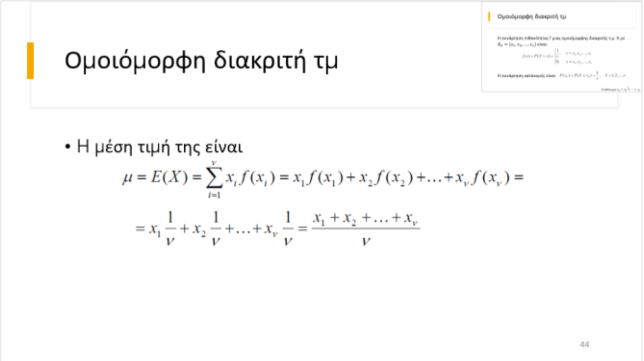
$$\sigma^2 = Var(X) = E(X^2) - [E(X)]^2 = E(X^2) - \mu^2$$

$E(X^2) \geq [E(X)]^2$, ή ισοδύναμα, $[E(X)]^2 \leq E(X^2)$

# Ομοιόμορφή τμ

- Η διακύμανση της X με βάση του του τύπου

$$Var(X) = E(X^2) - [E(X)]^2 = \frac{x_1^2 + x_2^2 + \ldots + x_\nu^2}{\nu} - \left( \frac{x_1 + x_2 + \ldots + x_\nu}{\nu} \right)^2$$


Ομοιόμορφη διακριτή τμ

- Η μέση τιμή της είναι

$$\mu = E(X) = \sum_{i=1}^{\nu} x_i f(x_i) = x_1 f(x_1) + x_2 f(x_2) + \ldots + x_\nu f(x_\nu) =$$

$$= x_1 \frac{1}{\nu} + x_2 \frac{1}{\nu} + \ldots + x_\nu \frac{1}{\nu} = \frac{x_1 + x_2 + \ldots + x_\nu}{\nu}$$

44

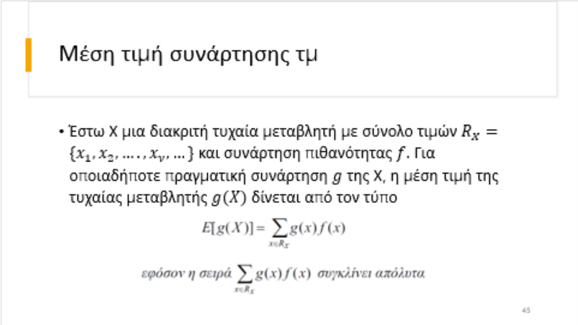
# Παράδειγμα (συν)

- Έστω Y η τυχαία μεταβλητή που εκφράζει τον αριθμό εργατικών ατυχημάτων που συμβαίνουν σε μια ημέρα στη βιομηχανική ζώνη Α. Στον πίνακα αντιστοιχίας που ακολουθεί φαίνεται η συνάρτηση πιθανότητας f της Y.

| $y$ | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| $f(y) = P(Y = y)$ | 0.366 | 0.370 | 0.182 | 0.062 | 0.016 | 0.004 |

$$E(Y) = \sum_{y \in R_Y} y f(y) = \sum_{y=0}^{5} y f(y) =$$

$$= 0 \cdot (0.366) + 1 \cdot (0.370) + 2 \cdot (0.182) + 3 \cdot (0.062) + 4 \cdot (0.016) + 5 \cdot (0.004) = 1.004.$$



## Μέση τιμή συνάρτησης τμ

- Έστω X μια διακριτή τυχαία μεταβλητή με σύνολο τιμών $R_X = \{x_1, x_2, \dots, x_\nu, \dots\}$ και συνάρτηση πιθανότητας $f$. Για οποιαδήποτε πραγματική συνάρτηση $g$ της X, η μέση τιμή της τυχαίας μεταβλητής $g(X)$ δίνεται από τον τύπο

$$E[g(X)] = \sum_{x \in R_X} g(x) f(x)$$

$$\text{εφόσον η σειρά } \sum_{x \in R_X} g(x) f(x) \text{ συγκλίνει απόλυτα}$$

45

# Παράδειγμα (συν)

Υπολογισμός της διακύμανσης

$$\sigma^2 = Var(Y) = E(Y^2) - [E(Y)]^2$$

$$E(Y^2) = \sum_{y \in R_y} y^2 f(y) = \sum_{y=0}^{5} y^2 f(y) =$$

$$= 0^2 \cdot (0.366) + 1^2 \cdot (0.370) + 2^2 \cdot (0.182) + 3^2 \cdot (0.062) + 4^2 \cdot (0.016) + 5^2 \cdot (0.004) = 2.02$$

Χρήσιμες ιδιότητες

$$\sigma^2 = Var(X) = E(X^2) - [E(X)]^2 = E(X^2) - \mu^2$$

$E(X^2) \geq [E(X)]^2$, ή ισοδύναμα, $[E(X)]^2 \leq E(X^2)$

65

$$\sigma^2 = Var(Y) = E(Y^2) - [E(Y)]^2 = 2.02 - (1.004)^2 = 1.01$$

$$\sigma = \sqrt{Var(Y)} = \sqrt{1.01} \cong 1$$

# Διακύμανση και η τυπική απόκλιση του γραμμικού συνδυασμού

- Αν X διακριτή τυχαία μεταβλητή και $\alpha, \beta$, πραγματικοί αριθμοί, τότε η διακύμανση και η τυπική απόκλιση του γραμμικού συνδυασμού $\alpha \mathrm{X} + \beta$, δίνονται από τους τύπους

$$Var(\alpha X + \beta) = \alpha^2 Var(X)$$

$$\sigma_{\alpha X + \beta} = |\alpha| \sigma_X .$$

# Παράδειγμα (συν)

- Ο αριθμός των προϊόντων που πουλάει μια επιχείρηση σε διάστημα μιας εβδομάδας είναι διακριτή τυχαία μεταβλητή, έστω Χ, με συνάρτηση πιθανότητας

$$f(x) = P(X = x) = \begin{cases} 0.1, & x = 0 \\ 0.2, & x = 1 \\ 0.6, & x = 2 \\ 0.1, & x = 3. \end{cases}$$

- Έστω ότι το κέρδος από την πώληση ενός προϊόντος είναι 200€ και ότι τα πάγια εβδομαδιαία έξοδα της επιχείρησης είναι 100€. Ενδιαφερόμαστε για το αναμενόμενο εβδομαδιαίο κέρδος της επιχείρησης

# Παράδειγμα (συν)

- Η μτ της Y είναι

$$E(200X - 100) = 200E(X) - 100 = 200 \cdot 1.7 - 100 = 240\ €.$$

- Η διακύμανση της Y είναι

$$V(200X - 100) = 200^2 V(X)$$

$$Var(X) = E(X^2) - [E(X)]^2 =$$

$$= 0^2 \cdot 0.1 + 1^2 \cdot 0.2 + 2^2 \cdot 0.6 + 3^2 \cdot 0.1 - (1.7)^2 = 3.5 - 2.89 = 0.61$$

$$Var(Y) = Var(200X - 100) = 200^2 Var(X) = 200^2 \cdot 0.61 = 24400\ €^2$$

$$\sigma_Y = \sqrt{24400} = 156.2\ €.$$

Διακύμανση και η τυπική απόκλιση του γραμμικού συνδυασμού

- Αν X διακριτή τυχαία μεταβλητή και $\alpha, \beta$, πραγματικοί αριθμοί, τότε η διακύμανση και η τυπική απόκλιση του γραμμικού συνδυασμού $\alpha X + \beta$, δίνονται από τους τύπους

$$Var(\alpha X + \beta) = \alpha^2 Var(X)$$

$$\sigma_{\alpha X+\beta} = |\alpha|\sigma_X .$$

69

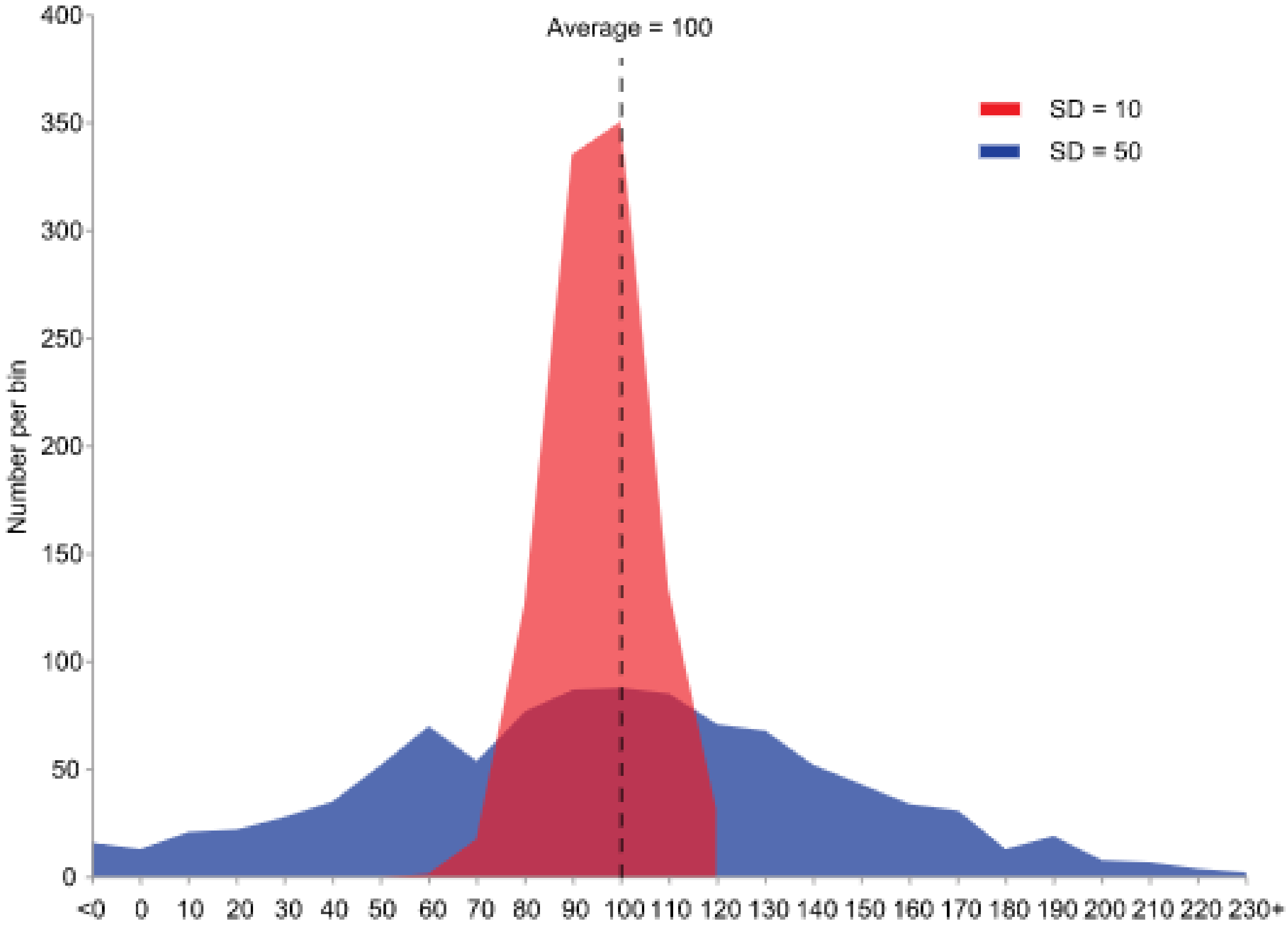
# Τυποποιημένη τμ

- Μια τυχαία μεταβλητή με μέση τιμή 0 και διακύμανση 1 (άρα και τυπική απόκλιση 1), ονομάζεται τυποποιημένη τυχαία μεταβλητή (standardized random variable)
- Κάθε τμ μπορεί να τυποποιηθεί αφαιρώντας την μέση τιμή και διαιρώντας με την τυπική απόκλιση

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

# Υπολογισμός της πιθανότητας μιας τ.μ. σε σχέση με την μέση τιμή και της τυπικής απόκλισης

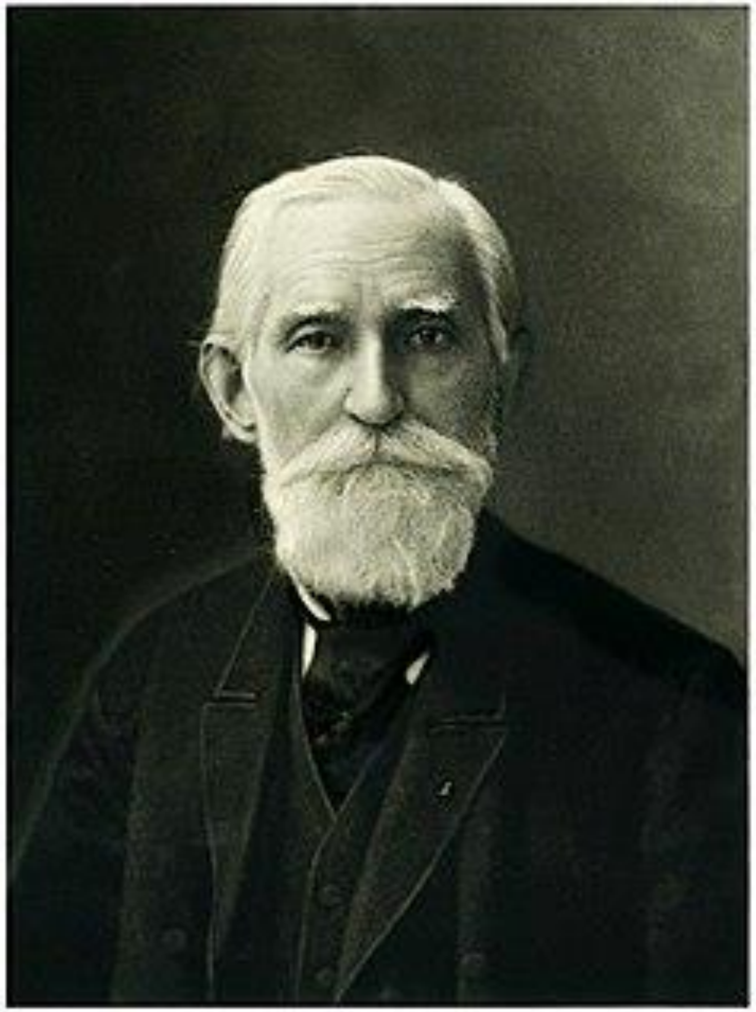
- Η μέση τιμή μιας τυχαίας μεταβλητής X είναι μια ένδειξη για τη θέση γύρω από την οποία παίρνει τιμές η τ.μ.
- Η τυπική απόκλιση μιας τυχαίας μεταβλητής X είναι ένα μέτρο της συγκέντρωσης (ή της διασποράς) των τιμών της γύρω από τη μέση τιμή της.



# Ανισότητα Chebyshev

- Αν Χ μια τυχαία μεταβλητή με μέση τιμή $\mu$ και διακύμανση $\sigma^2$ , τότε, για κάθε $c > 0$ , ισχύει ότι

$$P(|X - \mu| \geq c) \leq \frac{\sigma^2}{c^2}$$



# Ανισότητα Chebyshev (συν)

- Αν θέσουμε $c = k\sigma$, τότε

$$P(|X-\mu| \geq k\sigma) \leq \frac{\sigma^2}{k^2\sigma^2} = \frac{1}{k^2}$$

$$P(|X-\mu| < k\sigma) \geq 1-\frac{1}{k^2} \Leftrightarrow P(\mu - k\sigma < X < \mu + k\sigma) \geq 1-\frac{1}{k^2}$$

# Ανισότητα Chebyshev (παραδείγματα)

- Για οποιαδήποτε τυχαία μεταβλητή X, η πιθανότητα να πάρει τιμή:
  - στο διάστημα (μ – 2σ , μ + 2σ ) είναι τουλάχιστον 0.75
  - στο διάστημα (μ – 2.5 σ , μ + 2.5 σ ) είναι τουλάχιστον 0.84
  - στο διάστημα (μ – 3σ , μ + 3σ ) είναι τουλάχιστον 0.89

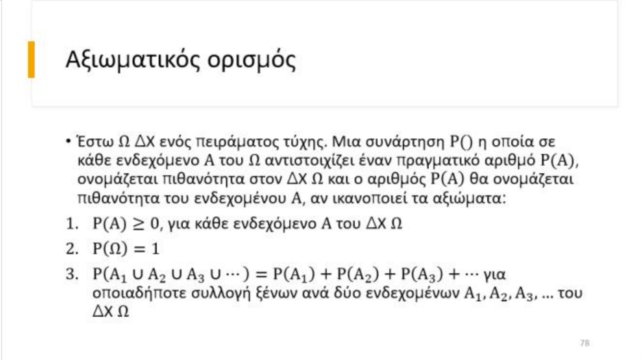
| **k** | **Min. % within *k* standard deviations of mean** |
|---|---|
| 1 | 0% |
| $\sqrt{2}$ | 50% |
| 1.5 | 55.56% |
| 2 | 75% |
| $2\sqrt{2}$ | 87.5% |
| 3 | 88.8889% |
| 4 | 93.75% |
| 5 | 96% |
| 6 | 97.2222% |
| 7 | 97.9592% |
| 8 | 98.4375% |
| 9 | 98.7654% |
| 10 | 99% |

# Backup

# Αξιωματικός ορισμός

- Έστω $\Omega$ ΔΧ ενός πειράματος τύχης. Μια συνάρτηση $\mathrm{P}()$ η οποία σε κάθε ενδεχόμενο $\mathrm{A}$ του $\Omega$ αντιστοιχίζει έναν πραγματικό αριθμό $\mathrm{P}(\mathrm{A})$, ονομάζεται πιθανότητα στον ΔΧ $\Omega$ και ο αριθμός $\mathrm{P}(\mathrm{A})$ θα ονομάζεται πιθανότητα του ενδεχομένου $\mathrm{A}$, αν ικανοποιεί τα αξιώματα:

1. $\mathrm{P}(\mathrm{A}) \geq 0$, για κάθε ενδεχόμενο Α του ΔΧ $\Omega$
2. $\mathrm{P}(\Omega) = 1$
3. $\mathrm{P}(\mathrm{A}_1 \cup \mathrm{A}_2 \cup \mathrm{A}_3 \cup \cdots) = \mathrm{P}(\mathrm{A}_1) + \mathrm{P}(\mathrm{A}_2) + \mathrm{P}(\mathrm{A}_3) + \cdots$ για οποιαδήποτε συλλογή ξένων ανά δύο ενδεχομένων $\mathrm{A}_1, \mathrm{A}_2, \mathrm{A}_3, \ldots$ του ΔΧ $\Omega$




# Ιδιότητες

- $P(\emptyset) = 0$
- Για οποιαδήποτε ξένα ανά δύο ενδεχόμενα $A_1, A_2, A_3, \ldots, A_\nu$ ισχύει $P(A_1 \cup A_2 \cup A_3 \cup \cdots \cup A_\nu) = P(A_1) + P(A_2) + P(A_3) + \cdots + P(A_\nu)$ (πεπερασμένη προσθετικότητα)
- Αν $A = \{\alpha_1, \alpha_2, \ldots\}$ είναι αριθμησίμως άπειρα ενδεχόμενο του ΔΧ Ω, τότε $P(A) = P(\{\alpha_1\}) + P(\{\alpha_2\}) + \cdots$ (ισχύει και για πεπερασμένα ενδεχόμενα)

# Δεσμευμένη πιθανότητα (ορισμός)

- Έστω ενδεχόμενα A και B του ΔΧ Ω, και P(B)>0, τότε η δεσμευμένη πιθανότητα του A δοθέντος του B (conditional probability) ισούται με

$$\mathrm{P}(\mathrm{A}|\mathrm{B}) = \frac{P(AB)}{P(B)}$$

- Εναλλακτικός συμβολισμός: $\mathrm{P}_B(A)$

# Συμπληρωματικοί ορισμοί

- Αν σε μια εκτέλεση πειράματος εμφανίστηκε το ενδεχόμενο $\omega$ και $\omega \in \mathrm{A}$, τότε λέμε το ενδεχόμενο $\mathrm{A}$ υλοποιήθηκε/συνέβη.
- Ο $\Omega$ περιέχει όλα τα ενδεχόμενα και οπότε είναι το «σίγουρο» ενδεχόμενο, ενώ το κενό σύνολο ($\emptyset$) είναι το αδύνατο ενδεχόμενο.
- Πεπερασμένος δειγματικός χώρος (finite SS): αν ένα πείραμα έχει πεπερασμένο σύνολο δυνατών αποτελεσμάτων ($\Omega = \{\omega_1, \omega_2, \dots, \omega_{\mathrm{N}}\}$).
- Αριθμήσιμος άπειρος ΔΧ (countably infinite SS): αν το πείραμα έχει άπειρα πολλά δυνατά αποτελέσματα, αλλά μπορούμε να τους αποδώσουμε ένα φυσικό αριθμό ($\Omega = \{\omega_1, \omega_2, \omega_3, \dots\}$).
- Συνεχής ΔΧ (continuous SS): αν το πείραμα έχει άπειρα πολλά δυνατά αποτελέσματα αλλά μη αριθμήσιμα.
- Διακριτός ΔΧ (discrete SS): Πεπερασμένος ή Αριθμήσιμος άπειρος ΔΧ