Πανεπιστήμιο Αιγαίου

Πιθανότητες και Στατιστική

Ελισάβετ Κωνσταντίνου

ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ
Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο
ΕΣΠΑ 2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ
Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

Άδειες Χρήσης

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Αιγαίου**» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.

Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο

ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

ΕΣΠΑ
2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

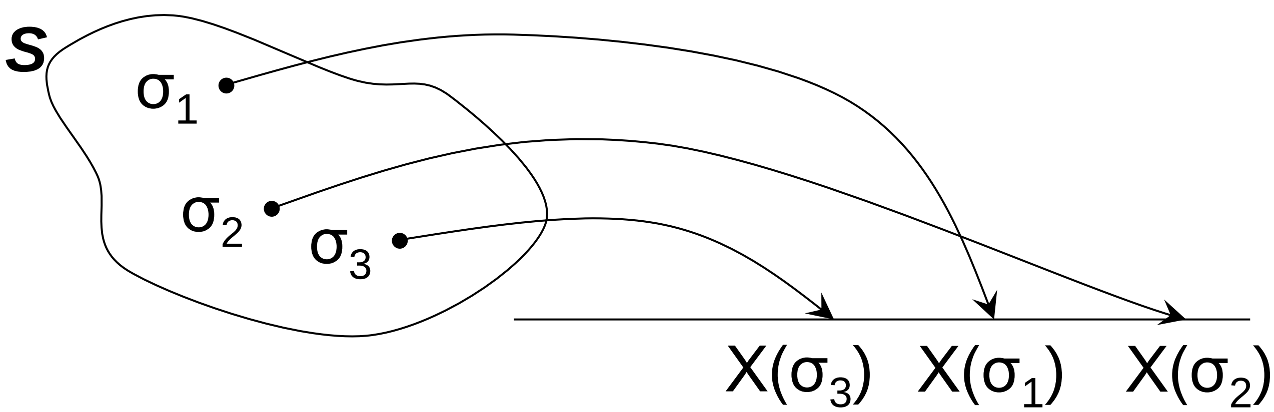
Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

Εργαλεία

- Τυχαίες μεταβλητές (διακριτές – συνεχείς)
- Συνάρτηση κατανομής
- Πυκνότητα πιθανότητας
- Διακριτή συνάρτηση πιθανότητας
- Ροπές (μέση τιμή, διασπορά κλπ)
- Συναρτήσεις τ.μ.
- Περιγραφική Στατιστική
- Εκτίμηση παραμέτρων
- Ελεγχος υποθέσεων

Τυχαίες μεταβλητές

- Πολλοί δειγματικοί χώροι περιλαμβάνουν μη αριθμητικά αποτελέσματα. Για διευκόλυνση της μελέτης τέτοιων χώρων, εισάγουμε την έννοια της τυχαίας μεταβλητής.
- ΟΡΙΣΜΟΣ: Μια τ.μ. Χ είναι μια συνάρτηση που αντιστοιχεί σε κάθε σημείο ενός δειγματικού χώρου ένα πραγματικό αριθμό, δηλ. Χ: $\sigma \in \boldsymbol{S} \rightarrow X(\sigma) \in \mathbf{R}$.



Τυχαίες μεταβλητές – παραδείγματα

- Αν ***S*** = {Κ, Γ} μια τ.μ. θα μπορούσε να είναι μια συνάρτηση Χ με Χ(Κ) = 1 και Χ(Γ) = 0.
  (Μια άλλη τ.μ. θα μπορούσε να έχει Χ(Κ) = 0 και Χ(Γ) = 1).
- Στο πείραμα της ρίψης ενός ζαριού, Χ(σ) = σ.
- Στο πείραμα της ρίψης δύο ζαριών, οι τιμές Χ(σ) είναι το άθροισμα των δύο αποτελεσμάτων.
- Σε ένα τεστ αντιστοιχούμε σε κάθε φοιτητή σ έναν αριθμό Χ(σ) (π.χ. 0 – 10) που είναι το αποτέλεσμα του τεστ.

Διακριτές και συνεχείς τ.μ.

- Διακριτές τ.μ.: ορίζονται σε διακριτούς χώρους (πεπερασμένους ή άπειρους αριθμήσιμους)
  $S = \{\sigma_1, \sigma_2, \sigma_3, \ldots\} \rightarrow \{x_1, x_2, x_3, \ldots\}$
  - Συνάρτηση κατανομής: $F(x_i) = \{\sigma \in \boldsymbol{S} \ / \ X(\sigma) \leq x_i\}$
  - Συνάρτηση πιθανότητας: $p(x_i) = P(\{\sigma \in \boldsymbol{S} \ / \ X(\sigma) = x_i\})$
- Συνεχείς τ.μ.: ορίζονται σε συνεχείς χώρους
  - Συνάρτηση κατανομής: $F(x) = \{\sigma \in \boldsymbol{S} \ / \ X(\sigma) \leq x\}$
  - Πυκνότητα πιθανότητας $f(x) = dF \ / \ dx$

Ενδεχόμενα που ορίζονται από μια τ.μ.

Αν Χ είναι μια τ.μ., μας ενδιαφέρουν δύο είδη ενδεχομένων:

1. Το ενδεχόμενο που περιλαμβάνει τα στοιχεία του ***S*** που αντιστοιχούν στον ίδιο πραγματικό αριθμό α:

   $$\{\sigma \in \boldsymbol{S} \ / \ X(\sigma) = \alpha\}$$

   Αν $X(\sigma_1) = X(\sigma_2) = X(\sigma_3) = \alpha$,

   αντί $P(\{$αποτέλεσμα σ_1 ή σ_2 ή $\sigma_3\}) = P(\{\sigma \in \boldsymbol{S} \ / \ X(\sigma) = \alpha\})$ γράφουμε $P(X = \alpha)$.

   Αντίστοιχα αντί $\{\sigma \in \boldsymbol{S} \ / \ X(\sigma) = \alpha\}$ θεωρούμε το ενδεχόμενο $\{X = \alpha\}$.

Ενδεχόμενα που ορίζονται από μια τ.μ.

2. Το ενδεχόμενο που περιλαμβάνει τα στοιχεία του ***S*** που αντιστοιχούν σε αριθμούς μικρότερους ή ίσους του α:

   $$\{\sigma \in \boldsymbol{S} \;/\; X(\sigma) \leq \alpha\}$$

   Συμβολίζουμε το ενδεχόμενο αυτό με $\{X \leq \alpha\}$.

   Επίσης με $P(X \leq \alpha)$ δηλώνουμε την $P(\{\sigma \in \boldsymbol{S} \;/\; X(\sigma) \leq \alpha\})$.

- **Ερώτημα:**

  Πώς μας βοηθούν οι τυχαίες μεταβλητές στον υπολογισμό πιθανοτήτων;

Παράδειγμα

- Ρίχνουμε ένα νόμισμα 3 φορές. Ο δειγματικός χώρος είναι ***S*** = {ΚΚΚ, ΚΚΓ, ΚΓΚ, ΓΚΚ, ΚΓΓ, ΓΚΓ, ΓΓΚ, ΓΓΓ}.
- Ορίζουμε μια τ.μ. Χ που αντιστοιχεί σε κάθε σ$\in$***S*** το πλήθος των Κ στο αποτέλεσμα, δηλ. Χ(σ) = 0, 1, 2 ή 3.

$\sigma \in \boldsymbol{S}$	$X(\sigma)$	Γεγονότα $\{X = \alpha\}$	$P(X = \alpha) = p(a)$
ΚΚΚ	3	$\{X = 3\} = \{ΚΚΚ\}$	$P(X = 3) = 1/8$
ΚΚΓ ΚΓΚ ΓΚΚ	2 2 2	$\{X = 2\} = \{ΚΚΓ, ΚΓΚ, ΓΚΚ\}$	$P(X = 2) = 3/8$
ΚΓΓ ΓΚΓ ΓΓΚ	1 1 1	$\{X = 1\} = \{ΚΓΓ, ΓΚΓ, ΓΓΚ\}$	$P(X = 1) = 3/8$
ΓΓΓ	0	$\{X = 0\} = \{ΓΓΓ\}$	$P(X = 0) = 1/8$
		$\{X = \alpha\} = \varnothing$ για $\alpha \neq 0, 1, 2, 3$	$P(X = \alpha) = 0$ για $\alpha \neq 0, 1, 2, 3$

$\sigma \in \boldsymbol{S}$	$X(\sigma)$	Γεγονότα $\{X \leq \alpha\}$	$P(X \leq \alpha) = F_X(\alpha)$
ΚΚΚ	3	$\{X \leq 3\} = \boldsymbol{S}$	$P(X \leq 3) = 1$
ΚΚΓ ΚΓΚ ΓΚΚ	2 2 2	$\{X \leq 2\} =$ {ΚΚΓ, ΚΓΚ, ΓΚΚ, ΚΓΓ, ΓΚΓ, ΓΓΚ, ΓΓΓ}	$P(X \leq 2) = 7/8$
ΚΓΓ ΓΚΓ ΓΓΚ	1 1 1	$\{X \leq 1\} =$ {ΚΓΓ, ΓΚΓ, ΓΓΚ, ΓΓΓ}	$P(X \leq 1) = 1/2$
ΓΓΓ	0	$\{X \leq 0\} =$ {ΓΓΓ}	$P(X \leq 0) = 1/8$
		$\{X \leq \alpha\} = \varnothing$ για $\alpha < 0$	$P(X \leq \alpha) = 0$ για $\alpha < 0$

Μελέτη περίπτωσης: ελαστικά Feng Shui

- Τυχαία μεταβλητή L:

  Για κάθε ελαστικό σ στο δείγμα, L(σ) είναι η διάρκεια ζωής (σε χιλιάδες χιλιόμετρα).
- Δείγμα 100.000 ελαστικών:

  $\sigma_1 \rightarrow 10.6110$

  $\sigma_2 \rightarrow 39.8714$

  $\sigma_3 \rightarrow 27.8039$

  $\sigma_4 \rightarrow 47.2726$

  (...)

  $\sigma_N \rightarrow 36.8855$

Μελέτη περίπτωσης (2)

Από τη μελέτη του παραπάνω δείγματος προέκυψε ότι:

- Η διάρκεια ζωής L(σ) για κάθε σ κυμαίνεται μεταξύ 10 και 50 χιλιάδων km.
- Για x πραγματικό αριθμό με $10 \le x \le 50$, το ποσοστό των ελαστικών με διάρκεια ζωής $L \le x$ μπορεί να προσεγγιστεί ως εξής:

  $P(L \le x) = F(x) = 0.025\,(x-10)$ για $10 \le x \le 50$

  Προφανώς $F(x) = 0$ για $x < 10$

  και $F(x) = 1$ για $x > 50$

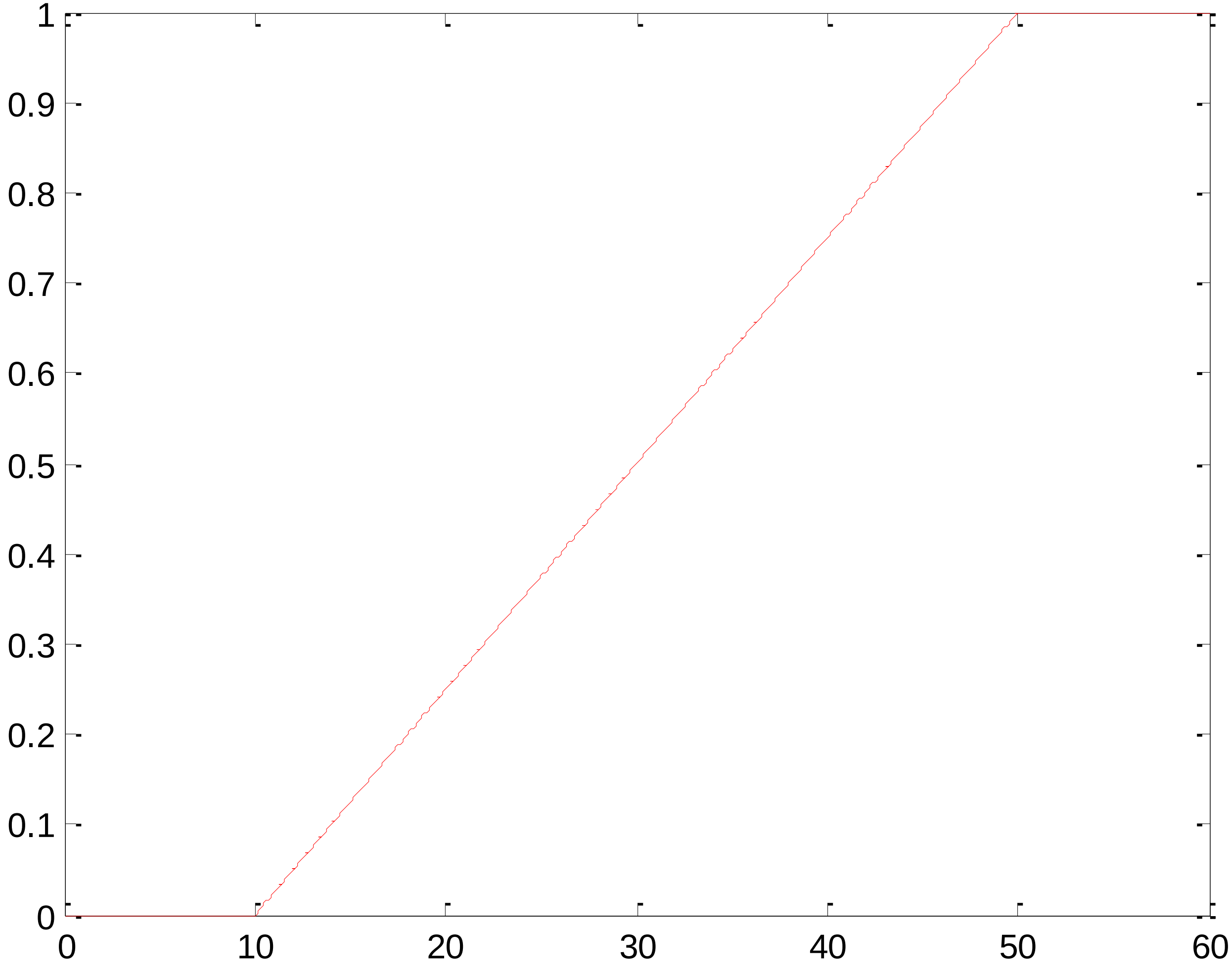
Μελέτη περίπτωσης (3)

Υπολογίστε τις εξής πιθανότητες:

- $P(L \leq 25) =$
- $P(L > 40) =$
- $P(30 < L \leq 40) =$

  $= P(L \leq 30) – P(L \leq 40) =$
- Η $F(x)$ γενικά λέγεται αθροιστική συνάρτηση κατανομής της τ.μ.
- Η συγκεκριμένη F λέγεται ομοιόμορφη.

Αθροιστική συνάρτηση κατανομής F(x)



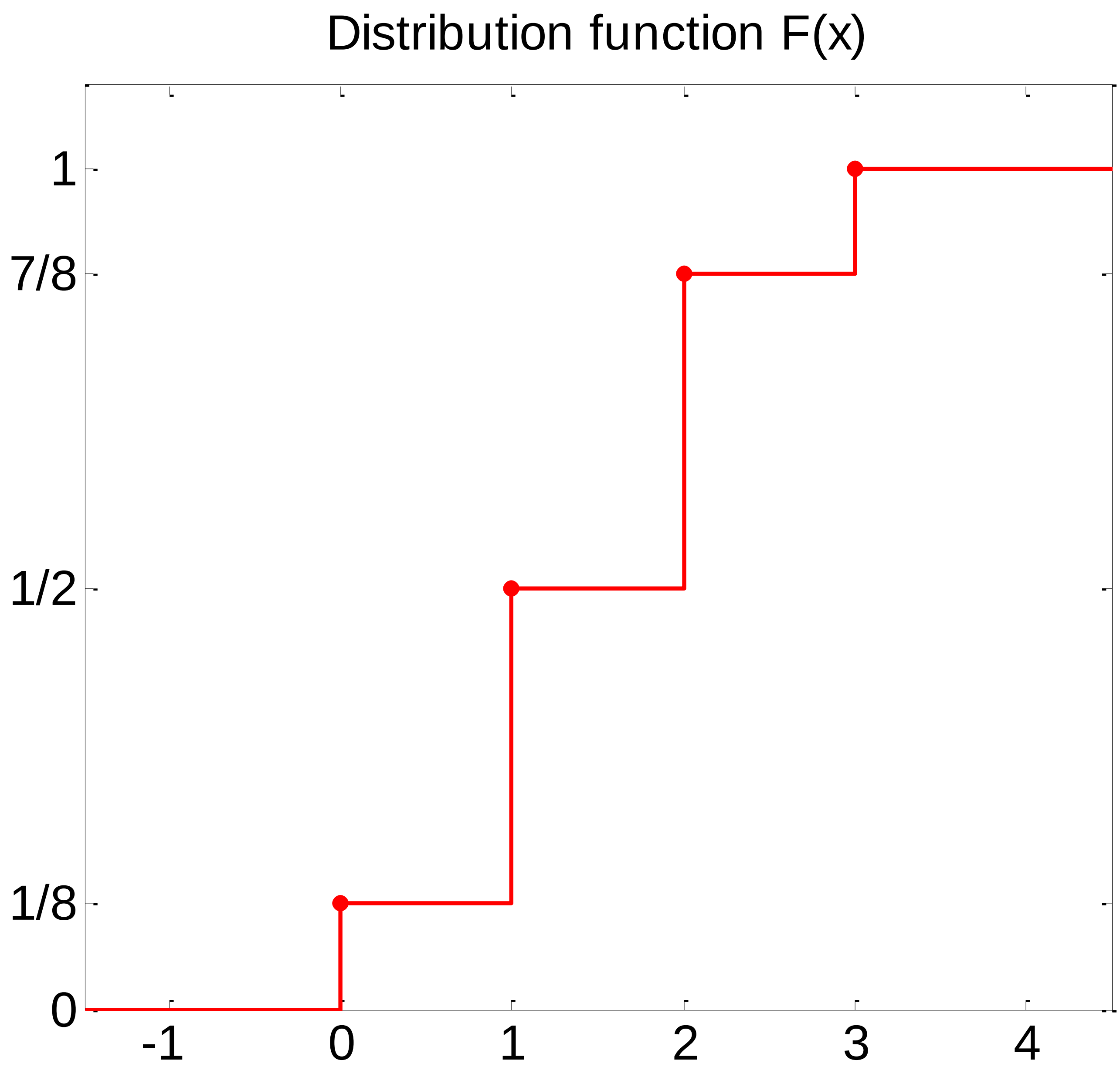
Συνάρτηση κατανομής μιας τ.μ.

- ΟΡΙΣΜΟΣ: Αθροιστική συνάρτηση κατανομής (cumulative distribution function, cdf) ή απλά συνάρτηση κατανομής μιας τ.μ. X είναι η εξής πραγματική συνάρτηση (**R**→**R**):

$$F(x) = P(X \leq x), \; -\infty < x < \infty$$

- Παράδειγμα: η σ.κ. του αριθμού των Κ σε τρεις ρίψεις είναι:

$$F(x) = \begin{cases} 1 & x \geq 3 \\ 7/8 & 2 \leq x < 3 \\ 1/2 & 1 \leq x < 2 \\ 1/8 & 0 \leq x < 1 \\ 0 & x < 0 \end{cases}$$


Distribution function F(x)
1
7/8
1/2
1/8
0
-1
0
1
2
3
4

Ιδιότητες της συνάρτησης κατανομής

$0 \leq F(x) \leq 1$ για κάθε $x \in \mathbf{R}$

$x_1 < x_2 \Rightarrow F(x_1) \leq F(x_2)$ (η F είναι αύξουσα συνάρτηση)

$$\lim_{x \to +\infty} F(x) = 1$$

$$\lim_{x \to -\infty} F(x) = 0$$

Ιδιότητες της συνάρτησης κατανομής (2)

$\lim_{x \to x_0+} F(x) = F(x_0)$ (η F είναι συνεχής από δεξιά)

- Η F δεν είναι απαραίτητα συνεχής από αριστερά (όπως στο παράδειγμα των 3 ρίψεων του νομίσματος).
- Παρατήρηση: αν μια συνάρτηση έχει τις προηγούμενες ιδιότητες, είναι σ.κ. κάποιας τ.μ.

Υπολογισμός πιθανοτήτων από την σ.κ.

- $P(X \leq \alpha) = F(\alpha)$ (από τον ορισμό)
- $P(X > \beta) = 1 – P(X \leq \beta) = 1 – F(\beta)$
- $P(\alpha < X \leq \beta) = P(X \leq \beta) – P(X \leq \alpha) = F(\beta) – F(\alpha)$
- $P(X = \alpha) = F(\alpha) – F(\alpha–)$

  (Με $F(\alpha–)$ εννούμε το όριο από αριστερά)

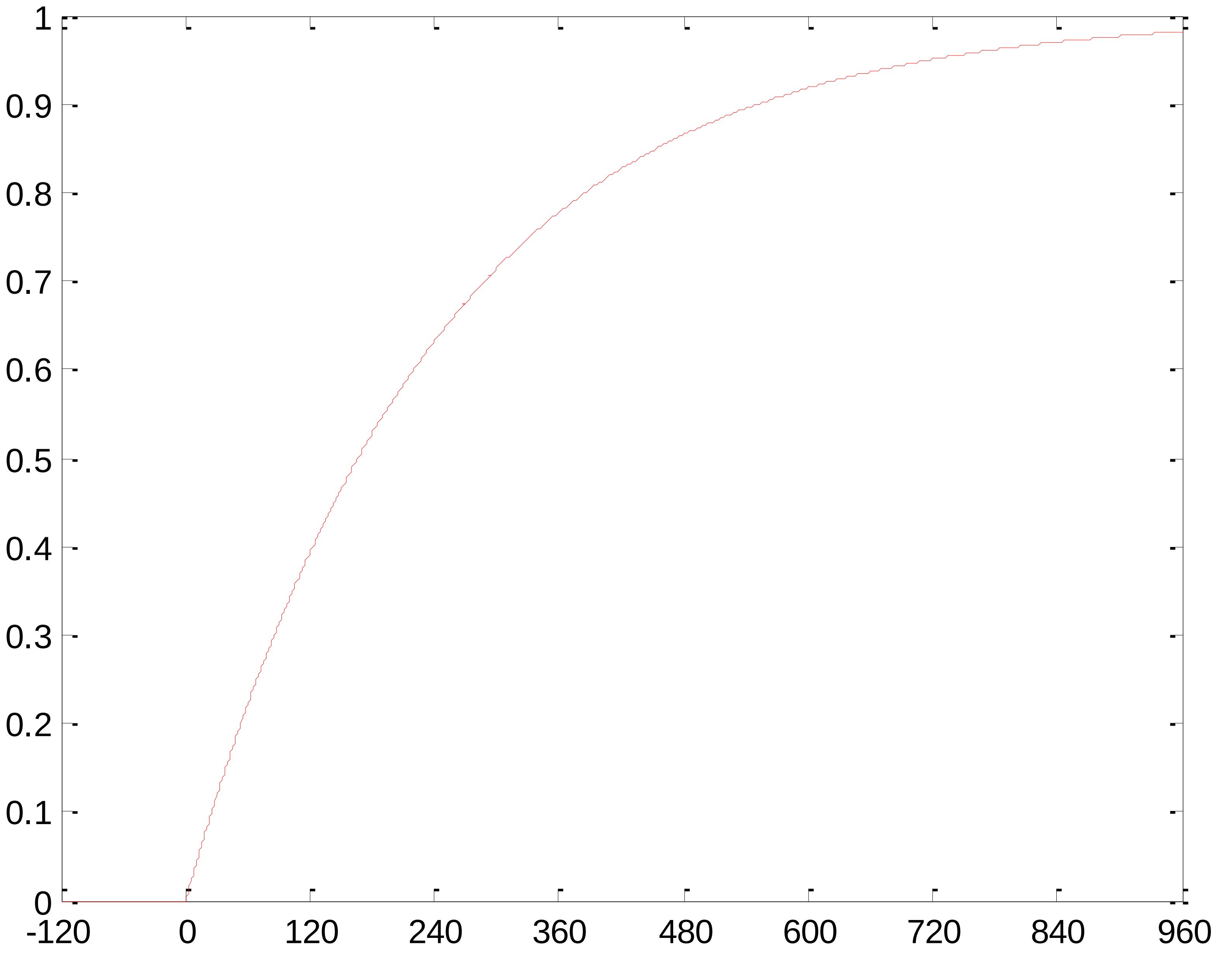
Με βάση τα παραπάνω μπορούμε να υπολογίσουμε την πιθανότητα οποιουδήποτε ενδεχομένου από την σ.κ., χωρίς να χρησιμοποιήσουμε τον δειγματικό χώρο.

Παράδειγμα

- Ο χρόνος μέχρι την επόμενη εμφάνιση βλάβης σε ένα σύστημα περιγράφεται από μια τ.μ. X με την εξής σ.κ.:

$$F(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-\lambda x} & x \geq 0 \end{cases}$$

  όπου το x εκφράζεται σε μέρες και $\lambda = 1 / 240$.
- Επαληθεύουμε εύκολα ότι η F έχει τις γνωστές ιδιότητες και άρα είναι σ.κ.
- Η πιθανότητα να έχουμε βλάβη μέσα στις επόμενες 120 μέρες είναι $P(X \leq 120) = F(120) = 1 - e^{-0.5} = 0.39$.
- Η πιθανότητα να μην έχουμε βλάβη τις επόμενες 480 μέρες είναι $P(X > 480) = 1 - F(480) = e^{-2} = 0.14$.



1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-120
0
120
240
360
480
600
720
840
960