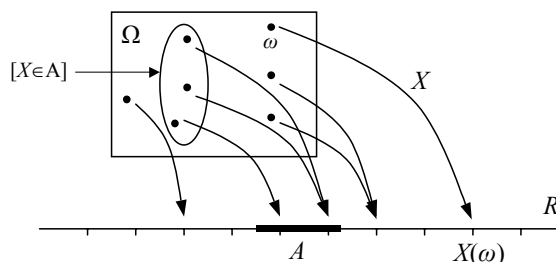


0. Σύντομη επισκόπηση θεωρίας πιθανοτήτων

A. Τυχαίες μεταβλητές

Τυχαία μεταβλητή καλείται μια μεταβλητή η τιμή της οποίας καθορίζεται από το αποτέλεσμα κάποιου στοχαστικού πειράματος. Αν Ω ο δειγματικός χώρος του στοχαστικού πειράματος, τότε η τιμή μιας τυχαίας μεταβλητής X καθορίζεται από το στοιχειώδες ενδεχόμενο $\omega \in \Omega$ που πραγματοποιήθηκε, δηλ. $X = X(\omega)$. Αυστηρότερα έχουμε τον επόμενο ορισμό.

Ορισμός. Κάθε απεικόνιση X από το δειγματικό χώρο Ω (σύνολο των δυνατών αποτελεσμάτων ενός πειράματος τύχης) σε κάποιο υποσύνολο των πραγματικών αριθμών R θα καλείται **τυχαία μεταβλητή (τ.μ.)**.



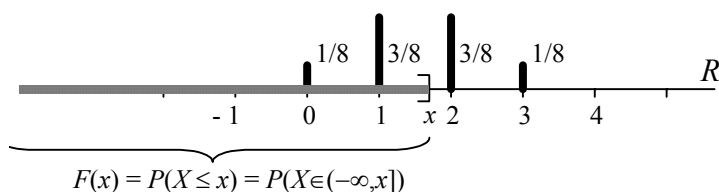
Η απεικόνιση αυτή μπορεί να θεωρηθεί ότι «μεταφέρει» πιθανότητα από τον Ω στο R . Για παράδειγμα στο $A \subseteq R$ μεταφέρεται η πιθανότητα των σημείων του Ω που μέσω της τ.μ. απεικονίζονται μέσα στο A (δηλ. των σημείων $\{\omega \in \Omega: X(\omega) \in A\}$). Δηλαδή,

$$P(X \in A) = P(\{\omega \in \Omega: X(\omega) \in A\}).$$

Με αυτό τον τρόπο, η συνολική πιθανότητα (1 ή 100%) «κατανέμεται» στα στοιχεία του συνόλου τιμών της τ.μ. X . Συνήθως δεν μας ενδιαφέρει από ποια στοιχεία του Ω έχει προέλθει η πιθανότητα που έχει κατανομηθεί στα διάφορα σημεία του R , αλλά μόνο η «κατανομή» της πιθανότητας στον R . Μία περιγραφή της «κατανομής» αυτής γίνεται μέσω της συνάρτησης κατανομής.

Ορισμός. Η συνάρτηση $F_X: R \rightarrow [0,1]$ με $F_X(x) = P(X \leq x)$, $x \in R$ καλείται (αθροιστική) συνάρτηση κατανομής (σ.κ.) της τ.μ. X .

Για παράδειγμα:



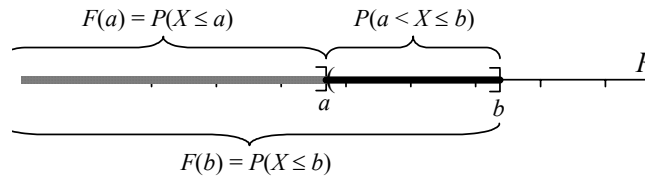
Υπενθυμίζουμε ορισμένες ιδιότητες της συνάρτησης κατανομής F μιας τ.μ.,

- 1) $0 \leq F(x) \leq 1$ για κάθε $x \in R$.
- 2) $\lim_{x \rightarrow \infty} F(x) = F(\infty) = 1$, $\lim_{x \rightarrow -\infty} F(x) = F(-\infty) = 0$,
- 3) Η F είναι μη φθίνουσα συνάρτηση, δηλαδή για κάθε $x > y$ ισχύει ότι $F(x) \geq F(y)$.

Επίσης, αν X είναι μία τ.μ. και $a < b \in R$ τότε,

$$P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a).$$

Σχηματικά:



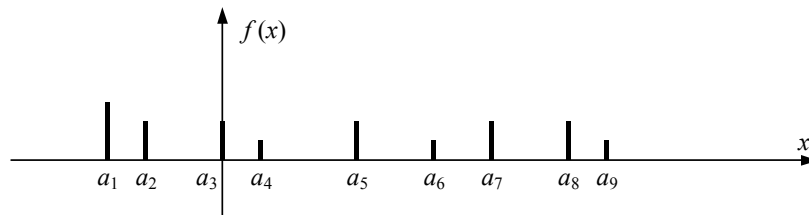
B. Διακριτές τυχαίες μεταβλητές

Διακριτές καλούνται οι τ.μ. που έχουν ως πεδίο τιμών κάποιο υποσύνολο του $\mathbb{Z}$ ή του $\mathbb{N}$ ή γενικότερα παίρνουν αριθμήσιμο πλήθος τιμών (αριθμήσιμο καλείται ένα σύνολο όταν μπορεί να γραφεί στην μορφή $\{a_1, a_2, \dots, a_n\}$ ή $\{a_1, a_2, \dots\}$).

Ορισμός. Έστω μία διακριτή τ.μ. με τιμές στο αριθμήσιμο $B = X(\Omega) \subseteq \mathbb{R}$. Η συνάρτηση $f_X: B \rightarrow [0, 1]$,

$$f_X(x) = P(X = x), \quad x \in B$$

θα καλείται συνάρτηση πιθανότητας (σ.π.) της τ.μ. X .



Αν π.χ. $X \in \{0, 1, 2, \dots\} = B$ τότε

$$\sum_{x=0}^{\infty} f_X(x) = 1, \quad F_X(x) = P(X \leq x) = \sum_{i=0}^x P(X = i) = \sum_{i=0}^x f_X(i), \quad x = 0, 1, \dots$$

$$f_X(x) = P(X = x) = F_X(x) - F_X(x-1), \quad x = 1, 2, 3, \dots, \quad f_X(0) = F_X(0).$$

Γ. Συνεχείς τυχαίες μεταβλητές

Συνεχείς καλούνται οι τ.μ. που έχουν ως πεδίο τιμών ένα διάστημα του $\mathbb{R}$, ή όλο το $\mathbb{R}$ ενώ επιπλέον έχουν παραγωγίσιμες συναρτήσεις κατανομής.

Ορισμός. Μία τ.μ. X (ή αντίστοιχα η κατανομή της) θα καλείται συνεχής αν υπάρχει μία μη αρνητική συνάρτηση f_X (δηλ. $f_X \geq 0$), ώστε

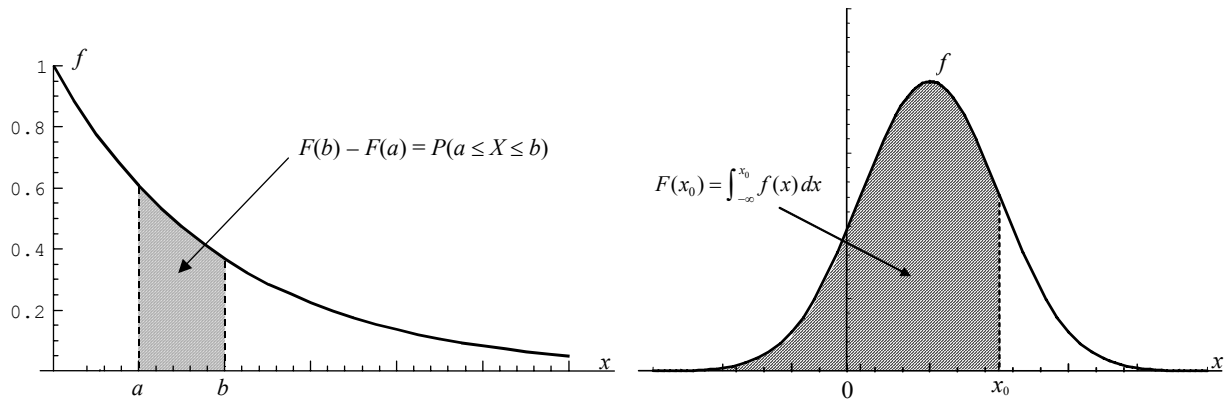
$$F_X(b) - F_X(a) = \int_a^b f_X(x) dx, \quad a, b \in \mathbb{R}$$

Η συνάρτηση f_X θα καλείται συνάρτηση πυκνότητας πιθανότητας (ή πυκνότητα) της συνεχούς τ.μ. X .

Υπενθυμίζεται επίσης ότι

$$\int_{-\infty}^{\infty} f_X(x) dx = 1 \quad \text{και} \quad F_X(x) = \int_{-\infty}^x f_X(t) dt, \quad \frac{d}{dx} F_X(x) = f_X(x).$$

Στην περίπτωση των συνεχών τ.μ., η συνολική πιθανότητα δεν κατανέμεται πάνω σε ένα αριθμήσιμο υποσύνολο $\{a_1, a_2, \dots\}$ του $\mathbb{R}$ (όπως στις διακριτές τ.μ.) αλλά «απλώνεται» με συνεχή τρόπο πάνω σε όλα τα σημεία ενός διαστήματος του $\mathbb{R}$ (ή ακόμη και σε ολόκληρο το $\mathbb{R}$) και η σ.π. f δείχνει την «πυκνότητα» με την οποία έχει κατανεμηθεί («απλωθεί») η συνολική πιθανότητα στον $\mathbb{R}$.



Δ. Παράμετροι θέσης και κλίμακας μιας κατανομής

Εισάγουμε δύο παραμέτρους που προσφέρουν σημαντική πληροφορία για τη μορφή της κατανομής μιας τ.μ. Την *μέση τιμή* η οποία προσδιορίζει το «βαρύκεντρο» της κατανομής (η τ.μ. παίρνει τιμές «γύρω» από αυτήν) και τη *διακύμανση* η οποία εκφράζει τη «μεταβλητότητα» της τ.μ. γύρω από τη μέση τιμή.

- Μέση τιμή

Μέση τιμή μιας τ.μ. X καλείται η ποσότητα

$$\mu_X = E(X) = \sum_{x=0}^{\infty} x f_X(x) \text{ (διακριτές τ.μ. στο } N), \quad \mu_X = E(X) = \int_{-\infty}^{\infty} x f_X(x) dx \text{ (συνεχείς τ.μ.)}$$

Γενικότερα αποδεικνύεται ότι η μέση τιμή μιας τ.μ. $Y = g(X)$ θα είναι

$$E(g(X)) = \sum_{x=0}^{\infty} g(x) f_X(x) \text{ (διακριτές τ.μ. στο } N), \quad E(g(X)) = \int_{-\infty}^{\infty} g(x) f_X(x) dx \text{ (συνεχείς τ.μ.)}$$

- Διασπορά

Η ποσότητα (και για διακριτές και για συνεχείς τ.μ.)

$$\sigma_X^2 = V(X) = E((X - \mu_X)^2) = E(X^2) - E(X)^2$$

καλείται διασπορά ή διακύμανση της τ.μ. X , ενώ η $\sigma_X = \sqrt{V(X)}$ καλείται τυπική απόκλιση της τ.μ. X .

- Ιδιότητες μέσης τιμής: α) $E(c) = c$, β) $E(aX + b) = aE(X) + b$

- Ιδιότητες διασποράς: α) $V(c) = 0$, β) $V(aX + b) = a^2 V(X)$,

Ε. Συνήθεις διακριτές κατανομές

Ε.1. Η ομοιόμορφη διακριτή κατανομή.

Ορισμός. Η διακριτή κατανομή με σ.π.

$$f(a_i) = \frac{1}{n}, i = 1, 2, \dots, n$$

καλείται διακριτή ομοιόμορφη κατανομή στο σύνολο $A = \{a_1, a_2, \dots, a_n\}$ (συμβολίζεται και με $DU(A)$).



Αν η τ.μ. X εκφράζει έναν «τυχαία» επιλεγμένο αριθμό από ένα πεπερασμένο σύνολο $A=\{a_1, a_2, \dots, a_n\}$ (δηλ. κάθε στοιχείο του A έχει την ίδια πιθανότητα εκλογής) τότε $X \sim DU(A)$. Αν π.χ. η τ.μ. $X \sim DU(A)$, $A = \{1, 2, \dots, n\}$ τότε

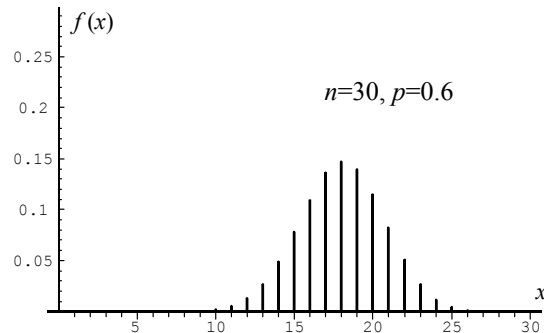
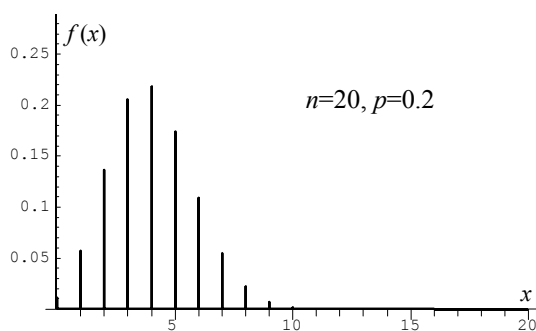
$$E(X) = \frac{n+1}{2}, \quad V(X) = \frac{n^2-1}{12}.$$

Ε.2. Η Διωνυμική κατανομή.

Ορισμός. Η διακριτή κατανομή με σ.π.

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \dots, n$$

καλείται διωνυμική κατανομή με παραμέτρους n, p (συμβολίζεται και με $B(n, p)$).



Υπενθυμίζεται ότι αν μία τ.μ. X εκφράζει το πλήθος των επιτυχιών σε n ανεξάρτητες δοκιμές σε κάθε μία από τις οποίες εμφανίζεται επιτυχία με πιθανότητα p και αποτυχία με πιθανότητα $1-p$ τότε $X \sim B(n, p)$. Αν $X \sim B(n, p)$ τότε

$$E(X) = np, \quad V(X) = np(1-p).$$

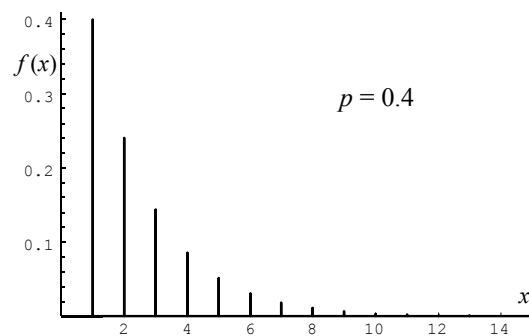
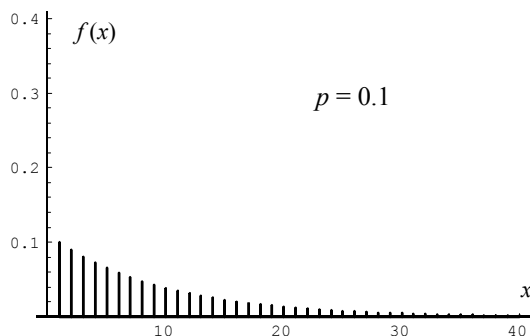
Για $n=1$ η κατανομή $B(1, p)$ είναι γνωστή και ως κατανομή Bernoulli. Αν $X \sim B(1, p)$ τότε $X \in \{0, 1\}$ και $P(X=0) = 1-p$ και $P(X=1) = p$ ενώ $E(X) = p$, $V(X) = p(1-p)$.

Ε.3. Η Γεωμετρική κατανομή

Ορισμός. Η διακριτή κατανομή με σ.π.

$$f(x) = (1-p)^{x-1} p, \quad x = 1, 2, \dots$$

καλείται γεωμετρική κατανομή με παράμετρο p (συμβολίζεται και με $Ge(p)$).



Υπενθυμίζεται ότι αν μία τ.μ. X μπορεί να θεωρηθεί ότι εκφράζει το πλήθος των δοκιμών μέχρι και την εμφάνιση της πρώτης επιτυχίας σε μία ακολουθία ανεξάρτητων δοκιμών με πιθανότητα επιτυχίας p τότε $X \sim Ge(p)$. Επίσης, αν $X \sim Ge(p)$ τότε

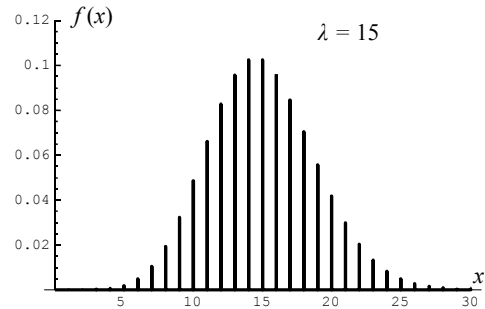
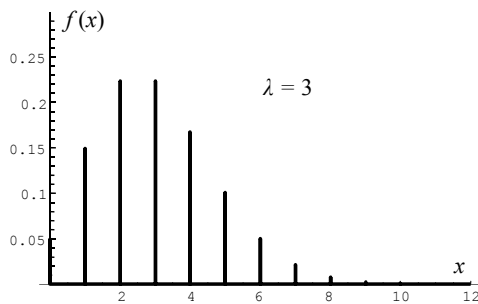
$$E(X) = \frac{1}{p}, \quad V(X) = \frac{1-p}{p^2}.$$

Ε.4. Η κατανομή Poisson.

Ορισμός. Η διακριτή κατανομή με σ.π.

$$f(x) = e^{-\lambda} \frac{\lambda^x}{x!}, \quad x = 0, 1, 2, \dots$$

καλείται κατανομή Poisson με παράμετρο λ (συμβολίζεται και με $Po(\lambda)$).



Υπενθυμίζεται ότι αν $n \rightarrow \infty$, $p = \lambda / n$ τότε η διωνυμική κατανομή $B(n, p)$ «συγκλίνει» στην κατανομή Poisson $Po(\lambda) = Po(np)$, δηλαδή,

$$\binom{n}{k} p^k (1-p)^{n-k} \rightarrow e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \dots$$

Με άλλα λόγια, αν μία τ.μ. X μπορεί να θεωρηθεί ότι εκφράζει το πλήθος των επιτυχιών σε ένα μεγάλο αριθμό δοκιμών ($n \rightarrow \infty$) με πιθανότητα επιτυχίας $p \rightarrow 0$ σε κάθε δοκιμή τότε προσεγγιστικά $X \sim Po(np)$. Αν $X \sim Po(\lambda)$ τότε

$$E(X) = \lambda, \quad V(X) = \lambda.$$

Ε.5. Η Υπεργεωμετρική κατανομή.

Ορισμός. Η διακριτή κατανομή με σ.π.

$$f(x) = \frac{\binom{A}{x} \binom{N-A}{n-x}}{\binom{N}{n}}, \quad x = \max\{0, n + A - N\}, \dots, \min\{A, n\}$$

καλείται υπεργεωμετρική κατανομή με παραμέτρους A, N, n ($A, n < N$). (συμβ. και με $h(A, N, n)$).

Υπενθυμίζεται ότι αν μία τ.μ. X εκφράζει το πλήθος των λευκών σφαιρών ανάμεσα σε n τυχαία επιλεγμένες σφαίρες (χωρίς επανάθεση) από μία κάλπη με A λευκές και $N-A$ μαύρες τότε η τ.μ. X θα ακολουθεί την υπεργεωμετρική κατανομή με παραμέτρους A, N, n . (υπενθυμίζεται ότι αν η επιλογή των n σφαιρών γίνει με επανάθεση, τότε η τ.μ. $X \sim B(n, A/N)$)

Αν μία τ.μ. $X \sim h(A, N, n)$ τότε

$$E(X) = n \frac{A}{N} \quad \text{και} \quad V(X) = n \frac{A}{N} \left(1 - \frac{A}{N}\right) \frac{N-n}{N-1}.$$

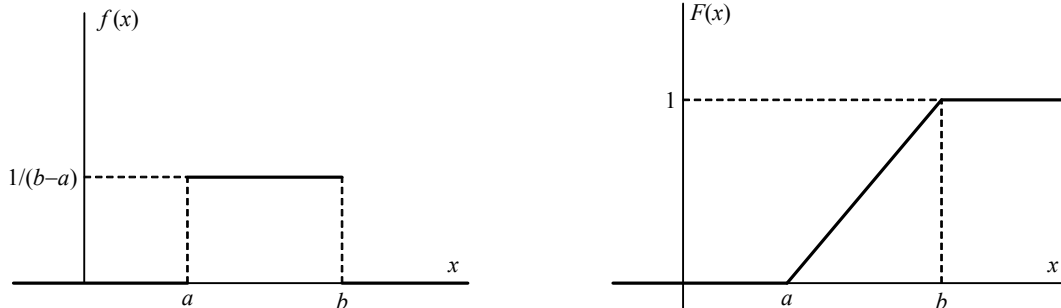
ΣΤ. Συνήθειες συνεχείς κατανομές

ΣΤ.1. Η ομοιόμορφη συνεχής κατανομή.

Ορισμός. Η συνεχής κατανομή με σ.π.π.

$$f(x) = \frac{1}{b-a}, \quad x \in [a, b], \quad f(x) = 0, \quad x \notin [a, b]$$

καλείται συνεχής ομοιόμορφη κατανομή στο διάστημα $[a, b]$ (συμβολίζεται και με $U(a, b)$).



Αν X εκφράζει έναν «τυχαία» επιλεγμένο αριθμό από ένα διάστημα $[a, b]$ (κάθε «σημείο» ή καλύτερα κάθε «απειροστό διάστημα» του $[a, b]$ έχει την ίδια πιθανότητα εκλογής) τότε $X \sim U(a, b)$. Αν $X \sim U(a, b)$ τότε

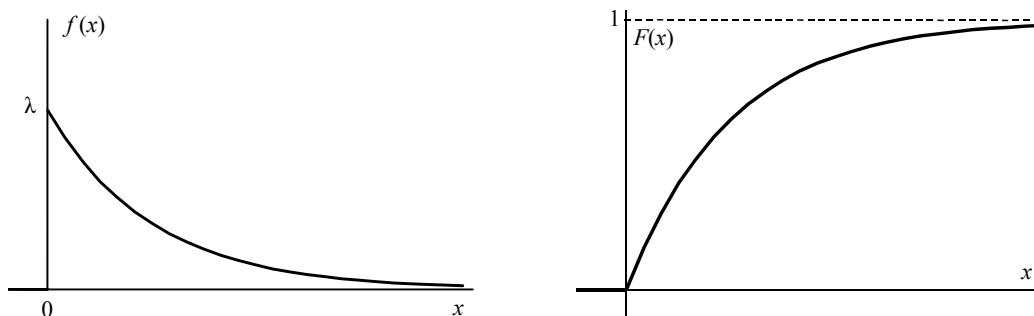
$$E(X) = \frac{a+b}{2}, \quad V(X) = \frac{(b-a)^2}{12}.$$

ΣΤ.2. Η εκθετική κατανομή.

Ορισμός. Η συνεχής κατανομή με σ.π.π.

$$f(x) = \lambda e^{-\lambda x}, \quad x \geq 0 \quad (\lambda > 0)$$

($f(x) = 0, x < 0$) καλείται εκθετική κατανομή με παράμετρο λ (συμβολίζεται και με $Exp(\lambda)$).



Η εκθετική κατανομή εμφανίζεται συνήθως σε περιπτώσεις όπου μελετάμε το χρόνο αναμονής μέχρι την πραγματοποίηση ενός γεγονότος. Πιο συγκεκριμένα, έστω ότι κάποια γεγονότα πραγματοποιούνται σε τυχαίες στιγμές στο χρόνο, δηλαδή σε κάθε πολύ μικρό χρονικό διάστημα μήκους h , η πραγματοποίηση ενός γεγονότος συμβαίνει με πιθανότητα περίπου λh ενώ είναι ανεξάρτητη από το τι έχει συμβεί στα υπόλοιπα διαστήματα. Αν X ο χρόνος αναμονής μέχρι την εμφάνιση του πρώτου τέτοιου γεγονότος τότε αποδεικνύεται ότι $X \sim Exp(\lambda)$.

Η συνάρτηση κατανομής της εκθετικής κατανομής είναι $F(x) = 1 - e^{-\lambda x}$, $x \geq 0$ ενώ $F(x) = 0$ αν $x < 0$. Η μέση τιμή και διασπορά της εκθετικής κατανομής είναι

$$E(X) = \frac{1}{\lambda} \quad \text{και} \quad V(X) = \frac{1}{\lambda^2}.$$