Στατιστική Ι

Γιώργος Τσιρογιάννης

Τμήμα Διοίκησης Επιχειρήσεων Αγροτικών Προϊόντων και Τροφίμων,

Πανεπιστήμιο Πατρών

Ακαδημαϊκό έτος: 2020-2021

Διάλεξη 7η

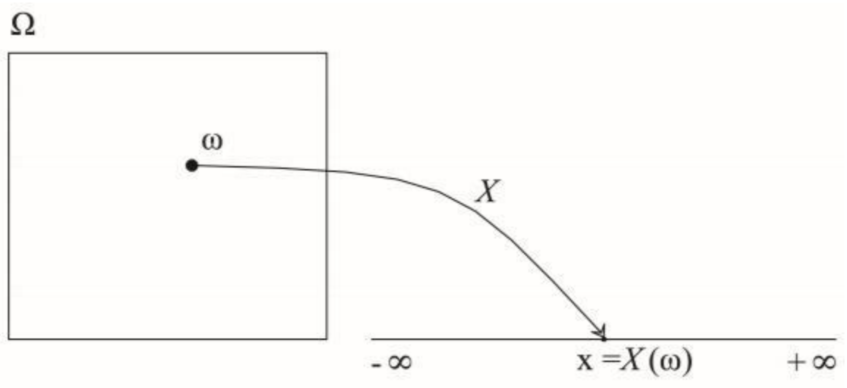
Συνεχείς τυχαίες μεταβλητές

Συνάρτηση κατανομής

Παράμετροι διακριτών τμ

Ανεξαρτησία τμ

5.4 – 5.5



Τυχαίες μεταβλητές

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

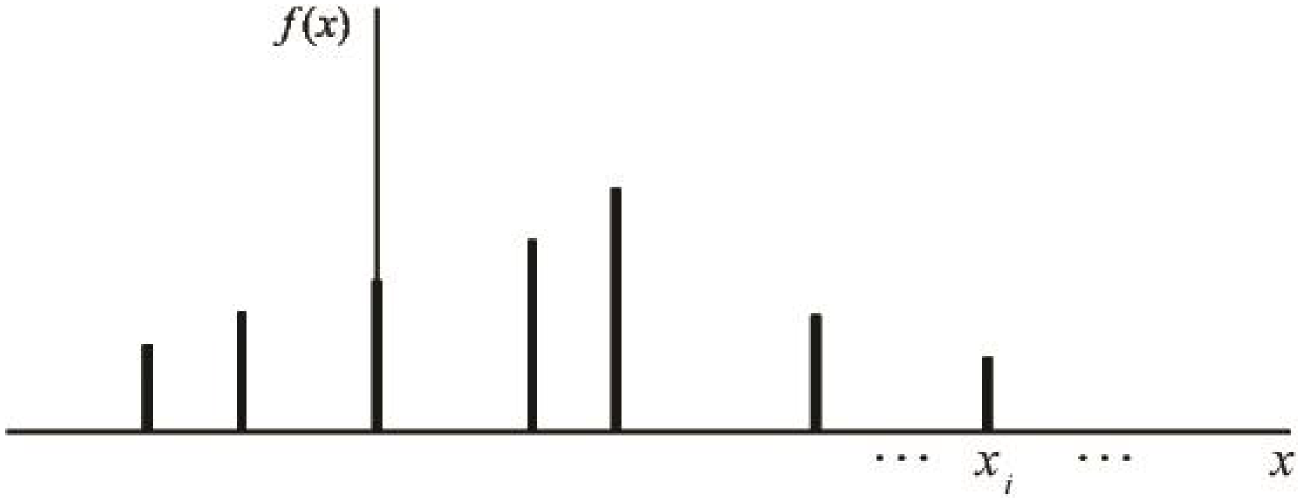
Συνάρτηση (μάζας) πιθανότητας (διακριτές τμ)

- Έστω X διακριτή τυχαία μεταβλητή με σύνολο τιμών R_X. Η πραγματική συνάρτηση f με

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

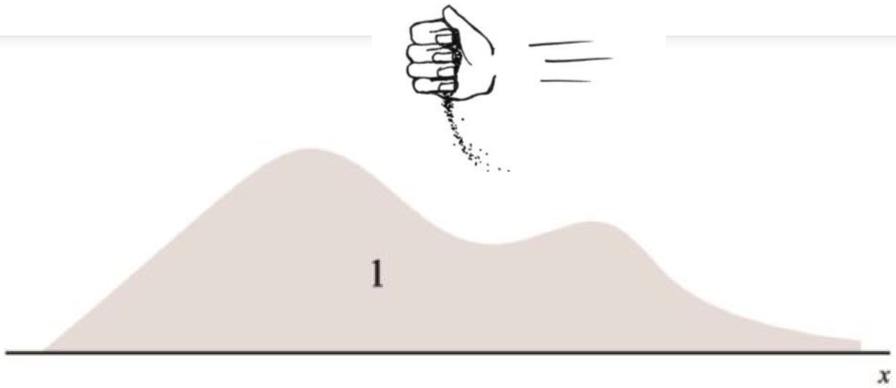
ονομάζεται συνάρτηση πιθανότητας ή συνάρτηση μάζας πιθανότητας της X (probability mass function).

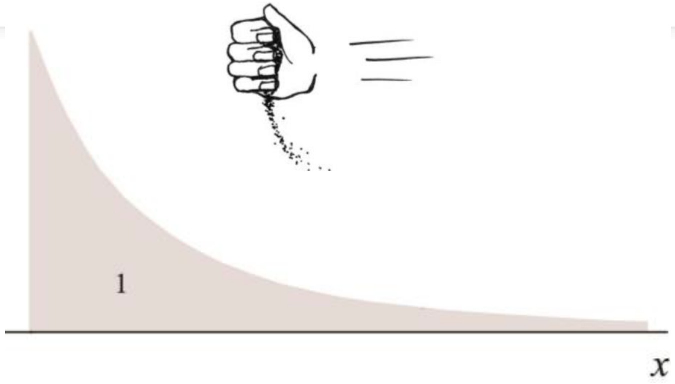
Συνάρτηση (μάζας) πιθανότητας (διακριτές τμ)

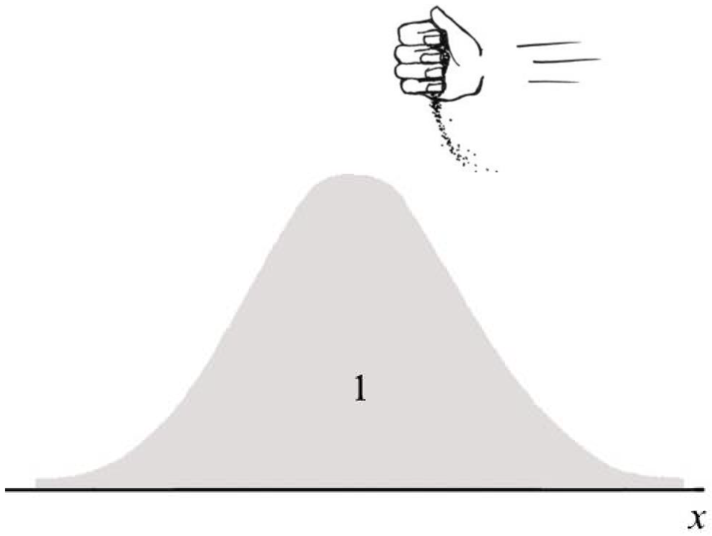


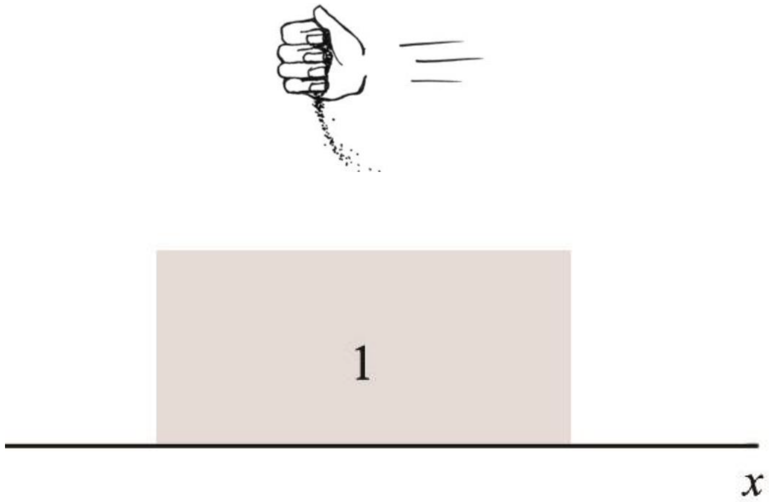
Συνεχείς τυχαίες μεταβλητές

Παραδείγματα εκχώρησης/κατανομής της συνολικής πιθανότητας σε συνεχές διάστημα τιμών

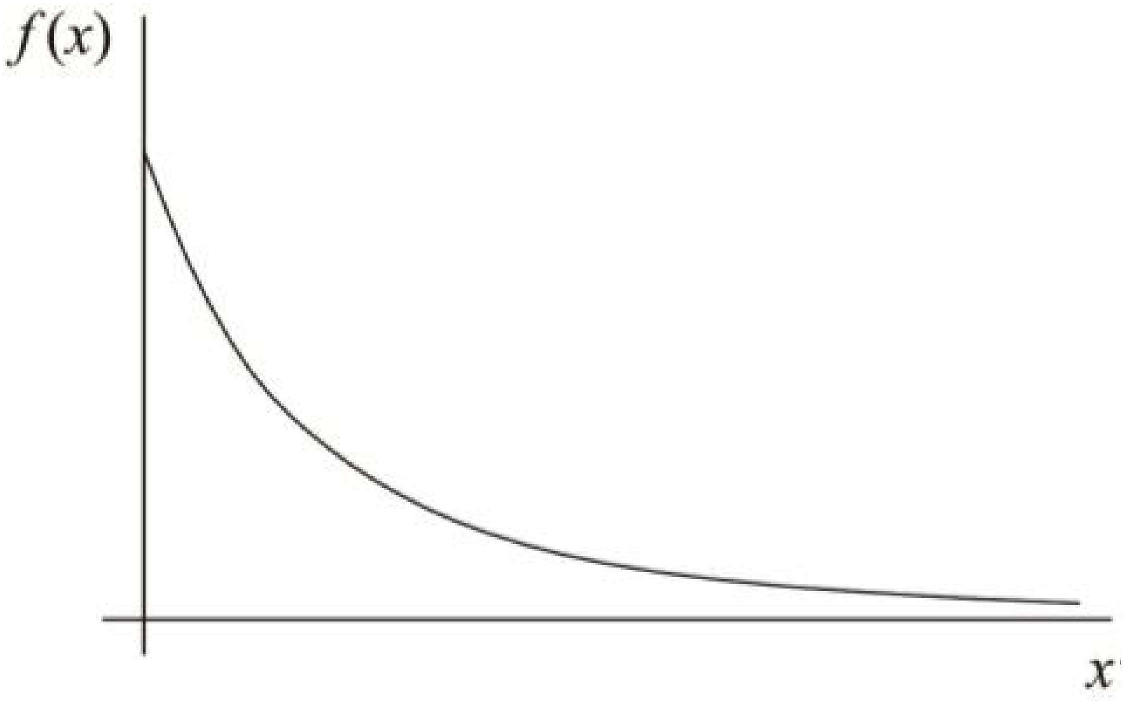




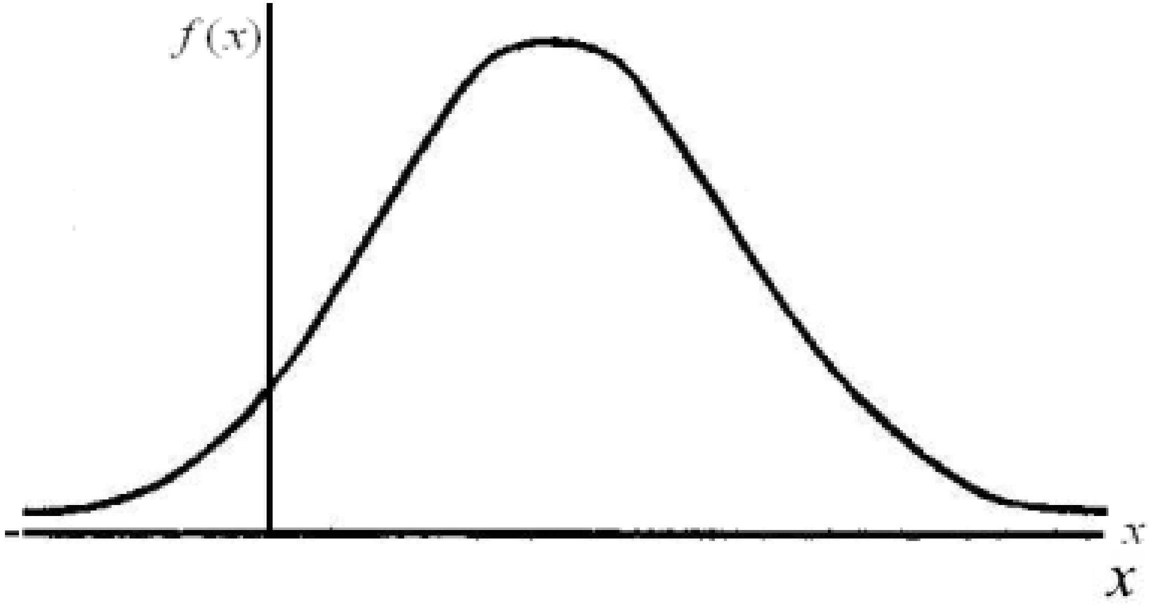




Παράδειγμα συνάρτησης πυκνότητας πιθανότητας



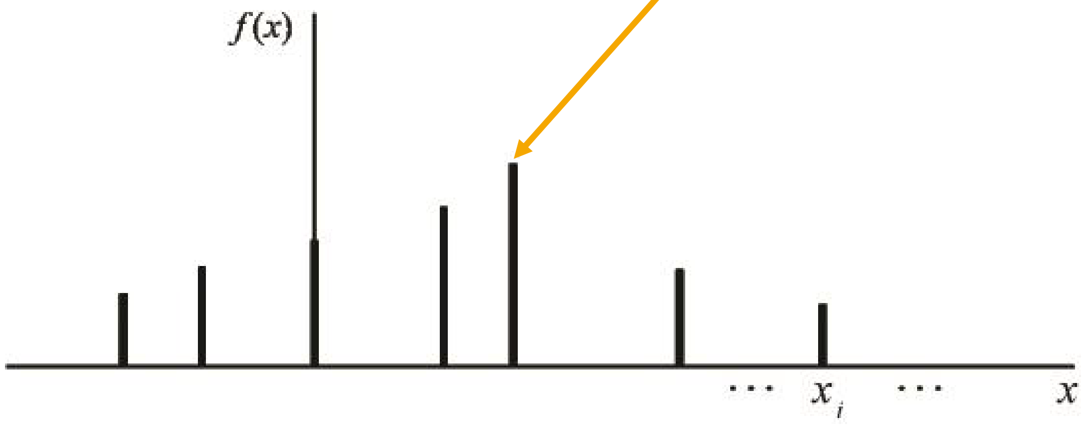
Παράδειγμα συνάρτησης πυκνότητας πιθανότητας



Συνάρτηση πιθανότητας: διακριτές vs συνέχεις τμ

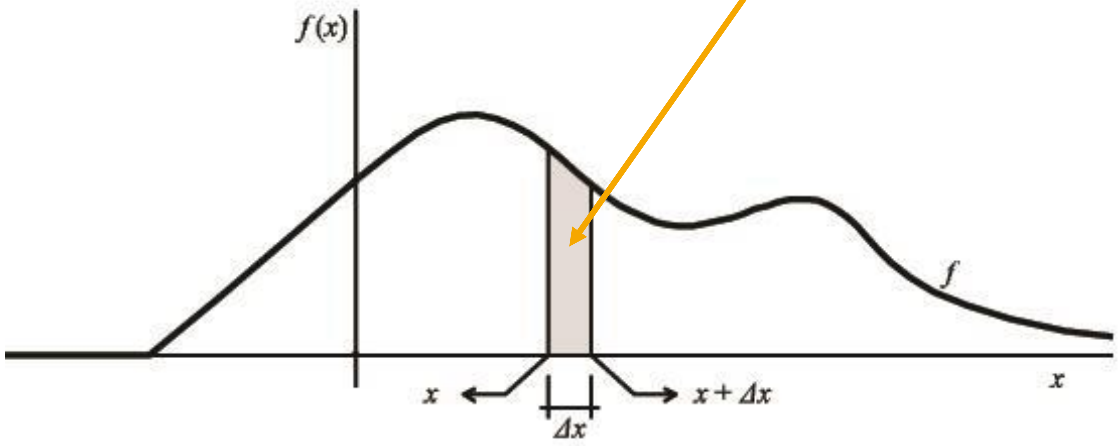
διακριτή

Αντιστοιχεί σε τιμή



συνεχής

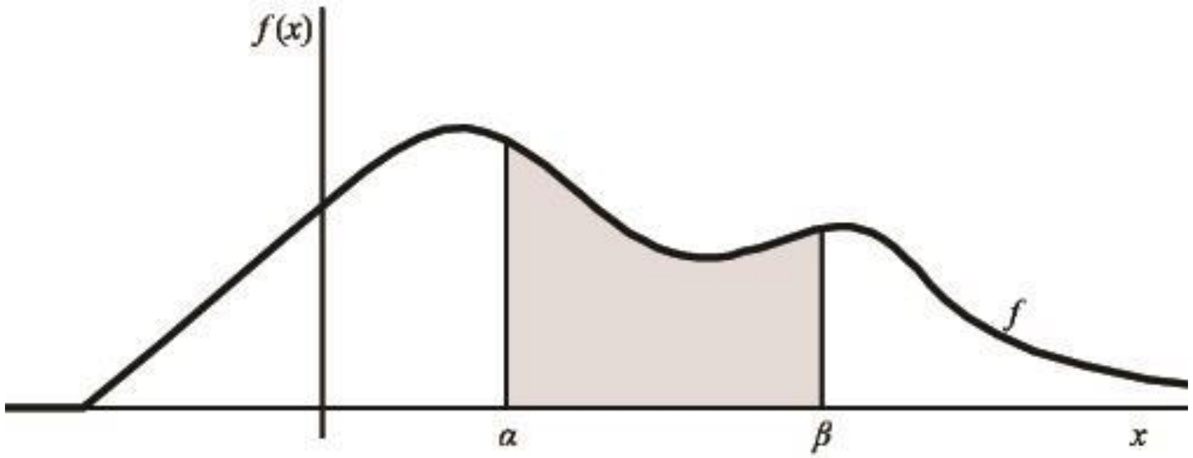
Αντιστοιχεί σε εμβαδό



$$P(x \leq X < x + \Delta x)$$

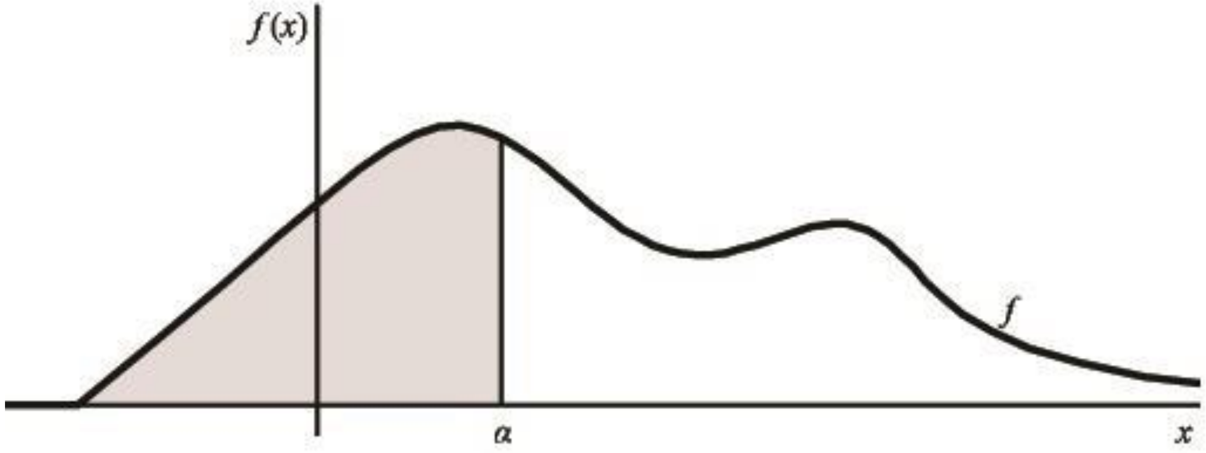
Πιθανότητες ενδεχομένων

$X \in [\alpha, \beta]$



$P(\alpha \le X \le \beta)$

$X \in (-\infty, \alpha]$



$P(X \le \alpha)$

Συνεχής τυχαία μεταβλητή και συνάρτηση πυκνότητας πιθανότητας

- Μια τυχαία μεταβλητή X λέγεται συνεχής αν υπάρχει μια μη αρνητική πραγματική συνάρτηση

$$f : R \to [0, +\infty)$$

τέτοια ώστε για κάθε υποσύνολο A των πραγματικών να ισχύει

$$P(X \in A) = \int_A f(x)dx$$

- Η συνάρτηση f ονομάζεται **συνάρτηση πυκνότητας πιθανότητας** ή απλά συνάρτηση πυκνότητας της X (probability density function, pdf).

Πρόταση

- Αν X συνεχής τυχαία μεταβλητή με συνάρτηση κατανομής F και συνάρτηση πυκνότητας f, τότε:

$$\alpha)\ \int_{-\infty}^{+\infty} f(x)dx = 1$$

$$\beta)\ F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt$$

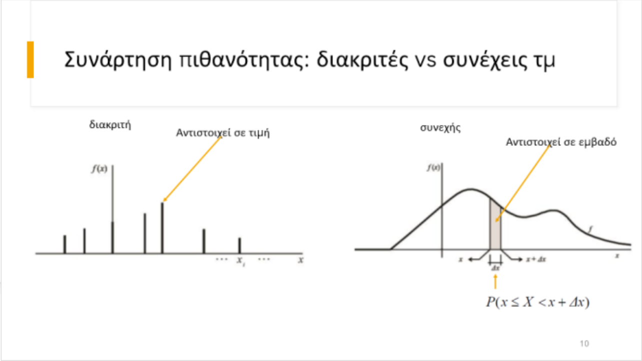
$$\gamma)\ f(x) = F'(x)$$ (στα σημεία συνέχειας της f)

$\delta)$ *Για οποιουσδήποτε πραγματικούς* α, β *με* $\alpha \leq \beta$

$$(i)\quad P(\alpha \leq X \leq \beta) = \int_{\alpha}^{\beta} f(x)dx = F(\beta) - F(\alpha).$$

$$(ii)\ P(\alpha < X < \beta) = P(\alpha \leq X < \beta) = P(\alpha < X \leq \beta) = P(\alpha \leq X \leq \beta)$$

Σχόλια



- Η πιθανότητα μια συνεχής τυχαία μεταβλητή Χ να πάρει μια συγκεκριμένη τιμή α , είναι ίση μηδέν για οποιοδήποτε α.
- Αυτό δεν σημαίνει ότι το ενδεχόμενο Χ=α είναι το αδύνατο ενδεχόμενο ∅.
- **Σημαίνει ότι αυτό το ενδεχόμενο α είναι απίθανο να συμβεί και όχι ότι είναι απραγματοποίητο.**

Ιδιότητες συνάρτησης πυκνότητας πιθανότητας

- Αν μια πραγματική συνάρτηση f ικανοποιεί τις συνθήκες

$$1. \quad f(x) \geq 0, \ \forall x \in (-\infty, +\infty)$$

$$2. \quad \int_{-\infty}^{+\infty} f(x)dx = 1$$

τότε αυτή η συνάρτηση μπορεί να χρησιμοποιηθεί ως συνάρτηση πυκνότητας μιας συνεχούς τυχαίας μεταβλητής X.

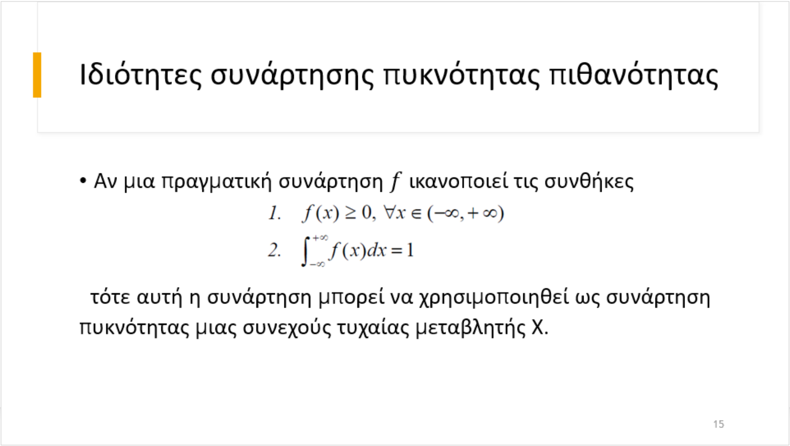
Παράδειγμα

- Μπορεί η συνάρτηση f με

$$f(x) = \begin{cases} \dfrac{3}{8}x^2, & 0 \le x \le 2 \\ 0 \quad, & \text{αλλού} \end{cases}$$

να χρησιμοποιηθεί ως συνάρτηση πυκνότητας μιας συνεχούς τυχαίας μεταβλητής;

Παράδειγμα

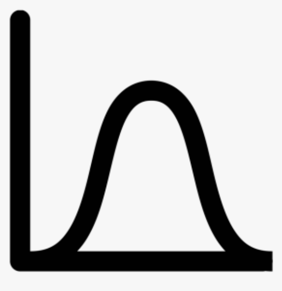


Ιδιότητες συνάρτησης πυκνότητας πιθανότητας

- Αν μια πραγματική συνάρτηση f ικανοποιεί τις συνθήκες
  1. $f(x) \geq 0,\ \forall x \in (-\infty, +\infty)$
  2. $\int_{-\infty}^{+\infty} f(x)dx = 1$

τότε αυτή η συνάρτηση μπορεί να χρησιμοποιηθεί ως συνάρτηση πυκνότητας μιας συνεχούς τυχαίας μεταβλητής X.

15



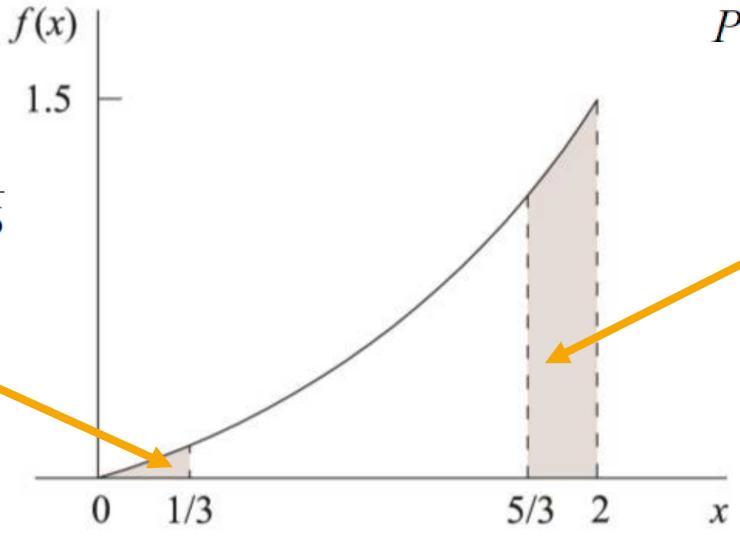
- Προφανώς θετική
- $\int_{-\infty}^{+\infty} f(x)dx = \int_{-\infty}^{0} 0dx + \int_{0}^{2} \frac{3}{8}x^2 dx + \int_{2}^{+\infty} 0dx = \frac{3}{8}\left[\frac{x^3}{3}\right]_0^2 = \frac{3}{8}\cdot\frac{8}{3} = 1$

Ισχύει και η δεύτερη συνθήκη και επομένως η f είναι μια πραγματική συνάρτηση κατάλληλη για να χρησιμοποιηθεί ως συνάρτηση πυκνότητας συνεχούς τυχαίας μεταβλητής.

Παράδειγμα

$$P\left(\frac{5}{3} \le X \le 2\right) = \int_{5/3}^{2} \frac{3}{8} x^2 dx = \frac{3}{8}\left[\frac{x^3}{3}\right]_{5/3}^{2} = \frac{91}{216}$$

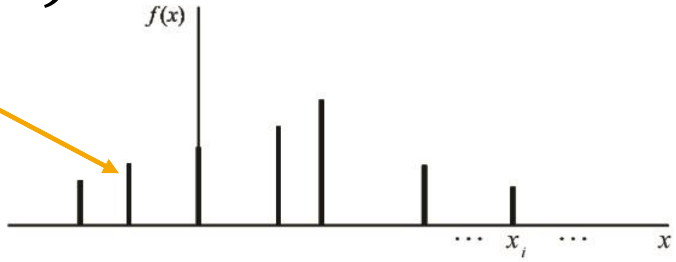
$$P\left(0 \le X \le \frac{1}{3}\right) = \int_{0}^{1/3} \frac{3}{8} x^2 dx = \frac{3}{8}\left[\frac{x^3}{3}\right]_{0}^{1/3} = \frac{1}{216}$$



Η τμ X κοντά στο 2 παίρνει τιμές με μεγαλύτερη πιθανότητα από ότι κοντά στο 0 αφού η πιθανότητα που εκχωρείται μέσω της f σε ένα διάστημα τιμών, εκφράζεται αντίστοιχο εμβαδόν κάτω από τη γραφική της παράσταση

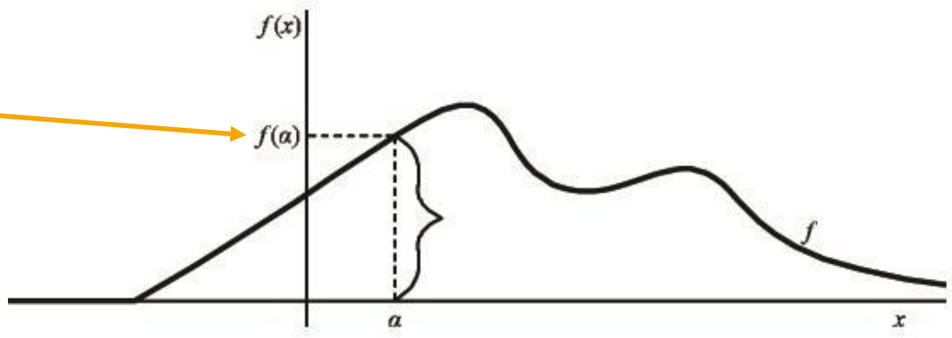
Σχόλια

- Η τιμή $f(x)$ της συνάρτησης πιθανότητας f μιας διακριτής τ.μ. X, εκφράζει όπως είδαμε την πιθανότητα $P(X = x)$



- Τί εκφράζει η $f(x)$ της συνάρτησης πυκνότητας πιθανότητας f μιας συνεχούς τ.μ. X;

- Η $f(x)$ εκφράζει **πυκνότητα πιθανότητας**, δηλαδή, είναι ένα μέτρο του πόσο πιθανό είναι να πάρει η X τιμή κοντά στο x



Η τιμής της συνάρτηση πυκνότητας πιθανότητας $f(x)$ δεν είναι πιθανότητα και μπορεί να πάρει τιμές μεγαλύτερες από το 1!

Ομοιόμορφή συνεχής τμ

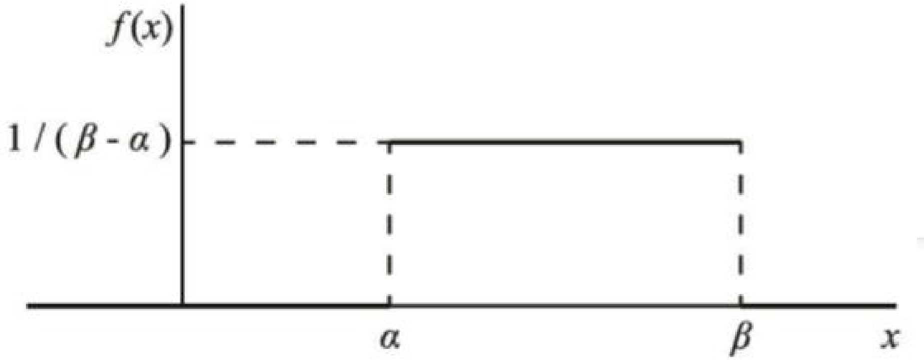
Έστω μια συνεχής τυχαία μεταβλητή Χ με συνάρτηση πυκνότητας

$$f(x) = \begin{cases} \dfrac{1}{\beta - \alpha}, & \alpha \leq x \leq \beta \\ 0, & \text{αλλού.} \end{cases}$$

Τότε η τυχαία μεταβλητή Χ ονομάζεται ομοιόμορφη στο διάστημα [α ,β] και η κατανομή της ομοιόμορφη κατανομή (uniform distribution) στο διάστημα [α ,β] και συμβολίζεται με U(α ,β)

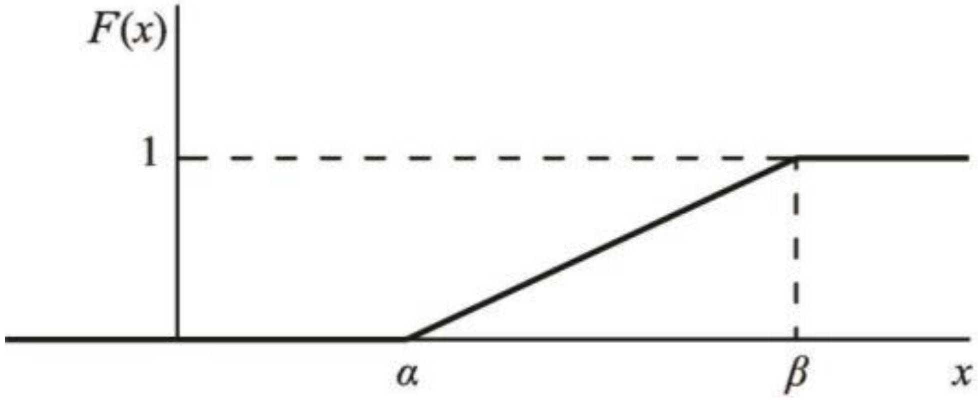
$$X \sim U(\alpha, \beta).$$

Ομοιόμορφή συνεχής τμ



CDF
συνάρτηση κατανομής

$$F(x)=\begin{cases}0, & -\infty < x < \alpha \\ \dfrac{x-\alpha}{\beta-\alpha}, & \alpha \le x < \beta \\ 1, & \beta \le x < +\infty.\end{cases}$$



Παράμετροι συνεχούς τυχαίας μεταβλητής

Παράμετροι θέσης συνεχούς τυχαίας μεταβλητής

Μέση τιμή συνεχούς τυχαίας μεταβλητής

- Έστω X μια συνεχής τυχαία μεταβλητή με συνάρτηση πυκνότητας f. Η μέση τιμή της X δίνεται από τον τύπο

$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx$$

εφόσον $\int_{-\infty}^{+\infty} |x| f(x) dx < +\infty$

Πρόταση

- Έστω X μια συνεχής τυχαία μεταβλητή με συνάρτηση πυκνότητας f. Για οποιαδήποτε πραγματική συνάρτηση g της X, η μέση τιμή της τυχαίας μεταβλητής $Y = g(x)$ δίνεται από τον τύπο

$$E[g(X)] = \int_{-\infty}^{+\infty} g(x) f(x) dx$$

Πρόταση

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

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

Προτάσεις

- Αν X συνεχής τυχαία μεταβλητή και α, β, δύο πραγματικοί αριθμοί, τότε

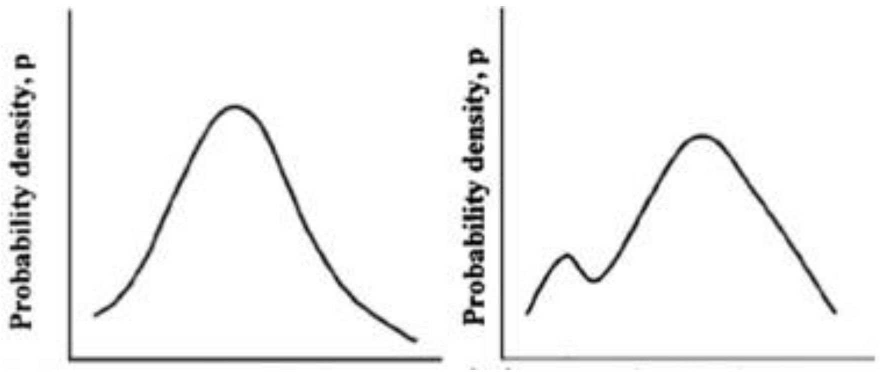
$$E(\alpha X + \beta) = \alpha E(X) + \beta$$

- Για οποιεσδήποτε διακριτές τυχαίες μεταβλητές $X_1, X_2, X_3, \dots, X_k$ των οποίων οι μέσες τιμές υπάρχουν και για οποιουσδήποτε πραγματικούς αριθμούς $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_k$ 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

- Μια τιμή x_0 της Χ λέγεται **κορυφή** της Χ αν το σημείο $(x_0, f(x_0))$ είναι μέγιστο (ολικό ή τοπικό) της f



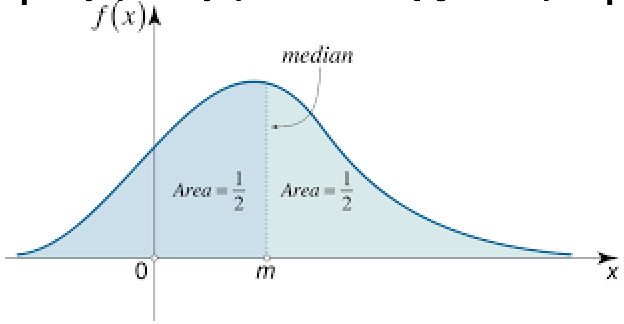
μονοκόρυφη δικόρυφη

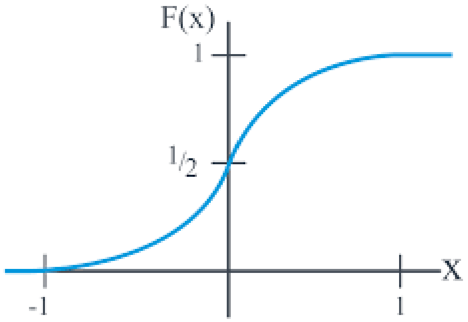
Διάμεσος

- Η λύση της εξίσωσης καλείται **διάμεσος** (median)

$$\int_{-\infty}^{\delta} f(x)dx = \frac{1}{2} \text{ ή } F(\delta) = \frac{1}{2}$$

όπου $f(x)$ και $F(x)$ οι συναρτήσεις πυκνότητας πιθανότητας και κατανομής της συνεχούς τμ X.



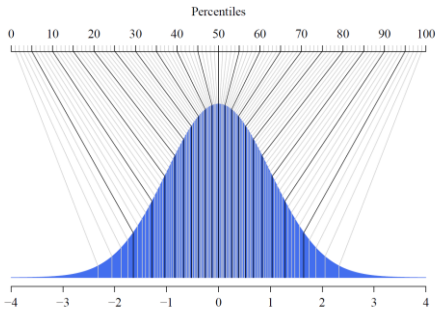


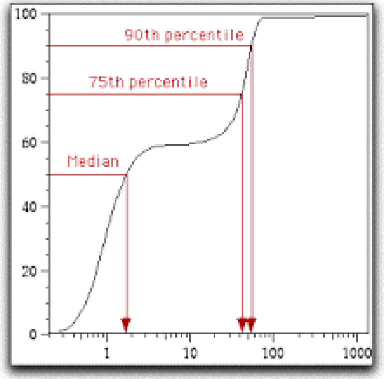
p-ποσοστιαίο σημείο

- Η λύση της εξίσωσης καλείται p-ποσοστιαίο σημείο (percentile)

$$\int_{-\infty}^{x_p} f(x)dx = p \text{ ή } F(x_p) = p$$

όπου $f(x)$ και $F(x)$ οι συναρτήσεις πυκνότητας πιθανότητας και κατανομής της συνεχούς τμ X.





Παράμετροι διασποράς συνεχούς τυχαίας μεταβλητής

Διακύμανση

- Έστω X μια συνεχής τυχαία μεταβλητή για την οποία υπάρχει η μέση τιμή $\mathrm{E(X)} = \mu$. Διακύμανση (variance) της X ονομάζεται η μέση τιμή της τυχαίας μεταβλητής $(\mathrm{X} - \mu)^2$ και συμβολίζεται με $Var(X)$ η σ^2.

$$\sigma^2 = Var(X) = E[(X-\mu)^2] = \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)dx$$

Τυπική απόκλιση

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

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

Χρήσιμη ιδιότητα

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

Διακύμανση και η τυπική απόκλιση του γραμμικού συνδυασμού

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

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

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

Τυποποιημένη συνεχής τμ

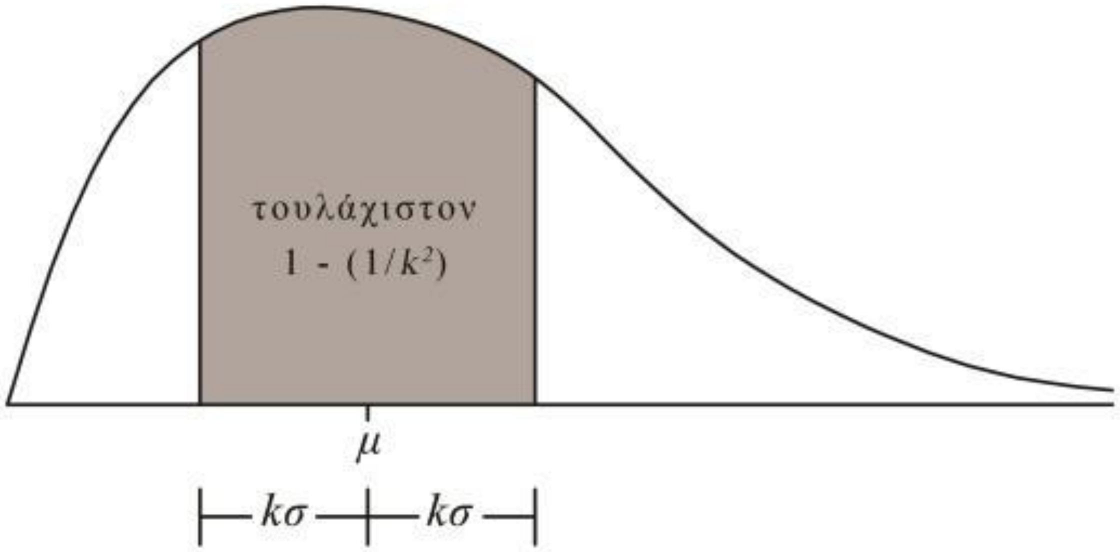
- Μια συνεχής τυχαία μεταβλητή με μέση τιμή 0 και διακύμανση 1 (άρα και τυπική απόκλιση 1), ονομάζεται τυποποιημένη τυχαία μεταβλητή (standardized random variable)
- Κάθε συνεχής τμ μπορεί να τυποποιηθεί αφαιρώντας την μέση τιμή και διαιρώντας με την τυπική απόκλιση

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

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%



Ενδοτεταρτημοριακό πλάτος/εύρος

- Ενδοτεταρτημοριακό έυρος ορίζεται ως η διαφορά $x_{0.75} - x_{0.25}$

- Ημιενδοτεταρτημοριακό εύρος ορίζεται ως η διαφορά $\frac{x_{0.75} - x_{0.25}}{2}$

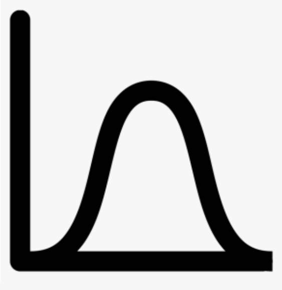
Παράδειγμα

- Μπορεί η τμ X με συνάρτηση πυκνότητας πιθανότητας f με

$$f(x)=\begin{cases}\dfrac{1}{\beta - \alpha}, & \alpha \leq x \leq \beta \\ 0, & \text{αλλού.}\end{cases}$$

Να υπολογισθούν οι παράμετροι θέσης και διασποράς της X

Παράδειγμα (συν)



- Μέση τιμή: $\mu = E(X) = \int_{-\infty}^{+\infty} x f(x)dx =$

$$\int_{\alpha}^{\beta} x \frac{1}{\beta-\alpha} dx = \frac{1}{\beta-\alpha} \cdot \left[\frac{x^2}{2}\right]_{\alpha}^{\beta} =$$

$$= \frac{1}{\beta-\alpha} \cdot \frac{\beta^2-\alpha^2}{2} = \frac{\alpha+\beta}{2}$$

Μέση τιμή συνεχούς τυχαίας μεταβλητής

- Έστω X μια συνεχής τυχαία μεταβλητή με συνάρτηση πυκνότητας f. Η μέση τιμή της X δίνεται από τον τύπο

$$E(X) = \int_{-\infty}^{+\infty} x f(x)dx$$

εφόσον $\int_{-\infty}^{+\infty} |x|\, f(x)dx < +\infty$

Μέση τιμή συνεχούς τυχαίας μεταβλητής

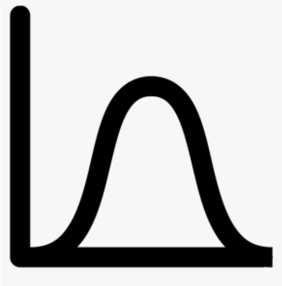
- Έστω X μια συνεχής τυχαία μεταβλητή με συνάρτηση πυκνότητας f. Η μέση τιμή της X δίνεται από τον τύπο

$$E(X) = \int_{-\infty}^{+\infty} x f(x)dx$$

εφόσον $\int_{-\infty}^{+\infty} |x|\, f(x)dx < +\infty$

Παράδειγμα (συν)

$$f(x)=\begin{cases}\dfrac{1}{\beta - \alpha}, & \alpha \le x \le \beta \\ 0, & \text{αλλού.}\end{cases}$$

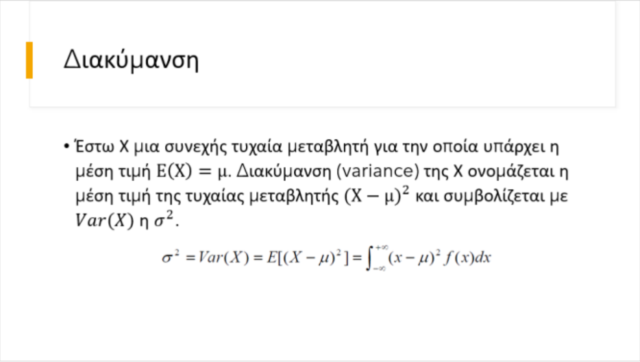


- Διακύμανση: $E(X^2)=\int_{-\infty}^{+\infty} x^2 f(x)dx =$

$$\int_{\alpha}^{\beta} x^2 \frac{1}{\beta-\alpha}dx = \frac{1}{\beta-\alpha}\cdot\left[\frac{x^3}{3}\right]_{\alpha}^{\beta} = \frac{1}{\beta-\alpha}\cdot\frac{\beta^3-\alpha^3}{3} =$$

$$= \frac{\beta^2+\alpha\beta+\alpha^2}{3}$$

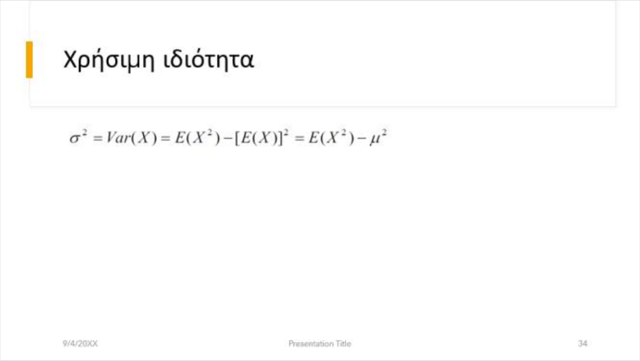
$$\sigma^2 = E(X^2)-\mu^2 = \frac{\beta^2+\alpha\beta+\alpha^2}{3} - \left(\frac{\alpha+\beta}{2}\right)^2 = \frac{(\beta-\alpha)^2}{12}$$



Διακύμανση

- Έστω X μια συνεχής τυχαία μεταβλητή για την οποία υπάρχει η μέση τιμή $E(X) = \mu$. Διακύμανση (variance) της X ονομάζεται η μέση τιμή της τυχαίας μεταβλητής $(X-\mu)^2$ και συμβολίζεται με $Var(X)$ η σ^2.

$$\sigma^2 = Var(X) = E[(X-\mu)^2] = \int_{-\infty}^{+\infty}(x-\mu)^2 f(x)dx$$



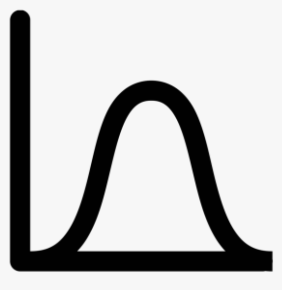
Χρήσιμη ιδιότητα

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

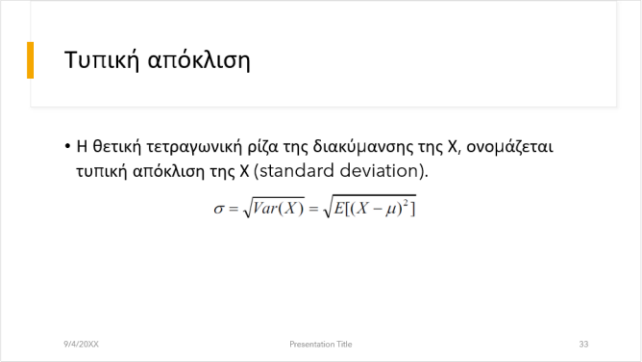
9/4/20XX Presentation Title 34

Παράδειγμα (συν)

$$f(x)=\begin{cases}\dfrac{1}{\beta-\alpha}, & \alpha\le x\le\beta\\ 0, & \text{αλλού.}\end{cases}$$



- Τυπική απόκλιση:



Τυπική απόκλιση

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

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

9/4/20XX Presentation Title 33

$$\sigma=\sqrt{Var(X)}=\sqrt{\frac{(\beta-\alpha)^2}{12}}=\frac{\sqrt{3}(\beta-\alpha)}{6}$$

Παράδειγμα (συν)

$$f(x)=\begin{cases}\dfrac{1}{\beta - \alpha}, & \alpha \le x \le \beta \\ 0, & \text{αλλού.}\end{cases}$$

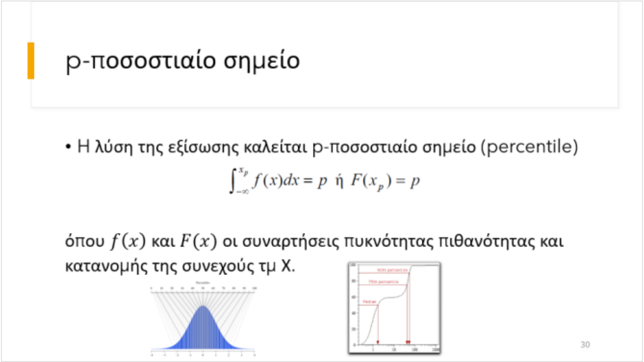
- p-ποσοστιαία σημεία:

$$\int_{-\infty}^{x_p} f(x)dx = p \Leftrightarrow \int_{\alpha}^{x_p} \frac{1}{\beta-\alpha}dx = p \Leftrightarrow$$

$$\frac{1}{\beta-\alpha}\cdot[x]_{\alpha}^{x_p} = p \Leftrightarrow \frac{1}{\beta-\alpha}\cdot(x_p-\alpha) = p$$

$$x_p = \alpha + (\beta-\alpha)p$$

$$x_{0.5} = \alpha + (\beta-\alpha)0.5,\ x_{0.25} = \alpha + (\beta-\alpha)0.25,\ x_{0.75} = \alpha + (\beta-\alpha)0.75$$


p-ποσοστιαίο σημείο

- Η λύση της εξίσωσης καλείται p-ποσοστιαίο σημείο (percentile)

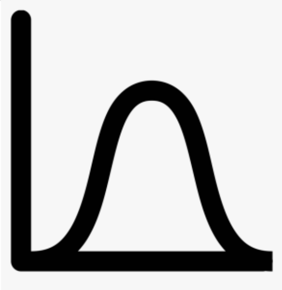
$$\int_{-\infty}^{x_p} f(x)dx = p \text{ ή } F(x_p) = p$$

όπου $f(x)$ και $F(x)$ οι συναρτήσεις πυκνότητας πιθανότητας και κατανομής της συνεχούς τμ X.

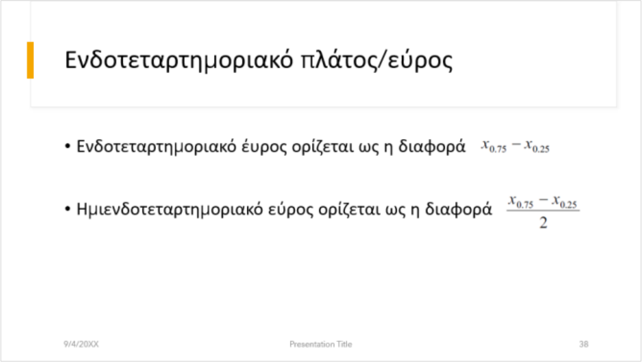
30

Παράδειγμα (συν)

$$f(x)=\begin{cases}\dfrac{1}{\beta - \alpha}, & \alpha \le x \le \beta \\ 0, & \text{αλλού.}\end{cases}$$



- Ενδοτεταρτημοριακό εύρος: $x_{0.75} - x_{0.25} = (\beta - \alpha)0.5$



Παράδειγμα

- Η ποσότητα καλίου που περιέχεται σε ένα πορτοκάλι μεσαίου μεγέθους είναι τυχαία μεταβλητή με μέση τιμή 550 mg. Επίσης, η ποσότητα καλίου που περιέχεται σε ένα ακτινίδιο μεσαίου μεγέθους είναι τυχαία μεταβλητή με μέση τιμή 300 mg. Επειδή σκεφτόμαστε να εντάξουμε σε ένα ημερήσιο διαιτολόγιο δύο πορτοκάλια και ένα ακτινίδιο μεσαίου μεγέθους, μας ενδιαφέρει η μέση ποσότητα καλίου που περιέχεται σε ένα τέτοιο συνδυασμό φρούτων.

Παράδειγμα (συν)

- Έστω X_1 η τυχαία μεταβλητή που εκφράζει την ποσότητα καλίου που περιέχεται σε ένα πορτοκάλι μεσαίου μεγέθους και X_2 η τυχαία μεταβλητή που εκφράζει την ποσότητα καλίου που περιέχεται σε ένα ακτινίδιο μεσαίου μεγέθους.
- Μας ενδιαφέρει η μέση τιμή της τυχαίας μεταβλητής που περιέχει 2 πορτοκάλια και 1 ακτινίδιο. Δηλαδή η τμ

$$Y = 2X_1 + X_2$$

- Έστω X_1 η τυχαία μεταβλητή που εκφράζει την ποσότητα καλίου που περιέχεται σε ένα πορτοκάλι μεσαίου μεγέθους και X_2 η τυχαία μεταβλητή που εκφράζει την ποσότητα καλίου που περιέχεται σε ένα ακτινίδιο μεσαίου μεγέθους.
- Μας ενδιαφέρει η μέση τιμή της τυχαίας μεταβλητής που περιέχει 2 πορτοκάλια και 1 ακτινίδιο. Δηλαδή η τμ

Παράδειγμα (συν)

- Την κατανομή της Y δεν την γνωρίζουμε.
- Αλλά γνωρίζουμε τις επιμέρους μέσες τιμές των X_1 και X_2

Προτάσεις

- Αν X συνεχής τυχαία μεταβλητή και α, β, δύο πραγματικοί αριθμοί, τότε
$$E(\alpha X+\beta)=\alpha E(X)+\beta$$
- Για οποιεσδήποτε διακριτές τυχαίες μεταβλητές $X_1, X_2, X_3, \dots, X_k$ των οποίων οι μέσες τιμές υπάρχουν και για οποιουσδήποτε πραγματικούς αριθμούς $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_k$ ισχύει:
$$E(\alpha_1 X_1+\alpha_2 X_2+\dots+\alpha_k X_k)=\alpha_1 E(X_1)+\alpha_2 E(X_2)+\dots+\alpha_k E(X_k)$$

27

$$E(Y)=2E(X_1)+E(X_2)=$$

$$2\cdot 550+300=1400\, mg$$

Ανεξαρτησία τυχαίων μεταβλητών

Ανεξαρτησία ενδεχομένων και πειραμάτων

- Δύο ενδεχόμενα είναι ανεξάρτητα αν η πραγματοποίηση του ενός δεν έχει καμία επίδραση στην πραγματοποίηση του άλλου.
- Γενικότερα, $\nu \geq 2$ ενδεχόμενα είναι ανεξάρτητα αν η πραγματοποίηση οποιουδήποτε από αυτά καθώς και η (ταυτόχρονη) πραγματοποίηση οποιουδήποτε αριθμού από αυτά δεν έχει καμία επίδραση στην πραγματοποίηση των υπολοίπων
- $\nu \geq 2$ πειράματα είναι ανεξάρτητα αν δεν «αλληλοεπηρεάζονται».

Δύο ανεξάρτητες τυχαίες μεταβλητές

- Δύο τυχαίες μεταβλητές X, Y ονομάζονται (στοχαστικά) **ανεξάρτητες** (independent random variables) αν τα ενδεχόμενα $\{X \in A\}$ και $\{Y \in B\}$ είναι ανεξάρτητα ενδεχόμενα για οποιαδήποτε (μετρήσιμα) υποσύνολα A, B του συνόλου των πραγματικών αριθμών, δηλαδή αν

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$$

- Προφανώς ισχύει

$$P(X \in A \mid Y \in B) = P(X \in A) \text{ και } P(Y \in B \mid X \in A) = P(Y \in B)$$

$\nu \geq 2$ ανεξάρτητες τυχαίες μεταβλητές

- $\nu \geq 2$ τυχαίες μεταβλητές $X_1, X_2 \ldots X_\nu$ ονομάζονται (στοχαστικά) ανεξάρτητες αν για οποιαδήποτε (μετρήσιμα) υποσύνολα $A_1, A_2 \ldots A_\nu$ του συνόλου των πραγματικών αριθμών ισχύει

$$P(X_1 \in A_1, X_2 \in A_2, \ldots, X_\nu \in A_\nu) = P(X_1 \in A_1) P(X_2 \in A_2) \ldots P(X_\nu \in A_\nu)$$

Ιδιότητες

- Αν X, Y ανεξάρτητες τυχαίες μεταβλητές και x, y, οποιοιδήποτε πραγματικοί αριθμοί τότε, για $A = (-\infty, x]$ και $\mathrm{B} = (-\infty, y]$ ισχύει $P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y)$
- Αν $\mathrm{X}_1, \mathrm{X}_2 \ldots \mathrm{X}_\nu$ ανεξάρτητες τυχαίες μεταβλητές και $\mathrm{x}_1, \mathrm{x}_2 \ldots \mathrm{x}_\nu$ οποιοιδήποτε πραγματικοί αριθμοί τότε

Σχόλιο

- Οι τυχαίες μεταβλητές που αναφέρονται σε πειράματα που εκτελούνται σε διαφορετικούς τόπους ή σε διαφορετικά άτομα, γενικά, σε διαφορετικές πειραματικές μονάδες, θεωρούνται ανεξάρτητες
- Δηλαδή η ανεξαρτησία συνήθως προκύπτει από τις συνθήκες παρά διαπιστώνεται από τον ορισμό
- Ανεξάρτητες θεωρούνται επίσης, οι τυχαίες μεταβλητές που αναφέρονται σε διαδοχικές ανεξάρτητες επαναλήψεις του ίδιου πειράματος, δηλαδή, σε **ανεξάρτητες δοκιμές**.

Ιδιότητες διακύμανσης ανεξαρτήτων τμ

- Για οποιεσδήποτε ανεξάρτητες τυχαίες μεταβλητές $X_1, X_2 \dots X_\nu$ για τις οποίες $\mathrm{E}(X^2)$ και οποιουσδήποτε πραγματικούς αριθμούς $\alpha_1, \alpha_2 \dots \alpha_\nu$ ισχύει

$$Var(\alpha_1 X_1 + \alpha_2 X_2 + \ldots + \alpha_k X_k) = \alpha_1^2 Var(X_1) + \alpha_2^2 Var(X_2) + \ldots + \alpha_k^2 Var(X_k)$$

Τυχαίο δείγμα
Στατιστικός πληθυσμός

Τυχαίο δείγμα

- Τυχαίο δείγμα στην στατιστική θεωρείται η βάση το οποίο οδηγούμαστε σε συμπεράσματα από το «μέρος» για το «όλο».
- Το δείγμα αποτελείται από ανεξάρτητες τυχαίες μεταβλητές.

Στατιστικός πληθυσμός

- **Πληθυσμός ή στατιστικός πληθυσμός** (statistical population), είναι οι δυνατές τιμές μιας τυχαίας μεταβλητής, δηλαδή η κατανομή των δυνατών τιμών μιας τυχαίας μεταβλητής.

Δειγματικός μέσος

- Ας θεωρήσουμε ένα τυχαίο δείγμα $X_1, X_2 \ldots X_\nu$ (δηλαδή ν τυχαίες μεταβλητές) από έναν πληθυσμό (δηλαδή κατανομή).
- Δειγματικός μέσος (sample mean) είναι η τμ που ορίζεται ως

$$\frac{1}{\nu}X_1 + \frac{1}{\nu}X_2 + \ldots + \frac{1}{\nu}X_\nu = \frac{1}{\nu}\sum_{i=1}^{\nu} X_i$$

- Συμβολίζεται $\overline{X}$

Διακύμανση δειγματικού μέσου

- Η διακύμανση $\mathrm{Var}(\overline{\mathrm{X}})$ του δειγματικού μέσου

$$\overline{X} = \frac{1}{\nu}\sum_{i=1}^{\nu} X_i$$

συμβολίζεται με $\sigma_{\overline{\mathrm{X}}}^2$ και ισούται με σ^2/ν δηλαδή

$$Var(\overline{X}) = \sigma_{\overline{X}}^2 = \frac{\sigma^2}{\nu}$$

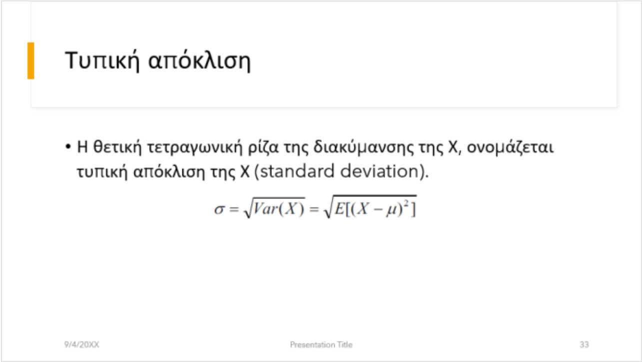
Σχόλια

- Η διακύμανση του δειγματικού μέσου είναι μικρότερη της διακύμανση του πληθυσμού
- Όσο το μέγεθος του δείγματος μεγαλώνει η διακύμανση του δειγματικού μέσου μειώνεται

Τυπικό σφάλμα

- Τυπικό σφάλμα είναι η τυπική απόκλιση του δειγματικού μέσου είναι ίση με $\sigma/\sqrt{\nu}$
- Συνήθως ονομάζεται τυπικό σφάλμα (standard error)

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{\nu}}$$



Μέση τιμή δειγματικού μέσου vs μέσης τιμής πληθυσμού

- Έστω $\mathrm{X}_1, \mathrm{X}_2 \ldots \mathrm{X}_\nu$ ένα τυχαίο δείγμα από μια κατανομή που έχει μέση τιμή μ και διακύμανση σ^2. Η μέση τιμή $\mathrm{E}(\overline{\mathrm{X}})$ του δειγματικού μέσου

$$\overline{X} = \frac{1}{\nu}\sum_{i=1}^{\nu} X_i$$

συμβολίζεται με $\mu_{\overline{\mathrm{X}}}$ και είναι ίση με τη μέση τιμή του πληθυσμού μ, δηλαδή

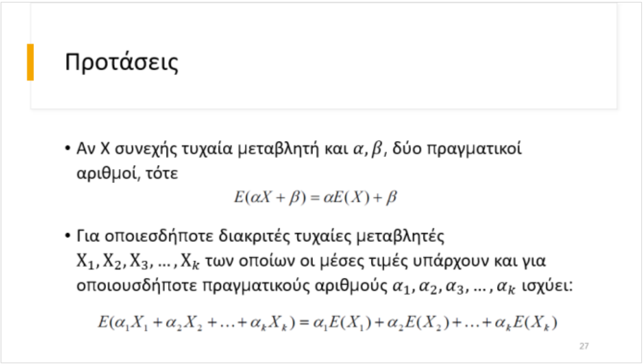
$$E(\overline{X}) = \mu_{\overline{X}} = \mu$$

Απόδειξη

$$E(\overline{X}) = E\left(\frac{1}{\nu}\sum_{i=1}^{\nu} X_i\right) = E\left(\frac{1}{\nu}X_1 + \frac{1}{\nu}X_2 + \ldots, + \frac{1}{\nu}X_\nu\right)$$

$$= \frac{1}{\nu}E(X_1) + \frac{1}{\nu}E(X_2) + \ldots + \frac{1}{\nu}E(X_\nu) =$$

$$= \frac{1}{\nu}\mu + \frac{1}{\nu}\mu + \ldots + \frac{1}{\nu}\mu = \nu\frac{1}{\nu}\mu = \mu$$



Backup