# ΑΝΑΛΥΣΗ ΧΡΟΝΟΣΕΙΡΩΝ ΚΑΙ ΔΕΔΟΜΕΝΩΝ ΠΕΡΙΒΑΛΛΟΝΤΙΚΩΝ ΚΙΝΔΥΝΩΝ

Καθ. Θεόδωρος Καρακασίδης
Δρ Αθανάσιος Φράγκου

**Διατμηματικό Πρόγραμμα Μεταπτυχιακών Σπουδών**
**«Βιώσιμη Διαχείριση Περιβαλλοντικών Αλλαγών και Κυκλική Οικονομία»**

# Κατανομές

- Σημαντικό ρόλο στην Στατιστική Ανάλυση  εκτός από την εύρεση πιθανοτήτων αποτελεί η συχνότητα εμφάνισης των μετρήσεων κατά τη διεξαγωγή μιας μελέτης καθώς και η κατανομή των μετρήσεων, ο τρόπος δηλαδή που οι τιμές «απλώνονται» στο επίπεδο (γραφική παράσταση).

# Κατανομές

- Τυχαία μεταβλητή Χ: Σύνολο των αποτελεσμάτων ενός πειράματος (σύνολο των δεδομένων)

Κατηγορίες τυχαίων μεταβλητών

- *Διακριτή* (πλήθος τηλεφωνικών κλήσεων, αριθμός ημερών βροχής σε ένα μήνα)

- *Συνεχής* ( ατμοσφαιρική πίεση, θερμοκρασία)

# Κατανομές

- Συνάρτηση κατανομής: ο τρόπος με τον οποίο κατανέμονται οι τιμές (μετρήσεις) της τυχαίας μεταβλητής.

$$\boldsymbol{F(x)} = P(X \leq x)$$

Μας δίνει όλες τις πληροφορίες που χρειαζόμαστε για την τυχαία μεταβλητή , δηλαδή (μεταφράζοντας την παραπάνω σχέση) ποια η πιθανότητα η τυχαία μεταβλητή X να παίρνει τιμές μικρότερες από ένα αριθμό x

Ισχύει

$$P(X > \kappa) = 1 - P(X \leq \kappa)$$

# Κατανομές

- Συνάρτηση κατανομής (Παράδειγμα): Σε ένα τηλεφωνικό κέντρο οι κλήσεις έρχονται με ρυθμό 10 κλήσεις ανά ώρα. Η πιθανότητα να έρθουν λιγότερες από δέκα είναι

$$\boldsymbol{F(x)} = P(X \leq \mathbf{10})$$

Η συνάρτηση κατανομής F(x) της οποίας τον τύπο γνωρίζουμε, θα μας δώσει το αποτέλεσμα (πληροφορία).
Εδώ η τυχαία μεταβλητή X είναι ο αριθμός των κλήσεων.

# Κατανομές
# Παράμετροι κατανομής

- Τρόπος με τον οποίο οι τιμές της τυχαίας μεταβλητής κατανέμονται.
- -Μέση τιμή (μέτρο θέσης) → EX
- -Διασπορά (μέτρο διασποράς) →varX

- **Η διασπορά αποτελεί δείκτη συγκέντρωσης της κατανομής των τιμών.**
- **Εάν η τιμή της διασποράς είναι μικρή, τότε αυξάνεται η πιθανότητα οι τιμές της τυχαίας μεταβλητής να είναι γύρω από τη μέση τιμή. Σε αντίθετη περίπτωση οι τιμές έχουν μεγάλη απόκλιση από τη μέση τιμή.**

- **Παρακάτω θα αναφερθούμε στις σημαντικότερες κατανομές για τους σκοπούς του σεμιναρίου**

# Διακριτές Κατανομές

# Διακριτές Κατανομές

- Τα δεδομένα μεγεθών όπως ο αριθμός ρίψεων ενός ζαριού, αριθμός ημερών του μήνα που βρέχει είναι διακριτές τυχαίες μεταβλητές -> ακολουθούν διακριτές κατανομές

- Κυριότερες διακριτές κατανομές

-Διωνυμική κατανομή

- Ομοιόμορφη διακριτή κατανομή

- Κατανομή Poisson

# Διακριτές Κατανομές
# Δυωνυμική

- Δυωνυμική κατανομή ακολουθεί η διεξαγωγή διαδικασίας ενός αριθμού ανεξαρτήτων δοκιμών με την ίδια πιθανότητα επιτυχίας σε κάθε δοκιμή.

- Οι τιμές της μεταβλητής Χ γίνονται μηδέν (0) ή ένα (1)

πχ «κεφαλή» για ν φορές ρίψη νομίσματος

- Συνάρτηση Πιθανότητας F

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x},$$
$$x = 0,1,2,\dots,n, \qquad 0 < p < 1$$

- Μέση Τιμή

$$\boldsymbol{EX = np}$$

- Διασπορά

$$\boldsymbol{varX = np(1-p)}$$

# Διακριτές Κατανομές
# Δυωνυμική - Παράδειγμα

- Μετά από απότομη χαλαζόπτωση καταστράφηκε το 20% της ποσότητας μήλων μιας μεγάλης παραγωγής. Ο παραγωγός για να εκτιμήσει το μέγεθος της φυσικής καταστροφής προκειμένου να ζητήσει το ανάλογο επίδομα από την πολιτεία, επιλέγει 4 τυχαία μήλα. Ποια η πιθανότητα να είναι τουλάχιστον ένα χαλασμένο;
- Να βρεθεί η μέση τιμή και η διασπορά του αριθμού των χαλασμένων μήλων.

# Διακριτές Κατανομές
# Δυωνυμική – Παράδειγμα (Λύση)

- Μεγάλη ποσότητα μήλων: Εάν επιλέξουμε ένα θα είναι χαλασμένο ή όχι (ανεξάρτητες δοκιμές με την ίδια πιθανότητα επιτυχίας σε κάθε δοκιμή)

  - Η πιθανότητα να υπάρχουν x χαλασμένα μήλα στο δείγμα των n μήλων δίνεται με τη διωνυμική κατανομή

$$P(X=x)=\binom{n}{x}p^x(1-p)^{n-x}, \qquad x=0,1,2,\ldots,n, \qquad 0<p<1$$

- Η πιθανότητα να υπάρχουν ένα χαλασμένο ή περισσότερα

$$P(X\geq 1)=1-P(X=0)=1-\binom{\mathbf{4}}{\mathbf{0}}(\mathbf{0.2})^{\mathbf{0}}(1-\mathbf{0.2})^{\mathbf{4-0}}=\mathbf{1-0,4094}$$
$$=\mathbf{0,5904}$$

- Μέση Τιμή $\boldsymbol{EX=np=4\cdot 0,2=0,8}$

- Διασπορά $\boldsymbol{varX=np(1-p)=4\cdot 0,2\cdot 0,8=0,64}$

# Διακριτές Κατανομές
# Ομοιόμορφη διακριτή

- Η υπό εξέταση τ.μ λαμβάνει πεπερασμένο πλήθος τιμών πχ για τη ρίψη ενός ζαριού η συνάρτηση πιθανότητας λαμβάνει τις τιμές 1,2,…,6 και η πιθανότητα είναι 1/6

- Συνάρτηση Πιθανότητας

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

# Διακριτές Κατανομές
# Poisson

- Γεγονός σε συγκεκριμένο χρονικό διάστημα με ένα μέσο ρυθμό εμφανίσεων λ όπου μετρούμε τον αριθμό τον εμφανίσεων X,

πχ ο αριθμός των ελαττωματικών λαμπτήρων που παράγει ένα εργοστάσιο ή το πλήθος των καταιγίδων σε ένα έτος.

- Η συνάρτηση πιθανότητας

$$P(X = x) = \begin{cases} \dfrac{e^{-\lambda}\lambda^{x}}{x!}, & x = 0,1,2,\ldots \\ 0\,, & \alpha\lambda\lambda o\acute{\upsilon} \end{cases}$$

- Μέση Τιμή

$$\boldsymbol{EX = \lambda}$$

- Διασπορά

$$\boldsymbol{varX = \lambda}$$

# Διακριτές Κατανομές
# Poisson Παράδειγμα

- Προκειμένου για τη μελέτη της επιρροής στο μικροκλίμα του αριθμού των κοιλάδων σε μια μεγάλη περιοχή οικολογικού ενδιαφέροντος μετρήθηκε και βρέθηκε ο αριθμός τους ότι ακολουθεί την κατανομή Poisson με λ=3,2 κοιλάδες ανά τετραγωνικό χιλιόμετρο.

Σε μια περιοχή έκτασης 2 τετρ. χιλιομέτρων ποια η πιθανότητα να υπάρχουν 2 κοιλάδες;

$$P(X=x)=\begin{cases} \dfrac{e^{-\lambda}\lambda^{x}}{x!}, & x=0,1,2,\ldots \\ 0\,, & \alpha\lambda\lambda o\acute{\upsilon} \end{cases}$$

# Διακριτές Κατανομές
# Poisson  Παράδειγμα

- Έστω X τ.μ. ο αριθμός των κοιλάδων στην περιοχή των 2 τετρ. χιλιομέτρων. Τότε η X ακολουθεί κατανομή Poisson με παράμετρο κ=2λ=6.4

$$P(X = \mathbf{2}) = \boldsymbol{e}^{-6.4}\frac{(\mathbf{6.4})^{\mathbf{2}}}{\mathbf{2}!} = \mathbf{0.034}$$

- Μέση Τιμή $\boldsymbol{EX = 6.4}$

- Διασπορά $\boldsymbol{varX = 6.4}$

# Συνεχείς Κατανομές

# Συνεχείς Κατανομές

- Τα δεδομένα μεγεθών όπως η ατμοσφαιρική πίεση, υγρασία, θερμοκρασία είναι συνεχείς τυχαίες μεταβλητές -> ακολουθούν συνεχείς κατανομές

- Κυριότερες συνεχείς κατανομές

- Ομοιόμορφη συνεχής κατανομή
- Κανονική κατανομή
- Student ή t – κατανομή
- Κατανομή $X_{\nu}^{2}$

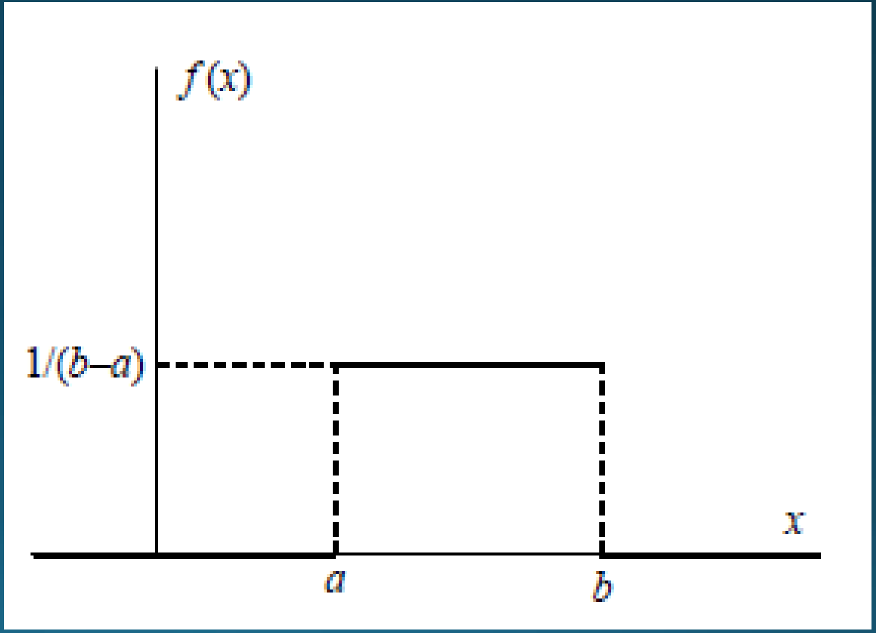
# Συνεχείς Κατανομές
# Ομοιόμορφη συνεχής

- Ομοιόμορφη κατανομή $U(a,b)$ακολουθεί μια διαδικασία κατά την οποία εντός του διαστήματος $(a,b)$ καθένα από τα n ισαπέχοντα σημεία του έχει πιθανότητα $\frac{1}{n}$

- **Συνάρτηση πυκνότητας πιθανότητας**

$$f(x) = \begin{cases} \dfrac{1}{b-\alpha}, & \alpha \le x \le b \\ 0 & \alpha\lambda\lambda o\acute{\upsilon} \end{cases}$$

**Γραφική Παράσταση κατανομής**




- Μέση Τιμή

$$\boldsymbol{EX = \frac{\alpha + b}{2}}$$

- Διασπορά

$$\boldsymbol{varX = \frac{(b-a)^2}{12}}$$

# Συνεχείς Κατανομές
# Κανονική Κατανομή

- Κάθε φυσική ποσότητα της οποίας η τιμή μπορεί να θεωρηθεί ότι διαμορφώνεται από ένα μεγάλο αριθμό (ανεξάρτητων) παραγόντων ακολουθεί προσεγγιστικά κανονική κατανομή.

- Παραδείγματα: το ύψος των παιδιών σε μία τάξη διαμορφώνεται από διάφορους φυσικούς παράγοντες (είδος διατροφής, μυϊκή άσκηση, DNA κλπ).

- Το βάρος των χαλικιών σε μια παραλία διαμορφώνεται από τις εκάστοτε καιρικές συνθήκες που επικρατούν κάθε χρόνο και επηρεάζουν την παραλία όπως ο κυματισμός της θάλασσας, η ποσότητα της βροχής, η σφοδρότητα των ανέμων κλπ)

# Συνεχείς Κατανομές
# Κανονική Κατανομή

- Συνάρτηση πυκνότητας πιθανότητας

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \qquad -\infty < x < \infty$$

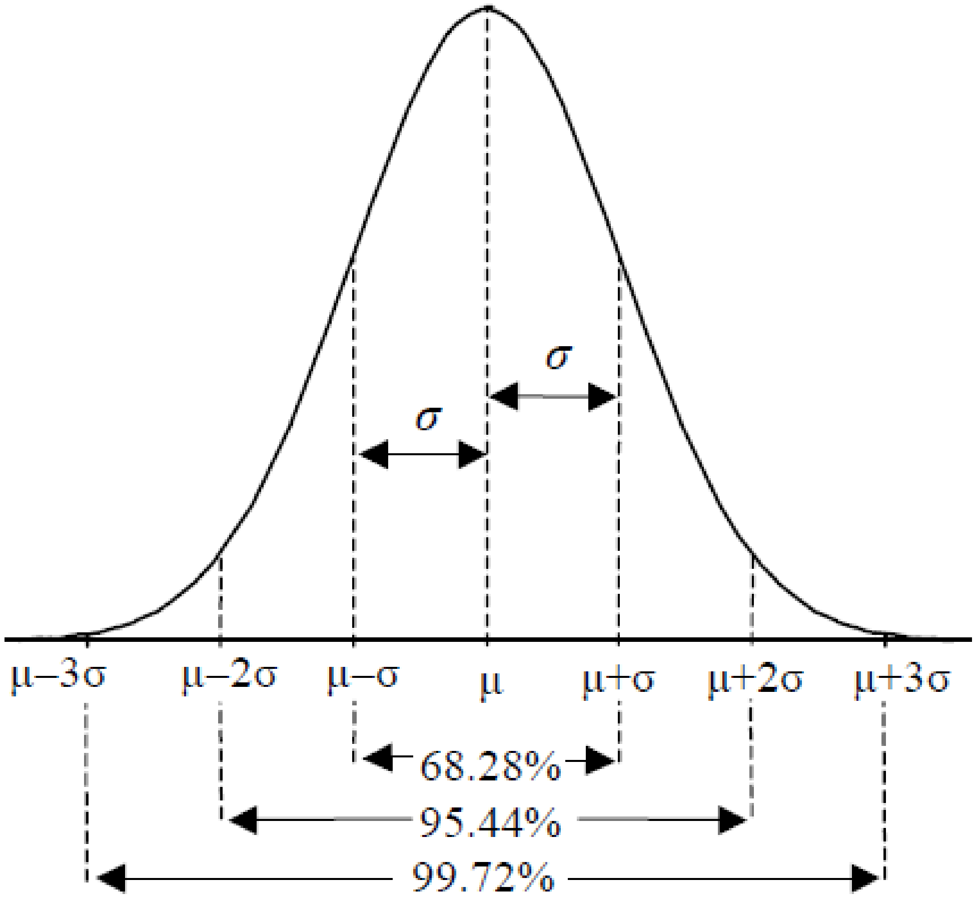
- Μέση Τιμή $\boldsymbol{EX = \mu}$

- Διασπορά $\boldsymbol{varX = \sigma}$

- Επειδή θα ασχοληθούμε περισσότερο με τη συγκεκριμένη κατανομή θα συμβολίζουμε $X \sim N(\mu, \sigma^2)$

# Συνεχείς Κατανομές
# Κανονική Κατανομή

- Αν μία τ.μ. $X$ ακολουθεί $N(\mu,\sigma^2)$ τότε παίρνει τιμές μεταξύ του $\mu-3\sigma$ και του $\mu+3\sigma$ με πιθανότητα σχεδόν 1, τιμές μεταξύ του $\mu-2\sigma$ και του $\mu+2\sigma$ με πιθανότητα περίπου 95% και τιμές μεταξύ του $\mu-\sigma$ και του $\mu+\sigma$ με πιθανότητα περίπου 68%.



# Συνεχείς Κατανομές
# Κανονική Κατανομή (τυπική)

- Πολλές φορές είναι χρήσιμο να τυποποιούμε τα δεδομένα μας ή να τα κανονικοποιούμε (μ=0,σ=1)

- Ειδική περίπτωση για μέση τιμή μηδέν και τυπική απόκλιση 1 (μ=0,σ=1)

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \qquad -\infty < z < +\infty$$

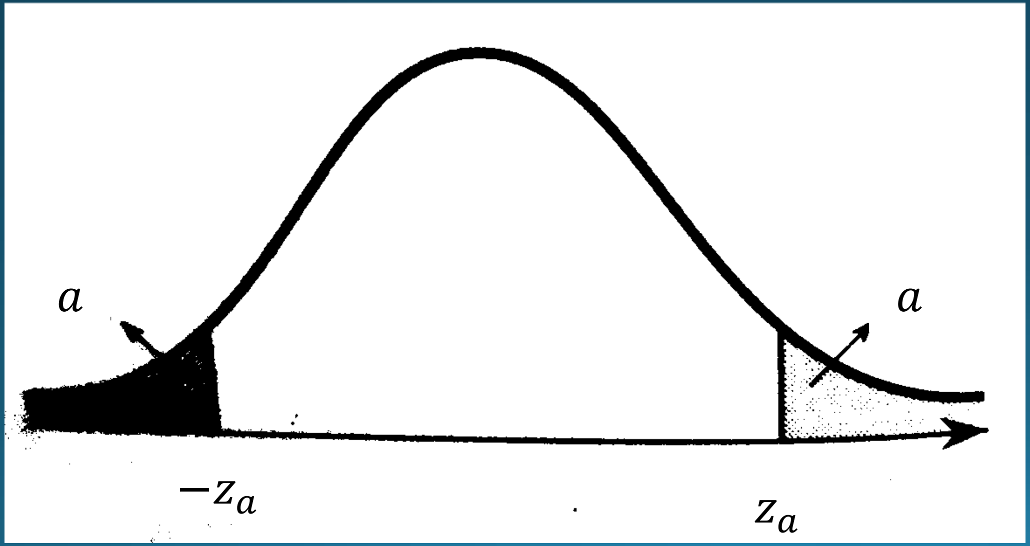
Η τυποποίηση γίνεται κυρίως σε δεδομένα που δεν έχουν το ίδιο εύρος τιμών.

# Συνεχείς Κατανομές
# Κανονική Κατανομή (τυπική)

Τυποποίηση(Ορισμός): Εάν μια τυχαία μεταβλητή X ακολουθεί την κανονική κατανομή, δηλαδή $X \sim N(\mu, \sigma^2)$, τότε η τυχαία μεταβλητή $Y = \frac{X-\mu}{\sigma}$ ακολουθεί την τυπική κανονική κατανομή, δηλαδή $Y \sim N(0,1)$

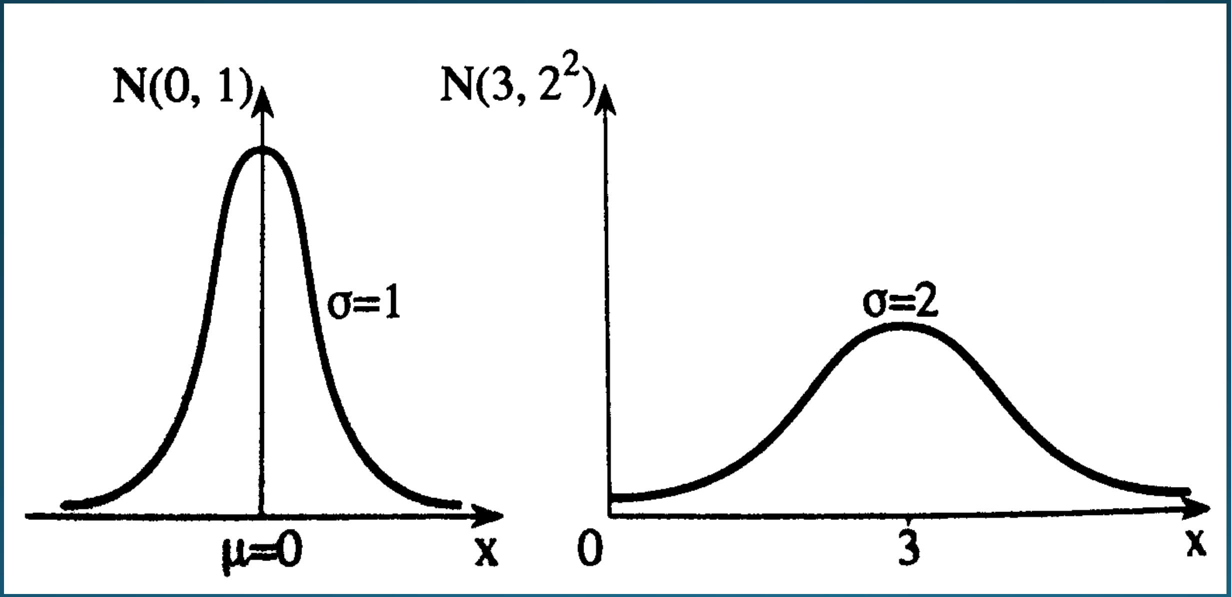
Για τον προσδιορισμό της πιθανότητας κάποιου γεγονότος του οποίου η τυχαία μεταβλητή Z ακολουθεί την κανονική κατανομή ισχύει:

$$P(Z \geq z_a) = a$$



# Συνεχείς Κατανομές
# Κανονική Κατανομή (τυπική)

Οι παραστάσεις της συνάρτησης πυκνότητας πιθανότητας της κατανομής $N(3, 2^2)$ και της τυποποιημένης $N(0,1)$

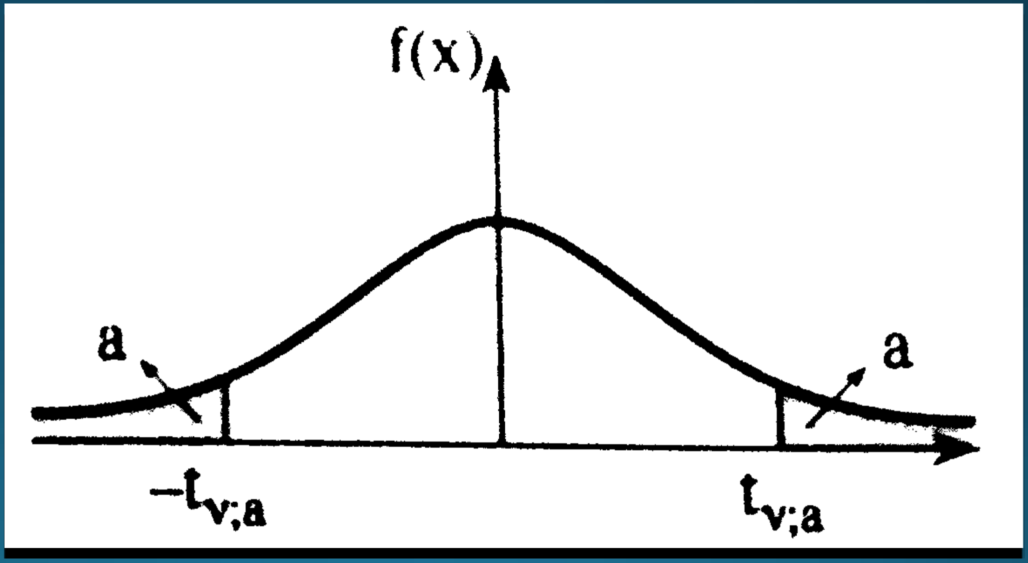


# Συνεχείς Κατανομές
# Student t-Κατανομή

- Συμμετρική ως προς τον άξονα των y και για αριθμό δείγματος n>30 συμπίπτει με την κανονική κατανομή.
- Χρήσιμη για έλεγχο υποθέσεων πχ εάν ισχύει η υπόθεση ότι το 10% των ημερών του έτους παρουσιάζεται ακραία μεταβολή στην ατμόσφαιρα με αποτέλεσμα κάποιο ακραίο φαινόμενο.

- **Συνάρτηση πυκνότητας πιθανότητας – Γραφική Παράσταση**

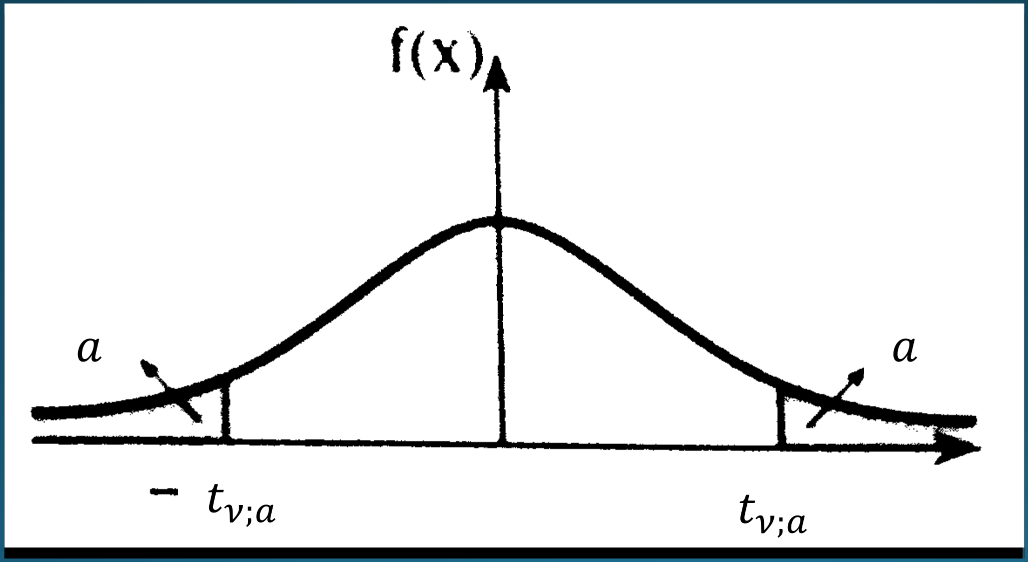
$$\boldsymbol{f}(x) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}$$



# Συνεχείς Κατανομές
# Student t-Κατανομή

Για τον προσδιορισμό της πιθανότητας κάποιου γεγονότος του οποίου η τυχαία μεταβλητή X ακολουθεί την κατανομή Student ισχύει:

$$P\left(X \geq t_{\nu;a}\right) = a$$

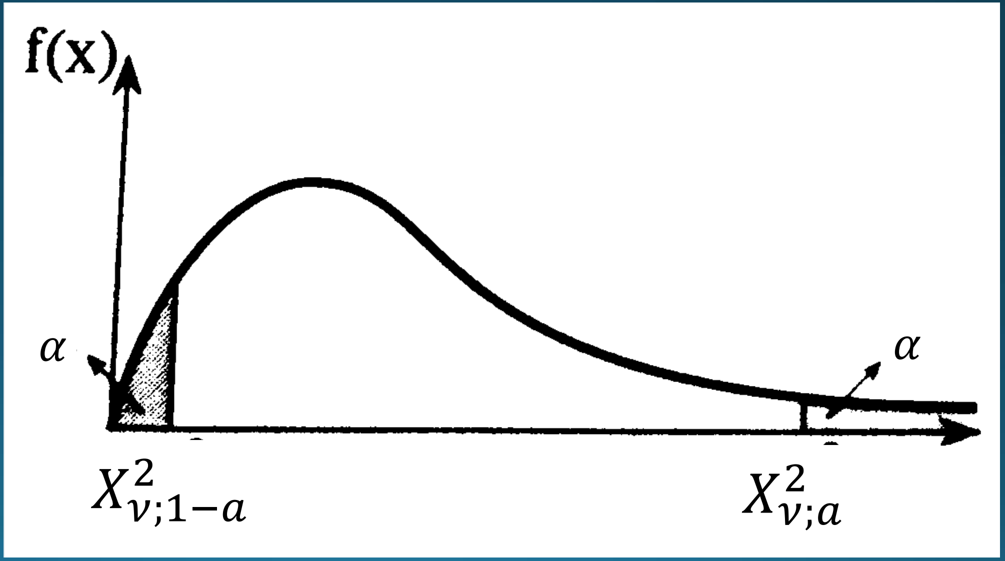


# Συνεχείς Κατανομές
# Κατανομή $X_\nu^2$

- Δεν παρουσιάζει συμμετρίες
- Χρήσιμη για test προσαρμογής και ανεξαρτησίας (θα ασχοληθούμε σε μελλοντικά κεφάλαια)

- **Συνάρτηση πυκνότητας πιθανότητας – Γραφική Παράσταση**

$$\boldsymbol{f}(x) = \left(\frac{1}{2}\right)^{\frac{\nu}{2}} \frac{1}{\Gamma\left(\frac{\nu}{2}\right)} x^{\frac{\nu}{2}-1} e^{-\frac{x}{2}}$$



# Συνεχείς Κατανομές
# Κατανομή$X_{\nu}^{2}$

Για τον προσδιορισμό της πιθανότητας κάποιου γεγονότος του οποίου η τυχαία μεταβλητή Χ ακολουθεί την κατανομή Χ2 ισχύει:

$$P\left(X \geq X_{\nu;a}^{2}\right) = a$$