Στατιστική Επιχειρήσεων Ι

Κεντρικό Οριακό Θεώρημα

Κεντρικό Οριακό Θεώρημα

- Αν X_1, X_2, …, X_N ανεξάρτητες τ.μ. που ακολουθούν την ίδια κατανομή με $E(X_i) = \mu$ και $Var(X_i) = \sigma^2$, τότε **για μεγάλα *N***, ***κατά προσέγγιση***, ισχύει:

$$\boldsymbol{\overline{X} = \frac{\sum_{i=1}^{N} X_i}{N} \sim N\left(\mu, \frac{\sigma^2}{N}\right)}$$

ή ισοδύναμα για το άθροισμα των X_1, X_2, …, X_N

$$\boldsymbol{S_N = \sum_{i=1}^{N} X_i \sim N(N\mu, \mathrm{N}\sigma^2)}$$

Ερμηνεία Κ.Ο.Θ.

- Το Κεντρικό Οριακό Θεώρημα μας βεβαιώνει ότι παρότι δε γνωρίζουμε τη μορφή της κατανομής των ανεξάρτητων τ.μ. του δείγματός μας, εντούτοις, γνωρίζουμε την κατανομή του δειγματικού μέσου τους, όταν τα δείγματά μας έχουν μέγεθος αρκετά μεγάλο

Παράδειγμα 1°

- Μας είναι γνωστό ότι τα μήλα στάρκιν που παράγονται στο οροπέδιο της Τεγέας έχουν μέσο βάρος μ=220 gr με τυπική απόκλιση σ=80 gr
- Στο συσκευαστήριο του τοπικού συνεταιρισμού τα μήλα συσκευάζονται σε κιβώτια των 60 μήλων και προωθούνται στα ψυγεία και την αγορά
- Μπορούμε να υπολογίσουμε ποιο ποσοστό (κατά προσέγγιση) των κιβωτίων περιέχει μήλα με μέσο βάρος μεταξύ 200 και 250 gr;

Προσέγγιση Διωνυμικής κατανομής από την κανονική (Θεώρημα De Moivre – Laplace)

- Αν $X \sim B(n, p)$, τότε για μεγάλα n (θεωρητικά $n \to +\infty$), κατά προσέγγιση,

$$\boldsymbol{X \sim N\big(np, np(1-p)\big)}$$

- Η προσέγγιση είναι ικανοποιητική αν $np \geq 5$ και $n(1\text{-}p) \geq 5$

Παράδειγμα 2ο

- Ο ιδανικός αριθμός πρωτοετών φοιτητών σε ένα πανεπιστημιακό τμήμα είναι 150
- Το τμήμα, γνωρίζοντας από προηγούμενη εμπειρία ότι από τους φοιτητές που γίνονται δεκτοί για εγγραφή μόνο το 30% παρακολουθεί τα μαθήματα, κάνει δεκτούς 450 φοιτητές
- Ποια είναι η πιθανότητα από τους 450 πρωτοετείς φοιτητές να παρακολουθούν τελικά τα μαθήματα περισσότεροι από 150;

Προσέγγιση της κατανομής Poisson από την κανονική

- Αν $X \sim P(\lambda)$, τότε για μεγάλα λ, κατά προσέγγιση,

$$\boldsymbol{X \sim N(\lambda, \lambda)}$$

- Η προσέγγιση είναι ικανοποιητική αν $\lambda \geq 10$

Παράδειγμα 3ο

- Σε μια καλλιέργεια κηπευτικών, έχει παρατηρηθεί ότι ο αριθμός των φυτών που δεν αναπτύσσονται (ξεραίνονται) είναι τ.μ. που ακολουθεί την κατανομή Poisson με μέση τιμή $\lambda=100$ φυτά / καλλιεργητική περίοδο
- Ποια είναι η πιθανότητα σε μια καλλιεργητική περίοδο ο αριθμός των φυτών που δε θα αναπτυχθούν να είναι τουλάχιστον 120;

Κατανομή χ^2 με n βαθμούς ελευθερίας

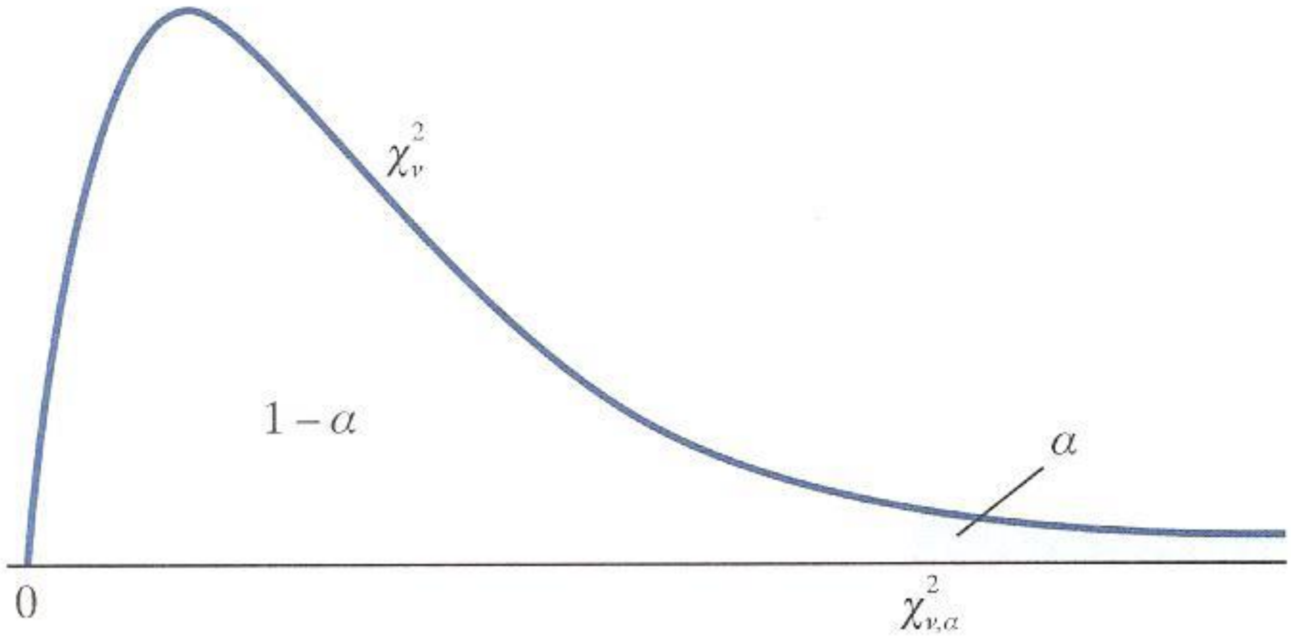
- Αν Z_i, $i = 1, 2 ,\ldots, n$ ανεξάρτητες τυποποιημένες κανονικές τ.μ., τότε η κατανομή

$$\boldsymbol{Z_1^2 + Z_2^2 + \ldots Z_n^2}$$

ονομάζεται κατανομή «χι-τετράγωνο» με n βαθμούς ελευθερίας και συμβολίζεται με $\chi_n{}^2$

Για μεγάλα n προσεγγίζεται ικανοποιητικά από την $\boldsymbol{N(n, 2n)}$

Κατανομή χ^2 με *n* βαθμούς ελευθερίας – καμπύλη



$P(X>\chi^2_{n;a})=a$

Η κατανομή *t* (ή Student) με *n* βαθμούς ελευθερίας

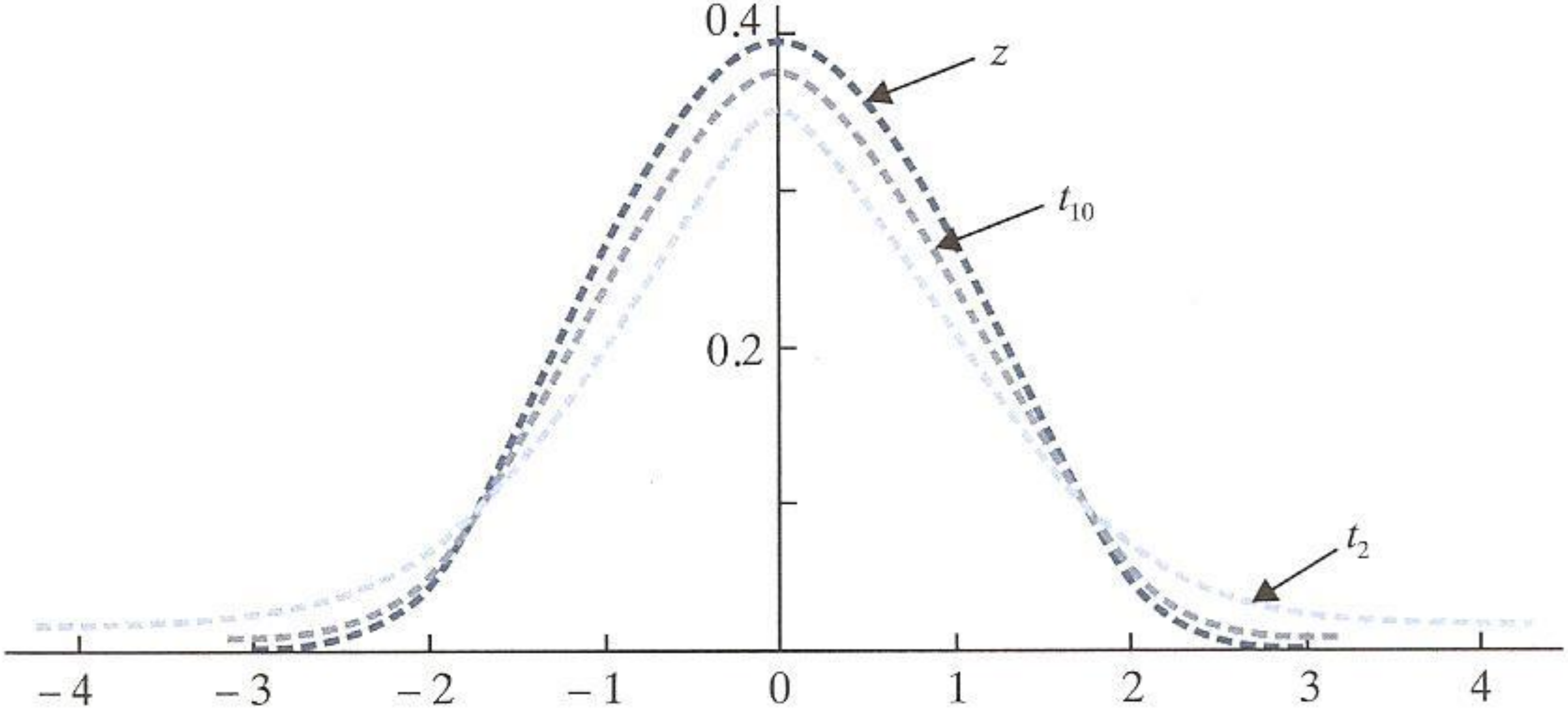
- Αν Z, S_n ανεξάρτητες τ.μ. με $Z \sim N(0,1)$ και $S_n \sim \chi_n^2$, τότε η κατανομή της τ.μ.

$$\frac{\boldsymbol{Z}}{\sqrt{\frac{\boldsymbol{S_n}}{\boldsymbol{n}}}}$$

ονομάζεται κατανομή *t* (ή κατανομή Student) με *n* βαθμούς ελευθερίας και συμβολίζεται με t_n

Για μεγάλα *n* προσεγγίζεται ικανοποιητικά από τη $\mathbf{N}\left(\mathbf{0}, \frac{\boldsymbol{n}}{\boldsymbol{n}-\mathbf{2}}\right)$

Η κατανομή *t* (ή Student) με *n* βαθμούς ελευθερίας – καμπύλη



Η συνάρτηση πυκνότητας της t_ν για $\nu = 2$ και $\nu = 10$ και της $N(0, 1)$.

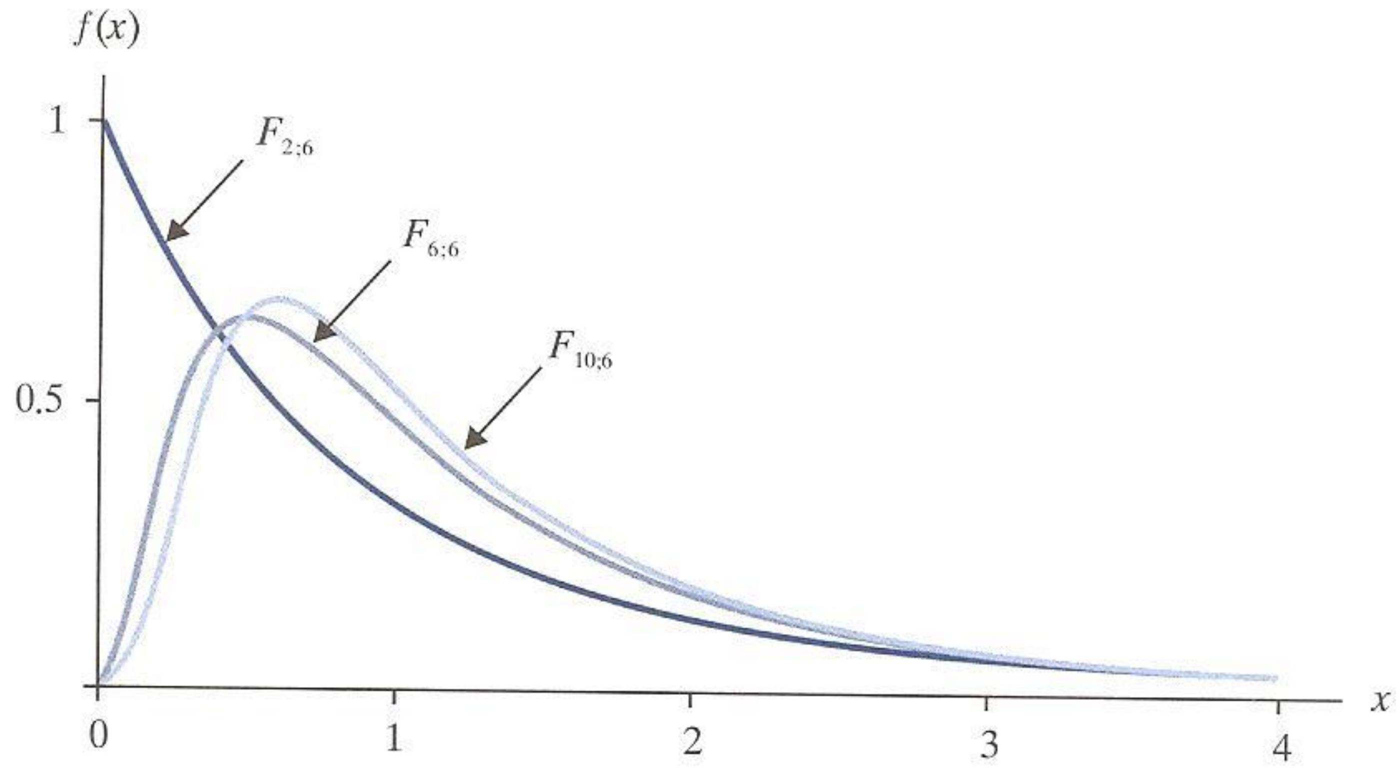
Η κατανομή *F* με *n* και *m* βαθμούς ελευθερίας

- Αν S_m, S_n δύο ανεξάρτητες τ.μ. με $S_n \sim \chi_n^2$ και $S_m \sim \chi_m^2$, τότε η κατανομή της τ.μ.

$$\frac{\frac{S_n}{n}}{\frac{S_m}{m}} = \frac{m}{n} \cdot \frac{S_n}{S_m}$$

ονομάζεται κατανομή *F* με *n* και *m* βαθμούς ελευθερίας και συμβολίζεται με $F_{n;m}$

Η κατανομή *F* με *n* και *m* βαθμούς ελευθερίας – καμπύλη



Η συνάρτηση πυκνότητας της $F_{\nu;m}$ για $\nu = 2,6,10$ και $m = 6$.