

Πιθανότητες & Στατιστική

Το Κεντρικό Οριακό Θεώρημα

Δημήτρης Μητσούδης

Δ.Α. Μητσούδης

Πιθανότητες & Στατιστική - Τμ. Ναυπηγών

1 / 9

Ανεξάρτητες, ισόνομες τ.μ.

Έστω $X_1, X_2, \dots, X_n$ ανεξάρτητες τ.μ. που ακολουθούν την ίδια κατανομή. Τότε οι $X_1, X_2, \dots, X_n$ λέγονται ανεξάρτητες, **ισόνομες** τ.μ. (= ακολουθούν την ίδια κατανομή).

Επίσης καλούνται **τυχαίο δείγμα μεγέθους n** .

(Στα αγγλικά: **independent, identically distributed (i.i.d.)**)

Δ.Α. Μητσούδης

Πιθανότητες & Στατιστική - Τμ. Ναυπηγών

2 / 9

Το κεντρικό οριακό θεώρημα

Το κεντρικό οριακό θεώρημα είναι ένα πολύ σημαντικό θεώρημα των Πιθανοτήτων που εξετάζει πώς συμπεριφέρεται ασυμπτωτικά το άθροισμα πολλών ανεξάρτητων τυχαίων μεταβλητών της μορφής

$$X_1 + X_2 + \dots + X_n,$$

για μεγάλο n (καθώς το $n \rightarrow \infty$).

Δ.Α. Μητσούδης

Πιθανότητες & Στατιστική - Τμ. Ναυπηγών

3 / 9

Το κεντρικό οριακό θεώρημα

Θεώρημα (Κεντρικό οριακό θεώρημα)

Έστω $X_1, X_2, \dots, X_n$ τυχαίο δείγμα από μια συνάρτηση κατανομής, τέτοιο ώστε

$$E[X_i] = \mu \quad \text{και} \quad \text{Var}[X_i] = \sigma^2, \quad i = 1, 2, \dots, n.$$

Τότε, για μεγάλο n , ο αποκαλούμενος δειγματικός μέσος

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$$

ακολουθεί κανονική κατανομή με μέσο μ και διασπορά σ^2/n .

Δηλαδή,

$$\bar{X} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right) \quad \text{και} \quad S_n = X_1 + X_2 + \dots + X_n \sim \mathcal{N}(n\mu, n\sigma^2).$$

Δ.Α. Μητσούδης

Πιθανότητες & Στατιστική - Τμ. Ναυπηγών

4 / 9

Παρατηρήσεις

Η σημασία του κεντρικού οριακού θεωρήματος είναι τεράστια τόσο από θεωρητική όσο και από πρακτική σκοπιά.

- Υποδεικνύει ότι το άθροισμα ενός **μεγάλου** αριθμού ανεξάρτητων τ.μ. ακολουθεί κατά προσέγγιση κανονική κατανομή.

Αυτό έχει εφαρμογή σε πολλές περιπτώσεις όπου ένα τυχαίο αποτέλεσμα προκύπτει ως το άθροισμα ενός μεγάλου αριθμού μικρών αλλά ανεξάρτητων τυχαίων παραγόντων. Π.χ., ο θόρυβος σε πολλά φυσικά ή μηχανικά συστήματα έχει αυτή την ιδιότητα.

- Μας επιτρέπει να υπολογίζουμε συγκεκριμένες πιθανότητες κάνοντας χρήση μόνο του πίνακα της αθροιστικής συνάρτησης της κανονικής κατανομής. Οι υπολογισμοί αυτοί απαιτούν απλά να γνωρίζουμε μέσους και διασπορές.

Στην πράξη ένα μέγεθος δείγματος $n > 30$ αρκεί για να πάρουμε ικανοποιητικές προσεγγίσεις.

Παρατηρήσεις (συνέχεια)

- Στην ειδική περίπτωση όπου οι τ.μ. $X_1, X_2, \dots, X_n$ ακολουθούν **κανονική κατανομή**, τότε το άθροισμα $S_n = X_1 + X_2 + \dots + X_n$ ακολουθεί πάντα κανονική κατανομή και το κεντρικό οριακό θεώρημα ισχύει και για πολύ **μικρές τιμές του n** .

Προσέγγιση με κανονική κατανομή με χρήση του Κ.Ο.Θ.

Έστω $S_n = X_1 + X_2 + \dots + X_n$, όπου οι X_i είναι ανεξάρτητες, ισόνομες τ.μ. με μέσο μ και διασπορά σ^2 . Αν το n είναι μεγάλο, τότε μπορούμε να προσεγγίσουμε την $P(S_n \leq c)$ ωςάν το S_n να ακολουθούσε κανονική κατανομή, εκτελώντας τα ακόλουθα βήματα:

- Υπολογίζουμε τον μέσο $n\mu$ και τη διασπορά $n\sigma^2$ του αθροίσματος S_n .
- Υπολογίζουμε την τυποποιημένη τιμή $z = (c - n\mu)/(\sigma\sqrt{n})$.
- Χρησιμοποιούμε την προσέγγιση

$$P(S_n \leq c) \approx \Phi(z),$$

όπου η $\Phi(\cdot)$ είναι η αθροιστική συνάρτηση της τυποποιημένης κανονικής κατανομής και οι τιμές της δίνονται από το γνωστό πίνακα.

Παράδειγμα

Φορτώνουμε σε ένα αεροπλάνο 100 πακέτα με βάρη που είναι ανεξάρτητες τυχαίες μεταβλητές ομοιόμορφα καταμεμημένες μεταξύ 2 και 22 kg. Υπολογίστε την πιθανότητα το συνολικό βάρος να υπερβεί τα 1310 kg.

Λύση: Θέλουμε να βρούμε την $P(S_{100} > 1310)$, όπου S_{100} είναι το συνολικό βάρος των 100 πακέτων. Αφού κάθε $X_i \sim U(2, 22)$, έχουμε ότι

$$\mu = \frac{2 + 22}{2} = 12, \quad \sigma^2 = \frac{(22 - 2)^2}{12} = 33.33.$$

Υπολογίζουμε την τυποποιημένη τιμή και έχουμε:

$$z = \frac{1310 - 100 \cdot 12}{\sqrt{33.33 \cdot 100}} = 1.91.$$

Από τον πίνακα της τυποποιημένης κανονικής κατανομής έχουμε ότι:

$$P(S_{100} \leq 1310) \approx \Phi(1.91) = 0.9719,$$

επομένως

$$P(S_{100} > 1310) = 1 - P(S_{100} \leq 1310) \approx 1 - 0.9719 = 0.0281,$$

δηλαδή η ζητούμενη πιθανότητα είναι περίπου 3%.