Εκτιμητές Παραμέτρων

Υλοποίηση τυχαίας μεταβλητής

Έστω ότι έχω μια τυχαία μεταβλητή X που περιγράφει κάποιο πείραμα τύχης. Αν το τυχαίο πείραμα εκτελεστεί θα πάρουμε κάποια συγκεκριμένη πραγματική τιμή για την τυχαία μεταβλητή X, π.χ. $5,51621891$. Αυτή η τιμή (δεν θα είναι τυχαία μεταβλητή πλέον) λέγεται μια υλοποίηση της τυχαίας μεταβλητής X.

Αντίστοιχα σε ένα δείγμα όπου έχω n τυχαίες μεταβλητές, αν εκτελεστεί το πείραμα θα πάρω n πραγματικές τιμές (παρατηρήσεις) οπότε θα έχω μια υλοποίηση του τυχαίου δείγματος. Στην πράξη γενικά παρατηρώ ένα μόνο από τα δυνατά δείγματα και ο σκοπός της στατιστικής επαγωγής είναι από n παρατηρήσεις να βγάλει συμπέρασμα για την κατανομή πιθανότητας που παράγει τα τυχαία δείγματα.

Αν μπορούσα να εκτελέσω το πείραμα άπειρες φορές, θα έπαιρνα άπειρα δείγματα, οπότε θα μπορούσα να βρω την κατανομή πιθανότητας που με ενδιαφέρει, π.χ. τον μέσο της κατανομής. Η κατανομή συχνοτήτων από όλες τις εκτελέσεις θα ήταν η κατανομή πιθανότητας που με ενδιαφέρει. Το ερώτημα είναι αφού μπορώ να παρατηρήσω μόνο ένα δείγμα αντί για άπειρα, είναι δυνατόν να βγάλω συμπέρασμα για τον πληθυσμό βάσει ενός μόνο δείγματος; Η απάντηση είναι ναι, και σε αυτό με βοηθάει η έννοια του τυχαίου δείγματος.

Τυχαίο Δείγμα

Έστω τυχαία μεταβλητή X. Αν έχω n τυχαίες μεταβλητές $X_1, X_2, \ldots, X_n$ όπου καθεμία έχει την ίδια κατανομή με τη X και είναι ανεξάρτητες, τότε αυτές οι τυχαίες μεταβλητές συνιστούν ένα τυχαίο δείγμα από την τυχαία μεταβλητή X.

Εκτιμητής παραμέτρου

Δεδομένου ενός τυχαίου δείγματος $X_1, X_2, \ldots, X_n$ από την τυχαία μεταβλητή X, μας ενδιαφέρει να εκτιμήσουμε κάποιες παραμέτρους αυτής της κατανομής, όπως ο μέσος $\mu = E(X)$ ή/και η διακύμανση $\sigma^2 = Var(X)$. Ένας εκτιμητής του μ γενικά θα είναι μια συνάρτηση του δείγματος οπότε θα είναι μια τυχαία μεταβλητή, η οποία προφανώς θα έχει κατανομή την οποία λέμε κατανομή δειγματοληψίας, καθώς μας δίνει εικόνα για την συμπεριφορά του εκτιμητή σε εναλλακτικά δείγματα. Για κάθε διαφορετικό δείγμα ο εκτιμητής θα δίνει μια διαφορετική τιμή, ή αλλιώς εκτίμηση. Καθώς γενικά μπορούμε να κατασκευάσουμε πολλούς εκτιμητές πρέπει να έχουμε κάποια κριτήρια με τα οποία να τους συγκρίνουμε. Σαν συναρτήσεις του τυχαίου δείγματος, θα εξαρτώνται φυσικά από την κατανομή πιθανότητας της X, οπότε η κατανομή πιθανότητας τους θα εξαρτάται και από τις άγνωστες παραμέτρους.

Έστω ένας εκτιμητής $\hat{\theta}(X_1, X_2, \ldots, X_n)$ ή για συντομία $\hat{\theta}$ για μια παράμετρο θ. Σημειώστε ότι για κάποια υλοποίηση του τυχαίου δείγματος, π.χ. $(x_1, x_2, \ldots, x_n)$ η εκτίμηση θα είναι $\hat{\theta}(x_1, x_2, \ldots, x_n)$ και θα είναι κάποια πραγματική τιμή (δεν θα εξαρτάται από άγνωστες παραμέτρους). Σημειώστε ότι για να είναι μια τυχαία μεταβλητή εκτιμητής κάποιας παραμέτρου, πρέπει να μπορεί να υπολογισθεί με βάση τα στοιχεία του δείγματος αποκλειστικά και μόνο, δηλαδή να μην εξαρτάται από άγνωστες παραμέτρους.

Επιθυμητές ιδιότητες εκτιμητή (σε μικρό δείγμα)

1. Αμεροληψία: $E(\hat{\theta}) = \theta$
   Δηλαδή κατά μέσο όρο (για πολλά δείγματα) να δίνει την πραγματική τιμή της παραμέτρου.

2. Αποτελεσματικότητα: $Var(\hat{\theta})$ είναι η μικρότερη δυνατή.
   Δηλαδή χρησιμοποιεί αποτελεσματικά την διαθέσιμη πληροφορία ώστε η διασπορά γύρω από

την τιμή μ να είναι όσο το δυνατόν μικρότερη.

Μεταξύ δύο αμερόληπτων εκτιμητών διαλέγουμε αυτόν με την μικρότερη διακύμανση

Αν πρέπει να συγκρίνουμε μεταξύ δύο μη αμερόληπτων εκτιμητών τότε μπορούμε να χρησιμοποιήσουμε το μέσο τετραγωνικό σφάλμα που λαμβάνει υπόψη την μεροληψία αλλά και την διακύμανση ενός εκτιμητή. Το Μέσο Τετραγωνικό Σφάλμα (Mean Squared Error) ενός εκτιμητή $\hat{\theta}$ για μια παράμετρο θ ορίζεται ως

$$MSE(\hat{\theta}) \quad = \quad E\left[(\hat{\theta}-\theta)^2\right]$$

το οποίο είναι ίσο με

$$MSE(\hat{\theta}) = Var(\hat{\theta}) + \left[E(\hat{\theta}) - \theta\right]^2$$

όπου $E(\hat{\theta}) - \theta$ είναι η μεροληψία του εκτιμητή, ή αλλιώς το σφάλμα μεροληψίας του εκτιμητή.

Αλλά πώς κατασκευάζουμε εκτιμητές; Υπάρχουν διάφορες μέθοδοι όπως η Μέθοδος των Ροπών που θα χρησιμοποιήσουμε.

Η Μέθοδος των Ροπών

Η μέθοδος των ροπών είναι μια μέθοδος κατασκευής εκτιμητών όπου εξισώνουμε την ροπή μιας κατανομή με την αντίστοιχη δειγματική ροπή. Δηλαδή, δεδομένου ενός τυχαίου δείγματος $X_1, X_2, \ldots, X_n$ από την τυχαία μεταβλητή X, αν μας ενδιαφέρει να εκτιμήσουμε την παράμετρο $\mu = E(X)$ που είναι η πρώτη ροπή της κατανομής της X μπορούμε να χρησιμοποιούμε την δειγματική ροπή που είναι η $\frac{1}{n}\sum_{i=1}^{n} X_i$ σαν εκτιμητή της πρώτης ροπής, δηλαδή να πάρουμε $\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

Για παράδειγμα αν έχω ένα τυχαίο δείγμα $X_1, X_2, \ldots, X_n$ από την τυχαία μεταβλητή $X \sim N(\mu, \sigma^2)$ όπου ισχύει ότι $\mu = E(X)$ και $\sigma^2 = E\left[(X - E(X))^2\right]$ ο εκτιμητής των ροπών θα προκύψει εξισώνοντας τις 2 πρώτες ροπές με τις δειγματικές ροπές οπότε θα έχω

$$\begin{aligned} \hat{\mu} &= \frac{1}{n}\sum_{i=1}^{n} X_i \\ \hat{\sigma}^2 &= \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \frac{1}{n}\sum_{i=1}^{n} X_i\right)^2. \end{aligned}$$

Στην πράξη αντικαθιστούμε της αναμενόμενες τιμές με μέσους όρους από το τυχαίο δείγμα.

Κατανομή δειγματοληψίας του δειγματικού μέσου

Όπως είπαμε πιο πάνω έστω τυχαίο δείγμα $X_1, X_2, \ldots, X_n$ από την τυχαία μεταβλητή X για την οποία $E(X) = \mu$ και $Var(X) = \sigma^2$. Ο δειγματικός μέσος ορίζεται ως

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Σημειώστε ότι κάθε φορά που λαμβάνουμε ένα δείγμα, οι τυχαίες μεταβλητές παίρνουν κάποιες τιμές,

$$X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n$$

οπότε για τη συγκεκριμένη υλοποίηση του δείγματος, $x_1, x_2, \ldots, x_n$ ο δειγματικός παίρνει μια συγκεκριμένη τιμή

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

Δηλαδή έχουμε:
$\bar{X}$: εκτιμητής του μ,
$\bar{x}$: εκτίμηση του μ.

Η αναμενόμενη τιμή του $\bar{X}$ θα είναι

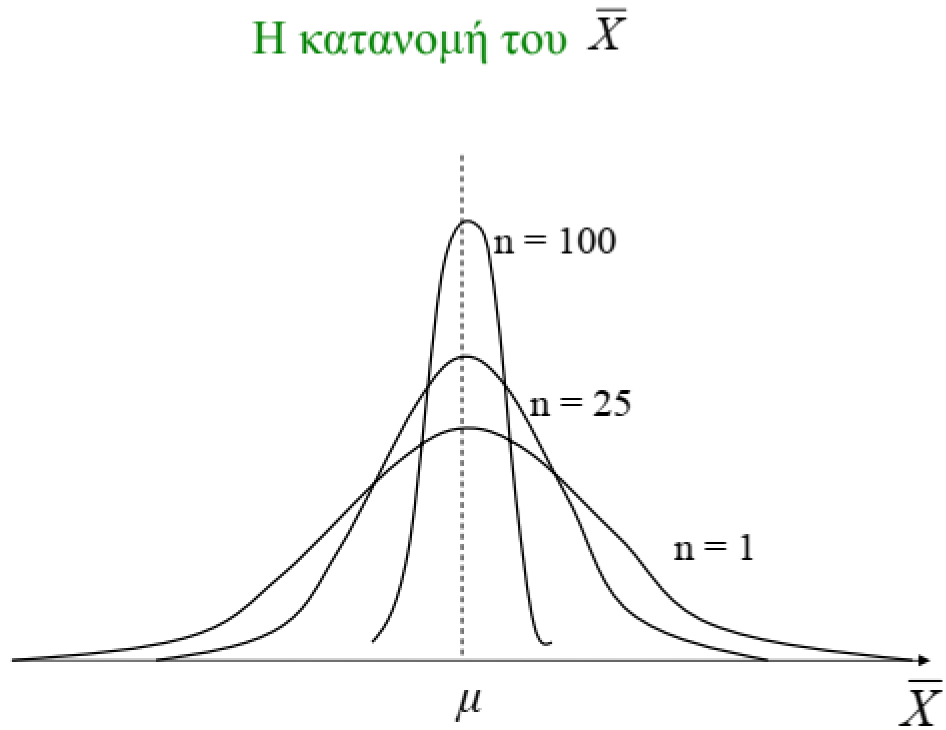
$$E(\bar{X}) = E(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n}\mu = \mu$$

οπότε έχει την ίδια αναμενόμενη τιμή με κάθε τυχαία μεταβλητή X_i που αποτελεί το τυχαίο δείγμα.

Η διακύμανση του $\bar{X}$ είναι

$$Var(\bar{X}) = Var(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n^2}Var(\sum_{i=1}^{n} X_i) = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}$$

οπότε η διακύμανση (διασπορά) του δειγματικού μέσου είναι μικρότερη από την διακύμανση κάθε τυχαίας μεταβλητής ξεχωριστά. Επίσης καθώς $n \to \infty$ έχουμε $Var(\bar{X}) \to 0$, το οποίο σημαίνει ότι μεγαλύτερα δείγματα βοηθούν να προσδιορίσουμε το άγνωστο μ με μεγαλύτερη ακρίβεια, καθώς η κατανομή του $\bar{X}$ θα τείνει να συγκεντρώνεται όλο και πιο κοντά γύρω από το μ.



Έχοντας υπολογίσει την αναμενόμενη τιμή και την διακύμανση του δειγματικού μέσου όμως δεν είναι αρκετό για να μπορώ να υπολογίζω πιθανότητες.

Κατανομή δειγματοληψίας του δειγματικού μέσου για κανονικό πληθυσμό

Αν επιπλέον ξέρουμε ότι η τυχαία μεταβλητή X ακολουθεί την κανονική κατανομή, οπότε $X \sim N(\mu, \sigma^2)$ τότε θα έχουμε:

$$\bar{X} \sim N(\mu, \sigma^2)$$

Άρα μπορούμε να υπολογίσουμε πιθανότητες που αφορούν το δειγματικό μέσο αν οι παράμετροι μ, σ^2 είναι γνωστές.

Παράδειγμα 1: $\mu = 5, \sigma^2 = 1, n = 16$ Σημειώνουμε ότι τυποποιώντας έχουμε $Z := \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} = \frac{\bar{X}-5}{1/\sqrt{16}} \sim N(0,1)$ οπότε μπορώ να υπολογίσω πιθανότητες όπως:

$$P(\bar{X} > 5.5) = P\left(\frac{\bar{X}-5}{1/\sqrt{16}} > \frac{5.5-5}{1/\sqrt{16}}\right) = P(Z > 2) = 1 - P(Z < 2) = 1 - \Phi(2) \simeq 1 - 0.9772 = 0.0228.$$

Παράδειγμα 2: μ **άγνωστο**, $\sigma^2 = 1, n = 16$ Σε αυτή την περίπτωση μπορώ να υπολογίζω πιθανότητες όπως π.χ.

$$P(\bar{X} > \mu + 0.5) = P\left(\frac{\bar{X}-\mu}{1/\sqrt{16}} > \frac{0.5}{1/\sqrt{16}}\right) = P(Z > 2) = 1 - P(Z < 2) = 1 - \Phi(2) \simeq 1 - 0.9772 = 0.0228.$$

Επίσης, καθώς για $Z \sim N(0,1)$ ισχύει (από τα πινακάκια) ότι $P(-2.57 \leq Z \leq 2.57) = 0.99$ έχω γενικά ότι:

$$\begin{aligned}
P(-2.57 \leq Z \leq 2.57) &= 0.99 \\
\Leftrightarrow P(-2.57 \leq \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \leq 2.57) &= 0.99 \\
\Leftrightarrow P\left(-2.57\frac{\sigma}{\sqrt{n}} \leq \bar{X} - \mu \leq 2.57\frac{\sigma}{\sqrt{n}}\right) &= 0.99 \\
\Leftrightarrow P\left(\mu - 2.57\frac{\sigma}{\sqrt{n}} \leq \bar{X} \leq \mu + 2.57\frac{\sigma}{\sqrt{n}}\right) &= 0.99
\end{aligned}$$

οπότε αφού $n = 16$ και $\sigma = 1$ θα έχω

$$P\left(\mu - 0.6425 \leq \bar{X} \leq \mu + 0.6425\right) = 0.99$$

δηλαδή με πιθανότητα 99% θα έχω ότι $\bar{X} = \mu \pm 0.6425$ δηλαδή με μεγάλη πιθανότητα η εκτίμηση $\bar{x}$ που θα πάρω για κάποιο δείγμα που θα παρατηρήσω θα βρίσκεται σε αυτό το διάστημα. Προφανώς αυτό το διάστημα θα επηρεάζεται από το μέγεθος του δείγματος και από το σ^2. Συγκεκριμένα, όσο μεγαλύτερο το δείγμα και όσο μικρότερη η διακύμανση τόσο μικρότερο το διάστημα, άρα θα αυξάνεται και η ακρίβεια της εκτίμησης που θα παίρνω (καθώς θα είναι πιο κοντά στο άγνωστο μ που ψάχνω).

Γενικότερα αν έχω για $Z \sim N(0,1)$ μπορώ να βρω πραγματική τιμή $z_{a/2}$ τέτοια ώστε $P(Z > z_{a/2}) = \frac{a}{2}$ όπου $0 \leq a \leq 1$ οπότε θα έχω γενικά λόγω συμμετρίας ότι

$$\begin{aligned}
P(-z_{a/2} \leq Z \leq z_{a/2}) &= 1 - \alpha \\
\Leftrightarrow P\left(\mu - z_{a/2}\frac{\sigma}{\sqrt{n}} \leq \bar{X} \leq \mu + z_{a/2}\frac{\sigma}{\sqrt{n}}\right) &= 1 - a
\end{aligned}$$

Αλλά γενικά το σ^2 θα είναι άγνωστο οπότε θα πρέπει να βρω εκτιμητή και για αυτό.

Εκτιμητές για την διακύμανση του πληθυσμού σ^2

Αν υποθέσουμε ότι μ γνωστό και σ^2 άγνωστο (πρακτικά δεν συμβαίνει) τότε ένας εκτιμητής για το σ^2 είναι ο παρακάτω:

$$S_\mu^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$$

ο οποίος είναι αμερόληπτος για το σ^2 καθώς

$$\begin{aligned}
E(S_\mu^2) &= E\left[\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2\right] \\
&= \frac{1}{n}\sum_{i=1}^{n}E(X_i - \mu)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}Var(X_i) \\
&= \frac{1}{n}\sum_{i=1}^{n}\sigma^2 \\
&= \sigma^2
\end{aligned}$$

Στην συνηθισμένη περίπτωση όπου το μ είναι άγνωστο, ένας εκτιμητής για την διακύμανση σ^2 είναι ο εκτιμητής των ροπών που ορίζεται ως:

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

Είναι αμερόληπτος ο παραπάνω εκτιμητής; Προσέξτε πρώτα ότι μπορούμε να τον γράψουμε στην ακόλουθη μορφή

$$\begin{aligned}
\hat{\sigma}^2 &= \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(X_i^2 - 2X_i\bar{X} + \bar{X}^2\right) \\
&= \frac{1}{n}\left(\sum_{i=1}^{n}X_i^2 - 2\bar{X}\sum_{i=1}^{n}X_i + \sum_{i=1}^{n}\bar{X}^2\right) \\
&= \frac{1}{n}\left(\sum_{i=1}^{n}X_i^2 - 2\bar{X}n\frac{1}{n}\sum_{i=1}^{n}X_i + n\bar{X}^2\right) \\
&= \frac{1}{n}\left(\sum_{i=1}^{n}X_i^2 - 2\bar{X}n\bar{X} + n\bar{X}^2\right) \\
&= \frac{1}{n}\left(\sum_{i=1}^{n}X_i^2 - n\bar{X}^2\right)
\end{aligned}$$

Οπότε έχουμε

$$E(\hat{\sigma}^2) = \frac{1}{n}E\left(\sum_{i=1}^{n}X_i^2 - n\bar{X}^2\right)$$

$$\begin{aligned} &= \frac{1}{n}\left[E\left(\sum_{i=1}^{n} X_i^2\right) - nE\left(\bar{X}^2\right)\right] \\ &\frac{1}{n}\left[\sum_{i=1}^{n} E(X_i^2) - nE\left(\bar{X}^2\right)\right] \end{aligned}$$

Αλλά έχουμε ότι $Var(X_i) = E(X_i^2) - [E(X_i)]^2 \Leftrightarrow \sigma^2 = E(X_i^2) - \mu^2 \Leftrightarrow E(X_i^2) = \mu^2 + \sigma^2$. Επίσης $Var(\bar{X}) = E(\bar{X}^2) - [E(\bar{X})]^2 \Leftrightarrow \frac{\sigma^2}{n} = E(\bar{X}^2) - \mu^2 \Leftrightarrow E(\bar{X}^2) = \mu^2 + \frac{\sigma^2}{n}$.

Αντικαθιστώντας έχουμε:

$$\begin{aligned} E(\hat{\sigma}^2) &= \frac{1}{n}\left[\sum_{i=1}^{n}(\mu^2 + \sigma^2) - n(\mu^2 + \frac{\sigma^2}{n})\right] \\ &= \frac{1}{n}\left[n(\mu^2 + \sigma^2) - n(\mu^2 + \frac{\sigma^2}{n})\right] \\ &= \frac{1}{n}\left[n\mu^2 + n\sigma^2 - n\mu^2 - \sigma^2\right] \\ &= \frac{n-1}{n}\sigma^2 \end{aligned}$$

άρα δεν είναι αμερόληπτος εκτιμητής. Γι' αυτό και συνήθως χρησιμοποιούμε τον ακόλουθο εκτιμητή, τον οποίο ονομάζουμε δειγματική διακύμανση:

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

ο οποίος θα είναι αμερόληπτος.

Κατανομή του $\frac{\bar{X}-\mu}{S/\sqrt{n}}$ για κανονικό πληθυσμό

Αφού έχουμε βρει τρόπο να εκτιμούμε την διακύμανση σ^2 το πρόβλημα τώρα είναι πώς μπορώ να υπολογίζω πιθανότητες που αφορούν τον δειγματικό μέσο όταν έχω τυχαίο δείγμα $X_1, X_2, \ldots, X_n$ από μία τυχαία μεταβλητή $X \sim N(\mu, \sigma^2)$ όπου μ και σ^2 άγνωστα, αλλά με ενδιαφέρει το μ. Σημειώστε ότι μπορώ να απαντήσω αυτό το ερώτημα όπως και πριν αρκεί να ξέρω την κατανομή του

$$T := \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

όπου αντικαθιστώ το άγνωστο σ με την ρίζα του εκτιμητή της δειγματικής διακύμανσης $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$. Καθώς το S είναι τυχαία μεταβλητή η κατανομή του παραπάνω λόγου δεν θα είναι τυποποιημένη κανονική κατανομή όπως πριν.

Έχει αποδειχθεί ότι $T \sim t_{n-1}$ δηλαδή ο παραπάνω λόγος ακολουθεί την κατανομή t (Student's t) με παράμετρο (βαθμοί ελευθερίας) $n-1$. Αυτό προκύπτει καθώς αποδεικνύεται ότι $\sum_{i=1}^{n}\frac{(X_i-\bar{X})^2}{\sigma^2} \sim \chi^2(n-1)$. Είναι εύκολο να δείτε ότι εξ' ορισμού $\sum_{i=1}^{n}\frac{(X_i-\mu)^2}{\sigma^2} = \sum_{i=1}^{n} Z_i^2 \sim \chi^2(n)$ καθώς $Z_i \sim N(0,1)$, οπότε το προηγούμενο αποτέλεσμα στην ουσία δείχνει ότι χρησιμοποιώντας το $\bar{X}$ αντί του μ μειώνουμε τους βαθμούς ελευθερίας κατά έναν, με άλλα λόγια χάνουμε μία παρατήρηση για να εκτιμήσουμε το μ.

Επομένως, καθώς η Student t είναι γνωστή κατανομή μπορώ να βρω πραγματική τιμή $t_{n-1,a/2}$ τέτοια ώστε $P(T > t_{n-1,a/2}) = a/2$ και αφού είναι συμμετρική γύρω από το 0 θα έχω:

$$\begin{aligned} P(-t_{n-1,a/2} \le T \le t_{n-1,a/2}) &= 1-a \\ \Leftrightarrow P\left(\mu - t_{n-1,a/2}\frac{S}{\sqrt{n}} \le \bar{X} \le \mu + t_{n-1,a/2}\frac{S}{\sqrt{n}}\right) &= 1-a \end{aligned}$$

Επιθυμητές ιδιότητες εκτιμητή (σε μεγάλο δείγμα)

1. Συνέπεια: $\hat{\theta}_n \to \theta$ καθώς $n \to \infty$
   Καθώς το μέγεθος του δείγματος μεγαλώνει ($n \to \infty$) ο εκτιμητής τείνει να είναι πιο κοντά στην παράμετρο που εκτιμά.

2. Ασυμπτωτική κανονικότητα: $\frac{\hat{\theta}_n - \theta}{Var(\hat{\theta}_n)} \sim N(0,1)$ καθώς $n \to \infty$
   Καθώς το μέγεθος του δείγματος μεγαλώνει ($n \to \infty$) η κατανομή του (τυποποιημένου) εκτιμητή πλησιάζει όλο και περισσότερο στην κανονική κατανομή.

Νόμος των Μεγάλων Αριθμών (Συνέπεια του Δειγματικού Μέσου) Αν $X_1, X_2, X_3, \ldots$ άπειρη ακολουθία ανεξάρτητων τυχαίων μεταβλητών όπου καθεμία έχει ίδια κατανομή με μια τυχαία μεταβλητή X όπου $E(X) = \mu$ (όπου $\mu < \infty$), τότε

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to E(X) = \mu, \quad \text{καθώς } n \to \infty$$

δηλαδή ο δειγματικός μέσος συγκλίνει στο μέσο του πληθυσμού!

Κεντρικό Οριακό Θεώρημα Αν $Var(X) = \sigma^2$ (όπου $\sigma^2 < \infty$) τότε

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1), \quad \text{καθώς } n \to \infty$$

δηλαδή για μεγάλο n προσεγγιστικά θα ισχύει ότι

$$\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$$

Επιπλέον, καθώς το σ^2 είναι γενικά άγνωστο θα έχουμε ότι για οποιονδήποτε συνεπή εκτιμητή του σ^2 όπως $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i$ (εκτιμητής της μεθόδου των ροπών) ή $\hat{\sigma}^2 = S^2 = \frac{1}{n-1}\sum_{i=1}^{n} X_i$ (δειγματική διακύμανση θα έχουμε

$$\frac{\bar{X} - \mu}{\hat{\sigma}/\sqrt{n}} \sim N(0,1), \quad \text{καθώς } n \to \infty$$

δηλαδή, για μεγάλο δείγμα, μπορούμε προσεγγιστικά να βγάζουμε την κατανομή του $\bar{X}$ όπως στην περίπτωση που έχουμε κανονικό πληθυσμό και γνωστή διακύμανση, χρησιμοποιώντας την εκτίμηση της διακύμανσης!