

ΔΙΑΣΤΗΜΑΤΑ ΕΜΠΙΣΤΟΣΥΝΗΣ

Έστω τυχαίο δείγμα n παρατηρήσεων από πληθυσμό του οποίου η κατανομή εξαρτάται από μία ή περισσότερες παραμέτρους, π.χ. μ .

Επειδή σε κάθε δείγμα αναμένεται διαφορετική τιμή του μ , είναι προτιμότερο να επιδιώκεται ο καθορισμός ενός διαστήματος εμπιστοσύνης, δηλαδή ενός διαστήματος π.χ. $\beta < \mu < \gamma$, στο οποίο περιέχεται η άγνωστη παράμετρος με πολύ μεγάλη πιθανότητα.

➤ Ορισμός: Έστω $X_1, X_2, \dots, X_n$ τυχαίο δείγμα μεγέθους n και θ η άγνωστη παράμετρος του πληθυσμού. Τότε για συντελεστή εμπιστοσύνης $1-\alpha$ το $100(1-\alpha)\%$ διάστημα εμπιστοσύνης ορίζεται από τη σχέση,

$$P(\beta < \theta < \gamma) = 1 - \alpha$$

Διάστημα Εμπιστοσύνης για τη Μέση Τιμή

A) Κανονικός πληθυσμός, διασπορά γνωστή (Δείγματα μικρά ή μεγάλα)

Έστω τυχαίο δείγμα μεγέθους n ενός πληθυσμού που ακολουθεί κανονική κατανομή. Τότε, η κατανομή του δειγματικού μέσου είναι επίσης κανονική. Ειδικότερα,

➤ Αν, $X \rightarrow N(\mu, \sigma^2)$

➤ Τότε, $\bar{X} \rightarrow N(\mu, \sigma^2 / n)$

ΔΙΑΣΤΗΜΑΤΑ ΕΜΠΙΣΤΟΣΥΝΗΣ

➤ Ορίζοντας, $Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \rightarrow N(0,1)$

Ζητάμε να βρούμε τον αριθμό b , έτσι ώστε στο διάστημα $[-b, b]$ η πιθανότητα $P\left(-b \leq \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \leq b\right)$ να είναι π.χ ίση με 95%

➤ Δηλαδή, $P\left(-b \leq \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \leq b\right) = 0.95$

$$\Rightarrow \Phi(b) - \Phi(-b) = 0.95 \quad \begin{matrix} \Phi(-b) = \\ 1 - \Phi(b) \end{matrix} \Rightarrow \Phi(b) = 0.975 \Rightarrow b = 1.96$$

ΔΙΑΣΤΗΜΑΤΑ ΕΜΠΙΣΤΟΣΥΝΗΣ

➤ Άρα,
$$P\left(-1.96 \leq \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \leq 1.96\right) = 0.95$$

➤ Τελικά,
$$P\left(\bar{X} - 1.96 \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + 1.96 \frac{\sigma}{\sqrt{n}}\right) = 0.95$$

Δηλαδή, σε διαδοχικές δειγματοληψίες με δείγματα μεγέθους n στις 95% των περιπτώσεων, το μ θα βρίσκεται σε ένα διάστημα $1.96 \sigma / \sqrt{n}$ μονάδες γύρω από τον δειγματικό μέσο.

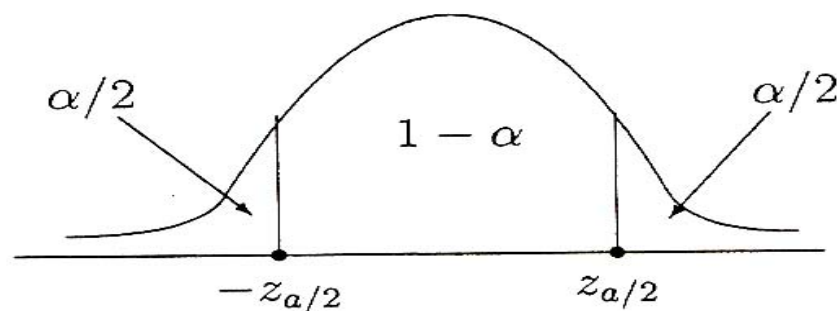
➤ Ορισμός: Ένα $100(1-\alpha)\%$ αμφίπλευρο διάστημα εμπιστοσύνης για το μ , για κανονικό πληθυσμό, γνωστή διασπορά, ανεξαρτήτως μεγέθους δείγματος, είναι

$$\bar{X} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}}, \quad \bar{X} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

ΔΙΑΣΤΗΜΑΤΑ ΕΜΠΙΣΤΟΣΥΝΗΣ

Γενικά, εάν ισχύει $Z \sim N(0,1)$ και για συντελεστή εμπιστοσύνης χρησιμοποιούμε το $1-\alpha$, τότε συμβολίζουμε $z_{\alpha/2}$ εκείνο το σημείο για το οποίο $P(Z > z_{\alpha/2}) = \alpha / 2$

$1 - \alpha$	0,80	0,85	0,90	0,95	0,99
α	0,20	0,15	0,10	0,05	0,01
$\alpha/2$	0,10	0,075	0,05	0,025	0,005
$1 - \alpha/2$	0,90	0,925	0,95	0,975	0,995
$z_{\alpha/2}$	1,28	1,44	1,645	1,96	2,58



Σχήμα 4.2: $100(1 - \alpha)\%$ διάστημα εμπιστοσύνης

Β) Κανονικός πληθυσμός, διασπορά άγνωστη

(Δείγματα μεγάλα, $n > 30$)

Στην περίπτωση αυτή αντικαθιστούμε το σ με τη δειγματική τυπική απόκλιση

- Ορισμός: Ένα $100(1-\alpha)\%$ αμφίπλευρο διάστημα εμπιστοσύνης για το μ , για κανονικό πληθυσμό, άγνωστη διασπορά, είναι

$$\bar{X} - z_{\alpha/2} \frac{s}{\sqrt{n}}, \quad \bar{X} + z_{\alpha/2} \frac{s}{\sqrt{n}}$$

$$\mu \in P\left(\bar{X} - z_{\alpha/2} \frac{s}{\sqrt{n}} \leq \mu \leq \bar{X} + z_{\alpha/2} \frac{s}{\sqrt{n}}\right) = 1 - \alpha$$

ΔΙΑΣΤΗΜΑΤΑ ΕΜΠΙΣΤΟΣΥΝΗΣ

Παράδειγμα 1. Εταιρία ενδιαφέρεται να εκτιμήσει το μέσο χρόνο εξυπηρέτησης κατ' οίκον, τον προηγούμενο χρόνο. Τυχαίο δείγμα 200 μετρήσεων στο αντίστοιχο χρονικό διάστημα, έδωσε μέση τιμή 23 λεπτά και τυπική απόκλιση 2,7 λεπτά. Να βρεθεί διάστημα εμπιστοσύνης 95% για τη μέση τιμή του χρόνου παράδοσης του περασμένου έτους.

Παράδειγμα 2. Εταιρία συγκοινωνιών ενδιαφέρεται να εκτιμήσει το μέσο χρόνο αναμονής των πελατών σε δοσμένη στάση στο κέντρο της πόλης. Τυχαίο δείγμα 120 μετρήσεων, έδωσε μέση τιμή 17 λεπτά και τυπική απόκλιση 7 λεπτά. Να βρεθεί διάστημα εμπιστοσύνης 99% για τη μέση τιμή του χρόνου αναμονής.παράδοσης του περασμένου έτους.

Διάστημα Εμπιστοσύνης για τη διαφορά δύο πληθυσμιακών Μέσων Τιμών

A) Κανονικός πληθυσμός, διασπορές γνωστές
(Δείγματα ανεξάρτητα)

- Ορισμός: Ένα $100(1-\alpha)\%$ διάστημα εμπιστοσύνης για τη διαφορά $\mu_1 - \mu_2$, για κανονικούς πληθυσμούς, και γνωστές διασπορές σ_1^2, σ_2^2 είναι το,

$$\bar{X}_1 - \bar{X}_2 \pm z_{\alpha/2} \sqrt{s_1^2 / n_1 + s_2^2 / n_2}$$

ΔΙΑΣΤΗΜΑΤΑ ΕΜΠΙΣΤΟΣΥΝΗΣ

B) Κανονικός πληθυσμός, διασπορές άγνωστες (Δείγματα ανεξάρτητα, μεγάλα $n > 30$)

Στην περίπτωση αυτή αντικαθιστούμε το σ με τη
δειγματική τυπική απόκλιση

- Ορισμός: Ένα $100(1-\alpha)\%$ διάστημα εμπιστοσύνης για τη διαφορά $\mu_1 - \mu_2$, για κανονικούς πληθυσμούς, και άγνωστές διασπορές είναι το,

$$\bar{X}_1 - \bar{X}_2 \pm z_{\alpha/2} \sqrt{\sigma_1^2 / n_1 + \sigma_2^2 / n_2}$$

ΔΙΑΣΤΗΜΑΤΑ ΕΜΠΙΣΤΟΣΥΝΗΣ

Παράδειγμα 3. Ερευνητής ενδιαφέρεται να εκτιμήσει τη διαφορά των επιπέδων του ορικού οξέως του ορού, μεταξύ ασθενών με μογκολισμό ή όχι. Σε ψυχιατρική κλινική από τυχαίο δείγμα 12 ατόμων με μογκολισμό υπολογίστηκε ότι $\bar{X}_1 = 4,5mg / 100ml$ ενώ σε γενικό νοσοκομείο σε τυχαίο δείγμα 15 ατόμων υπολογίστηκε ότι $\bar{X}_1 = 3,4mg / 100ml$. Να βρεθεί διάστημα εμπιστοσύνης 95% για τη διαφορά $\mu_1 - \mu_2$. Θεωρείστε ότι οι τιμές του ορικού οξέως στον ορό ακολουθούν κανονική κατανομή με διασπορές 1.

Παράδειγμα 4. Για την σύγκριση της οικονομικής κατάστασης των υπαλλήλων δύο εταιριών μελετήθηκε δείγμα 85 ατόμων από την πρώτη εταιρία με μέσο όρο οικογενειακού εισοδήματος 70000ευρώ και δείγμα 70 ατόμων από την δεύτερη εταιρία με μέσο εισόδημα 56000ευρώ. Αν τα αντίστοιχα δείγματα παρουσιάζουν τυπικές αποκλίσεις 7000 και 5500 ευρώ αντίστοιχα να υπολογίσετε διάστημα εμπιστοσύνης 99% για τη διαφορά $\mu_1 - \mu_2$

Προσδιορισμός του μεγέθους του δείγματος στην εκτίμηση της μέσης τιμής (Κανονικός πληθυσμός, διασπορά γνωστή)

- Πρόβλημα: Για ποια τιμή του μεγέθους n ενός δείγματος η απόκλιση $|\mu - \bar{X}|$ είναι μικρότερη από δοσμένο αριθμό d με πιθανότητα $1-\alpha$; Δηλαδή,

$$P(|\mu - \bar{X}| \leq d) = 1 - \alpha$$

- Ισοδύναμα

$$P(\bar{X} - d \leq \mu \leq \bar{X} + d) = 1 - \alpha$$

- Συγκρίνοντας, $d = z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \Rightarrow n = \left(\frac{\sigma z_{\alpha/2}}{d} \right)^2$

ΔΙΑΣΤΗΜΑΤΑ ΕΜΠΙΣΤΟΣΥΝΗΣ

Παράδειγμα 5. Ερευνητής ενδιαφέρεται να εκτιμήσει τη μέση τιμή ενός πληθυσμού ο οποίος είναι γνωστό ότι παρουσιάζει διασπορά 20 μονάδων. Ποιο μέγεθος δείγματος πρέπει να πάρει ώστε να είναι σίγουρος ότι το σφάλμα στην εκτίμηση της μέσης τιμής δε θα είναι μεγαλύτερο από 2,5, με πιθανότητα
α) 95%, β) 99%;