Κεφάλαιο 2

ΕΚΤΙΜΗΣΗ ΠΑΡΑΜΕΤΡΩΝ

Οι στατιστικές δείγματος που υπολογίζονται από τα δεδομένα που έχουν συλλεχθεί, όπως η δειγματική μέση τιμή $\bar{x}$ και η δειγματική διασπορά s^2, χρησιμοποιούνται για την *εκτίμηση* των σχετικών παραμέτρων πληθυσμού (μ και σ^2 αντίστοιχα).

Γενικά από τις παρατηρήσεις του δείγματος μπορούμε να υπολογίσουμε τη **σημειακή εκτίμηση** (point estimation) και την **εκτίμηση διαστήματος** (interval estimation) της παραμέτρου μιας τ.μ..

2.1 Σημειακή Εκτίμηση

Η *σημειακή εκτίμηση* μιας παραμέτρου είναι η στατιστική που υπολογίζουμε από το δείγμα, δηλαδή είναι μια τιμή, που υπολογίζεται με βάση τα δεδομένα του δείγματος και αντιπροσωπεύει την πραγματική τιμή της σχετικής παράμετρου του πληθυσμού. Για παράδειγμα, ο μέσος όρος των 25 τιμών αντοχής θραύσης που μετρήθηκαν σε δείγμα 25 σκυροδεμάτων αποτελεί μια σημειακή εκτίμηση της μέσης αντοχής θραύσης σκυροδέματος (δες Παράδειγμα 1.2).

Έστω X μια τ.μ. με αθροιστική συνάρτηση κατανομής $F_X(x;\theta)$ ή απλά $F(x;\theta)$ που εξαρτάται από την παράμετρο θ την οποία θέλουμε να εκτιμήσουμε. Έστω ακόμα ότι έχουμε παρατηρήσεις $\{x_1, \ldots, x_n\}$ της X από ένα δείγμα μεγέθους n. Τότε η σημειακή εκτίμηση της θ δίνεται από τη συνάρτηση $g(x_1, \ldots, x_n)$ των τιμών του δείγματος που λέγεται **εκτιμήτρια συνάρτηση**. Η **εκτιμήτρια** (estimator) της θ από το δείγμα είναι $\hat{\theta} = g(x_1, \ldots, x_n)$.

Επειδή οι παρατηρήσεις $\{x_1, \ldots, x_n\}$ αλλάζουν κάθε φορά που μελετάμε διαφορετικό δείγμα μεγέθους n, μπορούμε να υποθέσουμε ότι οι παρατηρήσεις $\{x_1, \ldots, x_n\}$ είναι τιμές των τ.μ. $\{X_1, \ldots, X_n\}$, που είναι ανεξάρτητες μεταξύ τους κι ακολουθούν την ίδια κατανομή $F(x;\theta)$. Άρα η παράμετρος $\hat{\theta}$ είναι συνάρτηση αυτών των τ.μ.. Για ευκολία θα χρησιμοποιούμε το συμβολισμό $\{x_1, \ldots, x_n\}$ και θεωρητικά (εννοώντας τις τ.μ. $\{X_1, \ldots, X_n\}$) και πρακτικά (εννοώντας τις παρατηρούμενες αριθμητικές τιμές αυτών των τ.μ.).

Είναι φανερό ότι για διαφορετικά δείγματα (διαφορετικές τιμές $\{x_1, \ldots, x_n\}$) η εκτιμήτρια συνάρτηση της παραμέτρου $\hat{\theta}$ παίρνει διαφορετικές τιμές, δηλαδή η $\hat{\theta}$ είναι η ίδια τ.μ. με κάποια κατανομή κι έχει μέση τιμή $\mu_{\hat{\theta}} = \mathrm{E}(\hat{\theta})$ και διασπορά $\sigma^2_{\hat{\theta}} = \mathrm{Var}(\hat{\theta})$.

Δύο σημαντικές παράμετροι μιας τ.μ. X που θέλουμε να εκτιμήσουμε είναι η μέση τιμή μ κι η διασπορά σ^2.

Εκτίμηση μέσης τιμής Είναι φυσικό ως εκτιμήτρια της μ να ορίσουμε τη *δειγματική μέση τιμή* που ορίσαμε στην Παράγραφο 1.2.1

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i. \tag{2.1}$$

Εκτίμηση διασποράς Όπως για τη μέση τιμή έτσι και για τη διασπορά σ^2 η εκτιμήτρια είναι η *δειγματική διασπορά* που ορίσαμε στην Παράγραφο 1.2.2 ως

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2. \tag{2.2}$$

Βέβαια μπορεί κάποιος να ορίσει την εκτιμήτρια της σ^2 ως

$$\tilde{s}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2. \tag{2.3}$$

Οι εκτιμήτριες s^2 και $\tilde{s}^2$ διαφέρουν μόνο ως προς το συντελεστή του αθροίσματος ($\frac{1}{n-1}$ και $\frac{1}{n}$ αντίστοιχα). Για μεγάλο n οι δύο εκτιμήτριες συγκλίνουν στην ίδια τιμή. Αναπτύσσοντας τα τετράγωνα της (2.2) έχουμε τον ισοδύναμο τύπο (όμοια για την (2.3), δες επίσης την (1.5) στην Παράγραφο 1.2.2)

$$s^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right), \tag{2.4}$$

που συνήθως χρησιμοποιούμε στους υπολογισμούς.

2.1.1 Κριτήρια καλών εκτιμητριών

Παραπάνω ορίσαμε κάπως αυθαίρετα εκτιμήτριες της μέσης τιμής μ και της διασποράς σ^2 χωρίς να γνωρίζουμε αν είναι 'καλές' εκτιμήτριες ή όχι. Γενικά όταν ορίζουμε μια εκτιμήτρια $\hat{\theta}$ κάποιας παραμέτρου θ θέλουμε να ελέγξουμε αν είναι κατάλληλη και γι αυτό θέτουμε κάποια κριτήρια ή ιδιότητες που πρέπει να πληρεί μια 'καλή' εκτιμήτρια. Παρακάτω περιγράφονται ορισμένες επιθυμητές ιδιότητες μιας εκτιμήτριας $\hat{\theta}$.

Αμεροληψία Η $\hat{\theta}$ είναι αμερόληπτη (unbiased) αν η μέση τιμή της είναι ίση με την παράμετρο θ, δηλαδή αν ισχύει

$$\mathrm{E}(\hat{\theta}) = \theta.$$

Αλλιώς λέγεται μεροληπτική με μεροληψία

$$b(\hat{\theta}) = \mathrm{E}(\hat{\theta}) - \theta.$$

Παράδειγμα 2.1 *Η δειγματική μέση τιμή $\bar{x}$ είναι αμερόληπτη εκτιμήτρια της μέσης τιμής μ μιας τ.μ. X ενός πληθυσμού.*

Έχουμε $\theta = \mu$ και $\hat{\theta} = \bar{x}$ και θέλουμε να δείξουμε ότι $\boldsymbol{E}(\bar{x}) = \mu$. Αυτό προκύπτει ως

$$\boldsymbol{E}(\bar{x}) = \boldsymbol{E}\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{E}(x_i) = \frac{1}{n}\sum_{i=1}^{n}\mu = \mu.$$

Παράδειγμα 2.2 *Για την εκτίμηση της διασποράς σ^2 μιας τ.μ. X ενός πληθυσμού μπορεί να δειχθεί ότι:*

1. *Η δειγματική διασπορά s^2 είναι αμερόληπτη εκτιμήτρια της σ^2, δηλαδή ισχύει*
$$E(s^2) = \sigma^2.$$

2. *Η δειγματική διασπορά $\tilde{s}^2$ που ορίστηκε στη (2.3) είναι μεροληπτική εκτιμήτρια της σ^2 με μεροληψία*
$$b(\tilde{s}^2) = -\frac{\sigma^2}{n}.$$

Συνέπεια Η ιδιότητα αυτή ορίζει πως όσο αυξάνει το μέγεθος του δείγματος τόσο μεγαλώνει η πιθανότητα η εκτίμηση να είναι 'κοντά' στην πραγματική τιμή της παραμέτρου, όπου το 'κοντά' σημαίνει ότι η διαφορά της εκτιμούμενης από την πραγματική τιμή της παραμέτρου είναι μικρότερη από κάποια αυθαίρετα μικρή απόσταση ϵ. Δηλαδή η $\hat{\theta}$ είναι συνεπής (consistent) αν ισχύει
$$\mathrm{P}(|\theta - \hat{\theta}| \leq \epsilon) \to 1 \quad \text{όταν} \quad n \to \infty,$$
όπου ϵ είναι αυθαίρετα μικρός θετικός αριθμός.

Παράδειγμα 2.3 *Η εκτιμήτρια $\bar{x}$ της μέσης τιμής μ μιας τ.μ. X είναι συνεπής. Αν όμως αντί της δειγματικής μέσης τιμής διαλέξουμε σαν εκτιμήτρια της μ τον αριθμητικό μέσο της μικρότερης και μεγαλύτερης τιμής του δείγματος*
$$x_d = \frac{x_{min} + x_{max}}{2},$$
τότε μπορεί να δειχθεί ότι η εκτιμήτρια x_d δεν είναι συνεπής εκτιμήτρια της μ.

Αποτελεσματικότητα Η αποτελεσματικότητα αναφέρεται στη διασπορά της εκτιμήτριας και δίνεται συγκριτικά. Μια εκτιμήτρια $\hat{\theta}_1$ της θ είναι πιο αποτελεσματική (effective) από μια άλλη εκτιμήτρια $\hat{\theta}_2$ αν έχει μικρότερη διασπορά, $\sigma^2_{\hat{\theta}_1} < \sigma^2_{\hat{\theta}_2}$.

Παράδειγμα 2.4 *Η δειγματική μέση τιμή $\bar{x}$ και η x_d είναι δύο εκτιμήτριες της μέσης τιμής μ κι έχουν διασπορές $\sigma^2_{\bar{x}}$ και $\sigma^2_{x_d}$ αντίστοιχα. Μπορεί να δειχθεί ότι $\sigma^2_{\bar{x}} < \sigma^2_{x_d}$ κι άρα η εκτιμήτρια $\bar{x}$ είναι πιό αποτελεσματική από τη x_d.*

Επάρκεια Μια εκτιμήτρια της παραμέτρου θ είναι επαρκής (adequate) όταν χρησιμοποιεί όλη την πληροφορία από το δείγμα που σχετίζεται με τη θ.

Παράδειγμα 2.5 *Η δειγματική μέση τιμή $\bar{x}$, εκτιμήτρια της μέσης τιμής μ μιας τ.μ. X, είναι επαρκής γιατί χρησιμοποιεί όλες τις παρατηρήσεις που μετρήθηκαν στο δείγμα, ενώ η x_d δεν είναι επαρκής γιατί χρησιμοποιεί μόνο δύο τιμές των παρατηρήσεων του δείγματος (x_{min} και x_{max}).*

Παρατηρήσεις

Από τα παραπάνω παραδείγματα βλέπουμε πως οι εκτιμήτριες, $\bar{x}$ για την παράμετρο μ και s^2 για την παράμετρο σ^2, που ορίσαμε αυθαίρετα, πληρούν όλες τις τέσσερις ιδιότητες και είναι καλές εκτιμήτριες.

Στον ορισμό των εκτιμητριών $\bar{x}$ και s^2 δεν κάναμε κάποια υπόθεση για την κατανομή της τ.μ. X κι άρα μπορούμε να τις χρησιμοποιήσουμε για οποιαδήποτε τ.μ. X που παρατηρούμε.

Στη συνέχεια θα υποθέσουμε πως η κατανομή της X είναι γνωστή (ως προς τη γενική μορφή της) αλλά δεν είναι γνωστή κάποια παράμετρο θ της κατανομής και θα δούμε πως μπορούμε γενικά να υπολογίσουμε την εκτιμήτρια της θ.

2.1.2 Μέθοδος υπολογισμού της σημειακής εκτίμησης

Υποθέτουμε ότι η τ.μ. X έχει κάποια γνωστή κατανομή, δηλαδή γνωρίζουμε τη γενική μορφή της αθροιστικής συνάρτησης κατανομής $F(x;\theta)$ και της $f(x;\theta)$, που είναι η συνάρτηση πυκνότητας πιθανότητας αν η X είναι συνεχής και η συνάρτηση μάζας πιθανότητας αν η X είναι διακριτή. Η παράμετρος θ της κατανομής είναι άγνωστη και θέλουμε να την εκτιμήσουμε από το δείγμα των παρατηρήσεων $\{x_1,\ldots,x_n\}$. Θα θεωρήσουμε επίσης το πρόβλημα να έχουμε περισσότερες από μία άγνωστες παραμέτρους.

Μέθοδος των Ροπών Για συνήθεις κατανομές, μια παράμετρος θ της κατανομής $F(x;\theta)$ σχετίζεται με τις δύο κύριες παραμέτρους μ και σ^2. Για παράδειγμα, για την κανονική κατανομή οι μ και σ^2 είναι οι μόνες δύο παράμετροι που καθορίζουν πλήρως της συνάρτηση της κατανομής. Για την ομοιόμορφη κατανομή σε διάστημα $[a,b]$, η σχέση των παραμέτρων της κατανομής a και b με τις μ και σ^2 δίνεται ως $\mu=\frac{a+b}{2}$ και $\sigma^2=\frac{(b-a)^2}{12}$.

Γενικά όταν υπάρχει κάποια σχέση που μας επιτρέπει να υπολογίσουμε την παράμετρο θ (ή τις παραμέτρους θ_1,θ_2) από τις μ και σ^2, τότε βρίσκουμε την εκτιμήτρια $\hat{\theta}$ (ή τις εκτιμήτριες $\hat{\theta}_1,\hat{\theta}_2$) ως εξής:

1. υπολογίζουμε τις εκτιμήσεις $\bar{x}$ και s^2 των μ και σ^2 αντίστοιχα,
2. αντικαθιστούμε τις εκτιμήσεις $\bar{x}$ και s^2 στην έκφραση της θ (ή των θ_1,θ_2) ως προς μ και σ^2.

Αυτή είναι η *μέθοδος των ροπών* (method of moments). Η ονομασία προκύπτει από τη χρήση των ροπών στην εκτίμηση των παραμέτρων: τη μέση τιμή μ που είναι η πρώτη ροπή και τη διασπορά σ^2 που είναι η δεύτερη κεντρική ροπή. Αν οι δύο αυτές ροπές δεν επαρκούν, δηλαδή έχουμε να εκτιμήσουμε περισσότερες από δύο παραμέτρους ή οι σχέσεις δε δίνουν μοναδικότητα λύσης για τις παραμέτρους, χρησιμοποιούμε και ροπές μεγαλύτερου βαθμού, αλλά δε θα ασχοληθούμε με τέτοια προβλήματα.

Παράδειγμα 2.6 *Στο Κεφάλαιο 1 μελετήσαμε ένα δείγμα από μετρήσεις της αντοχής θραύσης 25 σκυροδεμάτων κάποιου τύπου Α. Τα δεδομένα αυτά (καθώς και τα δεδομένα από 20 σκυροδέματα ενός άλλου τύπου Β που θα μελετήσουμε αργότερα) δίνονται στον Πίνακα 1.3. Για την αντοχή θραύσης σκυροδέματος τύπου Α είχαμε βρεί πως η δειγματική μέση τιμή $\bar{x}$ είναι*

$$\bar{x}=\frac{1}{25}\sum_{i=1}^{25}x_i=\frac{1}{25}141.8=5.67$$

και η δειγματική διασπορά s^2 είναι (όπου έχουμε πρώτα υπολογίσει πως $\sum_{i=1}^{25}x_i^2=813.3$)

$$s^2=\frac{1}{24}\left(\sum_{i=1}^{25}x_i^2-25\bar{x}^2\right)=\frac{1}{24}(813.3-25\cdot 5.67^2)=0.375.$$

*Με βάση αυτό το δείγμα η εκτίμηση της μέση τιμής μ είναι 5.67 **ksi** και της διασποράς σ^2 είναι $0.375\,($**ksi**$)^2$.*

Αν η τ.μ. X (αντοχή θραύσης σκυροδέματος) ακολουθεί κανονική κατανομή $N(\mu, \sigma^2)$, τότε είναι φανερό πως αυτές οι εκτιμήσεις περιγράφουν πλήρως την κανονική κατανομή της X με βάση αυτό το δείγμα.

Αν η τ.μ. X ακολουθεί ομοιόμορφη κατανομή σε κάποιο διάστημα $[a, b]$, τότε μπορούμε να εκτιμήσουμε τις παραμέτρους a και b από τις σχέσεις $\mu = \frac{a+b}{2}$ και $\sigma^2 = \frac{(b-a)^2}{12}$. Αντικαθιστούμε τις μ και σ^2 με τις εκτιμήσεις $\bar{x}$ και s^2 και λύνουμε το παρακάτω σύστημα εξισώσεων

$$\begin{array}{rcl} \bar{x} & = & \frac{a+b}{2} \\ s^2 & = & \frac{(b-a)^2}{12} \end{array} \quad \Rightarrow \quad \begin{array}{rcl} \hat{a} & = & \bar{x} - \sqrt{3}s \\ \hat{b} & = & \bar{x} + \sqrt{3}s \end{array} \quad \Rightarrow \quad \begin{array}{l} \hat{a} = 4.61 \\ \hat{b} = 6.73 \end{array}$$

Μέθοδος της Μεγίστης Πιθανοφάνειας Η μέθοδος αυτή δίνει την εκτίμηση που έχει τη μέγιστη πιθανοφάνεια, δηλαδή δίνει την τιμή της παραμέτρου η οποία, μεταξύ όλων των δυνατών τιμών της παραμέτρου, είναι η πιό πιθανή με βάση το δείγμα.

Η συνάρτηση πυκνότητας πιθανότητας για κάποια τιμή $X = x_i$ που εξαρτάται από κάποια παράμετρο θ παίρνει την τιμή $f(x_i; \theta)$ (αν η X είναι διακριτή τότε αυτή η τιμή εκφράζει την πιθανότητα $\mathrm{P}(X = x_i)$). Επειδή $\{x_1, \ldots, x_n\}$ είναι ανεξάρτητες, η πιθανότητα να τις παρατηρήσουμε σ' ένα τυχαίο δείγμα μεγέθους n δίνεται από τη **συνάρτηση πιθανόφανειας** (likelihood function) ως προς θ

$$L(x_1, \ldots, x_n; \theta) = f(x_1; \theta) \cdots f(x_n; \theta).$$

Στα προβλήματα εκτίμησης, θεωρούμε τα $\{x_1, \ldots, x_n\}$ δεδομένα και ενδιαφερόμαστε για τη θ. Άν λοιπόν $L(x_1, \ldots, x_n; \theta_1) > L(x_1, \ldots, x_n; \theta_2)$ για δύο τιμές θ_1 και θ_2 της θ, τότε η τιμή θ_1 είναι πιο αληθοφανής από τη θ_2 γιατί δίνει μεγαλύτερη πιθανότητα να παρατηρήσουμε τα $\{x_1, \ldots, x_n\}$. Θέλουμε λοιπόν να βρούμε την 'πιό αληθοφανή' τιμή της θ, δηλαδή την τιμή $\hat{\theta}$ που μεγιστοποιεί τη $L(x_1, \ldots, x_n; \theta)$ ή καλύτερα (για ευκολότερους υπολογισμούς) τη $\log L(x_1, \ldots, x_n; \theta)$. Άρα η **εκτιμήτρια μεγίστης πιθανοφάνειας** (maximum likelihood estimator) $\hat{\theta}$ βρίσκεται από τη σχέση

$$\frac{\partial \log L(x_1, \ldots, x_n; \theta)}{\partial \theta} = 0. \tag{2.5}$$

Άν θέλουμε να εκτιμήσουμε δύο ή περισσότερες παραμέτρους $\theta_1, \ldots, \theta_m$, η συνάρτηση πιθανόφανειας είναι $L(x_1, \ldots, x_n; \theta_1, \ldots, \theta_m)$ και οι εκτιμήτριες $\hat{\theta}_1, \ldots, \hat{\theta}_m$ βρίσκονται λύνοντας το σύστημα των m εξισώσεων

$$\frac{\partial \log L(x_1, \ldots, x_n; \theta_1, \ldots, \theta_m)}{\partial \theta_j} = 0 \quad \text{για} \quad j = 1, \ldots, m. \tag{2.6}$$

Παράδειγμα 2.7 *Έχουμε ένα τυχαίο δείγμα $\{x_1, \ldots, x_n\}$ από κανονική κατανομή $N(\mu, \sigma^2)$ και θέλουμε να εκτιμήσουμε τη μέση τιμή μ θεωρώντας τη σ^2 γνωστή. Η συνάρτηση πυκνότητας πιθανότητας της κανονικής κατανομής είναι*

$$f_X(x; \mu) \equiv f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Η συνάρτηση πιθανόφανειας (για την οποία μόνο η παράμετρος μ είναι άγνωστη) είναι

$$L(x_1, \ldots, x_n; \mu) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2\right],$$

όπου $\exp(x) \equiv e^x$*. Ο λογάριθμος της συνάρτησης πιθανόφανειας είναι*

$$\log L(x_1, \ldots, x_n; \mu) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2.$$

Η εκτιμήτρια μεγίστης πιθανοφάνειας $\hat{\mu}$ *βρίσκεται μηδενίζοντας την παράγωγο της* $\log L$ *(σχέση (2.5))*

$$\frac{\partial \log L}{\partial \mu} = 0 \quad \Rightarrow \quad \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \mu) = 0 \tag{2.7}$$

που δίνει τη λύση

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x},$$

δηλαδή είναι ίδια με την εκτιμήτρια $\bar{x}$ *της μέσης τιμής* μ *που ορίσαμε για οποιαδήποτε κατανομή της τ.μ.* X.

Παράδειγμα 2.8 *Άς υποθέσουμε στο προηγούμενο παράδειγμα πως κι η διασπορά* σ^2 *είναι άγνωστη. Τότε στην παραπάνω εξίσωση (2.7) προστίθεται κι η εξίσωση*

$$\frac{\partial \log L}{\partial \sigma^2} = 0 \quad \Rightarrow \quad -\frac{n}{2\sigma^2} + \frac{1}{\sigma^4}\sum_{i=1}^{n}(x_i - \mu)^2 = 0. \tag{2.8}$$

Η επίλυση του συσήματος των εξισώσεων (2.7) και (2.8) δίνει την ίδια λύση για τη μ *και για τη* σ^2 *είναι*

$$\widehat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{\mu})^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2.$$

Οι εκτιμήτριες μεγίστης πιθανοφάνειας λοιπόν για τη μέση τιμή μ *και τη διασπορά* σ^2 *μιας τ.μ. που ακολουθεί κανονική κατανομή είναι απλά η δειγματική μέση τιμή και διασπορά αντίστοιχα, αλλά για τη διασπορά έχουμε τη 'μεροληπτική' δειγματική διασπορά* $\tilde{s}^2$ *(σχέση (2.3)). Ασυμπτωτικά όμως (για μεγάλο* n*) η εκτιμήτρια* $\widehat{\sigma^2} = \tilde{s}^2$ *είναι αμερόληπτη.*

Η μέθοδος μεγίστης πιθανοφάνειας είναι η καλύτερη μέθοδος εκτίμησης αν γνωρίζουμε την κατανομή της τ.μ. X και μπορεί να εφαρμοσθεί σε οποιοδήποτε πρόβλημα εκτίμησης παραμέτρων.

Μέθοδος των Ελαχίστων Τετραγώνων Αυτή η μέθοδος εφαρμόζεται μόνο στην περίπτωση που οι άγνωστες παράμετροι εμφανίζονται σε σχέσεις τυχαίων μεταβλητών και οι σχέσεις αυτές είναι *γραμμικές* ως προς τις παραμέτρους που θέλουμε να εκτιμήσουμε. Μια απλή περίπτωση είναι να έχουμε μια τ.μ. Y και η κάθε τιμή της y να δίνεται από τη σχέση

$$y = \theta_1 x_1 + \cdots + \theta_m x_m + \epsilon,$$

όπου οι τιμές $x_1, \dots, x_m$ είναι γνωστές, $\theta_1, \dots, \theta_m$ είναι οι άγνωστες παράμετροι και ϵ είναι μια άλλη τ.μ. με $\mathrm{E}(\epsilon) = 0$. Θα ασχοληθούμε με τη μέθοδο αυτή στο Κεφάλαιο 4.

Παρατηρήσεις

Η μέθοδος μεγίστης πιθανοφάνειας μπορεί να εφαρμοσθεί για οποιοδήποτε θ αν γνωρίζουμε την κατανομή $F_X(x;\theta)$ ενώ η μέθοδος των ροπών δεν εφαρμόζεται αν η θ δε μπορεί να υπολογισθεί από τις ροπές.

Η εκτίμηση μεγίστης πιθανοφάνειας έχει όλες τις ιδιότητες καλής εκτιμήτριας, δηλαδή είναι αμερόληπτη (ασυμπτωτικά, δηλαδή η μεροληψία $b(\theta)$ τείνει στο μηδέν για μεγάλα n), συνεπής, αποτελεσματική κι επαρκής. Γι αυτό κι αυτή η μέθοδος είναι η καλύτερη μέθοδος εκτίμησης παραμέτρων αν γνωρίζουμε την κατανομή $F_X(x;\theta)$.

2.2 Εκτίμηση Διαστήματος Εμπιστοσύνης

Η σημειακή εκτίμηση $\hat{\theta}$ από κάποιο δείγμα δεν περιέχει καμιά πληροφορία για την ακρίβεια της εκτίμησης της θ. Η εκτίμηση $\hat{\theta}$ είναι μια τιμή που δε γνωρίζουμε πόσο κοντά είναι στην πραγματική τιμή της θ κι επίσης η τιμή αυτή αλλάζει με το δείγμα. Για παράδειγμα, υπολογίζουμε τη δειγματική μέση τιμή $\bar{x}$ από ένα τυχαίο δείγμα μεγέθους n. Άν πάρουμε ένα άλλο τυχαίο δείγμα ίδιου μεγέθους, η τιμή της $\bar{x}$ θα είναι διαφορετική. Μπορεί να είναι πιό κοντά ή πιό μακριά στην πραγματική τιμή της μ απ' ότι αυτή από το προηγούμενο δείγμα. Γενικά λοιπόν η εκτιμήτρια $\hat{\theta}$ εξαρτάται από το δείγμα και είναι λοιπόν τ.μ. με κάποια κατανομή. Γι αυτό στην εκτίμηση της θ είναι σημαντικό εκτός από τη σημειακή εκτίμηση $\hat{\theta}$ να υπολογίσουμε και διάστημα $[\theta_1, \theta_2]$ που να μπορούμε να πούμε με μεγάλη εμπιστοσύνη ότι θα περιέχει την πραγματική τιμή της παραμέτρου θ. Στη συνέχεια θα δούμε τέτοια διαστήματα εμπιστοσύνης για διάφορες παραμέτρους, αρχίζοντας από τη μέση τιμή.

2.2.1 Διάστημα εμπιστοσύνης της μέσης τιμής μ

Η σημειακή εκτίμηση της μέσης τιμής μ μιας τ.μ. X είναι η δειγματική μέση τιμή $\bar{x}$ (σχέση (2.1)) που είναι κι αμερόληπτη εκτιμήτρια της μ, δηλαδή η μέση τιμή της $\bar{x}$ είναι η πραγματική τιμή μ που είναι άγνωστη όμως σε μας, $\mu_{\bar{x}} = \mathrm{E}(\bar{x}) = \mu$. Παρ' όλο που η εκτιμήτρια $\bar{x}$ είναι διαφορετική από δείγμα σε δείγμα, επειδή η $\bar{x}$ είναι συνεπής εκτιμήτρια, όταν αυξάνεται το μέγεθος n του δείγματος η $\bar{x}$ πλησιάζει τη μέση τιμή μ. Η διασπορά λοιπόν της $\bar{x}$ θα πρέπει να εξαρτάται από το n έτσι ώστε καθώς ο αριθμός των παρατηρήσεων μεγαλώνει η διασπορά να μικραίνει. Πράγματι για τη διασπορά $\sigma_{\bar{x}}^2$ έχουμε

$$\sigma_{\bar{x}}^2 = \mathrm{Var}(\bar{x}) = \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}(x_i) = \frac{1}{n^2}(n\sigma^2) = \frac{\sigma^2}{n}, \tag{2.9}$$

δηλαδή η διασπορά της εκτιμήτριας $\bar{x}$ είναι ανάλογη της διασποράς σ^2 της X κι αντιστρόφως ανάλογη του αριθμού των παρατηρήσεων n. Στην παραπάνω σχέση υποθέσαμε πως οι παρατηρήσεις $x_1, \dots, x_n$ είναι ανεξάρτητες (εννοώντας και πάλι τις τ.μ. $X_1, \dots, X_n$). Την τυπική απόκλιση (τετραγωνική ρίζα της διασποράς) $\sigma_{\bar{x}} = \sigma/\sqrt{n}$ της $\bar{x}$ θα την ονομάζουμε **σταθερό σφάλμα** (standard error) της εκτιμήτριας $\bar{x}$.

Η τ.μ. $\bar{x}$ λοιπόν έχει κάποια κατανομή με μέση τιμή $\mu_{\bar{x}} = \mu$ και διασπορά $\sigma^2_{\bar{x}} = \sigma^2/n$. Στη συνέχεια, για να ορίσουμε διάστημα εμπιστοσύνης της μ θα βασιστούμε στην κατανομή της $\bar{x}$. Για να υποθέσουμε κάποια γνωστή κατανομή για τη $\bar{x}$ χρειάζεται να ελέγξουμε το μέγεθος του δείγματος, αν η κατανομή της X είναι κανονική κι αν γνωρίζουμε τη διασπορά της.

Γνωστή διασπορά Θεωρούμε εδώ ότι γνωρίζουμε τη διασπορά σ^2 της τ.μ. X στον πληθυσμό. Για την κατανομή της $\bar{x}$ διακρίνουμε δύο περιπτώσεις:

1. Αν η τ.μ. X ακολουθεί κανονική κατανομή $\mathrm{N}(\mu, \sigma^2)$ **ή** το δείγμα είναι μεγάλο ($n > 30$) τότε και η τ.μ. $\bar{x}$ ακολουθεί κανονική κατανομή $\mathrm{N}(\mu, \sigma^2/n)$.

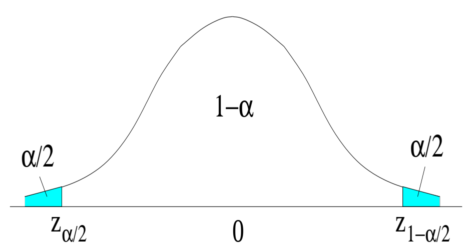
2. Αν δε συμβαίνει το παραπάνω, δηλαδή αν η τ.μ. X δεν ακολουθεί κανονική κατανομή **και** το δείγμα είναι μικρό τότε γενικά δε γνωρίζουμε την κατανομή της $\bar{x}$.

Αν το δείγμα είναι μεγάλο η κανονική κατανομή της $\bar{x}$ δίνεται από το *Κεντρικό Οριακό Θεώρημα*, ενώ αν η τ.μ. X ακολουθεί κανονική κατανομή τότε και το άθροισμα τέτοιων τ.μ. $X_1, \ldots, X_n$ ακολουθεί κανονική κατανομή κι έτσι προκύπτει πως και η $\bar{x}$ ακολουθεί κανονική κατανομή.

Υποθέτοντας ότι η $\bar{x}$ ακολουθεί κανονική κατανομή $\mathrm{N}(\mu, \sigma^2/n)$, η τ.μ. z που προκύπτει από τον απλό μετασχηματισμό $z \equiv \frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$ ακολουθεί την τυπική κανονική κατανομή

$$\bar{x} \sim \mathrm{N}(\mu, \sigma^2/n) \quad \Rightarrow \quad z \equiv \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \sim \mathrm{N}(0,1).$$

Για την τυπική κανονική κατανομή μπορούμε εύκολα να ορίσουμε ένα διάστημα $[z_{\alpha/2}, z_{1-\alpha/2}]$, στο οποίο θα ανήκει η z με κάποια δοθείσα πιθανότητα $1-\alpha$, όπως φαίνεται στο Σχήμα 2.1. Ο λόγος που χρησιμοποιούμε το δείκτη $\alpha/2$ για την κρίσιμη τιμή $z_{\alpha/2}$ φαίνεται απο το Σχήμα 2.1. Τα άκρα του διαστήματος, $z_{\alpha/2}$ και $z_{1-\alpha/2}$, λέγονται *κρίσιμες τιμές*. Οι δείκτες $\alpha/2$ και $1-\alpha/2$



Σχήμα 2.1: Συνάρτηση πυκνότητας πιθανότητας της τυπικής κανονικής κατανομής και οι ουρές της για κάποιο α.

δηλώνουν τις τιμές της αθροιστικής συνάρτησης για $z_{\alpha/2}$ και $z_{1-\alpha/2}$ αντίστοιχα, δηλαδή ισχύει

$$\begin{aligned} \Phi(z_{\alpha/2}) &= \mathrm{P}(z < z_{\alpha/2}) = \alpha/2 \\ \Phi(z_{1-\alpha/2}) &= \mathrm{P}(z < z_{1-\alpha/2}) = 1-\alpha/2 \end{aligned}$$

όπου $\Phi(z)$ είναι η αθροιστική συνάρτηση της τυπικής κανονικής κατανομής. Άρα η πιθανότητα να είναι $z < z_{\alpha/2}$ ή $z > z_{1-\alpha/2}$ είναι α. Οι δύο σκιασμένες περιοχές στο Σχήμα 2.1 κατέχουν μαζί ποσοστό $\alpha\%$ του συνολικού εμβαδού του ολοκληρώματος της συνάρτησης πυκνότητας πιθανότητας. Αντίστοιχα η πιθανότητα να συμβαίνει $z \in [z_{\alpha/2}, z_{1-\alpha/2}]$ είναι $1-\alpha$. Γενικά λοιπόν ισχύει

$$\mathrm{P}(z_{\alpha/2} < z \leq z_{1-\alpha/2}) = \Phi(z_{1-\alpha/2}) - \Phi(z_{\alpha/2}) = 1-\alpha.$$

Επειδή η συνάρτηση πυκνότητας πιθανότητας της τυπικής κανονικής κατανομής είναι συμμετρική ως προς το 0 ισχύει

$$z_{\alpha/2} = -z_{1-\alpha/2}.$$

Άρα στην ουσία για να ορίσουμε το διάστημα $[z_{\alpha/2}, z_{1-\alpha/2}]$ χρειαζόμαστε μία μόνο κρίσιμη τιμή.

Ανακεφαλαιώνοντας, βρήκαμε ότι σε κάθε δοθείσα πιθανότητα $1-\alpha$ αντιστοιχεί ένα διάστημα τιμών της τ.μ. z που ορίζεται από την κρίσιμη τιμή $z_{1-\alpha/2}$ που την υπολογίζουμε ως $z_{1-\alpha/2} = \Phi^{-1}(1-\alpha/2)$ από τον στατιστικό πίνακα της τυπικής κανονικής κατανομής. Για παράδειγμα για $\alpha = 0.05$ το διάστημα $[-1.96, 1.96]$ περιέχει την τ.μ. z με πιθανότητα 0.95, όπου $z_{0.975} = \Phi^{-1}(0.975) = 1.96$. [Ο συμβολισμός που χρησιμοποιούμε εδώ, όπως και για άλλες κατανομές που θα δούμε παρακάτω, βασίζεται στην αρχή της αντιστοίχισης του δείκτη της κρίσιμης τιμής στην τιμή της αντίστοιχης αθροιστικής συνάρτησης. Σε κάποια βιβλία η θετική κρίσιμη τιμή $z_{1-\alpha/2}$ συμβολίζεται ως $z_{\alpha/2}$ και η αρνητική κρίσιμη τιμή $z_{\alpha/2}$ συμβολίζεται ως $-z_{\alpha/2}$].

Θέλουμε να μετασχηματίσουμε το διάστημα $[z_{\alpha/2}, z_{1-\alpha/2}]$ για πιθανότητα $1-\alpha$ στο αντίστοιχο διάστημα που περιέχει την παραμέτρο μ. Γι αυτό λύνουμε τις σχέσεις

$$z_{\alpha/2} = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \qquad z_{1-\alpha/2} = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$$

ως προς μ και βρίσκουμε τα άκρα του διαστήματος για τη μέση τιμή μ

$$\bar{x} \pm z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \quad \rightarrow \quad \left[\bar{x} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}, \quad \bar{x} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right]. \tag{2.10}$$

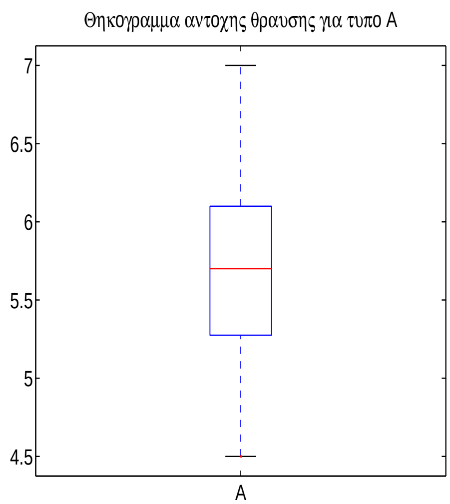
Το διάστημα αυτό υπολογίστηκε για κάποια δοθείσα πιθανότητα $1-\alpha$ που είναι το προκαθορισμένο **επίπεδο εμπιστοσύνης** (confidence level) και λέγεται **διάστημα εμπιστοσύνης** (confidence interval) της μ σε επίπεδο $1-\alpha$. Η ερμηνεία αυτού του διαστήματος θέλει προσοχή. Σε μια πρώτη προσέγγιση θα λέγαμε ότι σημαίνει ‘με πιθανότητα (εμπιστοσύνη) $1-\alpha$ η μέση τιμή μ βρίσκεται μέσα σ' αυτό το διάστημα’, το οποίο δεν είναι ορθό αφού η μ είναι σταθερό μέγεθος κι όχι τ.μ. για να μιλάμε για την τιμή της με πιθανότητες. Το μέγεθος που αλλάζει (ανάλογα με το δείγμα) είναι το διάστημα και σ' αυτό πρέπει να αναφέρεται η πιθανότητα ή εμπιστοσύνη. Σωστότερη λοιπόν ερμηνεία είναι ότι ‘αν χρησιμοποιούσαμε πολλά τέτοια διαστήματα από διαφορετικά δείγματα, ποσοστό $(1-\alpha)\%$ απ' αυτά θα περιείχαν τη μ’ ή ‘με $1-\alpha$ πιθανότητα (εμπιστοσύνη) το διάστημα αυτό θα περιέχει την πραγματική μ’.

Συνοπτικά η διαδικασία για τον προσδιορισμό του διαστήματος εμπιστοσύνης της μ είναι

Επιλογή του επιπέδου εμπιστοσύνης $1-\alpha$
$\Downarrow$
Υπολογισμός του $z_{1-\alpha/2}$ από τον αντίστοιχο στατιστικό πίνακα για την τυπική κανονική κατανομή
$\Downarrow$
Υπολογισμός του διαστήματος $[\bar{x} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}]$, όπου το σ είναι γνωστό και το $\bar{x}$ είναι η δειγματική μέση τιμή των n παρατηρήσεων.

Παράδειγμα 2.9 *Θέλουμε να εκτιμήσουμε διάστημα εμπιστοσύνης σε επίπεδο* 95% *για τη μέση αντοχή θραύσης σκυροδέματος τύπου A από τα δεδομένα του Πίνακα 1.3. Υποθέτουμε ότι από παλιότερες μετρήσεις γνωρίζουμε ότι η διασπορά για την αντοχή θραύσης σκυροδέματος είναι* $\sigma^2 = 0.38\,(ksi)^2$. *Εξετάζοντας τη δειγματική κατανομή αντοχής θραύσης από τα δεδομένα μας, π.χ. σχεδιάζοντας το ιστόγραμμα των δεδομένων του Πίνακα 1.3 (όπως κάναμε στο Σχήμα 1.2) ή το θηκόγραμμα που παρουσιάζεται στο Σχήμα 2.2, βλέπουμε ότι φαίνεται να είναι κανονική (είναι συμμετρική και δεν έχει μακριές ουρές). Ειδικά για το θηκόγραμμα φαίνεται να τηρούνται τα τρία χαρακτηριστικά που συνιστούν κανονική κατανομή όπως τα θέσαμε στην Παράγραφο 1.2:*

- *Η διάμεσος δεν τείνει προς το* Q_1 *ή το* Q_3.
- *Οι μύστακες έχουν περίπου το ίδιο μήκος.*
- *Δεν υπάρχουν ακραίες τιμές ή ύποπτες ακραίες τιμές.*



Σχήμα 2.2: Θηκόγραμμα των δεδομένων αντοχής θραύσης σκυροδέματος τύπου A του Πίνακα 1.3.

Άρα μπορούμε να υποθέσουμε ότι η αντοχή θραύσης σκυροδέματος X *ακολουθεί κανονική κατανομή* $N(\mu, 0.38)$ *και τότε η δειγματική μέση τιμή* $\bar{x}$ *της αντοχής θραύσης σκυροδέματος ακολουθεί επίσης κανονική κατανομή* $N(\mu, 0.38/25)$, *όπου* $n = 25$ *είναι το μέγεθος του δείγματος.*

Για την εκτίμηση του διάστηματος εμπιστοσύνης ακολουθούμε τα βήματα:

1. *Το επίπεδο εμπιστοσύνης είναι* $1-\alpha = 0.95$ *(*$\alpha = 0.05$*).*
2. *Από τον στατιστικό πίνακα έχουμε* $z_{0.975} = \Phi^{-1}(0.975) = 1.96$.
3. *Το διάστημα για τη μέση τιμή* μ *είναι (έχουμε υπολογίσει ότι* $\bar{x} = 5.67$ *ksi, δες Παράδειγμα 2.6)*

$$5.67 \pm 1.96\frac{\sqrt{0.38}}{5} \quad \rightarrow \quad [5.43,\ 5.91].$$

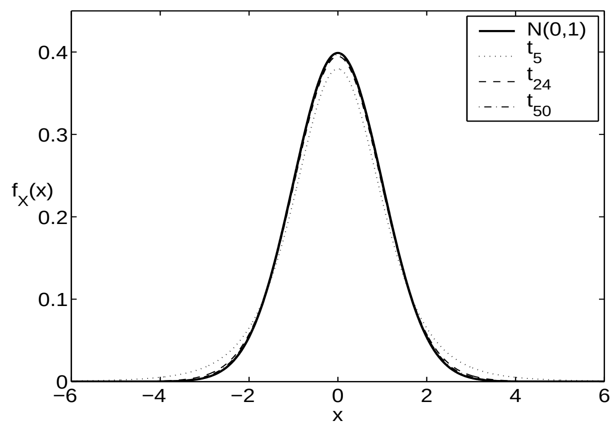
Άρα το 95% διάστημα εμπιστοσύνης της μέσης αντοχής θραύσης σκυροδέματος με βάση αυτό το δείγμα είναι $[5.43,\ 5.91]$. *Μπορούμε να πούμε ότι η σημειακή εκτίμηση* $\bar{x} = 5.67$ *είναι αρκετά ακριβής αφού το αντίστοιχο 95% διάστημα εμπιστοσύνης είναι αρκετά μικρό.*

Άγνωστη διασπορά Γενικά η διασπορά σ^2 της τ.μ. X είναι άγνωστη και την εκτιμούμε από το δείγμα με την s^2 (π.χ. δες την (2.4) που χρησιμοποιούμε για τους υπολογισμούς). Αν το δείγμα είναι μεγάλο ($n > 30$) η εκτιμήτρια s^2 είναι αρκετά ακριβής κι απλά μπορούμε να αντικαταστήσουμε τη διασπορά σ^2 με τη δειγματική διασπορά s^2 στην παραπάνω διαδικασία για να πάρουμε το διάστημα εμπιστοσύνης για τη μ. Αν όμως το δείγμα είναι μικρό τότε η προσέγγιση δεν είναι καλή και το διάστημα μπορεί να είναι αρκετά ανακριβές.

Για μικρό n και υποθέτοντας πως η X ακολουθεί κανονική κατανομή, η τ.μ. t που ορίζεται ως $t \equiv \frac{\bar{x}-\mu}{s/\sqrt{n}}$ ακολουθεί την κατανομή student ή την t–κατανομή με $n-1$ βαθμούς ελευθερίας

$$t \equiv \frac{\bar{x}-\mu}{s/\sqrt{n}} \sim \mathrm{t}_{n-1}.$$

Η κατανομή αυτή μοιάζει με την τυπική κανονική κατανομή και την προσεγγίζει καθώς αυξάνει ο αριθμός των βαθμών ελευθερίας όπως φαίνεται στο Σχήμα 2.3. Για μεγάλο n η προσέγγιση



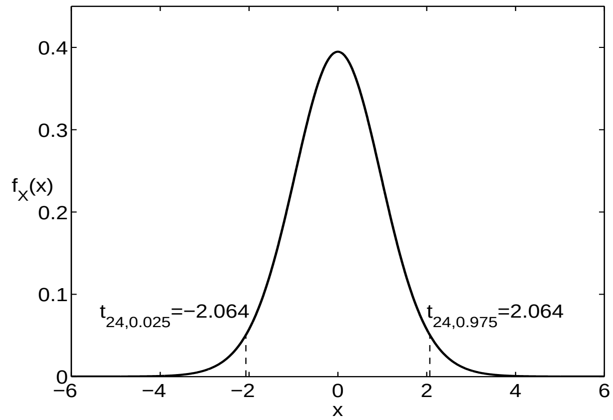
Σχήμα 2.3: Συνάρτηση πυκνότητας πιθανότητας για την τυπική κανονική κατανομή και την κατανομή student με 5, 24 και 50 βαθμούς ελευθερίας, όπως δείχνει το επεξήγημα των γραμμών.

είναι πολύ καλή κι οι τιμές των t και z είναι πρακτικά ίδιες.

Η διαδικασία για την εκτίμηση του διαστήματος εμπιστοσύνης είναι ίδια όπως στην περίπτωση της γνωστής διασποράς, αλλά η άγνωστη διασπορά σ^2 αντικαθίσταται από τη δειγματική διασπορά s^2 που υπολογίζεται από το δείγμα κι η κρίσιμη τιμή είναι $t_{n-1,1-\alpha/2}$ αντί για $z_{1-\alpha/2}$ και τη βρίσκουμε από το στατιστικό πίνακα για την κατανομή student. [Η κρίσιμη τιμή της t ορίζεται από το $1-\alpha/2$ αλλά και από τους βαθμούς ελευθερίας $n-1$. Στο Σχήμα 2.4 παρουσιάζεται η κατανομή student για $n-1=24$ βαθμούς ελευθερίας κι οι κρίσιμες τιμές της για $1-\alpha=0.95$.] Το διάστημα εμπιστοσύνης σε επίπεδο $1-\alpha$ είναι

$$\bar{x} \pm t_{n-1,1-\alpha/2}\frac{s}{\sqrt{n}}. \tag{2.11}$$

Παράδειγμα 2.10 *Για το προηγούμενο παράδειγμα υποθέτουμε πως η διασπορά είναι άγνωστη (που είναι και η πιό πιθανή περίπτωση για το πραγματικό πρόβλημα). Το δείγμα είναι αρκετά μεγάλο κι ίσως θα μπορούσαμε να υποθέσουμε πως η $\bar{x}$ σαν τ.μ. ακολουθεί κανονική κατανομή*



Σχήμα 2.4: Συνάρτηση πυκνότητας πιθανότητας για την κατανομή student με 24 βαθμούς ελευθερίας κι η δεξιά κι αριστερή κρίσιμη τιμή της για $1-\alpha = 0.95$.

και να χρησιμοποιήσουμε στους υπολογισμούς του διαστήματος εμπιστοσύνης την κρίσιμη τιμή $z_{1-\alpha/2}$. Για να εκτιμήσουμε όμως το διάστημα εμπιστοσύνης για τη μέση αντοχή θραύσης με την καλύτερη δυνατή ακρίβεια χρησιμοποιούμε την κρίσιμη τιμή $t_{n-1,1-\alpha/2}$ από την κατανομή student (δες Σχήμα 2.4). Η δειγματική διασπορά έχει βρεθεί να είναι $s^2 = 0.375\,(\textbf{ksi})^2$ (δες Παράδειγμα 2.6). Από το στατιστικό πίνακα για την κατανομή student, για $1-\alpha/2 = 0.975$ και $n-1 = 24$, βρίσκουμε $t_{24,\,0.975} = 2.064$ και το διάστημα εμπιστοσύνης για τη μ είναι

$$5.67 \pm 2.064\frac{\sqrt{0.375}}{5} \quad \rightarrow \quad [5.42, 5.92].$$

Η χρήση της κρίσιμης τιμής $t_{24,\,0.975} = 2.064$ υπαγορεύεται από τη χρήση της εκτίμησης s^2 αντί της πραγματικής διασποράς σ^2 που δεν τη γνωρίζουμε. Αν αποφασίζαμε εσφαλμένα να χρησιμοποιήσουμε στους παραπάνω υπολογισμούς τη $z_{0.975} = 1.96$ της κανονικής κατανομής θα βρίσκαμε το διάστημα $[5.52, 5.82]$ πού είναι φυσικά πιό μικρό αφού $z_{0.975} < t_{24,\,0.975}$. Το διάστημα όμως αυτό δεν είναι ακριβές γιατί κάναμε την υπόθεση για κανονική κατανομή της $\bar{x}$ που δεν εφαρμόζεται στην περίπτωση που η διασπορά είναι άγνωστη και το δείγμα είναι μικρό.

Γενικά όταν τον n δεν είναι μεγάλο το διάστημα εμπιστοσύνης της $\bar{x}$ που παίρνουμε υποθέτοντας ότι η $\bar{x}$ ακολουθεί κατανομή student είναι πιό μεγάλο από αυτό που παίρνουμε υποθέτοντας ότι η $\bar{x}$ ακολουθεί κανονική κατανομή και η διασπορά παραμένει η ίδια.

Μη κανονική κατανομή και μικρά δείγματα Σε κάποιες περιπτώσεις το δείγμα μπορεί να είναι μικρό **και** η κατανομή της τ.μ. X που παρατηρούμε να μην είναι κανονική. [Όταν δε ξέρουμε τίποτε για την κατανομή της X αυτό το εκτιμούμε από τα δεδομένα του δείγματος, π.χ. από τη μορφή του ιστογράμματος ή του θηκογράμματος των δεδομένων.] Σε μια τέτοια περίπτωση (κι ανεξάρτητα από το αν η διασπορά είναι γνωστή ή όχι) δε μπορούμε να υποθέσουμε πως η $\bar{x}$ ακολουθεί κάποια συγκεκριμένη κατανομή και στη συνέχεια να εκτιμήσουμε διάστημα εμπιστοσύνης για τη μέση τιμή μ.

Σε τέτοιες περιπτώσεις ενδιαφερόμαστε για τη διάμεσο $\delta = \delta_X$ της τ.μ. X αντί της $\mu = \mu_X$. Η σημειακή εκτίμηση $\tilde{x}$ της δ είναι απλά η κεντρική τιμή των παρατηρήσεων διαταγμένες σε αύξουσα σειρά (δες Παράγραφο 1.2.1). Το διάστημα εμπιστοσύνης για τη δ βρίσκεται με τη μέθοδο Wilcoxon που βασίζεται στην τάξη των παρατηρήσεων παρά στις τιμές τους. Αυτή η εκτίμηση διαστήματος εμπιστοσύνης λέγεται **μη παραμετρική** (επειδή δεν υποθέτουμε κάποια κατανομή και τις παραμέτρους της για την εκτιμήτρια).

Παρατηρήσεις

Γενικά θα θέλαμε το διάστημα εμπιστοσύνης να είναι όσο το δυνατό μικρότερο για να έχουμε όσο το δυνατόν πιό ακριβή σημειακή εκτίμηση. Το διάστημα εμπιστοσύνης εξαρτάται από:

- την κατανομή και τη διασπορά σ^2 της τ.μ. X [αυτά δε μπορούμε να τα αλλάξουμε, είναι χαρακτηριστικά της τ.μ. που μελετάμε].
- το μέγεθος n του δείγματος [αύξηση του n έχει σαν αποτέλεσμα τη μείωση του εύρους του διαστήματος, που είναι φυσικά επιθυμητό αλλά όχι πάντοτε εφικτό, αφού η απόκτηση πολλών παρατηρήσεων μπορεί να είναι εργασία επίπονη και πολυέξοδη].
- το επίπεδο εμπιστοσύνης $1-\alpha$ [αυτό το καθορίζουμε εμείς, αλλά δε θα θέλαμε να μικρήνουμε το διάστημα μειώνοντας την εμπιστοσύνη μας σ' αυτό γιατί τότε τα αποτελέσματά μας δε θα έχουν την επιθυμητή στατιστική σημαντικότητα]. Το επίπεδο εμπιστοσύνης που συνήθως χρησιμοποιείται στην πράξη είναι 95%.

Στον Πίνακα 2.1 συνοψίζεται η εκτίμηση του διαστήματος εμπιστοσύνης της μ στις διάφορες περιπτώσεις.

διασπορά	*κατανομή της* X	n	*κατανομή της* $\bar{x}$	*διάστημα εμπιστοσύνης*
γνωστή	κανονική		$z \equiv \dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}} \sim \mathrm{N}(0,1)$	$\bar{x} \pm z_{1-\alpha/2}\dfrac{\sigma}{\sqrt{n}}$
γνωστή	μη κανονική	μεγάλο	$z \equiv \dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}} \sim \mathrm{N}(0,1)$	$\bar{x} \pm z_{1-\alpha/2}\dfrac{\sigma}{\sqrt{n}}$
γνωστή	μη κανονική	μικρό	—	—
άγνωστη		μεγάλο	$z \equiv \dfrac{\bar{x}-\mu}{s/\sqrt{n}} \sim \mathrm{N}(0,1)$	$\bar{x} \pm z_{1-\alpha/2}\dfrac{s}{\sqrt{n}}$
άγνωστη	κανονική	μικρό	$t \equiv \dfrac{\bar{x}-\mu}{s/\sqrt{n}} \sim \mathrm{t}_{n-1}$	$\bar{x} \pm t_{n-1,1-\alpha/2}\dfrac{s}{\sqrt{n}}$
άγνωστη	μη κανονική	μικρό	—	—

Πίνακας 2.1: Εκτίμηση του διαστήματος εμπιστοσύνης της μ ανάλογα με τη γνώση της διασποράς και κατανομής της τ.μ. X και το μέγεθος n του δείγματος.

Εύρος διαστήματος εμπιστοσύνης Πολλές φορές πριν να κάνουμε το πείραμα και συλλέξουμε τις μετρήσεις προκαθορίζουμε ένα συγκεκριμένο εύρος για το δ.ε. ή ζητάμε το εύρος του δ.ε. να μην ξεπερνάει κάποιο ανώτατο όριο για να έχουν νόημα τα αποτελέσματα. Για να το πετύχουμε αυτό χωρίς να αλλάξουμε τη σημαντικότητα των στατιστικών αποτελεσμάτων, βρίσκουμε το μέγεθος n του δείγματος που μας δίνει αυτό το εύρος του δ.ε. Αυτό υπολογίζεται θέτοντας το εύρος του δ.ε. ίσο με την τιμή που ζητάμε και λύνοντας την εξίσωση ως προς το n. Για παράδειγμα, ας υποθέσουμε πως το δείγμα είναι μικρό και η τ.μ. X ακολουθεί κανονική κατανομή με άγνωστη διασπορά, χρησιμοποιούμε δηλαδή τη σχέση (2.11) για να υπολογίσουμε το δ.ε. της μέσης τιμής μ. Το εύρος του δ.ε. είναι

$$w = 2\, t_{n-1,1-\alpha/2} \cdot \frac{s}{\sqrt{n}} \tag{2.12}$$

και λύνοντας ως προς n βρίσκουμε ότι για να είναι το εύρος του δ.ε. ίσο με w πρέπει το δείγμα να έχει μέγεθος

$$n = \left(2\, t_{n-1,1-\alpha/2} \cdot \frac{s}{w}\right)^2. \tag{2.13}$$

Στην παραπάνω σχέση η τιμή $t_{n-1,1-\alpha/2}$ δεν είναι γνωστή αφού το n είναι άγνωστο και ζητούμενο. Θα πρέπει λοιπόν να χρησιμοποιήσουμε την τιμή $t_{n-1,1-\alpha/2}$ που συμφωνεί καλύτερα με το αποτέλεσμα για το n από την (2.13). Αν το n παίρνει μεγάλες τιμές το παραπάνω πρόβλημα δεν υφίσταται αφού για μεγάλα n η κρίσιμη τιμή $t_{n-1,1-\alpha/2}$ είναι πρακτικά ίδια με την αντίστοιχη κρίσιμη τιμή της τυπικής κανονικής κατανομής $z_{1-\alpha/2}$. Άρα για μεγάλο n η σχέση (2.13) δίνεται ως

$$n = \left(2\, z_{1-\alpha/2} \cdot \frac{s}{w}\right)^2. \tag{2.14}$$

Παράδειγμα 2.11 *Στο προηγούμενο παράδειγμα για την αντοχή θραύσης σκυροδέματος υπολογίσαμε το 95% δ.ε. της μέσης αντοχής θραύσης κάνοντας χρήση της κατανομής student και βρήκαμε ότι είναι $[5.42, 5.92]$ ksi. Το εύρος του δ.ε. είναι $w = 5.92 - 5.42 = 0.50$ ksi. Αν θέλουμε να το μειώσουμε στο μισό (δηλαδή σε $w = 0.25$ ksi) τότε αντί για 25 δοκίμια πρέπει να χρησιμοποιήσουμε*

$$n = \left(2 \cdot 1.96 \cdot \frac{\sqrt{0.375}}{0.25}\right)^2 = 92.2 \simeq 93,$$

δηλαδή πρέπει να αυξήσουμε το δείγμα σε 93 δοκίμια.

2.2.2 Διάστημα εμπιστοσύνης της διασποράς σ^2

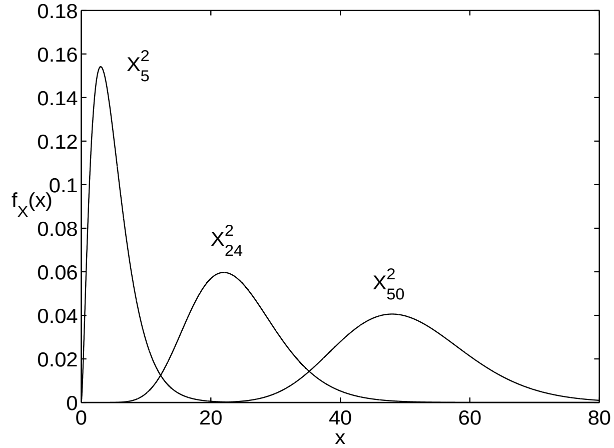
Όπως για να βρούμε διάστημα εμπιστοσύνης για τη μέση τιμή μ ορίσαμε πρώτα την κατανομή της εκτιμήτριας $\bar{x}$ έτσι και για να βρούμε διάστημα εμπιστοσύνης για τη διασπορά σ^2 ορίζουμε πρώτα την κατανομή της αμερόληπτης εκτιμήτριας s^2 της σ^2 (π.χ. από τη σχέση (2.4)). Γνωρίζουμε ότι η $\chi^2 \equiv \frac{(n-1)s^2}{\sigma^2}$ ακολουθεί την κατανομή $\mathcal{X}^2$ με $n-1$ βαθμούς ελευθερίας

$$\chi^2 \equiv \frac{(n-1)s^2}{\sigma^2} \sim \mathcal{X}^2_{n-1}.$$

Στο Σχήμα 2.5 παρουσιάζεται η μορφή της κατανομής $\mathcal{X}^2$ για χαρακτηριστικούς βαθμούς ελευθερίας. Για λίγους βαθμούς ελευθερίας η κατανομή $\mathcal{X}^2$ είναι αρκετά λοξή και γίνεται πιο συμμετρική καθώς αυξάνουν οι βαθμοί ελευθερίας.

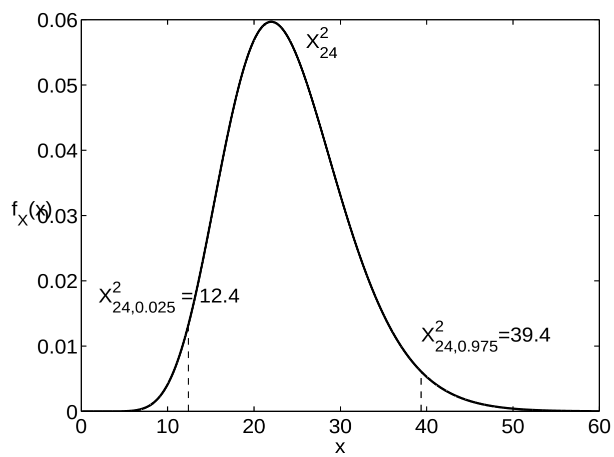
Για δοθείσα πιθανότητα $1-\alpha$ μπορούμε να βρούμε από τον στατιστικό πίνακα για τη κατανομή $\mathcal{X}^2$ τις δύο κρίσιμες τιμές $\chi^2_{n-1,\alpha/2}$ και $\chi^2_{n-1,1-\alpha/2}$ για τις οποίες ισχύει

$$\mathrm{P}(\chi^2_{n-1,\alpha/2} < \chi^2 < \chi^2_{n-1,1-\alpha/2}) = 1-\alpha. \tag{2.15}$$



Σχήμα 2.5: Συνάρτηση πυκνότητας πιθανότητας για την κατανομή $\mathcal{X}^2$ με 5, 24 και 50 βαθμούς ελευθερίας.

Επειδή η $\mathcal{X}^2$ δεν είναι συμμετρική έχουμε δύο κρίσιμες τιμές: την αριστερή κρίσιμη τιμή $\chi^2_{n-1,\alpha/2}$ που είναι τέτοια ώστε $\mathrm{P}(\chi^2 < \chi^2_{n-1,\alpha/2}) = \alpha/2$ και τη δεξιά κρίσιμη τιμή $\chi^2_{n-1,1-\alpha/2}$ που είναι τέτοια ώστε $\mathrm{P}(\chi^2 < \chi^2_{n-1,1-\alpha/2}) = 1-\alpha/2$. [Οι δείκτες $\alpha/2$ και $1-\alpha/2$ των κρίσιμων τιμών είναι κι οι τιμές της αντίστοιχης αθροιστικής συνάρτησης κατανομής.] Στο Σχήμα 2.6 παρουσιάζεται η κατανομή $\mathcal{X}^2$ για $n-1 = 24$ βαθμούς ελευθερίας καθώς κι οι κρίσιμες τιμές της για $1-\alpha = 0.95$.



Σχήμα 2.6: Συνάρτηση πυκνότητας πιθανότητας για την κατανομή $\mathcal{X}^2$ με 24 βαθμούς ελευθερίας κι η δεξιά κι αριστερή κρίσιμη τιμή της για $1-\alpha = 0.95$.

Στη σχέση (2.15), λύνοντας τις δύο ανισότητες

$$\chi^2_{n-1,\alpha/2} < \frac{(n-1)s^2}{\sigma^2} \leq \chi^2_{n-1,1-\alpha/2}$$

ως προς σ^2 βρίσκουμε το $(1-\alpha)\%$ διάστημα εμπιστοσύνης για τη σ^2

$$\left[\frac{(n-1)s^2}{\chi^2_{n-1,1-\alpha/2}}, \frac{(n-1)s^2}{\chi^2_{n-1,\alpha/2}}\right], \tag{2.16}$$

όπου η δειγματική διασπορά s^2 υπολογίζεται από το δείγμα των n παρατηρήσεων.

Το 95% δ.ε. για την τυπική απόκλιση σ έχει σαν άκρα τις τετραγωνικές ρίζες των αντίστοιχων άκρων του 95% δ.ε. για τη διασπορά σ^2.

Παράδειγμα 2.12 *Από τα δεδομένα για την αντοχή θραύσης σκυροδέματος τύπου Α του Πίνακα 1.3 θέλουμε να εκτιμήσουμε τη διασπορά σ^2 της αντοχής θραύσης σκυροδέματος τύπου Α. Η σημειακή εκτίμηση βρέθηκε να είναι $s^2 = 0.375\,(\mathbf{ksi})^2$. Για $n-1 = 24$ και $\alpha = 0.05$ από τον στατιστικό πίνακα για τη $\mathcal{X}^2$ βρίσκουμε $\chi^2_{24,\,0.025} = 12.4$ και $\chi^2_{24,\,0.975} = 39.4$ (δες Σχήμα 2.6). Το 95% διάστημα εμπιστοσύνης για τη σ^2 είναι*

$$\left[\frac{24\cdot 0.375}{39.4},\, \frac{24\cdot 0.375}{12.4}\right] = [0.228,\, 0.726].$$

Το 95% δ.ε. για την τυπική απόκλιση σ της αντοχής θραύσης σκυροδέματος είναι

$$[\sqrt{0.228},\, \sqrt{0.726}] = [0.478,\, 0.852].$$

Στο παραπάνω παράδειγμα παρατηρούμε ότι η εκτίμηση $s^2 = 0.375\,(\mathbf{ksi})^2$ είναι πιό κοντά στο αριστερό άκρο του διαστήματος εμπιστοσύνης. Γενικά το διάστημα εμπιστοσύνης για τη διασπορά σ^2 δεν είναι συμμετρικό ως προς τη σημειακή εκτίμηση s^2 κι αυτό γιατί η κατανομή $\mathcal{X}^2$ δεν είναι συμμετρική όπως είναι η κανονική κατανομή κι η κατανομή student. Όσο μεγαλώνουν όμως οι βαθμοί ελευθερίας (δηλαδή το μέγεθος δείγματος) η $\mathcal{X}^2$ κατανομή προσεγγίζει την κανονική κατανομή (δες Σχήμα 2.5). Γι αυτό για πολύ μεγάλα δείγματα το διάστημα εμπιστοσύνης μπορεί να υπολογισθεί κι από άλλο τύπο που περιέχει κρίσιμες τιμές της τυπικής κανονικής κατανομής.

2.2.3 Διάστημα εμπιστοσύνης της αναλογίας p

Σε αρκετά προβλήματα τα δεδομένα δεν είναι αριθμητικές τιμές μιας τ.μ. του πληθυσμού αλλά δυαδικές τιμές, δηλαδή κάποιο στοιχείο του πληθυσμού έχει μια ιδιότητα (επιτυχία ή 1) ή δεν την έχει (αποτυχία ή 0). Ο λόγος των στοιχείων του πληθυσμού που πληρούν την ιδιότητα προς το σύνολο όλων των στοιχείων του πληθυσμού λέγεται **αναλογία** p. [Η αναλογία p είναι η πιθανότητα 'επιτυχίας σε μια δοκιμή' όταν αναφερόμαστε σε ακολουθίες Bernoulli.] Ένα παράδειγμα είναι η αναλογία κοριτσιών στις Πολυτεχνικές Σχολές της Ελλάδας ή η αναλογία σκουριασμένων ράβδων χάλυβα σε μια αποθήκη. Σε πολλές περιπτώσεις θέλουμε να εκτιμήσουμε την αναλογία p από ένα δείγμα μεγέθους n. Η σημειακή εκτίμηση της p είναι απλά $\hat{p} = \frac{m}{n}$, ο λόγος των 'επιτυχίων' m στο δείγμα προς το πλήθος n των στοιχείων του δείγματος.

Γνωρίζουμε ότι για μεγάλο n η κατανομή της εκτιμήτριας $\hat{p}$ είναι κανονική με μέση τιμή $\mathrm{E}(\hat{p}) = p$ και διασπορά $\mathrm{Var}(\hat{p}) = \frac{p(1-p)}{n}$, δηλαδή ισχύει

$$\hat{p} \sim \mathrm{N}(p, \frac{p(1-p)}{n}).$$

Ακολουθώντας την ίδια διαδικασία όπως για τη μέση τιμή με γνωστή διασπορά, με τη βοήθεια του μετασχηματισμού

$$z \equiv \frac{\hat{p}-p}{\sqrt{p(1-p)/n}}$$

καταλήγουμε στο διάστημα εμπιστοσύνης για την αναλογία p

$$\hat{p} \pm z_{1-\alpha/2}\sqrt{\frac{p(1-p)}{n}}.$$

Αντικαθιστώντας στο τυπικό σφάλμα την αναλογία p με την δειγματική αναλογία $\hat{p}$ έχουμε το διάστημα εμπιστοσύνης για την p από το δείγμα

$$\hat{p} \pm z_{1-\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}. \tag{2.17}$$

Το εύρος του δ.ε. της αναλογίας p είναι

$$w = 2\, z_{1-\alpha/2} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}. \tag{2.18}$$

Λύνοντας την παραπάνω σχέση ως προς n βρίσκουμε

$$n = \left(\frac{2\, z_{1-\alpha/2}}{w}\right)^2 \hat{p}(1-\hat{p}).$$

Πριν από τη μελέτη του δείγματος δε γνωρίζουμε την εκτίμηση της αναλογίας p. Άρα από την παραπάνω σχέση δε μπορούμε να βρούμε τον αριθμό των παρατηρήσεων που θα χρειαστούμε για να υπολογίσουμε δ.ε. της p εύρους w. Μπορεί να δειχθεί όμως πως η μέγιστη τιμή που μπορεί να πάρει η έκφραση $\hat{p}(1-\hat{p})$ είναι 0.25. Άρα το μέγεθος δείγματος που χρειαζόμαστε για να εξασφαλίσουμε μέγιστο εύρος w του δ.ε. είναι

$$n = \left(\frac{2z_{1-\alpha/2}}{w}\right)^2 0.25 = \left(\frac{z_{1-\alpha/2}}{w}\right)^2. \tag{2.19}$$

Παράδειγμα 2.13 *Για να εκτιμήσουμε την αναλογία σκουριασμένων ραβδών χάλυβα μιας αποθήκης πήραμε ένα δείγμα από $n = 100$ ράβδους και βρήκαμε $m = 12$ σκουριασμένες.*

Η σημειακή εκτίμηση για την αναλογία p των σκουριασμένων ραβδών είναι $\hat{p} = \frac{12}{100} = 0.12$. Το 95% διάστημα εμπιστοσύνης είναι ($z_{1-\alpha/2} = z_{0.975} = 1.96$)

$$0.12 \pm 1.96\sqrt{\frac{0.12 \cdot (1-0.12)}{100}} \quad \rightarrow \quad [0.056,\, 0.184].$$

Άρα σε 95% επίπεδο εμπιστοσύνης και με βάση το δείγμα μας μπορούμε να πούμε ότι το διάστημα $[0.056,\, 0.184]$ θα περιέχει την πραγματική αναλογία σκουριασμένων ραβδών χάλυβα, την οποία θα μπορούσαμε να τη βρούμε αν καταμετρούσαμε όλες τις ράβδους στην αποθήκη.

Έστω ότι θέλουμε να εκτιμήσουμε το διάστημα εμπιστοσύνης της αναλογίας p των σκουριασμένων ραβδών στο ίδιο επίπεδο εμπιστοσύνης 95% αλλά με μικρότερο εύρος $w = 0.05$ (ή αν αναφερόμαστε σε ποσοστά το εύρος να είναι 5 εκατοστιαίες μονάδες). Από την σχέση (2.19) βρίσκουμε ότι το μέγιστο μέγεθος n του δείγματος που θα χρειαστούμε είναι (αντικαθιστώντας τις τιμές $z_{1-\alpha/2} = z_{0.975} = 1.96$ και $w = 0.05$)

$$n = \left(\frac{1.96}{0.05}\right)^2 = 1536.6 \simeq 1537,$$

δηλαδή για να μειώσουμε το εύρος του 95% διαστήματος εμπιστοσύνης της αναλογίας σκουριασμένων ραβδών περίπου στο 1/3 πρέπει να αυξήσουμε το μέγεθος του δείγματος πάνω από 15 φορές.

2.2.4 Διάστημα εμπιστοσύνης της διαφοράς δύο μέσων τιμών $\mu_1 - \mu_2$

Θέλουμε να εκτιμήσουμε τη διαφορά των μέσων τιμών μ_1 και μ_2 δύο ανεξάρτητων τ.μ. X_1 και X_2 έχοντας δύο δείγματα μεγέθους n_1 και n_2 από τον πληθυσμό της X_1 και τον πληθυσμό της X_2 αντίστοιχα.

Η σημειακή εκτίμηση της διαφοράς $\mu_1 - \mu_2$ είναι απλά η διαφορά των δειγματικών μέσων τιμών $\bar{x}_1 - \bar{x}_2$. Για το διάστημα εμπιστοσύνης της διαφοράς $\mu_1 - \mu_2$ πρέπει να ελέγξουμε την κατανομή της εκτιμήτριας $\bar{x}_1 - \bar{x}_2$, όπως κάναμε στην εκτίμηση διαστήματος εμπιστοσύνης για τη μέση τιμή μ.

Γνωστές διασπορές Θεωρούμε ότι γνωρίζουμε τις διασπορές σ_1^2 και σ_2^2 των X_1 και X_2. Υποθέτουμε επίσης ότι το δείγμα είναι αρκετά μεγάλο **ή** η κατανομή των X_1 και X_2 είναι κανονική. Τότε η εκτιμήτρια $\bar{x}_1 - \bar{x}_2$ ακολουθεί κανονική κατανομή με μέση τιμή $\mu_1 - \mu_2$ και διασπορά $\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$. Αν οι X_1 και X_2 είναι *ομοσκεδαστικές*, δηλαδή $\sigma_1^2 = \sigma_2^2 = \sigma^2$, τότε η διασπορά είναι $\sigma^2(\frac{1}{n_1} + \frac{1}{n_2})$.

Η διαδικασία είναι ίδια όπως για την εύρεση του διαστήματος εμπιστοσύνης για τη μέση τιμή μ αν κάνουμε τις παρακάτω αντικαταστάσεις:

$$
\begin{array}{rccc}
\text{εκτιμήτρια} & \bar{x} & \longrightarrow & \bar{x}_1 - \bar{x}_2 \\
\text{μέση τιμή εκτιμήτριας} & \mu & \longrightarrow & \mu_1 - \mu_2 \\
\text{διασπορά εκτιμήτριας} & \dfrac{\sigma^2}{n} & \longrightarrow & \dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}.
\end{array}
$$

Επομένως το διάστημα εμπιστοσύνης της διαφοράς $\mu_1 - \mu_2$ είναι

$$
(\bar{x}_1 - \bar{x}_2) \pm z_{1-\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}. \tag{2.20}
$$

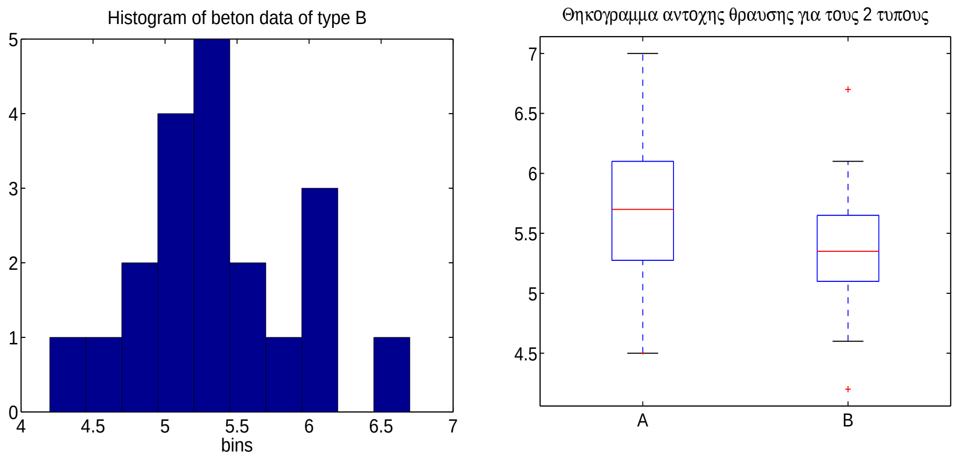
Στην πράξη, όταν εκτιμάμε το διάστημα εμπιστοσύνης της διαφοράς $\mu_1 - \mu_2$ είναι γιατί θέλουμε να διαπιστώσουμε αν κατά μέσο όρο η μια τ.μ. είναι διαφορετική (μεγαλύτερη ή μικρότερη) από την άλλη και αν ναι να εκτιμήσουμε το μέγεθος αυτής της διαφοράς. Το διάστημα εμπιστοσύνης λοιπόν το ερμηνεύουμε ως εξής:

- Αν περιέχει το μηδέν τότε δε μπορούμε να πούμε ότι οι μέσες τιμές των τ.μ. X_1 και X_2 διαφέρουν για το επίπεδο εμπιστοσύνης που χρησιμοποιήσαμε και με βάση τα συγκεκριμένα δεδομένα.

- Αν είναι θετικό τότε μπορούμε να πούμε πως για το επίπεδο εμπιστοσύνης που χρησιμοποιήσαμε η τ.μ. X_1 είναι κατά μέσο όρο μεγαλύτερη από τη X_2 κατά ένα ποσό που κυμαίνεται στα όρια του διαστήματος που εκτιμήσαμε. Ανάλογα ερμηνεύουμε το διάστημα εμπιστοσύνης όταν είναι αρνητικό.

Παράδειγμα 2.14 *Ενδιαφερόμαστε να διαπιστώσουμε αν η αντοχή θραύσης σκυροδέματος ενός τύπου Α είναι κατά μέσο όρο διαφορετική από αυτή ενός άλλου τύπου Β. Γι αυτό θέλουμε να εκτιμήσουμε το 95% διάστημα εμπιστοσύνης για τη διαφορά των μέσων τιμών μ_1 και μ_2 της αντοχής θραύσης σκυροδέματος του τύπου Α και Β αντίστοιχα. Δίνεται ότι η διασπορά είναι κοινή και γνωστή και είναι $\sigma^2 = 0.38$ (ksi)^2.*

Στον Πίνακα 1.3 έχουμε 25 μετρήσεις αντοχής θραύσης σκυροδέματος από τον τύπο Α και 20 από τον τύπο Β. Στο Παράδειγμα 2.9 είδαμε με τη βοήθεια ιστογράμματος (Σχήμα 1.2) και θηκογράμματος (Σχήμα 2.2) πως η αντοχή θραύσης του σκυροδέματος τύπου Α θα μπορούσε να ακολουθεί κανονική κατανομή. Από το ιστόγραμμα και το θηκόγραμμα του Σχήματος 2.7 μπορούμε να πούμε το ίδιο και για τον τύπο Β. Άρα μπορούμε να υποθέσουμε πως οι τ.μ. αντοχής



Σχήμα 2.7: Ιστόγραμμα των δεδομένων αντοχής θραύσης σκυροδέματος του τύπου Β του Πίνακα 1.3 και θηκογράμματα των δεδομένων αντοχής θραύσης σκυροδέματος των δύο τύπων.

θραύσης X_1 και X_2 και για τους δύο πληθυσμοούς, δηλαδή τους δύο τύπους σκυροδέματος, ακολουθούν κανονική κατανομή. Συγκρίνοντας τα θηκογράμματα για τον τύπο Α και Β (δες Σχήμα 2.2) φαίνεται ότι η κεντρική τάση (εδώ διάμεσος) της αντοχής θραύσης σκυροδέματος τύπου Β να είναι χαμηλότερη από αυτή του τύπου Α, αλλά ίσως όχι σημαντικά αφού οι δύο θήκες (τα διαστήματα των 50% κεντρικών τιμών του κάθε δείγματος) επικαλύπτονται κατά μεγάλο ποσοστό. Στη συνέχεια θα υπολογίσουμε αν αυτή η διαφορά είναι στατιστικά σημαντική κάνοντας χρήση του διαστήματος εμπιστοσύνης της $\mu_1 - \mu_2$.

Οι δειγματικές μέσες τιμές υπολογίζονται σε $\bar{x}_1 = 5.67$ ksi και $\bar{x}_2 = 5.38$ ksi. Η διαφορά τους είναι $\bar{x}_1 - \bar{x}_2 = 0.29$ ksi. Από τη σχέση (2.20), όπου $z_{1-\alpha/2} = z_{0.975} = 1.96$ και $\sigma_1^2 = \sigma_2^2 = \sigma^2 = 0.38$ βρίσκουμε ότι το 95% διάστημα εμπιστοσύνης είναι

$$0.29 \pm 1.96\sqrt{0.38 \cdot \left(\frac{1}{25} + \frac{1}{20}\right)} \quad \rightarrow \quad [-0.073,\, 0.653].$$

Συμπεραίνουμε λοιπόν πως σε επίπεδο εμπιστοσύνης 95% δε μπορούμε να πούμε πως οι δύο τύποι σκυροδέματος διαφέρουν σημαντικά ως προς τη μέση αντοχή θραύσης. Το διάστημα $[-0.073,\, 0.653]$ είναι σχεδόν θετικό αλλά δε μας επιτρέπει να συμπεράνουμε στατιστικά σημαντική διαφορά. Για τέτοια οριακά αποτελέσματα η μόνη λύση είναι να αυξήσουμε τα δείγματα και να κάνουμε πάλι την εκτίμηση.

Άγνωστες διασπορές Συνήθως όταν δε γνωρίζουμε τις μ_1, μ_2 αγνοούμε και τις σ_1^2, σ_2^2. Όταν το δείγμα είναι μεγάλο ($n > 30$) μπορούμε να αντικαταστήσουμε στη σχέση (2.20) τις σ_1^2, σ_2^2 με τις δειγματικές διασπορές s_1^2, s_2^2 και να εκτιμήσουμε έτσι το διάστημα εμπιστοσύνης για τη διαφορά $\mu_1 - \mu_2$.

Όταν όμως το μέγεθος του ενός ή και των δύο δειγμάτων είναι μικρό η εκτίμηση του διαστήματος εμπιστοσύνης είναι πιο περίπλοκη. Στην περίπτωση που οι κατανομές των X_1 και X_2 δε δίνονται (ή φαίνονται από τα δεδομένα) να είναι κανονικές, δε μπορούμε γενικά να προσδιορίσουμε την κατανομή της διαφοράς $\bar{x}_1 - \bar{x}_2$ για να εκτιμήσουμε έτσι το διάστημα εμπιστοσύνης. Σε μια τέτοια περίπτωση πρέπει να καταφύγουμε σε μή-παραμετρική εκτίμηση.

Στη συνέχεια υποθέτουμε πως οι κατανομές των X_1 και X_2 είναι κανονικές κι επιπλέον ομοσκεδαστικές ($\sigma_1^2 = \sigma_2^2 = \sigma^2$). Σ' αυτήν την περίπτωση ορίζουμε πρώτα τη *δειγματική κοινή διασπορά* s^2 ως συνάρτηση των s_1^2 και s_2^2

$$s^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}. \tag{2.21}$$

Μπορεί να δειχθεί ότι η s^2 είναι αμερόληπτη εκτιμήτρια της κοινής διασποράς σ^2. Με τη βοήθεια της s^2 η εκτίμηση της διασποράς της διαφοράς $\bar{x}_1 - \bar{x}_2$ είναι $s^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)$. Για την εκτιμήτρια $\bar{x}_1 - \bar{x}_2$ μπορούμε να ορίσουμε τον παρακάτω μετασχηματισμό που ακολουθεί κατανομή student με n_1+n_2-2 βαθμούς ελευθερίας

$$t \equiv \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim \mathrm{t}_{n_1+n_2-2}$$

από το οποίο προκύπτει πως το $(1-\alpha)\%$ διάστημα εμπιστοσύνης της διαφοράς $\mu_1 - \mu_2$ είναι

$$(\bar{x}_1 - \bar{x}_2) \pm t_{n_1+n_2-2,1-\alpha/2}\, s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}. \tag{2.22}$$

Παράδειγμα 2.15 *Για το προηγούμενο παράδειγμα ας υποθέσουμε πως η διασπορά αντοχής θραύσης για τους δύο τύπους σκυροδέματος είναι άγνωστη. Από τα δύο ιστογράμματα στα Σχήματα 1.2 και 2.7 καθώς και τα θηκογράμματα στο Σχήμα 2.7 μπορούμε να δεχτούμε οτι οι τ.μ. αντοχής θραύσης και για τους δύο τύπους ακολουθούν κανονική κατανομή και μάλιστα έχουν την ίδια διασπορά (το εύρος των τιμών των δύο δειγμάτων είναι περίπου το ίδιο). [Τα θηκογράμματα δείχνουν κάποια διαφορά στη διασπορά των τιμών μεταξύ των δύο τύπων αλλά για να αποκλίσουμε ότι οι διασπορές μπορεί να είναι ίσες θα πρέπει η μια θήκη να είναι τουλάχιστον διπλάσια της άλλης.] Δεχόμαστε λοιπόν ότι η κατανομή της αντοχής θραύσης και για τα δύο σκυροδέματα είναι κανονική, οι διασπορές είναι ίδιες κι επειδή επιπλέον τα δείγματα είναι σχετικά μικρά συμπεραίνουμε ότι η διαφορά $\bar{x}_1 - \bar{x}_2$ ακολουθεί κατανομή student.*

Η δειγματική διασπορά για τον τύπο A είναι $s_1^2 = 0.375\,(\mathrm{ksi})^2$ και για τον τύπο B είναι $s_2^2 = 0.326\,(\mathrm{ksi})^2$ (σχετικά κοντά). Εφαρμόζοντας τη σχέση (2.21) βρίσκουμε ότι η δειγματική κοινή διασπορά είναι $s^2 = 0.353\,(\mathrm{ksi})^2$ κι επομένως $s = 0.594$ ksi (όπου $n_1 = 25$ και $n_2 = 20$). Οι βαθμοί ελευθερίας είναι $n_1+n_2-2 = 43$ και για επίπεδο εμπιστοσύνης 95% βρίσκουμε από τον στατιστικό πίνακα για την κατανομή student την κρίσιμη τιμή $t_{43,\,0.975} = 2.02$ (πολύ κοντά στην αντίστοιχη κρίσιμη τιμή $z_{0.975} = 1.96$ της τυπικής κανονικής κατανομής γιατί οι βαθμοί ελευθερίας είναι πολλοί). Το 95% διάστημα εμπιστοσύνης είναι ($\bar{x}_1 - \bar{x}_2 = 0.29$)

$$0.29 \pm 2.02 \cdot 0.594\sqrt{\frac{1}{25}+\frac{1}{20}} \quad \rightarrow \quad [-0.07,\, 0.65].$$

Όπως και πριν που η διασπορά ήταν γνωστή δε μπορούμε να πούμε πως οι μέσες τιμές αντοχής θραύσης των δύο τύπων σκυροδέματος διαφέρουν με στατιστική σημαντικότητα.

Παρατηρήσεις

Οι παράγοντες που επηρεάζουν τον υπολογισμό του διαστήματος εμπιστοσύνης της $\mu_1 - \mu_2$ είναι οι ίδιες όπως για τη μ. Σ' αυτές προστίθεται και ο παράγοντας της ισότητας των διασπορών των X_1 και X_2. Στον Πίνακα 2.2 συνοψίζεται η εκτίμηση του διαστήματος εμπιστοσύνης της $\mu_1 - \mu_2$ στις διάφορες περιπτώσεις. Υπάρχει φανερή αντιστοιχία των περιπτώσεων για

διασπορές των X_1,X_2	*κατανομή των* X_1,X_2	n_1, n_2	*κατανομή* της $\bar{x}_1 - \bar{x}_2$	*διάστημα εμπιστοσύνης*
γνωστές	κανονική		$z \equiv \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \sim \mathrm{N}(0,1)$	$(\bar{x}_1 - \bar{x}_2) \pm z_{1-\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$
γνωστές	μη κανονική	μεγάλα	$z \equiv \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \sim \mathrm{N}(0,1)$	$(\bar{x}_1 - \bar{x}_2) \pm z_{1-\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$
γνωστές	μη κανονική	μικρά	—	—
άγνωστες άνισες/ίσες		μεγάλα	$z \equiv \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}} \sim \mathrm{N}(0,1)$	$(\bar{x}_1 - \bar{x}_2) \pm z_{1-\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$
άγνωστες ίσες	κανονική	μικρά	$t \equiv \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim \mathrm{t}_{n_1+n_2-2}$	$(\bar{x}_1 - \bar{x}_2) \pm t_{n_1+n_2-2,1-\alpha/2}\, s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$
άγνωστες ίσες	μη κανονική	μικρά	—	—
άγνωστες άνισες		μικρά	—	—

Πίνακας 2.2: Εκτίμηση του διαστήματος εμπιστοσύνης της διαφοράς $\mu_1 - \mu_2$ ανάλογα με τη γνώση των διασπορών και κατανομών των τ.μ. X_1 και X_2 καθώς και των μεγεθών n_1 και n_2 των αντιστοίχων δειγμάτων.

το διάστημα εμπιστοσύνης της μέσης τιμής (Πίνακας 2.1) και για το διάστημα εμπιστοσύνης της διαφοράς δύο μέσων τιμών (Πίνακας 2.2, οι έξι πρώτες σειρές). Για τη διαφορά μέσων τιμών υπάρχει ακόμα η περίπτωση των άνισων κι αγνώστων διασπορών σε συνδιασμό με μικρά δείγματα (τελευταία σειρά του Πίνακα 2.2) για την οποία δεν μπορούμε να καθορίσουμε την κατανομή της εκτιμήτριας $\bar{x}_1 - \bar{x}_2$ και απ' αυτήν να βρούμε το διάστημα εμπιστοσύνης. Σ' αυτήν την περίπτωση ούτε η μη παραμετρική εκτίμηση μπορεί να δώσει διάστημα εμπιστοσύνης (για τη διάμεσο).

Για το διάστημα εμπιστοσύνης της διαφοράς μέσων τιμών υποθέσαμε οτι οι τ.μ. X_1 και X_2 είναι ανεξάρτητες. Δε θα ασχοληθούμε με την περίπτωση που οι X_1 και X_2 είναι εξαρτημένες (όταν δηλαδή έχουμε *ζευγαρωτές παρατηρήσεις*) γιατί τέτοια προβλήματα δεν παρουσιάζονται συχνά στη μηχανική.

2.2.5 Διάστημα της διαφοράς δύο αναλογιών $p_1 - p_2$

Η εκτίμηση διαστήματος εμπιστοσύνης της διαφοράς δύο αναλογιών $p_1 - p_2$ παρουσιάζεται όταν θέλουμε να συγκρίνουμε δύο πληθυσμούς ως προς μια ιδιότητα, δηλαδή αν η αναλογία p_1 των στοιχείων που πληρούν μια ιδιότητα στον ένα πληθυσμό είναι διαφορετική, κι αν ναι κατά πόσο, από την αντίστοιχη αναλογία p_2 του άλλου πληθυσμού. Για παράδειγμα, θέλουμε να δούμε κατά πόσο διαφέρει η αναλογία κοριτσιών στην Πολυτεχνική Σχολή και στη Φυσικομαθηματική ή θέλουμε να διερευνήσουμε αν οι αναλογίες σκουριασμένων ραβδών χάλυβα σε δύο αποθήκες διαφέρουν και κατά πόσο.

Η σημειακή εκτίμηση της διαφοράς $p_1 - p_2$ δίνεται από τη διαφορά των $\hat{p}_1 = \frac{m_1}{n_1}$ και $\hat{p}_2 = \frac{m_2}{n_2}$, όπου m_1 είναι το πλήθος 'επιτυχιών' στο δείγμα μεγέθους n_1 από τον πρώτο πληθυσμό και m_2 είναι το πλήθος 'επιτυχιών' στο δείγμα μεγέθους n_2 από το δεύτερο πληθυσμό. Γνωρίζουμε ότι για μεγάλα n_1 και n_2 η κατανομή της εκτιμήτριας $\hat{p}_1 - \hat{p}_2$ είναι κανονική με μέση τιμή $\mathrm{E}(\hat{p}_1 - \hat{p}_2) = p_1 - p_2$ και διασπορά $\mathrm{Var}(\hat{p}_1 - \hat{p}_2) = \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}$ κι άρα

$$z \equiv \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\dfrac{p_1(1-p_1)}{n_1} + \dfrac{p_2(1-p_2)}{n_2}}} \sim \mathrm{N}(0,1).$$

Αντικαθιστώντας στον τύπο της διασποράς τις άγνωστες αναλογίες p_1 και p_2 με τις δειγματικές εκτιμήσεις $\hat{p}_1$ και $\hat{p}_2$ βρίσκουμε το $(1-\alpha)\%$ διάστημα εμπιστοσύνης της διαφοράς $p_1 - p_2$

$$(\hat{p}_1 - \hat{p}_2) \pm z_{1-\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}. \tag{2.23}$$

Για την εκτίμηση της διασποράς της εκτιμήτριας $\hat{p}_1 - \hat{p}_2$, μπορούμε να υποθέσουμε ότι οι αναλογίες p_1 και p_2 δε διαφέρουν παρά πολύ έτσι ώστε να μπορούμε να χρησιμοποιήσουμε την εκτίμηση της κοινής αναλογίας $\hat{p}$ που ορίζεται ως

$$\hat{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2}.$$

Τότε η διασπορά της $\hat{p}_1 - \hat{p}_2$ δίνεται ως $\hat{p}(1-\hat{p})(\frac{1}{n_1} + \frac{1}{n_2})$ και το διάστημα εμπιστοσύνης της διαφοράς $p_1 - p_2$ δίνεται από τη σχέση

$$(\hat{p}_1 - \hat{p}_2) \pm z_{1-\alpha/2}\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}. \tag{2.24}$$

Παράδειγμα 2.16 *Για να εκτιμήσουμε αν υπάρχει διαφορά στο ποσοστό σκουριασμένων ραβδών χάλυβα δύο αποθηκών πήραμε ένα δείγμα 100 ραβδών από την πρώτη αποθήκη και βρήκαμε 12 σκουριασμένες κι ένα δείγμα 120 ραβδών από τη δεύτερη αποθήκη και βρήκαμε 26 σκουριασμένες.*

Η εκτίμηση από τα δείγματα για τις αναλογίες σκουριασμένων ραβδών είναι $\hat{p}_1 = \frac{12}{100} = 0.12$ για την πρώτη αποθήκη και $\hat{p}_2 = \frac{26}{120} = 0.217$ για τη δεύτερη αποθήκη. Η διαφορά είναι $\hat{p}_1 - \hat{p}_2 = -0.097$ (σε ποσοστό 9.7%) κι από το 95% διάστημα εμπιστοσύνης μπορούμε να κρίνουμε αν είναι στατιστικά σημαντική. Από τη σχέση (2.23) και για $z_{1-\alpha/2} = 1.96$ βρίσκουμε

$$-0.097 \pm 1.96\sqrt{\frac{0.12 \cdot 0.88}{100} + \frac{0.217 \cdot 0.783}{120}} \quad \to \quad [-0.198,\ 0.004].$$

Το 95% διάστημα εμπιστοσύνης περιέχει οριακά το μηδέν και γι αυτό συμπεραίνουμε πως με αυτά τα δείγματα και σ' αυτό το επίπεδο εμπιστοσύνης η δειγματική διαφορά των αναλογιών αν και είναι σε ποσοστό περίπου 10% δεν είναι στατιστικά σημαντική. Φαίνεται όμως το ποσοστό σκουριασμένων ραβδών στη πρώτη αποθήκη να είναι μικρότερο. Ίσως θα έπρεπε να αυξήσουμε το δείγμα μας και να πάρουμε έτσι στενότερο διάστημα εμπιστοσύνης για να διαπιστώσουμε αν πράγματι περιέχει το 0 ή όχι, αν δηλαδή η διαφορά των δύο αναλογιών είναι στατιστικά σημαντική.

Τα διαστήματα εμπιστοσύνης, πέρα από το ότι δίνουν ένα εύρος τιμών για την παράμετρο που μελετάμε, μας δίνουν τη δυνατότητα να ελέγξουμε αν μια συγκεκριμένη τιμή της παραμέτρου είναι πιθανή βλέποντας αν ανήκει στο διάστημα εμπιστοσύνης. Έτσι, στο τελευταίο παράδειγμα, ελέγχοντας αν η τιμή μηδέν ανήκει στο διάστημα εμπιστοσύνης μας επιτρέπει στην ουσία να ελέγξουμε αν οι δύο αναλογίες διαφέρουν σημαντικά ή όχι.

Στο επόμενο κεφάλαιο θα παρουσιάσουμε τέτοιους στατιστικούς ελέγχους για τις ίδιες παραμέτρους για τις οποίες εκτιμήσαμε διαστήματα εμπιστοσύνης.