Σημειακή εκτίμηση και εκτίμηση με διάστημα

Αρκετά τρόφιμα περιέχουν το ιχνοστοιχείο σελήνιο το οποίο, όταν προσλαμβάνεται σε μικρές ποσότητες ημερησίως, έχει ευεργετική επίδραση στην υγεία. Ένας φοιτητής, στο πλαίσιο της πτυχιακής του εργασίας, πρέπει να μελετήσει την ημερήσια ποσότητα σεληνίου που προσλαμβάνουν οι ενήλικες μέσω της διατροφής τους σε δύο συγκεκριμένες περιοχές της χώρας, έστω Α και Β, όπου κατεξοχήν καταναλώνονται τρόφιμα τοπικής παραγωγής (η περιεκτικότητα των τροφίμων σε σελήνιο ποικίλει ανάλογα με την περιοχή παραγωγής τους). Η ημερήσια ποσότητα σεληνίου που προσλαμβάνει ο άνθρωπος από τις τροφές είναι προφανώς τυχαία μεταβλητή και από τη βιβλιογραφία ο φοιτητής γνωρίζει ότι αυτή ακολουθεί κανονική κατανομή. Δε γνωρίζει όμως τις τιμές των παραμέτρων της, τόσο για την περιοχή Α όσο και για την περιοχή Β και ειδικότερα για τους ενήλικες αυτών των περιοχών. Δηλαδή, αν Χ η τυχαία μεταβλητή που εκφράζει την ημερήσια ποσότητα σεληνίου που προσλαμβάνουν από τη διατροφή τους οι ενήλικες στην περιοχή Α και Υ η τυχαία μεταβλητή που εκφράζει την ημερήσια ποσότητα σεληνίου που προσλαμβάνουν από τη διατροφή τους οι ενήλικες στην περιοχή Β, ο φοιτητής γνωρίζει ότι $X \sim N(\mu_1, \sigma_1^2)$ *και* $Y \sim N(\mu_2, \sigma_2^2)$ *όμως δε γνωρίζει τις τιμές των παραμέτρων τους* μ_1, σ_1^2 *και* μ_2, σ_2^2, *αντίστοιχα. Για το σκοπό αυτό επέλεξε με μια τυχαία διαδικασία 16 ενήλικες από κάθε περιοχή και για καθέναν υπολόγισε/μέτρησε (με κάποια μέθοδο) την ημερήσια ποσότητα σεληνίου που προσλαμβάνει μέσω της διατροφής του. Μετά την επεξεργασία των δεδομένων που συγκέντρωσε, προέκυψε ότι ο δειγματικός μέσος και η δειγματική τυπική απόκλιση για το δείγμα από την περιοχή Α αντίστοιχα είναι* $\bar{x} = 75\mu gr$ *και* $s_1 = 8.4\mu gr$ *και για το δείγμα από την περιοχή Β, αντίστοιχα είναι* $\bar{y} = 87\mu gr$ *και* $s_2 = 9\mu gr$. *Με βάση τα ευρήματα σε αυτά τα δύο τυχαία δείγματα, τι μπορεί άραγε να συμπεράνει ο φοιτητής για την άγνωστη μέση τιμή* μ_1, *δηλαδή, για τη μέση ημερήσια ποσότητα σεληνίου που προσλαμβάνουν οι ενήλικες στην περιοχή Α, και για την, άγνωστη επίσης, μέση τιμή* μ_2, *αλλά και για το πώς αυτές συγκρίνονται, δηλαδή, αν είναι ίσες ή διαφέρουν και ποια είναι μεγαλύτερη (και πόσο) αν διαφέρουν. Επίσης, τι μπορεί να συμπεράνει για τις άγνωστες διακυμάνσεις* σ_1^2 *και* σ_2^2 *και για το πώς αυτές συγκρίνονται.*

Αυτά τα ερωτήματα είναι τυπικά παραδείγματα ερωτημάτων στα οποία δίνει απαντήσεις η *στατιστική συμπερασματολογία*. Στην ενότητα αυτή, θα δούμε (... επιτέλους) πώς, για παράδειγμα, ο φοιτητής μπορεί, με βάση τα ευρήματα στο *τυχαίο δείγμα* από την περιοχή Α, να «πει κάτι» για τη *μέση τιμή* μ_1 της *τυχαίας μεταβλητής Χ* της οποίας δε γνωρίζει επακριβώς την κατανομή (γνωρίζει μόνο την οικογένεια κατανομών στην οποία ανήκει) ή τι μπορεί να πει για το αν οι *μέσες τιμές* μ_1 και μ_2 είναι ίσες ή διαφέρουν (και πόσο) και αντίστοιχα για τις *διακυμάνσεις* σ_1^2 και σ_2^2. Θα δούμε επίσης, τι απαντήσεις μπορεί να δώσει σε αυτά τα ερωτήματα ακόμη και αν δε γνωρίζει σε ποια οικογένεια κατανομών ανήκει η κατανομή της *τυχαίας μεταβλητής Χ* (και της *Υ)*.

11.1 Εκτιμήτριες συναρτήσεις και μέθοδοι εκτίμησης

Ένα από τα γενικότερα θέματα που τίθενται με τα ερωτήματα που απασχολούν το φοιτητή, είναι το εξής: ***πώς μπορούμε να εκτιμήσουμε μία ή περισσότερες άγνωστες παραμέτρους της κατανομής μιας τυχαίας μεταβλητής με βάση ένα τυχαίο δείγμα από αυτήν;***

Εφόσον θεωρούμε ότι το *τυχαίο δείγμα* που έχουμε στη διάθεσή μας είναι ένα αντιπροσωπευτικό δείγμα των δυνατών τιμών της τυχαίας μεταβλητής που μελετάμε, είναι λογικό να εκτιμήσουμε τις άγνωστες παραμέτρους της μέσω «κατάλληλων» *στατιστικών συναρτήσεων* που θα κατασκευάσουμε και θα μπορούμε (όπως έχουμε εξηγήσει) να υπολογίσουμε από το δείγμα. Για παράδειγμα, για να εκτιμήσουμε την άγνωστη *μέση τιμή* μ μιας *τυχαίας μεταβλητής*, έστω *X*, είναι λογικό να χρησιμοποιήσουμε τον *δειγματικό μέσο* $\overline{X}$ που είναι μια συνάρτηση του δείγματος και μπορεί να υπολογισθεί από αυτό.

Κάθε *στατιστική συνάρτηση* που χρησιμοποιείται για την εκτίμηση μιας άγνωστης *παραμέτρου* ενός *πληθυσμού* (δηλαδή, της *κατανομής μιας τ.μ.*) ονομάζεται ***εκτιμήτρια συνάρτηση*** ή ***εκτιμήτρια (estimator)*** της άγνωστης *παραμέτρου*. Η τιμή της *εκτιμήτριας* για συγκεκριμένη πραγματοποίηση του *τυχαίου δείγματος,* ονομάζεται ***εκτίμηση (estimation)*** της άγνωστης *παραμέτρου*.

Έτσι, στο παράδειγμά μας, αν ως *εκτιμήτρια* της άγνωστης *μέσης τιμής* μ_1, της *τυχαίας μεταβλητής X*, επιλέξουμε τον *δειγματικό μέσο*

$$\overline{X} = \frac{1}{16}\sum_{i=1}^{16} X_i$$

τότε, η τιμή του $\overline{x} = 75\mu gr$ για το συγκεκριμένο *τυχαίο δείγμα* που πήρε ο φοιτητής από την περιοχή Α, μπορεί να χρησιμοποιηθεί ως μια *εκτίμηση* της άγνωστης *μέσης τιμής* μ_1 της *X,* δηλαδή, της άγνωστης *μέσης ημερήσιας ποσότητας σεληνίου που προσλαμβάνουν οι ενήλικες στην περιοχή Α μέσω της διατροφής τους.* Αντίστοιχα, η τιμή $\overline{y} = 87\mu gr$ του *δειγματικού μέσου*

$$\overline{Y} = \frac{1}{16}\sum_{i=1}^{16} Y_i$$

μπορεί να χρησιμοποιηθεί ως μια *εκτίμηση* της άγνωστης *μέσης τιμής* μ_2.

11.1.1 Σημειακή εκτίμηση

Η *εκτίμηση* μιας άγνωστης *παραμέτρου* ενός πληθυσμού μέσω της τιμής μιας *εκτιμήτριας* (για συγκεκριμένη πραγματοποίηση ενός *τυχαίου δείγματος* από τον *πληθυσμό*), ονομάζεται ***σημειακή εκτίμηση (point estimation)*** της *παραμέτρου* και η *εκτιμήτρια* ονομάζεται ***σημειακή εκτιμήτρια (point estimator)***.

Είναι προφανές, ότι από κάθε πραγματοποίηση του *τυχαίου δείγματος,* παίρνουμε μια διαφορετική (εν γένει) *σημειακή εκτίμηση* της άγνωστης *παραμέτρου* του *πληθυσμού*. Έτσι, αν ο φοιτητής έπαιρνε ένα άλλο *τυχαίο δείγμα* μεγέθους 16 από την περιοχή Α και εύρισκε, για παράδειγμα, *δειγματικό μέσο* $\overline{x} = 74.5\mu gr$, αυτή θα ήταν μια άλλη *σημειακή εκτίμηση* της *μέσης τιμής* μ_1 του *πληθυσμού*.

Επίσης, αν ως *εκτιμήτρια* της άγνωστης *πληθυσμιακής μέσης τιμής,* ο φοιτητής δεν επέλεγε τον *δειγματικό μέσο* $\overline{X}$ αλλά κατασκεύαζε μια άλλη *στατιστική συνάρτηση*, για παράδειγμα, την

$$T = \frac{X_{\min} + X_{\max}}{2}$$

δηλαδή, το ημιάθροισμα της μικρότερης και της μεγαλύτερης τιμής του δείγματος, προφανώς, θα έπαιρνε μια διαφορετική *σημειακή εκτίμηση* από αυτήν που πήρε, για την ίδια πραγματοποίηση του *τυχαίου δείγματος,* από τον *δειγματικό μέσο*.

Γεννώνται επομένως δύο εύλογα ερωτήματα.
α) Υπάρχει κάποια μέθοδος για να καθορίζουμε/κατασκευάζουμε εκτιμήτριες με βάση κάποια κριτήρια και όχι αυθαίρετα;

*β) Ποιο είναι το **«σφάλμα της εκτίμησης»** που κάνουμε, δηλαδή, πόσο κοντά στην πραγματική τιμή της παραμέτρου βρίσκεται η τιμή της εκτιμήτριας που υπολογίζουμε από την πραγματοποίηση ενός τυχαίου δείγματος;*

Σε ότι αφορά το πρώτο ερώτημα η απάντηση (ασφαλώς) είναι ναι. Μια πολύ γνωστή μέθοδος κατασκευής *εκτιμητριών* είναι η ***μέθοδος μέγιστης πιθανοφάνειας (method of maximum likelihood)*** που προτάθηκε από τον *Ronald A. Fisher (1922)*. Μια άλλη, επίσης πολύ γνωστή μέθοδος/τεχνική *εκτίμησης,* είναι η ***μέθοδος των ροπών (method of moments)*** που προτάθηκε από τον *Karl Pearson (1894).* Επίσης, για τις *εκτιμήτριες* έχουν ορισθεί επιθυμητές ιδιότητες, ή αλλιώς, κριτήρια «καλής συμπεριφοράς» τους, όπως ***αμεροληψία***, ***συνέπεια*** και ***αποτελεσματικότητα***. Ας δούμε το νόημα αυτών των ιδιοτήτων (στις μεθόδους εκτίμησης δε θα αναφερθούμε).

11.1.2 Ιδιότητες των εκτιμητριών

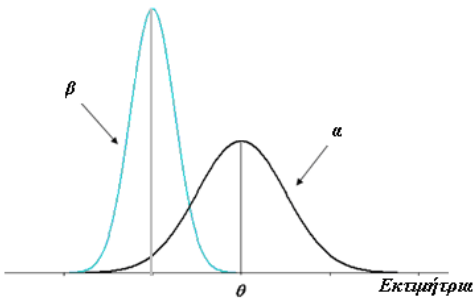
Είναι προφανές, ότι μια *εκτιμήτρια* μπορεί να χαρακτηρισθεί «καλή» ή όχι, δηλαδή, μπορούμε να θεωρήσουμε ότι έχει ή δεν έχει «καλές» ιδιότητες, ανάλογα με τη συμπεριφορά της σε επαναλαμβανόμενες δειγματοληψίες (από δείγμα σε δείγμα), η οποία περιγράφεται από την κατανομή της, ή αλλιώς, από τη ***δειγματοληπτική κατανομή***.

Μια επιθυμητή/καλή ιδιότητα μιας *εκτιμήτριας* θα μπορούσε να είναι η εξής: σε επαναλαμβανόμενες δειγματοληψίες, κατά μέσο όρο να εκτιμά σωστά την άγνωστη *παράμετρο*, δηλαδή, ούτε να την υπερεκτιμά ούτε να την υποεκτιμά συστηματικά. Η ιδιότητα αυτή ονομάζεται ***αμεροληψία (unbiasedness)***.

Μια *εκτιμήτρια*, έστω $\hat{\theta}$, μιας *παραμέτρου* θ, ονομάζεται ***αμερόληπτη (unbiased)***, αν η *μέση τιμή* της είναι ίση με την αληθή/πραγματική τιμή της *παραμέτρου*, δηλαδή, αν

$$E[\hat{\theta}] = \theta .$$

Στο *Σχήμα 11.1.1*, η κατανομή (α) είναι η κατανομή μιας *εκτιμήτριας* μιας *παραμέτρου* θ και η κατανομή (β) είναι η κατανομή μιας άλλης *εκτιμήτριας* της ίδιας *παραμέτρου* θ. Προφανώς, η (α) είναι κατανομή *αμερόληπτης εκτιμήτριας* της *παραμέτρου* θ ενώ η (β), που είναι μετατοπισμένη προς τα αριστερά της αληθούς τιμής της θ, είναι κατανομή *μη αμερόληπτης εκτιμήτριας* της θ και τείνει να υποεκτιμά τη θ.

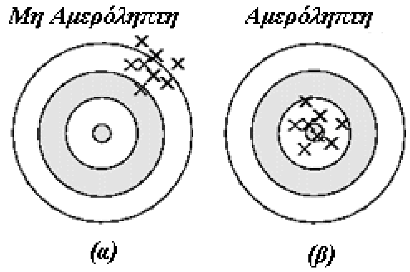


Σχήμα 11.1.1
Η εκτιμήτρια (α) είναι αμερόληπτη εκτιμήτρια της θ
ενώ η (β) είναι μη αμερόληπτη εκτιμήτρια της θ

Στο *Σχήμα 11.1.2* δίνουμε την έννοια της *αμεροληψίας* πιο παραστατικά. Σκεφθείτε ότι ένας σκοπευτής στοχεύει με βέλη προς ένα στόχο και θεωρείστε ότι ο στόχος (το κέντρο του στόχου) είναι η αληθής/πραγματική τιμή μιας *παραμέτρου* θ και ο σκοπευτής είναι η *εκτιμήτρια*.

Οι επαναλαμβανόμενες ρίψεις του βέλους, κατ' αναλογία, αντιστοιχούν στις επαναλαμβανόμενες δειγματοληψίες και το σημείο πρόσπτωσης του βέλους, στην *εκτίμηση* της θ, που φυσικά εξαρτάται από τον σκοπευτή (την *εκτιμήτρια*).

Στην περίπτωση του σκοπευτή-εκτιμήτριας (α) το σημείο πρόσπτωσης του βέλους (η *εκτίμηση*), στις επαναλαμβανόμενες ρίψεις (δειγματοληψίες), βρίσκεται συστηματικά σε συγκεκριμένη περιοχή μακριά από το κέντρο του στόχου (πάνω δεξιά), ενώ στην περίπτωση (β) βρίσκεται γύρω/κοντά στο κέντρο του στόχου (στην αληθή τιμή της *παραμέτρου*).



Σχήμα 11.2
Στην περίπτωση (α) το βέλος πέφτει συστηματικά σε συγκεκριμένη περιοχή μακριά από το κέντρο του στόχου (πάνω δεξιά), ενώ στην περίπτωση (β) πέφτει κοντά στο κέντρο του στόχου

Ας δούμε αν ο *δειγματικός μέσος*

$$\overline{X} = \frac{1}{\nu}\sum_{i=1}^{\nu} X_i$$

είναι *αμερόληπτη εκτιμήτρια* της *μέσης τιμής,* μ, ενός *πληθυσμού,* δηλαδή, αν *κατά μέσο όρο* εκτιμά σωστά τη *μέση τιμή* του *πληθυσμού.*

Στο *Πόρισμα 5.5.2* δείξαμε ότι

$$E(\overline{X}) = E(X) = \mu \,.$$

Επομένως, για οποιονδήποτε πληθυσμό, με *μέση τιμή* μ και *διακύμανση* σ^2, ο *δειγματικός μέσος*

$$\overline{X} = \frac{1}{\nu}\sum_{i=1}^{\nu} X_i$$

(για οποιοδήποτε μέγεθος δείγματος ν) είναι *αμερόληπτη εκτιμήτρια* της *πληθυσμιακής μέσης τιμής* μ.

Επίσης, για οποιονδήποτε πληθυσμό, με *μέση τιμή* μ και *διακύμανση* σ^2, αποδεικνύεται ότι η *δειγματική διακύμανση*

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

(για οποιοδήποτε μέγεθος δείγματος ν) είναι μια *αμερόληπτη εκτιμήτρια* της *πληθυσμιακής διακύμανσης* σ^2.

Πράγματι,

$$E[S^2] = E\left[\frac{1}{\nu-1}\sum_{i=1}^{\nu}(X_i - \overline{X})^2\right] = E\left[\frac{1}{\nu-1}\left(\sum_{i=1}^{\nu}X_i^2 - \nu\overline{X}^2\right)\right] =$$

$$= \frac{1}{\nu-1}\left(\sum_{i=1}^{\nu}E[X_i^2] - \nu E[\overline{X}^2]\right) = \frac{1}{\nu-1}\left(\sum_{i=1}^{\nu}(\mu^2+\sigma^2) - \nu(Var(\overline{X}) + (E[\overline{X}])^2)\right) =$$

$$= \frac{1}{\nu-1}\left(\nu(\mu^2+\sigma^2) - \nu\frac{\sigma^2}{\nu} - \nu\mu^2\right) = \sigma^2.$$

Σημείωση 11.1.1: *Στην προηγούμενη απόδειξη χρησιμοποιήσαμε τη γνωστή από το Α΄ Μέρος σχέση* $Var(X) = \sigma^2 = E[X^2] - \mu^2$ *καθώς και το ότι για οποιονδήποτε πληθυσμό με μέση τιμή* μ και *διακύμανση* σ^2, *η διακύμανση του δειγματικού μέσου για οποιοδήποτε μέγεθος δείγματος* ν, *είναι (όπως δείξαμε στο Πόρισμα 5.5.1) ίση με* σ^2/ν, *δηλαδή*

$$Var(\overline{X}) = \sigma^2_{\overline{X}} = \frac{\sigma^2}{\nu}.$$

■

Παρατηρήσεις 11.1.1: *α) Αν ως εκτιμήτρια της διακύμανσης* σ^2 *ενός πληθυσμού χρησιμοποιήσουμε τη στατιστική συνάρτηση*

$$S_*^2 = \frac{1}{\nu}\sum_{i=1}^{\nu}(X_i - \overline{X})^2$$

εύκολα μπορεί να αποδειχθεί ότι

$$E[S_*^2] = \frac{\nu-1}{\nu}\sigma^2$$

δηλαδή, η S_*^2 *δεν είναι αμερόληπτη εκτιμήτρια της διακύμανσης* σ^2, *αφού κατά μέσο όρο υποεκτιμά τη* σ^2. *Αυτός είναι και ο λόγος που στην Περιγραφική Στατιστική, ως δειγματική διακύμανση ορίσαμε τη στατιστική συνάρτηση*

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

και όχι την S_*^2, *που θα ήταν και πιο «λογικό».*

β) Εύκολα μπορεί να αποδειχθεί (δείτε το ως άσκηση) ότι αν η πληθυσμιακή μέση τιμή μ *είναι* ***γνωστή*** *τότε η στατιστική συνάρτηση*

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

είναι μια αμερόληπτη_εκτιμήτρια της διακύμανσης σ^2.

■

Ο *δειγματικός μέσος* $\overline{X}$, όπως αποδείξαμε, είναι μια *αμερόληπτη εκτιμήτρια* της *πληθυσμιακής μέσης τιμής* μ. Όμως, και η *στατιστική συνάρτηση*

$$T = \frac{X_1 + X_\nu}{2}$$

αν χρησιμοποιηθεί ως *εκτιμήτρια* της *πληθυσμιακής μέσης τιμής* μ, θα την εκτιμά κατά μέσο όρο σωστά, αφού,

$$E[T] = E\left[\frac{X_1 + X_\nu}{2}\right] = \frac{1}{2}(E[X_1] + E[X_\nu]) = \frac{1}{2}(\mu + \mu) = \mu$$

δηλαδή, είναι και αυτή μια *αμερόληπτη εκτιμήτρια* της μ.

Παρατηρείστε επίσης, ότι για οποιουσδήποτε πραγματικούς αριθμούς $\alpha_i,\ i = 1,2,\ldots,\nu$ με $\sum_{i=1}^{\nu} \alpha_i = 1$, η *στατιστική συνάρτηση*

$$T = \sum_{i=1}^{\nu} a_i X_i$$

μπορεί να χρησιμοποιηθεί ως *αμερόληπτη εκτιμήτρια* της *μέσης τιμής* μ οποιουδήποτε *πληθυσμού (δείτε το ως μια απλή άσκηση)*.

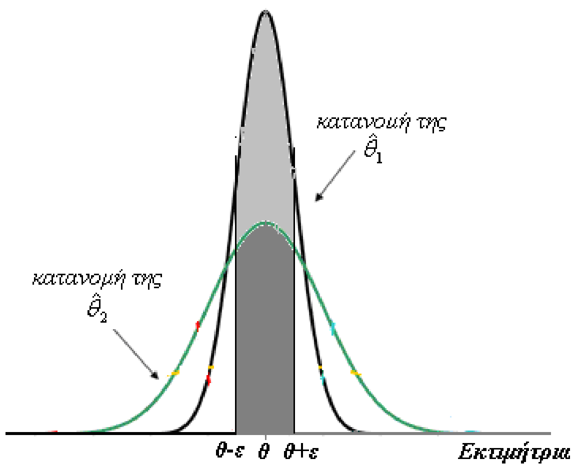
Άραγε, μεταξύ *αμερόληπτων εκτιμητριών* της ίδιας *παραμέτρου* ενός *πληθυσμού*, έχει σημασία ποια θα επιλέξουμε ή μήπως δεν έχει; Η απάντηση είναι ότι έχει. Η *αμεροληψία* δεν είναι η μόνη επιθυμητή ιδιότητα για μια *εκτιμήτρια*. Δείτε στο *Σχήμα 11.1.3* τη *συνάρτηση πυκνότητας* μιας *αμερόληπτης εκτιμήτριας* $\hat{\theta}_1$ μιας *παραμέτρου* θ και τη *συνάρτηση πυκνότητας* μιας άλλης, *αμερόληπτης* επίσης, *εκτιμήτριας* $\hat{\theta}_2$ της ίδιας *παραμέτρου* θ. Η *εκτιμήτρια* $\hat{\theta}_1$ έχει προφανώς μικρότερη *διακύμανση* από την *εκτιμήτρια* $\hat{\theta}_2$. Όπως φαίνεται από το σχήμα, για $\varepsilon > 0$,

$$P(\theta - \varepsilon \le \hat{\theta}_1 \le \theta + \varepsilon) > P(\theta - \varepsilon \le \hat{\theta}_2 \le \theta + \varepsilon)$$

δηλαδή, η $\hat{\theta}_1$ έχει μεγαλύτερη πιθανότητα από την $\hat{\theta}_2$ να πάρει τιμή που να βρίσκεται σε απόσταση το πολύ ε από την αληθή τιμή της παραμέτρου θ. Επομένως, η εκτιμήτρια $\hat{\theta}_1$ συμπεριφέρεται καλύτερα από την $\hat{\theta}_2$ γιατί έχει μικρότερη *διακύμανση*. Έτσι, μια *αμερόληπτη εκτιμήτρια* είναι επιθυμητό *(δείτε και το Σχήμα 11.1.4)* να έχει μικρή *διακύμανση*.

Μια *αμερόληπτη εκτιμήτρια*, $\hat{\theta}_1$ μιας *παραμέτρου* θ, ονομάζεται ***πιο αποτελεσματική (more efficient)*** από μια άλλη *αμερόληπτη εκτιμήτρια,* $\hat{\theta}_2$ της ίδιας *παραμέτρου* θ, αν

$$Var(\hat{\theta}_1) < Var(\hat{\theta}_2).$$



Σχήμα 11.1.3

Η εκτιμήτρια $\hat{\theta}_1$ *είναι πιο αποτελεσματική από την* $\hat{\theta}_2$

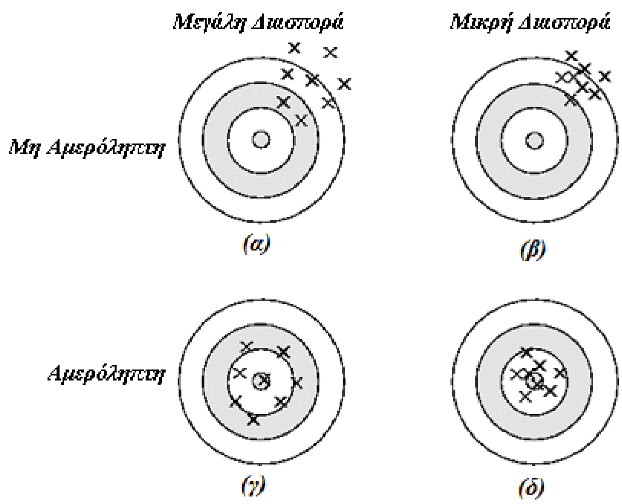
Εύκολα μπορεί να αποδειχθεί ότι από τις *εκτιμήτριες*

$$T = \sum_{i=1}^{\nu} a_i X_i \text{ , με } \sum_{i=1}^{\nu} \alpha_i = 1$$

της *μέσης τιμής* μ ενός *πληθυσμού,* τη μικρότερη *διακύμανση* έχει αυτή για την οποία $\alpha_i = 1/\nu$, $i = 1,2,...,\nu$, δηλαδή, ο *δειγματικός μέσος*

$$\overline{X} = \frac{1}{\nu}\sum_{i=1}^{\nu} X_i \text{ .}$$

Προφανώς, είναι ενδιαφέρον και φυσικά επιθυμητό, από όλες τις *αμερόληπτες εκτιμήτριες* μιας *παραμέτρου* να προσδιορίσουμε εκείνη με την *ελάχιστη διακύμανση*, δηλαδή, να προσδιορίσουμε την ***αμερόληπτη εκτιμήτρια ελάχιστης διασποράς*** της *παραμέτρου*. Όμως, για το πώς μπορούμε να προσδιορίζουμε μια τέτοια *εκτιμήτρια* δε θα επεκταθούμε περισσότερο.



Σχήμα 11.1.4

Επιθυμητή είναι η περίπτωση (δ) αφού τα σημεία πρόσπτωσης του βέλους βρίσκονται κοντά στο κέντρο του στόχου όπως βέβαια και στην περίπτωση (γ), όμως επιπλέον είναι περισσότερο συγκεντρωμένα γύρω από το κέντρο. Στις περιπτώσεις (α) και (β) βρίσκονται συστηματικά σε συγκεκριμένη περιοχή μακριά από το κέντρο του στόχου. Στην περίπτωση (α) είναι και περισσότερο διασκορπισμένα.

Ολοκληρώνουμε αυτή τη σύντομη αναφορά σε κάποιες από τις επιθυμητές ιδιότητες των *εκτιμητριών,* με την ιδιότητα της ***συνέπειας (consistency)*** η οποία κατατάσσεται στις ιδιότητες που αναφέρονται στην *ασυμπτωτική* συμπεριφορά των *εκτιμητριών*, δηλαδή, στο πώς συμπεριφέρεται μια *εκτιμήτρια* όταν το μέγεθος του δείγματος αυξάνεται (θεωρητικά όταν $\nu \to +\infty$). Δε θα αναφερθούμε εκτενώς, ούτε θα δώσουμε τον αυστηρό ορισμό της *συνέπειας*. Θα κάνουμε μόνο μια σύντομη αναφορά στο νόημά της και θα δώσουμε ένα παράδειγμα.

Όταν, προηγουμένως, αναφερθήκαμε στην ιδιότητα της *αποτελεσματικότητας,* προσπαθήσαμε να εξηγήσουμε γιατί είναι επιθυμητό να έχει μια *αμερόληπτη εκτιμήτρια* μικρή *διακύμανση*. Μπορεί να αποδειχθεί ότι αν η *διακύμανση* μιας *αμερόληπτης εκτιμήτριας* (ακριβέστερα, μια *ακολουθίας αμερόληπτων εκτιμητριών*) συγκλίνει στο μηδέν όταν το μέγεθος του δείγματος αυξάνεται ($\nu \to +\infty$), τότε αυξάνεται η πιθανότητα η *εκτίμηση* να βρίσκεται «κοντά» στην τιμή της *παραμέτρου*. Αυτή η ιδιότητα ονομάζεται ***συνέπεια***.

Μια *εκτιμήτρια* ονομάζεται ***συνεπής (consistent)*** αν, καθώς αυξάνεται το μέγεθος του δείγματος, η *εκτίμηση* γίνεται καλύτερη/ακριβέστερη, δηλαδή, οι τιμές της *εκτιμήτριας* πλησιάζουν (συγκλίνουν στοχαστικά) στην τιμή της *παραμέτρου*.

Εύκολα μπορεί να αποδειχθεί ότι ο *δειγματικός μέσος* είναι *συνεπής εκτιμήτρια* της *πληθυσμιακής μέσης τιμής* (αρκεί να θυμηθούμε ότι είναι *αμερόληπτη εκτιμήτρια* της *πληθυσμιακής μέσης τιμής* και ότι η *διακύμανση* της, σ^2/ν, τείνει στο μηδέν όταν $\nu \to +\infty$ και εφόσον, φυσικά, η *πληθυσμιακή διακύμανση,* σ^2, είναι πεπερασμένη). Δηλαδή, όσο αυξάνεται το μέγεθος του δείγματος, αυξάνεται η πιθανότητα οι τιμές του *δειγματικού μέσου,* $\overline{X}$, να βρίσκονται κοντά, να συγκλίνουν, στην *πληθυσμιακή μέση τιμή,* μ. Σημειώνουμε, τέλος, ότι και η *αμερόληπτη δειγματική διακύμανση,* S^2, είναι *συνεπής εκτιμήτρια* της *πληθυσμιακής διακύμανσης* σ^2.

Ας δούμε τώρα πώς μπορούμε να απαντήσουμε στο δεύτερο από τα δύο ερωτήματα που θέσαμε προηγουμένως *(στην ενότητα 11.1.1),* δηλαδή για το τι μπορούμε να πούμε για το ***σφάλμα*** που κάνουμε όταν εκτιμάμε μια άγνωστη *παράμετρο* ενός *πληθυσμού* με βάση την πληροφορία που παίρνουμε από ένα *τυχαίο δείγμα* τιμών του.

11.1.3 Εκτίμηση με διάστημα

Μια *σημειακή εκτίμηση* μιας άγνωστης *παραμέτρου* ενός *πληθυσμού*, όπως είδαμε, είναι ένας συγκεκριμένος αριθμός. Με τις επιθυμητές ιδιότητες των *εκτιμητριών,* ουσιαστικά ζητάμε οι τιμές της *εκτιμήτριας* (σε επαναλαμβανόμενες δειγματοληψίες) να βρίσκονται με μεγάλη πιθανότητα κοντά στην πραγματική/αληθή τιμή της *παραμέτρου* Ζητάμε δηλαδή, ***το σφάλμα της εκτίμησης* να είναι μικρό, με μεγάλη πιθανότητα**. Μπορούμε όμως να κάνουμε κάτι ακόμη, ένα επιπλέον βήμα. Με βάση πάντα την πληροφορία που παίρνουμε από ένα *τυχαίο δείγμα*, αντί να δίνουμε ένα μόνο αριθμό (μια *σημειακή εκτίμηση*) ως εκτίμηση της άγνωστης *παραμέτρου* του *πληθυσμού*, να δίνουμε ένα *διάστημα εμπιστοσύνης,* δηλαδή, ένα διάστημα το οποίο να περιέχει, με δεδομένη εμπιστοσύνη (πιθανότητα), την εκτιμώμενη *παράμερο.* Μια τέτοια *εκτίμηση* λέγεται ***εκτίμηση με διάστημα (interval estimation)***.

Ένα ***100(1 – α)% διάστημα εμπιστοσύνης (confidence interval)*** ($0 < \alpha < 1$), για μια *παράμετρο* ενός *πληθυσμού*, είναι ένα διάστημα που υπολογίζεται από ένα *τυχαίο δείγμα* από τον *πληθυσμό* και έχει πιθανότητα $1-\alpha$ να περιέχει την πραγματική τιμή της *παραμέτρου*. Η πιθανότητα $1-\alpha$ ονομάζεται ***συντελεστής εμπιστοσύνης (confidence coefficient)*** του διαστήματος.

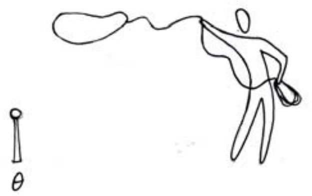
Έτσι, κατασκευάζοντας για μια άγνωστη *παράμετρο* ενός *πληθυσμού,* ένα $100(1-\alpha)\%$ *διάστημα εμπιστοσύνης,* δίνουμε ως εκτίμησή της, όχι έναν αριθμό (μια *σημειακή εκτίμηση*) αλλά δύο αριθμούς που ορίζουν ένα διάστημα το οποίο περιέχει την άγνωστη *παράμετρο* με ορισμένη πιθανότητα, ίση με $1-\alpha$. Η πιθανότητα αυτή ασφαλώς ορίζεται να είναι μεγάλη (συνήθως, ίση με 0.90 ή 0.95 ή 0.98 ή 0.99).

Διευκρινίζουμε και επισημαίνουμε, ότι λέγοντας *«το διάστημα εμπιστοσύνης υπολογίζεται από ένα τυχαίο δείγμα»,* εννοούμε ότι τα άκρα του είναι *στατιστικές συναρτήσεις* (συναρτήσεις του δείγματος που δεν περιέχουν άγνωστες παραμέτρους), δηλαδή, μπορεί να υπολογισθεί **αποκλειστικά** από το συγκεκριμένο δείγμα που, κάθε φορά, έχουμε στη διάθεσή μας. Είναι προφανές, ότι από πραγματοποίηση σε πραγματοποίηση του δείγματος, το ***διάστημα εμπιστοσύνης*** εν γένει μεταβάλλεται, είναι δηλαδή ***τυχαία μεταβλητή***.

Πριν εξηγήσουμε πώς κατασκευάζεται, ας δούμε πώς πρέπει να ερμηνεύεται ένα $100(1-\alpha)\%$ *διάστημα εμπιστοσύνης.*

Σύμφωνα με τον ορισμό που δώσαμε προηγουμένως, ένα $100(1-\alpha)\%$ *διάστημα εμπιστοσύνης* μιας *παραμέτρου* ενός πληθυσμού, περιέχει την τιμή της *παραμέτρου* με πιθανότητα $1-\alpha$. Αυτό, σύμφωνα με την ερμηνεία της πιθανότητας ως *οριακή σχετική συχνότητα* (ορισμός της πιθανότητας κατά *von Mises*), σημαίνει/έχει την έννοια ότι: σε μεγάλο αριθμό επαναλήψεων του πειράματος *«παίρνω ένα τυχαίο δείγμα μεγέθους ν από τον πληθυσμό και κατασκευάζω για την άγνωστη παράμετρο ένα* $100(1-\alpha)\%$ *διάστημα εμπιστοσύνης*», ποσοστό $1-\alpha$ των δειγμάτων, θα δώσουν διάστημα που θα περιέχει την τιμή της *παραμέτρου* και ποσοστό α των δειγμάτων θα δώσουν διάστημα που δεν θα περιέχει την τιμή της παραμέτρου. Βέβαια, όταν πάρουμε ένα συγκεκριμένο *τυχαίο δείγμα*, **ποτέ** δεν είμαστε βέβαιοι/δεν ξέρουμε αν το *διάστημα εμπιστοσύνης* που θα κατασκευάσουμε από αυτό, περιέχει/εγκλωβίζει ή όχι την τιμή της *παραμέτρου* (εκτός αν γνωρίζαμε την τιμή της παραμέτρου). Αυτό, φυσικά, δε μειώνει την αξία ενός *διαστήματος εμπιστοσύνης*. Η *εμπιστοσύνη* μας στην *εκτίμηση* της *παραμέτρου* με αυτό, είναι ίση με $1-\alpha$, είναι σαφώς ορισμένη και έχει την έννοια που εξηγήσαμε προηγουμένως.

Ας δούμε και μια πιο παραστατική ερμηνεία της *εκτίμησης με διάστημα εμπιστοσύνης* που πιστεύουμε ότι θα βοηθήσει στην κατανόηση. Σκεφθείτε ότι προσπαθείτε, πετώντας ένα λάσο, να «εγκλωβίσετε» έναν πακτωμένο/σταθερό πάσσαλο (να περάσετε τη θηλιά του λάσου στον πάσσαλο).



Ο πάσσαλος αναπαριστά την *παράμετρο* που θέλετε να εκτιμήσετε και η θηλιά του λάσου το *διάστημα εμπιστοσύνης*. Κάθε φορά που πετάτε το λάσο προς τον πάσσαλο, ελπίζετε ότι θα καταφέρετε να τον «εγκλωβίσετε». Όμως αυτό, κάποιες φορές συμβαίνει και κάποιες όχι. Κατ' αναλογία, κάθε φορά που παίρνετε ένα *τυχαίο δείγμα* και κατασκευάζετε από αυτό ένα *διάστημα εμπιστοσύνης* για να εκτιμήσετε μια *παράμετρο*, ελπίζετε ότι θα «εγκλωβίσετε» την *παράμετρο* στο διάστημα που κατασκευάσατε, αλλά όπως συμβαίνει με τη θηλιά και τον πάσσαλο, κάποιες φορές επιτυγχάνετε και κάποιες αποτυγχάνετε. Το ποσοστό των «επιτυχιών» σας, σε επαναλαμβανόμενες δειγματοληψίες, είναι ο *συντελεστής εμπιστοσύνης*.

Παρατήρηση 11.1.2: *Κάποιος θα μπορούσε να δώσει την εξής ερμηνεία για ένα* $100(1-\alpha)\%$ *διάστημα εμπιστοσύνης: η παράμετρος έχει* $1-\alpha$ *πιθανότητα να βρίσκεται στο συγκεκριμένο διάστημα που κατασκευάσαμε. Προφανώς μια τέτοια ερμηνεία είναι άστοχη/λάθος γιατί η παράμετρος είναι, άγνωστη μεν, αλλά σταθερός αριθμός και επομένως ή ανήκει ή δεν ανήκει στο συγκεκριμένο διάστημα (πιθανότητες αποδίδουμε στις τιμές τυχαίων μεταβλητών).*

■

Ας δούμε τώρα, πώς μπορούμε να κατασκευάσουμε ένα $100(1-\alpha)\%$ *διάστημα εμπιστοσύνης* για τη *μέση τιμή* μ ενός *πληθυσμού*.

11.2 Διάστημα εμπιστοσύνης για τη μέση τιμή ενός πληθυσμού

Ως *εκτιμήτρια* της *μέσης τιμής* μ θα χρησιμοποιήσουμε τον *δειγματικό μέσο* $\overline{X}$ αφού όπως διαπιστώσαμε έχει καλές ιδιότητες και επιπλέον, όπως είδαμε στο Α΄ Μέρος και στο *10ο Κεφάλαιο*, υπάρχουν αποτελέσματα της *Θεωρίας Πιθανοτήτων* και της *Στατιστικής* που μας δίνουν επακριβώς ή με ικανοποιητική προσέγγιση την κατανομή του. Συγκεκριμένα, έχουμε στη διάθεσή μας αποτελέσματα σχετικά με την κατανομή του *δειγματικού μέσου* $\overline{X}$

α) όταν ο πληθυσμός είναι κανονικός

β) όταν το μέγεθος του δείγματος είναι μεγάλο (και ο πληθυσμός όχι κατ' ανάγκη κανονικός).

Για τις περιπτώσεις αυτές ας δούμε πώς μπορούμε να κατασκευάσουμε ένα $100(1-\alpha)\%$ *διάστημα εμπιστοσύνης* για την εκτίμηση της *μέσης τιμής* μ ενός *πληθυσμού.*

11.2.1 Ο πληθυσμός είναι κανονικός

Θα διακρίνουμε την περίπτωση που η *διακύμανση* του πληθυσμού είναι γνωστή από την περίπτωση που δεν είναι γνωστή.

(α) Ο πληθυσμός είναι κανονικός με γνωστή διακύμανση

Έστω *τυχαίο δείγμα* $X_1, X_2, \ldots, X_\nu$, μεγέθους ν (οτιδήποτε, μικρό ή μεγάλο), από έναν κανονικό πληθυσμό με άγνωστη *μέση τιμή* μ (την οποία θέλουμε να εκτιμήσουμε) και με γνωστή *διακύμανση* σ^2.

Επειδή $X_i \sim N(\mu, \sigma^2)$, $i = 1, 2, \ldots, \nu$, από τη θεωρία πιθανοτήτων γνωρίζουμε ότι η κατανομή του *δειγματικού μέσου* $\overline{X}$ είναι *κανονική* με *μέση τιμή* μ και *διακύμανση* σ^2/ν, δηλαδή

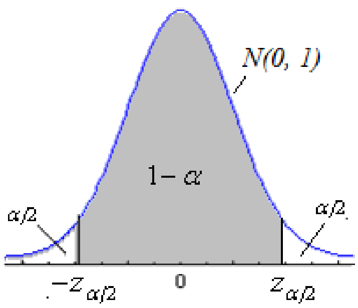
$$\overline{X} \sim N(\mu, \frac{\sigma^2}{\nu})$$

και επομένως

$$\frac{(\overline{X}-\mu)\sqrt{n}}{\sigma} = Z \sim N(0,1).$$

Επίσης γνωρίζουμε, ότι για την *τυποποιημένη κανονική κατανομή* μπορούμε να ορίσουμε ένα συμμετρικό γύρω από τη *μέση τιμή* της (από το 0) διάστημα, στο οποίο η Z παίρνει τιμές με δοθείσα πιθανότητα $1-\alpha$ *(θυμηθείτε πώς ορίζουμε και συμβολίζουμε ένα άνω α -ποσοστιαίο σημείο στην τυποποιημένη κανονική κατανομή).* Έτσι, όπως φαίνεται και στο *Σχήμα 11.2.1*, η Z παίρνει τιμές στο συμμετρικό διάστημα $[-z_{\alpha/2},\ z_{\alpha/2}]$ με πιθανότητα $1-\alpha$. Δηλαδή

$$P(-z_{\alpha/2} \le Z \le z_{\alpha/2}) = 1-\alpha .$$



Σχήμα 11.2.1

$P(-z_{\alpha/2} \le Z \le z_{\alpha/2}) = 1-\alpha$

Επομένως

$$P(-z_{\alpha/2} \le \frac{(\overline{X}-\mu)\sqrt{\nu}}{\sigma} \le z_{\alpha/2}) = 1-\alpha$$

και επειδή

$$-z_{\alpha/2} \le \frac{(\overline{X}-\mu)\sqrt{\nu}}{\sigma} \le z_{\alpha/2} \Leftrightarrow -z_{\alpha/2}\frac{\sigma}{\sqrt{\nu}} \le \overline{X}-\mu \le z_{\alpha/2}\frac{\sigma}{\sqrt{\nu}} \Leftrightarrow$$

$$-\overline{X}-z_{\alpha/2}\frac{\sigma}{\sqrt{\nu}} \le -\mu \le -\overline{X}+z_{\alpha/2}\frac{\sigma}{\sqrt{\nu}} \Leftrightarrow \overline{X}-z_{\alpha/2}\frac{\sigma}{\sqrt{\nu}} \le \mu \le \overline{X}+z_{\alpha/2}\frac{\sigma}{\sqrt{\nu}}$$

τα ενδεχόμενα

$$-z_{\alpha/2} \le \frac{(\overline{X}-\mu)\sqrt{\nu}}{\sigma} \le z_{\alpha/2} \text{ και } \overline{X}-z_{\alpha/2}\frac{\sigma}{\sqrt{\nu}} \le \mu \le \overline{X}+z_{\alpha/2}\frac{\sigma}{\sqrt{\nu}}$$

συμβαίνουν με την ίδια πιθανότητα, έτσι

$$P(\overline{X}-z_{\alpha/2}\frac{\sigma}{\sqrt{\nu}} \le \mu \le \overline{X}+z_{\alpha/2}\frac{\sigma}{\sqrt{\nu}}) = 1-\alpha .$$

Αυτό σημαίνει, ότι το *τυχαίο διάστημα*

$$\overline{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{\nu}}$$

περιέχει τη *μέση τιμή* μ του *πληθυσμού* με πιθανότητα $1-\alpha$ και επομένως, για συγκεκριμένη πραγματοποίηση του *τυχαίου δείγματος*, μας δίνει ένα $100(1-\alpha)\%$ *διάστημα εμπιστοσύνης* για τη *μέση τιμή* μ του *πληθυσμού.*

Έτσι, κάθε φορά που έχουμε στη διάθεσή μας μια συγκεκριμένη πραγματοποίηση $x_1, x_2, \ldots, x_\nu$ του *τυχαίου δείγματος* $X_1, X_2, \ldots, X_\nu$, υπολογίζουμε την τιμή, $\overline{x}$, της $\overline{X}$ και αφού καθορίσουμε έναν επιθυμητό *συντελεστή εμπιστοσύνης* $1-\alpha$, μπορούμε να κατασκευάσουμε ένα $100(1-\alpha)\%$ *διάστημα εμπιστοσύνης*, αφού η *διακύμανση,* σ^2, του *πληθυσμού* μάς είναι γνωστή και η τιμή $z_{\alpha/2}$ δίνεται από τον πίνακα της *τυποποιημένης κανονικής κατανομής* (με αντίστροφη ανάγνωση).

Κατασκευάζοντας για τη *μέση τιμή* μ του *πληθυσμού*, με βάση μια συγκεκριμένη πραγματοποίηση $x_1, x_2, \ldots, x_\nu$ του *τυχαίου δείγματος*, ένα συγκεκριμένο $100(1-\alpha)\%$ *διάστημα εμπιστοσύνης*

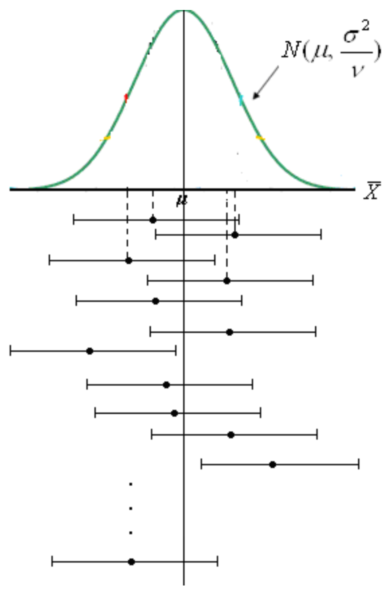
$$\overline{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{\nu}}$$

δεν είμαστε ποτέ βέβαιοι, όπως ήδη έχουμε εξηγήσει, αν το διάστημα αυτό περιέχει ή όχι την άγνωστη *μέση τιμή* μ. Γνωρίζουμε όμως, ότι η πιθανότητα να συμβεί αυτό είναι $1-\alpha$ και έχει την έννοια, όπως επίσης εξηγήσαμε, ότι αν επαναλάβουμε την προηγούμενη διαδικασία πολλές φορές, ποσοστό $1-\alpha$ των δειγμάτων θα δώσουν διάστημα που θα περιέχει τη *μέση τιμή* μ του *πληθυσμού* και ποσοστό α των δειγμάτων θα δώσουν διάστημα που δε θα την περιέχει.

Στο *Σχήμα 11.2.2*, δείχνουμε πιο παραστατικά αυτό ακριβώς. Θεωρείστε ότι επαναλάβαμε την προηγούμενη διαδικασία 100 φορές και κατασκευάσαμε (για συγκεκριμένο/γνωστό σ και συγκεκριμένο μέγεθος δείγματος ν) 100 *διαστήματα εμπιστοσύνης* με 0.95 *συντελεστή εμπιστοσύνης* το καθένα, δηλαδή 100 διαστήματα

$$\overline{x} \pm 1.96\frac{\sigma}{\sqrt{\nu}} .$$

Τα πέντε από αυτά ($0.05 \cdot 100 = 5$), δηλαδή ποσοστό 0.05, περιμένουμε να μην περιέχουν την *πληθυσμιακή μέση τιμή* μ. Παρατηρείστε ότι τα κέντρα των διαστημάτων που δεν περιέχουν την μ *(στο σχήμα φαίνονται 2 από τα 5 που περιμένουμε)*, είναι αρκετά απομακρυσμένα από το κέντρο της *δειγματοληπτικής κατανομής* (της κατανομής της $\overline{X}$).



Σχήμα 11.2.2
95% διαστήματα εμπιστοσύνης για τη μέση τιμή μ ενός πληθυσμού.
Ποσοστό 0.95 από αυτά αναμένουμε να «εγκλωβίζουν» τη μέση τιμή μ
και ποσοστό 0.05 να αποτυγχάνουν να την «εγκλωβίσουν»

Παρατήρηση 11.2.1: *Όπως φαίνεται και στο Σχήμα 11.2.2, το πλάτος των* $100(1-\alpha)\%$ *διαστημάτων εμπιστοσύνης,* $\bar{x} \pm z_{\alpha/2}\,\sigma/\sqrt{\nu}$*, από δείγμα σε δείγμα του ιδίου μεγέθους δεν αλλάζει, παραμένει σταθερό και είναι ίσο με* $2z_{\alpha/2}\,\sigma/\sqrt{\nu}$*. Αυτό συμβαίνει γιατί στον τύπο αυτό δεν εμφανίζεται κάποια ποσότητα που μεταβάλλεται από δείγμα σε δείγμα του ιδίου μεγέθους (η τυπική απόκλιση του πληθυσμού* σ*, είναι ένας μοναδικός/σταθερός αριθμός).*

■

Περιθώριο σφάλματος

Όταν εκτιμάμε τη *μέση τιμή* μ ενός *πληθυσμού* με τον *δειγματικό μέσο* $\overline{X}$, το ***σφάλμα της εκτίμησης (error of estimation)*** που κάνουμε, για συγκεκριμένο κάθε φορά δείγμα, είναι $\mu - \bar{x}$ και προφανώς, από δείγμα σε δείγμα αυτό αλλάζει. Είναι δηλαδή, *τυχαία μεταβλητή* ($\mu - \overline{X}$). Δίνοντας ως *εκτίμηση* της *πληθυσμιακής μέσης τιμής* μ, ένα $100(1-\alpha)\%$ *διάστημα εμπιστοσύνης*, $\bar{x} \pm z_{\alpha/2}\,\sigma/\sqrt{\nu}$, ουσιαστικά δίνουμε με προκαθορισμένη πιθανότητα ίση με $1-\alpha$, ένα *περιθώριο σφάλματος* της *εκτίμησης* ίσο με $z_{\alpha/2}\,\sigma/\sqrt{\nu}$. Αν και είναι προφανές τι εννοούμε με την όρο *περιθώριο σφάλματος*, καθώς και γιατί στην περίπτωση που εξετάζουμε αυτό είναι $z_{\alpha/2}\,\sigma/\sqrt{\nu}$ (με πιθανότητα $1-\alpha$), ας δούμε πώς ορίζεται και πώς τα προηγούμενα προκύπτουν πιο «φορμαλιστικά».

Ως ***περιθώριο σφάλματος (margin of error)*** της *εκτίμησης* της *πληθυσμιακής μέσης τιμής* μ από τον *δειγματικό μέσο* $\overline{X}$, ορίζουμε το μέγιστο σφάλμα, με πιθανότητα $1-\alpha$, που μπορεί να κάνουμε. Δηλαδή, ένα θετικό αριθμό *d*, τέτοιον ώστε

$$P(|\mu - \overline{X}| \le d) = 1-\alpha .$$

Επομένως, ισοδύναμα, έχουμε

$$P(-d \le \overline{X} - \mu \le d) = 1-\alpha \text{ ή } P\left(\frac{-d}{\sigma/\sqrt{\nu}} \le Z \le \frac{d}{\sigma/\sqrt{\nu}}\right) = 1-\alpha .$$

Άρα

$$\frac{d}{\sigma/\sqrt{\nu}} = z_{\alpha/2} \Leftrightarrow d = z_{\alpha/2}\frac{\sigma}{\sqrt{\nu}}.$$

Έτσι, εκτιμώντας την άγνωστη *μέση τιμή* μ με ένα, για παράδειγμα, 95% *διάστημα εμπιστοσύνης*, το *περιθώριο σφάλματος*, με πιθανότητα 0.95, είναι $1.96\sigma/\sqrt{\nu}$. Δηλαδή, ο *δειγματικός μέσος,* με πιθανότητα 0.95, βρίσκεται σε απόσταση το πολύ ίση με $1.96\sigma/\sqrt{\nu}$ (αριστερά ή δεξιά) από την *πληθυσμιακή μέση τιμή*.

Σημείωση 11.2.1: *Όπως διαπιστώσαμε, η τυπική απόκλιση* $\sigma_{\overline{X}} = \sigma/\sqrt{\nu}$ *του δειγματικού μέσου* $\overline{X}$*, συνδέεται άμεσα με το σφάλμα της εκτίμησης της πληθυσμιακής μέσης τιμής από τον δειγματικό μέσο. Έτσι εξηγείται γιατί η τυπική απόκλιση* $\sigma_{\overline{X}}$*, ονομάσθηκε* ***τυπικό σφάλμα (standard error)****.*

Εμπιστοσύνη vs ακρίβεια και επιλογή μεγέθους δείγματος

Είναι φανερό ότι όσο μεγαλύτερο *συντελεστή εμπιστοσύνης* ορίζουμε, δηλαδή, όσο μεγαλύτερη θέλουμε να είναι η πιθανότητα το διάστημα που θα κατασκευάσουμε να περιέχει την *πληθυσμιακή μέση τιμή*, τόσο το πλάτος/εύρος του διαστήματος, έστω ℓ, γίνεται μεγαλύτερο. Επομένως, αυξάνοντας το *συντελεστή εμπιστοσύνης*, αυξάνεται το *περιθώριο σφάλματος,* ή αλλιώς, μειώνεται η *ακρίβεια* της εκτίμησης. Παρατηρείστε τον τύπο που δίνει το *διάστημα εμπιστοσύνης*. Το πλάτος/εύρος του διαστήματος είναι

$$\ell = 2z_{\alpha/2}\frac{\sigma}{\sqrt{\nu}} = 2d$$

δηλαδή, καθορίζεται από το *περιθώριο σφάλματος*. Αν αυξήσουμε το *συντελεστή εμπιστοσύνης*, $1-\alpha$, η τιμή $z_{\alpha/2}$ θα μεγαλώσει γιατί είναι αύξουσα συνάρτηση του $1-\alpha$ (ή αλλιώς, φθίνουσα συνάρτηση του α) και επειδή το σ είναι σταθερό/μοναδικό για τον *πληθυσμό* και το ν είναι καθορισμένο, το *περιθώριο σφάλματος* και επομένως το πλάτος του διαστήματος, αναπόφευκτα, θα μεγαλώσει και θα χάσουμε σε ακρίβεια.

Εφόσον όταν απαιτούμε μεγαλύτερη *εμπιστοσύνη* αυξάνεται το *περιθώριο σφάλματος* (χάνουμε σε *ακρίβεια*), για να ελέγξουμε και τα δύο αυτά μεγέθη, μπορούμε να κάνουμε το εξής. Στη φάση σχεδιασμού της έρευνας, δηλαδή, πριν πάρουμε το δείγμα, καθορίζουμε, τόσο το επιθυμητό *επίπεδο εμπιστοσύνης* $1-\alpha$*,* όσο και το ανεκτό *περιθώριο σφάλματος d*. Έτσι, αξιοποιώντας τον τύπο

$$d = z_{\alpha/2}\frac{\sigma}{\sqrt{\nu}}$$

λύνουμε ως προς το μέγεθος του δείγματος ν,

$$\nu = \left(\frac{\sigma z_{\alpha/2}}{d}\right)^2 \text{ ή } \nu = \left(\frac{2\sigma z_{\alpha/2}}{\ell}\right)^2$$

και με αυτόν τον τρόπο βρίσκουμε το απαιτούμενο *μέγεθος δείγματος*, για να πετύχουμε, στο επιθυμητό *επίπεδο εμπιστοσύνης*, το ανεκτό *περιθώριο σφάλματος*.

Πριν συνεχίσουμε με άλλες περιπτώσεις υπολογισμού *διαστημάτων εμπιστοσύνης*, ας δούμε ένα παράδειγμα.

Παράδειγμα 11.2.1: *Μια αυτόματη μηχανή συσκευάζει καλαμπόκι σε τσουβάλια των 25kg. Το βάρος του καλαμποκιού που συσκευάζεται ανά τσουβάλι είναι τυχαία μεταβλητή, έστω X, η οποία, σύμφωνα με τον κατασκευαστή της μηχανής, ακολουθεί* *μια κανονική κατανομή με διακύμανση $\sigma^2 = 1.5^2 Kg^2$. Ο υπεύθυνος παραγωγής αμφιβάλλει για το μέσο βάρος του καλαμποκιού που συσκευάζεται ανά τσουβάλι και προκειμένου να το εκτιμήσει, επέλεξε τυχαία 15 τσουβάλια από την παραγωγή μιας ημέρας, τα ζύγισε και βρήκε ότι έχουν μέσο βάρος $24.8kg$. Με βάση την πληροφορία από το τυχαίο δείγμα που πήρε ο υπεύθυνος παραγωγής, τι μπορούμε να συμπεράνουμε για το μέσο βάρος του καλαμποκιού που συσκευάζεται ανά τσουβάλι από τη μηχανή;*

Απάντηση: Ως μια *εκτίμηση (σημειακή)* της *μέσης τιμής* μ, της *τυχαίας μεταβλητής Χ* μπορούμε προφανώς, να χρησιμοποιήσουμε το *δειγματικό μέσο* $\bar{x} = 24.8kg$ και να συμπεράνουμε ότι η τιμή της μ είναι περίπου ίση με 24.8*kg*. Βέβαια, ένα άλλο *τυχαίο δείγμα* του ιδίου μεγέθους δεν περιμένουμε να δώσει ακριβώς την ίδια *εκτίμηση*, μπορεί να δώσει 25*kg*, ή 25.1*kg* ή 24.5*kg* ή κάποια άλλη *εκτίμηση*. Γνωρίζουμε όμως ότι ο *δειγματικός μέσος* έχει καλές ιδιότητες, όπως ότι σε επαναλαμβανόμενες δειγματοληψίες κατά μέσο όρο εκτιμά σωστά τη μ και ότι όσο αυξάνεται το μέγεθος του δείγματος, παίρνει τιμές κοντά στην πραγματική τιμή της μ με πιο μεγάλη πιθανότητα.

Επιπλέον μπορούμε να κατασκευάσουμε ένα *διάστημα εμπιστοσύνης* με ικανοποιητική *εμπιστοσύνη*, για παράδειγμα 0.95, δηλαδή ένα διάστημα το οποίο με πιθανότητα 0.95 να περιέχει την πραγματική τιμή της *μέσης τιμής*, μ. Επειδή το δείγμα μας προέρχεται από *κανονικό πληθυσμό* με γνωστή *διακύμανση* $\sigma^2 = 1.5^2 kg^2$, ένα *95% διάστημα εμπιστοσύνης* για τη *μέση τιμή*, μ, της *τυχαίας μεταβλητής Χ,* δηλαδή, για το *μέσο βάρος του καλαμποκιού που συσκευάζεται ανά τσουβάλι από τη μηχανή*, είναι

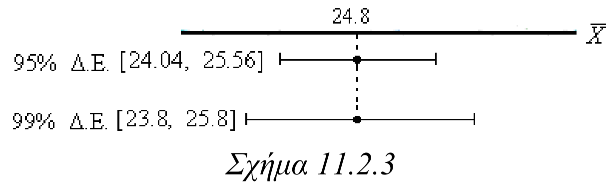
$$\bar{x} \pm z_{0.05/2}\frac{\sigma}{\sqrt{\nu}} \text{ ή } 24.8 \pm z_{0.025}\frac{1.5}{\sqrt{15}} \text{ ή } 24.8 \pm 1.96\frac{1.5}{\sqrt{15}} \text{ ή } 24.8 \pm 0.76 .$$

Μπορούμε επομένως να συμπεράνουμε ότι το διάστημα [24.04, 25.56] έχει 0.95 πιθανότητα να περιέχει το πραγματικό *μέσο βάρος του καλαμποκιού που συσκευάζεται ανά τσουβάλι από τη μηχανή.* Το *περιθώριο σφάλματος* της *εκτίμησης,* με πιθανότητα 0.95, είναι 0.76*kg*.

Αν θέλουμε να έχουμε μεγαλύτερη *εμπιστοσύνη* στην εκτίμησή μας, ας πούμε 0.99, κατασκευάζουμε από το ίδιο δείγμα ένα *99% διάστημα εμπιστοσύνης*. Έτσι, παίρνουμε

$$\bar{x} \pm z_{0.01/2}\frac{\sigma}{\sqrt{\nu}} \text{ ή } 24.8 \pm z_{0.005}\frac{1.5}{\sqrt{15}} \text{ ή } 24.8 \pm 2.58\frac{1.5}{\sqrt{15}} \text{ ή } 24.8 \pm 1 .$$

Συνεπώς, τώρα μπορούμε να συμπεράνουμε ότι το διάστημα [23.8, 25.8] έχει 0.99 πιθανότητα να περιέχει το πραγματικό *μέσο βάρος του καλαμποκιού που συσκευάζεται ανά τσουβάλι από τη μηχανή*. Όμως, το *περιθώριο σφάλματος* αυτής της *εκτίμησης* με πιθανότητα 0.99*,* είναι 1*kg,* δηλαδή, μεγάλωσε και επομένως χάσαμε σε *ακρίβεια εκτίμησης*. Παρατηρείστε ότι το πλάτος του *99% διαστήματος εμπιστοσύνης* είναι $2 \cdot 1kg = 2kg$ ενώ του *95% διαστήματος εμπιστοσύνης* είναι $2 \cdot 0.76kg = 1.52kg$ *(δείτε και το Σχήμα 11.2.3).*



Σχήμα 11.2.3
Το 99% διάστημα εμπιστοσύνης είναι πιο πλατύ από το αντίστοιχο 95% διάστημα εμπιστοσύνης

Αν θέλουμε η *εμπιστοσύνη* στην εκτίμησή μας για το *μέσο βάρος του καλαμποκιού που συσκευάζεται ανά τσουβάλι από τη μηχανή* να είναι 0.99 και το *σφάλμα της εκτίμησης* να είναι, για παράδειγμα, το πολύ $\pm 0.5kg$, δηλαδή, το *περιθώριο σφάλματος* με πιθανότητα 0.99 να είναι $d = 0.5kg$ (καθορισμένο εκ των προτέρων), τότε το απαιτούμενο μέγεθος του δείγματος που πρέπει να επιλέξουμε είναι

$$\nu = \left(\frac{\sigma\, z_{\alpha/2}}{d}\right)^2 = \left(\frac{1.5 z_{0.005}}{0.5}\right)^2 \cong 60 .$$

■

Στην περίπτωση που ακολουθεί όπως και στις επόμενες, δίνουμε μόνο τους τύπους υπολογισμού των *διαστημάτων εμπιστοσύνης*. Η μέθοδος που εφαρμόζουμε για να κατασκευάσουμε αυτά τα *διαστήματα εμπιστοσύνης* είναι ανάλογη με αυτήν που εφαρμόσαμε προηγουμένως. Φυσικά, αξιοποιούμε τα αντίστοιχα, κατά περίπτωση, αποτελέσματα των *Πιθανοτήτων* και της *Στατιστικής* για τις *δειγματοληπτικές κατανομές* που αναφέραμε στο *10^o Κεφάλαιο* (για δείγματα από *κανονική κατανομή* ή για *μεγάλα δείγματα*). Σημειώνουμε επίσης, ότι όσα προηγουμένως αναφέραμε για το πώς επηρεάζεται η *ακρίβεια* της *εκτίμησης* από το *επίπεδο εμπιστοσύνης,* ισχύουν και στις περιπτώσεις που ακολουθούν. Μόνο οι σχετικοί τύποι αλλάζουν.

(β) Ο πληθυσμός είναι κανονικός με άγνωστη διακύμανση

Στην περίπτωση (α) προηγουμένως, υπολογίσαμε τον τύπο που δίνει ένα $100(1-\alpha)\%$ *διάστημα εμπιστοσύνης* για τη *μέση τιμή* μ ενός *κανονικού πληθυσμού* του οποίου γνωρίζουμε τη *διακύμανση* σ^2. Όμως, σε πρακτικά προβλήματα, η *διακύμανση* του πληθυσμού, συνήθως δεν είναι γνωστή.

Έστω λοιπόν, τυχαίο δείγμα $X_1, X_2, \ldots, X_\nu$ μεγέθους ν (οτιδήποτε, μικρό ή μεγάλο) από μια *κανονική κατανομή* με άγνωστη *μέση τιμή* μ (την οποία θέλουμε να εκτιμήσουμε) και με άγνωστη *διακύμανση* σ^2. Στην περίπτωση αυτή, εκτιμάμε την άγνωστη *πληθυσμιακή διακύμανση* σ^2 από την *αμερόληπτη δειγματική διακύμανση*

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

και επειδή (όπως είδαμε στο *10^o Κεφάλαιο*) η *τυχαία μεταβλητή*

$$T = \frac{(\overline{X}-\mu)\sqrt{\nu}}{S}$$

ακολουθεί *t-κατανομή* με $\nu - 1$ *βαθμούς ελευθερίας,* δηλαδή

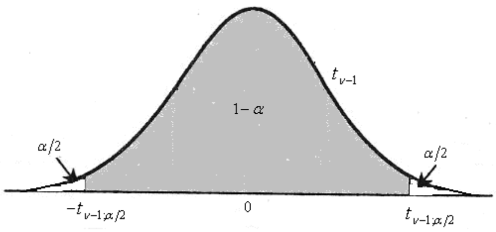
$$T = \frac{(\overline{X}-\mu)\sqrt{\nu}}{S} \sim t_{\nu-1}$$

εργαζόμενοι όπως στην περίπτωση (α) προηγουμένως, εύκολα μπορούμε να αποδείξουμε ότι το τυχαίο διάστημα

$$\overline{X} \pm t_{\nu-1;\alpha/2}\frac{S}{\sqrt{\nu}},$$

περιέχει τη *μέση τιμή* μ του *πληθυσμού* με πιθανότητα $1-\alpha$ και επομένως, για συγκεκριμένη πραγματοποίηση του *τυχαίου δείγματος*, μας δίνει ένα $100(1-\alpha)\%$ *διάστημα εμπιστοσύνης* για τη *μέση τιμή* μ του *πληθυσμού.*

Υπενθυμίζουμε, ότι με $t_{\nu-1;\alpha/2}$ συμβολίζεται το *άνω* $\alpha/2$*-ποσοστιαίο σημείο* της *t-κατανομής* με $\nu - 1$ *βαθμούς ελευθερίας*, το οποίο για διάφορες τιμές του α και του ν, δίνεται από σχετικό πίνακα *(δείτε και το Σχήμα 11.2.4).*



Σχήμα 11.2.4
$P(-t_{\nu-1;\alpha/2} \le T \le t_{\nu-1;\alpha/2}) = 1-\alpha$

Παρατηρείστε ότι το τυχαίο διάστημα

$$\overline{X} \pm t_{\nu-1;\alpha/2} \frac{S}{\sqrt{\nu}}$$

από δείγμα σε δείγμα (του ιδίου μεγέθους) μεταβάλλεται, εν γένει, όχι μόνο ως προς την τιμή του *δειγματικού μέσου* $\overline{X}$, αλλά και ως προς την τιμή της *δειγματικής τυπικής απόκλισης S*. Δηλαδή, το πλάτος/εύρος

$$\ell = 2t_{\nu-1;\alpha/2} \frac{S}{\sqrt{\nu}},$$

των $100(1-\alpha)\%$ *διαστημάτων εμπιστοσύνης* που κατασκευάζουμε είναι *τυχαία μεταβλητή* αφού από δείγμα σε δείγμα μεταβάλλεται (ενώ στην περίπτωση (α) προηγουμένως όπως είδαμε είναι σταθερό). Έτσι, και ο προσδιορισμός κατάλληλου μεγέθους δείγματος για να πετύχουμε προκαθορισμένο *περιθώριο σφάλματος* (στον επιθυμητό *συντελεστή εμπιστοσύνης*), δε μπορεί να γίνει όπως στην περίπτωση (α). Λύση στο πρόβλημα αυτό έχει δώσει ο *Charles Stein (1945)*, με μια *διαδικασία δισταδιακής δειγματοληψίας (Stein's two-stage scheme)*, όμως δε θα επεκταθούμε περισσότερο.

Ας επανέλθουμε τώρα στα ερωτήματα που απασχολούν το φοιτητή (στο εισαγωγικό παράδειγμα του κεφαλαίου).

Παράδειγμα 11.2.2: *Το τυχαίο δείγμα μεγέθους $\nu = 16$ που πήρε ο φοιτητής από την περιοχή Α, προέρχεται από κανονικό πληθυσμό με άγνωστη μέση τιμή μ_1 και άγνωστη διακύμανση σ_1^2. Δηλαδή, για την τυχαία μεταβλητή X που εκφράζει την ημερήσια ποσότητα σεληνίου που προσλαμβάνουν από τη διατροφή τους οι ενήλικες στην περιοχή Α, έχουμε $X \sim N(\mu_1, \sigma_1^2)$. Επειδή ο δειγματικός μέσος και η δειγματική τυπική απόκλιση είναι αντίστοιχα, $\overline{x} = 75\mu gr$ και $s_1 = 8.4\mu gr$, ένα 95% διάστημα εμπιστοσύνης για την άγνωστη μέση ημερήσια ποσότητα σεληνίου που προσλαμβάνουν από τη διατροφή τους οι ενήλικες στην περιοχή Α, είναι*

$$75 \pm t_{15;0.025} \frac{8.4}{\sqrt{16}} \text{ ή } 75 \pm 2.131 \frac{8.4}{\sqrt{16}} \text{ ή } 75 \pm 4.475\,.$$

Επομένως, μπορούμε να συμπεράνουμε ότι το διάστημα $[70.52, 79.48]$ έχει 0.95 πιθανότητα να περιέχει την άγνωστη *μέση* τιμή μ_1, δηλαδή, την άγνωστη *μέση ημερήσια ποσότητα σεληνίου που προσλαμβάνουν από τη διατροφή τους οι ενήλικες στην περιοχή Α*. Το *εύρος* αυτού του διαστήματος είναι ίσο με $2 \cdot 4.475\mu gr$. Αν θέλουμε, για τον ίδιο *συντελεστή εμπιστοσύνης*, καλύτερη ακρίβεια (διάστημα μικρότερου εύρους) πρέπει να πάρουμε μεγαλύτερου μεγέθους δείγμα. Για να υπολογίσουμε το μέγεθος του δείγματος που απαιτείται για να πετύχουμε, σε δεδομένο *επίπεδο εμπιστοσύνης* ένα προκαθορισμένο ανεκτό *εύρος* για το *διάστημα εμπιστοσύνης,* μπορούμε να εφαρμόσουμε τη *διαδικασία δισταδιακής δειγματοληψίας Stein*.

Βέβαια, όπως εξηγήσαμε στα προηγούμενα, μπορούμε να βελτιώσουμε την *ακρίβεια* της εκτίμησης αν από το ίδιο δείγμα κατασκευάσουμε ένα *διάστημα εμπιστοσύνης* με μικρότερο *συντελεστή εμπιστοσύνης,* για παράδειγμα, 0.90. Πράγματι, το 90% *διάστημα εμπιστοσύνης* που παίρνουμε από το συγκεκριμένο δείγμα, είναι

$$75 \pm t_{15;0.05}\frac{8.4}{\sqrt{16}} \text{ ή } 75 \pm 1.753\frac{8.4}{\sqrt{16}} \text{ ή } 75 \pm 3.68$$

και το *εύρος* τώρα είναι μικρότερο (είναι ίσο με $2 \cdot 3.68 \mu gr$).

Εργαζόμενοι με τον ίδιο τρόπο μπορούμε να δώσουμε *εκτίμηση με διάστημα εμπιστοσύνης* και για την άγνωστη *μέση τιμή* μ_2, δηλαδή, για τη *μέση ημερήσια ποσότητα σεληνίου που προσλαμβάνουν από τη διατροφή τους οι ενήλικες στην περιοχή Β (δείτε το ως άσκηση).*

■

11.2.2 Το μέγεθος του δείγματος είναι μεγάλο

Θα διακρίνουμε και πάλι την περίπτωση που η *διακύμανση* του πληθυσμού είναι γνωστή από την περίπτωση που δεν είναι γνωστή.

(α) Το μέγεθος του δείγματος είναι μεγάλο και η διακύμανση του πληθυσμού είναι γνωστή

Αν το τυχαίο δείγμα $X_1, X_2, \ldots, X_\nu$ προέρχεται από οποιαδήποτε κατανομή (όχι κατ' ανάγκη κανονική), με *μέση τιμή* μ και *διακύμανση* σ^2, τότε όπως είδαμε στο Α΄ Μέρος (*Κ.Ο.Θ.*), για μεγάλο μέγεθος δείγματος ν (εν γένει, $\nu \geq 30$), **κατά προσέγγιση** έχουμε

$$\overline{X} \sim N(\mu, \frac{\sigma^2}{\nu}).$$

Επομένως, στην περίπτωση αυτή, για συγκεκριμένη πραγματοποίηση, $x_1, x_2, \ldots, x_\nu$, του *τυχαίου δείγματος* $X_1, X_2, \ldots, X_\nu$, ένα $100(1-\alpha)\%$ *διάστημα εμπιστοσύνης* για τη *μέση τιμή* μ του *πληθυσμού,* όπως στην περίπτωση *11.2.1α*, δίνεται από τον τύπο

$$\overline{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{\nu}}.$$

Βέβαια, επειδή η κατανομή της $\overline{X}$ δεν είναι *κανονική* αλλά **προσεγγίζεται** ικανοποιητικά από την *κανονική*, το διάστημα αυτό είναι *συντελεστή εμπιστοσύνης* $1-\alpha$ κατά προσέγγιση. Δηλαδή, έχει πιθανότητα περίπου $1-\alpha$, να περιέχει την *άγνωστη πληθυσμιακή μέση τιμή* μ.

(β) Το μέγεθος του δείγματος είναι μεγάλο και η διακύμανση του πληθυσμού είναι άγνωστη

Αν το δείγμα $X_1, X_2, \ldots, X_\nu$ προέρχεται από οποιαδήποτε κατανομή (όχι κατ' ανάγκη κανονική), με *μέση τιμή* μ και *διακύμανση* σ^2, τότε *(όπως είδαμε στο 10^o Κεφάλαιο)* για μεγάλο μέγεθος δείγματος ν (εν γένει, $\nu \geq 30$), **κατά προσέγγιση** έχουμε

$$\frac{(\overline{X}-\mu)\sqrt{\nu}}{S} \sim N(0,1).$$

Επομένως, στην περίπτωση αυτή, για συγκεκριμένη πραγματοποίηση $x_1, x_2, \ldots, x_\nu$ του *τυχαίου δείγματος* $X_1, X_2, \ldots, X_\nu$, ένα $100(1-\alpha)\%$ *διάστημα εμπιστοσύνης* για τη *μέση τιμή* μ του *πληθυσμού,* δίνεται από τον τύπο

$$\bar{x} \pm z_{\alpha/2} \frac{s}{\sqrt{\nu}}.$$

Βέβαια, όπως και στην προηγούμενη περίπτωση *11.2.2α*, το διάστημα αυτό είναι *συντελεστή εμπιστοσύνης* $1-\alpha$ κατά προσέγγιση. Δηλαδή, έχει πιθανότητα περίπου $1-\alpha$, να περιέχει την *άγνωστη πληθυσμιακή μέση τιμή* μ.

Για διευκόλυνσή μας συνοψίζουμε τις προηγούμενες περιπτώσεις στον *Πίνακα 11.2.1.*

Πληθυσμός	*Διακύμανση του πληθυσμού* (σ^2)	*Μέγεθος του δείγματος* (ν)	$100(1-\alpha)\%$ *Δ. Ε. για τη μέση τιμή,* μ*, του πληθυσμού*
Κανονικός	*Γνωστή*	*Οτιδήποτε*	$\bar{X} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{\nu}}$ *(1)*
Οποιοσδήποτε	*Γνωστή*	*Μεγάλο*	
Οποιοσδήποτε	*Άγνωστη*	*Μεγάλο*	$\bar{X} \pm z_{\alpha/2} \frac{S}{\sqrt{\nu}}$ *(2)*
Κανονικός	*Άγνωστη*	*Οτιδήποτε*	$\bar{X} \pm t_{\nu-1;\alpha/2} \frac{S}{\sqrt{\nu}}$ *(3)*
Όχι Κανονικός	*Γνωστή ή Άγνωστη*	*Μικρό*	?

Πίνακας 11.2.1

$100(1-\alpha)\%$ *διαστήματα εμπιστοσύνης για τη μέση τιμή* μ *ενός πληθυσμού από ένα τυχαίο δείγμα μεγέθους* ν

Σημείωση 11.2.2: *α) Είναι προφανές, όμως το επισημαίνουμε, ότι σε κάθε περίπτωση το δείγμα θεωρείται τυχαίο. β) Για μεγάλο δείγμα από κανονική κατανομή με άγνωστη διακύμανση, μπορεί να εφαρμοσθεί ή ο τύπος (2) ή ο τύπος (3) αφού στην περίπτωση αυτή δίνουν ίδιο διάστημα (για μεγάλα* ν*,* $t_{\nu-1;\alpha/2} \cong z_{\alpha/2}$*). Όμως, για μικρό δείγμα από κανονική κατανομή με άγνωστη διακύμανση, το διάστημα (3) αποτελεί μοναδική επιλογή μας. γ) Ο τύπος (3) εφαρμόζεται με την προϋπόθεση ότι ο πληθυσμός είναι κανονικός, εντούτοις, έχει διαπιστωθεί ότι στην πράξη δίνει καλά αποτελέσματα ακόμη και όταν ο πληθυσμός δεν είναι κανονικός. Δηλαδή, η πιθανότητα το διάστημα αυτό να περιέχει την πληθυσμιακή μέση τιμή δεν απέχει πολύ από* $1-\alpha$*, αρκεί βέβαια, ο πληθυσμός να μην απέχει δραματικά από το να είναι κανονικός (σοβαρή ασυμμετρία, περισσότερες από μια κορυφές, κτλ.) και το δείγμα να μην είναι πολύ μικρό.*

■

Παράδειγμα 11.2.3 (συνέχεια του Παραδείγματος 9.1.3): *Στον Πίνακα 11.2.2 φαίνεται για κάθε μια από 50 τυχαία επιλεγμένες γαλακτοπαραγωγές αγελάδες, ο χρόνος (σε μήνες) από την πρώτη εκδήλωση μιας συγκεκριμένης ασθένειας που προσβάλλει τις αγελάδες μέχρι την επανεμφάνισή της.*

2.1	1.7	0.8	0.8	4.1	8.7	1.4	2.9	1.9	2.7
4.4	2.2	5.5	7.0	1.8	0.2	1.0	0.9	4.0	0.7
2.0	6.5	0.7	4.3	0.2	5.6	2.4	1.4	1.3	1.2
0.5	3.9	7.4	3.3	8.8	0.3	2.0	5.7	0.8	2.6
9.9	1.6	2.8	1.0	0.6	1.3	0.8	5.9	0.9	0.4

Πίνακας 11.2.2

Οι χρόνοι επανεμφάνισης (σε μήνες) μιας ασθένειας σε 50 γαλακτοπαραγωγές αγελάδες

Στα Παραδείγματα 9.1.3&9.1.4 και στην Άσκηση 9.13 μελετήσαμε την κατανομή του συγκεκριμένου δείγματος. Ας δούμε τώρα τι μπορούμε να πούμε, με βάση το συγκεκριμένο δείγμα, για τον πληθυσμό από τον οποίο αυτό προέρχεται. Δηλαδή, τι *μπορούμε να πούμε για την κατανομή της τυχαίας μεταβλητής, έστω X, που εκφράζει το χρόνο επανεμφάνισης της ασθένειας.*

Η κατανομή της τυχαίας μεταβλητής X μας είναι άγνωστη. Δε γνωρίζουμε ούτε τη μορφή της (αν είναι, για παράδειγμα, κανονική) ούτε κάποια παράμετρό της όπως τη μέση τιμή της ή τη διακύμανσή της.

Έστω λοιπόν μ η άγνωστη μέση τιμή της X, δηλαδή, ο άγνωστος μέσος χρόνος επανεμφάνισης της ασθένειας. Με βάση το τυχαίο δείγμα τιμών της X που έχουμε στη διάθεσή μας, μπορούμε πλέον να κατασκευάσουμε ένα *95% διάστημα εμπιστοσύνης* για την άγνωστη μέση τιμή μ, δηλαδή, να πάρουμε μια εκτίμηση για το μέσο χρόνο επανεμφάνισης της ασθένειας.

Πράγματι, επειδή το δείγμα είναι μεγάλο *(*$\nu = 50 \geq 30$*)* και η διακύμανση σ^2 της X (του πληθυσμού) μας είναι άγνωστη, με βάση όσα προηγουμένως εξηγήσαμε *(περίπτωση 11.2.2β)*, ένα *95% διάστημα εμπιστοσύνης* για την άγνωστη μέση τιμή μ της X δίνεται από τον τύπο

$$\overline{X} \pm z_{\alpha/2} \frac{S}{\sqrt{\nu}}.$$

Εύκολα βρίσκουμε *(Άσκηση 9.13)* ότι ο δειγματικός μέσος και η δειγματική τυπική απόκλιση για το συγκεκριμένο δείγμα αντίστοιχα είναι $\overline{x} = 2.82\,\mu\acute{\eta}\nu\varepsilon\varsigma$ και $s = 2.51\,\mu\acute{\eta}\nu\varepsilon\varsigma$ και επομένως, με βάση το συγκεκριμένο δείγμα, το ζητούμενο *95% διάστημα εμπιστοσύνης*, είναι

$$2.82 \pm z_{0.05/2} \frac{2.51}{\sqrt{50}} \text{ ή } 2.82 \pm z_{0.025} \frac{2.51}{\sqrt{50}} \text{ ή } 2.82 \pm 1.96 \frac{2.51}{\sqrt{50}} \text{ ή } 2.82 \pm 0.7.$$

Αυτό σημαίνει ότι το διάστημα [2.12, 3.52] έχει 0.95 πιθανότητα να περιέχει τον άγνωστο μέσο χρόνο επανεμφάνισης της ασθένειας.

■

11.3 Διάστημα εμπιστοσύνης για το διωνυμικό ποσοστό

Έστω $X_1, X_2, \ldots, X_\nu$ *τυχαίο δείγμα* από *πληθυσμό* που ακολουθεί κατανομή *Bernoulli* με παράμετρο p (πιθανότητα επιτυχίας) και επομένως με *μέση τιμή* $\mu = p$ και *διακύμανση* $\sigma^2 = p(1-p)$. Στο *10^o Κεφάλαιο* εξηγήσαμε ότι για την *τυχαία μεταβλητή*

$$\frac{X_1 + X_2 + \ldots + X_\nu}{\nu}$$

η οποία εκφράζει το ποσοστό των επιτυχιών στο δείγμα $\hat{P}$, για μεγάλα ν, **κατά προσέγγιση** έχουμε

$$\hat{P} = \frac{X_1 + X_2 + \ldots + X_\nu}{\nu} \sim N\left(p, \frac{p(1-p)}{\nu}\right).$$

Η προσέγγιση αυτή είναι ικανοποιητική αν $\nu p \geq 5$ **και** $\nu(1-p) \geq 5$ ή $\nu p(1-p) \geq 10$ *(θυμηθείτε ότι το πόσο μεγάλο πρέπει να είναι το* ν *εξαρτάται και από το p).*

Επομένως, **κατά προσέγγιση** έχουμε

$$P(-z_{\alpha/2} \leq \frac{\hat{P} - p}{\sqrt{\frac{p(1-p)}{\nu}}} \leq z_{\alpha/2}) = 1 - \alpha$$

και χρησιμοποιώντας ως *εκτιμήτρια* της άγνωστης *διακύμανσης* $\sigma^2 = p(1-p)$, την $\hat{P}(1-\hat{P})$, εύκολα προκύπτει ότι το διάστημα

$$\hat{P} \pm z_{\alpha/2}\sqrt{\frac{\hat{P}(1-\hat{P})}{\nu}}$$

είναι ένα **κατά προσέγγιση** $100(1-\alpha)\%$ *διάστημα εμπιστοσύνης* για την παράμετρο *p (ποσοστό στον πληθυσμό).*

Έτσι, για συγκεκριμένη πραγματοποίηση $x_1, x_2, \ldots, x_\nu$ του *τυχαίου δείγματος* $X_1, X_2, \ldots, X_\nu$, με ποσοστό επιτυχιών στο δείγμα $\hat{p} = (x_1 + x_2 + \ldots + x_\nu)/\nu$, το διάστημα

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

είναι ένα κατά προσέγγιση $100(1-\alpha)\%$ *διάστημα εμπιστοσύνης* για το *ποσοστό p στον πληθυσμό*, εφόσον, $\nu\hat{p} \geq 5$ **και** $\nu(1-\hat{p}) \geq 5$.

Σημείωση 11.3.1: *Αν το μέγεθος του δείγματος δεν είναι μεγάλο, η κανονική προσέγγιση που χρησιμοποιήσαμε δε μπορεί να εφαρμοσθεί. Στην περίπτωση αυτή μπορεί να χρησιμοποιηθεί η ακριβής κατανομή της στατιστικής συνάρτησης* $X_1 + \ldots + X_\nu = \nu\hat{P} \sim B(\nu, p)$. *Όμως, στην πράξη, σε προβλήματα εκτίμησης ποσοστού, συνήθως δεν αντιμετωπίζουμε πρόβλημα μικρών δειγμάτων.*

■

Παράδειγμα 11.3.1: *Ένας ερευνητής, προκειμένου να εκτιμήσει το ποσοστό των ατόμων σε έναν πληθυσμό που έχουν ομάδα αίματος Α, επέλεξε με βάση ένα σχέδιο τυχαίας δειγματοληψίας 200 άτομα από τον πληθυσμό αυτό και βρήκε ότι ομάδα αίματος Α έχουν τα 89.*

Προφανώς, πρόκειται για δειγματοληψία από έναν πληθυσμό *Bernoulli* με παράμετρο *p (μέση τιμή)*, την οποία ο ερευνητής θέλει να εκτιμήσει με βάση ένα *τυχαίο δείγμα* $X_1, X_2, \ldots, X_{200}$, μεγέθους 200. Η τιμή του *δειγματικού μέσου (δειγματικού ποσοστού)*

$$\hat{P} = \frac{X_1 + X_2 + \ldots + X_{200}}{200}$$

για τη συγκεκριμένη πραγματοποίηση του δείγματος είναι

$$\hat{p} = \frac{89}{200} = 0.445$$

και επειδή $\nu\hat{p} = 200 \cdot 0.445 = 89 \geq 5$ και $\nu(1-\hat{p}) = 200 \cdot 0.555 = 111 \geq 5$, το διάστημα

$$0.445 \pm z_{0.025}\sqrt{\frac{0.445 \cdot 0.555}{200}} \text{ ή } 0.445 \pm 1.96 \cdot 0.035 \text{ ή } 0.445 \pm 0.0686$$

είναι ένα κατά προσέγγιση 95% *διάστημα εμπιστοσύνης* για την παράμετρο *p (ποσοστό στον πληθυσμό).*

Επομένως, το διάστημα [0.376, 0.514] έχει 0.95 πιθανότητα να περιέχει το ποσοστό των ατόμων στον πληθυσμό που έχει ομάδα αίματος Α.

■

Παράδειγμα 11.3.2 (συνέχεια του Παραδείγματος 6.1.1): *Στο Παράδειγμα 6.1.1 είδαμε ότι σε μια καλλιέργεια τουλίπας, από 52 βολβούς τουλίπας που τυχαία επελέγησαν, βλάστησαν μόνο οι 38 και γι' αυτό ο αγρότης που τους καλλιεργεί αμφιβάλλει για το αν το ποσοστό των βολβών που βλαστάνουν είναι πράγματι 90% όπως ισχυρίζεται ο γεωπόνος του φυτώριου από το οποίο τους αγόρασε.*

Η απάντηση που δώσαμε (στο *Παράδειγμα 6.1.1*) δικαιώνει όπως είδαμε τις αμφιβολίες του αγρότη γιατί όπως δείξαμε *«με την υπόθεση ότι αληθεύει ο ισχυρισμός του γεωπόνου, η πιθανότητα να βλαστήσουν μόνο 38 από τους 52 βολβούς είναι περίπου μηδενική».*

Ας δούμε τώρα πώς μπορούμε, με βάση το *δειγματικό ποσοστό* που έχουμε στη διάθεσή μας, να κατασκευάσουμε ένα *διάστημα εμπιστοσύνης* για το ποσοστό p των βολβών που βλαστάνουν (επί του συνόλου της παραγωγής του φυτώριου) και να πάρουμε έτσι για το p μια εκτίμηση με δεδομένη εμπιστοσύνη, έστω 95%.

Το *δειγματικό ποσοστό* είναι

$$\hat{p} = \frac{38}{52} = 0.73$$

και επειδή $\nu\hat{p} = 52 \cdot 0.73 = 38 \geq 5$ και $\nu(1-\hat{p}) = 52 \cdot 0.27 = 14 \geq 5$, το διάστημα

$$0.73 \pm z_{0.025}\sqrt{\frac{0.73 \cdot 0.27}{52}} \text{ ή } 0.73 \pm 1.96 \cdot 0.062 \text{ ή } 0.73 \pm 0.12$$

είναι ένα κατά προσέγγιση 95% *διάστημα εμπιστοσύνης* για το ποσοστό p των βολβών του φυτωρίου που βλαστάνουν *(ποσοστό στον πληθυσμό).*

Επομένως, το διάστημα $[0.61,\ 0.85]$ έχει 0.95 πιθανότητα να περιέχει το ποσοστό των βολβών (επί του συνόλου της παραγωγής) που βλαστάνουν. Παρατηρείστε ότι δεν περιέχει την τιμή 0.9, δηλαδή, οι αμφιβολίες του αγρότη και πάλι δικαιώνονται.

■

Παρατήρηση 11.3.1: *Το εύρος* $2z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/\nu}$ ενός $100(1-\alpha)\%$ *διαστήματος εμπιστοσύνης* για το *διωνυμικό ποσοστό p, προφανώς δεν είναι σταθερό από δείγμα σε δείγμα. Όμως, για* $\hat{p} = 1/2$*, το γινόμενο* $\hat{p}(1-\hat{p})$ *γίνεται μέγιστο και επομένως το περιθώριο σφάλματος της εκτίμησης με πιθανότητα* $1-\alpha$ *είναι*

$$d = z_{\alpha/2}\sqrt{\frac{1}{4}\frac{1}{\nu}} \text{ ή } d = z_{\alpha/2}\frac{1}{2\sqrt{\nu}}.$$

Αν επομένως προκαθορίσουμε ένα ανεκτό περιθώριο σφάλματος d, με πιθανότητα $1-\alpha$*, τότε για την εκτίμηση του ποσοστού στον πληθυσμό, απαιτείται μέγεθος δείγματος*

$$\nu = \frac{1}{4}\left(\frac{z_{\alpha/2}}{d}\right)^2.$$

Αν έχουμε λόγους να πιστεύουμε ότι το άγνωστο πληθυσμιακό ποσοστό είναι περίπου ίσο με κάποια τιμή, έστω p'*, τότε, το απαιτούμενο μέγεθος δείγματος είναι*

$$\nu = p'(1-p')\left(\frac{z_{\alpha/2}}{d}\right)^2.$$

Έτσι, αν στο παράδειγμά μας θέλουμε το σφάλμα της εκτίμησης, με πιθανότητα 0.95, να μην ξεπερνά, για παράδειγμα, το 0.03, τότε πρέπει να πάρουμε δείγμα μεγέθους

$$\nu = \frac{1}{4}\left(\frac{1.96}{0.03}\right)^2 = 1067.1 \cong 1068.$$

Αν εκτιμάμε/θεωρούμε ότι το ποσοστό στον πληθυσμό είναι περίπου 0.40, τότε το απαιτούμενο μέγεθος δείγματος είναι

$$\nu = p'(1-p')\left(\frac{z_{\alpha/2}}{d}\right)^2 = 0.40 \cdot 0.60 \cdot \left(\frac{1.96}{0.03}\right)^2 \cong 1025.$$

11.4 Διάστημα εμπιστοσύνης για τη διακύμανση ενός κανονικού πληθυσμού

Έστω $X_1, X_2, \ldots, X_\nu$ *τυχαίο δείγμα* από *κανονικό πληθυσμό*, δηλαδή, $X_i \sim N(\mu, \sigma^2)$ για κάθε $i = 1, 2, \ldots, \nu$. Στο *10^ο Κεφάλαιο* είδαμε ότι για οποιοδήποτε *μέγεθος δείγματος* ν

$$\frac{(\nu-1)S^2}{\sigma^2} \sim \chi^2_{\nu-1}$$

όπου

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

η *αμερόληπτη δειγματική διακύμανση*. Επομένως (δείτε και το *Σχήμα 11.4.1*),

$$P\left(\chi^2_{\nu-1;1-\alpha/2} \le \frac{(\nu-1)S^2}{\sigma^2} \le \chi^2_{\nu-1;\alpha/2}\right) = 1-\alpha$$

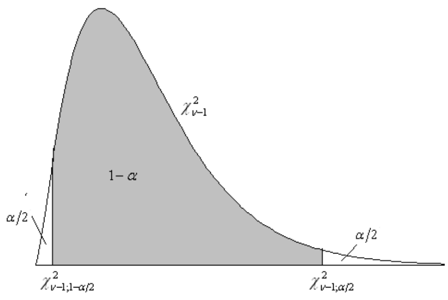
ή ισοδύναμα

$$P\left(\frac{\nu-1}{\chi^2_{\nu-1;\alpha/2}}S^2 \le \sigma^2 \le \frac{\nu-1}{\chi^2_{\nu-1;1-\alpha/2}}S^2\right) = 1-\alpha\,.$$

Αυτό σημαίνει ότι για συγκεκριμένη πραγματοποίηση $x_1, x_2, \ldots, x_\nu$, του *τυχαίου δείγματος* $X_1, X_2, \ldots, X_\nu$, το διάστημα

$$\left[\frac{\nu-1}{\chi^2_{\nu-1;\alpha/2}}s^2,\quad \frac{\nu-1}{\chi^2_{\nu-1;1-\alpha/2}}s^2\right]$$

είναι ένα $100(1-\alpha)\%$ *διάστημα εμπιστοσύνης* για τη *διακύμανση* σ^2 του *πληθυσμού.*



Σχήμα 11.4.1

$$P(\chi^2_{\nu-1;1-\alpha/2} \le (\nu-1)S^2/\sigma^2 \le \chi^2_{\nu-1;\alpha/2}) = 1-\alpha$$

Προφανώς, ένα $100(1-\alpha)\%$ *διάστημα εμπιστοσύνης* για την *τυπική απόκλιση* σ *του πληθυσμού* είναι

$$\left[s\sqrt{\frac{\nu-1}{\chi^2_{\nu-1;\alpha/2}}},\quad s\sqrt{\frac{\nu-1}{\chi^2_{\nu-1;1-\alpha/2}}}\right].$$

Υπενθυμίζουμε ότι με $\chi^2_{\nu-1;\alpha/2}$ και $\chi^2_{\nu-1;1-\alpha/2}$ συμβολίζουμε, αντίστοιχα, το *άνω* $\alpha/2$ και το *άνω* $1-\alpha/2$ *ποσοστιαίο σημείο* της *κατανομής* $\chi^2_{\nu-1}$ τα οποία δίνονται, για διάφορες τιμές του α και του ν από σχετικό πίνακα *(δείτε και το Σχήμα 11.4.1).*

Παράδειγμα 11.4.1 (συνέχεια του Παραδείγματος 11.2.2): *Επειδή για την τυχαία μεταβλητή Χ που εκφράζει την ημερήσια ποσότητα σεληνίου που προσλαμβάνουν από τη διατροφή τους οι ενήλικες στην περιοχή Α, έχουμε ότι* $X \sim N(\mu_1, \sigma_1^2)$ *και η* *δειγματική τυπική απόκλιση είναι* $s_1 = 8.4 \mu gr$*, ένα 95% διάστημα εμπιστοσύνης για την άγνωστη διακύμανση* σ_1^2 *της X, είναι το διάστημα*

$$\left[\frac{15}{\chi^2_{15;0.025}}\cdot 8.4^2,\ \ \frac{15}{\chi^2_{15;0.975}}\cdot 8.4^2\right] \text{ ή } \left[\frac{15}{27.488}\cdot 8.4^2,\ \ \frac{15}{6.262}\cdot 8.4^2\right] \text{ ή } [38.5,\ 169.0].$$

Επομένως, μπορούμε να συμπεράνουμε ότι με πιθανότητα 0.95 το διάστημα $[38.5,\ 169.0]$ περιέχει την άγνωστη *διακύμανση* σ_1^2 της X.

Επίσης, μπορούμε να συμπεράνουμε ότι με πιθανότητα 0.95, το διάστημα $[\sqrt{38.5},\ \sqrt{169.0}]$ ή $[6.2,\ 13]$ περιέχει την άγνωστη *τυπική απόκλιση* σ_1 της X.

Ως άσκηση, μπορείτε να κατασκευάσετε ένα 99% *διάστημα εμπιστοσύνης* για την άγνωστη *διασπορά* σ_2^2 της Y.

■

Σημείωση 11.4.1: *Το διάστημα εμπιστοσύνης που δώσαμε για τη διακύμανση* σ^2 *ενός κανονικού πληθυσμού, όπως είδαμε, κατασκευάσθηκε «αδιαφορώντας» για τη μέση τιμή του πληθυσμού* μ*. Αν η μέση τιμή του πληθυσμού* μ*, είναι γνωστή, χρησιμοποιώντας την τυχαία μεταβλητή,*

$$\frac{\nu S'^2}{\sigma^2}, \text{ όπου, } S'^2 = \frac{1}{\nu}\sum_{i=1}^{\nu}(X_i - \mu)^2$$

(αντί την $(\nu-1)S^2/\sigma^2$ *), με ανάλογο τρόπο (επειδή,* $\nu S'^2/\sigma^2 \sim \chi^2_\nu$*) προκύπτει ότι για συγκεκριμένη πραγματοποίηση* $x_1, x_2, \ldots, x_\nu$*, του τυχαίου δείγματος* $X_1, X_2, \ldots, X_\nu$*, το διάστημα*

$$\left[\frac{\nu}{\chi^2_{\nu;\alpha/2}}s'^2,\ \ \frac{\nu}{\chi^2_{\nu;1-\alpha/2}}s'^2\right]$$

είναι ένα $100(1-\alpha)\%$ *διάστημα εμπιστοσύνης για τη διακύμανση* σ^2 *του πληθυσμού.*

■

Συχνά απαιτείται να εκτιμήσουμε τη διαφορά των *μέσων τιμών* δύο πληθυσμών ή το *λόγο των διακυμάνσεών* τους. Ας δούμε πώς μπορούμε να κατασκευάσουμε *διαστήματα εμπιστοσύνης* για αυτές τις περιπτώσεις.

11.5 Διάστημα εμπιστοσύνης για τη διαφορά των μέσων τιμών δύο πληθυσμών

Ο φοιτητής στο εισαγωγικό παράδειγμα αυτού του κεφαλαίου ενδιαφέρεται να εκτιμήσει πόσο διαφέρει η μέση ημερήσια ποσότητα σεληνίου που προσλαμβάνουν οι ενήλικες μέσω της διατροφής τους στις δύο περιοχές A και B. Το τμήμα ποιοτικού ελέγχου μιας παραγωγικής μονάδας ενδιαφέρεται να εκτιμήσει πόσο διαφέρει το ποσοστό των ελαττωματικών προϊόντων που παράγονται από τη γραμμή παραγωγής A από το ποσοστό των ελαττωματικών προϊόντων που παράγονται από τη γραμμή παραγωγής B. Σε μια μελέτη για την περιεκτικότητα των ζαχαρότευτλων σε σάκχαρα, πρέπει να εκτιμήσουμε πόσο διαφέρει η μέση περιεκτικότητα σε σάκχαρα μιας ποικιλίας ζαχαρότευτλων από τη μέση περιεκτικότητα μιας άλλης ποικιλίας ζαχαρότευτλων.

Μια απάντηση σε ερωτήματα όπως αυτά μπορεί προφανώς να δοθεί με την κατασκευή κατάλληλου διαστήματος εμπιστοσύνης. Συγκεκριμένα, με την κατασκευή διαστήματος εμπιστοσύνης για τη διαφορά των *μέσων τιμών* των αντίστοιχων πληθυσμών.

Για να κατασκευάσουμε ένα $100(1-\alpha)\%$ *διάστημα εμπιστοσύνης* για τη *διαφορά των μέσων τιμών* δύο πληθυσμών εργαζόμαστε όπως εργασθήκαμε στα προηγούμενα για την κατασκευή ενός $100(1-\alpha)\%$ *διαστήματος εμπιστοσύνης* για τη *μέση τιμή* ενός πληθυσμού. Με μια βέβαια διαφορά. Αντί για ένα δείγμα χρησιμοποιούμε δύο, ένα από κάθε πληθυσμό. Ας δούμε ένα παράδειγμα που δόθηκε από τον *Student* (στην εργασία *The probable error of a mean, Biometrika, 6, 1-25, 1908*).

Παράδειγμα 11.5.1: *Σε δέκα ασθενείς που πάσχουν από αϋπνία δόθηκε μια φαρμακευτική αγωγή, έστω A, και για καθέναν καταγράφηκε η παράταση του ύπνου (σε ώρες). Στους ίδιους ασθενείς δόθηκε και μια άλλη φαρμακευτική αγωγή, έστω B, και επίσης καταγράφηκε η παράταση του ύπνου για κάθε ασθενή (σε ώρες). Ας συμβολίσουμε με X την τυχαία μεταβλητή που εκφράζει την παράταση του ύπνου σε ώρες όταν οι ασθενείς παίρνουν τη φαρμακευτική αγωγή A και με Y την τυχαία μεταβλητή που εκφράζει την παράταση του ύπνου σε ώρες όταν οι ασθενείς παίρνουν τη φαρμακευτική αγωγή B. Στον πίνακα που ακολουθεί φαίνονται τα δεδομένα που συγκεντρώθηκαν.*

Αύξων αριθμός ασθενούς (i)	1	2	3	4	5	6	7	8	9	10
Παράταση ύπνου (σε ώρες) υπό τη φαρμακευτική αγωγή A (x_i)	1.9	0.8	1.1	0.1	-0.1	4.4	5.5	1.6	4.6	3.4
Παράταση ύπνου (σε ώρες) υπό τη φαρμακευτική αγωγή B (y_i)	0.7	-1.6	-0.2	-1.2	-0.1	3.4	3.7	0.8	0.0	2.0

Αν μ_1 η *μέση τιμή* της *τυχαίας μεταβλητής X*, δηλαδή, η *μέση παράταση του ύπνου ασθενών που πάσχουν από αϋπνία όταν παίρνουν την αγωγή A* και μ_2 η *μέση τιμή* της *τυχαίας μεταβλητής Y*, δηλαδή, η *μέση παράταση του ύπνου ασθενών που πάσχουν από αϋπνία όταν παίρνουν την αγωγή B,* το ζητούμενο είναι να κατασκευάσουμε, από τα τυχαία δείγματα $X_1, X_2, \ldots, X_{10}$ και $Y_1, Y_2, \ldots, Y_{10}$ ένα $100(1-\alpha)\%$ *διάστημα εμπιστοσύνης* για τη *μέση τιμή* $\mu_1 - \mu_2$, της διαφοράς $X - Y$ των *τυχαίων μεταβλητών X* και *Y*.

Είναι εύλογο, ως *εκτιμήτρια* της διαφοράς $\mu_1 - \mu_2$ να χρησιμοποιήσουμε τη διαφορά $\overline{X} - \overline{Y}$ των αντίστοιχων *δειγματικών μέσων*. Όμως δε μπορούμε να αξιοποιήσουμε τα σχετικά αποτελέσματα της *Θεωρίας Πιθανοτήτων* και της *Στατιστικής* για την κατανομή της $\overline{X} - \overline{Y}$ που παρουσιάσαμε στο *10ο Κεφάλαιο* γιατί τα δύο δείγματα δεν έχουν ληφθεί το ένα ανεξάρτητα από το άλλο. Βέβαια, τα ζεύγη

$$(X_1, Y_1),\ (X_2, Y_2),\ \ldots, (X_{10}, Y_{10})$$

είναι μεταξύ τους ανεξάρτητα αφού αναφέρονται σε διαφορετικούς ασθενείς τυχαία επιλεγμένους, όμως τα X_i και Y_i εντός του ίδιου ζεύγους δεν είναι ανεξάρτητα (δε μπορούν να θεωρηθούν ανεξάρτητα) αφού αναφέρονται στον ίδιο ασθενή.

Για να κατασκευάσουμε επομένως, ένα $100(1-\alpha)\%$ *διάστημα εμπιστοσύνης* για τη *διαφορά των μέσων τιμών* δύο πληθυσμών από δύο δείγματα, ένα από κάθε πληθυσμό, πρέπει να διακρίνουμε την περίπτωση που τα δύο δείγματα *είναι ανεξάρτητα* από την περίπτωση που τα δύο δείγματα *δεν είναι ανεξάρτητα*. Υπενθυμίζουμε ότι στην πράξη, η ανεξαρτησία (είτε ενδεχομένων, είτε μεταβλητών, είτε πειραμάτων) συνήθως αναγνωρίζεται και διαπιστώνεται από την περιγραφή των μεταβλητών και από τις συνθήκες εκτέλεσης των πειραμάτων και δεν απαιτείται να ελεγχθούν/αποδειχθούν οι αντίστοιχες συνθήκες ανεξαρτησίας *(ξαναδείτε και την Ενότητα 5.5 του 5^{ου} Κεφαλαίου).*

11.5.1 Εξαρτημένα δείγματα/Ζευγαρωτές παρατηρήσεις

Έστω ότι από κάθε *δειγματοληπτική ή πειραματική μονάδα* παίρνουμε ένα *ζεύγος παρατηρήσεων* (X_i, Y_i), $i = 1,2,\ldots,\nu$, δηλαδή, έστω ότι από τα δύο δείγματα $X_1, X_2,\ldots, X_\nu$ και $Y_1, Y_2,\ldots, Y_\nu$ δημιουργούνται ν *ζεύγη παρατηρήσεων,* $(X_1,Y_1), (X_2,Y_2),\ldots,(X_\nu,Y_\nu)$, τα οποία είναι μεταξύ τους ανεξάρτητα, όμως τα X_i και Y_i εντός του ίδιου ζεύγους δεν είναι ανεξάρτητα. Σε επόμενα κεφάλαια (στο *12^o* και στο *13^o*) θα εξηγήσουμε πολύ αναλυτικά σε ποιες περιπτώσεις (και γιατί) ενδείκνυται να εργαζόμαστε με τέτοια δείγματα τα οποία στη βιβλιογραφία αναφέρονται ως *ζευγαρωτές παρατηρήσεις* ***(paired samples, matched samples, dependent samples,)*** αλλά και πώς/με ποια κριτήρια πρέπει να δημιουργούνται τα *ζεύγη*. Στο σημείο αυτό αναφέρουμε μόνο ότι ένα *ζεύγος παρατηρήσεων* μπορεί να ληφθεί από την ίδια *πειραματική ή δειγματοληπτική μονάδα* **πριν** και **μετά** από κάποια επέμβαση ή υπό δύο διαφορετικές επεμβάσεις σε κάποια χρονική απόσταση (όπως στο προηγούμενο παράδειγμα), αλλά και από δύο όμοιες ή σχεδόν όμοιες μονάδες.

Παίρνοντας για κάθε ζεύγος (X_i, Y_i) τη διαφορά

$$D_i = X_i - Y_i$$

(η οποία στο παράδειγμά μας εκφράζει τη διαφορά των επιδράσεων των δύο φαρμάκων στον ασθενή *i*), δημιουργούμε ένα *τυχαίο δείγμα*

$$D_1, D_2,\ldots, D_\nu$$

το οποίο μπορούμε να θεωρήσουμε ότι προέρχεται από ένα (θεωρητικό) πληθυσμό, τον πληθυσμό των διαφορών, με *μέση τιμή*

$$\mu_D = \mu_1 - \mu_2$$

όπου μ_1 η *μέση τιμή* της *Χ* και μ_2 η *μέση τιμή* της *Υ* και επομένως, για την κατασκευή ενός $100(1-\alpha)\%$ *διαστήματος εμπιστοσύνης* για τη *μέση τιμή* $\mu_D = \mu_1 - \mu_2$ ισχύει ό,τι έχουμε αναφέρει για την κατασκευή ενός $100(1-\alpha)\%$ *διαστήματος εμπιστοσύνης* για τη *μέση τιμή* **ενός** *πληθυσμού* από **ένα** *δείγμα.*

Έστω για παράδειγμα, ότι ο *πληθυσμός των διαφορών* είναι *κανονικός* με άγνωστη *διακύμανση* σ_D^2, και *μέση τιμή* $\mu_D = \mu_1 - \mu_2$ την οποία θέλουμε να εκτιμήσουμε με ένα $100(1-\alpha)\%$ *διάστημα εμπιστοσύνης*. Έστω επίσης

$$\overline{D} = \overline{X} - \overline{Y} \text{ και } S_D^2 = \frac{1}{\nu - 1}\sum_{i=1}^{\nu}(D_i - \overline{D})^2$$

ο *μέσος* και η *αμερόληπτη διακύμανση* του δείγματος των **διαφορών** $D_1, D_2,\ldots, D_\nu$ αντίστοιχα. Τότε με βάση όσα αναφέραμε στα προηγούμενα, το διάστημα

$$\overline{D} \pm t_{\nu-1;\alpha/2}\frac{S_D}{\sqrt{\nu}}$$

είναι ένα $100(1-\alpha)\%$ *διάστημα εμπιστοσύνης* για τη *μέση τιμή* μ_D του πληθυσμού των διαφορών δηλαδή για τη διαφορά $\mu_1 - \mu_2$.

Επομένως, για συγκεκριμένη πραγματοποίηση $(x_1, y_1), (x_2, y_2), \ldots, (x_\nu, y_\nu)$ των $(X_1,Y_1), (X_2,Y_2),\ldots,(X_\nu,Y_\nu)$, το διάστημα

$$\overline{d} \pm t_{\nu-1;\alpha/2}\frac{s_D}{\sqrt{\nu}} \text{ ή } \overline{x} - \overline{y} \pm t_{\nu-1;\alpha/2}\frac{s_D}{\sqrt{\nu}}$$

είναι ένα $100(1-\alpha)\%$ *διάστημα εμπιστοσύνης* για τη διαφορά $\mu_1 - \mu_2$ των *μέσων τιμών* των δύο πληθυσμών με *ζευγαρωτές παρατηρήσεις*.

Ας επανέλθουμε τώρα στο *Παράδειγμα 11.5.1.*

Παράδειγμα 11.5.1 (συνέχεια): Για κάθε ζεύγος (x_i, y_i), $i = 1, 2, ..., 10$ υπολογίζουμε τη διαφορά, $d_i = x_i - y_i$, η οποία εκφράζει τη διαφορά των επιδράσεων των δύο φαρμάκων στον ασθενή *i*.

Αύξων αριθμός ασθενούς (i)	1	2	3	4	5	6	7	8	9	10
x_i	1.9	0.8	1.1	0.1	-0.1	4.4	5.5	1.6	4.6	3.4
y_i	0.7	-1.6	-0.2	-1.2	-0.1	3.4	3.7	0.8	0.0	2.0
$d_i = x_i - y_i$	1.2	2.4	1.3	1.3	0.0	1.0	1.8	0.8	4.6	1.4

Η τιμή $\bar{d}$ του *δειγματικού μέσου* $\bar{D} = \bar{X} - \bar{Y}$ και η τιμή s_D^2 της *δειγματικής διακύμανσης*

$$S_D^2 = \frac{1}{9}\sum_{i=1}^{10}(D_i - \bar{D})^2$$

για τη συγκεκριμένη πραγματοποίηση του τυχαίου δείγματος των διαφορών $D_1, D_2, \ldots, D_{10}$, είναι αντίστοιχα

$$\bar{d} = \frac{\sum_{i=1}^{10} d_i}{10} = \frac{1.2 + 2.4 + \ldots + 4.6 + 1.4}{10} = \frac{15.8}{10} = 1.58$$

και

$$s_D^2 = \frac{\sum_{i=1}^{10}(d_i - 1.58)^2}{9} = \frac{(1.2-1.58)^2 + (2.4-1.58)^2 + \ldots + (1.4-1.58)^2}{9} = 1.51.$$

Έτσι, με την παραδοχή ότι ο πληθυσμός των διαφορών είναι *κανονικός*, το διάστημα

$$\bar{d} \pm t_{\nu-1;0.025}\frac{s_D}{\sqrt{\nu}} \text{ ή } 1.58 \pm t_{9;0.025}\frac{1.23}{\sqrt{10}} \text{ ή } 1.58 \pm 2.262 \cdot 0.39 \text{ ή } 1.58 \pm 0.88$$

είναι ένα 95% *διάστημα εμπιστοσύνης* για τη διαφορά $\mu_1 - \mu_2$ με βάση τα συγκεκριμένα δείγματα.

*Επομένως, για να κατασκευάσουμε **με ζευγαρωτές παρατηρήσεις** διαστήματα εμπιστοσύνης για τη διαφορά των μέσων τιμών $\mu_1 - \mu_2$ δύο πληθυσμών, εργαζόμαστε με ένα δείγμα, αυτό των διαφορών $D_1, D_2, \ldots, D_\nu$ και ισχύει ό,τι έχουμε αναφέρει για την κατασκευή διαστημάτων εμπιστοσύνης για τη μέση τιμή μ ενός πληθυσμού με ένα δείγμα.*

■

Ας δούμε τώρα πώς εργαζόμαστε όταν τα δύο δείγματα λαμβάνονται το ένα ανεξάρτητα από το άλλο, όπως στο εισαγωγικό παράδειγμα.

11.5.2 Ανεξάρτητα δείγματα

Έστω δύο πληθυσμοί *A* και *B*, με *μέση τιμή* και *διακύμανση*, μ_1, σ_1^2 και μ_2, σ_2^2, αντίστοιχα. Έστω επίσης, δύο ***ανεξάρτητα*** *τυχαία δείγματα* $X_1, X_2, \ldots, X_{\nu_1}$ και $Y_1, Y_2, \ldots, Y_{\nu_2}$ (μεγέθους ν_1 και ν_2 αντίστοιχα), ένα από κάθε *πληθυσμό*.

Ας συμβολίσουμε με $\overline{X}$ και S_1^2 τον *μέσο* και τη *διακύμανση* του δείγματος $X_1, X_2, \ldots, X_{\nu_1}$ από τον *πληθυσμό Α* και με $\overline{Y}$ και S_2^2 τον *μέσο* και τη *διακύμανση* του δείγματος $Y_1, Y_2, \ldots, Y_{\nu_2}$ από τον *πληθυσμό Β*. Δηλαδή

$$\overline{X} = \frac{1}{\nu_1}\sum_{i=1}^{\nu_1} X_i\,, \quad S_1^2 = \frac{1}{\nu_1 - 1}\sum_{i=1}^{\nu_1}(X_i - \overline{X})^2$$

και

$$\overline{Y} = \frac{1}{\nu_2}\sum_{j=1}^{\nu_2} Y_j\,,\ S_2^2 = \frac{1}{\nu_2 - 1}\sum_{j=1}^{\nu_2}(Y_j - \overline{Y})^2\,.$$

Προφανώς, ως *εκτιμήτρια* της διαφοράς $\mu_1 - \mu_2$, θα χρησιμοποιήσουμε την αντίστοιχη διαφορά $\overline{X} - \overline{Y}$, των *δειγματικών μέσων* για την κατανομή της οποίας μπορούμε να αξιοποιήσουμε τα σχετικά αποτελέσματα της *Θεωρίας Πιθανοτήτων* και της *Στατιστικής* που αναφέραμε στο *10^o Κεφάλαιο* αφού τα δύο δείγματα είναι μεταξύ τους *ανεξάρτητα*.

Με βάση όσα αναφέραμε στο *10^o Κεφάλαιο* και εργαζόμενοι όπως εργασθήκαμε για την κατασκευή *διαστημάτων εμπιστοσύνης* για τη *μέση τιμή* ή τη *διακύμανση* ενός *πληθυσμού* από ένα *τυχαίο δείγμα*, εύκολα παίρνουμε κατά περίπτωση τα *διαστήματα εμπιστοσύνης που φαίνονται στον Πίνακα 11.5.1.*

Πληθυσμοί	*Διακυμάνσεις των πληθυσμών* $\sigma_1^2,\ \sigma_2^2$	*Μεγέθη των δειγμάτων* ν_1, ν_2	$100(1-\alpha)\%$ *Δ. Ε. για τη διαφορά* $\mu_1 - \mu_2$ *των μέσων τιμών των πληθυσμών*
Κανονικοί	*Γνωστές*	*Οτιδήποτε*	$(\overline{X} - \overline{Y}) \pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{\nu_1} + \frac{\sigma_2^2}{\nu_2}}$ *(1)*
Οποιοιδήποτε	*Γνωστές*	*Μεγάλα*	
Οποιοιδήποτε	*Άγνωστες*	*Μεγάλα*	$(\overline{X} - \overline{Y}) \pm z_{\alpha/2}\sqrt{\frac{S_1^2}{\nu_1} + \frac{S_2^2}{\nu_2}}$ *(2)*
Κανονικοί	*Άγνωστες και **ίσες***	*Οτιδήποτε*	$(\overline{X} - \overline{Y}) \pm t_{\nu;\alpha/2} S\sqrt{\frac{1}{\nu_1} + \frac{1}{\nu_2}}$ *(3)* *Όπου,* $\nu = \nu_1 + \nu_2 - 2$ *και* $S = \sqrt{\frac{(\nu_1 - 1)S_1^2 + (\nu_2 - 1)S_2^2}{\nu_1 + \nu_2 - 2}}$
Όχι Κανονικοί	*Γνωστές ή άγνωστες (ίσες ή άνισες)*	*Μικρά*	*?*

Πίνακας 11.5.1

$100(1-\alpha)\%$ *Διαστήματα εμπιστοσύνης για τη διαφορά* $\mu_1 - \mu_2$ *των μέσων τιμών δύο πληθυσμών από δύο ανεξάρτητα δείγματα μεγέθους* ν_1 *και* ν_2 *αντίστοιχα.*

Σημείωση 11.5.1: *α) Σε κάθε περίπτωση, τα δείγματα πρέπει να είναι τυχαία. β) Για μικρά δείγματα από κανονικούς πληθυσμούς με άγνωστες διακυμάνσεις, έχει προταθεί προσεγγιστικός τύπος που δεν προϋποθέτει ότι οι άγνωστες διακυμάνσεις είναι ίσες, όμως αν και χρησιμοποιείται, δεν είναι γενικής αποδοχής. γ) Ο τύπος (2) και ο τύπος*

(1) (όταν οι πληθυσμοί δεν είναι κανονικοί) δίνουν $100(1-\alpha)\%$ *διαστήματα εμπιστοσύνης* ***κατά προσέγγιση***.

■

Ας δούμε πάλι το εισαγωγικό παράδειγμα.

Παράδειγμα 11.5.2 (συνέχεια του Παραδείγματος 11.2.2): Το *τυχαίο δείγμα* μεγέθους $\nu_1 = 16$ που πήρε ο φοιτητής από την περιοχή Α, προέρχεται από *κανονικό πληθυσμό* με άγνωστη *μέση τιμή* μ_1 και άγνωστη *διακύμανση* σ_1^2, δηλαδή, $X_1, X_2, \ldots, X_{16} \sim N(\mu_1, \sigma_1^2)$ και αντίστοιχα, το *τυχαίο δείγμα* μεγέθους $\nu_2 = 16$ που πήρε από την περιοχή Β προέρχεται από *κανονικό* επίσης *πληθυσμό* με άγνωστη *μέση τιμή* μ_2 και άγνωστη *διακύμανση* σ_2^2, δηλαδή, $Y_1, Y_2, \ldots, Y_{16} \sim N(\mu_2, \sigma_2^2)$.

Επειδή οι δύο πληθυσμοί είναι *κανονικοί* με *διακυμάνσεις* άγνωστες και τα δύο δείγματα είναι μικρά, με την παραδοχή ότι οι άγνωστες *διακυμάνσεις* είναι ίσες, ένα 95% *διάστημα εμπιστοσύνης* για τη διαφορά, $\mu_1 - \mu_2$, των *μέσων τιμών* των δύο *πληθυσμών* είναι

$$(\bar{x} - \bar{y}) \pm t_{\nu_1+\nu_2-2;0.025} \cdot s\sqrt{\frac{1}{\nu_1} + \frac{1}{\nu_2}} \ ,$$

όπου, $\bar{x} = 75\mu gr$, $\bar{y} = 87\mu gr$ και

$$s = \sqrt{\frac{(\nu_1 - 1)s_1^2 + (\nu_2 - 1)s_2^2}{\nu_1 + \nu_2 - 2}} = \sqrt{\frac{15 \cdot 8.4^2 + 15 \cdot 9^2}{30}} = 8.7 \, .$$

Επομένως, ένα 95% *διάστημα εμπιστοσύνης* για τη διαφορά, $\mu_1 - \mu_2$, από τα δύο *ανεξάρτητα τυχαία δείγματα* που πήρε ο φοιτητής, είναι το διάστημα

$$(75 - 87) \pm t_{30;0.025} \cdot 8.7 \cdot \sqrt{\frac{1}{16} + \frac{1}{16}} \ \text{ ή } -12 \pm 1.96 \cdot 3.08 \ \text{ ή } -12 \pm 6.04 \, .$$

Έτσι, με βάση τα δύο *ανεξάρτητα τυχαία δείγματα* που πήρε ο φοιτητής, μπορούμε να συμπεράνουμε ότι το διάστημα $[-18.04, -5.96]$ έχει 0.95 πιθανότητα να περιέχει την τιμή της διαφοράς $\mu_1 - \mu_2$ των *μέσων τιμών* των δύο *πληθυσμών*. Ασφαλώς, το συμπέρασμα αυτό είναι αξιόπιστο υπό την προϋπόθεση ότι η παραδοχή που κάναμε (ότι οι άγνωστες *διακυμάνσεις* είναι ίσες) πράγματι ισχύει. Στη συνέχεια *(στην Ενότητα 11.7)* θα δούμε πώς μπορούμε να ελέγξουμε μια τέτοια παραδοχή.

Παρατηρείστε ότι το 95% διάστημα εμπιστοσύνης $[-18.04, -5.96]$ *που υπολογίσαμε για τη διαφορά* $\mu_1 - \mu_2$ *των μέσων τιμών* των δύο *πληθυσμών δεν περιέχει το 0. Μπορούμε επομένως να συμπεράνουμε ότι* ***οι δύο μέσες τιμές δεν είναι ίσες μεταξύ τους και η πιθανότητα το συμπέρασμα αυτό να είναι λάθος είναι το πολύ 5%***.

■

11.6 Διάστημα εμπιστοσύνης για τη διαφορά δύο διωνυμικών ποσοστών με δύο ανεξάρτητα τυχαία δείγματα

Παράδειγμα 11.6.1: *Στο περιοδικό journal of Biology δημοσιεύθηκαν τα αποτελέσματα μιας έρευνας για το ποσοστό* p_1 *των ψαριών στη Μεσόγειο και το ποσοστό* p_2 *των ψαριών στον Ατλαντικό που έχουν προσβληθεί από παράσιτα. Στη Μεσόγειο, από 588 τυχαία επιλεγμένα ψάρια που εξετάσθηκαν βρέθηκαν μολυσμένα από παράσιτα τα 211 ενώ στον Ατλαντικό, από 123 τυχαία επιλεγμένα ψάρια που εξετάσθηκαν βρέθηκαν μολυσμένα από παράσιτα τα 26. Μπορούμε, με βάση τα δύο* *συγκεκριμένα ανεξάρτητα τυχαία δείγματα, να κατασκευάσουμε ένα 95% διάστημα εμπιστοσύνης για την εκτίμηση της διαφοράς $p_1 - p_2$ των δύο ποσοστών;*

<u>*Απάντηση:*</u> Είναι προφανές, ότι τα δύο *ανεξάρτητα τυχαία δείγματα* προέρχονται από δύο *πληθυσμούς Bernoulli* με *παραμέτρους* p_1 και p_2 αντίστοιχα.

Τίθεται επομένως το γενικότερο πρόβλημα: αν $X_1, X_2, \ldots, X_{\nu_1}$ και $Y_1, Y_2, \ldots, Y_{\nu_2}$ δύο *ανεξάρτητα τυχαία δείγματα* (μεγέθους ν_1 και ν_2 αντίστοιχα) που το καθένα προέρχεται από μια κατανομή *Bernoulli* με *παράμετρο (μέση τιμή)* p_1 και p_2 αντίστοιχα, πώς μπορούμε να κατασκευάσουμε ένα $100(1-\alpha)\%$ *διάστημα εμπιστοσύνης* για τη διαφορά $p_1 - p_2$;

Οι *δειγματικοί μέσοι*

$$\overline{X} = \frac{X_1 + X_2 + \ldots + X_{\nu_1}}{\nu_1} \text{ και } \overline{Y} = \frac{Y_1 + Y_2 + \ldots + Y_{\nu_2}}{\nu_2}$$

οι οποίοι, όπως έχουμε εξηγήσει, εκφράζουν τα αντίστοιχα *δειγματικά ποσοστά* $\hat{P}_1$ και $\hat{P}_2$, ακολουθούν κατά προσέγγιση *κανονική κατανομή* και συγκεκριμένα

$$\overline{X} \equiv \hat{P}_1 \sim N\left(p_1, \frac{p_1(1-p_1)}{\nu_1}\right), \text{ εφόσον } \nu_1 p_1 \geq 5 \textbf{ και } \nu_1(1-p_1) \geq 5$$

και $\overline{Y} \equiv \hat{P}_2 \sim N\left(p_2, \dfrac{p_2(1-p_2)}{\nu_2}\right)$, εφόσον $\nu_2 p_2 \geq 5$ **και** $\nu_2(1-p_2) \geq 5$.

Επειδή δε, τα δύο *τυχαία δείγματα* είναι *ανεξάρτητα,* για την *τυχαία μεταβλητή* $\hat{P}_1 - \hat{P}_2$ κατά προσέγγιση έχουμε

$$\hat{P}_1 - \hat{P}_2 \sim N\left(p_1 - p_2, \frac{p_1(1-p_1)}{\nu_1} + \frac{p_2(1-p_2)}{\nu_2}\right)$$ και επομένως, κατά προσέγγιση

$$\frac{(\hat{P}_1 - \hat{P}_2) - (p_1 - p_2)}{\sqrt{\dfrac{p_1(1-p_1)}{\nu_1} + \dfrac{p_2(1-p_2)}{\nu_2}}} \sim N(0, 1) .$$

Άρα, κατά προσέγγιση έχουμε

$$P\left(-z_{\alpha/2} \leq \frac{(\hat{P}_1 - \hat{P}_2) - (p_1 - p_2)}{\sqrt{\dfrac{p_1(1-p_1)}{\nu_1} + \dfrac{p_2(1-p_2)}{\nu_2}}} \leq z_{\alpha/2}\right) = 1 - \alpha$$

και χρησιμοποιώντας ως εκτιμήτρια της άγνωστης *διακύμανσης*

$$\frac{p_1(1-p_1)}{\nu_1} + \frac{p_2(1-p_2)}{\nu_2}$$

την

$$\frac{\hat{P}_1(1-\hat{P}_1)}{\nu_1} + \frac{\hat{P}_2(1-\hat{P}_2)}{\nu_2},$$

εύκολα προκύπτει ότι το διάστημα

$$(\hat{P}_1 - \hat{P}_2) \pm z_{\alpha/2} \sqrt{\frac{\hat{P}_1(1-\hat{P}_1)}{\nu_1} + \frac{\hat{P}_2(1-\hat{P}_2)}{\nu_2}}$$

είναι ένα **κατά προσέγγιση** $100(1-\alpha)\%$ *διάστημα εμπιστοσύνης* για τη διαφορά $p_1 - p_2$. Έτσι, για συγκεκριμένη πραγματοποίηση των δύο δειγμάτων, με *ποσοστά επιτυχιών* $\hat{p}_1$ και $\hat{p}_2$ αντίστοιχα, το διάστημα

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{\nu_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{\nu_2}}$$

είναι ένα **κατά προσέγγιση** $100(1-\alpha)\%$ *διάστημα εμπιστοσύνης* για τη διαφορά $p_1 - p_2$ δύο *διωνυμικών ποσοστών (ή των μέσων τιμών δύο κατανομών Bernoulli)* εφόσον $\nu_i \hat{p}_i \geq 5$ ***και*** $\nu_i(1-\hat{p}_i) \geq 5$, $i = 1, 2$.

Μπορούμε πλέον να απαντήσουμε και στο ερώτημα του παραδείγματός μας. Τα δύο *δειγματικά ποσοστά* είναι αντίστοιχα

$$\hat{p}_1 = \frac{211}{588} = 0.36 \text{ και } \hat{p}_2 = \frac{26}{123} = 0.21.$$

Επειδή, $\nu_1\hat{p}_1 = 211 \geq 5$ **,** $\nu_1(1-\hat{p}_1) = 377 \geq 5$, $\nu_2\hat{p}_2 = 26 \geq 5$ και $\nu_2(1-\hat{p}_2) = 97 \geq 5$, το διάστημα

$$(\hat{p}_1 - \hat{p}_2) \pm z_{0.025}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{\nu_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{\nu_2}} \quad \text{ή}$$

$$(0.36 - 0.21) \pm 1.96\sqrt{\frac{0.36 \cdot 0.64}{588} + \frac{0.21 \cdot 0.79}{123}} \quad \text{ή}$$

$$0.15 \pm 0.08$$

είναι ένα κατά προσέγγιση 95% *διάστημα εμπιστοσύνης* για τη διαφορά $p_1 - p_2$.

Επομένως, το διάστημα $[0.07, 0.23]$ έχει πιθανότητα περίπου 0.95 να περιέχει τη διαφορά $p_1 - p_2$ των δύο *πληθυσμιακών ποσοστών*.

■

Ας δούμε μια τελευταία περίπτωση κατασκευής *διαστήματος εμπιστοσύνης*.

11.7 Διάστημα εμπιστοσύνης για το λόγο των διακυμάνσεων δύο κανονικών πληθυσμών

Το συμπέρασμα στο *Παράδειγμα 11.5.2* όπως εξηγήσαμε είναι αξιόπιστο υπό την προϋπόθεση ότι οι άγνωστες *διακυμάνσεις,* σ_1^2 και σ_2^2, των δύο *πληθυσμών* είναι ίσες. Δηλαδή, όταν ο λόγος των *διακυμάνσεων* των δύο πληθυσμών είναι ίσος με 1. Είναι χρήσιμο επομένως να μπορούμε να κατασκευάσουμε ένα *διάστημα εμπιστοσύνης* για το λόγο των *διακυμάνσεων* δύο *πληθυσμών*. Βέβαια, η κατασκευή *διαστήματος εμπιστοσύνης* για το λόγο των *διακυμάνσεων* δύο *πληθυσμών*, δεν προκύπτει μόνο ως ένα δευτερογενές πρόβλημα. Σε πολλά προβλήματα (ιδιαίτερα σε προβλήματα ποιοτικού ελέγχου και ελέγχου προδιαγραφών) μας ενδιαφέρει αυτή καθαυτή η σύγκριση των *διακυμάνσεων* δύο πληθυσμών.

Έστω λοιπόν δύο *ανεξάρτητα τυχαία δείγματα* $X_1, X_2, \ldots, X_{\nu_1} \sim N(\mu_1, \sigma_1^2)$ και $Y_1, Y_2, \ldots, Y_{\nu_2} \sim N(\mu_2, \sigma_2^2)$. Επειδή (όπως είδαμε στο *10^o Κεφάλαιο*)

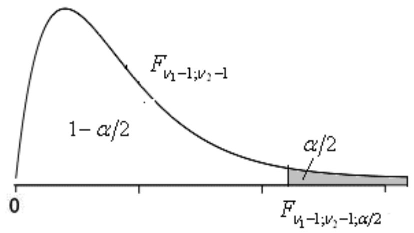
$$\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F_{\nu_1-1;\nu_2-1}$$

εύκολα αποδεικνύεται ότι το διάστημα

$$\left[\frac{1}{F_{\nu_1-1;\nu_2-1;\alpha/2}}\frac{S_1^2}{S_2^2},\ F_{\nu_2-1;\nu_1-1;\alpha/2}\frac{S_1^2}{S_2^2}\right]$$

είναι ένα $100(1-\alpha)\%$ *διάστημα εμπιστοσύνης* για το λόγο των *διακυμάνσεων,* σ_1^2/σ_2^2 **,** δύο *κανονικών πληθυσμών.*

Υπενθυμίζουμε ότι με $F_{\nu_1-1;\nu_2-1;\alpha/2}$ συμβολίζουμε το *άνω* $\alpha/2$*- ποσοστιαίο σημείο* της κατανομής $F_{\nu_1-1;\nu_2-1}$ *(δείτε και το Σχήμα 11.7.1).*



Σχήμα 11.7.1
Το άνω $\alpha/2$*-ποσοστιαίο σημείο της κατανομής* $F_{\nu_1-1;\nu_2-1}$

Έτσι, για το παράδειγμά μας, ένα 95% *διάστημα εμπιστοσύνης* για το λόγο των *διακυμάνσεων,* σ_1^2/σ_2^2 , των δύο *πληθυσμών* από τα δύο *ανεξάρτητα τυχαία δείγματα* που πήρε ο φοιτητής είναι το διάστημα

$$\left[\frac{1}{F_{15;15;0.025}}\frac{8.4^2}{9^2},\ F_{15;15;0.025}\frac{8.4^2}{9^2}\right] \text{ ή } \left[\frac{1}{2.86}\frac{8.4^2}{9^2},\ 2.86\frac{8.4^2}{9^2}\right] \text{ ή } [0.304,\ 2.49].$$

Με βάση επομένως, τα δύο *ανεξάρτητα τυχαία δείγματα* που πήρε ο φοιτητής, μπορούμε να συμπεράνουμε ότι το διάστημα, $[0.304,\ 2.49]$ έχει 0.95 πιθανότητα να περιέχει την τιμή του λόγου των *διακυμάνσεων,* σ_1^2/σ_2^2 , των δύο *πληθυσμών.*

Παρατηρείστε ότι το 95% διάστημα εμπιστοσύνης $[0.304,\ 2.49]$ *για το λόγο των διακυμάνσεων* σ_1^2/σ_2^2 *των δύο πληθυσμών περιέχει την τιμή 1.* ***Μπορούμε άραγε να συμπεράνουμε κάτι για τις διακυμάνσεις των δύο πληθυσμών;***

■

11.8 Πάνω (κάτω) φράγμα εμπιστοσύνης

Σε πρακτικά προβλήματα και εφαρμογές, παρουσιάζονται περιπτώσεις όπου αντί να κατασκευάσουμε ένα διάστημα εμπιστοσύνης για μια άγνωστη παράμετρο, δηλαδή, αντί να δώσουμε ένα πάνω και ένα κάτω φράγμα για την άγνωστη παράμετρο, μας αρκεί (ή είναι προτιμότερο) να υπολογίσουμε ένα μόνο φράγμα (πάνω ή κάτω). Σε αυτές τις περιπτώσεις, η εκτίμηση γίνεται με ένα $100(1-\alpha)\%$ ***πάνω (ή κάτω) φράγμα εμπιστοσύνης (upper or lower confidence bound)***. Για παράδειγμα, αν ο υπεύθυνος παραγωγής στο *Παράδειγμα 11.2.1* ανησυχεί ότι η ποσότητα καλαμποκιού που συσκευάζεται ανά τσουβάλι ξεπερνάει κάποιο ανώτερο όριο, αρκεί να υπολογίσει ένα 95% πάνω φράγμα εμπιστοσύνης.

Για να κατασκευάσουμε ένα $100(1-\alpha)\%$ πάνω (ή κάτω) φράγμα εμπιστοσύνης για τη μέση τιμή ενός πληθυσμού ή για τη διαφορά των μέσων τιμών δύο πληθυσμών κτλ, εργαζόμαστε όπως στα προηγούμενα. Για παράδειγμα, εύκολα προκύπτει (δείτε το ως μια απλή άσκηση) ότι ένα $100(1-\alpha)\%$ *πάνω φράγμα εμπιστοσύνης* για τη μέση τιμή ενός κανονικού πληθυσμού με άγνωστη διακύμανση από ένα τυχαίο δείγμα μεγέθους ν είναι

$$\bar{x} + t_{\nu-1;\alpha} \frac{s}{\sqrt{\nu}}.$$

Παρατηρείστε ότι είναι μικρότερο από το δεξιό άκρο

$$\bar{x} + t_{\nu-1;\alpha/2} \frac{s}{\sqrt{\nu}}$$

του αντίστοιχου $100(1-\alpha)\%$ *διαστήματος εμπιστοσύνης*. Σκεφθείτε τι σημαίνει αυτό για την ακρίβεια της εκτίμησης.

■

Προβλήματα για εξάσκηση στα θέματα που παρουσιάσαμε σε αυτό το κεφάλαιο δίνονται στο τέλος του κεφαλαίου που ακολουθεί . Για αυτοαξιολόγηση στις βασικές έννοιες του κεφαλαίου δείτε επίσης τις ερωτήσεις κατανόησης που ακολουθούν.

Ερωτήσεις κατανόησης

1. *Από έναν πληθυσμό επιλέξατε ένα τυχαίο δείγμα μεγέθους ν και κατασκευάσατε ένα 95% διάστημα εμπιστοσύνης για την άγνωστη μέση τιμή του. Αν από το ίδιο δείγμα κατασκευάσετε για τη μέση τιμή του πληθυσμού ένα άλλο διάστημα εμπιστοσύνης με μεγαλύτερο συντελεστή εμπιστοσύνης, η ακρίβεια της εκτίμησης θα αυξηθεί, θα μειωθεί ή μήπως θα παραμείνει ίδια;*
2. *Κατασκευάσατε ένα διάστημα εμπιστοσύνης για έναν άγνωστο πληθυσμιακό μέσο και διαπιστώσατε ότι ουσιαστικά δεν είναι ιδιαίτερα χρήσιμο γιατί είναι πολύ πλατύ. Τι από τα παρακάτω θα κάνατε για να αντιμετωπίσετε αυτό το πρόβλημα. α) Αύξηση του συντελεστή εμπιστοσύνης. β) Επανάληψη της δειγματοληψίας με μεγαλύτερο μέγεθος δείγματος. γ) Επανάληψη της δειγματοληψίας με μικρότερο μέγεθος δείγματος.*
3. *Κατασκευάσατε ένα 95% διάστημα εμπιστοσύνης για έναν άγνωστο πληθυσμιακό μέσο. Αυτό σημαίνει ότι δώσατε για τον άγνωστο πληθυσμιακό μέσο α) μια σημειακή εκτίμηση με ακρίβεια 95% β) μια εκτίμηση με διάστημα και με ακρίβεια 95% γ) μια εκτίμηση με διάστημα και περιθώριο σφάλματος 5% δ) όλα τα προηγούμενα ε) τίποτε από τα προηγούμενα.*
4. *Ένας ερευνητής πήρε από έναν πληθυσμό ένα τυχαίο δείγμα μεγέθους ν και κατασκεύασε ένα 95% διάστημα εμπιστοσύνης για την άγνωστη μέση τιμή του πληθυσμού. Αυτό σημαίνει ότι α) το διάστημα αυτό έχει 0.95 πιθανότητα να περιέχει την τιμή της μέσης τιμής του πληθυσμού β) η μέση τιμή του πληθυσμού έχει 0.95 πιθανότητα να βρίσκεται μέσα σε αυτό το διάστημα γ) αν από τον πληθυσμό πάρουμε διαφορετικά τυχαία δείγματα μεγέθους ν (το καθένα) και από κάθε δείγμα κατασκευάσουμε (με την ίδια διαδικασία) ένα 95% διάστημα εμπιστοσύνης για τη μέση τιμή του πληθυσμού τότε ποσοστό 95% αυτών των δειγμάτων θα δώσουν διάστημα που θα περιέχει τη μέση τιμή του πληθυσμού δ) το διάστημα αυτό περιέχει το 95% του πληθυσμού ε) τίποτε από τα προηγούμενα στ) όλα τα προηγούμενα. Ποια (ή ποιες) από τις έξι εκδοχές είναι σωστή (σωστές);*
5. *Από έναν πληθυσμό πήρατε ένα τυχαίο δείγμα μεγέθους ν_1 και κατασκευάσατε για την άγνωστη μέση τιμή του ένα 95% διάστημα εμπιστοσύνης. Από τον ίδιο πληθυσμό πήρατε ένα ακόμη τυχαίο δείγμα μεγέθους $\nu_2 > \nu_1$ και κατασκευάσατε ένα επίσης 95% διάστημα εμπιστοσύνης για τη μέση τιμή του πληθυσμού. Ποιο από τα δύο διαστήματα δίνει ακριβέστερη εκτίμηση της μέσης τιμής του πληθυσμού; Μήπως οι δύο εκτιμήσεις είναι το ίδιο ακριβείς (με 95% ακρίβεια);*

■