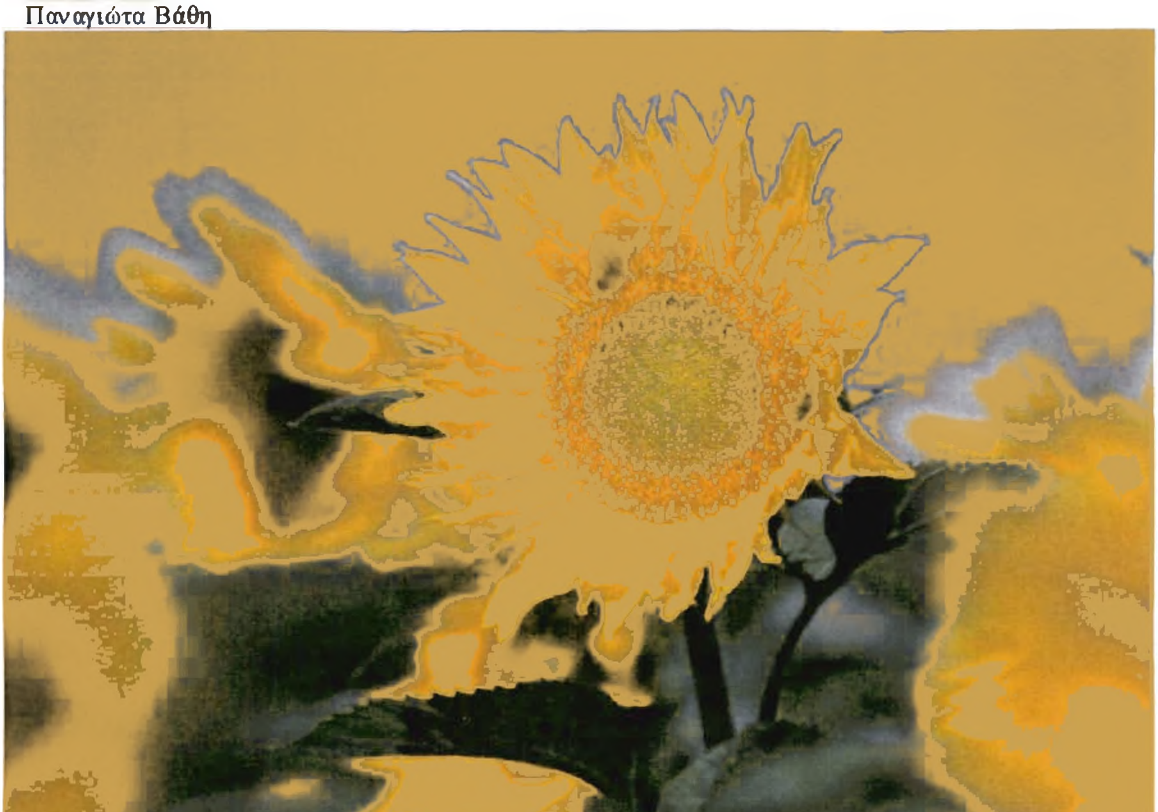
ΑΝΩΤΑΤΟ ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ ΜΕΣΟΛΟΓΓΙΟΥ
ΣΧΟΛΗ ΔΙΟΙΚΗΣΗΣ ΚΑΙ ΟΙΚΟΝΟΜΙΑΣ
ΤΜΗΜΑ ΔΙΟΙΚΗΣΗΣ ΚΟΙΝΩΝΙΚΩΝ - ΣΥΝΕΤΑΙΡΙΣΤΙΚΩΝ ΕΠΙΧΕΙΡΗΣΕΩΝ ΚΑΙ ΟΡΓΑΝΩΣΕΩΝ

Πτυχιακή εργασία της Αικατερίνης Σπυρίδων ΑΜ:8661

Θέμα:

Εκτιμητική στατιστική και έλεγχοι υποθέσεων οικονομικών στοιχείων της Τράπεζας Κύπρου της πενταετίας 2002 έως 2006

<u>Εισηγήτρια:</u>
Παναγιώτα Βάθη



Μεσολόγγι 2010

Τις θερμές μου ευχαριστίες στην καθηγήτρια μου κ.Βάθη Παναγιώτα για την πολύτιμη βοήθεια και στήριξη που μου παρείχε.

ΕΙΣΑΓΩΓΗ

Στα πλαίσια της παρούσας εργασίας, πραγματοποιείται η μελέτη στατιστικών στοιχειών με τις μεθόδους της στατιστικής συμπερασματολογίας και συγκεκριμένα την εκτιμητική και τους ελέγχους υποθέσεων.

Βασικό αντικείμενο διαπραγμάτευσης αποτελούν πραγματικά στατιστικά δεδομένα,που αφορούν τη δραστηριότητα του καταστήματος της Τράπεζας Κύπρου στα Ιωάννινα.

Η δομή της εν λόγω εργασίας αποτελείται από την εισαγωγή.τα δυο κεφάλαια,το συμπέρασμα και την βιβλιογραφία.

Στο πρώτο κεφάλαιο γίνεται αναφορά στην απαιτούμενη στάθμη γνώσεων,παρουσιάζοντας εννοιολογικά και διεξοδικά το υπόβαθρο της εκτιμητικής θεωρίας με έμφαση σε εφαρμογές δειγματοληπτικών κατανομών.Στη συνέχεια παρουσιάζεται εκτενώς το αντικείμενο των ελέγχων υποθέσεων,καθώς παρατίθενται επίσης εφαρμογές όπως οι έλεγχοι μέσων τιμών,διακυμάνσεων κ.τ.λ.

Στο δεύτερο κεφάλαιο περιγράφεται η επιλεγμένη επιχείρηση(Τράπεζα Κύπρου),δίνοντας τις απαραίτητες πληροφορίες που έχουν να κάνουν με την ιστορική διαδρομή,το αντικείμενο και το κύρος των δραστηριοτήτων της.

Σημαντικό ενδιαφέρον παρουσιάζει η ανάλυση των στατιστικών στοιχείων που παραχωρήθηκαν από την Τράπεζα και αντιστοιχούν στην χρονική περίοδο 2002-2006. Με αφορμή την οικονομική κρίση διερευνάται κατά πόσο επηρεάστηκε το συγκεκριμένο κατάστημα και κατ'επέκταση η περιοχή της Ηπείρου από αυτή.

Η μεθοδολογία που ακολουθήθηκε είναι η συγκέντρωση πληροφοριών με κατ'ιδίαν συνεντεύξεις-συναντήσεις με στελέχη της τράπεζας.και συγκέντρωση στατιστικών δεδομένων μέσω της βιβλιογραφίας και διαδικτυακών τόπων.

Θα πρέπει να αναφερθεί η δυσχέρεια στην συλλογή των στοιχείων από την τράπεζα,η οποία προκύπτει από το κλίμα υψηλού ανταγωνισμού,την διοικητική διάρθρωση του τραπεζικού συστήματος και του τρόπου αντίληψης των τραπεζικών στελεχών όσον αφορά την προώθηση των προϊόντων της τράπεζας.

1 ΚΕΦΑΛΑΙΟ

Γενικά

Η Στατιστική ορίζεται ως ένας κλάδος των Μαθηματικών, ο οποίος ασχολείται με την συλλογή, οργάνωση και ερμηνεία αριθμητικών δεδομένων, με απώτερο σκοπό την εξαγωγή συμπερασμάτων. Τα δεδομένα αυτά μπορούν να προέρχονται είτε από το πλήρες σύνολο των στοιχείων ενός πληθυσμού είτε από ένα υποσύνολο του.

Δύο πολύ βασικοί όροι που χρησιμοποιούνται στην Στατιστική είναι ο **πληθυσμός** και το **δείγμα.** Με τον όρο πληθυσμός εννοείται ένα σύνολο υποκείμενων ή αντικειμένων ή μιας πειραματικής διαδικασίας. Με τον όρο δείγμα εννοείται ένα υποσύνολο του πληθυσμού.

Η εξαγωγή συμπερασμάτων για έναν πληθυσμό, μπορεί να γίνει από δεδομένα που προέρχονται από ολόκληρο τον πληθυσμό ή από ένα δείγμα του. Στις περισσότερες των περιπτώσεων το ενδιαφέρον εστιάζεται σε πληθυσμούς των οποίων το πλήθος είναι πολύ μεγάλο ή μη αριθμήσιμο, οπότε είναι συχνά αδύνατη η συλλογή δεδομένων από ολόκληρο τον πληθυσμό. Σε αυτές τις περιπτώσεις η μελέτη του πληθυσμού γίνεται με την συλλογή δεδομένων που προέρχονται από το δείγμα του.

Το δείγμα είναι το μέσο από το οποίο μπορούν να εξαχθούν συμπεράσματα για ολόκληρο τον πληθυσμό. Οι ιδιότητες των στοιχείων ενός πληθυσμού ονομάζονται **χαρακτηριστικά** π.χ. η ηλικία, το βάρος, το φύλλο, το εισόδημα κ.τ.λ.

Το σύνολο των τιμών ενός χαρακτηριστικού, μετρούμενο σε όλα τα στοιχεία ενός πληθυσμού, ορίζει μια **τυχαία μεταβλητή.**

Παρατέθηκαν κάποιες έννοιες που αφορούν την γενική σημασιολογία της Στατιστικής ώστε να είναι πιο κατανοητοί οι δύο κλάδοι της Στατιστικής Συμπερασματολογίας , η εκτιμητική και οι έλεγχοι υποθέσεων που θα αναλύσουμε σε αυτό το κεφάλαιο.

1.1)Η ***εκτιμητική*** είναι η μία από τις δύο γενικές περιοχές ενδιαφέροντος της στατιστικής συμπερασματολογίας. Η μεγάλη περιοχή είναι οι *έλεγχοι υποθέσεων*. Αντικείμενο της εκτιμητικής είναι ο προσδιορισμός .των παραμέτρων ενός πληθυσμού, χρησιμοποιώντας ως βάση αναφοράς κατάλληλες στατιστικές συναρτήσεις.

Οι μέθοδοι της εκτιμητικής διακρίνονται σε αυτές που αφορούν τη *σημειακή εκτίμηση (point estimation)* μιας πληθυσμιακής παραμέτρου και σε αυτές που αφορούν την *εκτίμηση διαστήματος (interval estimation).* Στις σημειακές εκτιμήσεις χρησιμοποιούνται κατάλληλες στατιστικές συναρτήσεις προκειμένου να προσδιοριστεί ένας απλός αριθμός ο οποίος θα εκτιμά κάθε φορά μια πληθυσμιακή παράμετρο. Για παράδειγμα, μια δειγματική μέση τιμή $\bar{x}$ μπορεί να χρησιμοποιηθεί προκειμένου να εκτιμηθεί η αντίστοιχη πληθυσμιακή μέση τιμή μ. Το πρόβλημα σε μια σημειακή εκτίμηση του είδους είναι ότι διαφορετικά δείγματα του ιδίου πληθυσμού έχουν συνήθως διαφορετικές μέσες τιμές και, επομένως, κατά τον προσδιορισμό της αντίστοιχης πληθυσμιακής μέσης τιμής ενυπάρχει το στοιχείο της αβεβαιότητας.

Γενικά οι σημειακές εκτιμήσεις δεν παρέχουν καμία πληροφορία ως προς τη μεταβλητότητα της εκτιμήτριας συνάρτησης. Δεν είναι γνωστό, δηλαδή, πόσο κοντά μπορεί να βρίσκεται μια δειγματική μέση τιμή x στην αντίστοιχη πληθυσμιακή μέση τιμή μ που εκτιμά. Ενώ, για παράδειγμα, μια δειγματική μέση τιμή όσο μεγαλύτερο είναι το μέγεθος του δείγματος, μια σημειακή εκτίμηση δεν λαμβάνει υπόψη της το μέγεθος του δείγματος. Το κενό σε αυτήν την περάτωση καλύπτεται από την εκτίμηση διαστήματος της παραμέτρου!! Η εκτίμηση διαστήματος/παρέχει, ένα διάστημα δυνατών τιμών, εντός του οποίου μπορεί να περιλαμβάνεται η αντίστοιχη πληθυσμιακή παράμετρος στην προκειμένη περίπτωση η μέση τιμή μ- με ένα συγκεκριμένο βαθμό βεβαιότητας (ή αντίστροφα αβεβαιότητας). Ένα διάστημα τιμών του είδους ονομάζεται *διάστημα εμπιστοσύνης (confidence interval).*

Σημειακή εκτίμηση πληθυσμιακών παραμέτρων

Η εκτίμηση των πληθυσμιακών παραμέτρων βασίζεται, όπως ήδη αναφέρθηκε, στη χρήση κατάλληλων στατιστικών συναρτήσεων. Η εκλογή της κατάλληλης στατιστικής συνάρτησης εξαρτάται κάθε φορά από την πληθυσμιακή παράμετρο που πρόκειται να εκτιμηθεί. *Εκτιμήτρια συνάρτηση μιας πληθυσμιακής παράμετρος είναι μια στατιστική συνάρτηση, της οποίας οι τιμές παρέχουν προσεγγίσεις της πληθυσμιακής παραμέτρου. Μία συγκεκριμένη τιμή της συνάρτησης αυτής αποτελεί σημειακή εκτίμηση της παραμέτρου. Για παράδειγμα* η στατιστική συνάρτηση $\overline{X} = \frac{\sum x_i}{n}$,η οποία ορίζεται σε όλα τα δείγματα μεγέθους n ενός πληθυσμού, είναι εκτιμήτρια συνάρτηση της αντίστοιχης πληθυσμιακής μέσης τιμής μ. Μια τιμή αυτής της συνάρτησης, η οποία προκύπτει από ένα δείγμα μεγέθους n του πληθυσμού, αποτελεί σημειακή εκτίμηση της πληθυσμιακής μέσης τιμής.

Γενικότερα, έστω θ μια πληθυσμιακή παράμετρος την οποία θέλουμε να εκτιμήσουμε και $\hat{\theta}$ η αντίστοιχη εκτιμήτρια συνάρτηση. Όπως έχουμε δει στα προηγούμενα, η δειγματοληπτική κατανομής της εκτιμήτριας συνάρτησης μπορεί να είναι γνωστή, όπως συμβαίνει στην περίπτωση μιας μέσης τιμής ή μιας αναλογίας, ή στις περιπτώσεις της διαφοράς δυο μέσων τιμών ή δύο αναλογιών. *Η εκτιμήτρια συνάρτηση* $\hat{\theta}$ *θα ονομάζεται αμερόληπτη εκτιμήτρια (unbiased estimator) της παραμέτρου* θ, *αν η μέση τιμή της δειγματοληπτικής κατανομής της ισούται με* θ, *δηλαδή, αν* $E(\hat{\theta}) = \theta$. Στην περίπτωση αυτήν λέμε ότι η αντίστοιχη σημειακή εκτίμηση προκύπτει διαμέσου μιας *αμερόληπτης εκτιμητικής διαδικασίας. Στο* Σχήμα 1 εμφανίζονται οι δειγματοληπτικές κατανομές δύο εκτιμητριών της ίδιας πληθυσμιακής παραμέτρου θ, από τις οποίες η μία είναι αμερόληπτη εκτιμήτρια της θ, ενώ η άλλη δεν είναι.

Γνωρίζουμε ήδη από τα προηγούμενα ότι, για τις δειγματοληπτικές κατανομές της δειγματικής μέσης τιμής μιας δειγματικής αναλογίας και των διαφορών δύο μέσων τιμών ή δύο αναλογιών ισχύει

$$E(\overline{X}) = \mu_{\overline{X}} = \mu, \qquad E(\hat{p}) = \mu_{\hat{p}} = p$$

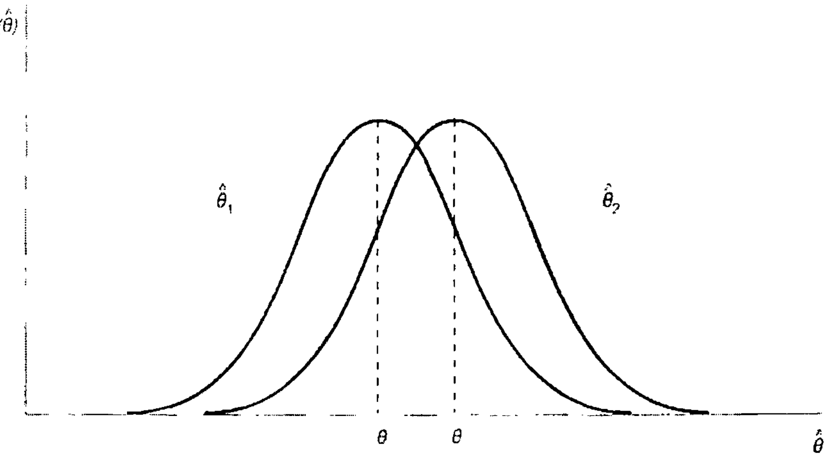
και

$$E(\overline{X}_1 - \overline{X}_2) = \mu_{\overline{x}_1 - \overline{x}_2} = \mu_1 - \mu_2, \qquad E(\hat{p}_1 - \hat{p}_2) = \mu_{\hat{p}_1 - \hat{p}_2} = p_1 - p_2$$

Επομένως, μπορούμε να ισχυριστούμε ότι η δειγματική μέση τιμή, μια δειγματική αναλογία και οι διαφορές δύο δειγματικών μέσων τιμών ή δύο δειγματικών αναλογιών, είναι αμερόληπτες εκτιμήτριες των αντίστοιχων

Σχήμα 1

Δειγματοληπτικές κατανομές δυο εκτιμητριών της παραμέτρου *θ*, από τις οποίες η $\hat{\theta}_1$ είναι



αμερόληπτη εκτιμήτρια της *θ*, ενώ η $\hat{\theta}_2$ δεν είναι πληθυσμιακών παραμέτρων. Αντίστοιχη πρόταση μπορεί να διατυπωθεί, όπως θα δούμε και στη συνέχεια, και για τη δειγματική διακύμανση. Ισχύει, δηλαδή, $E(S^2) = \sigma^2$, όπου $S^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$ η στατιστική συνάρτηση που ορίζεται σε όλα τα δείγματα μεγέθους *η* ενός πληθυσμού και η οποία συμβολίζει τη δειγματική διακύμανση.

Μια εκτιμήτρια η οποία δεν είναι αμερόληπτη λέγεται μεροληπτική (biased). Το μέγεθος της μεροληψίας (bias) σε μια τέτοια περίπτωση ισούται με τη διαφορά της μέσης τιμής της δειγματοληπτικής κατανομής της εκτιμήτριας συνάρτησης και της αντίστοιχης πληθυσμιακής παραμέτρου που αυτή εκτιμά, δηλαδή με E(θ) - θ. Στο Σχήμα 1, η διαφορά $\theta' - \theta$ είναι η μεροληψία της $\hat{\theta}_2$ κατά την εκτίμηση της *θ*.

Σε πολλές περιπτώσεις είναι δυνατόν να υπάρχουν περισσότερες της μίας αμερόληπτες εκτιμήτριες της ίδιας πληθυσμιακής παραμέτρου. Σε αυτές τις περιπτώσεις είναι λογικό να προτιμάται εκείνη η εκτιμήτρια, της όποιας η δειγματοληπτική κατανομή έχει τη μικρότερη διασπορά γύρω α-από την πληθυσμιακή παράμετρο.

Ειδικότερα, αν $\hat{\theta}_1$ *και* $\hat{\theta}_2$ *είναι δύο αμερόληπτες εκτιμήτριες της πληθυσμιακής παραμέτρου θ, οι οποίες προκύπτουν από δείγματα του ιδίου μεγέθους, τότε:*

i. *η εκτιμήτρια* $\hat{\theta}_1$ *λέγεται ότι είναι πιο αποτελεσματική (efficient) από την* $\hat{\theta}_2$, αν

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

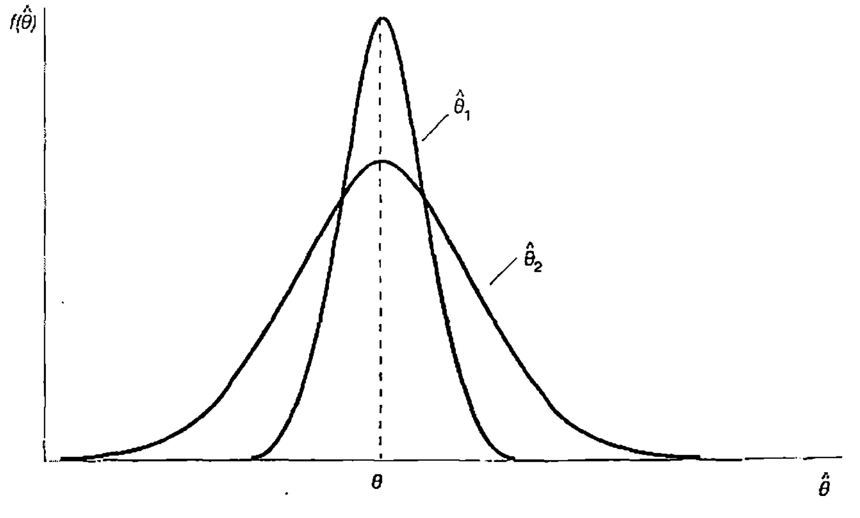
ii. *η σχετική αποτελεσματικότητα (relative efficiency) μιας αμερόληπτης -εκτιμήτριας $\hat{\theta}_1$ σε σχέση με μία άλλη $\hat{\theta}_2$ ορίζεται από το λόγο των διακυμάνσεών τους, δηλαδή το λόγο* $\dfrac{\mathrm{Var}(\hat{\theta}_2)}{\mathrm{Var}(\hat{\theta}_1)}$

Σε περιπτώσεις πολλών αμερόληπτων εκτιμητριών της ιδίας παραμέτρου, η εκτιμήτρια με τη μικρότερη διακύμανση ονομάζεται *πιο αποτελεσματική (most efficient)* ή *αμερόληπτη εκτιμήτρια ελάχιστης διακύμανσης (minimum variance unbiased estimator).* Στο Σχήμα εμφανίζονται οι δειγματοληπτικές κατανομές των αμερόληπτων εκτιμητριών $\hat{\theta}_2$ και $\hat{\theta}_2$ της παραμέτρου *θ,* από τις οποίες η $\hat{\theta}_1$ είναι πιο αποτελεσματική από τη $\hat{\theta}_2$.

Το γενικότερο πρόβλημά της επιλογής της κατάλληλης εκτιμήτριας για τη σημειακή εκτίμηση μιας πληθυσμιακής παραμέτρου, είναι συχνά αρκετά σύνθετο και η επίλυση του ανάγεται σε αναλυτικές ή πιθανό θεωρητικές μεθόδους. Μεταξύ των αναλυτικών μεθόδων περιλαμβάνεται *η μέθοδος των ροπών της μέγιστης πιθανοφάνειας* και *η μέθοδος των ελαχίστων τετραγώνων.* Στις πιθανοθεωρητικές μεθόδους, η επιλογή μιας αμερόληπτης εκτιμήτριας βασίζεται στη «βέλτιστη ικανοποίηση» ορισμένων ιδιοτήτων, τις οποίες απαιτείται να έχει η εκτιμήτρια συνάρτηση, όπως *της αμεροληψίας, της συνέπειας (consistency), της ελάχιστης διακύμανσης* κ.ά.

Σχήμα

Δειγματοληπτικές κατανομές δυο αμερόληπτων εκτιμητριών της παραμέτρου θ, από τις οποίες η $\hat{\theta}_1$ είναι πιο αποτελεσματική από τη $\hat{\theta}_2$



Παράδειγμα

Έστω X η τυχαία μεταβλητή των μέσων τιμών όλων των δυνατών δειγμάτων μεγέθους *η*, που μπορούν να ληφθούν από έναν πληθυσμό με μέση τιμή μ και διακύμανση σ^2. Η τυχαία μεταβλητή X, όπως ήδη αναφέρθηκε, είναι αμερόληπτη εκτιμήτρια της πληθυσμιακής μέσης τιμής, με διακύμανση

$$\mathrm{Var}\left(\overline{\mathrm{X}}\right) = \sigma^2_{\bar{x}} = \frac{\sigma^2}{n}$$

Ως εναλλακτική εκτιμήτρια της πληθυσμιακής μέσης τιμής μπορεί να χρησιμοποιηθεί η τυχαία μεταβλητή των διαμέσων όλων των δυνατών δειγμάτων του ιδίου μεγέθους *η*. Αποδεικνύεται ότι η τυχαία αυτή μεταβλητή είναι επίσης αμερόληπτη εκτιμήτρια της παραμέτρου μ και ότι η διακύμανση της ισούται με

$$Var(\Delta\iota\alpha\mu\acute{\varepsilon}\sigma o\upsilon) \quad \frac{\pi}{2}\left(\frac{\sigma^2}{n}\right) \approx \left(\frac{\sigma^2}{n}\right)$$

Επομένως, η δειγματική μέση τιμή είναι πιο αποτελεσματική εκτιμήτρια από τη δειγματική διάμεσο, ενώ η σχετική αποτελεσματικότητα της σε σχέση με τη διάμεσο είναι ίση με

$$\frac{Var(\Delta\iota\alpha\mu\acute{\varepsilon}\sigma o\upsilon)}{Var\left(\overline{X}\right)} = 1{,}57.$$

Δηλαδή, η διακύμανση της δειγματοληπτικής κατανομής της διαμέσου είναι κατά 57% μεγαλύτερη από αυτήν της μέσης τιμής. Αυτή η σχετική αναποτελεσματικότητα της διαμέσου ως εκτιμήτριας σε σχέση με τη μέση τιμή είναι και ο βασικός λόγος για τον οποίο η μέση τιμή προτιμάται κατά κανόνα έναντι της διαμέσου, στις εκτιμητικές διαδικασίες της στατιστικής συμπερασματολογίας.

Διαστήματα εμπιστοσύνης πληθυσμιακών μεσών τιμών

Οι εκτιμήσεις διαστήματος των πληθυσμιακών παραμέτρων δίδονται ν διά μέσου των διαστημάτων εμπιστοσύνης. Σε αντίθεση με τις σημειακές εκτιμήσεις, όπου οι πληθυσμιακοί παράμετροι εκτιμώνται με μία και μόνο αριθμητική τιμή, τα διαστήματα εμπιστοσύνης είναι σύνολα τιμών, στο εσωτερικό των οποίων περιλαμβάνεται με μια συγκεκριμένη πιθανότητα κάθε φορά, η ζητούμενη πληθυσμιακή παράμετρος. Για να κατασκευαστεί ένα διάστημα εμπιστοσύνης μιας πληθυσμιακής παραμέτρου, είναι απαραίτητο να είναι γνωστή η δειγματοληπτική κατανομή της στατιστικής συνάρτησης η οποία εκτιμά αμερόληπτα την παράμετρο. Έτσι, για έναν πληθυσμό με μέση τιμή μ και τυπική απόκλιση σ, είναι γνωστό από το κεντρικό οριακό θεώρημα ότι η κατανομή της τυχαίας μεταβλητής

$Z = \dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}}$ όπου $\overline{X}$ η δειγματική μέση τιμή,

είναι τυπική κανονική, αν ο πληθυσμός ακολουθεί

και είναι κατά προσέγγιση τυπική κανονική, αν ο πληθυσμός δεν ακολουθεί την κανονική κατανομή, αλλά το n είναι επαρκώς μεγάλο. Για την τυπική κανονική κατανομή είναι γνωστό ότι 95% των τιμών της περιλαμβάνεται στο διάστημα από $-1{,}96$ έως $1{,}96$. Δηλαδή η πιθανότητα η τυχαία μεταβλητή Z να παίρνει τιμές μεταξύ -1,96 και 1,96 είναι

$P(-1{,}96 \leq Z \leq 1{,}96) = 0{,}95$

ή ισοδύναμα

$$P\left(-1{,}96 \leq \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \leq 1{,}96\right) = 0{,}95.$$

Από την προηγούμενη σχέση, πολλαπλασιάζοντας και τους τρεις όρους της α-νισότητας με το τυπικό σφάλμα $\frac{\sigma}{\sqrt{n}}$ προκύπτει

$$P\left(-1{,}96 \le \frac{\sigma}{\sqrt{n}} \le \overline{X}-\mu \le 1{,}96\frac{\sigma}{\sqrt{n}}\right)=0{,}95$$

ή

$$P\left(-1{,}96 \le \frac{\sigma}{\sqrt{n}}-\overline{X} \le -\mu \le 1{,}96\frac{\sigma}{\sqrt{n}}-\overline{X}\right)=0{,}95.$$

Πολλαπλασιάζοντας τους όρους της τελευταίας ανισότητας με -1, προκύπτει

$$P\left(1{,}96\frac{\sigma}{\sqrt{n}}+\overline{X} \ge \mu \ge -1{,}96\frac{\sigma}{\sqrt{n}}+\overline{X}\right)=0{,}95$$

ή αλλιώς

$$P\left(\overline{X}-1{,}96\frac{\sigma}{\sqrt{n}} \le \mu \le \overline{X}+1{,}96\frac{\sigma}{\sqrt{n}}\right)=0{,}95.$$

Στην παραπάνω σχέση, οι ποσότητες $\overline{X}-1{,}96\frac{\sigma}{\sqrt{n}}$ και $\overline{X}+1{,}96\frac{\sigma}{\sqrt{n}}$ ονομάζονται 95% *όρια εμπιστοσύνης της πληθυσμιακής μέσης τιμής*, ενώ το διάστημα $\left(\overline{X}-1{,}96\frac{\sigma}{\sqrt{n}},\ \overline{X}+1{,}96\frac{\sigma}{\sqrt{n}}\right)$ ονομάζεται 95% *διάστημα εμπιστοσύνης της πληθυσμιακής μέσης τιμής*. Ένα διάστημα του είδους ουσιαστικά δηλώνει ότι, αν επιλέξουμε 100 τυχαία δείγματα μεγέθους *n* από έναν πληθυσμό και υπολογίσουμε για κάθε ένα από αυτά το 95% διάστημα εμπιστοσύνης, αναμένουμε στα 95 από τα 100 διαστήματα να περιλαμβάνεται η πραγματική πληθυσμιακή μέση τιμή.

Αν και τα 95% διαστήματα εμπιστοσύνης χρησιμοποιούνται συχνότερα στην πράξη, εντούτοις δεν περιοριζόμαστε αποκλειστικά σε αυτά, όταν πρόκειται να δώσουμε μια εκτίμηση διαστήματος για τη μέση τιμή μ ενός | πληθυσμού. Σε περιπτώσεις που επιθυμούμε να έχουμε μεγαλύτερο βαθμό βεβαιότητας στην εκτίμηση της πληθυσμιακής μέσης τιμής, μπορούμε να κατασκευάσουμε ένα 99% διάστημα εμπιστοσύνης αντί του 95% διαστήματος.

Εφόσον 99% των τιμών της τυπικής κανονικής κατανομής περιλαμβάνεται στο διάστημα -2,58 έως 2,58, το 99% διάστημα εμπιστοσύνης για την πληθυσμιακή μέση τιμή μ δίδεται από τον τύπο

$$\left(\overline{X} - 2{,}58\frac{\sigma}{\sqrt{n}},\ \overline{X} + 2{,}58\frac{\sigma}{\sqrt{n}}\right).$$

Δηλαδή, 99 στα 100 διαστήματα του είδους, υπολογιζόμενα σε 100 ανεξάρτητα τυχαία δείγματα μεγέθους η του πληθυσμού, αναμένεται να περιλαμβάνουν την πληθυσμιακή μέση τιμή μ. Όπως ήταν αναμενόμενο, ένα 9,9% διάστημα εμπιστοσύνης είναι μεγαλύτερης έκτασης από ένα 95% διάστημα. Όσο δηλαδή το εύρος των τιμών ενός διαστήματος εμπιστοσύνης περιορίζεται τόσο μικρότερος είναι και ο βαθμός βεβαιότητας που δίνει το διάστημα στο ενδεχόμενο να περιλαμβάνει τη μέση τιμή μ.

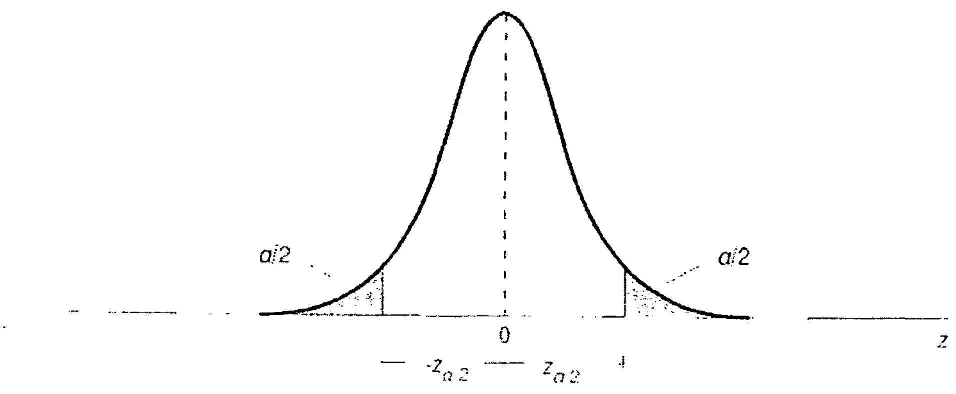
Ένα γενικευμένο διάστημα εμπιστοσύνης για την πληθυσμιακή μέση τιμή μπορεί να ληφθεί με τη βοήθεια των τιμών της τυπικής κανονικής κατανομής. Έστω $z_{\alpha/2}$ η τιμή της τυπικής κανονικής κατανομής που ορίζει μια περιοχή με εμβαδόν α/2 στο δεξιό άκρο της κατανομής και $-z_{\alpha/2}$ η τιμή που ορίζει την αντίστοιχη περιοχή με εμβαδόν α/2 στον αριστερό άκρο της τυπικής κανονικής κατανομής. Το διάστημα

$$\left(\overline{X} - z_{a/2}\frac{\sigma}{\sqrt{n}},\ \overline{X} + z_{a/2}\frac{\sigma}{\sqrt{n}}\right),$$

είναι ένα *100(1 - α)% διάστημα εμπιστοσύνης*. Η ποσότητα 1 - α ονομάζεται *συντελεστής εμπιστοσύνης* του διαστήματος. Ένα διάστημα του είδους έχει 100(1 - α) πιθανότητα να περιλαμβάνει την πληθυσμιακή μέση τιμή μ. Αν π.χ. α = 0,05, τότε $\alpha/2 = 0{,}025$ ενώ $z_{0,025} = 1{,}96$ και $-z_{0,025} = -1{,}96$.

Σχήμα

Περιοχές της τυπικής κανονικής κατανομής που ορίζονται από τις τιμές $z_{\alpha/2}$ και $-z_{\alpha/2}$



Δηλαδή, για $\alpha = 0{,}05$, το διάστημα που ορίζεται είναι το γνωστό 95% διάστημα εμπιστοσύνης$\left(\overline{X} - 1{,}96\frac{\sigma}{\sqrt{n}},\ \overline{X} + 1{,}96\frac{\sigma}{\sqrt{n}}\right)$,

Αν επιθυμούμε να υπολογίσουμε ένα πιο «σφιχτό» διάστημα χωρίς να . μειώσουμε το επίπεδο εμπιστοσύνης, χρειαζόμαστε περισσότερες πληροφορίες για την πληθυσμιακή μέση τιμή, δηλαδή χρειαζόμαστε ένα δείγμα μεγαλύτερου μεγέθους. Όσο το μέγεθος *η* του δείγματος αυξάνει, το *τυπικό* σφάλμα $\frac{\sigma}{\sqrt{n}}$ ελαττώνεται, γεγονός που οδηγεί στον προσδιορισμό περισσότερο «συμπαγών», γύρω από την πληθυσμιακή μέση τιμή, διαστημάτων εμπιστοσύνης. Αν π.χ. θεωρήσουμε το 95% διάστημα εμπιστοσύνης

$$\left(\overline{X} - 1{,}96\frac{\sigma}{\sqrt{n}},\ \overline{X} + 1{,}96\frac{\sigma}{\sqrt{n}}\right),$$

τότε για δείγμα μεγέθους *η* — 10 τα αντίστοιχα όρια εμπιστοσύνης είναι $\overline{X} \pm 1{,}96\left(\frac{\sigma}{\sqrt{10}}\right)$.

Αν το μέγεθος του δείγματος είναι 100, τα όρια εμπιστοσύνης είναι

$$\overline{X} \pm 1{,}96\left(\frac{\sigma}{\sqrt{100}}\right).$$

ενώ όταν το μέγεθος του δείγματος είναι 1000, τα όρια εμπιστοσύνης είναι

$$\overline{X} \pm 1{,}96\left(\frac{\sigma}{\sqrt{1000}}\right).$$

Συνοψίζοντας τους υπολογισμούς των παραπάνω ορίων εμπιστοσύνης προκύπτει ο ακόλουθος πίνακας:

n	95% όρια εμπιστοσύνης για τη μέση τιμή μ	Εύρος διαστήματος
10	$\overline{X} \pm 0{,}620\sigma$	1,240 σ
100	$\overline{X} \pm 0{,}193\sigma$	0.392σ
100	$\overline{X} \pm 0{,}062\sigma$	0,124σ

Όσο αυξάνει, δηλαδή, το μέγεθος των δειγμάτων που λαμβάνονται από τον πληθυσμό, η μεταβλητότητα της τυχαίας μεταβλητής της εκτιμήτριας συνάρτησης δηλαδή του πληθυσμιακού μέσου *μ*- γίνεται μικρότερη. Η ενυπάρχουσα μεταβλητότητα της πληθυσμιακής κατανομής, όμως, η οποία προσδιορίζεται από την πληθυσμιακή τυπική απόκλιση *σ,* εξακολουθεί να επηρεάζει το εύρος των εκτιμήσεων που παρέχουν τα διαστήματα εμπιστοσύνης.

Παράδειγμα

Μια εταιρεία κατασκευής μπίρας γνωρίζει από τον έλεγχο της παραγωγής ότι οι ποσότητες της μπίρας στα κουτιά ακολουθούν την κανονική κατανομή με τυπική απόκλιση σ =12 gr. Σε μια διαδικασία ποιοτικού ελέγχου, επιλέχθηκαν τυχαία 100 κουτιά μπίρας από την παραγωγή ενός μήνα και βρέθηκε ότι το μέσο βάρος του περιεχομένου τους είναι 250 gr. Να υπολογιστεί το 95% διάστημα εμπιστοσύνης για την πραγματική μέση τιμή του βάρους της ποσότητας μπίρας που περιέχεται στα κουτιά.

Το 95% διάστημα εμπιστοσύνης για την πληθυσμιακή μέση τιμή μ είναι το

$$\left(\overline{\mathrm{X}} - 1{,}96\frac{\sigma}{\sqrt{n}},\ \overline{\mathrm{X}} + 1{,}96\frac{\sigma}{\sqrt{n}}\right).$$

Επομένως, για $\overline{\mathrm{X}} = 250$, έχουμε

$$\left(250 - 1{,}96\frac{12}{\sqrt{100}} 2{,}50 + 1{,}96\frac{12}{\sqrt{100}}\right) = (247{,}6,\quad 252{,}4).$$

Άρα, η πραγματική μέση τιμή του βάρους της μπίρας στα κουτιά κυμαίνεται από 247,6 έως 252,4 γραμμάρια με βεβαιότητα 95%.

Μονόπλευρα διαστήματα εμπιστοσύνης

Σε ορισμένες περιπτώσεις, το ενδιαφέρον περιορίζεται στον προσδιορισμό μόνο ενός ανώτερου ή μόνο ενός κατώτερου ορίου (όχι όμως και των δύο) για την πληθυσμιακή μέση τιμή. Ένα ανώτερο ή ένα κατώτερο όριο του είδους ορίζει ένα διάστημα τιμών προς μία κατεύθυνση, στο οποίο με μια συγκεκριμένη βεβαιότητα κάθε φορά θα περιλαμβάνεται η πληθυσμιακή μέση τιμή. Ένα διάστημα τιμών οριζόμενο μόνο ως προς μία κατεύθυνση ονομάζεται *μονόπλευρο διάστημα εμπιστοσύνης (one-sided confidence interval).* Για την κατασκευή ενός μονόπλευρου διαστήματος εμπιστοσύνης για την πληθυσμιακή μέση τιμή, χρησιμοποιούνται τα διαστήματα τιμών της τυπικής κανονικής κατανομή!

Έστω ότι οι αποδοχές των υπαλλήλων μιας μεγάλης επιχείρησης ακολουθούν την κανονική κατανομή με τυπική απόκλιση σ =12000 δρχ. Για την εκτίμηση του μέσου μισθού των υπαλλήλων της παραπάνω επιχείρησης παίρνουμε ένα δείγμα 25 υπαλλήλων. Από το δείγμα αυτό προκύπτει $\bar{x}$ = 81.000 δρχ.

Ποια είναι τα όρια εμπιστοσύνης μέσα στα οποία περιμένουμε με πιθανότητα 95% ότι θα βρίσκεται ο μέσος μισθός μ του συνόλου των εργαζομένων στην επιχείρηση.

Λύση

Έχουμε ότι σ = 12000, n = 25, $\bar{x}$ = 81000, 1 – α =0,95 και α =0,05.

Από τον πίνακα της κανονικής τυπικής κατανομής για επίπεδο σημαντικότητας α =5% βρίσκουμε ότι $z_{\alpha/2}$ = 1,96. Το τυπικό σφάλμα εκτίμησης του μέσου θα είναι:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{12000}{5} = 2400$$

Επομένως, ο ζητούμενος μέσος μισθός μ του συνόλου των εργαζομένων θα περιλαμβάνεται στο διάστημα:

$$\bar{x} = z_{a/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{a/2}\frac{\sigma}{\sqrt{n}} \text{ ή}$$

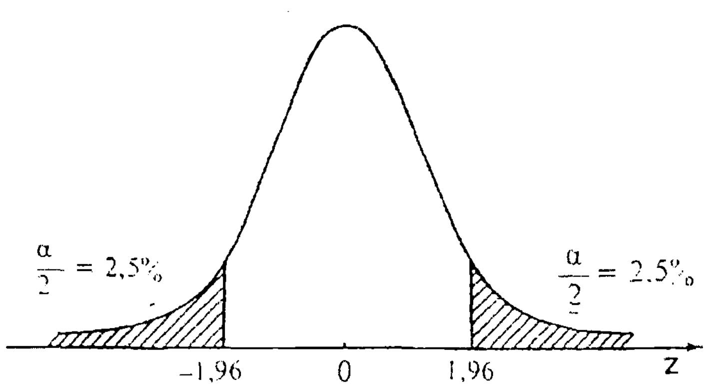
$$81000 - 1,96 \times 2400 < \mu < 81000 + 1,96 \times 2400 \text{ ή}$$

$$76296 < \mu < 85704$$

Επομένως, το 95% διάστημα εμπιστοσύνης είναι 76296 < μ < 85704 και δεν πρέπει να γράφουμε:

$$P\{76296 < \mu < 85704\} = 0,95$$

γιατί τα όρια εμπιστοσύνης δεν είναι πλέον τυχαίες μεταβλητές.



$P\{z < -z_{\alpha/2} = -1,96\} = 0,025$

Έστω ότι από έναν κανονικό πληθυσμό, του οποίου δεν γνωρίζουμε τη διακύμανση (σ^2), παίρνουμε ένα δείγμα μεγέθους 14 παρατηρήσεων (n = 14). Από το δείγμα αυτό βρέθηκε: $\bar{x}$ = 52,52 και S = 3,37.

Ζητείται το 95% διάστημα εμπιστοσύνης για το μέσο αριθμό μ του πληθυσμού.

Λύση

Στην παραπάνω περίπτωση η διακύμανση σσ^2 είναι άγνωστη και ταυτόχρονα το n = 14 < 30, επομένως θα χρησιμοποιήσουμε την κατανομή t του Student. Από τους πίνακες της κατανομής t με βαθμούς ελευθερίας ν = n – 1 = 14 – 1 = 13 και επίπεδο σημαντικότητας α = 5%, βρίσκουμε ότι $t_{5\%,13}$ = 2,16 και επομένως το 95% διάστημα εμπιστοσύνης για το μέσο μ του πληθυσμού θα είναι:

$$\bar{x} - t_{a/2,v}\frac{S}{\sqrt{n}} < \mu < \bar{x} + t_{a/2,v}\frac{S}{\sqrt{n}} \Rightarrow$$

$$52{,}52 - 2{,}16\frac{3{,}37}{\sqrt{14}} < \mu < 52{,}52 + 2{,}16\frac{3{,}37}{\sqrt{14}} \text{ ή}$$

$$50{,}58 < \mu < 54{,}47.$$

Προκειμένου να βρεθεί η διαφορά του βάρους μεταξύ των σπουδαστών της Ανωτάτης Εμπορικής και του Πανεπιστημίου Πειραιώς, έχει ληφθεί από την πρώτη δείγμα n_1 = 1200 σπουδαστών και από τη δεύτερη n_2 = 250 σπουδαστών.

Από τα δείγματα αυτά προέκυψε:

$\bar{x}_1$ = 72 Kgr για τους Σπουδαστές της Εμπορικής

$\bar{x}_2$ = 70 kgr για τους σπουδαστές του Πανεπ. Πειραιώς

Αν θεωρηθεί γνωστό ότι η κατανομή των βαρών στις δύο Σχολές είναι κανονική και ότι σ_1^2 και σ_2^2 είναι γνωστές, δηλαδή σ_1^2 = 100 και σ_2^2 = 81, ζητείται το διάστημα εμπιστοσύνης της διαφοράς των δύο μέσων $\mu_1 - \mu_2$ σε επίπεδο σημαντικότητας 5%.

Λύση

Επειδή σ_1^2 και σ_2^2 είναι γνωστές, για τον προσδιορισμό του διαστήματος εμπιστοσύνης της διαφοράς των δύο μέσων, εφαρμόζουμε τη σχέση:

$$\left(\bar{x}_1 - \bar{x}_2\right) - Z_a\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \leq \mu_1 - \mu_2 \leq \left(\bar{x}_1 - \bar{x}_2\right) + Z_a\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Με αντικατάσταση παίρνουμε:

$$(72-70)-1{,}96\sqrt{\frac{100}{1200}+\frac{81}{250}} \le \mu_1-\mu_2 \le (72-70)+1{,}96\sqrt{\frac{100}{1200}+\frac{81}{250}}$$

$$\text{ή } 2-1{,}96\cdot 0{,}4073 \le \mu_1-\mu_2 \le 2+1{,}96\cdot 0{,}4073 \text{ ή}$$

$$1{,}20 \le \mu_1-\mu_2 \le 2{,}80.$$

Για να συγκρίνουμε δύο μηχανές που κατασκευάζουν νήματα, παίρνουμε ένα δείγμα n = 10 μονάδων από κάθε μηχανή, ελέγχουμε την αντοχή και βρίσκουμε τα παρακάτω στοιχεία.

Μηχανή Α: 380, 382, 362, 382, 364, 392, 360, 358, 400, 402

Μηχανή Β: 398, 440, 396, 428, 420, 446, 420, 420, 410, 390

Αν υποθέσουμε ότι η κατανομή της αντοχής του νήματος είναι κανονική με $\sigma_1^2=\sigma_2^2$, αλλά άγνωστες τις διακυμάνσεις, να βρεθεί το διάστημα εμπιστοσύνης της διαφοράς των δύο μέσων σε επίπεδο εμπιστοσύνης (1 – α) 100%.

Λύση

Όταν $\sigma_1^2=\sigma_2^2$ αλλά είναι άγνωστες και το μέγεθος του δείγματος είναι μικρό, δεν χρησιμοποιούμε την τυποποιημένη κανονική κατανομή, αλλά την κατανομή του Student με βαθμούς ελευθερίας $v = n_1 + n_2 - 2$, το δε διάστημα εμπιστοσύνης της διαφοράς $\mu_1 - \mu_2$ δίνεται από τη σχέση:

$$\left(\overline{X}_1-\overline{X}_2\right)-t_{v,\alpha/2}\cdot S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \le \mu_1-\mu_2 \le \left(\overline{X}_1-\overline{X}_2\right)-t_{v,\alpha/2}\cdot S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

Ένα ποσοστό ηλεκτρικών λαμπτήρων, που παράγονται από το εργοστάσιο Α είναι ελαττωματικό. Για την εκτίμηση του άγνωστου αυτού ποσοστού παίρνουμε ένα τυχαίο δείγμα από 100 λαμπτήρες και βρίσκουμε 12 ελαττωματικούς. Να βρεθεί το 95% διάστημα εμπιστοσύνης, μέσα στο οποίο θα βρίσκεται το αληθινό ποσοστό ελαττωματικών λαμπτήρων του συνολικού πληθυσμού.

Λύση

Το διάστημα εμπιστοσύνης θα είναι:

$$\hat{p} - z_{a/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} \leq P \leq \hat{p} + z_{a/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} = 0{,}95$$

$$0{,}12 - 1{,}96\sqrt{\frac{0{,}12 \cdot 0{,}88}{100}} \leq P \leq 0{,}12 + 1{,}96\sqrt{\frac{0{,}12 \cdot 0{,}88}{100}}$$

$$0{,}12 - 1{,}96 \cdot 0{,}0325 \leq \mathrm{P} \leq 0{,}12 + 1{,}96 \cdot 0{,}0325 \text{ ή}$$

$$0{,}0563 \leq \mathrm{P} \leq 0{,}1837$$

Προκειμένου να μελετήσουμε το μηνιαίο εισόδημα των κατοίκων μιας πόλης αποφασίζουμε να πάρουμε ένα τυχαίο δείγμα.

Ερωτάται: α) Ποιο μέγεθος δείγματος θα πρέπει να πάρουμε, ώστε το δειγματοληπτικό σφάλμα να μην υπερβαίνει τις ±1000 δραχμές δεδομένου ότι σ = 1800 δρχ., σε επίπεδο σημαντικότητας 5%.

β) Ποιο πρέπει να είναι το συνολικό κόστος της έρευνας, αν τα γενικά έξοδα (έστω ότι είναι σταθερά) ανέρχονται σε 5.000.000 δρχ. και τα έξοδα για τη συμπλήρωση κάθε ερωτηματολογίου, που αντιστοιχεί σε κάθε μονάδα δείγματος, ανέρχεται σε 500 δραχμές.

γ) Αν όμως διαθέτουμε για την έρευνα 6.000.000 δρχ. ποιο πρέπει να είναι το δειγματοληπτικό σφάλμα;

Λύση

α) Το μέγεθος του δείγματος θα είναι:

$$\mathrm{n} = z_{a/2}^2 \frac{\sigma^2}{\kappa^2} = 1{,}96\frac{1800}{1000^2} = 1244$$

β)Κ = Γ + κ · λ = 5.000.000 + 1244 · 500 = 5.622.000

Άρα, το συνολικό κόστος θα είναι 5.622.000 δρχ.

γ) Αν διαθέσουμε για την έρευνα 6.000.000 δρχ., θα έχουμε:

$K = \Gamma + \kappa \cdot \lambda \Rightarrow 6.000.000 = 5.000.000 + 500 \cdot n \Rightarrow n = 2000$. Το δειγματοληπτικό σφάλμα θα είναι:

$$= \pm z_{a/2}\frac{\sigma}{\sqrt{n}} \Rightarrow Z_{a/2}\frac{\sigma}{\sqrt{n}} = \pm 1{,}96\frac{1800}{\sqrt{2000}} = \pm 789 \text{ δρχ.}$$

Δοθέντος ότι το δείγμα έχει αυξηθεί, το δειγματοληπτικό σφάλμα έχει μειωθεί.

Εναλλακτικές εκτιμήτριες της κεντρικής τάσης μιας κατανομής

Όπως έχει ήδη αναφερθεί, η μέση τιμή -σε αντίθεση με τη διάμεσο-είναι «ευαίσθητη» στην ύπαρξη παράτυπων τιμών (outliers). Για το λόγο αυτόν, όταν σε μια κατανομή υπάρχουν παράτυπες τιμές, είναι δυνατόν να προτιμάται η διάμεσος ως μέτρο κεντρικής τάσης αντί της μέσης τιμής. Για τον ίδιο λόγο, όταν εκτιμάται η κεντρική τάση μιας πληθυσμιακής κατανομής, είναι δυνατόν να χρησιμοποιείται η δειγματική διάμεσος ως εκτιμήτρια της πληθυσμιακής διαμέσου.

Εκτιμήτριες συναρτήσεις, οι οποίες παραμένουν ανεπηρέαστες από την ύπαρξη παράτυπων τιμών, ονομάζονται *ανθεκτικές εκτιμήτριες (robust estimators).* Εκτός της διαμέσου, ανθεκτική εκτιμήτρια της πληθυσμιακής μέσης τιμής θεωρείται και η λεγόμενη *«ξακρισμένη» μέση τιμή (trimmed mean).* Μια μέση τιμή του είδους υπολογίζεται εφόσον αφαιρέσουμε από το σύνολο των τιμών της κατανομής ένα συγκεκριμένο ποσοστό των τιμών της από την αριστερή και από την δεξιά πλευρά της - συνήθως 5% από αριστερά και 5% από δεξιά. Η μέση τιμή που υπολογίζεται στο σύνολο των υπόλοιπων τιμών είναι η «ξακρισμένη» μέση τιμή.

Η κατανομή του Student

Κατά των υπολογισμό των διαστημάτων εμπιστοσύνης μιας πληθυσμιακής μέσης τιμής μ, θεωρήθηκε ότι η πληθυσμιακή τυπική απόκλιση σ ήταν γνωστή. Στην πραγματικότητα, όμως, είναι απίθανο, όταν η πληθυσμιακή μέση τιμή είναι άγνωστη, να είναι γνωστή η αντίστοιχη πληθυσμιακή τυπική απόκλιση. Σε αυτές τις περιπτώσεις τα διαστήματα εμπιστοσύνης που μπορούν να κατασκευαστούν είναι παρόμοια με αυτά που έχουμε μέχρι στιγμής ορίσει, μόνο που αντί της τυπικής κανονικής κατανομής χρησιμοποιείται μια άλλη γνωστή θεωρητική κατανομή, *η κατανομή του Student ή αλλιώς κατανομή t.* Το όνομα Student είναι ψευδώνυμο του W.S Gossett *(1876-1937),* ο οποίος πρώτος μελέτησε τη συγκεκριμένη κατανομή.

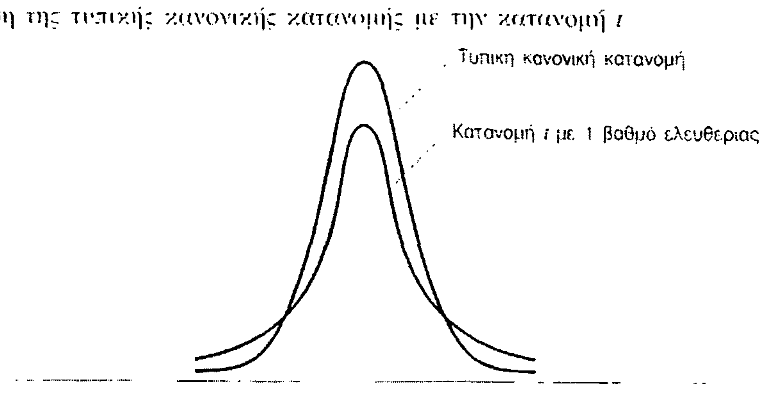
Κατά τον προσδιορισμό των διαστημάτων εμπιστοσύνης μιας πληθυσμιακής μέσης τιμής, θεωρήθηκε ότι η κατανομή της τυχαίας μεταβλητής $Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$ είναι τυπική κανονική, αν ο πληθυσμός ακολουθεί την κανονική κατανομή, και είναι κατά προσέγγιση τυπική κανονική, αν ο πληθυσμός δεν ακολουθεί την κανονική κατανομή, αλλά το n είναι επαρκώς μεγάλο. Όταν όμως η πληθυσμιακή τυπική απόκλιση είναι άγνωστη, είναι λογικό αυτή να εκτιμάται με τη βοήθεια της τυπικής απόκλισης s ενός τυχαίου δείγματος του πληθυσμού. Σε μια τέτοια περίπτωση, όμως, η ποσότητα $t = \frac{\overline{X} - \mu}{s/\sqrt{n}}$ παύει πλέον να ακολουθεί επακριβώς την τυπική κανονική κατανομή, διότι στην ενυπάρχουσα δειγματοληπτική μεταβλητότητα της τυχαίας μεταβλητής $\overline{X}$ προστίθεται και η μεταβλητότητα της δειγματικής διακύμανσης s. Εφόσον η τιμή της s αλλάζει από δείγμα σε δείγμα, είναι απαραίτητο να λαμβάνεται υπόψη η μεταβλητότητα της, ιδιαίτερα όταν το δείγμα στο οποίο βασίζονται οι εκτιμήσεις είναι μικρό.

Αν η μεταβλητή είναι κανονικά κατανεμημένη και ένα δείγμα μεγέθους n επιλεχθεί τυχαία από τον πληθυσμό, η δειγματοληπτική κατανομή της τυχαίας μεταβλητής $t = \frac{\overline{X} - \mu}{s/\sqrt{n}}$

είναι γνωστή ως *κατανομή του Student με n - 1 βαθμούς ελευθερίας* και συμβολίζεται Όπως και η τυπική κανονική κατανομή, η κατανομή t είναι κωδωνοειδής και συμμετρική γύρω από το 0, ενώ η συνολική επιφάνεια της περιοχής που βρίσκεται κάτω από την καμπύλη της είναι ίση με τη μονάδα. Οι «ουρές» της κατανομής του Student, όμως, είναι πιο πεπλατυσμένες από τις αντίστοιχες της τυπικής κανονικής κατανομής, διότι οι ακραίες τιμές είναι συχνότερες στη συγκεκριμένη κατανομή παρά στην τυπική κανονική. Γενικότερα, η διακύμανση της κατανομής t είναι μεγαλύτερη του 1, προσεγγίζοντας όμως το 1 όσο το μέγεθος του δείγματος μεγαλώνει. Για $n > 3$, η διακύμανση της κατανομής t ορίζεται από

την ποσότητα $\frac{n-1}{n-3}$.

Η σχετικά πεπλατυσμένη μορφή της κατανομής του Student, ουσιαστικά αντικατοπτρίζει την επιπλέον μεταβλητότητα που ενσωματώνεται στην ποσότητα t, λόγω του γεγονότος ότι η πληθυσμιακή τυπική απόκλιση σ δεν είναι γνωστή, αλλά κάθε φορά εκτιμάται από την αντίστοιχη δειγματική s. Εκτός τούτου, η κατανομή t προσδιορίζεται με τη βοήθεια ενός επιπλέον χαρακτηριστικού, το οποίο ονομάζεται *βαθμοί ελευθερίας (degrees of freedom).* Οι βαθμοί ελευθερίας μετρούν τις στοιχειώδεις μονάδες πληροφορίας



που εμπερικλείονται στα δειγματικά δεδομένα και οι οποίες χρήσιμο-ποιούνται κατά τον υπολογισμό της δειγματικής διακύμανσης. Οι βαθμοί ελευθερίας είναι η - 1 στην προκειμένη περίπτωση, διότι αν είναι γνωστή

η δειγματική μέση τιμή $\bar{x} = \dfrac{x_1 + x_2 + \ldots + x_n}{n}$ απαιτούνται n - 1 από

τις τιμές x_1, x_2, ...x_n προκειμένου να υπολογιστεί η δειγματική διακύμανση

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

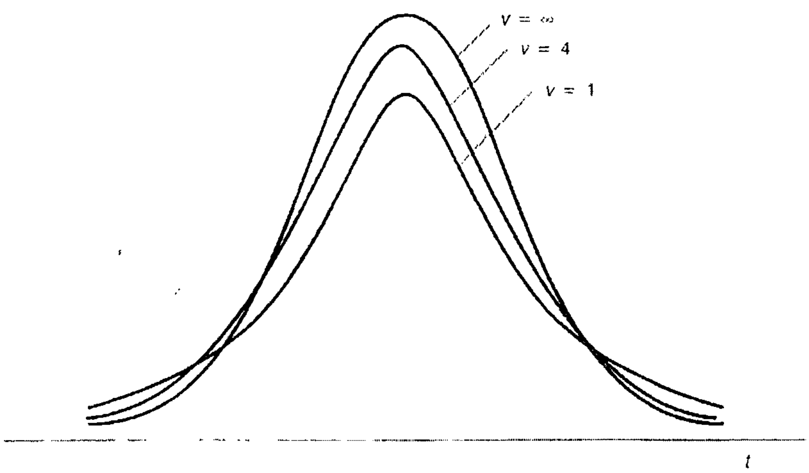
Στην πραγματικότητα, η κατανομή *t* δεν είναι μία μόνο κατανομή, αλλά μια οικογένεια κατανομών, κάθε μία από τις οποίες αντιστοιχεί σε συγκεκριμένους βαθμούς ελευθερίας. Οι κατανομές με τους λιγότερους βαθμούς ελευθερίας είναι πιο πεπλατυσμένες, ενώ όσο οι βαθμοί ελευθερίας αυξάνουν, οι κατανομές *t* προσεγγίζουν την τυπική κανονική κατανομή. Αυτό συμβαίνει, διότι όσο το μέγεθος του δείγματος αυξάνει (επομένως αυξάνουν οι βαθμοί ελευθερίας), η δειγματική τυπική απόκλιση *s* εκτιμά καλύτερα την αντίστοιχη πληθυσμιακή *σ,* ενώ όταν το δείγμα είναι *μάρα* πολύ μεγάλο η δειγματική τυπική απόκλιση τείνει να συμπέσει με την αντίστοιχη πληθυσμιακή.

Η κατανομή ί, όπως και η τυπική κανονική κατανομή, είναι πινακο-ποιημένη. Κατά τη χρήση πίνακα λαμβάνονται υπόψη οι βαθμοί ελευθερίας *ν* και ο συντελεστής εμπιστοσύνης που μας ενδιαφέρει κάθε φορά.

Η γενική μορφή ενός 100(1 - *α)%* διαστήματος εμπιστοσύνης, που ορίζεται με τη βοήθεια της κατανομής t_v με *ν* βαθμούς ελευθερίας, είναι

Σχήμα

Η κατανομή t για διάφορους βαθμούς ελευθερίας v



$$\left(\overline{X}-t_{v,\alpha/2}\frac{s}{\sqrt{n}},\ \overline{X}+t_{v,\alpha/2}\frac{s}{\sqrt{n}}\right),$$

όπου $t_{v\alpha/2}$ η τιμή που ορίζει μια περιοχή με εμβαδόν α/2 στο δεξιό άκρο της κατανομής του Student και $-t_{v\ \alpha/2}$ η τιμή που ορίζει την αντίστοιχη περιοχή με εμβαδόν α/2 στο αριστερό άκρο.

Ας σημειωθεί ότι για την έγκυρη χρήση της κατανομής t απαιτείται η πληθυσμιακή κατανομή να είναι κανονική. Εμπειρικά έχει αποδειχθεί ότι μεσαίου βαθμού εκτροπές της πληθυσμιακής κατανομής από την κανονικότητα είναι ανεκτές κατά τη χρήση της κατανομής του Student. Γενικότερα, ως ελάχιστη απαίτηση για τη χρήση της κατανομής t στον προσδιορισμό διαστημάτων εμπιστοσύνης θεωρείται η κωδωνοειδής μορφή της πληθυσμιακής κατανομής.

Παράδειγμα

Από ένα τυχαίο δείγμα 1562 σπουδαστών τριτοβάθμιων ιδρυμάτων ζητήθηκε να αξιολογηθεί συνολικά η ποιότητα των σπουδών τους, χρησιμοποιώντας μια κλίμακα ικανοποίησης από το 0 μέχρι το 100. Η μέση τιμή του βαθμού ικανοποίησης που υπολογίστηκε από τις απαντήσεις των σπουδαστών ήταν 70,32 και η τυπική απόκλιση 10,12. Να υπολογιστεί το 99% διάστημα εμπιστοσύνης για την πληθυσμιακή μέση τιμή του βαθμού ικανοποίησης.

Θεωρώντας ότι η πληθυσμιακή κατανομή του βαθμού ικανοποίησης είναι κατά προσέγγιση κανονική, μπορούμε να υπολογίσουμε το 99% διάστημα εμπιστοσύνης για την πληθυσμιακή μέση τιμή μ, χρησιμοποιώντας τα διαστήματα τιμών της κατανομής t. Έχουμε, επομένως,

$$\left(\overline{X} - t_{1561,\,0,005}\frac{s}{\sqrt{n}} \le \mu \le \overline{X} + t_{1561,\,0,005}\frac{s}{\sqrt{n}}\right)$$

Συμβουλευόμενοι τον Πίνακα της κατανομής t και χρησιμοποιώντας ττ^ μέση τιμή και την τυπική απόκλιση του βαθμού ικανοποίησης του δείγματος, παίρνουμε

$$\left(70{,}32 - 2{,}576\frac{10{,}12}{\sqrt{1562}} \le \mu \le 70{,}32 + 2{,}576\frac{10{,}12}{\sqrt{1562}}\right)$$

δηλαδή

$(69{,}66 \le \mu \le 70{,}98)$.

Επομένως, το 99% διάστημα εμπιστοσύνης του μέσου βαθμού ικανοποίησης των σπουδαστών είναι

$(69{,}66,\ 70{,}98)$

Διαστήματα εμπιστοσύνης για τη διαφορά δύο πληθυσμιακών μεσών τιμών

Σε ορισμένες περιπτώσεις είναι αναγκαίο να εκτιμηθεί, με τη βοήθεια ενός διαστήματος εμπιστοσύνης, η διαφορά δύο πληθυσμιακών μέσων τιμών. Σε μια τέτοια περίπτωση, οι διαδικασίες που αναπτύχθηκαν για τον υπολογισμό των διαστημάτων εμπιστοσύνης μιας πληθυσμιακής μέσης τιμής επεκτείνονται αντίστοιχα και για τις δύο πληθυσμιακές μέσες τιμές.

Πληθυσμιακές διακυμάνσεις ίσες

Αν οι πληθυσμιακές διακυμάνσεις θεωρηθούν ίσες, τότε και οι δύο δειγματικές διακυμάνσεις μπορούν να χρησιμοποιηθούν ως εκτιμήτριες της κοινής πληθυσμιακής διακύμανσης. Ως εκτιμήτρια συνάρτηση S_p της κοινής πληθυσμιακής διακύμανσης μπορεί να χρησιμοποιηθεί η σταθμισμένη μέση τιμή των δειγματικών διακυμάνσεων. Η στάθμιση γίνεται ως προς τους βαθμούς ελευθερίας κάθε κατανομής. Η εκτίμηση της πληθυσμιακής διακύμανσης από τα δειγματικά δεδομένα γίνεται, επομένως, από την ποσότητα

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

Το τυπικό σφάλμα της διαφοράς των δύο πληθυσμιακών μέσων τιμών δίδεται από τη σχέση

$$\sigma_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{s_P^2}{n_1} + \frac{s_P^2}{n_2}}$$

Άρα, τα 100(1 - *ά)%* όρια εμπιστοσύνης για τη διαφορά των πληθυσμιακών μέσων μ_1 - μ_2 είναι

$$\left(\overline{X}_1 - \overline{X}_2\right) \pm t_{(n_1+n_2-2), a/2} = \sqrt{\frac{s_P^2}{n_1} + \frac{s_P^2}{n_2}}$$

όπου $n_1 + n_2 - 2$ είναι οι βαθμοί ελευθερίας της κατανομής *t*.

Διαστήματα εμπιστοσύνης για μια πληθυσμιακή αναλόγιο

Για την κατασκευή ενός διαστήματος εμπιστοσύνης μιας πληθυσμιακής! αναλογίας *p,* εργαζόμαστε με τρόπο παρόμοιο όπως και στην περίπτωση της πληθυσμιακής μέσης τιμής. Ένα δείγμα επιλέγεται τυχαία από τον πληθυσμό και η δειγματική αναλογία *p* που μας ενδιαφέρει υπολογίζεται. Η δειγματική αυτή αναλογία, όπως έχει ήδη αναφερθεί είναι σημειακή εκτίμηση της πληθυσμιακής.

Διαστήματα εμπιστοσύνης για τη διαφορά δύο πληθυσμιακών αναλογιών

Αντίστοιχα με τα διαστήματα εμπιστοσύνης για μια πληθυσμιακή αναλογία υπολογίζονται και τα διαστήματα εμπιστοσύνης για τη διαφορά |δύο πληθυσμιακών αναλογιών. Μια αμερόληπτη εκτιμήτρια της διαφοράς $p_1 - p_2$ δύο πληθυσμιακών αναλογιών είναι η διαφορά $\hat{p}_1 - \hat{p}_2$ των δειγματικών αναλογιών που υπολογίζονται σε όλα τα δυνατά ζεύγη τυχαίων δειγμάτων συγκεκριμένου μεγέθους, τα οποία μπορούν να εξαχθούν από τους δύο πληθυσμούς.

Προσδιορισμός του μεγέθους του δείγματος για την εκτίμηση μέσων τιμών

Το ζήτημα του μεγέθους του δείγματος τίθεται πάντοτε κατά την πραγματοποίηση μιας έρευνας ή μιας πειραματικής διαδικασίας. Είναι σημαντικό και είναι αναγκαίο να αντιμετωπίζεται με την απαιτούμενη βαρύτητα. Ένα δείγμα μεγαλύτερο από αυτό που απαιτείται για το σκοπό μιας διερεύνησης είναι πιθανά δαπανηρό και χρονοβόρο κατά τη λήψη του, ενώ ένα μικρό δείγμα συχνά οδηγεί σε αποτελέσματα μειωμένης ισχύος, με ελάχιστο ενδιαφέρον από πρακτικής απόψεως. Θα δούμε αρχικά πώς μπορούμε να προσδιορίσουμε το μέγεθος του δείγματος για την εκτίμηση μιας πληθυσμιακής μέσης τιμής και, στη συνέχεια, θα επεκτείνουμε τη διαδικασία αυτήν και στην περίπτωση της εκτίμησης μιας πληθυσμιακής αναλογίας. Με απλές αναγωγές οι διαδικασίες αυτές επεκτείνονται και για τον προσδιορισμό του μεγέθους δειγμάτων σε πιο σύνθετες διερευνήσεις.

Σκοπός των εκτιμήσεων διαστήματος είναι η κατασκευή διαστημάτων ε μπιστοσύνης μικρού εύρους και μεγάλης αξιοπιστίας. Από τη γενική μορφή ενός διαστήματος εμπιστοσύνης προκύπτει ότι το εύρος του εξαρτάται από την τιμή της κατάλληλης κάθε φορά θεωρητικής κατανομής –που ορίζει έναν συντελεστή εμπιστοσύνης ίσο με 100 (1 - α)% και το τυπικό σφάλμα της εκτιμούμενης πληθυσμιακής παραμέτρου. Στην περίπτωση μιας πληθυσμιακής μέσης τιμής, οι δύο αυτές συνιστώσες ενός διαστήματος εμπιστοσύνης ορίζουν το γινόμενο $z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$ το διπλάσιο του οποίου δίνει το εύρος του 100(1 - *α)%* διαστήματος εμπιστοσύνης. Αν η ποσότητα $z_{\alpha/2}$ είναι σταθερή, δηλαδή αν ορίσουμε έναν συγκεκριμένο συντελεστή εμπιστοσύνης για το διάστημα, το εύρος του τελευταίου θα ελαττώνεται όσο το τυπικό σφάλμα ελαττώνεται επίσης. Επειδή το τυπικό σφάλμα είναι ίσο με $\frac{\sigma}{\sqrt{n}}$ και η τυπική απόκλιση του πληθυσμού σ είναι σταθερή, για να προκύψει ένα μικρό τυπικό σφάλμα απαιτείται ένα μεγάλο δείγμα. Το ζήτημα είναι πόσο μεγάλο πρέπει να είναι το δείγμα; Η απάντηση στο τελευταίο ερώτημα εξαρτάται από το μέγεθος της πληθυσμιακής τυπικής απόκλισης, τον απαιτούμενο συντελεστή εμπιστοσύνης του διαστήματος και το επιθυμητό εύρος του διαστήματος.

Αν υποθέσουμε ότι θέλουμε ένα 100(1 - α)% διάστημα εμπιστοσύνης που θα εκτείνεται κατά d μονάδες και από τις δύο πλευρές της δειγματικής μέσης τιμής, τότε, εφόσον η δειγματοληψία είναι με επανατοποθέτηση από άπειρο πληθυσμό ή από επαρκώς μεγάλο πληθυσμό, μπορούμε να θέσουμε

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

Επιλύοντας την παραπάνω ισότητα ως προς n, έχουμε

$$n = \frac{z_{a/2}^2 \sigma^2}{d^2}.$$

Όταν η δειγματοληψία είναι χωρίς επανατοποθέτηση από πεπερασμένο πληθυσμό, απαιτείται η διόρθωση πεπερασμένου πληθυσμού κατά τον υπολογισμό της ποσότητας d

$$d = z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}.$$

Η σχέση αυτή επιλυόμενη ως προς n δίνει

$$n = \frac{N z_{a/2}^2 \sigma^2}{d^2 (N-1) + z_{a/2}^2 \sigma^2}$$

Ο υπολογισμός του μεγέθους του δείγματος με βάση τις προηγούμενες σχέσεις απαιτεί τη γνώση της πληθυσμιακής διακύμανσης σ^2, αλλά αυτό -όπως έχει ήδη αναφερθεί- είναι μάλλον σπάνιο να συμβαίνει. Συνήθως η πληθυσμιακή διακύμανση εκτιμάται. Η εκτίμηση της πληθυσμιακής διακύμανσης μπορεί να γίνει με τους εξής τρόπους:

1. Μια προκαταρκτική πιλοτική μελέτη (pilot study) μπορεί να γίνει στον πληθυσμό και η διακύμανση του δείγματος της μελέτης αυτής να χρησιμοποιηθεί ως εκτίμηση της πληθυσμιακής διακύμανσης σ^2. Οι παρατηρήσεις μιας τέτοιας μελέτης μπορούν να χρησιμοποιηθούν και ως στοιχεία του τελικού δείγματος που θα πάρουμε. Δηλαδή, αν n είναι το μέγεθος του τελικού δείγματος, n_1 το μέγεθος του δείγματος της πιλοτικής μελέτης, ο αριθμός των παρατηρήσεων που θα πάρουμε από τον πληθυσμό μετά την πιλοτική μελέτη θα είναι $n - n_1 = n_2$.
2. Μπορεί να υπάρχουν εκτιμήσεις της πληθυσμιακής διακύμανσης σ^2 από προηγούμενες μελέτες.
3. Αν θεωρηθεί ο πληθυσμός από τον οποίο παίρνουμε το δείγμα κανονικός, τότε

μπορεί η πληθυσμιακή διακύμανση να εκτιμηθεί από τη σχέση $\sigma = \frac{εύρος}{6}$ (εφόσον το εύρος των τιμών της κατανομής ισούται με περίπου 6 τυπικές αποκλίσεις). Η μέθοδος αυτή απαιτεί κάποια γνώση για την ελάχιστη και μέγιστη δυνατή τιμή που μπορεί να πάρει η μεταβλητή στον πληθυσμό.

Παράδειγμα

Προκειμένου να εκτιμηθεί η επίδραση ενός αντιβιοτικού στην ανάπτυξη ενός συγκεκριμένου βακτηριδίου, μετρήθηκε η μέση ποσότητα βακτηριδίων σε καλλιέργειες στις οποίες εφαρμόστηκε σταθερή ποσότητα αντιβιοτικού. Από προηγούμενη εμπειρία επάνω στην επίδραση του αντιβιοτικού στο συγκεκριμένο τύπο βακτηριδίου, είναι γνωστό ότι η τυπική απόκλιση της ποσότητας των βακτηριδίων ανά καλλιέργεια είναι 13cm^2. Να προσδιοριστεί ο αριθμός των καλλιεργειών που πρέπει να χρησιμοποιηθούν, προκειμένου να εκτιμηθεί η μέση ποσότητα βακτηριδίων ανά καλλιέργεια, χρησιμοποιώντας ένα 99% διάστημα εμπιστοσύνης συνολικού εύρους $2d = 6$cm^2.

Εφόσον θέλουμε να προσδιορίσουμε ένα 99% διάστημα εμπιστοσύνης εύρους 6cm^2, ο αριθμός των καλλιεργειών που πρέπει να χρησιμοποιήσουμε είναι

$$n = \frac{z_{a/2}^2 \sigma^2}{d^2} = \frac{(2{,}58)^2 (13)^2}{3^2} = 124{,}9$$

Στρογγυλοποιώντας στον αμέσως μεγαλύτερο ακέραιο, καταλήγουμε ότι απαιτούνται 125 καλλιέργειες.

Προσδιορισμός του μεγέθους του δείγματος για την εκτίμηση αναλογιών

Με τρόπο αντίστοιχο με αυτόν που είδαμε προηγουμένως, μπορούμε να προσδιορίσουμε το μέγεθος του δείγματος και σε περιπτώσεις που ενδιαφερόμαστε για την εκτίμηση μιας πληθυσμιακής αναλογίας p.

Θεωρώντας ότι η δειγματοληψία που πρόκειται να κάνουμε είναι τυχαία και υποθέτοντας ότι ισχύουν οι προϋποθέσεις που διασφαλίζουν την κατά προσέγγιση κανονικότητα της δειγματοληπτικής κατανομής της p, οδηγούμαστε στον εξής τύπο για το μέγεθος η του δείγματος:

$$\frac{z_{a/2}^2 p(1-p)}{d^2}$$

όπου 100(1 - α)% είναι ο συντελεστής εμπιστοσύνης που θέλουμε για την εκτίμηση της πληθυσμιακής αναλογίας και *1Δ* είναι το εύρος του διαστήματος εμπιστοσύνης. Ο τύπος αυτός ισχύει, εφόσον η δειγματοληψία είναι με επανατοποθέτηση από άπειρο πληθυσμό ή από επαρκώς μεγάλο πληθυσμό.

Σε περίπτωση που η δειγματοληψία είναι χωρίς επανατοποθέτηση από πεπερασμένο πληθυσμό, στον προηγούμενο τύπο υπεισέρχεται η διόρθωση πεπερασμένου πληθυσμού

$$n = \frac{Nz_{a/2}^2 p(1-p)}{d^2(N-1) + z_{a/2}^2 p(1-p)}$$

Όταν το μέγεθος του πληθυσμού N είναι μεγάλο σε σχέση με το μέγεθος του δείγματος η (δηλαδή όταν $\frac{n}{N} \leq 0{,}05$), η διόρθωση πεπερασμένου πληθυσμού μπορεί να αγνοείται.

Και οι δυο προηγούμενοι τύποι για το μέγεθος του δείγματος απαιτούν τη γνώση της πληθυσμιακής αναλογίας ρ. Κάτι που προφανώς δεν είναι δυνατόν να συμβαίνει, εφόσον η τιμή της ρ είναι και το ζητούμενο της δειγματοληψίας. Μια λύση στο πρόβλημα είναι η χρήση μιας εκτίμησης της ρ αντί της ίδιας της ρ. Ορισμένες φορές, για παράδειγμα, μπορεί να είναι γνωστό ένα πιθανό ανώτερο όριο για την πληθυσμιακή αναλογία, το οποίο μπορεί να χρησιμοποιηθεί αντί της ρ στους συγκεκριμένους τύπους. Αν δεν υπάρχει κάποια γνώση για τη ζητούμενη πληθυσμιακή αναλογία, μπορεί να χρησιμοποιηθεί αντί αυτής η τιμή 0,5. Όταν στους παραπάνω τύπους τεθεί $\rho = 0{,}5$, τότε το μέγεθος του δείγματος η που προκύπτει είναι το μέγιστο που μπορεί να υπολογιστεί, εφόσον η ποσότητα $\rho(1 - \rho)$ παίρνει τη μέγιστη τιμή της για ρ - 0,5. Η χρήση της τιμής 0,5 αντί της πληθυσμιακής αναλογίας ρ δίνει σαφώς επαρκές μέγεθος δείγματος για την εκτίμηση που θέλουμε, είναι όμως δυνατό να δίνει και πολύ μεγαλύτερο από το απαιτούμενο. Σε μια τέτοια περίπτωση θα καταλήγαμε στην επιλογή δείγματος μεγαλυτέρου από όσο χρειαζόμαστε για την ακρίβεια της εκτίμησης και, άρα, πιο ακριβό όσον αφορά τον κόστος του ή πιο χρονοβόρο όσον αφορά τη χρονική διάρκεια της λήψης του. Για τους λόγους αυτούς είναι καλό να αποφεύγεται η χρήση της τιμής 0,5 αντί του ρ, εκτός και αν δεν υπάρχει άλλη δυνατότητα επιλογής.

Παράδειγμα

Μια δειγματοληπτική έρευνα σχεδιάστηκε προκειμένου να προσδιοριστεί η αναλογία των ατόμων του ελληνικού πληθυσμού που είναι χρήστες του διαδικτύου. Από προηγούμενη μελέτη είναι γνωστό ότι η αναλογία (%) των ατόμων που κάνουν χρήση του διαδικτύου είναι 10%. Να προσδιοριστεί το μέγεθος του δείγματος που πρέπει να ληφθεί, ώστε να διασφαλίσουμε κατά τον προσδιορισμό της πληθυσμιακής αναλογίας ένα 99% διάστημα εμπιστοσύνης εύρους $2\delta = 0{,}05$.

Επειδή το μέγεθος του πληθυσμού είναι μεγάλο σε σχέση με το μέγεθος του δείγματος που πρόκειται να προσδιοριστεί, η διόρθωση πεπερασμένου πληθυσμού μπορεί να αγνοηθεί. Χρησιμοποιώντας την αναλογία
των χρηστών από την προηγούμενη μελέτη, καταλήγουμε ότι το μέγεθος του δείγματος πρέπει να είναι

$$n = \frac{(2{,}58)^2(0{,}10)(0{,}90)}{(0{,}025)^2} = 958{,}5.$$

Δηλαδή με στρογγυλοποίηση στον πλησιέστερο προς τα άνω ακέραιο-959 άτομα.

1.2)ΕΛΕΓΧΟΙ ΥΠΟΘΕΣΕΩΝ

Έχει ήδη αναφερθεί ότι, εκτός της εκτιμητικής, το δεύτερο μεγάλο πεδίο ενδιαφέροντος της στατιστικής συμπερασματολογίας είναι οι έλεγχοι υποθέσεων. Όπως και στην εκτιμητική, έτσι και στους ελέγχους υποθέσεων οι διαδικασίες που ακολουθούνται αφορούν την εξαγωγή συμπερασμάτων για τις παραμέτρους ενός πληθυσμού, έχοντας ως βάση αναφοράς κατάλληλες στατιστικές συναρτήσεις. Ενώ όμως στην εκτιμητική τα συμπεράσματα βασίζονται σε σημειακές εκτιμήσεις και στην κατασκευή διαστημάτων εμπιστοσύνης για τις πληθυσμιακές παραμέτρους, στους ελέγχους υποθέσεων τα συμπεράσματα προκύπτουν από τη διαμόρφωση και τον έλεγχο συγκεκριμένων υποθέσεων που αφορούν τις πληθυσμιακές παραμέτρους. Και στους ελέγχους υποθέσεων, τη βάση για την εξαγωγή συμπερασμάτων αποτελούν τα στατιστικά περιγραφικά μέτρα που προέρχονται από τυχαία δείγματα.

Γενικές έννοιες

Για την πραγματοποίηση ενός ελέγχου, συνήθως ξεκινάμε από την υπόθεση ότι μια πληθυσμιακή παράμετρος, π.χ. ένας πληθυσμιακός μέσος μ, παίρνει μία συγκεκριμένη αριθμητική τιμή, έστω μ_0. Μια τέτοια υπόθεση ονομάζεται *μηδενική υπόθεση (null hypothesis)* και συμβολίζεται H_0.

Δηλαδή,

$H_0 : \mu = \mu_0$.

Μία δεύτερη πρόταση, ακριβώς αντίστροφη της μηδενικής, διατυπώνεται και συμβολίζεται H_A. Η υπόθεση αυτή ονομάζεται *εναλλακτική υπόθεση (alternative hypothesis)*.

Δηλαδή,

$H_A : \neq \mu_0$.

Από κοινού η μηδενική και η εναλλακτική υπόθεση καλύπτουν όλες τις δυνατές τιμές που μπορεί να πάρει η πληθυσμιακή μέση τιμή μ. Επομένως, διαζευκτικά, μία από τις δύο θα είναι αληθής, ενώ η απόρριψη της μίας οδηγεί υποχρεωτικά στην αποδοχή της άλλης.

Αφού διαμορφωθούν οι υποθέσεις του ελέγχου, ένα τυχαίο δείγμα μεγέθους η επιλέγεται από τον πληθυσμό για τον οποίο ενδιαφερόμαστε. Στη συνέχεια, η μέση τιμή χ του δείγματος συγκρίνεται με την τιμή μ_0 που έχει προκαθοριστεί ως τιμή του πληθυσμιακού μέσου (σύμφωνα με τη μηδενική υπόθεση). Η σύγκριση αυτή αποβλέπει στο να διερευνήσει κατά πόσο η διαφορά μεταξύ της χ και της μ_0 μπορεί πιθανολογικά να είναι αποδεκτή ή όχι. Η σύγκριση γίνεται με τη βοήθεια μιας στατιστικής συνάρτησης, η οποία ακολουθεί μια εκ των προτέρων γνωστή κατανομή. Μια συνάρτηση του είδους μπορεί -για την περίπτωση της μέσης τιμής- να είναι η τυχαία μεταβλητή

$$Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$$

η οποία ως γνωστό -εφόσον το μέγεθος του δείγματος n είναι επαρκώς μεγάλο- ακολουθεί κατά προσέγγιση την τυπική κανονική κατανομή.

Αν υπάρχουν επαρκείς ενδείξεις ότι το δείγμα που έχει ληφθεί δεν μπορεί να προέρχεται από έναν πληθυσμό με μέση τιμή μ_0, η μηδενική υπόθεση H_0 : μ — μ_0 απορρίπτεται. Αυτό συμβαίνει όταν αποδεχόμενοι την H_0 ως αληθή, η πιθανότητα να προκύψει μια δειγματική μέση τιμή τόσο ή και περισσότερο ακραία -σε σχέση με την υποτιθέμενη πληθυσμιακή μέση τιμή μ_0- όσο και η παρατηρούμενη χ, είναι επαρκώς μικρή. Σε αυτήν την περίπτωση τα δεδομένα είναι ασυμβίβαστα με τη μηδενική υπόθεση και πιθανολογικά συμβαδίζουν περισσότερο με την εναλλακτική. Με λογικό επομένως τρόπο, καταλήγουμε στο συμπέρασμα ότι η πληθυσμιακή μέση τιμή δεν μπορεί να είναι ίση με τη μ_0. Επιπλέον, διατηρώντας την κοινά αποδεκτή φρασεολογία που χρησιμοποιείται σε αυτές τις περιπτώσεις, λέμε ότι το αποτέλεσμα του ελέγχου είναι *στατιστικά σημαντικό.*

Αντίθετα, αν δεν υπάρχουν επαρκείς ενδείξεις για να αμφισβητηθεί η εγκυρότητα της μηδενικής υπόθεσης, δεν μπορεί να απορριφθεί. Σε μια τέτοια περίπτωση, καταλήγουμε ότι η πληθυσμιακή μέση τιμή μπορεί να είναι ίση με τη μ_0, χωρίς όμως αυτό να σημαίνει ότι η μηδενική υπόθεση είναι αποδεκτή. Εξακολουθεί να ισχύει το ενδεχόμενο, η πληθυσμιακή μέση τιμή να είναι διαφορετική από τη μ_0, απλά το τυχαίο δείγμα που έχει ληφθεί δεν μπορεί να το επιβεβαιώσει. Κάτι τέτοιο μπορεί να συμβαίνει όταν π.χ. το δείγμα που χρησιμοποιείται είναι πολύ μικρό.

Επίπεδο σημαντικότητας ενός ελέγχου

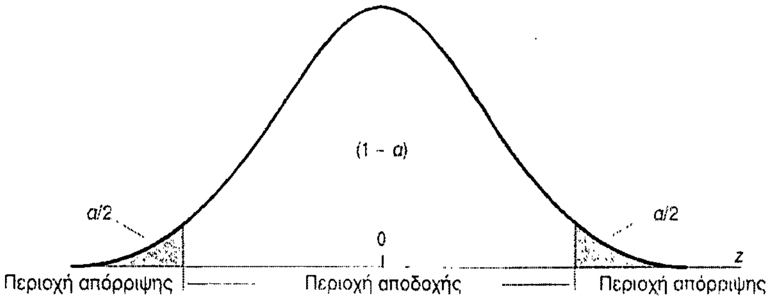
Ειπώθηκε προηγουμένως ότι, αν η πιθανότητα να προκύπτει μια δειγματική μέση τιμή, τόσο ή και περισσότερο ακραία όσο και η παρατηρούμενη χ είναι επαρκώς μικρή, μπορεί να απορριφθεί τη μηδενική υπόθεση. Λέγοντας πιθανότητα «επαρκώς μικρή» εννοούμε συνήθως πιθανότητα μικρότερη ή ίση από 0,05. Δηλαδή, μπορούμε να απορρίψουμε τη μηδενική υπόθεση H_0, όταν η πιθανότητα το δείγμα που έχουμε πάρει να προέρχεται από έναν πληθυσμό με μέση τιμή μ_0 είναι μικρότερη ή ίση από 0,05 (ή αλλιώς 5%). Στην περίπτωση του ελέγχου μιας πληθυσμιακής μέσης τιμής, αυτό ουσιαστικά μεταφράζεται στην ισχύ της πρότασης

$$Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} \geq z_{0,025} \text{ ή } Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} \leq - z_{0,025},$$

όπου $z_{0,025}$ η τιμή της τυπικής κανονικής κατανομής που ορίζει μια περιοχή με εμβαδόν 0,025 στο δεξιό άκρο της και $- z_{0,025}$ η τιμή που ορίζει μία περιοχή με εμβαδόν 0,025 στο αριστερό άκρο. Σε ορισμένες περιπτώσεις, προκειμένου ο έλεγχος να είναι πιο αυστηρός, μπορούμε να επιλέξουμε ως πιθανότητα εσφαλμένης απόρριψης της μηδενικής υπόθεσης την 0,01 (δηλαδή $Z > z_{0,025}$ ή $Z < - z_{0,025}$).

Γενικότερα, για την απόρριψη ή μη της μηδενικής υπόθεσης H_0, η δειγματοληπτική κατανομή της στατιστικής συνάρτησης του ελέγχου χωρίζεται σε δυο περιοχές: σε μία η οποία ονομάζεται *περιοχή απόρριψης (rejection area)* και μία άλλη η οποία ονομάζεται περιοχή αποδοχής *(acceptance area).* Οι τιμές της δειγματοληπτικής κατανομής που είναι σε τόσο ακραία θέση -σε σχέση με τη μέση τιμή τους- ώστε να έχουν μικρή πιθανότητα να προκύψουν, ορίζουν την περιοχή απόρριψης, ενώ οι υπόλοιπες ορίζουν την περιοχή αποδοχής. Το εμβαδόν της περιοχής απόρριψης συμβολίζεται με το ελληνικό γράμμα α και ονομάζεται *επίπεδο σημαντικότητας του* ελέγχου. Δηλαδή το επίπεδο σημαντικότητας ενός ελέγχου, ουσιαστικά ορίζει την πιθανότητα να προκύψει μια τιμή για τη συνάρτηση του ελέγχου τόσο ακραία -σε σχέση με τη μέση τιμής της- ώστε αυτή να βρίσκεται στην περιοχή απόρριψης. Αν η τιμή της στατιστικής συνάρτησης του ελέγχου για τα δειγματικά δεδομένα βρίσκεται στην περιοχή απόρριψης, τότε η μηδενική υπόθεση απορρίπτεται. Αν βρεθεί στην περιοχή αποδοχής, η μηδενική υπόθεση δεν απορρίπτεται. Το επίπεδο σημαντικότητας ενός ελέγχου ουσιαστικά αντιστοιχεί σε αυτό που στα διαστήματα εμπιστοσύνης ορίστηκε ως συντελεστής εμπιστοσύνης και συνήθως προκαθορίζεται στο 0,05 ή στο 0,01 (γενικότερα κάτω από 0,05).

Περιοχές απόρριψης και αποδοχής για τον έλεγχο μιας πληθυσμιακής μέσης τιμής απόρριψης και η περιοχή αποδοχής για τον έλεγχο μιας πληθυσμιακής μέσης τιμής με επίπεδο σημαντικότητας α.



Είδη σφαλμάτων κατά τους ελέγχους υποθέσεων

Από άποψη λογικής, ο έλεγχος μιας μέσης τιμής συνοψίζεται στις εξής διαδικασίες. Ξεκινάμε θεωρώντας ότι ισχύει η μηδενική υπόθεση

$$H_0 : \mu = \mu_0$$

και ελέγχουμε την ισχύ της πρότασης πιθανολογικά, με τη βοήθεια των ενδείξεων που παρέχονται από ένα τυχαίο δείγμα του πληθυσμού. Βασιζόμενοι στα δειγματικά δεδομένα, η μηδενική υπόθεση είτε απορρίπτεται είτε δεν απορρίπτεται. Υπάρχουν, επομένως, δυο καταστάσεις στις οποίες μπορούμε να βρεθούμε και στις οποίες το εξαγόμενο συμπέρασμα είναι σωστό: όταν η πληθυσμιακή μέση τιμή ισούται με μ_0 και η μηδενική υπόθεση $\mu = \mu_0$ δεν απορρίπτεται, και όταν η πληθυσμιακή μέση τιμή είναι διάφορη της μ_0 και η μηδενική υπόθεση απορρίπτεται.

Αντίστροφα, υπάρχουν δύο είδη σφαλμάτων που μπορούν να προκύψουν κατά τον έλεγχο: είτε να απορριφθεί η μηδενική υπόθεση μ - μ_0, ενώ στην πραγματικότητα αυτή ισχύει, είτε να μην απορριφθεί ενώ στην πραγματικότητα αυτή δεν ισχύει. Το πρώτο από τα σφάλματα αυτά ονομάζεται *σφάλμα τύπου I*, ενώ το δεύτερο ονομάζεται *σφάλμα τύπου II.* Η πιθανότητα να γίνει ένα σφάλμα τύπου I είναι ουσιαστικά το επίπεδο σημαντικότητας α του ελέγχου. Δηλαδή,

α = P (η H_0 απορρίπτεται | η H_0 ισχύει).

Δυνατά αποτελέσματα ενός ελέγχου με βάση την ορθότητα ή μη των συμπερασμάτων του

	Πληθυσμός	
Αποτέλεσμα του ελέγχου	$\mu = \mu_0$	$\mu \neq \mu_0$
Η H_0 δεν απορρίπτεται	Σωστό	Λάθος (Σφάλμα τύπου II)
Η H_0 απορρίπτεται	Λάθος (Σφάλμα τύπου I)	Σωστό

Αν, επομένως, επρόκειτο να γίνουν ανεξάρτητοι επαναλαμβανόμενοι έλεγχοι έχοντας ως επίπεδο σημαντικότητας $\alpha = 0{,}05$, στους 5% από αυτούς αναμένεται να απορριφθεί εσφαλμένα η μηδενική υπόθεση.

Η πιθανότητα να γίνει ένα σφάλμα τύπου II συμβολίζεται με το ελληνικό γράμμα β. Δηλαδή,

$\beta = P($η H_0 δεν απορρίπτεται | η H_0 δεν ισχύει$)$.

Κάθε φορά που απορρίπτεται η μηδενική υπόθεση, υπάρχει πάντα ο κίνδυνος αυτή να ισχύει και, επομένως, να διαπράξουμε ένα σφάλμα τύπου I. Επιπλέον, όταν η μηδενική υπόθεση δεν απορρίπτεται, ο κίνδυνος αποδοχής μιας εσφαλμένης υπόθεσης υπάρχει εξίσου. Κατά τη διαδικασία ενός ελέγχου μπορούμε να διασφαλίσουμε την ελαχιστοποίηση του α, αλλά δεν μπορούμε να ασκήσουμε κανέναν έλεγχο επί του β, αν και είναι γνωστό ότι στις περισσότερες των περιπτώσεων είναι μεγαλύτερο του α. Κανένα όμως από τα δύο αυτά είδη σφαλμάτων δεν είναι γνωστό αν έχει διαπραχθεί κατά τη διαδικασία του ελέγχου, εφόσον η αληθινή κατάσταση των πραγμάτων αγνοείται. Αν ο έλεγχος οδηγήσει στην απόρριψη της μηδενικής υπόθεσης, διασφαλίζεται ότι η πιθανότητα για σφάλμα τύπου I είναι επαρκώς μικρή. Αν όμως δεν απορριφθεί η μηδενική υπόθεση, η πιθανότητα να υπάρχει σφάλμα τύπου II, είναι άγνωστη.

Η πιθανότητα να προκύψει μια δειγματική μέση τιμή τόσο ή και περισσότερο ακραία όσο και η παρατηρούμενη χ, δεδομένου ότι η μηδενική υπόθεση $H_0 : \mu — \mu_0$ είναι αληθής, συμβολίζεται με το λατινικό γράμμα p (p-value). Η τιμή p συγκρίνεται κάθε φορά με το προκαθορισμένο επίπεδο σημαντικότητας του ελέγχου α, προκειμένου να αποφασιστεί αν θα απορριφθεί ή όχι η μηδενική υπόθεση. Αν η τιμή του p είναι μικρότερη ή ίση με το α, η μηδενική υπόθεση H_0 απορρίπτεται. Αν η τιμή του p είναι μεγαλύτερη από το α, η μηδενική υπόθεση δεν απορρίπτεται. Συνήθως, μαζί με το τελικό συμπέρασμα ενός ελέγχου αναφέρεται και η αντίστοιχη τιμή p που αφορά τον έλεγχο.

Αμφίπλευρος έλεγχος μιας πληθυσμιακής μέσης τιμής

Κατά τον έλεγχο ενός πληθυσμιακού μέσου, το ενδιαφέρον εστιάζεται για άλλη μια φορά στη δειγματοληπτική κατανομή της μέσης τιμής. Θεωρώντας ότι η συνεχής τυχαία μεταβλητή *X*, για την οποία γίνεται ο έλεγχος, έχει μέση τιμή μ_0 και (γνωστή) τυπική απόκλιση σ, τότε (συμφωνά πάντα με το, κεντρικό οριακό θεώρημα), η κατανομή της τυχαίας μεταβλητής

$$Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$$

είναι κατά προσέγγιση τυπική κανονική, εφόσον η τιμή του η είναι επαρκώς μεγάλη. Για μια δεδομένη επομένως δειγματική μέση τιμή $\bar{x}$, μπορούμε

να υπολογίσουμε την αντίστοιχη τιμή της *Z*, η οποία είναι η $z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$

Στη συνέχεια προσδιορίζουμε την πιθανότητα να πάρουμε για την τυχαία μεταβλητή Ζ μια τιμή τόσο ακραία ή και περισσότερο ακραία από την παρατηρούμενη $z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$. Η συγκεκριμένη πιθανότητα μπορεί να βρεθεί είτε από τους πίνακες της τυπικής κανονικής κατανομής είτε μπορεί να προσδιοριστεί με τη βοήθεια κατάλληλου λογισμικού που εξειδικεύεται στη στατιστική ανάλυση δεδομένων (π.χ. SPSS, SAS, S-PLUS κ.λπ.). Λέγοντας περισσότερο ακραία τιμή από τη $z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$, εννοούμε πιο απομακρυσμένη, στην κατεύθυνση πάντα της εναλλακτικής υπόθεσης.

Όταν η πληθυσμιακή τυπική απόκλιση σ δεν είναι γνωστή, μπορεί να εκτιμηθεί -όπως έχουμε ήδη αναφέρει- από τη δειγματική τυπική απόκλιση s. Αν η κατανομή της μεταβλητής *X* είναι κανονική, τότε η τυχαία μεταβλητή

$$Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$$

ακολουθεί την κατανομή t με η-*1* βαθμούς ελευθερίας. Σε αυτήν την περίπτωση, μπορούμε πάλι να υπολογίσουμε την τιμή της t για τη δειγματική μέση τιμή $\bar{x}$, και χρησιμοποιώντας κατάλληλο στατιστικό λογισμικό να υπολογίσουμε την πιθανότητα να πάρουμε μια δειγματική μέση τιμή πιο ακραία από την παρατηρούμενη $\bar{x}$. Η συγκεκριμένη διαδικασία ελέγχου είναι γνωστή ως *t-test,* επειδή βασίζεται στην κατανομή *t.* Μια σειρά από δοκιμασίες που βασίζονται στην κατανομή *t,* όπως θα δούμε στη συνέχεια, φέρουν την ονομασία *t-tests.*

Παράδειγμα

Η μέση τιμή της χοληστερόλης του ορού του αίματος σε ένα δείγμα 332 ενηλίκων ανδρών είναι $\bar{x}$ = 217 mg/dl. Η τυπική απόκλιση της κατανομής της χοληστερόλης στον αντίστοιχο πληθυσμό των ανδρών είναι γνωστή και ισούται με 37mg/dl. Θα ελέγξουμε αν η μέση τιμή της χοληστερόλης στον πληθυσμό των ανδρών διαφέρει από την τιμή μ_0 = 220 mg/dl, η οποία θεωρείται το ανώτερο φυσιολογικό όριο τιμών της χοληστερόλης σε ενήλικες.

Η μηδενική υπόθεση στον συγκεκριμένο έλεγχο διαμορφώνεται ως εξής:

H_0: $\mu = 220$ mg/dl.

Εφόσον η μέση τιμή μ του πληθυσμού των ανδρών ελέγχεται αν γενικά είναι διάφορη από τη μ_0, χωρίς να προσδιορίζεται ακριβώς η κατεύθυνση του ελέγχου (γεγονός που σημαίνει ότι η απόκλιση της $\bar{x}$ από τη μ_0 ελέγχεται και από τις δύο πλευρές της δειγματοληπτικής κατανομής της $\bar{X}$), η εναλλακτική υπόθεση είναι

H_A: $\mu \neq 220$ mg/dl.

Το ερώτημα επομένως που τίθεται είναι αν το δείγμα των 332 ενήλικων ανδρών μπορεί να προέρχεται ή όχι από έναν πληθυσμό ο οποίος έχει μέση τιμή χοληστερόλης μ = 220 mg/dl. Για να απαντήσουμε στο συγκεκριμένο ερώτημα, υπολογίζουμε την τιμή της τυχαίας μεταβλητής Ζ για τα δειγματικά δεδομένα

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{217 - 220}{37/\sqrt{332}}$$

Αν ισχύει η μηδενική υπόθεση, τότε η συγκεκριμένη μεταβλητή ακολουθεί την τυπική κανονική κατανομή. Το εμβαδόν της περιοχής που βρίσκεται αριστερά της τιμής $z = -1{,}477$ -το οποίο είναι και η πιθανότητα να πάρουμε μία τιμή για την τυπική κανονική κατανομή ίση με -1,477 ή και μικρότερη- είναι 0,0694. Η περιοχή δεξιά της τιμής z = 1,477 έχει εμβαδόν επίσης 0,0694. Επομένως, συνολικά το εμβαδόν των δύο περιοχών στις άκρες της κατανομής είναι 0,1388. Το εμβαδόν αυτό ορίζει την πιθανότητα p του ελέγχου, δηλαδή την πιθανότητα να προκύψει μια δειγματική μέση τιμή τόσο ή και περισσότερο ακραία -προς κάθε κατεύθυνση της δειγματοληπτικής κατανομής της Ζ- όσο και η παρατηρούμενη $\bar{x}$. Εφόσον $p > 0{,}05$, δεν μπορούμε να απορρίψουμε τη μηδενική υπόθεση. Άρα δεν μπορούμε να ισχυριστούμε ότι η μέση τιμή της χοληστερόλης στον πληθυσμό των ανδρών είναι διάφορη της τιμής $\mu_0 = 220$ mg/dl.

Ισοδυναμία των ελέγχων υποθέσεων με τα διαστήματα εμπιστοσύνης

Αν και δεν είναι άμεσα εμφανές, υπάρχει ισοδυναμία από μαθηματικής απόψεως μεταξύ των διαστημάτων εμπιστοσύνης και των ελέγχων υποθέσεων. Εφόσον ένας έλεγχος είναι αμφίπλευρος, οποιαδήποτε τιμή της στατιστικής συνάρτησης Z του ελέγχου η οποία βρίσκεται μεταξύ των τιμών -1,9.6 και 1,96 οδηγεί στον προσδιορισμό μιας πιθανότητας p μεγαλύτερης της 0,05 (π.χ. η τιμή p - 0,1388 που προέκυψε στο προηγούμενο παράδειγμα). Σε μια τέτοια περίπτωση η μηδενική υπόθεση δεν μπορεί να απορριφθεί: Από την άλλη πλευρά, η μηδενική υπόθεση H_Q μπορεί να απορριφθεί για κάθε τιμή της Z η οποία είναι μικρότερη από -1,96 ή μεγαλύτερη από 1,96. Οι τιμές -1,96 και 1,96, επειδή ουσιαστικά μπορεί να χρησιμοποιηθούν για την απόρριψη ή μη της μηδενικής υπόθεσης, ονομάζονται *κρίσιμες τιμές (critical values)* της στατιστικής συνάρτησης του ελέγχου.

Μονόπλευροι έλεγχοι υποθέσεων

Ένας έλεγχος μπορεί να διερευνά την ισχύ της μηδενικής υπόθεσης, είτε σε αντιδιαστολή με τις δύο δυνατές κατευθύνσεις που μπορεί να εμφανίζει η απόκλιση του πληθυσμιακού μέσου από τη μ_0 είτε σε αντιδιαστολή με μία μόνο προκαθορισμένη κατεύθυνση. Η κατεύθυνση του ελέγχου προσδιορίζει και τη θέση της περιοχής απόρριψης, δηλαδή αν η περιοχή αυτή θα είναι στη μία πλευρά της αντίστοιχης δειγματοληπτικής κατανομής ή θα είναι και στις δύο πλευρές της. Αν ο έλεγχος διερευνά την απόκλιση της πληθυσμιακής μέσης τιμής μ από τη μ_0 προς συγκεκριμένη κατεύθυνση, π.χ. $\mu > \mu_0$, η μηδενική υπόθεση εξακολουθεί να είναι η

$H_0: \mu = \mu_0,$

ενώ εναλλακτική είναι η

$H_A: \mu > \mu_0.$

Μόνο που τώρα η απόρριψη της μηδενικής υπόθεσης εμπερικλείει και την απόρριψη κάθε άλλης ισότητας $\mu = \mu'_o$, όπου μ Q $< \mu_0$. Και αυτό διότι σε έναν έλεγχο του είδους, αποδεχόμενοι την ισχύ της μηδενικής υπόθεσης $H_o : \mu — \mu_0$, η πιθανότητα (σύμφωνα με την κατεύθυνση του ελέγχου) να πάρουμε μια δειγματική μέση τιμή τόσο μεγάλη ή και μεγαλύτερη από την παρατηρούμενη από τα δειγματικά δεδομένα χ, μικραίνει όσο η διαφορά $\chi - \mu_0$ μεγαλώνει, όσο δηλαδή η δειγματική μέση τιμή χ

είναι μεγαλύτερη από την υποτιθέμενη μέση τιμή μ_0. Επειδή όμως η δειγματική μέση τιμή είναι δεδομένη, αν η άγνωστη πληθυσμιακή μέση τιμή θεωρηθεί ότι είναι μ — $\mu'_o < \mu_0$, η πιθανότητα απόρριψης ρ' που αφορά τον έλεγχο $H_o : \mu$ -μ'_0 είναι ακόμη μικρότερη από την αντίστοιχη πιθανότητα απόρριψης/? που αφορά τον έλεγχο $H_o : \mu = \mu_0$. Άρα η απόρριψη της $H_o : \mu = \mu_0$ εμπερικλείει και την απόρριψη κάθε άλλης μηδενικής υπόθεσης $H_o : \mu = \mu'_o$, όπου $\mu'_o < \mu_0$. Επομένως, η εναλλακτική υπόθεση που προκύπτει υποχρεωτικά είναι η $H_A : \mu > \mu_0$.

Σε έναν μονόπλευρο έλεγχο, επομένως, η περιοχή απόρριψης βρίσκεται εξ ολοκλήρου στην ουρά της κατανομής προς την οποία εμφανίζεται και η απόκλιση της πληθυσμιακής μέσης τιμής μ από την προκαθορισμένη τιμή μ_0.

Ισχύς ενός ελέγχου

Είδαμε προηγουμένως ότι δυο είδη σφαλμάτων μπορούν να προκύψουν κατά τον έλεγχο μιας υπόθεσης: είτε αυτή να απορριφθεί, ενώ στην πραγματικότητα ισχύει, είτε να μην απορριφθεί, ενώ στην πραγματικότητα δεν ισχύει. Το πρώτο από αυτά τα σφάλματα ονομάζεται σφάλμα τύπου I και είναι στην ουσία το επίπεδο σημαντικότητας *α* του ελέγχου, ενώ το δεύτερο ονομάζεται σφάλμα τύπου II και συμβολίζεται με το ελληνικό γράμμα *β*. Κατά τη διαδικασία ενός ελέγχου μπορούμε να ελαχιστοποιήσουμε to *α*, αλλά δεν μπορούμε να ασκήσουμε καμία επίδραση επί του β.

Έλεγχος της διαφοράς δύο μεσών τιμών

Συχνά ο έλεγχος ενός πληθυσμιακού μέσου δεν γίνεται ως προς μία προκαθορισμένη αριθμητική τιμή, αλλά ως προς μία άλλη πληθυσμιακή μέση τιμή. Δηλαδή η διαφορά που εντοπίζεται μεταξύ δυο διαφορετικών δειγματικών μέσων ελέγχεται ως προς το ενδεχόμενο να είναι τυχαία ή να οφείλεται στη διαφορά που όντως υπάρχει μεταξύ των αντιστοίχων πληθυσμιακών μέσων τιμών. Ένας έλεγχος του είδους είναι παρόμοιος από πολλές απόψεις με τον έλεγχο ενός πληθυσμιακού μέσου ως προς μία προκαθορισμένη αριθμητική τιμή. Η μηδενική υπόθεση είναι συνήθως αυτή που ορίζεται από την ισότητα των δύο πληθυσμιακών μέσων τιμών $\mu_1 = \mu_2$, ενώ η εναλλακτική είναι αυτή που ορίζεται από τη διαφορά $\mu_1 \neq \mu_2$ ή από μία εκ των δύο ανισοτήτων: $\mu_1 > \mu_2$ ή $\mu_1 < \mu_2$. Αποδεχόμενοι την ισχύ της μηδενικής υπόθεσης και ανάλογα με την εναλλακτική που τίθεται κάθε φορά, υπολογίζουμε την πιθανότητα να πάρουμε ένα ζεύγος δειγματικών μέσων τιμών τόσο διαφορετικών μεταξύ τους ή και ακόμη περισσότερο όσο και των παρατηρούμενων από τα διαθέσιμα δειγματικά δεδομένα. Αν η πιθανότητα αυτή είναι επαρκώς μικρή, απορρίπτουμε τη μηδενική υπόθεση και καταλήγουμε ότι οι δύο πληθυσμιακοί μέσοι διαφέρουν πράγματι μεταξύ τους. Όπως και στους προηγούμενους ελέγχους που έχουμε δει, πρέπει να προκαθορίζεται το επίπεδο σημαντικότητας *α*, καθώς επίσης και η κατεύθυνση του ελέγχου. Η στατιστική συνάρτηση που θα χρησιμοποιηθεί στον έλεγχο εξαρτάται από τη μορφή των δεδομένων. Αν δηλαδή αυτά προέρχονται από ανεξάρτητα τυχαία δείγματα ή από δείγματα εξαρτημένα κατά ζεύγη.

Ο Δήμος Αθηνών έχει παραγγείλει μια μεγάλη ποσότητα ηλεκτρικών λαμπτήρων. Το εργοστάσιο κατασκευής ισχυρίζεται ότι η μέση διάρκεια ζωής των λαμπτήρων είναι 1000 ώρες. Ο Υπεύθυνος των αγορών παίρνει ένα τυχαίο δείγμα από 25 λαμπτήρες. Αν η μέση διάρκεια ζωής τους είναι 976 ώρες ή περισσότερες, η παραγγελία γίνεται δεκτή, αν όμως η μέση διάρκεια ζωής τους είναι μικρότερη από 976 ώρες, η παραγγελία επιστρέφεται, δηλαδή αν $\overline{X} \geq$ 976 γίνεται η παραλαβή, ενώ αν $\overline{X}$ < 976 επιστρέφεται η παραγγελία.

Λύση

Το πρόβλημα που παρουσιάζεται στην προκειμένη περίπτωση είναι: Με βάση τα στοιχεία ενός δείγματος, μπορεί να επιστραφεί η παραγγελία, ενώ οι λαμπτήρες είναι καλοί ή να παραληφθούν οι λαμπτήρες, ενώ οι λαμπτήρες είναι ελαττωματικοί.

Ας υποθέσουμε ότι η παραγγελία είναι καλή, αν η μέση διάρκεια ζωής τους ' είναι-μ = 1000 και είναι κακή, αν μ < 976 ώρες, δηλαδή:

καλή αν $\mu \geq 1000$ ώρες

κακή αν $\mu < 976$ ώρες

Η απόφαση της παραλαβής ή της επιστροφής εξαρτάται από το αποτέλεσμα του δείγματος, γι' αυτό είναι δυνατόν να έχουμε τις περιπτώσεις:

α) Ενώ το εμπόρευμα είναι καλό ($\mu \geq 1000$), το αποτέλεσμα του δείγματος δείχνει ότι το φορτίο πρέπει να επιστραφεί ($\overline{X}$ < 976). Το εμπόρευμα επιστρέφεται και επομένως ο υπεύθυνος υπάλληλος του Δήμου κάνει σφάλμα πρώτου είδους.

6) Ενώ το εμπόρευμα είναι κακό (μ < 976) το δείγμα δείχνει $\overline{X} \geq$ 976 και το εμπόρευμα γίνεται δεκτό" επομένως ο υπάλληλος κάνει σφάλμα δευτέρου είδους.

Αν υποθέσουμε ότι η τυπική απόκλιση των ωρών λειτουργίας των λαμπτήρων στον πληθυσμό είναι σ = 60, τότε το σφάλμα *πρώτου* είδους θα είναι:

$$\alpha = P\ \{\text{απορρίπτουμε την } H_0 \,/\, \text{ενώ η } H_0 \text{ είναι σωστή}\} =$$

$$P\{\overline{X} < 976\} = P\left\{\frac{\overline{X}-\mu}{\sigma_{\overline{x}}} < \frac{976-1000}{\frac{60}{5}}\right\} =$$

$$= P\left\{Z < -\frac{24}{12}\right\} = P\{Z < -2\} = 1 - F(2) =$$

$$1 - 0{,}9772 = 0{,}0228 = \alpha$$

Η πιθανότητα να δεχθούμε το εμπόρευμα που είναι καλό είναι ίση με:

$$1-\alpha = 1 - 0{,}0228 = 0{,}972$$

Η πιθανότητα να κάνουμε σφάλμα δεύτερου είδους υπολογίζεται ως εξής: Αν υποθέσουμε ότι:

$$\mu = 975 \text{ και } \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{60}{\sqrt{25}} = 12$$

Αν και οι λαμπτήρες δεν είναι καλοί, ο αρμόδιος υπάλληλος θα τις παραλάβει αν το δείγμα δώσει $\overline{X} \geq$ 976 ώρες. Επομένως θα έχουμε:

$$\mathrm{B} = \mathrm{P}\left\{\overline{\mathrm{X}} \geq 976\right\} = \mathrm{P}\left\{\frac{\overline{\mathrm{X}} - \mu_1}{\sigma_{\bar{x}}} < \frac{976 - 975}{12}\right\} =$$

$$= P\left\{Z \geq \frac{1}{12}\right\} \text{(πιθανότητα δευτέρου είδους)}$$

Η πιθανότητα να επιστρέψουμε το εμπόρευμα, το οποίο είναι κακό, είναι ίση με:

$$\gamma = 1 - \beta = 1 - 0{,}083 = 0{,}917$$

Από τα παραπάνω παρατηρούμε ότι η τιμή του α και 6 εξαρτάται από την τιμή του μέσου του δείγματος και από το μέγεθος του δείγματος. Όταν η πιθανότητα α αυξάνει, η πιθανότητα 6 ελαττώνεται και αντίστροφα. Ο μόνος τρόπος να μειώσουμε και τα δύο είδη σφαλμάτων είναι να αυξήσουμε το μέγεθος του δείγματος.

Στην πράξη θεωρούμε πιο σημαντικό να αποφεύγουμε το σφάλμα πρώτου είδους σε σχέση με το σφάλμα δευτέρου είδους και επομένως το κριτήριο που πρέπει να χρησιμοποιήσουμε πρέπει να είναι εκείνο που με μικρότερη πιθανότητα μας οδηγεί στο να απορρίψουμε μια αληθινή υπόθεση.

Έστω ότι ένας επιχειρηματίας εμπορεύεται ένα προϊόν, το οποίο συσκευάζεται μηχανικά σε πακέτα με προκαθορισμένο από τις προδιαγραφές βάρος 120 γραμμαρίων. Από την πείρα έχει διαπιστωθεί ότι η τυπική απόκλιση είναι σ = 5 γραμμάρια. Έστω ότι για τον έλεγχο του μηχανισμού του βάρους και συ-σκευασίας των πακέτων παίρνουμε ένα δείγμα από 16 πακέτα για να διαπιστώσουμε αν η συσκευασία λειτουργεί σύμφωνα με τις προδιαγραφές με πιθανότητα σφάλματος 5%. Η κατανομή του βάρους ακολουθεί την κανονική κατανομή. Αν x = 117 να ελεγχθεί αν η συσκευασία λειτουργεί σύμφωνα με τις προδιαγραφές.

Λύση

Η υπόθεση θα είναι:

$$H_0: \mu = 120$$

$$H_1: \mu \neq 120$$

Δηλαδή γίνεται δεκτό ότι το μέσο βάρος των πακέτων του πληθυσμού είναι μ = 120 γραμμάρια και ότι η εκτίμηση $\bar{x} = 117$ του δείγματος διαφέρει μόνο για λόγους που οφείλονται στις διακυμάνσεις της δειγματοληψίας.

Η εναλλακτική υπόθεση H_1 είναι H_1: $\mu \neq 120$, υπόθεση που θα γίνει δεκτή αν απορριφθεί η υπόθεση μηδέν H_0. Επειδή μας ενδιαφέρει να γνωρίζουμε αν έχουμε αποκλίσει είτε προς τα πάνω, που επιφέρουν ζημιά στον επιχειρηματία, είτε διαφορά προς τα κάτω, περίπτωση που αδικεί τους πελάτες, θα εφαρμόσουμε δίπλευρο έλεγχο.

1. Χρησιμοποιούμε το κριτήριο:

$$|Z| = \frac{|\bar{x} - \mu_0|}{\frac{\sigma}{\sqrt{n}}} = \frac{|117 - 120|}{\frac{5}{4}} = \frac{|117 - 120|}{1{,}25} = \frac{|3|}{1{,}25} = |2{,}4|$$

Από τους πίνακες έχουμε:

$$-Z_{\alpha/2} = -\,1{,}96 \text{ και } Z_{\alpha/2} = 1{,}96$$

Επειδή:

$$Z = -\,2{,}4 < Z_{2,5\%} = -\,1{,}96 \text{ και } Z = +\,2{,}4 > Z_{2,5\%} = 1{,}96$$

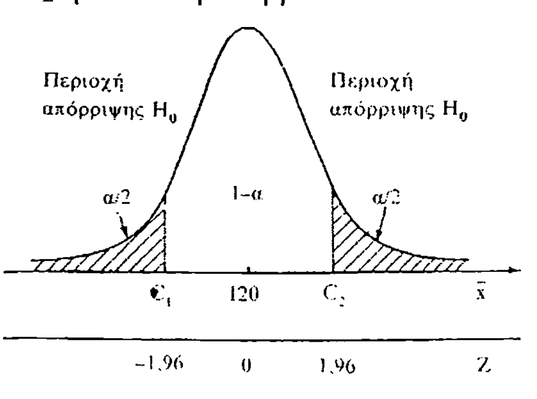
η υπόθεση απορρίπτεται.

2) Χρησιμοποιούμε τις κριτικές τιμές C_1 και C_2:

$$C_1 = 120 - 1{,}96 \;\; 5/4 = 117{,}55$$

$$C_2 = 120 + 1{,}96 \; 5/4 = 122{,}45$$

Επειδή $\bar{x} < C_1$ και $\bar{x} > C_2$ η υπόθεση απορρίπτεται.



Από έναν κανονικό πληθυσμό πήραμε ένα δείγμα με τιμές 4, 3, 7, 8. Να ελεγχθεί η υπόθεση:

$$H_0: \mu = 10$$

$$H_1: \mu < 10$$

σε επίπεδο σημαντικότητας 5%

Λύση:

Υπολογίζουμε πρώτα το δειγματικό μέσο:

$$\bar{x} = \frac{\sum x_i}{n} = \frac{4+3+7+8}{4} = \frac{22}{4} = 5{,}5$$

και στη συνέχεια τη δειγματική διακύμανση:

$$S^2 = \frac{\sum\left(x_i - \bar{x}\right)}{n-1} = 5{,}67 \quad \text{και } S = \sqrt{5{,}67} = 2{,}38$$

Στη συνέχεια, επειδή η διακύμανση του πληθυσμού είναι άγνωστη και το μέγεθος του δείγματος μικρό, χρησιμοποιούμε σαν κριτήριο τη μεταβλητή t του Student, δηλαδή:

$$t = \frac{\bar{x} - \mu_0}{\frac{S}{\sqrt{n}}} = \frac{5{,}5 - 10}{\frac{2{,}38}{\sqrt{4}}} = -3{,}78$$

Από τους πίνακες της κατανομής t για βαθμούς ελευθερίας:

$$v = n - 1 = 4 - 1 = 3 \text{ και } \alpha = 5\%$$

βρίσκουμε:

$$t_{2\alpha,v} = t_{0,10,3} = -\,2{,}353$$

Επειδή:

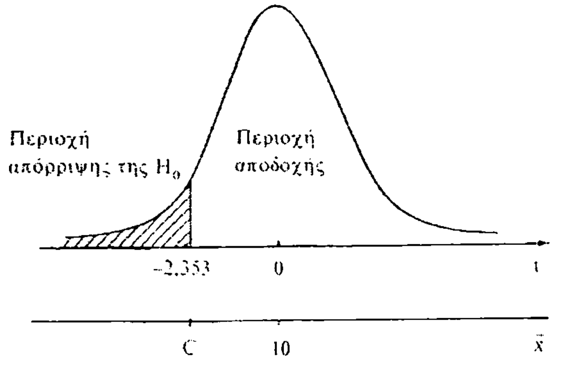
$$t = -\,3{,}78 < t_{0,10,3} = -\,2{,}353$$

απορρίπτουμε την υπόθεση H_0 και δεχόμαστε την εναλλακτική υπόθεση.

Επίσης την παραπάνω υπόθεση μπορούμε να την ελέγξουμε και ως εξής:

$$C = \mu_0 - t_{2\alpha,v}\,\frac{S}{\sqrt{n}} = 10 - 2{,}353\,\frac{2{,}38}{2} = 7{,}20$$

Επειδή $\bar{x} = 5{,}5 < C = 7{,}20$ η υπόθεση H_0 απορρίπτεται.



Σε ένα δήμο προσφέρονται 2 τύποι ηλεκτρικών λαμπτήρων για αγορά. Ο δειγματικός έλεγχος από κάθε τύπο λαμπτήρων μεγέθους n1 = 20 και n2 = 20 μας έδωσε: μέση διάρκεια ζωής του πρώτου τύπου $\bar{x}_1$ = 1000 ώρες και του δευτέρου τύπου $\bar{x}_2$ = 950 ώρες, επίσης τυπική απόκλιση S_1 = 30 ώρες και S_2 = 40 ώρες. Να ελεγχθεί η υπόθεση:

α) $H_0 : \mu_1 = \mu_2$ έναντι της $H_1 : \mu_1 \neq \mu_2$ και

β) $H_0 : \mu_1 = \mu_2$ έναντι της $H_1 : \mu_1 > \mu_2$, σε επίπεδο σημαντικότητας 5%.

Λύση

Επειδή δε γνωρίζουμε τις διακυμάνσεις σ_1^2 και σ_2^2 του πληθυσμού και το μέγεθος του δείγματος είναι μικρό n = 20, θα χρησιμοποιήσουμε σαν κριτήριο ελέγχου τη σχέση:

$$\alpha)\ t = \frac{|\bar{x}_1 + \bar{x}_2|}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{|1000 - 950|}{\sqrt{\frac{30}{20} + \frac{40}{20}}} = 4{,}5$$

Από τους πίνακες της κατανομής t με βαθμούς ελευθερίας:

$$v = n_1 + n_2 - 2 = 20 - 2 = 38,$$

επίπεδο σημαντικότητας 5% και για δίπλευρο έλεγχο θα έχουμε $t_{v,\alpha} = t_{38,5\%}$ □2.

Επειδή $t > t_{v,\alpha}$ ή 4,5 > 2 η υπόθεση απορρίπτεται.

β) Για την υπόθεση $H_0 : \mu_1 = \mu_2$ έναντι $H_1 : \mu_1 > \mu_2$ θα χρησιμοποιήσουμε το ίδιο κριτήριο όπως και στο πρώτο ερώτημα, αλλά εδώ ο έλεγχος είναι μονόπλευρος και η τιμή από τον πίνακα της κατανομής t θα είναι $t_{v,2\alpha} = t_{38,0,10} = 1{,}7$.

Επειδή 1,7 < 4,5 η υπόθεση απορρίπτεται.

2 ΚΕΦΑΛΑΙΟ
ΤΡΑΠΕΖΑ ΚΥΠΡΟΥ

Τράπεζα Κύπρου: Σύντομη ιστορική αναφορά

Το συγκρότημα της τράπεζας Κύπρου Ελλάδας είναι μέλος της τράπεζας Κύπρου που ιδρύθηκε στην Κύπρο το 1899, όπου και κατέχει εκεί το μεγαλύτερο μερίδιο αγοράς. Στα πλαίσια της διεθνούς παρουσίας του συγκροτήματος, η Ελλάδα αποτελεί τον κυριότερο άξονα επέκτασης του, καθώς έχει δημιουργηθεί ένας όμιλος όμοιος με αυτόν της Κύπρου.

Η τράπεζα Κύπρου δραστηριοποιείται στην Ελλάδα τα τελευταία 17 χρόνια, μέσα στα οποία σημειώνει μία ανοδική πορεία και γρήγορη ανάπτυξη. Η ανοδική πορεία και η ταχύτατη ανάπτυξη είναι στοιχεία που παρατηρούνται και από το γεγονός ότι από 17 καταστήματα που λειτουργούσαν το 1999, αυξηθήκαν σε 138 σήμερα , ανάλογη είναι και η αύξηση του προσωπικού του συγκροτήματος όπου από 796 άτομα το 1999 ανέρχονται στα 3000 σήμερα.

Το 2000 η τράπεζα Κύπρου Ελλάδας εισήχθη στο χρηματιστήριο Αθηνών , δημιουργώντας νέες προοπτικές για την τράπεζα στην Ελλάδα, δίνοντας το δικαίωμα στο ευρύ ελληνικό κοινό (και κυρίως στους πελάτες της) να συμμετάσχει στην εξέλιξη του αναπτυξιακού πλάνου του ομίλου.

Κατάστημα Ιωαννίνων-στατιστικά στοιχεία της τετραετίας 2002-2006

Ένα από τα καταστήματα του ομίλου Ελλάδας είναι και το κατάστημα Ιωαννίνων που βρίσκεται στο κέντρο της πόλης επί της οδού Αβέρωφ 45.

Όπως και τα υπόλοιπα καταστήματα του ομίλου προσφέρει ένα πλήρες φάσμα χρηματοοικονομικών υπηρεσιών όπως τραπεζικές υπηρεσίες, leasing, asset management, διαχείριση αμοιβαίων κεφαλαίων, χρηματιστηριακές υπηρεσίες, factoring, καθώς και ασφαλιστικές υπηρεσίες.

Τα βασικά χαρακτηριστικά που εδραίωσαν την τράπεζα Κύπρου στα Ιωάννινα και κέρδισε την εμπιστοσύνη πολλών πολιτών αλλά και φοιτητών της πόλης είναι η ποιότητα εξυπηρέτησης, η ευελιξία και η παροχή πρωτοποριακών προϊόντων.

Για την καλύτερη και γρηγορότερη εξυπηρέτηση των πελατών εφαρμόζεται ο διαχωρισμός της πελατειακής βάσης σε τρεις κύριους τομείς : τις μεγάλες επιχειρήσεις, τις μικρομεσαίες επιχειρήσεις και τους ιδιώτες.

Τον κορμό της πελατείας του καταστήματος αποτελούν κυρίως οι μικρομεσαίες επιχειρήσεις και οι έμποροι της πόλης. Σε μικρότερα ποσοστά είναι οι φοιτητές, οι οποίοι έχουν κυρίως τραπεζικές συναλλαγές όπως αναλήψεις και καταθέσεις μετρητών καθώς επίσης κάνουν εκτεταμένη χρήση του internet banking. Πολλοί μισθωτοί και δημόσιοι υπάλληλοι αποτελούν σημαντικό κομμάτι του πελατειακού κορμού όσον αφορά προϊόντα τα οποία σχετίζονται με χορηγήσεις όπως στεγαστικά, καταναλωτικά δάνεια, πιστωτικές κάρτες,factoring (αγορά αυτοκινήτου) καθώς επίσης και καταθέσεις λόγω του ότι η πληρωμή της μισθοδοσίας τους γίνεται μέσω της συγκεκριμένης τράπεζας και μέσω ελκυστικών επιτοκίων και προγραμμάτων κάνουν σε ικανοποιητικό βαθμό καταθέσεις.

Όσον αφορά τις μικρομεσαίες επιχειρήσεις ,τους εμπόρους και τους βιοτέχνες απορροφούν επενδυτικά προγράμματα, επιχειρηματικά δάνεια ,leasing,μπλοκ επιταγών, αμοιβαία κεφάλαια σε μεγάλο ποσοστό. Σχεδόν η πλειοψηφεία των επιχειρήσεων αυτής της κατηγορίας διατηρεί τουλάχιστον έναν λογαριασμό στην τράπεζα Κύπρου.

Μεγάλες επιχειρήσεις όπως εταιρείες που ανήκουν στην κατηγορία της <<βαριάς βιομηχανίας>> είναι ελάχιστες στην περιοχή και κάνουν τις συναλλαγές τους μέσω διαδικτύου.

Χορηγήσεις αποτελούν οι πιστωτικές κάρτες, τα δάνεια κάθε μορφής όπως στεγαστικά,καταναλωτικά,επιχειρηματικά,toleasing,tofactoring(αγορά αυτοκινήτου), χρηματιστηριακές συναλλαγές.

Παρακάτω παρατίθεντο στατιστικά στοιχεία της τετραετίας 2002-2006:

Από το 2003 παρατηρείται σταδιακή άνοδος στα **στεγαστικά δάνεια** της τάξης του 15%.Η ανοδική πορεία συνεχίζεται και το 2004 όπου το ποσοστό φτάνει στο 42% καθώς και το 2005 αγγίζει το 50%.Παρατηρώντας τον πίνακα βλέπουμε μία κάμψη το 2006, η σταδιακή αυτή μείωση έγινε πιο έντονη τις επόμενες χρονιές.

Η σταδιακή αύξηση των χορηγήσεων των στεγαστικών δανείων οφείλεται σε μία σειρά από ευνοϊκές ρυθμίσεις αλλά και πολιτικοοικονομικές συγκυρίες.Ήδη από το 2002 η Ελλάδα διανύει μια περίοδο προσαρμογής ως μέλος της Ενιαίας

Νομισματικής Ένωσης και άλλαξε το νόμισμα από δραχμή σε ευρώ.

Η νομισματική αλλαγή είχε ποικίλα αποτελέσματα θετικά και αρνητικά.Ένα από τα θετικά ήταν ότι οι τράπεζες μείωσαν τα επιτόκια και χορηγούσαν δάνεια και άλλα τραπεζικά προϊόντα με πολύ καλούς όρους και ανταγωνιστικά προγράμματα.Πολύ σημαντικό ρόλο στο να αυξηθούν οι χορηγήσεις στεγαστικών διαδραμάτισαν και άλλες παράμετροι όπως οι φθηνές οικοδομές,η χορήγηση επιδότησης στεγαστικού δανείου από την εργατική κατοικία καθώς και τα μέτρα που έλαβε το κράτος για να προτρέψει τους πολίτες να αγοράσουν σπίτι όπως φοροελαφρύνσεις και φοροαπαλλαγές.

Στα **επιχειρηματικά** παρατηρούμε μια αύξηση από το 2003 που από 18% που ήταν το 2002 αυξήθηκε στο 24%.Γενικά από το 2003 σημειώνεται μια σταθερή άνοδος η οποία το 2004 έφτασε το 32% και το 2005 το 38% όπου και παρέμεινε το 2006.

Παράμετροι που συντέλεσαν στην άνοδο αυτή είναι μία σειρά από μεγάλες επενδύσεις που έγιναν σε διάφορους τομείς, λόγω των ολυμπιακών αγώνων του 2004, όπως ο τουρισμός κατασκευές κτλ.Αυτή την χρονική περίοδο κατασκευάστηκαν στην Ελλάδα μεγάλα οδικά έργα όπως η Εγνατία οδός η οποία ενώνει τις πόλεις της βόρειας Ελλάδας και της Ηπείρου ζεύξη Ρίου-Αντιρρίου αποτελεί τον συνδετικό κρίκο μεταξύ Πελοποννήσου και Στερεάς Ελλάδας.Τα οδικά αυτά έργα βοήθησαν στο να αναπτυχθούν επιχειρήσεις στην επαρχία διότι πλέον το μεταφορικό κόστος μειώθηκε αισθητά.

Επίσης η κρατική επιχορήγηση με την μορφή επικουρικών προγραμμάτων σε ήδη υπάρχουσες επιχειρήσεις αλλά και επιδοτήσεις νέων επιχειρηματιών ώθησε πολλούς επιχειρηματίες να επεκτείνουν τις επιχειρήσεις τους καταφεύγοντας σε επιχειρηματικά επιδοτούμενα δάνεια.

Σταδιακή αύξηση παρατηρείται στα **καταναλωτικά δάνεια-πιστωτικές κάρτες**, από το 2003 της τάξης του 17% σε σχέση με το 2002 που ήταν 15%.Η άνοδος αυτή συνεχίζεται και το 2004 όπου φτάνει το 28% και το 2005 το 45%, ενώ το 2006 παρατηρείται μια πτώση της τάξεως του 5% δηλ. στα 40%.

Αρχικά οι καταναλωτές ήταν διστακτικοί ως προς τον δανεισμό καταναλωτικών δανείων και χρήση πιστωτικών καρτών,αυτό συνέβαινε διότι τα επιτόκια ήταν υψηλά. Από το 2003 αρχίζουν δειλά δειλά να στρέφονται στα καταναλωτικά και στις πιστωτικές κάρτες και αυτό γιατί η τράπεζα προσέφερε τα δάνεια με χαμηλότερα επιτόκια και οι όροι αποπληρωμής έγιναν πιο ευέλικτοι και ελαστικοί π.χ. ένας πελάτης που παίρνει καταναλωτικό δάνειο μπορεί να αρχίσει την αποπληρωμή του μετά από έξι μήνες ή και τον επόμενο χρόνο.

Άλλη μια παράμετρος που έπαιξε σημαντικό ρόλο στην αύξηση των χορηγήσεων ήταν και το γεγονός ότι οι αυτοκινητοβιομηχανίες στον βωμό του ανταγωνισμού προσφέρουν τα προϊόντα τους σε ελκυστικές τιμές και ανταγωνιστικά προγράμματα αποπληρωμής (factoring).

Όσον αφορά τις πιστωτικές κάρτες ισχύει ότι και για τα καταναλωτικά δάνεια,δηλαδή μειώθηκαν τα επιτόκια και επιπλέον οι καταναλωτές κατανόησαν σε μεγάλο βαθμό τον τρόπο χρήσης της πιστωτικής.

Επιπλέον πολλές εταιρείες (π.χ. κινητής τηλεφωνίας,ηλεκτρονικών υπολογιστών κτλ.) διαθέτουν τα προϊόντα τους σε χαμηλότερες τιμές εάν γίνεται χρήση πιστωτικής για την αποπληρωμή.

Στις **καταθέσεις** συμπεριλαμβάνονται κάθε είδους καταθέσεις όπως ταμιευτηρίου, όψεως, χρηματιστηριακά.

Οι μικρές διακυμάνσεις που παρατηρούνται οφείλονται στο γεγονός ότι οι πελάτες επηρεασμένοι από την ρευστή οικονομική κατάσταση(αύξηση του πληθωρισμού,αύξηση του κόστους εργασίας ,άνοδος του ποσοστού της ανεργίας)ήταν επιφυλακτικοί στο να επενδύσουν στο χρηματιστήριο.Επίσης λόγω ανόδου το πληθωρισμού σε συνδυασμό με το ότι δεν αυξάνονται οι μισθοί δεν είναι σε θέση να κάνουν σημαντικές καταθέσεις.Παρατηρήθηκε και το φαινόμενο το οποίο υπήρξε σημαντική παράμετρος,ότι πολύ καταναλωτές χρεώθηκαν με δάνεια και πιστωτικές κάρτες και ανταπεξέρχονται με δυσκολία στις υποχρεώσεις τους έτσι δεν μπορούν να προβούν σε καταθετικά προγράμματα.

Στις **καταθέσεις ταμιευτηρίου** στις οποίες συμπεριλαμβάνονται κάθε είδους καταθέσεις και αποταμιεύσεις καταναλωτών,όχι κάποιας συγκεκριμένης ομάδας αλλά ενός ευρύτερου φάσματος πελατών.Γενικά δεν παρατηρήθηκε μεγάλη κινητικότητα δηλαδή μεγάλες αυξομειώσεις διότι δεν συντελέστηκε κάποιο πολύ σημαντικό γεγονός όπως οικονομική κρίση ώστε να υπάρξουν σημαντικές μεταβολές στις καταθέσεις.Η παράμετρος που αποτελεί και κινητήριο μοχλό για τις καταθέσεις είναι οι προσφορές που κάνει η τράπεζα για να προσελκύσει πελατεία, όπως μεγάλα επιτόκια επί της κατάθεσης του ποσού(όσο μεγαλύτερη κατάθεση,ευνοϊκότερο επιτόκιο) και κατόπιν συμφωνίας με τον εκάστοτε πελάτη σε ορισμένες περιπτώσεις.

Έτσι από το 2003 αρχίζει σταδιακά να αυξάνεται το ποσοστό σε 30% και παραμένει αμετάβλητο το 2004.Το 2005 όμως αρχίζει πάλι μια ανοδική πορεία της τάξεως του 32% και συνεχίζεται η άνοδος και το 2006 όπου αγγίζει το 35%,οι αυξήσεις αυτές σημειώθηκαν διότι οι καταναλωτές

Καταθέσεις όψεως, είναι ένας ειδικός λογαριασμός ο οποίος προσφέρεται από τους χρηματοπιστωτικούς οργανισμούς και ειδικά από τις τράπεζες για την διευκόλυνση των δικαιούχων τους οι οποίοι ως επί το πλείστον είναι επιχειρήσεις.

Τα κυριότερα χαρακτηριστικά των λογαριασμών όψεως είναι

Α) Η μεγάλη κινητικότητά τους καθώς οι επιχειρήσεις τους χρησιμοποιούν ως μέσο αποπληρωμής καθημερινών χρεών, λογαριασμών κοινής ωφέλειας καθώς επίσης και για αποπληρωμή μισθοδοτικών καταστάσεων.

Β) Άλλο ένα χαρακτηριστικό των λογαριασμών όψεως είναι οι αυξημένες χρεώσεις της τράπεζας για τις υπηρεσίες που προσφέρει στον δικαιούχο.

Γ) Οι λογαριασμοί όψεως σχετίζονται άμεσα όπως είπαμε με τις επιχειρήσεις οπότε και συνίστανται στη χρήση επιταγών.

Όπως παρατηρούμε από τα στοιχεία παραπάνω οι λογαριασμοί όψεως από το 2002 μέχρι το 2006 σημείωσαν μια μικρή αύξηση της τάξης του 7%, ενώ κατά τα έτη 2004 , 2005, 2006 παρέμειναν σταθεροί σε ποσοστό του 42%, γεγονός που αποδεικνύει κατά κάποιο τρόπο τις απαρχές της επερχόμενης κρίσης και τις δυσκολίες που αντιμετώπισαν πολλές επιχειρήσεις όσο αφορά στα χρηματοοικονομικά τους αποθέματα και τα ταμειακά τους διαθέσιμα μην μπορώντας έτσι να ανταπεξέλθουν στις καθημερινές τους υποχρεώσεις σημειώνοντας μείωση και στους λογαριασμούς όψεως. Άξιο αναφοράς είναι επίσης ότι λόγω του αυξημένου δείκτη χρηματοοικονομικής μόχλευσης δηλ. χρήσης ξένων κεφαλαίων ως προς ίδια κεφάλαια οι επιχειρήσεις οδηγήθηκαν στην πτώχευση.

Η οικονομική κρίση στο κατάστημα Ιωαννίνων της τράπεζας και στην Ήπειρο γενικότερα.

Όσο αφορά στην χώρα μας, θα λέγαμε πως για πρώτη ίσως φορά μας έσωσε το γεγονός του ότι δεν έχουμε βαριά βιομηχανία και πολλές εξαγωγές προϊόντων. Οπότε, οι μόνοι τομείς που επλήγεισαν περισσότερο από τους άλλους είναι αυτοί της παροχής υπηρεσιών, του γενικότερου εμπορίου και κυρίως οι τράπεζες, που όλοι ξέρουμε την ως τώρα πολιτική τους στην Ελλάδα όσο αφορά την έκδοση δανείων και καρτών.

Συγκεκριμενοποιώντας για την περιοχή της Ηπείρου θα λέγαμε πως λόγω του ότι η

Ήπειρος είναι αποκομμένη σχετικά με την υπόλοιπη Ελλάδα και καθώς η νοοτροπία των ντόπιων καταναλωτών ευτυχώς όπως αποδεικνύεται είναι ακόμα του στυλ «αγοράζουμε μόνο και πάντα με μετρητά εκτός ελαχίστων και εξαιρετικών περιπτώσεων» και επειδή επίσης η οικονομία του νομού Ιωαννίνων στηρίζεται ως επί το πλείστον στην γεωργία και κτηνοτροφία, ο νομός δεν επλήγη τόσο όσο άλλες περιοχές ή χώρες.

Το κατάστημα της Τράπεζας Κύπρου της πόλης των Ιωαννίνων όπως δείχνουν τα λιγοστά στοιχεία τα οποία καταφέραμε και συλλέξουμε εφόσον οι αρμόδιοι ακόμη μετρούν τις ζημιές από την πρόσφατη κρίση, δείχνουν πως το συγκεκριμένο κατάστημα δεν είχε τεράστιες αποκλίσεις από τις προηγούμενες αποδόσεις του σε όλους τους τομείς των τραπεζικών υπηρεσιών. Αυτό οφείλεται στους λόγους που αναλύσαμε πιο πάνω. Φυσικά μπορούμε να μιλήσουμε μόνο για μια ιδέα και περιγραφικά για την έκταση αυτών των αποκλίσεων καθώς τα απαραίτητα στοιχεία για μια τέτοια έρευνα δίνονται στη διάθεση τρίτων έπειτα από αρκετό καιρό και για ευνόητους λόγους. Μην ξεχνάμε ότι οι ειδικοί ακόμη μετρούν την έκταση των ζημιών από την κρίση σε όλους τους τομείς, κάτι που φυσικά θα φανεί μετά από λίγα χρόνια. Ελπίζουμε όλοι στην ανάκαμψη της οικονομίας τόσο της χώρας μας όσο και της παγκόσμιας με την οποία είμαστε άρρηκτα συνδεδεμένοι μέσω του φαινομένου των ανοιχτών οικονομιών και της παγκοσμιοποίησης.-

ΣΥΜΠΕΡΑΣΜΑΤΑ

Η έρευνα για την πραγματοποίηση αυτής της εργασίας, αποδείχτηκε πολύτιμη για .Με αφορμή την συγγραφή της, μου δόθηκε το έναυσμα για να έρθω σε επαφή με μία επιστήμη όπως η στατιστική, την οποία την συναντάμε πολύ συχνά σε διάφορους τομείς της ζωής μας, όπως έρευνες για διάφορα είδη τροφίμων,ιατρικές ανακαλύψεις,δημοσκοπήσεις κ.τ.λ.

Επίσης η λεπτομερής μελέτη δύο βασικών πεδίων της στατιστικής συμπερασματολογίας όπως η εκτιμητική και οι έλεγχοι υποθέσεων συντέλεσαν στο να είναι πιο κατανοητή πλέον η μελέτη και τεκμηρίωση μιας οικονομικής έκθεσης και η διεξαγωγή σωστών συμπερασμάτων από αυτή.

Η έρευνα σχετικά με το κατάστημα της Τράπεζας Κύπρου στα Ιωάννινα κατέστησε πιο σαφή την εικόνα οργάνωσης και λειτουργίας τόσο ενός τοπικού καταστήματος αλλά έδωσε και μια πολύ σημαντική ιδέα για το πώς λειτουργεί το τραπεζικό σύστημα και βάσει ποιών προϋποθέσεων παίρνονται αποφάσεις για κατάργηση ή αναζήτηση νέων ευέλικτων και προσοδοφόρων τραπεζικών προϊόντων.

Η διαμόρφωση της πολιτικής προσέλκυσης πελατείας επηρεάζεται κυρίως από τις κοινωνικές και οικονομικές εξελίξεις αλλά και από την πολιτική των ανταγωνιστικών τραπεζών.

Η αποτύπωση της πορείας ενός τραπεζικού καταστήματος γίνεται μέσω των στατιστικών στοιχείων.Παρεκλίσεις ή καθυστερήσεις από τους στόχους έχουν ως αποτέλεσμα την αλλαγή της πολιτικής του καταστήματος.

Οι στόχοι δεν αποτελούν <<ηθικές αρχές>> αλλά ποσοτικοποιημένα μεγέθη,δηλαδή νούμερα που για να εκπληρωθούν απαραίτητη προϋπόθεση είναι η συνεχής παρατήρηση και επεξεργασία.

ΒΙΒΛΙΟΓΡΑΦΙΑ

1) Κάκουλος Θ.(1972)Στατιστική,Θεωρία και Εφαρμογές.Αθήνα

2) Γναρδέλης Χ.(2003)Εφαρμοσμένη Στατιστική.Αθήνα.

3) www.bankofcyprus.gr .....Τράπεζα της Κύπρου

4) www.naftemporiki.gr.........Η Ναυτεμπορική

5) www.ekke.gr .... Εθνικό Κέντρο Κοινωνικών Ερευνών

6) www.bankofgreece.gr .....Τράπεζα της Ελλάδος

7) www.eixe.gr ....Εθνικό Ινστιτούτο Χρηματοπιστωτικών Ερευνών

8) www.capital.gr ....Το Κεφάλαιο

9) www.economics.gr .... Οικονομική Επιθεώρηση

10) www.hba.gr .... Ένωση Ελληνικών Τραπεζών