Κεφάλαιο 9

Έλεγχοι υποθέσεων

9.1 Εισαγωγή

Όταν παίρνουμε ένα ή περισσότερα *τυχαία δείγμα από κανονικούς πληθυσμούς* έχουμε τη δυνατότητα να υπολογίζουμε στατιστικά, όπως μέσους όρους, δειγματικές διασπορές ή αναλογίες. Αυτά τα στατιστικά αφορούν μόνο τα δείγματα από τα οποία προέρχονται τα στατιστικά. Έχει όμως ιδιαίτερη σημασία και αξία το να μπορούμε να βγάζουμε συμπεράσματα για τις παραμέτρους των πληθυσμών από τους οποίους πήραμε τα δείγματα. Για παράδειγμα, αν υπολογίσουμε τον μέσο όρο μιας μεταβλητής και αυτός έχει την τιμή 5, είναι εύλογο να αναρωτηθούμε αν είναι 5 ή διαφορετική του 5 η μέση τιμή της μεταβλητής για ολόκληρο τον πληθυσμό. Ή αν σε ένα γκάλοπ βρούμε ότι το κόμμα Α φαίνεται ότι θα το ψήφιζαν το 30% των ερωτωμένων, θα είναι τόσο το ποσοστό και σε ολόκληρο τον πληθυσμό των ψηφοφόρων ή θα διαφέρει; Ή ακόμη αν επιχειρήσουμε να βελτιώσουμε τη διδασκαλία ενός μαθήματος και στις εξετάσεις που ακολουθούν, οι βαθμοί που γράφουν οι φοιτητές σε ένα τυχαίο δείγμα που πήραμε, είναι κατά μέσο όρο μία μονάδα μεγαλύτεροι από παλιότερους βαθμούς, τότε μπορούμε να πούμε ότι γενικότερα βελτιώθηκε η επίδοση των φοιτητών στο μάθημα; Για να απαντήσουμε θα πρέπει να κάνουμε ελέγχους υποθέσεων. Ελέγχους υποθέσεων κάνουμε όταν παίρνουμε στατιστικά από δείγματα και στη συνέχεια βγάζουμε συμπεράσματα για τις αντίστοιχες παραμέτρους των πληθυσμών από τους οποίους προέρχονται αυτά τα δείγματα. Δεν επικεντρωνόμαστε στα στατιστικά των δειγμάτων, αλλά μας ενδιαφέρει να αποφασίζουμε για το αν γίνεται συνολικά στους πληθυσμούς από τους οποίους προέρχονται τα δείγματα. Έτσι παίρνουμε δείγματα από πληθυσμούς, υπολογίζουμε τα στατιστικά τους που μας ενδιαφέρουν και που εκτιμούν τις αντίστοιχες παραμέτρους των πληθυσμών από τους οποίους προέρχονται τα δείγματα, και τέλος με βάση θεωρητικές ιδιότητες συμπεραίνουμε το τι συμβαίνει στους πληθυσμούς.

Οι έλεγχοι δίνουν συμπεράσματα που αφορούν τον πληθυσμό και αυτά τα συμπεράσματα ισχύουν γενικά αλλά όχι απόλυτα. Το τελευταίο σημαίνει ότι μπορούμε να συμπεραίνουμε ότι κάτι θα ισχύει ή δεν θα ισχύει για τον πληθυσμό, αλλά πάντα θα υπάρχει μια μικρή πιθανότητα να πέφτουμε έξω. Συχνά αυτή προσδιορίζεται στο 5%.

Ο τρόπος που εργαζόμαστε είναι να δημιουργούμε δύο μαθηματικές προτάσεις που ονομάζονται στατιστικές υποθέσεις ή απλά υποθέσεις. Είναι πάντα δύο και μέσω μιας διαδικασίας ελέγχου αποφασίζουμε αν θα απορρίψουμε τη μία ή όχι. Έτσι στους ελέγχους υποθέσεων διατυπώνουμε δύο υποθέσεις, **τη μηδενική** που συνήθως εκφράζει ισότητα ή ομοιομορφία και **την εναλλακτική** που είναι μια εναλλακτική πρόταση αλλά όχι κατ' ανάγκη η πλήρως συμπληρωματική υπόθεση. Η εναλλακτική υπόθεση μπορεί να εκφράζει τη μη ισότητα (τη διαφορά) και τότε έχουμε *δίπλευρο έλεγχο*. Ή μπορεί να εκφράζει ανισότητα (την περίπτωση του μικρότερου ή του μεγαλύτερου), οπότε έχουμε *μονόπλευρο έλεγχο*.

Σε κάθε περίπτωση πάντα επιλέγουμε **μία** εναλλακτική υπόθεση, όχι και τις τρεις: διάφορο, μεγαλύτερο ή μικρότερο. Το ποια θα επιλέξουμε έχει να κάνει με τη διατύπωση της άσκησης ή το ζητούμενο της στατιστικής ανάλυσης που επιχειρούμε. Δεν επιλέγουμε να δουλέψουμε με όλες τις εναλλακτικές υποθέσεις.

Είναι σημαντικό να έχουμε κατά νου ότι πάντα ο έλεγχος γίνεται ως προς τη μηδενική υπόθεση. Δηλαδή πάντα ελέγχουμε αν ισχύει η μηδενική έναντι της εναλλακτικής. Απορρίπτουμε τη μηδενική υπόθεση ή δεν την απορρίπτουμε. Αν την απορρίψουμε, τότε ισχύει η εναλλακτική. Όλη η κουβέντα γίνεται για τη μηδενική υπόθεση: απορρίπτουμε ή δεν απορρίπτουμε τη μηδενική υπόθεση.

Βοηθάει να έχουμε επίσης κατά νου ότι για «γραμμικά» στατιστικά όπως μέσους όρους και ποσοστά χρησιμοποιούμε την κανονική κατανομή (για μεγάλα δείγματα $n>30$) ή την κατανομή t του Student (για μικρά δείγματα). Για τετραγωνικά στατιστικά όπως τις διασπορές χρησιμοποιούμε τη X^2 κατανομή και τέλος για τους λόγους διασπορών χρησιμοποιούμε την F κατανομή.

Οι έλεγχοι υποθέσεων ονομάζονται και έλεγχοι σημαντικότητας. Οι ερωτήσεις με τις οποίες ασχολούμαστε είναι κατά πόσο, για παράδειγμα, οι μέσες τιμές δύο πληθυσμών είναι ίσες μεταξύ τους ή αν υπάρχει στατιστικά σημαντική διαφορά μεταξύ τους. *Στατιστικά σημαντική* διαφορά είναι η διαφορά που τεκμηριώνεται με τη *χρήση ενός* στατιστικού (από το αποτέλεσμα ενός τύπου) όταν αυτό συγκρίνεται με μια κρίσιμη τιμή μιας στατιστικής κατανομής η οποία εξαρτάται από *μια στάθμη σημαντικότητας α.* Το α εκφράζει την πιθανότητα να κάνουμε λάθος αν απορρίψουμε τη μηδενική υπόθεση ενώ δεν έπρεπε. Η στάθμη σημαντικότητας α γράφεται και ως σσ α και οι τιμές της είναι συνήθως 0,05 ή λιγότερο.

9.2 Έλεγχος σημαντικότητας για τη μέση τιμή (μεγάλα δείγματα)

Ο πρώτος έλεγχος που θα δούμε είναι και ο ευκολότερος. Παίρνουμε ένα μεγάλο τυχαίο δείγμα ($n>30$) και η μεταβλητή Χ που μετράμε ακολουθεί την κανονική κατανομή $N(\mu,\sigma^2)$. Τις παραμέτρους όμως αυτές μ και σ^2 της κανονικής κατανομής δεν τις γνωρίζουμε. Αυτό που γνωρίζουμε είναι το μέγεθος του δείγματος, τον μέσο όρο των παρατηρήσεων της μεταβλητής στο δείγμα και τη δειγματική διασπορά. Βασισμένοι μόνο στα στατιστικά που έχουμε, θέλουμε να ελέγξουμε αν η μέση τιμή του πληθυσμού είναι ίση με μια συγκεκριμένη και δοσμένη τιμή μ_0. Συνήθως από παλιές έρευνες γνωρίζουμε πόση ήταν η μέση τιμή μ_0 και θέλουμε να δούμε αν και σήμερα η μέση τιμή ισούται με αυτή την ποσότητα. Το μ_0 είναι επίσης γνωστό γιατί μπορεί να αποτελεί μια ιδανική τιμή ή μπορεί να αποτελεί την τιμή ενός άλλου πληθυσμού με τον οποίο θέλουμε να συγκρίνουμε τον δικό μας.

Υπολογίζουμε ή μας δίνεται ο μέσος όρος και η τυπική απόκλιση της Χ στο δείγμα, $\bar{x}$ και s αντίστοιχα.

Η μηδενική υπόθεση είναι πάντα της ισότητας. Ως εναλλακτική μπορούμε να επιλέξουμε μία (μόνο μία) από τρεις: $\mu > \mu_0, \mu < \mu_0$ ή $\mu \neq \mu_0$. Αν επιλέξουμε μία από τις δύο πρώτες κάνουμε μονόπλευρο έλεγχο, αν επιλέξουμε την τελευταία, κάνουμε δίπλευρο έλεγχο.

Στον έλεγχο υποθέσεων κάνουμε τρία βήματα:

1. Γράφουμε τις υποθέσεις, τη μηδενική και την κατάλληλη εναλλακτική.
2. Υπολογίζουμε ένα στατιστικό χρησιμοποιώντας έναν τύπο, βάσει του οποίου θα βγάλουμε συμπεράσματα.
3. Ορίζουμε την απορριπτική περιοχή, δηλαδή συγκρίνουμε το στατιστικό που βρήκαμε στο προηγούμενο βήμα, με μια κρίσιμη τιμή (χρησιμοποιούμε γι' αυτόν το λόγο την κατάλληλη θεωρητική κατανομή), και αν το στατιστικό την υπερβαίνει απορρίπτουμε τη μηδενική υπόθεση.

Άρα για την περίπτωσή μας:

Βήμα 1

Γράφουμε τις κατάλληλες υποθέσεις, τη μηδενική και μία από τις τρεις εναλλακτικές

$$\mathrm{H}_0: \mu = \mu_0$$

$$\begin{aligned} \mathrm{H}_1: \mu &> \mu_0 \\ \mu &< \mu_0 \\ \mu &\neq \mu_0 \end{aligned}$$

Βήμα 2

Υπολογίζουμε το στατιστικό z το οποίο ακολουθεί την τυπική κανονική κατανομή

$$\mathrm{z} = \frac{\bar{\mathrm{x}} - \mu_0}{\frac{\mathrm{s}}{\sqrt{\mathrm{n}}}}$$

Βήμα 3

Επιλέγουμε μια στάθμη σημαντικότητας α (αν δεν μας δίνεται θεωρούμε α=0,05)

1. Αν η εναλλακτική είναι $\mathrm{H}_1: \mu > \mu_0$ τότε απορρίπτουμε την H_0 αν $\mathrm{z} > \mathrm{z}_\alpha$.
2. Αν η εναλλακτική είναι $\mathrm{H}_1: \mu < \mu_0$ τότε απορρίπτουμε την

H_0 αν $z < -z_\alpha$ *(δηλαδή στις περιπτώσεις 1 και 2 απορρίπτουμε την H_0 αν $|z| > z_\alpha$).*

3. Αν η εναλλακτική είναι $H_1: \mu \neq \mu_0$ απορρίπτουμε την H_0 αν $|z| > z_{\frac{\alpha}{2}}$.

Για το τελευταίο διαιρούμε το α στο μισό γιατί έχουμε δίπλευρο έλεγχο. Χωρίζουμε δηλαδή τη στάθμη σημαντικότητας σε δύο μισά.

Ας εξηγήσουμε τι ακριβώς σημαίνει η στάθμη σημαντικότητας, το α. Η τιμή του z δείχνει αν η μέση τιμή ισούται με μ_0 ή όχι. Όσο πιο μεγάλο είναι το z τόσο πιο σίγουροι είμαστε ότι η μέση τιμή διαφέρει από την ποσότητα με την οποία τη συγκρίνουμε. Αν η διαφορά είναι μεγάλη (μεγάλο z) τότε η μέση τιμή του πληθυσμού διαφέρει από την ποσότητα με την οποία τη συγκρίνουμε, και τότε απορρίπτουμε τη μηδενική υπόθεση. Πόσο μεγάλη πρέπει όμως να είναι η διαφορά (το z) ώστε να απορρίψουμε την ισότητα;

Το z είναι μεγάλο (ή μικρό – δηλαδή έχει ακραία τιμή) αν ανήκει στο 5% (ή γενικά στο 100α%) των μεγαλύτερων (ακραίων) τιμών της κατανομής την οποία ακολουθεί. Οπότε εμείς βρίσκουμε ποια είναι η κρίσιμη τιμή που καθορίζει το 5% των παρατηρήσεων της κατανομής των z, και αν το συγκεκριμένο z είναι μεγαλύτερο από αυτή την κρίσιμη τιμή, τότε είναι μεγάλο και απορρίπτουμε τη μηδενική υπόθεση της ισότητας.

9.3 Έλεγχος σημαντικότητας για τη μέση τιμή (μικρά δείγματα)

Αν στην προηγούμενη περίπτωση έχουμε μικρό τυχαίο δείγμα (n≤30) και η μεταβλητή X που μετράμε ακολουθεί την κανονική κατανομή $N(\mu,\sigma^2)$, τότε τα βήματα διαμορφώνονται ως εξής:

Βήμα 1

Γραφούμε τις κατάλληλες υποθέσεις, τη μηδενική και μία από τις τρεις εναλλακτικές

$$H_0: \mu = \mu_0$$
$$H_1: \mu > \mu_0$$
$$\mu < \mu_0$$
$$\mu \neq \mu_0$$

Βήμα 2

Υπολογίζουμε το στατιστικό t το οποίο ακολουθεί την κατανομή t του Student

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

Βήμα 3

Επιλέγουμε μια στάθμη σημαντικότητας α (αν δεν μας δίνεται θεωρούμε α=0,05)

1. Αν $H_1: \mu > \mu_0$ τότε απορρίπτουμε την H_0 αν $t > t_{n-1;\alpha}$.
2. Αν $H_1: \mu < \mu_0$ τότε απορρίπτουμε την H_0 αν $t < -t_{n-1;\alpha}$ (δηλαδή στις περιπτώσεις 1 και 2 απορρίπτουμε την H_0 αν $|t| > t_{n-1;\alpha}$).
3. Αν $H_1: \mu \neq \mu_0$ απορρίπτουμε την H_0 αν $|t| > t_{n-1;\frac{\alpha}{2}}$. Χωρίζουμε δηλαδή τη στάθμη σημαντικότητας σε δύο μισά, στις δύο ουρές της συμμετρικής κατανομής Student.

Παράδειγμα 9.1:

Ας υποθέσουμε ότι ο μέσος όρος της βαθμολογίας τυχαίου δείγματος 200 φοιτητών στο μάθημα της στατιστικής είναι 7,5 και η τυπική απόκλιση 3. Μέχρι πέρσι η μέση τιμή βαθμολογίας ήταν 8. Μπορούμε να πούμε σε σσ α=2% ότι η μέση τιμή της βαθμολογίας των φοιτητών στο μάθημα της στατιστικής έπεσε;
Πώς θα δουλεύαμε αν το δείγμα είχε μέγεθος 20 άτομα;

Λύση:

Γράφουμε τις υποθέσεις

$$H_0: \mu = \mu_0$$
$$H_1: \mu < \mu_0$$

ή

$$H_0: \mu = 8$$
$$H_1: \mu < 8$$

Η εναλλακτική δηλώνει ότι ελέγχουμε αν η μέση τιμή έπεσε, σύμφωνα με το ερώτημα της άσκησης.

Στο πρώτο ερώτημα της άσκησης έχουμε μεγάλο δείγμα και θα πάρουμε κανονική κατανομή, ενώ στη δεύτερη περίπτωση με τα 20 άτομα θα χρησιμοποιήσουμε την κατανομή Student.

Όταν λοιπόν n=200 υπολογίζουμε το στατιστικό z το οποίο ακολουθεί την τυπική κανονική κατανομή

$$z = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{7{,}5 - 8}{\frac{3}{\sqrt{200}}} = -0{,}01179$$

Επειδή α=0,02 απορρίπτουμε την H_0 αν $z < -z_\alpha$.

Από τον πίνακα της τυπικής κανονικής κατανομής $z_\alpha = z_{0,02} = 2{,}054$.

Και επειδή
$z = \frac{\bar{x}-\mu_0}{\frac{s}{\sqrt{n}}} = \frac{7{,}5-8}{\frac{3}{\sqrt{200}}} = -0{,}01179$ δεν είναι μικρότερο από το
$-z_\alpha = -z_{0,02} = -2{,}054$

δεν μπορούμε να απορρίψουμε τη μηδενική υπόθεση. Δηλαδή δεν μπορούμε να συμπεράνουμε ότι η μέση βαθμολογία έπεσε. Με άλλα λόγια, μπορεί στο δείγμα να συμβαίνει κάτι τέτοιο αλλά αυτό δεν μπορούμε να το γενικεύσουμε για ολόκληρο τον πληθυσμό.

Αν τώρα είχαμε n=20 υπολογίζουμε το στατιστικό t το οποίο ακολουθεί τη Student κατανομή (για μικρά δείγματα)

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{7{,}5 - 8}{\frac{3}{\sqrt{20}}} = -0{,}03727$$

Επειδή α=0,02 απορρίπτουμε την H_0 αν $t < -t_{n-1;\alpha}$
το $t_{n-1;\alpha} = t_{20-1;0,02} = t_{19;0,02} = 2{,}205$.

Και επειδή
$t = \frac{\bar{x}-\mu_0}{\frac{s}{\sqrt{n}}} = \frac{7{,}5-8}{\frac{3}{\sqrt{200}}} = -0{,}03727$ δεν είναι μικρότερο από το

$-t_{n-1;\alpha} = -t_{19;0,02} = -2{,}205$

δεν μπορούμε να απορρίψουμε τη μηδενική υπόθεση, δηλαδή δεν μπορούμε να συμπεράνουμε ότι η μέση βαθμολογία έπεσε.

Σημειώστε ότι αν για ένα μεγάλο δείγμα (n=200) δεν απορρίπτεται η μηδενική υπόθεση, τότε μάλλον δεν θα απορρίπτεται και για ένα μικρό δείγμα (n=20) χρησιμοποιώντας τα ίδια στατιστικά.

9.4 Έλεγχος σημαντικότητας για τη διασπορά

Παίρνουμε ένα τυχαίο δείγμα με μέγεθος n, στο οποίο η μεταβλητή X που μετράμε ακολουθεί κανονική κατανομή. Υπολογίζουμε τη δειγματική διασπορά s^2. Θέλουμε να ελέγξουμε αν η διασπορά του πληθυσμού ισούται με μια συγκεκριμένη τιμή έστω ${\sigma_0}^2$

Βήμα 1

Γράφουμε τις κατάλληλες υποθέσεις, τη μηδενική και μία από τις τρεις εναλλακτικές

$$
\begin{gathered}
H_0: \sigma^2 = {\sigma_0}^2 \\
H_1: \sigma^2 > {\sigma_0}^2 \\
\sigma^2 < {\sigma_0}^2 \\
\sigma^2 \neq {\sigma_0}^2
\end{gathered}
$$

Βήμα 2

Υπολογίζουμε το στατιστικό x^2 το οποίο ακολουθεί τη X^2 κατανομή με n-1 βαθμούς ελευθερίας

$$x^2 = \frac{(n-1)s^2}{\sigma_0^2}$$

Βήμα 3

Επιλέγουμε μια στάθμη σημαντικότητας α (αν δεν μας δίνεται θεωρούμε α=0,05)

1. Αν η εναλλακτική είναι $H_1: \sigma^2 > {\sigma_0}^2$ τότε απορρίπτουμε την H_0 αν $x^2 > x^2_{n-1;\alpha}$.
2. Αν η εναλλακτική είναι $H_1: \sigma^2 < {\sigma_0}^2$ τότε απορρίπτουμε την H_0

αν $x^2 < x^2_{n-1;1-\alpha}$.

3. Αν η εναλλακτική είναι $H_1: \sigma^2 \neq {\sigma_0}^2$ τότε απορρίπτουμε την H_0 αν $x^2 > x^2_{n-1;\frac{\alpha}{2}}$ ή $x^2 < x^2_{n-1;1-\frac{\alpha}{2}}$.

Το τελευταίο ($x^2 > x^2_{n-1;\frac{\alpha}{2}}$ ή $x^2 < x^2_{n-1;1-\frac{\alpha}{2}}$) αφορά τις δύο ουρές της X^2 κατανομής. Επειδή όπως είπαμε στο Κεφάλαιο 6, η X^2 κατανομή δεν είναι συμμετρική, αν χωρίσουμε το α σε δύο μισά, οι αντίστοιχες κρίσιμες τιμές για τον δίπλευρο έλεγχο δεν είναι όπως στην τυπική κανονική ή τη Student κατανομή, που είναι συμμετρικές.

9.5 Έλεγχος σημαντικότητας για το p της δυωνυμικής κατανομής (μεγάλα δείγματα)

Παίρνουμε ένα μεγάλο τυχαίο δείγμα στο οποίο η μεταβλητή X παίρνει την τιμή 0 ή 1 (αποτυχία ή επιτυχία). Υπολογίζουμε το ποσοστό p που επιθυμούμε (των «ευνοϊκών» περιπτώσεων ή επιτυχιών). Θέλουμε να ελέγξουμε αν η αντίστοιχη αναλογία στον πληθυσμό έχει μια συγκεκριμένη τιμή. Συχνά για να υπολογίσουμε το ποσοστό p υπολογίζουμε το $\frac{x}{n}$, διαιρούμε δηλαδή το πλήθος x των ευνοϊκών περιπτώσεων διά το μέγεθος του δείγματος n. Το δείγμα είναι συνήθως μεγάλο, δεδομένου ότι μιλάμε για ποσοστό και αναλογία. Αυτό σημαίνει ότι θα χρησιμοποιήσουμε την κανονική κατανομή.

Βήμα 1

Γράφουμε τις κατάλληλες υποθέσεις, τη μηδενική και μία από τις τρεις εναλλακτικές

$$H_0: p = p_0$$
$$H_1: p > p_0$$
$$p < p_0$$
$$p \neq p_0$$

Βήμα 2

Υπολογίζουμε το στατιστικό

$$z = \frac{\frac{x}{n} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Βήμα 3

Επιλέγουμε μια στάθμη σημαντικότητας α (αν δεν μας δίνεται θεωρούμε α=0,05)

1. Αν η εναλλακτική είναι $H_1: p > p_0$ τότε απορρίπτουμε την H_0 αν $z > z_\alpha$.
2. Αν η εναλλακτική είναι $H_1: p < p_0$ τότε απορρίπτουμε την H_0 αν $z < -z_\alpha$.
3. Αν η εναλλακτική είναι $H_1: p \neq p_0$ τότε απορρίπτουμε την H_0 αν $|z| > z_{\frac{\alpha}{2}}$.

Παράδειγμα 9.2:

Αν σε τυχαίο δείγμα με n=1000 έχουμε βρει ότι x=300 άτομα έχουν μια συγκεκριμένη άποψη, μπορούμε να πούμε σε στάθμη σημαντικότητας α=5% ότι το αντίστοιχο ποσοστό (αναλογία) του πληθυσμού δεν διαφέρει από το 35%;

Λύση:

$$\hat{p} = \frac{x}{n} = \frac{300}{1000} = 0{,}30$$

Προσέξτε ότι γράφουμε δεκαδικούς αριθμούς και όχι ποσοστά. Διατυπώνουμε τις υποθέσεις, τη μηδενική και την κατάλληλη εναλλακτική

$$H_0: p = 0{,}35$$
$$H_1: p \neq 0{,}35$$

Υπολογίζουμε

$$z = \frac{\frac{x}{n} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} z = \frac{0{,}30 - 0{,}35}{\sqrt{\frac{0{,}35(1-0{,}35)}{1000}}} = -3{,}31497$$

Απορρίπτουμε την H_0 αν $|\mathrm{z}| > \mathrm{z}_{\frac{\alpha}{2}}$

$$\mathrm{z}_{\frac{\alpha}{2}} = \mathrm{z}_{0,025} = 1{,}96$$

Βλέπουμε ότι όντως $|\mathrm{z}| = 3{,}31497 > \mathrm{z}_{\frac{\alpha}{2}}$.

Οπότε απορρίπτουμε τη μηδενική υπόθεση. Η αναλογία διαφέρει από την τιμή 0,35=35%. Στις πράξεις καλό είναι να χρησιμοποιούμε ακρίβεια τεσσάρων ή και πέντε δεκαδικών όταν υπολογίζουμε ποσότητες με ποσοστά.

9.6 Σύγκριση των διασπορών δύο πληθυσμών

Αυτός ο έλεγχος έχει σημασία γιατί τον κάνουμε ως ενδιάμεσο στάδιο στον έλεγχο ισότητας μέσων τιμών σε μικρά δείγματα, τον οποίο θα δούμε παρακάτω.

Παίρνουμε δύο ανεξάρτητα τυχαία δείγματα από πληθυσμούς στους οποίους η μεταβλητή X που μετράμε ακολουθεί κανονικές κατανομές. Υπολογίζουμε τις δειγματικές διασπορές s_1^2 και s_2^2. Θέλουμε να ελέγξουμε αν οι διασπορές των δύο πληθυσμών σ_1^2, σ_2^2 είναι ίσες μεταξύ τους.

Γενικός κανόνας όταν συγκρίνουμε διασπορές είναι ότι τις διαιρούμε, δεν τις αφαιρούμε.

Επίσης βολεύει να υπολογίζουμε το στατιστικό $\frac{\mathrm{s}_1^2}{\mathrm{s}_2^2}$ βάζοντας στον αριθμητή τη μεγαλύτερη από τις δύο δειγματικές διασπορές. Ας υποθέσουμε λοιπόν ότι $\mathrm{s}_1^2 > \mathrm{s}_2^2$. Η s_1^2 θα μπει στον αριθμητή.

Βήμα 1

Γράφουμε τις κατάλληλες υποθέσεις, τη μηδενική και μία από τις τρεις εναλλακτικές (στην πραγματικότητα αν υποθέσουμε ότι βάλαμε $\mathrm{s}_1^2 > \mathrm{s}_2^2$ οι εναλλακτικές είναι δύο, του διάφορου και του μεγαλύτερου). Ας δούμε πρώτα την εναλλακτική του μεγαλύτερου. Οι υποθέσεις είναι

$$\mathrm{H}_0: \frac{\sigma_1^2}{\sigma_2^2} = 1$$

$$\mathrm{H}_1: \frac{\sigma_1^2}{\sigma_2^2} > 1$$