Εφαρμοσμένη Στατιστική

Δημήτριος Μπάγκαβος

Τμήμα Μαθηματικών και Εφαρμοσμένων Μαθηματικών
Πανεπιστήμιο Κρήτης

14 Μαρτίου 2018

Εισαγωγή: Τυπικό παράδειγμα στατιστικού ελέγχου υποθέσεων.

- Ένας νέος τύπος τσιγάρων βρίσκεται στο στάδιο ποιοτικού ελέγχου.
- Αν το τμήμα ποιοτικού ελέγχου της καπνοβιομηχανίας παραγωγής, ενδιαφέρεται να γνωρίζει τη μέση ποσότητα νικοτίνης που περιέχεται στα νέου τύπου τσιγάρα, μπορεί να υπολογίσει ένα διάστημα εμπιστοσύνης και να πάρει έτσι μια εκτίμηση για την άγνωστη μέση ποσότητα νικοτίνης.
- Στην περίπτωση όμως, που ενδιαφέρεται να γνωρίζει μόνο αν στα νέου τύπου τσιγάρα η μέση ποσότητα νικοτίνης δεν υπερβαίνει ένα μέγιστο επιτρεπτό όριο, τότε πρέπει να κάνει κατάλληλο στατιστικό έλεγχο υποθέσεων ώστε να μπορεί να αποφασίσει μεταξύ των υποθέσεων:
  - Η μέση ποσότητα νικοτίνης δεν υπερβαίνει το μέγιστο επιτρεπτό όριο (H_0).
  - Η μέση ποσότητα νικοτίνης υπερβαίνει το μέγιστο επιτρεπτό όριο (H_1).
- Ο γενικός κανόνας είναι πάντα είναι η εφαρμογή σε στοχαστικά προβλήματα απόφασης μεταξύ δύο εναλλακτικών υποθέσεων.

Εισαγωγή: Τυπικό παράδειγμα στατιστικού ελέγχου υποθέσεων.

- Η γενική ιδέα της διαδικασίας στατιστικού ελέγχου υποθέσεων είναι η εξής:
  - θέτουμε ως μηδενική υπόθεση (H_0) αυτή για την οποία αμφιβάλουμε, αυτή που αμφισβητείται,
  - εξετάζουμε αν ένα τυχαίο δείγμα που παίρνουμε από τον πληθυσμό συνηγορεί-δίνει αποδείξεις υπέρ της απόρριψής της, έναντι της εναλλακτικής (H_1).
- Δηλαδή, η H_0, απορρίπτεται ή δεν απορρίπτεται με βάση το τι παρατηρείται στο τυχαίο δείγμα που πήραμε από τον πληθυσμό.
- Πιο συγκεκριμένα, υποθέτοντας ότι η H_0 είναι αληθής, αν αυτό που παρατηρείται στο δείγμα είναι ακραίο, δηλαδή, αν έχει πολύ μικρή πιθανότητα να συμβεί, τότε απορρίπτουμε την H_0.
- Σε αντίθετη περίπτωση, δηλαδή, αν αυτό που παρατηρείται στο δείγμα δεν είναι ακραίο-σπάνιο (όταν είναι αληθής η H_0) τότε το δείγμα που πήραμε δε μας δίνει αρκετές ενδείξεις για την απόρριψη της H_0 και «αποτυγχάνουμε να την απορρίψουμε».
- Βέβαια, με αυτή τη στρατηγική παίρνουμε «ρίσκο», γιατί και τα ακραία, έστω και με πολύ μικρή πιθανότητα, μπορεί να συμβούν.

Εισαγωγή: Τυπικό παράδειγμα στατιστικού ελέγχου υποθέσεων.

- Είναι φανερό, ότι για να προχωρήσουμε πρέπει να αποσαφηνιστεί:
  - τι εννοούμε επακριβώς όταν λέμε «αυτό που παρατηρείται στο δείγμα»; Πώς εκφράζεται; Μπορεί να μετρηθεί-ποσοτικοποιηθεί;
  - Πώς κρίνουμε ότι «αυτό που παρατηρείται στο δείγμα» είναι ή όχι «ακραίο»; Δηλαδή, με ποιον σαφή κανόνα θεωρείται το παρατηρούμενο στο δείγμα «ακραίο»;
- Επίσης, πρέπει να απαντήσουμε στα εύλογα ερωτήματα:
  - Πώς υπολογίζονται οι πιθανότητες σφάλματος τύπου I και σφάλματος τύπου II;
  - Μπορούν να ελαχιστοποιηθούν; Σχετίζονται με κάποιο τρόπο· Μπορούμε να τις θέσουμε υπό τον έλεγχό μας;
- Για να απαντήσουμε στα ερωτήματα αυτά, ας χρησιμοποιήσουμε ένα συγκεκριμένο παράδειγμα. Θα μας βοηθήσει στην κατανόηση. Παράδειγμα: Το όριο αντοχής ενός τύπου καλωδίων είναι τυχαία μεταβλητή X, με μέση τιμή $\mu = 1500$Kgr και τυπική απόκλιση $\sigma = 175$Kgr. Το εργοστάσιο που κατασκευάζει αυτόν τον τύπο καλωδίων ισχυρίζεται ότι βελτίωσε τα υλικά που χρησιμοποιεί και πλέον το όριο αντοχής των καλωδίων έχει αυξηθεί.

Τυπικό παράδειγμα στατιστικού ελέγχου υποθέσεων.

- Για να ελεγχθεί ο ισχυρισμός του εργοστασίου, ως μηδενική υπόθεση θέτουμε την $H_0 : \mu = 1500$ Kgr, δηλαδή, αυτήν η οποία αμφισβητείται από τον ισχυρισμό που ελέγχουμε.
- Γενικά, η H_0 δηλώνει ότι στον πληθυσμό η κατάσταση παραμένει αμετάβλητη, δεν υπάρχει αλλαγή/διαφορά ή αλλιώς, ότι η ανεξάρτητη μεταβλητή δεν έχει επίδραση στην εξαρτημένη μεταβλητή για τον πληθυσμό (στο παράδειγμά μας, ότι η βελτίωση των υλικών δεν έχει επίδραση στο όριο αντοχής των καλωδίων).
- Ένας δεύτερος κανόνας για τον καθορισμό της H_0 που έχει επίσης καθιερωθεί στη διεθνή επιστημονική πρακτική, είναι ο εξής:
  - Ως μηδενική υπόθεση θέτουμε την υπόθεση της οποίας η λανθασμένη απόρριψη εγκυμονεί τους περισσότερους κινδύνους.
  - Δηλαδή, αυτή που απαιτεί μεγαλύτερη προστασία από σφάλμα τύπου I.
- Ως εναλλακτική θέτουμε την $H_1 : \mu > 1500$ Kgr, δηλαδή, η H_1 δηλώνει ότι η βελτίωση των υλικών επηρεάζει, και ειδικότερα αυξάνει, το όριο αντοχής των καλωδίων.
- Γενικά, η H_1 δηλώνει ότι στον πληθυσμό υπάρχει αλλαγή/διαφορά ή αλλιώς, ότι η ανεξάρτητη μεταβλητή έχει επίδραση στην εξαρτημένη μεταβλητή για τον πληθυσμό.

Τύποι στατιστικών ελέγχων.

- Ο έλεγχος που μόλις διατυπώσαμε, είναι ένας μονόπλευρος και ειδικότερα δεξιόπλευρος έλεγχος. Γενικότερα, οι έλεγχοι

$$H_0 : \mu \leq \mu_0, \;\; H_0 : \mu \geq \mu_0$$
$$H_1 : \mu > \mu_0, \;\; H_1 : \mu < \mu_0$$

  ονομάζονται μονόπλευροι έλεγχοι (δεξιόπλευρος και αριστερόπλευρος αντίστοιχα) και ο έλεγχος,

$$H_0 : \mu = \mu_0, H_1 : \mu \neq \mu_0$$

  ονομάζεται αμφίπλευρος.
- Σημειώνουμε, επίσης, ότι τα δύο σύνολα τιμών της παραμέτρου που ελέγχουμε (στο παράδειγμά μας, της μ) που ορίζουν οι δύο υποθέσεις, πρέπει προφανώς να είναι ξένα μεταξύ τους (ή το ένα άρνηση του άλλου).
- Τέλος, υπογραμμίζουμε ότι και οι δύο υποθέσεις αναφέρονται στον πληθυσμό γι'αυτό δηλώνονται με όρους παραμέτρων του πληθυσμού.

Κατασκευή του τεστ.

- Στο παράδειγμά μας, αφού διατυπώσαμε την υπόθεση ότι η άγνωστη μέση τιμή του πληθυσμού των ορίων αντοχής των καλωδίων μετά τη βελτίωση των υλικών είναι 1500Kgr ($H_0 : \mu = 1500$Kgr), παίρνουμε ένα τυχαίο δείγμα καλωδίων από το σύνολο της παραγωγής του εργοστασίου και μετράμε το όριο αντοχής κάθε καλωδίου του δείγματος.
- Για τις ανάγκες του παραδείγματος, έστω ότι ένα τυχαίο δείγμα $X_1, X_2, \ldots, X_n$ μεγέθους $n = 50$, μας έδωσε τις μετρήσεις $x_1, x_2, \ldots, x_{50}$ με $\bar{X} = 1550$Kgr.
- Η «εμπειρία», δηλαδή, αυτό που παρατηρείται στο δείγμα, συμφωνεί άραγε με την υπόθεση $H_0 : \mu = 1500$Kgr, (δηλαδή με ό,τι αυτή συνεπάγεται για το δείγμα) ή μήπως δίνει αποδείξεις εναντίον της H_0 και υπέρ της H_1;
- Για να απαντήσουμε σε αυτό το ερώτημα, πρέπει, πρώτα απ'όλα, να κατασκευάσουμε/επιλέξουμε μια κατάλληλη στατιστική συνάρτηση $T = T(X_1, X_2, \ldots, X_n)$, δηλαδή, μια συνάρτηση του δείγματος, ώστε να ποσοτικοποιήσουμε «αυτό που παρατηρείται στο δείγμα» και η οποία, υπό την H_0, δηλαδή όταν ισχύει η H_0, να ακολουθεί γνωστή κατανομή (χωρίς άγνωστες παραμέτρους) ώστε να μπορούμε να υπολογίσουμε τις απαιτούμενες για τον έλεγχο πιθανότητες.

Κατασκευή του τεστ.

- Στο παράδειγμά μας, που αφορά στον έλεγχο της μέσης τιμής, μ, του πληθυσμού, είναι λογικό να επιλέξουμε ως στατιστική συνάρτηση T, τη δειγματική μέση τιμή

Πρόταση 1, εβδ. 4

$$\bar{X} = n^{-1} \sum_{i=1}^{n} X_i \sim N(\mu, n^{-1}\sigma^2) = N(1500, 50^{-1}175^2)$$

Κανονικοποιώντας τις τιμές του δείγματος, έχουμε ότι

Σελ. 20, εβδ. 3

$$Z = \frac{\bar{X} - 1500}{175/\sqrt{50}} = 2.02 \sim N(0,1)$$

οπότε αυτό που παρατηρείται στο δείγμα εκφράζεται από την $Z = 2.02$. Πως όμως ορίζουμε το ακραίο;

- Αν η $H_0 : \mu = 1500\text{Kgr}$ είναι αληθής, είναι λογικό να αναμένουμε ότι η τιμή της $\bar{X}$ στο δείγμα να είναι κοντά στην τιμή 1500.
- Αντίθετα, αν η H_0 δεν είναι αληθής, αναμένουμε η μέση τιμή του δείγματος να είναι μακριά (προς την κατεύθυνση της H_1, δηλαδή δεξιότερα) του 1500.

Κατασκευή του τεστ.

- Ένας λογικός, επομένως, έλεγχος είναι ο εξής: ορίζουμε μια τιμή c με βάση την οποία θα κρίνεται αν η δειγματική μέση τιμή βρίσκεται μακριά από την $\mu = 1500$ Kgr, δηλαδή θα θεωρείται ακραία.
- Έτσι, αν στο παράδειγμά μας επιλέξουμε $c = 1530$ Kgr τότε επειδή $\bar{X} = 1550 > 1530$, αυτό που παρατηρείται στο δείγμα κρίνεται ακραίο και η H_0 απορρίπτεται.
- Το κριτήριο αυτό είναι φυσικά λογικό, όμως, πόσο λογική-εύλογη είναι η αυθαίρετη τιμή $c = 1530$ Kgr με την οποία οριοθετήσαμε τις ακραίες από της μη ακραίες δειγματικές μέσες τιμές.
- Αν, για παράδειγμα, επιλέξουμε $c = 1570$ Kgr, τότε $\bar{X} = 1550 < 1570$ δηλαδή τώρα το παρατηρούμενο στο δείγμα δεν κρίνεται ακραίο και το δείγμα δεν υποστηρίζει απόρριψη της H_0.
- Τίθεται, επομένως, το ερώτημα: πώς επιλέγουμε την τιμή της σταθεράς c;
- Πριν απαντήσουμε σε αυτό το εύλογο ερώτημα, ας υπολογίσουμε την πιθανότητα να κάνουμε σφάλμα τύπου I στην περίπτωση που επιλέξουμε $c = 1530$ Kgr και αντίστοιχα στην περίπτωση που επιλέξουμε $c = 1570$ Kgr.

Κατασκευή του τεστ.

- Για $c = 1530$Kgr, έχουμε:

$$P(\text{σφάλμα τύπου I}) = P(\text{απόρριψη της } H_0 | \text{αληθής η } H_0)$$

$$= P(\bar{X} \geq 1530 | \mu = 1500) = P\left(\frac{\bar{X} - 1500}{175/\sqrt{50}} \geq \frac{1530 - 1500}{175/\sqrt{50}}\right)$$

$$= P(Z \geq 1.21) = 1 - \Phi(1.21) = 0.1131$$

Ομοίως για $c = 1570$Kgr

$$= P(\bar{X} \geq 1570 | \mu = 1500) = P\left(\frac{\bar{X} - 1500}{175/\sqrt{50}} \geq \frac{1570 - 1500}{175/\sqrt{50}}\right)$$

$$= P(Z \geq 2.83) = 1 - \Phi(2.83) = 0.0023$$

Ακόμα πιο γενικά, για οποιοδήποτε c,

$$P(\bar{X} \geq c | \mu = 1500) = P\left(\frac{\bar{X} - 1500}{175/\sqrt{50}} \geq \frac{c - 1500}{175/\sqrt{50}}\right)$$

$$= P\left(Z \geq \frac{c - 1500}{175/\sqrt{50}}\right) = 1 - \Phi\left(\frac{c - 1500}{175/\sqrt{50}}\right).$$

Κατασκευή του τεστ.

- Από τα παραπάνω, είναι φανερό ότι η τιμή της σταθεράς c επηρεάζει (ακριβέστερα, καθορίζει) την πιθανότητα σφάλματος τύπου I που κάνουμε. Έτσι, με κριτήριο τον έλεγχο του μεγέθους του σφάλματος τύπου I (θυμηθείτε και πώς ορίζουμε την H_0), μπορούμε να επιλέξουμε την τιμή της c ως εξής:
- Ορίζουμε ένα μέγιστο ανεκτό μέγεθος σφάλματος τύπου I και με βάση αυτό υπολογίζουμε την τιμή της c. Με αυτό τον τρόπο, καθορίζουμε έναν απολύτως σαφή κανόνα για να κρίνουμε αν αυτό που παρατηρείται στο δείγμα, δηλαδή η τιμή της στατιστικής συνάρτησης T (στο συγκεκριμένο παράδειγμα της $\bar{X}$ ή ισοδύναμα της Z) είναι «ακραία» ή όχι, και πλέον, αποφασίζουμε για την απόρριψη ή τη μη απόρριψη της H_0, με κριτήριο ένα προκαθορισμένο μέγεθος σφάλματος τύπου I.
- Το ανεκτό επίπεδο σφάλματος τύπου I που προκαθορίζουμε, συμβολίζεται με α και ονομάζεται επίπεδο σημαντικότητας του ελέγχου (γιατί από αυτό προκύπτει η τιμή της c που ορίζει αν αυτό που παρατηρείται στο δείγμα είναι σημαντικό-σημαντική απόδειξη για να υποστηρίξει την απόρριψη της H_0). Συνήθως το επίπεδο σημαντικότητας, α, ορίζεται ίσο με 0.01 ή 0.05.

Κατασκευή του τεστ.

- Για να ολοκληρώσουμε το παράδειγμά μας, θέτοντας επίπεδο σημαντικότητας $\alpha = 0.05$, πρέπει να επιλέξουμε τιμή c τέτοια ώστε:

$$P(\bar{X} \geq c|\mu = 1500) \leq 0.05 \Leftrightarrow P\left(\frac{\bar{X} - 1500}{175/\sqrt{50}} \geq \frac{c - 1500}{175/\sqrt{50}}\right) \leq 0.05$$

$$\Leftrightarrow P\left(Z \geq \frac{c - 1500}{175/\sqrt{50}}\right) \leq 0.05 \Leftrightarrow 1 - \Phi\left(\frac{c - 1500}{175/\sqrt{50}}\right) \leq 0.05$$

$$\Phi\left(\frac{c - 1500}{175/\sqrt{50}}\right) \geq 0.95 \Leftrightarrow \frac{c - 1500}{175/\sqrt{50}} \geq z_{0.05} = 1.645$$

$$\Leftrightarrow c \geq 1500 + 1.645\frac{175}{\sqrt{50}} = 1540.8$$

- Έτσι, επιλέγοντας $c = 1540.8$ έχουμε $\bar{X} = 1550 > 1540.8$ οπότε απορρίπτουμε την H_0 με επίπεδο σημαντικότητας $\alpha = 0.05$.
- Ισοδύναμα, αν ως στατιστική συνάρτηση επιλέξουμε την $Z = (\bar{X} - 1500)\sqrt{50}/175$ έχουμε

$$P(Z \geq c) \leq 0.05 \Rightarrow c = z_{0.05} = 1.645.$$

Κατασκευή του τεστ.

- Δηλαδή, ως τιμή της c επιλέγουμε το $\alpha = 0.05$ άνω ποσοστιαίο σημείο της τυποποιημένης κανονικής κατανομής, $z_{0.05}$, και επειδή η τιμή της στατιστικής συνάρτησης στο δείγμα, $z = 175^{-1}(1550 - 1500)\sqrt{50} = 2.02 > z_{0.05} = 1.645$ απορρίπτουμε την H_0 σε επίπεδο σημαντικότητας 0.05.
- Αν η φύση του προβλήματος που εξετάζουμε επιβάλλει μεγαλύτερη «προστασία» από σφάλμα τύπου Ι, (από εσφαλμένη απόρριψη της H_0), πρέπει να είμαστε πιο «συντηρητικοί» στην απόρριψη της H_0.
- Αυτό το επιτυγχάνουμε καθορίζοντας μικρότερο ανεκτό επίπεδο σφάλματος τύπου Ι, δηλαδή, μικρότερο επίπεδο σημαντικότητας.
- Έτσι, στο παράδειγμά μας, αν επιβάλλεται πιο αυστηρός έλεγχος του ισχυρισμού του εργοστασίου, κάνουμε τον έλεγχο σε μικρότερο επίπεδο σημαντικότητας, δηλαδή, κάνουμε τον έλεγχο με μικρότερη ανοχή σε εσφαλμένη απόρριψη της H_0, π.χ. με $\alpha = 0.01$. Στην περίπτωση πρέπει $c \geq 1557.7$ Επαληθεύστε το ως άσκηση.
- Η σταθερά c ονομάζεται κρίσιμη τιμή ή όριο απόρριψης γιατί με βάση αυτή κρίνεται αν μια τιμή της στατιστικής συνάρτησης T είναι ακραία ή όχι.

Χρήση του p-value.

- Με δεδομένο ότι η $H_0 : \mu = 1500$ είναι αληθής, υπολογίζουμε την πιθανότητα να εμφανισθεί η τιμή $\bar{X} = 1550$ Kgr που εμφανίσθηκε στο δείγμα ή κάποια μεγαλύτερή της (δηλαδή, προς την κατεύθυνση της H_1).
- Ζητάμε την πιθανότητα $P(\bar{X} \geq 1550|\mu = 1500)$ και επειδή γνωρίζουμε την κατανομή της $\bar{X}$ έχουμε,

$$P(\bar{X} \geq 1500|\mu = 1500) = P\left(\frac{\bar{X} - 1500}{175/\sqrt{50}} \geq \frac{1550 - 1500}{175/\sqrt{50}}\right)$$
$$= P(Z \geq 2.02) = 1 - \Phi(2.02) = 0.0217$$

- Αυτή είναι η p-value δηλ. η πιθανότητα να εμφανισθεί η τιμή της στατιστικής συνάρτησης ελέγχου που εμφανίσθηκε στο παράδειγμά μας ή κάποια πιο μακριά (πιο ακραία), προς την κατεύθυνση της H_1, δεδομένου ότι η H_0 είναι αληθής.
- Επομένως, όσο πιο μικρή είναι η p-value τόσο ισχυρότερες ενδείξεις εναντίον της H_0 προκύπτουν από το συγκεκριμένο τυχαίο δείγμα.

Στατιστικοί έλεγχοι υποθέσεων για τη μέση τιμή, μ, ενός πληθυσμού.

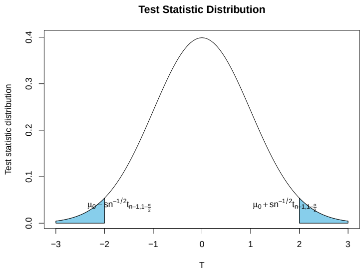
- Συμπερασματικά, η p-value δείχνει πόσο πιθανή είναι η εμφάνιση του αποτελέσματος που πήραμε δεδομένου ότι ισχύει η H_0.

- Από τους πιο συνηθισμένους ελέγχους είναι αυτός της υπόθεσης, $H_0 : \mu = \mu_0$, δηλαδή, της υπόθεσης ότι η άγνωστη μέση τιμή, μ ενός πληθυσμού έχει τιμή μ_0.
- Ειδικότερα, θα δώσουμε τη στατιστική συνάρτηση ελέγχου στις ακόλουθες περιπτώσεις όπου ο πληθυσμός του οποίου ελέγχουμε τη μέση τιμή ακολουθεί αντίστοιχα,
  1. κανονική κατανομή με γνωστή διασπορά
  2. κανονική κατανομή με άγνωστη διασπορά
  3. οποιαδήποτε κατανομή με άγνωστη διασπορά και το μέγεθος δείγματος είναι μεγάλο.

Κανονικός πληθυσμός με γνωστή διασπορά.

- Έστω τυχαίο δείγμα $X_1, X_2, \ldots, X_n$ από ένα πληθυσμό που ακολουθεί κανονική κατανομή με γνωστή διασπορά σ^2 και μέση τιμή $\mu = \mu_0$ (ελεγχόμενη), οπότε $X_i \sim N(\mu_0, \sigma^2), i = 1, 2, \ldots, n$
- Ως γνωστόν $\bar{X} \sim N(\mu_0, n^{-1}\sigma^2)$ άρα $\sigma^{-1}(\bar{X} - \mu_0)\sqrt{n} = Z \sim N(0, 1)$
- Αφού σ^2 γνωστή, η Z μπορεί να υπολογιστεί από το δείγμα.
- Με επίπεδο σημαντικότητας α απορρίπτουμε την $H_0 : \mu = \mu_0$

  - Έναντι της $H_1 : \mu > \mu_0$ για $\bar{X} \geq \mu_0 + \sigma n^{-1/2} z_{1-\alpha}$
  - Έναντι της $H_1 : \mu < \mu_0$ για $\bar{X} \leq \mu_0 - \sigma n^{-1/2} z_{1-\alpha}$
  - Έναντι της $H_1 : \mu \neq \mu_0$ για $\bar{X} \leq \mu_0 - \sigma n^{-1/2} z_{\frac{\alpha}{2}}$ ή $\bar{X} \geq \mu_0 + \sigma n^{-1/2} z_{1-\frac{\alpha}{2}}$



- Στα επόμενα, ακολουθούν λεπτομέρειες που εξηγούν τον μηχανισμό βάση του οποίου δουλεύει το στατιστικό τεστ.

Επεξήγηση του μηχανισμού του ελέγχου υποθέσεων.

- Πως δουλεύει το τεστ:
  - Αποφασίσαμε ότι η $\bar{X}$ είναι κατάλληλος εκτιμητής του μέσου όρου γιατί είναι αμερόληπτος και έχει ελάχιστη διασπορά.
  - Επίσης ξέρουμε από τη θεωρία ότι η $\bar{X}$ ακολουθεί κανονική κατανομή (άθροισμα κανονικών κατανομών είναι κανονική) με συγκεκριμένη διασπορά και διακύμανση.
- Με βάση τα παραπάνω, ξέρω ότι για διαφορετικά δείγματα $X_1, X_2, \ldots, X_n$ θα παίρνω διαφορετικές τιμές για το $\bar{X}$ οι οποίες όμως θα κατανέμονται στον άξονα των x σύμφωνα με το σχήμα της $N(\mu_0, n^{-1}\sigma^2)$ (που είναι η κατανομή του $\bar{X}$).
  - Άρα περιμένω για κάποιο δείγμα για το οποίο το $\bar{X}$ να πέφτει μέσα στην περιοχή αποδοχής να δεχτώ την H_0 γιατί θεωρώ ότι η τιμή που πήρα δεν είναι ακραία, άρα αυτό που παρατηρώ μάλλον θα ξανασυμβεί.
  - Αντίθετα, αν η τιμή του $\bar{X}$ πέφτει στην περιοχή απόρριψης, τότε θεωρώ ότι η τιμή που πήρα δεν είναι τόσο συνηθισμένη (δεν είναι πιθανό να ξανασυμβεί), άρα η υπόθεση μου δεν ισχύει.
- Ακριβώς η ίδια λογική εφαρμόζεται και στα επόμενα, και σε κάθε στατιστικό τεστ.

Επεξήγηση του μηχανισμού του ελέγχου υποθέσεων.

- Αυτό που αλλάζει από τεστ σε τεστ είναι η συνάρτηση με την οποία ελέγχουμε την υπόθεση, άρα και η κατανομή της, άρα και οι περιοχές απόρριψης / αποδοχής
- Οι παρακάτω αρχές όμως είναι κοινές σε κάθε τεστ:
    - Η συνάρτηση με την οποία κάνουμε τον έλεγχο, εκφράζει την ιδιότητα που θέλουμε να ελέγξουμε όταν ισχύει η μηδενική υπόθεση, π.χ. $\bar{X} = \mu_0$.
    - Η ίδια συνάρτηση έχει την ιδιότητα όσο πιο πολύ απομακρύνομαι από την μηδενική τόσο πιο πολύ οι τιμές της γίνονται ακραίες σε σχέση με τις τιμές που παίρνω από υπό την μηδενική υπόθεση.
- Τα όρια της περιοχή απόρριψης / αποδοχής καθορίζονται από το πόσο σίγουροι θέλουμε να είμαστε για το αποτέλεσμα.
    - Το κοινό σημείο στον καθορισμό κάθε περιοχής αποδοχής / απόρριψης είναι τα ποσοστημόρια (πόσο % της κατανομής συμπεριλαμβάνουμε / αφήνουμε από έξω) της κατανομής ώστε να είμαστε σίγουροι για το αποτέλεσμα.
    - Π.χ. στον δίπλευρο έλεγχο του τεστ που μόλις είδαμε οποιαδήποτε τιμή της Z πέφτει μεταξύ του αρχικού α% και του τελικού $(1-\alpha)$% της κατανομής σημαίνει αποδοχή του τεστ.

Επεξήγηση του μηχανισμού του ελέγχου υποθέσεων.

- Γιατί όμως έχω άλλο διάστημα αποδοχής (π.χ. στον δίπλευρο έλεγχο) και άλλο στο σχήμα;
- Απάντηση: γιατί έχω δυο διαφορετικούς αλλά ισοδύναμους τρόπους με τον οποίο μπορώ να κάνω τον έλεγχο.
- Τρόπος 1: Εφόσον ξέρω ότι $Z = \sigma^{-1}(\bar{X} - \mu_0)\sqrt{n} \sim N(0,1)$ μπορώ αντί να υπολογίσω μόνο το $\bar{X}$, να υπολογίσω μια και καλή την ποσότητα $\sigma^{-1}(\bar{X} - \mu_0)\sqrt{n}$ η οποία όμως ξέρω ότι ακολουθεί την τυπική κανονική κατανομή.
- Σε αυτή την περίπτωση ότι αποτέλεσμα πάρω, το συγκρίνω κατευθείαν με τα ποσοστημόρια $z_\alpha, z_{1-\alpha}$ της κανονικής τα όρια των μπλε περιοχών στο σχήμα.
- Τρόπος 2: Υπολογίζω μόνο την τιμή της $\bar{X}$ - σε αυτή την περίπτωση έχουμε και πάλι ότι

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1) \Leftrightarrow \mu_0 - \frac{\sigma}{\sqrt{n}} z \leq \bar{X} \leq \mu_0 + \frac{\sigma}{\sqrt{n}} z$$

όπου και πάλι το z ορίζεται από το δεξί / αριστερό ποσοστημόριο που ορίζουν τα άκρα της κατανομής μέσα στα οποία είμαστε διατεθειμένοι να δεχθούμε την H_0

Επεξήγηση του μηχανισμού του ελέγχου υποθέσεων.

- Ποια η σχέση της περιοχής απόρριψης / αποδοχής με το p value;
  - Π.χ. για τον δίπλευρο έλεγχο το p value είναι το εμβαδόν της κατανομής που είναι έξω από το διάστημα $(z_{\alpha/2}, z_{1-\alpha/2})$.
  - Αυτό το γεγονός είναι σε συμφωνία με το ότι όσο πιο μικρή είναι η p-value τόσο ισχυρότερες οι ενδείξεις εναντίον της H_0 σύμφωνα με το συγκεκριμένο τυχαίο δείγμα.
- Ισοδύναμα,
$$\text{p value} = P(Z \leq z_{\alpha/2} \text{ ή } Z \geq z_{1-\alpha/2})$$
- Επανερχόμαστε τώρα στο κομμάτι του ελέγχου υποθέσεων για διαφορές παραμέτρους του πληθυσμού.
  - Η υπόθεση που κάναμε ότι η διασπορά, του πληθυσμού είναι γνωστή, δεν είναι μια ιδιαίτερα ρεαλιστική υπόθεση.
  - Στην πράξη, η διασπορά του πληθυσμού συνήθως είναι άγνωστη.
  - Οι δύο περιπτώσεις που ακολουθούν, αναφέρονται στο πώς εργαζόμαστε όταν η διασπορά του πληθυσμού είναι άγνωστη.

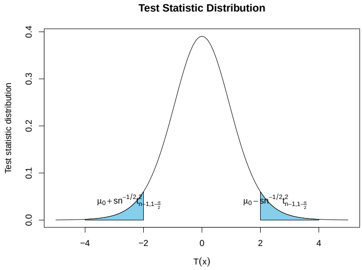
Κανονικός πληθυσμός με άγνωστη διασπορά.

- Έστω τυχαίο δείγμα $X_1, X_2, \ldots, X_n$ ανεξάρτητων, ισόνομων παρατηρήσεων από ένα πληθυσμό που ακολουθεί κανονική κατανομή με άγνωστη διασπορά σ^2 και μέση τιμή $\mu = \mu_0$.
- Έστω s^2 η δειγματική διασπορά. Έχουμε ότι

$$T = \frac{(\bar{X} - \mu_0)\sqrt{n}}{S} \sim t_{n-1} \text{ (Ορισμός } t \text{ κατανομής)}$$

- Με επίπεδο σημαντικότητας α απορρίπτουμε την $H_0 : \mu = \mu_0$

  - Έναντι της $H_1 : \mu > \mu_0$ για $\bar{X} \geq \mu_0 + sn^{-1/2} t_{n-1,1-\alpha}$
  - Έναντι της $H_1 : \mu < \mu_0$ για $\bar{X} \leq \mu_0 - sn^{-1/2} t_{n-1,1-\alpha}$
  - Έναντι της $H_1 : \mu \neq \mu_0$ για $\bar{X} \leq \mu_0 - sn^{-1/2} t_{n-1,1-\frac{\alpha}{2}}$ η $\bar{X} \geq \mu_0 + sn^{-1/2} t_{n-1,1-\frac{\alpha}{2}}$



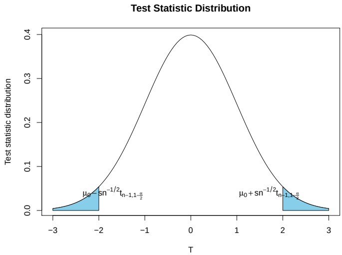
- η κατανομή t είναι γνωστή και ως κατανομή Student και οι σχετικοί έλεγχοι στατιστικών υποθέσεων ονομάζονται t-tests.

Πληθυσμός από άγνωστη $F(x)$ και $\sigma^2, n \to +\infty$.

- Έστω τ.δ. $X_1, X_2, \dots, X_n$ ανεξάρτητων, ισόνομων παρατηρήσεων από πληθυσμό οποιασδήποτε κατανομής με άγνωστη διασπορά σ^2 και μέση τιμή $\mu = \mu_0$.
- Θεωρούμε ότι $n \geq 30$. Έστω s^2 η δειγματική διασπορά.

$$T = \frac{(\bar{X} - \mu_0)\sqrt{n}}{s} \overset{\text{προσεγγιστικά}}{\to} Z \sim N(0,1)$$

- Με επίπεδο σημαντικότητας α απορρίπτουμε την $H_0 : \mu = \mu_0$
  - Έναντι της $H_1 : \mu > \mu_0$ για $\bar{X} \geq \mu_0 + \sigma n^{-1/2} z_{1-\alpha}$
  - Έναντι της $H_1 : \mu < \mu_0$ για $\bar{X} \leq \mu_0 - \sigma n^{-1/2} z_{1-\alpha}$
  - Έναντι της $H_1 : \mu \neq \mu_0$ για $\bar{X} \leq \mu_0 - \sigma n^{-1/2} z_{1-\frac{\alpha}{2}}$ ή $\bar{X} \geq \mu_0 + \sigma n^{-1/2} z_{1-\frac{\alpha}{2}}$



- Η κατανομή του T προσεγγίζεται από την κανονική. Όσο μεγαλύτερο το δείγμα, τόσο καλύτερη η προσέγγιση.
- Ερώτηση: Πότε προτιμάμε αυτόν τον έλεγχο και πότε το t-test;

Ασκήσεις

Άσκηση 1: Η θεωρητική μέση ετήσια παραγωγή ενός προϊόντος είναι 4000Kgr. Θέλουμε να ελέγξουμε αν ισχύει αυτός ο ισχυρισμός. Για το σκοπό αυτό και με βάση ένα σχέδιο τυχαίας δειγματοληψίας, επιλέγουμε ένα δείγμα 40 προϊόντων οπότε η μέση παραγωγή, βρέθηκε 3910Kgr με τυπική απόκλιση 250Kgr.

Απάντηση: Βάση στατιστικού ελέγχου θα ελέγξουμε, σε επίπεδο σημαντικότητας $\alpha = 0.05$, αν αυτό που παρατηρήθηκε στο δείγμα διαφέρει από τη θεωρητική μέση ετήσια παραγωγή X.

Ο δειγματικός μέσος είναι $\bar{X} = 3910$Kgr. Η υπόθεση που ελέγχουμε είναι: $H_0 : \mu = 4000$ έναντι της $H_1 : \mu \neq 4000$.

Δεν έχουμε κάποια ένδειξη για την κατανομή του δείγματος οπότε καλύτερα ταιριάζει η τρίτη περίπτωση - επίσης σε αυτό συνηγορεί και το γεγονός ότι $n = 40 > 30$ οπότε ισχύει η προσέγγιση από την κανονική κατανομή. Έτσι, η συνάρτηση ελέγχου είναι η

$$T = \frac{(\bar{X} - \mu_0)\sqrt{n}}{s} = \frac{(3910 - 4000)\sqrt{40}}{250} = -2.28 \overset{\text{προσεγγ.}}{\rightarrow} Z \sim N(0,1)$$

Με βάση την κρίσιμη περιοχή $|-2.28| > z_{0.05/2} = 1.96$ οπότε απορρίπτουμε την H_0 με επίπεδο σημαντικότητας $\alpha = 0.05$.

Ασκήσεις

Άσκηση 2: Από έναν πληθυσμό που ακολουθεί κανονική κατανομή, πήραμε ένα δείγμα μεγέθους $n = 9$, με $\bar{X} = 60$ μονάδες και $s = 12$ μονάδες. Θέλουμε να κάνουμε, σε επίπεδο σημαντικότητας $\alpha = 0.05$, τον έλεγχο της μηδενικής υπόθεσης : $H_0 : \mu = 65$ έναντι της εναλλακτικής: $H_1 : \mu \neq 65$.

Απάντηση: Σε αντίθεση με την Άσκηση 1, εδώ ξέρουμε ότι ο πληθυσμός είναι κανονικός (αλλά δεν μας δίνουν την διασπορά, μόνο τη δειγματική διασπορά) οπότε το πιο κατάλληλο είναι το t-test.
Η στατιστική συνάρτηση ελέγχου είναι η

$$T = \frac{(60 - 65)\sqrt{9}}{12} = -1.25 < t_{8,0.05/2} = -2.306$$

οπότε η τιμή της συνάρτησης είναι εκτός της περιοχής ελέγχου και έτσι με επίπεδο σημαντικότητας $\alpha = 0.05$, δεν απορρίπτουμε τη μηδενική υπόθεση.

Ερώτηση: Μη απορρίπτοντας την: $H_0 : \mu = 65$ αποδείξαμε άραγε ότι είναι αληθής;

Απάντηση: Όχι! Δεν αποδείξαμε ότι $\mu = 65$. Απλώς αποτύχαμε να απορρίψουμε την $H_0 : \mu = 65$. Γι΄ αυτό, στο συμπέρασμα δε γράψαμε ότι αποδεχόμαστε τη μηδενική υπόθεση αλλά ότι δεν την απορρίπτουμε.

Ασκήσεις

Άσκηση 3: Από έναν πληθυσμό με άγνωστη κατανομή και άγνωστη διασπορά, πήραμε ένα τυχαίο δείγμα μεγέθους $n = 36$. Από παλαιότερες έρευνες είναι γνωστό ότι η μέση τιμή του πληθυσμού είναι $\mu = 83$, όμως υπάρχουν υπόνοιες ότι έχει αλλάξει. Το δείγμα που πήραμε έδωσε, $\bar{X} = 86.2$ και $s = 10$. α) Να γίνει, σε επίπεδο σημαντικότητας $\alpha = 0.05$, κατάλληλος στατιστικός έλεγχος για τη μέση τιμή του πληθυσμού. β) Αν αλλαγή της μέσης τιμής σημαίνει μόνο αύξηση, αλλάζει κάτι στον έλεγχο που πρέπει να κάνουμε; Στο συμπέρασμα;

Απάντηση: Ο έλεγχος είναι προφανώς ο $H_0 : \mu = 83$ έναντι $H_1 : \mu \neq 83$. Η άσκηση δεν μας δίνει στοιχεία για την κατανομή του πληθυσμού αλλά μας δίνει μέγεθος δείγματος $n = 36 > 30$ οπότε είμαστε στην τρίτη περίπτωση. Έτσι,

$$|z| = \frac{(\bar{X} - \mu_0)\sqrt{n}}{s} = \frac{(86.2 - 83)\sqrt{36}}{10} = 1.92 < 1.96$$

οπότε σε επίπεδο σημαντικότητας $\alpha = 0.05$ η H_0 δεν απορρίπτεται. Στην περίπτωση β πρέπει να κάνουμε μονόπλευρο-δεξιόπλευρο έλεγχο: $H_0 : \mu = 83$ έναντι $H_1 : \mu > 83$. Η περιοχή απόρριψης είναι η $z \geq z_{1-0.05} = 1.645$. Εμείς έχουμε ότι $z = 1.92 \geq 1.645$, οπότε έχουμε στατιστικά σημαντικές αποδείξεις ότι η μέση τιμή έχει αυξηθεί.

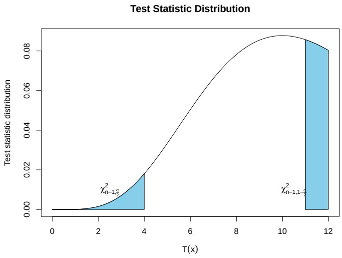
Έλεγχος υποθέσεων για την ισότητα διασποράς

- Έστω τυχαίο δείγμα $X_1, X_2, \ldots, X_n$, ανεξάρτητων, ισόνομων παρατηρήσεων από ένα πληθυσμό με κανονική κατανομή με διασπορά σ^2. Έστω s η δειγματική διασπορά. Η συνάρτηση

$$T = \frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1} \text{ ορισμός } \chi^2 : \sum Z_i^2, Z_i \sim N(\mu, \sigma^2)$$

- Για να ελέγξουμε την $H_0 : \sigma^2 = \sigma_0^2$, η τιμή της T μπορεί να υπολογιστεί από το δείγμα (αφού $\sigma^2 = \sigma_0^2$).
- Με επίπεδο σημαντικότητας α απορρίπτουμε την $H_0 : \sigma^2 = \sigma_0^2$:

  - Έναντι της $H_1 : \sigma^2 \geq \sigma_0^2$ για $(n-1)s^2\sigma^{-2} \geq \chi^2_{n-1,1-\alpha}$
  - Έναντι της $H_1 : \sigma^2 \leq \sigma_0^2$ για $(n-1)s^2\sigma^{-2} \leq \chi^2_{n-1,\alpha}$
  - Έναντι της $H_1 : \sigma^2 \neq \sigma_0^2$ για $(n-1)s^2\sigma^{-2} \geq \chi^2_{n-1,\alpha/2}$ η $(n-1)s^2\sigma^{-2} \leq \chi^2_{n-1,1-\alpha/2}$



Στατιστικός έλεγχος για τη διαφορά δυο μέσων τιμών: ανεξάρτητα δείγματα.

- Με δεδομένα δείγματα $X_1, \ldots, X_n$ και $Y_1, \ldots, Y_n$ από δύο πληθυσμούς, θέλουμε να κάνουμε έλεγχο της μηδενικής υπόθεσης $H_0 : \mu_X - \mu_Y = 0$ έναντι της εναλλακτικής $H_1 : \mu_X - \mu_Y \neq 0$.
- Ως συνάρτηση για τον έλεγχο φυσικά χρησιμοποιούμε την $T = \bar{X} - \bar{Y}$.
- Μπορούμε να προσεγγίσουμε την κατανομή της διαφοράς, $\bar{X} - \bar{Y}$ των αντίστοιχων δειγματικών μέσων, στις ακόλουθες περιπτώσεις:
  - όταν τα δείγματα προέρχονται από κανονικούς πληθυσμούς με γνωστές διασπορές
  - όταν τα δείγματα προέρχονται από κανονικούς πληθυσμούς με άγνωστες αλλά ίσες διασπορές
  - όταν τα δείγματα είναι μεγάλα και οι διασπορές των πληθυσμών είναι γνωστές (και οι πληθυσμοί όχι κατ' ανάγκη κανονικοί)
  - όταν τα δείγματα είναι μεγάλα και οι διασπορές των πληθυσμών είναι άγνωστες (και οι πληθυσμοί όχι κατ' ανάγκη κανονικοί).

(α) κανονικά δείγματα, γνωστές διασπορές.

- Με δεδομένα δείγματα από δύο κανονικούς πληθυσμούς $X_1, \ldots, X_{n_1} \sim N(\mu_1, \sigma_1^2)$ και $Y_1, \ldots, Y_n \sim N(\mu_2, \sigma_2^2)$, με γνωστές διασπορές, θέλουμε να κάνουμε έλεγχο της μηδενικής υπόθεσης $H_0 : \mu_X - \mu_Y = \delta$ έναντι της εναλλακτικής $H_1 : \mu_X - \mu_Y \neq \delta$.
- Η συνάρτηση για τον έλεγχο είναι η $T = \bar{X} - \bar{Y}$. Εφόσον σ_1^2, σ_2^2 γνωστά έχουμε

$$\bar{X} - \bar{Y} \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right) \Rightarrow \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \equiv z \sim N(0, 1).$$

Με επίπεδο σημαντικότητας α απορρίπτουμε την $H_0 : \mu_X - \mu_Y = \delta$

- Έναντι της $H_1 : \mu_X - \mu_Y > \delta$ για $z \geq z_{1-\alpha}$
- Έναντι της $H_1 : \mu_X - \mu_Y < \delta$ για $z \leq -z_{1-\alpha}$
- Έναντι της $H_1 : \mu_X - \mu_Y \neq \delta$ για $|z| \geq z_{1-\alpha/2}$

Ερώτηση: Γιατί $-z_{1-\alpha} = z_{\alpha}$;

(β) κανονικά δείγματα, άγνωστες, αλλά ίσες διασπορές.

- Με δεδομένα δείγματα $X_1, \ldots, X_{n_1} \sim N(\mu_1, \sigma_1^2)$ και $Y_1, \ldots, Y_{n_2} \sim N(\mu_2, \sigma_2^2)$, με άγνωστες αλλά ίσες διασπορές,
- Θέλουμε να κάνουμε έλεγχο της μηδενικής υπόθεσης $H_0 : \mu_X - \mu_Y = \delta$ έναντι της εναλλακτικής $H_1 : \mu_X - \mu_Y \neq \delta$.
- Μία εκτιμήτρια της κοινής διασποράς των δύο δειγμάτων είναι η

$$S^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2)}{n_1 + n_2 - 2}$$

Η στατιστική συνάρτηση που χρησιμοποιούμε για τον έλεγχο είναι

$$T = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n_1+n_2-2}.$$

Με επίπεδο σημαντικότητας α απορρίπτουμε την $H_0 : \mu_X - \mu_Y = \delta$

- Έναντι της $H_1 : \mu_X - \mu_Y > \delta$ για $T \geq t_{n_1+n_2-2,1-\alpha}$
- Έναντι της $H_1 : \mu_X - \mu_Y < \delta$ για $T \leq t_{n_1+n_2-2,\alpha}$
- Έναντι της $H_1 : \mu_X - \mu_Y \neq \delta$ για $|T| \geq t_{n_1+n_2-2,1-\frac{\alpha}{2}}$

(γ) μεγάλα ανεξάρτητα δείγματα, οποιοσδήποτε πληθυσμός.

- Με δεδομένα δείγματα $X_1, \ldots, X_{n_1}$ και $Y_1, \ldots, Y_{n_2}$, όπου $n_1, n_2 \to +\infty$,
- Αν τα σ_1^2, σ_2^2 είναι γνωστά τότε εφαρμόζεται κατά προσέγγιση ο έλεγχος της περίπτωσης α.
- Για άγνωστες διασπορές αν S_1^2, S_2^2 είναι οι αντίστοιχες δειγματικές διασπορές η συνάρτηση ελέγχου είναι η

$$d = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \overset{\text{προσεγγιστικά}}{\sim} N(0,1).$$

Με επίπεδο σημαντικότητας α απορρίπτουμε την $H_0 : \mu_X - \mu_Y = \delta$

- Έναντι της $H_1 : \mu_X - \mu_Y > \delta$ για $d \geq z_{1-\alpha}$
- Έναντι της $H_1 : \mu_X - \mu_Y < \delta$ για $d \leq z_{\alpha}$
- Έναντι της $H_1 : \mu_X - \mu_Y \neq \delta$ για $|d| \geq z_{1-\alpha/2}$

Ασκήσεις.

Άσκηση 1: Σε πρόσφατο δειγματοληπτικό έλεγχο, ένα δείγμα 100 παροχών νερού από το κέντρο της πόλης έδωσε μέση συγκέντρωση μολύβδου 36ppm με τυπική απόκλιση 6ppm και ένα δείγμα 90 παροχών από τα ανατολικά προάστια έδωσε μέση συγκέντρωση 34.1ppm με τυπική απόκλιση 5.9ppm. Να ελεγχθεί αν η διαφορά είναι στατιστικά σημαντική η αν προέκυψε προέκυψε τυχαία.

Απάντηση: Πρέπει να κάνουμε έλεγχο της $H_0 : \mu_X - \mu_Y = 0$ έναντι της $H_1 : \mu_X - \mu_Y \neq 0$. Τα δύο δείγματα μπορούν να θεωρηθούν ανεξάρτητα μεταξύ τους. Επίσης, οι κατανομές και οι διασπορές των πληθυσμών από τους οποίους προέρχονται τα δείγματα είναι άγνωστες.

Τα μεγέθη των δειγμάτων είναι μεγάλα ($n_X = 100 \geq 30$ και $n_Y = 90 \geq 30$) οπότε πρόκειται για έλεγχο της περίπτωσης (γ) και επομένως, σε επίπεδο σημαντικότητας $\alpha = 5\%$, έχουμε

$$|z| = \frac{\bar{X} - \bar{Y} - 0}{\sqrt{\frac{s_X^2}{n_X} + \frac{s_Y^2}{n_Y}}} = 2.2 > z_{1-0.05/2} = z_{1-0.025} = 1.96$$

οπότε σε επίπεδο σημαντικότητας $\alpha = 5\%$, απορρίπτουμε την H_0 δηλαδή έχουμε στατιστικά σημαντικές αποδείξεις ότι η μέση συγκέντρωση μολύβδου στο πόσιμο νερό διαφέρει στις δύο περιοχές της πόλης.

Ασκήσεις.

Άσκηση 2: Την τελευταία χρονιά, τα κτήματα 10 αγροτών από μία περιοχή είχαν μέση στρεμματική απόδοση 8.5Kgr/στρέμμα με τυπική απόκλιση 1.2Kgr/στρέμμα. Η μέση στρεμματική απόδοση στα κτήματα 15 αγροτών άλλης ήταν την τελευταία χρονιά 11Kgr/στρέμμα με τυπική απόκλιση 1.1Kgr/στρέμμα. Η διαφορά που παρατηρείται μεταξύ των δύο δειγμάτων είναι στατιστικά σημαντική ή μήπως οφείλεται στην τύχη;

Απάντηση: Έστω μ_A η μέση στρεμματική απόδοση από την πρώτη περιοχή και μ_B από τη δεύτερη. Ελέγχουμε την $H_0 : \mu_A - \mu_B = 0$ έναντι της $H_1 : \mu_A - \mu_B < 0$ με $\alpha = 5\%$. Τα δύο δείγματα μπορούν να θεωρηθούν ανεξάρτητα μεταξύ τους αλλά είναι μικρά και δεν γνωρίζουμε αν οι πληθυσμοί είναι κανονικοί (ούτε κάτι για τις διασπορές τους). Έλλειψη άλλης πληροφορίας θα χρησιμοποιήσουμε την περίπτωση (β) κάνοντας όμως δύο παραδοχές: τα δείγματα προέρχονται από κανονικούς πληθυσμούς με ίσες διασπορές.

$$t = \frac{\bar{X} - \bar{Y} - \delta}{s\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}} \leq -t_{n_A+n_B-2,0.05} = -5.37 < -1.714 = t_{23,0.05}$$

όπου $s = (n_A + n_B - 2)^{-1}((n_A - 1)s_A^2 + (n_B - 1)s_B^2) = 1.14$. Άρα η μέση στρεμματική απόδοση της ποικιλίας Α είναι μικρότερη από αυτή της Β.

Ασκήσεις.

Άσκηση 3: 5 άτομα ακολουθούν ένα διαιτολόγιο (Α) ενώ άλλα 5 ακολουθούν το διαιτολόγιο Β. Τρεις μήνες μετά η μέση μείωση του βάρους για το Α είναι $\mu_A = 3.1$Kgr με $s_A = 1.3$Kgr και για το Β, η μέση μείωση ήταν $\mu_B = 1.7$Kgr με $s_B = 1.2$Kgr. Το διαιτολόγιο Α είναι πιο αποτελεσματικό από το Β; ($\alpha = 5\%$).

Απάντηση: Ελέγχουμε την

$$H_0 : \mu_A - \mu_B = 0 \text{ έναντι της } H_1 : \mu_A - \mu_B > 0$$

Για τους ίδιους λόγους με την Άσκηση 2 είμαστε στην περίπτωση (β) οπότε

$$t = \frac{\bar{X} - \bar{Y} - \delta}{s\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}} = \frac{3.1 - 1.7 - 0}{1.25\sqrt{\frac{1}{5} + \frac{1}{5}}} = 1.77 < 1.860 = t_{8,1-0.05}$$

όπου

$$s^2 = (n_A + n_B - 2)^{-1}((n_A - 1)s_A^2 + (n_B - 1)s_B^2) = 1.57 \Rightarrow s = 1.25$$

Οπότε σε επίπεδο σημαντικότητας $\alpha = 5\%$ δεν απορρίπτουμε τη μηδενική υπόθεση ότι τα δύο διαιτολόγια έχουν την ίδια αποτελεσματικότητα.

(δ) εξαρτημένα δείγματα (ζεύγη παρατηρήσεων).

- Με δεδομένα δυο ισομεγέθη δείγματα $X_1, \ldots, X_n$ και $Y_1, \ldots, Y_n$,
- Για τα οποία θεωρούμε ότι τα $(X_i, Y_i), i = 1, \ldots, n$ είναι μεταξύ τους ανεξάρτητα ενώ εντός του ίδιου ζεύγους, τα X_i και Y_i είναι εξαρτημένα.
- Σχηματίζοντας τα ζεύγη $D_i = X_i - Y_i$, μπορούμε να εργαστούμε με ένα δείγμα το οποίο θεωρούμε ότι προέρχεται από πληθυσμό με μέση τιμή $\mu_D = \mu_X - \mu_Y$.
- Θέλουμε ισοδύναμα να ελέγξουμε $H_0 : \mu_D = \delta$. Αν οι διασπορές είναι άγνωστες, έχουμε

$$g = \frac{(\bar{D} - \delta)\sqrt{n}}{S_D}$$

Με επίπεδο σημαντικότητας α απορρίπτουμε την $H_0 : \mu_X - \mu_Y = \delta$

- Έναντι της $H_1 : \mu_D > \delta$ για $g \geq t_{n-1,1-\alpha}$
- Έναντι της $H_1 : \mu_D < \delta$ για $d \leq t_{n-1,\alpha}$
- Έναντι της $H_1 : \mu_D \neq \delta$ για $|g| \geq t_{n-1,1-\alpha/2}$

Άσκηση.

Επιλέγουμε 10 αγρούς σε δέκα διαφορετικές τοποθεσίες και κάθε αγρό τον χωρίζουμε σε δύο αγροτεμάχια ίδιου σχήματος και ίδιου εμβαδού. Στο ένα αγροτεμάχιο κάθε αγρού καλλιεργείται σιτάρι ποικιλίας A, και στο άλλο σιτάρι της ποικιλίας B τα οποία επιλέγονται με τυχαίο τρόπο. Υποθέτουμε ότι υπάρχουν οι ίδιες καλλιεργητικές συνθήκες και ίδιες συνθήκες συγκομιδής. Αν για κάθε ζεύγος η διαφορά των παρατηρήσεων είναι $d_i = 45, 30, 35, -40, 50, 50, 20, 35, 25, 30$ να ελεγχθεί η $H_0 : \mu_A - \mu_B = 0$ έναντι της $H_1 : \mu_A - \mu_B \neq 0$ όπου μ_A, μ_B η μέση απόδοση των A, B αντίστοιχα.

Απάντηση: Παρατηρούμε ότι $\bar{d} = 28$ και $s^2 = (n-1)^{-1} \sum_{i=1}^{10} (d_i - \bar{d})^2 = 673.33$. Οπότε η τιμή της στατιστικής συνάρτησης είναι

$$g = \frac{(28-0)\sqrt{10}}{\sqrt{673.33}} = 3.41 \geq t_{9,1-0.025} = 2.262$$

οπότε είμαστε μέσα στην περιοχή απόρριψης, και συνεπώς σε επίπεδο σημαντικότητας $\alpha = 0.05$ τα πειραματικά δεδομένα υποστηρίζουν ότι η μέση απόδοση της ποικιλίας A διαφέρει από τη μέση απόδοση της ποικιλίας B. Η πιθανότητα αυτό το συμπέρασμα να είναι λάθος είναι το πολύ 0.05.