

Ενότητα 3

Έλεγχος υπόθεσης. Σύγκριση μέσων τιμών

Εκτός από τις μέσες τιμές, τυπικές αποκλίσεις κλπ, θέλουμε να βρούμε κατά πόσον αυτές οι παρατηρούμενες τάσεις εξαρτώνται από συγκεκριμένες συνθήκες ή προϋποθέσεις. Δηλαδή διατυπώνουμε μια θεωρία ή νόμο που συνοψίζει αυτό που βλέπουμε, υπό μορφή **υπόθεσης**. Η υπόθεση αυτή πρέπει να ελεγχθεί σε σχέση με άλλες πιθανές εξηγήσεις. Η πιο απλή εξήγηση που φυσικά πρέπει να αποκλείσουμε πρώτη είναι τα αποτελέσματα να προέκυψαν κατά τύχη. Άρα, θα διατυπώσουμε μια αρχική "υπόθεση αναφοράς", τη **μηδενική υπόθεση** (null hypothesis - συνήθως γράφεται συμβολικά ως H_0) κατά την οποία τα αποτελέσματα που πήραμε ήταν τυχαία.

Ελέγχουμε κατά πόσον, δηλαδή *με ποια πιθανότητα* είναι δυνατό να συμβεί αυτό. Για το σκοπό αυτό, ορίζουμε ένα ποσοστό, ανάλογο με την εμπιστοσύνη που συζητήθηκε στην προηγούμενη ενότητα, αρκετά μικρό, π.χ. 5% ή 1%. Αν η πιθανότητα ικανοποίησης της μηδενικής υπόθεσης για το δείγμα μας είναι κάτω από αυτό το ποσοστό, τότε τα αποτελέσματα είναι **σημαντικά** (significant) και μπορούμε να πούμε ότι η δική μας **εναλλακτική** υπόθεση, υπερτερεί της υπόθεσης ότι προέκυψαν τυχαία.

Σύγκριση μέσων τιμών

Αν πάρουμε δύο δείγματα, υπολογίσουμε τις μέσες τιμές και τυπικές αποκλίσεις αυτών και μπορέσουμε να δείξουμε ότι οι δύο κατανομές προέρχονται από τον ίδιο πληθυσμό, τότε το αποτέλεσμα μπορεί να είναι σημαντικό. Εξάλλου, αν ένα δείγμα γνωστού πληθυσμού συγκριθεί με δείγμα που πιστεύεται ότι είναι από τον ίδιο πληθυσμό, θα ήταν σημαντικό να ελεγχθεί αυτή η υπόθεση της κοινής τους προέλευσης.

Έλεγχοι αυτού του είδους γίνονται με δύο τρόπους:

- υποθέτουμε ότι οι μέσες τιμές των δύο κατανομών διαφέρουν και ελέγχουμε πόσο λανθασμένη είναι η υπόθεση.
- πιο απλό είναι να υποθέσουμε ότι οι μέσες τιμές ταυτίζονται και να ελέγξουμε την ισχύ *αυτής* της υπόθεσης. Αυτή είναι η μηδενική υπόθεση για το συγκεκριμένο πρόβλημα.

Η διατύπωση της μηδενικής υπόθεσης σε συνοπτική μορφή, μπορεί να γραφεί έτσι:

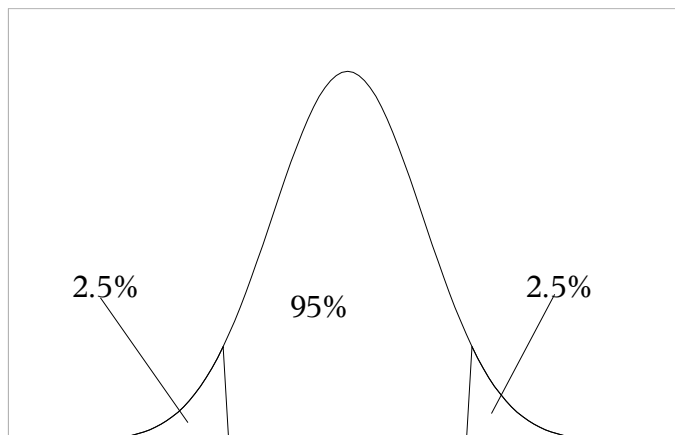
$$H_0: \mu_1 = \mu_2$$

ή $H_0: \mu_1 - \mu_2 = 0$

ή $H_0: \mu_2 - \mu_1 = 0$

Όπως και με την τιμή του μέσου ενός δείγματος, στην προηγούμενη ενότητα, έτσι και εδώ, μπορούμε να ορίσουμε ένα επίπεδο εμπιστοσύνης που απαιτούμε να ισχύει για να γίνει αποδεκτή η μηδενική υπόθεση, αλλιώς απορρίπτεται. Η μόνη διαφορά είναι ότι οι γνωστές μετασχηματισμένες ποσότητες (z , t) βασίζονται όχι στις μεμονωμένες μέσες τιμές αλλά στη διαφορά τους $\Delta\mu = \mu_1 - \mu_2$ η οποία κανονικά πρέπει να είναι μηδέν (ουσιαστικά, υποθέτουμε ότι η $\Delta\mu$ προέρχεται από μια κατανομή με κορυφή στο μηδέν).

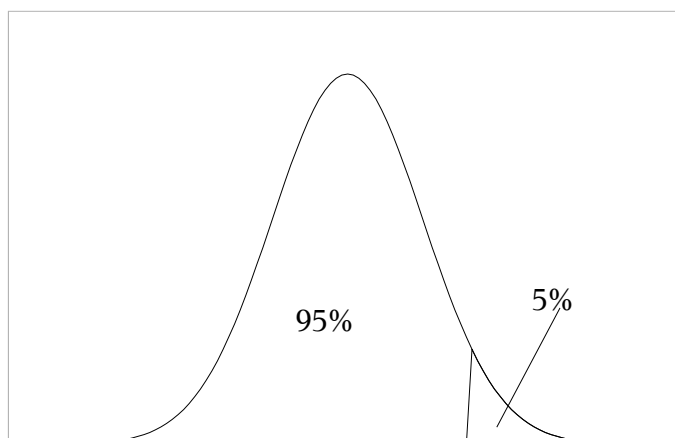
Ο παραπάνω έλεγχος λειτουργεί και προς τις δύο κατευθύνσεις, δηλαδή οι αποκλίσεις που διαψεύδουν την υπόθεση μπορεί να είναι είτε προς θετικές είτε προς αρνητικές τιμές. Έτσι, αν ορίσουμε 95% εμπιστοσύνη, το υπόλοιπο 5% ή 0.05 θα μοιράζεται στις δύο άκρες της ουράς της κατανομής που κάθε μία θα έχει εμβαδόν 0.025, όπως στο παρακάτω σχήμα:



Αυτού του είδους ο έλεγχος λέγεται **αμφίπλευρος** (two tailed test). Υπάρχει όμως και ο **μονόπλευρος** έλεγχος, είτε **από αριστερά** είτε **από δεξιά**. Στην περίπτωση π.χ. μονόπλευρου ελέγχου από δεξιά, η μηδενική υπόθεση γίνεται:

$$H_0: \mu_1 > \mu_2$$

και για απαίτηση εμπιστοσύνης 95% έχουμε την παρακάτω εικόνα:



Οι περιοχές εμβαδού 0.025 στα άκρα της ουράς, για τον αμφίπλευρο έλεγχο και η μία περιοχή με εμβαδό 0.05 για το μονόπλευρο, αποτελούν την **κρίσιμη περιοχή** κάθε περίπτωσης. Αν το στατιστικό ελέγχου οδηγήσει σε αυτή την περιοχή, θεωρούμε ότι η απόκλιση από το σενάριο της μηδενικής υπόθεσης είναι πολύ μεγάλη για να έχει προκύψει κατά τύχη.

Η εφαρμογή της μεθόδου σκιαγραφείται με ένα παράδειγμα χρήσης της κανονικής κατανομής. Στη συνέχεια θα αναφερθούν παραδείγματα και με άλλες κατανομές.

Παράδειγμα 1

Δίνονται δύο δείγματα με τα ακόλουθα χαρακτηριστικά:

	<u>Δείγμα 1</u>	<u>Δείγμα 2</u>
Μέση τιμή:	92.3	88.7
Μέγεθος δείγματος:	92	112
Τυπική Απόκλιση:	9.8	7.3

Θέλουμε να δούμε κατά πόσον οι δύο μέσες τιμές προέρχονται από την ίδια κατανομή. Αν ισχύει αυτό, τότε η παρατηρούμενη διαφορά τους είναι τυχαία και χωρίς σημασία, διαφορετικά είναι ένα αποτέλεσμα που έχει σημασία. Η μηδενική υπόθεση, λοιπόν, είναι:

$$H_0: \bar{x}_1 - \bar{x}_2 = 0$$

Άρα υποθέτουμε ότι η διαφορά μεταξύ των μέσων τιμών προέρχεται από μια κατανομή με μέση τιμή ίση με μηδέν.

Για να ελέγξουμε την υπόθεση θεωρούμε ότι τα δύο δείγματα είναι δείγματα της ίδιας κατανομής και τα χειριζόμαστε ως τέτοια, δηλαδή πρέπει να έχουν την ίδια μέση τιμή και τυπική απόκλιση. Αυτό ουσιαστικά σημαίνει ότι θεωρούμε τη διαφορά $\mu_1 - \mu_2$ ως μία μεταβλητή που έχει μέση τιμή 0 και συγκεκριμένη τυπική απόκλιση. Η τελευταία δίνεται από τη σχέση

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\sigma_{\bar{x}_1}^2 + \sigma_{\bar{x}_2}^2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Στη συγκεκριμένη περίπτωση θα βρούμε:

$$\sigma_{x_1 - x_2} = (9.8^2 / 91 + 7.3^2 / 111)^{1/2} = (1.05 + 0.48)^{1/2} = 1.24$$

Θεωρώντας κανονική κατανομή, υπολογίζουμε το $z = (92.3 - 88.7) / 1.24 = 2.9$

Ανατρέχοντας στους πίνακες βρίσκουμε ότι το συνολικό εμβαδόν (δηλ. το διπλάσιο από αυτό που δίνεται στον πίνακα λόγω συμμετρίας) είναι 0.976. Αυτό σημαίνει ότι υπάρχει διαφορά μεταξύ των μέσων τιμών με πιθανότητα 97.6%. Αν τα δείγματα ήταν από την ίδια κατανομή, τότε η παρατηρούμενη διαφορά, $z=2.9$, θα ήταν τυχαία διακύμανση. Αλλά η πιθανότητα να συμβεί αυτό θα ήταν $(100 - 97.6) / 2 = 1.2\%$ (διαιρούμε με το 2 γιατί θετικό z σημαίνει ότι είμαστε στη δεξιά ουρά).

Αν είχαμε ορίσει π.χ. εμπιστοσύνη 95%, η κρίσιμη περιοχή σε κάθε ουρά της κατανομής θα ήταν 2.5%. Αυτό σημαίνει ότι η παρατηρούμενη διαφορά των μέσων τιμών πρέπει να έχει πιθανότητα να προκύψει τυχαία, πάνω από 2.5% για να δεχτούμε ότι μπορεί να μην έχει σημασία. Με άλλα λόγια, η διαφορά είναι πολύ μεγάλη για να προέκυψε τυχαία (όπως είπαμε, υπάρχει μόνο 1.2% πιθανότητα να συμβαίνει κάτι τέτοιο). Άρα, η μηδενική υπόθεση απορρίπτεται με εμπιστοσύνη μεγαλύτερη από 95% και με αυτό το επίπεδο εμπιστοσύνης θεωρούμε ότι τα δείγματα προέρχονται από διαφορετικές κατανομές.

Άλλα προβλήματα

Το πώς ακριβώς διατυπώνουμε γενικότερα για άλλα προβλήματα, τη μηδενική υπόθεση εξαρτάται από τη φύση του προβλήματος. Π.χ. αν έχουμε δύο ενδεχόμενα, τότε αυτά μπορεί να είναι ανεξάρτητα, χωρίς να ευνοείται το ένα έναντι του άλλου, πράγμα που συνεπάγεται για το καθένα πιθανότητα $P = 50\%$. Στη συνέχεια δίνουμε παράδειγμα με χρήση της δυωνυμικής κατανομής. Σκοπός μας είναι να προσδιοριστεί μια αναλογία πληθυσμού (ποσοστό εμφάνισης) με συγκεκριμένο επίπεδο εμπιστοσύνης.

Παράδειγμα 2

Διεξάγεται έρευνα για το αν οι μύγες των φρούτων έλκονται από το μέλι ή από το ξύδι. Διατυπώνεται η εξής μηδενική υπόθεση H_0 : έλκονται εξίσου, άρα $P = 50\%$. Το δείγμα είναι n μύγες που πιάνονται και το πείραμα έγκειται στην καταγραφή του ποσοστού p των εντόμων που πιάστηκαν με το μέλι. Η εναλλακτική υπόθεση είναι $P > 50\%$. Ορίζουμε εμπιστοσύνη 2% (δηλαδή δεχόμαστε τα αποτελέσματα ως σημαντικά αν συμβαίνουν με πιθανότητα 98% ή είναι κάτω από 2% πιθανότητα να συμβεί το παρατηρούμενο αποτέλεσμα κατά τύχη).

Ο έλεγχος που θα κάνουμε είναι μονόπλευρος, δηλαδή θα θεωρήσουμε την κρίσιμη περιοχή ως την "ουρά" της κατανομής με τις περισσότερες επιτυχίες (συλλήψεις μυγών από το μέλι). Από

πίνακες της διωνυμικής κατανομής για $p = 50\%$ και μέγεθος δείγματος $n = 15$, βρίσκουμε ότι για 11 επιτυχίες (να πιαστούν με μέλι) η πιθανότητα είναι 1.76% ενώ για 12 έως 15, το άθροισμα των πιθανοτήτων είναι $0.42\% < 2\%$. Από την άλλη για 10 επιτυχίες υπερβαίνει από μόνο του το 5% (είναι 0.059). Το άθροισμα των πιθανοτήτων για επιτυχίες από 11 έως 15 είναι λίγο πάνω από το όριο που θέσαμε (2.18%). Άρα, πρέπει να πιαστούν τουλάχιστον 12 μύγες με μέλι για να πούμε ότι είναι κάτω από 2% πιθανό να γίνει κατά τύχη.

Το σύνολο των πιθανών αποτελεσμάτων που είναι σημαντικά αποτελεί την κρίσιμη περιοχή. Αν πιάστηκαν π.χ. 13 μύγες τότε είναι μέσα στην κρίσιμη περιοχή και η εναλλακτική υπόθεση δικαιώνεται έναντι της μηδενικής.

Παραμένοντας στο ίδιο παράδειγμα, εξετάζουμε μια συνεχή μεταβλητή, και συγκεκριμένα τη διάρκεια της ζωής των εντόμων. Έστω ότι από την προηγούμενη πείρα μας γνωρίζουμε πως οι μύγες ζούνε γύρω στις 12 μέρες με τυπική απόκλιση πληθυσμού $\sigma = 2$. Αυτή είναι η νέα μηδενική υπόθεση $H_0: \mu = 12$.

Έστω η εναλλακτική υπόθεση H_1 : μέση τιμή $\mu \neq 12$.

Παίρνουμε μεγάλα δείγματα, των 50, και βρίσκουμε τυπική απόκλιση της κατανομής δειγματοληψίας ίση με $2/\sqrt{50}=0.28$ (βλ. Κεντρικό Οριακό Θεώρημα).

Θέτουμε επίπεδο σημαντικότητας 1% , δηλαδή εμπιστοσύνη 99% και βρίσκουμε από τον πίνακα της τυπικής κανονικής κατανομής $z = \pm 2.58$

Έστω ότι από το δείγμα βρίσκουμε 12.5 μέρες. Αυτό αντιστοιχεί σε $z = 1.77$ άρα είναι μέσα στο διάστημα εμπιστοσύνης της μηδενικής υπόθεσης και επιβεβαιώνουμε έτσι, ότι οι φρουτόμυγες ζούνε γύρω στις 12 μέρες.

Σημαντική σημείωση: απαιτείται προσεκτικός **σχεδιασμός πειράματος** ώστε να αποκλειστούν άλλες πηγές επίδρασης στα αποτελέσματα, π.χ. αν ο άνεμος έσπρωχνε τις μύγες προς το μέλι. Το ίδιο ισχύει γενικότερα για κάθε είδους παρατήρηση.

Η σημασία του σχεδιασμού πειράματος φαίνεται σε ορισμένες πάγιες πρακτικές που ακολουθούνται όταν γίνονται έρευνες πάνω σε φάρμακα: όταν δοκιμάζεται ένα νέο φάρμακο σε ομάδα ασθενών, διεξάγεται και το λεγόμενο τυφλό πείραμα (δηλαδή μία ισάριθμη ομάδα παίρνει ένα ακίνδυνο ψευδοφάρμακο, "placebo", χωρίς τη δραστική ουσία) για να εξακριβωθεί μέσω της σύγκρισης, ο υποκειμενικός παράγοντας (αυθυποβολή των ασθενών που νομίζουν ότι θεραπεύονται). Αλλά επειδή το ίδιο πρόβλημα υπάρχει και με τους ερευνητές που μπορεί να υποβάλλουν στον εαυτό τους την ιδέα ότι βλέπουν βελτίωση την οποία μπορεί ακόμη και να τη μεταδώσουν στους ασθενείς που παίρνουν το φάρμακο, πιο σωστή μεθοδολογία είναι το λεγόμενο διπλό τυφλό πείραμα όπου ούτε οι ερευνητές ξέρουν ποια ομάδα παίρνει το αληθινό φάρμακο.

Υπάρχουν δύο είδη σφαλμάτων που θέλουμε να δούμε την πιθανότητά τους:

- Τύπου I (λέγεται και "α") = εσφαλμένη θεωρία που επιβεβαιώνουμε κατά τύχη
- Τύπου II (λέγεται και "β") = ορθή θεωρία που διαψεύδουμε κατά τύχη.

Όσο μεγαλώνει η πιθανότητα για το ένα μικραίνει για το άλλο (επίπεδο σημαντικότητας, μέγεθος δείγματος, διασπορά πληθυσμού). Επομένως αυτά τα δύο είναι ανταγωνιστικά και δε μπορούμε να τα μειώσουμε συγχρόνως. Θα πρέπει να αποφασίσουμε κατά περίπτωση.

Σχετικά παραδείγματα περιλαμβάνουν τα εξής:

- Φαρμακευτική έρευνα - καλύτερα να κρατήσω ένα φάρμακο που δε μου κάνει για περαιτέρω έρευνα παρά να πετάξω ένα χρήσιμο (απαιτείται μείωση σφάλματος τύπου "β")
- Διατύπωση νέων θεωριών - καλύτερα να διαψεύσω κατά τύχη μια ορθή θεωρία παρά να

επαληθεύσω κατά τύχη μια λανθασμένη (απαιτείται μείωση σφάλματος τύπου "α")

Διαφορά Μέσων

Επανερχόμαστε στη σύγκριση μέσων τιμών από δύο δείγματα, για να συζητήσουμε το θέμα γενικότερα. Για τις διαφορές μέσων τιμών από δύο διαφορετικά δείγματα υπάρχουν δύο δυνατά σενάρια:

- τα δείγματα μπορούν να διαταχθούν σε αντιστοιχία ένα-με-ένα, π.χ. ζεύγη παρατηρήσεων τύπου "πριν/μετά".
- ανεξάρτητα δείγματα (ίσως και με διαφορετικό μέγεθος) όπου δε γίνεται ή δεν έχει νόημα η ένα με ένα αντιστοίχιση των παρατηρήσεων

Πρώτη περίπτωση: δείγματα διατάξιμα κατά ζεύγη.

Παράδειγμα 3:

Συγκεκριμένη δίαιτα δοκιμάζεται σε δείγμα από $n = 49$ άτομα. Καταγράφουμε τα βάρη των ατόμων πριν και μετά και παίρνουμε τις διαφορές τους: $\Delta B = \text{Βάρος (μετά)} - \text{Βάρος (πριν)}$. Αν η δίαιτα δεν είχε αποτέλεσμα τότε περιμένω $\mu(\Delta B) = 0$. Αυτή είναι και η *μηδενική υπόθεση*. Αυτή θα τη συγκρίνω με μια **εναλλακτική** υπόθεση: *έχασαν βάρος, άρα $\mu(\Delta B) < 0$*

Αν ισχύει η μηδενική υπόθεση περιμένω κανονική κατανομή των ΔB_i . Αλλά η τυπική απόκλιση, σ , της κατανομής είναι άγνωστη, άρα θα βασιστώ σε εκτίμησή της βάσει της τυπικής κατανομής του δείγματος. $\sigma_{\bar{x}} = s / \sqrt{n}$ με τη βοήθεια του οποίου βρίσκω το z (αφού μέγεθος δείγματος > 30 , επιτρέπεται να χρησιμοποιήσω την κανονική κατανομή)

Έστω επίπεδο εμπιστοσύνης $\alpha = 0.05$

Βρίσκω κρίσιμη περιοχή $z \leq -1.65$

Έστω ότι $n = 49$, $\bar{x} = -5$ και $s = 3.5$ επομένως $\sigma_{\bar{x}} = 3.5/7 = 0.5$ με αποτέλεσμα

$$z = (-5 - 0) / 0.5 = -10$$

το οποίο είναι σαφώς μέσα στην κρίσιμη περιοχή, άρα η δίαιτα έφερε αποτέλεσμα.

Αν το δείγμα ήταν κάτω των 30 ατόμων θα χρησιμοποιούσαμε κατανομή του t υποθέτοντας κανονική κατανομή του πληθυσμού.

Παράδειγμα 3: σφάλμα τύπου II.

Επειδή τα επίπεδα εμπιστοσύνης είναι λίγο πολύ αυθαίρετα, υπάρχει ο κίνδυνος να απορριφθούν παρατηρήσεις που ισχύουν, ως τυχαίες. Χρειάζεται προσοχή και διεύρυνση των δειγμάτων αν υπάρχει σχετική υποψία. Π.χ. παρατηρήσεις σε ζεύγη διδύμων που έχουν ανατραφεί χωριστά (σε οικογένεια και ίδρυμα) για την επίπτωση στο δείκτη ευφυΐας IQ έδωσε τα παρακάτω αποτελέσματα:

Ζεύγη	Οικογένεια	Ίδρυμα	Διαφορά
1	105	95	-10
2	95	83	-12
3	103	103	0
4	98	96	-2
5	103	97	-6

Διατυπώνονται, η μηδενική υπόθεση: $\mu = 0$ (δεν υπάρχει διαφορά στο δείκτη ευφυΐας) και η εναλλακτική υπόθεση: $\mu \neq 0$, ενώ ορίζεται επίπεδο σημαντικότητας $\alpha = 5\%$.

Το δείγμα είναι μικρό (< 30). Υποθέτοντας κανονική κατανομή πληθυσμού, η κατανομή του t για το συγκεκριμένο επίπεδο σημαντικότητας, δίνει κρίσιμη περιοχή $|t| \geq 2.78$ ($df = 5 - 1 = 4$ και μοιράζοντας το 0.05 στις δύο ουρές της κατανομής βρίσκω $P(4; 0.025) = 2.776 \approx 2.78$)

Υπολογίζοντας τη μέση τιμή και τη διασπορά της διαφοράς των μέσων τιμών βρίσκω ότι $\bar{x} = -6, s = 5.1 \Rightarrow t = (\bar{x} - \mu) / s \sqrt{n} = -6 / 2.3 = 2.61$, δηλαδή εκτός της κρίσιμης περιοχής, άρα το αποτέλεσμα κρίνεται ως μη σημαντικό.

ΑΛΛΑ το δείγμα είναι *πολύ* μικρό και η απόρριψη είναι *οριακή*. Αν πάρουμε εμπιστοσύνη $\alpha = 10\%$ βρίσκουμε $|t| \geq 2.13$ που περιλαμβάνει το δείγμα. Με την πιο αυστηρή εμπιστοσύνη υπάρχει κίνδυνος να απορρίψουμε μια σωστή θεωρία!

Δεύτερη περίπτωση: ανεξάρτητα δείγματα.

Το πιο κλασικό παράδειγμα είναι αυτό του πραγματικού δείγματος συγκρινόμενο με τυφλό δείγμα. Εδώ δεν έχει νόημα να κάνουμε αντιστοιχία και να πάρουμε διαφορές γιατί δεν υπάρχει συγκεκριμένη αντιστοιχία.

Πώς εργαζόμαστε: Οι παρατηρήσεις δίνουν δύο μέσες τιμές μ_1, μ_2 . Διατυπώνουμε τη μηδενική υπόθεση $H_0: \mu_1 = \mu_2$ και την εναλλακτική υπόθεση $H_1: \mu_1 \neq \mu_2$ ή $\mu_1 > \mu_2$ ή $\mu_1 < \mu_2$, ανάλογα με τη θεωρία μας. Θέτουμε επίπεδο σημαντικότητας $\alpha = 1\%$ ή 5% (ή όσο θέλουμε).

Χρησιμοποιούμε z ή t , ανάλογα με το μέγεθος του δείγματος και την “κανονικότητα” του πληθυσμού, αλλά χρησιμοποιούμε τους εξής ορισμούς:

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad \text{ή} \quad t = \frac{\bar{x}_1 - \bar{x}_2}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad \text{όπου} \quad s = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

Παράδειγμα 4:

Στον πίνακα δίνονται στοιχεία για το ύψος φυτών που καλλιεργούνται με και χωρίς συγκεκριμένο λίπασμα.

	n	$\bar{x}$	s
1	100	27	5
2	100	29	4

Βρίσκουμε $z = 3.13$. Αν $\alpha = 1\%$, τότε $|z| \leq 2.33$ και αν η εναλλακτική είναι $\mu_1 < \mu_2$, τότε βρισκόμαστε στην κρίσιμη περιοχή.

Παράδειγμα 5:

Έστω ότι προμηθευόμαστε ανταλλακτικά από διαφορετικούς προμηθευτές για λογαριασμό μιας βιομηχανικής μονάδας όπου εργαζόμαστε. Θέλουμε να δούμε αν υπάρχουν σημαντικές διαφορές απόδοσης, ποιότητας κλπ για να αποφασίσουμε με ποιον θα συνεχίσουμε να συνεργαζόμαστε στο μέλλον.

	n	$\bar{x}$	s
1	50	150	10
2	100	153	5

Διατυπώνουμε τη μηδενική υπόθεση $H_0: \mu_1 = \mu_2$ και την εναλλακτική $H_1: \mu_1 \neq \mu_2$ με επίπεδο σημαντικότητας $\alpha = 1\%$.

Η κρίσιμη περιοχή βρίσκεται ότι είναι $|z| \geq 2.58$

Από τα δεδομένα βρίσκουμε $z = 2.0$ που δεν είναι στην κρίσιμη περιοχή άρα η διαφορά κρίνεται μη σημαντική.

Παραδοχές ή προϋποθέσεις που πρέπει να ισχύουν για να εφαρμοστεί η παραπάνω μεθοδολογία για ανεξάρτητα δείγματα, είναι:

- κανονική κατανομή *αμφότερων πληθυσμών*
- ίση τυπική απόκλιση $\sigma_1 = \sigma_2 = \sigma$ αμφότερων πληθυσμών

Φυσικά, δε γνωρίζω τα χαρακτηριστικά των πληθυσμών με ακρίβεια γιατί αν τα γνώριζα δε θα χρειαζόμουν τα δείγματα. Οπότε καταφεύγω σε εκτίμηση, π.χ. κάνω ιστόγραμμα των δειγμάτων. Αν το ένα είναι ασύμμετρο ή $s_1 \gg s_2$ τότε πιθανότατα παραβιάζονται οι παραδοχές.

Παρόμοια εργαζόμαστε και για έλεγχο με το t κατά *Student*: προϋποτίθενται κανονικά δείγματα, με ίσες διασπορές, $\sigma_1 = \sigma_2$. Στην πράξη, θέλουμε δείγματα παρόμοιου μεγέθους χωρίς μεγάλη ασυμμετρία στην κατανομή τους.

Διαφορές Διασπορών

Πολλές φορές θέλουμε να ξέρουμε αν οι διακυμάνσεις διαφέρουν σημαντικά από δείγμα σε δείγμα επειδή π.χ. οι μεγάλες διακυμάνσεις είναι ανεπιθύμητες. Για παράδειγμα, εργαζόμαστε σε μια βιομηχανία που παράγει κάποιο υλικό και θέλουμε να προμηθευτούμε πρώτη ύλη της οποίας η σύσταση έχει διακυμάνσεις που μπορούμε να ανεχθούμε μέχρι ένα όριο πάνω από το οποίο θα επιταχυνθούν ανεπιθύμητες αντιδράσεις που θα αλλοιώσουν το τελικό προϊόν. Για το σκοπό αυτό, συγκρίνουμε δείγματα διαφορετικών προμηθευτών για να αποφανθούμε για τις διαφορές των διασπορών τους.

Η γενική μεθοδολογία για τα προβλήματα αυτού του είδους συνίσταται στο να ορίσουμε το λόγο

$$F = \frac{s_1^2}{s_2^2} \quad \text{όπου } s_1 > s_2.$$

Αν η κατανομή του πληθυσμού είναι κανονική, αποδεικνύεται ότι ο λόγος F ακολουθεί τη λεγόμενη "κατανομή του F ". Αυτή δίνεται σε πίνακες με γραμμές και στήλες ανάλογα με τους βαθμούς ελευθερίας $df_i = n_i - 1$ όπου σε κάθε θέση του πίνακα δίνεται η τιμή της ουράς (δηλαδή ανάποδα απ' ο,τι για την κανονική κατανομή) για επίπεδο σημαντικότητας 5% και 1%.

Σε περίπτωση που δεν ενδιαφερόμαστε ειδικά για $\sigma_1 > \sigma_2$ (μία ουρά) αλλά γενικά για $\sigma_1 \neq \sigma_2$ (δύο ουρές) τότε θα χρησιμοποιήσουμε το διπλάσιο του ποσοστού εμπιστοσύνης ή σημαντικότητας για να πάμε στον πίνακα, π.χ. αν θέσουμε $\alpha = 1\%$ θα πάμε στην αντίστοιχη τιμή για τα δεδομένα df_1, df_2 αλλά θα θεωρήσουμε ότι το αποτέλεσμα δίνεται με σημαντικότητα 2% να προκύψει τυχαία.