

ΚΕΦΑΛΑΙΟ 19

ΕΛΕΓΧΟΙ ΥΠΟΘΕΣΕΩΝ ΓΙΑ ΤΗΝ ΣΥΓΚΡΙΣΗ ΜΕΣΩΝ ΤΙΜΩΝ ΚΑΙ ΑΝΑΛΟΓΙΩΝ ΔΥΟ ΚΑΝΟΝΙΚΩΝ ΠΛΗΘΥΣΜΩΝ

Όταν ενδιαφερόμαστε να συγκρίνουμε δύο πληθυσμούς, η φυσιολογική προσέγγιση είναι να προσπαθήσουμε να συγκρίνουμε τις μέσες τιμές ή τις τιμές κάποιου άλλου μέτρου θέσης των δύο αυτών πληθυσμών.

Σύμφωνα με όσα έχουμε δει μέχρι τώρα, η στατιστική διαδικασία που ενδείκνυται είναι ο έλεγχος στατιστικών υποθέσεων που θα αναφέρεται στις μέσες τιμές, ή κάποιες αναλογίες, των δύο υπό μελέτη πληθυσμών.

Η μεθοδολογία αυτή αποτελεί μια επέκταση των όσων έχουμε δει μέχρι τώρα για ελέγχους υποθέσεων ενός πληθυσμού.

ΣΥΓΚΡΙΣΗ ΜΕΣΩΝ ΤΙΜΩΝ ΔΥΟ ΚΑΝΟΝΙΚΩΝ ΠΛΗΘΥΣΜΩΝ

Συνοπτικά μπορούμε να πούμε τα εξής:

Αν έχουμε δύο κανονικούς πληθυσμούς με παραμέτρους (μ_X, σ_X^2) και (μ_Y, σ_Y^2) αντίστοιχα, οι υποθέσεις που ενδεχομένως ενδιαφερόμαστε να ελέγξουμε μπορεί να έχουν την παρακάτω μορφή,

$$H_0 : \mu_X = \mu_Y \quad (\mu \equiv \mu_X - \mu_Y = 0)$$

$$H_1 : \mu_X < \mu_Y \quad (\mu < 0) \quad H_1 : \mu_X \neq \mu_Y \quad (\mu \neq 0) \quad H_1 : \mu_X > \mu_Y \quad (\mu > 0)$$

όπου $\mu = \mu_X - \mu_Y$.

Αυτές είναι οι δυνατές περιπτώσεις για ελέγχους απλών υποθέσεων.

Για την περίπτωση συνθέτων υποθέσεων θα έχουμε τις περιπτώσεις,

$$H_0 : \mu_X \geq \mu_Y \quad (\mu \geq 0)$$

$$H_0 : \mu_X \leq \mu_Y \quad (\mu \leq 0)$$

$$H_1 : \mu_X < \mu_Y \quad (\mu < 0)$$

$$H_1 : \mu_X > \mu_Y \quad (\mu > 0)$$

Η επιλογή της στατιστικής συνάρτησης ελέγχου συνδέεται με το κατά πόσον τα τυχαία δείγματα που θα χρησιμοποιηθούν είναι ανεξάρτητα ή οι παρατηρήσεις τους ενός “συνδέονται” με τις

παρατηρήσεις του άλλου. Όπως και στο πρόβλημα της σύγκρισης μέσω τιμών πληθυσμών μέσω διαστημάτων εμπιστοσύνης, κάνουμε διάκριση των δύο αυτών περιπτώσεων στην συνέχεια. Συγκεκριμένα, οι τέσσερις πρώτες ενότητες που ακολουθούν (Α, Β, Γ, Δ) βασίζονται στην χρησιμοποίηση ανεξαρτήτων δειγμάτων, ενώ η πέμπτη (Ε) ενότητα στην χρησιμοποίηση δειγμάτων που έχουν μια συγκεκριμένη μορφή εξάρτησης (οι παρατηρήσεις επιλέγονται κατά ζεύγη).

Α. Περίπτωση Γνωστών Διακυμάνσεων (Ανεξάρτητα Δείγματα)

Σύμφωνα με όσα έχουμε δει μέχρι τώρα, ως στατιστική συνάρτηση ελέγχου θα χρησιμοποιήσουμε στην περίπτωση αυτή είτε την συνάρτηση $\bar{X} - \bar{Y}$ (που είναι η αμερόληπτη εκτιμήτρια της διαφοράς $\mu_X - \mu_Y$) είτε, ισοδύναμα, την αντίστοιχη τυποποιημένη συνάρτηση,

$$Z = \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sqrt{\sigma_X^2/n + \sigma_Y^2/m}}$$

Κάτω από την μηδενική υπόθεση, η στατιστική συνάρτηση ελέγχου θα γίνει,

$$Z_0 = \frac{\bar{X} - \bar{Y}}{\sqrt{\sigma_X^2/n + \sigma_Y^2/m}}$$

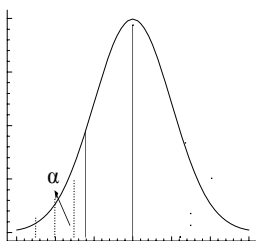
Έτσι, για τις διάφορες περιπτώσεις των εναλλακτικών υποθέσεων, όπως προκύπτει και από τα σχήματα που ακολουθούν, έχουμε

$$H_0 : \mu = \mu_X - \mu_Y = 0$$

$$H_1 : \mu = \mu_X - \mu_Y < 0$$

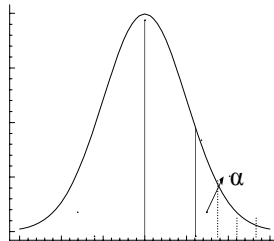
$$H_1 : \mu = \mu_X - \mu_Y > 0$$

$$H_1 : \mu = \mu_X - \mu_Y \neq 0$$

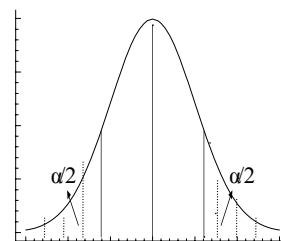


$$-Z_{1-\alpha}$$

Συγκεκριμένα,



$$Z_{1-\alpha}$$



$$-Z_{1-\alpha/2} \quad Z_{1-\alpha/2}$$

α) Αν $H_1: \mu = \mu_X - \mu_Y < 0$, τότε απορρίπτουμε την μηδενική υπόθεση σε επίπεδο σημαντικότητας α αν

$$Z_0 < -Z_{1-\alpha}$$

ή, ισοδύναμα, αν

$$\bar{X} - \bar{Y} < -Z_{1-\alpha} \sqrt{\sigma_X^2/n + \sigma_Y^2/m}$$

β) Αν $H_1: \mu = \mu_X - \mu_Y > 0$, τότε απορρίπτουμε την μηδενική υπόθεση σε επίπεδο σημαντικότητας α αν

$$Z_0 > Z_{1-\alpha}$$

ή, ισοδύναμα, αν

$$\bar{X} - \bar{Y} > Z_{1-\alpha} \sqrt{\sigma_X^2/n + \sigma_Y^2/m}$$

γ) Αν $H_1: \mu = \mu_X - \mu_Y \neq 0$, τότε απορρίπτουμε την μηδενική υπόθεση σε επίπεδο σημαντικότητας α αν

$$Z_0 < -Z_{1-\alpha/2} \quad \text{ή} \quad Z_0 > Z_{1-\alpha/2}$$

ή, ισοδύναμα, αν

$$\bar{X} - \bar{Y} < -Z_{1-\alpha/2} \sqrt{\sigma_X^2/n + \sigma_Y^2/m} \quad \text{ή} \quad \bar{X} - \bar{Y} > Z_{1-\alpha/2} \sqrt{\sigma_X^2/n + \sigma_Y^2/m}$$

Σημείωση 1: Η έννοια του παρατηρούμενου επιπέδου σημαντικότητας (της τιμής p) επεκτείνεται και στην περίπτωση αυτή με προφανή τρόπο.

Σημείωση 2: Στην περίπτωση συνθέτων υποθέσεων [$H_0: \mu_X \geq \mu_Y$ ($\mu \geq 0$) ή $H_0: \mu_X \leq \mu_Y$ ($\mu \leq 0$)], οι έλεγχοι γίνονται με τον ίδιο τρόπο με την διαφορά ότι, όπως και στους ελέγχους που αναφέρονται σε έναν πληθυσμό που εξετάσαμε, ο ορισμός του α και της p -τιμής αναφέρεται στο μέγιστο των πιθανοτήτων που ορίζουν τις έννοιες αυτές στις απλές υποθέσεις ως προς τις διάφορες τιμές του $\mu_X - \mu_Y$ κάτω από την μηδενική υπόθεση.

B. Περίπτωση Αγνώστων Ίσων Διακυμάνσεων ($\sigma_X^2 = \sigma_Y^2$)

(Ανεξάρτητα Δείγματα)

Ακολουθώντας την διαδικασία που χρησιμοποιήσαμε στα διαστήματα εμπιστοσύνης, κάνουμε χρήση της σταθμισμένης εκτιμήτριας της διασποράς S_p^2 .

Η στατιστική συνάρτηση ελέγχου θα είναι και πάλι η $\bar{X} - \bar{Y}$ ή ισοδύναμα, η

$$T = \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sqrt{S_p^2/n + S_p^2/m}}$$

η οποία, κάτω από τη μηδενική υπόθεση, γίνεται

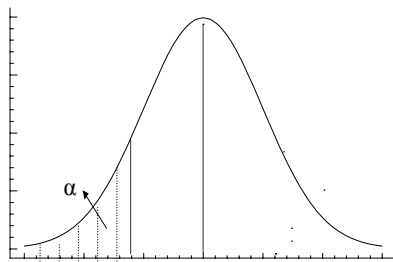
$$T_0 = \frac{\bar{X} - \bar{Y}}{\sqrt{S_p^2/n + S_p^2/m}}$$

Έτσι, για τις τρεις μορφές εναλλακτικής υπόθεσης γράφουμε, όπως προκύπτει και από τα σχήματα που ακολουθούν,

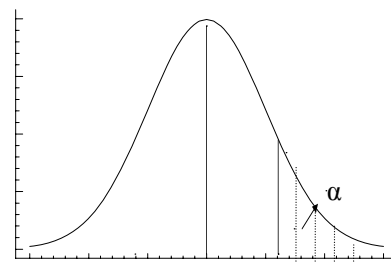
$$H_0 : \mu = \mu_X - \mu_Y = 0$$

$$H_1 : \mu = \mu_X - \mu_Y < 0$$

$$H_1 : \mu = \mu_X - \mu_Y > 0$$

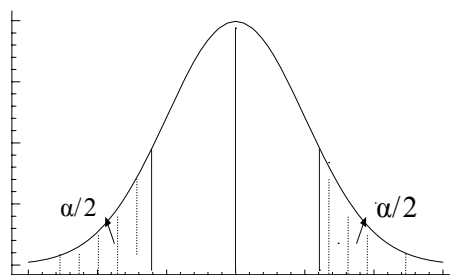


$$-t_{n+m-2, 1-\alpha}$$



$$t_{n+m-2, 1-\alpha}$$

$$H_1 : \mu = \mu_X - \mu_Y \neq 0$$



$$-t_{n+m-2, 1-\alpha/2} \quad t_{n+m-2, 1-\alpha/2}$$

Για κάθε μία από τις περιπτώσεις αυτές έχουμε αντίστοιχα,

α) Αν $H_1 : \mu = \mu_X - \mu_Y < 0$, τότε απορρίπτουμε την μηδενική υπόθεση σε επίπεδο σημαντικότητας α αν

$$T_0 < -t_{n+m-2, 1-\alpha}$$

ή, ισοδύναμα, αν

$$\bar{X} - \bar{Y} < -t_{n+m-2, 1-\alpha} \sqrt{S_p^2/n + S_p^2/m}$$

β) Αν $H_1 : \mu = \mu_X - \mu_Y > 0$, τότε απορρίπτουμε την μηδενική υπόθεση σε επίπεδο σημαντικότητας α αν

$$T_0 > t_{n+m-2, 1-\alpha}$$

ή, ισοδύναμα, αν

$$\bar{X} - \bar{Y} > t_{n+m-2, 1-\alpha} \sqrt{S_p^2/n + S_p^2/m}$$

γ) Αν $H_1 : \mu = \mu_X - \mu_Y \neq 0$, τότε απορρίπτουμε την μηδενική υπόθεση σε επίπεδο σημαντικότητας α αν

$$T_0 < -t_{n+m-2, 1-\alpha/2} \quad \text{ή} \quad T_0 > t_{n+m-2, 1-\alpha/2}$$

ή, ισοδύναμα, αν

$$\bar{X} - \bar{Y} < -t_{n+m-2, 1-\alpha/2} \sqrt{S_p^2/n + S_p^2/m} \quad \text{ή}$$

$$\bar{X} - \bar{Y} > t_{n+m-2, 1-\alpha/2} \sqrt{S_p^2/n + S_p^2/m}$$

Σημείωση: Η προηγηθείσα συμπερασματολογία ισχύει έστω και αν οι υπό μελέτη πληθυσμοί δεν ακολουθούν την κανονική κατανομή με την προϋπόθεση ότι τα μεγέθη των δειγμάτων από τους δύο πληθυσμούς είναι αρκετά μεγάλα.

Γ. Περίπτωση Αγνώστων Ανίσων Διακυμάνσεων ($\sigma_X^2 \neq \sigma_Y^2$)

Μεγάλα (Ανεξάρτητα) Δείγματα

Όπως και στα διαστήματα εμπιστοσύνης, αν τα μεγέθη n και m των δειγμάτων είναι μεγάλα, χρησιμοποιούμε τις εκτιμήτριες S_X^{*2} και S_Y^{*2} των διακυμάνσεων σ_X^2 και σ_Y^2 των δύο πληθυσμών (αντίστοιχα

τις S_X^2, S_Y^2) και αξιοποιούμε το γεγονός ότι, για μεγάλα δείγματα, η στατιστική συνάρτηση ελέγχου

$$Z = \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sqrt{S_X^{*2}/n + S_Y^{*2}/m}}$$

$$= \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sqrt{S_X^2/(n-1) + S_Y^2/(m-1)}}$$

ακολουθεί, σύμφωνα με το κεντρικό οριακό θεώρημα, κατά προσέγγιση την $N(0,1)$ κατανομή και, επομένως, κάτω από την μηδενική υπόθεση, η στατιστική συνάρτηση ελέγχου θα είναι

$$Z_0 = \frac{\bar{X} - \bar{Y}}{\sqrt{S_X^{*2}/n + S_Y^{*2}/m}}$$

$$= \frac{\bar{X} - \bar{Y}}{\sqrt{S_X^2/(n-1) + S_Y^2/(m-1)}}$$

Οι έλεγχοι γίνονται όπως και στην περίπτωση Α.

Δ. Περίπτωση Αγνώστων Ανίσων Διακυμάνσεων ($\sigma_X^2 \neq \sigma_Y^2$)

Μικρά (Ανεξάρτητα) Δείγματα

Όταν τα δεδομένα του προβλήματος δεν επιτρέπουν την υπόθεση της ισότητας των διακυμάνσεων των πληθυσμών ούτε είναι δυνατόν να χρησιμοποιηθούν μεγάλα δείγματα, δεν μπορούμε να χρησιμοποιήσουμε καμμία από τις μεθόδους που αναπτύξαμε προηγουμένως. Στην περίπτωση αυτή, αντιμετωπίζουμε το πρόβλημα που στη Στατιστική ονομάζεται πρόβλημα των Behrens-Fisher.

Όπως είδαμε και στην ανάπτυξη της θεωρίας των διαστημάτων εμπιστοσύνης, στην περίπτωση αυτή χρησιμοποιείται και πάλι η στατιστική συνάρτηση

$$T = \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sqrt{S_X^{*2}/n + S_Y^{*2}/m}} = \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sqrt{S_X^2/(n-1) + S_Y^2/(m-1)}}$$

που, όμως στην περίπτωση αυτή, έχει βαθμούς ελευθερίας

$$v \approx \frac{\left(\frac{S_X^{*2}}{n} + \frac{S_Y^2}{m} \right)^2}{\frac{\left(\frac{S_X^2}{n} \right)^2}{n-1} + \frac{\left(\frac{S_Y^2}{m} \right)^2}{m-1}}$$

Ο αριθμός v είναι πάντοτε μικρότερος ή ίσος από το $n+m-2$, από τον αριθμό δηλαδή των βαθμών ελευθερίας που χρησιμοποιούμε στην συνήθη στατιστική συνάρτηση t με την σταθμισμένη διασπορά. Προφανώς, όσο τα μεγέθη των δειγμάτων αυξάνουν το v αυξάνει επίσης και η κατανομή t συγκλίνει προς την κανονική κατανομή, οπότε μπορούμε να χρησιμοποιήσουμε κατά προσέγγιση την μέθοδο που αναπτύξαμε προηγουμένως.

Η μέθοδος αυτή ονομάζεται πολλές φορές και προσέγγιση του Satterthwaite.

Επομένως, στην περίπτωση αυτή, ο κανόνας απόφασης διαμορφώνεται ως εξής:

α) Αν $H_1 : \mu = \mu_X - \mu_Y < 0$, τότε απορρίπτουμε την μηδενική υπόθεση σε επίπεδο σημαντικότητας α αν

$$T_0 < -t_{v, 1-\alpha}$$

ή, ισοδύναμα, αν

$$\bar{X} - \bar{Y} < -t_{v, 1-\alpha} \sqrt{S_X^{*2}/n + S_Y^{*2}/m}$$

β) Αν $H_1 : \mu = \mu_X - \mu_Y > 0$, τότε απορρίπτουμε την μηδενική υπόθεση σε επίπεδο σημαντικότητας α αν

$$T_0 > t_{v, 1-\alpha}$$

ή, ισοδύναμα, αν

$$\bar{X} - \bar{Y} > t_{v, 1-\alpha} \sqrt{S_X^{*2}/n + S_Y^{*2}/m}$$

γ) Αν $H_1 : \mu = \mu_X - \mu_Y \neq 0$, τότε απορρίπτουμε την μηδενική υπόθεση σε επίπεδο σημαντικότητας α αν

$$T_0 < -t_{v, 1-\alpha/2} \quad \text{ή} \quad T_0 > t_{v, 1-\alpha/2}$$

ή, ισοδύναμα, αν

$$\bar{X} - \bar{Y} < -t_{v, 1-\alpha} \sqrt{S_X^{*2}/n + S_Y^{*2}/m} \quad \text{ή}$$

$$\bar{X} - \bar{Y} > t_{v, 1-\alpha} \sqrt{S_X^{*2}/n + S_Y^{*2}/m}$$

όπου v δίνεται από τον τύπο που προηγήθηκε.

Ε. Περίπτωση Παρατηρήσεων κατά Ζεύγη

Στην περίπτωση αυτή, έχουμε n ζευγάρια παρατηρήσεων (X_i, Y_i) , $i=1,2,\dots,n$, τις οποίες μεταχειριζόμαστε με τρόπο ίδιο με αυτόν που χρησιμοποιήσαμε στην περίπτωση που η σύγκριση των μέσων τιμών των δύο πληθυσμών έγινε μέσω της κατασκευής ενός διαστήματος εμπιστοσύνης. Θεωρώντας δηλαδή και πάλι τις διαφορές παρατηρήσεων $D_i=X_i-Y_i$, $i=1,2,\dots,n$, οδηγούμαστε σε ένα πρόβλημα n παρατηρήσεων D_1, D_2, \dots, D_n . Κατά συνέπεια, το πρόβλημα της σύγκρισης των μέσων των δύο πληθυσμών με τον έλεγχο της υπόθεσης $H_0: \mu_x=\mu_y$ γίνεται ισοδύναμο με το πρόβλημα ελέγχου υπόθεσης για την μέση τιμή ενός πληθυσμού με μέση τιμή $\mu_D=\mu_x-\mu_y$. Συγκεκριμένα, η μηδενική υπόθεση γίνεται ισοδύναμη με την υπόθεση $H_0: \mu_D=0$, ο έλεγχος της οποίας μπορεί να γίνει, σύμφωνα με όσα αναπτύχθηκαν στο προηγούμενο κεφάλαιο, με βάση το δείγμα D_1, D_2, \dots, D_n και με την χρήση της ελεγχοσυνάρτησης

$$T_0 = \frac{\bar{D}}{S_D^*/\sqrt{n}} \sim t_{n-1}$$

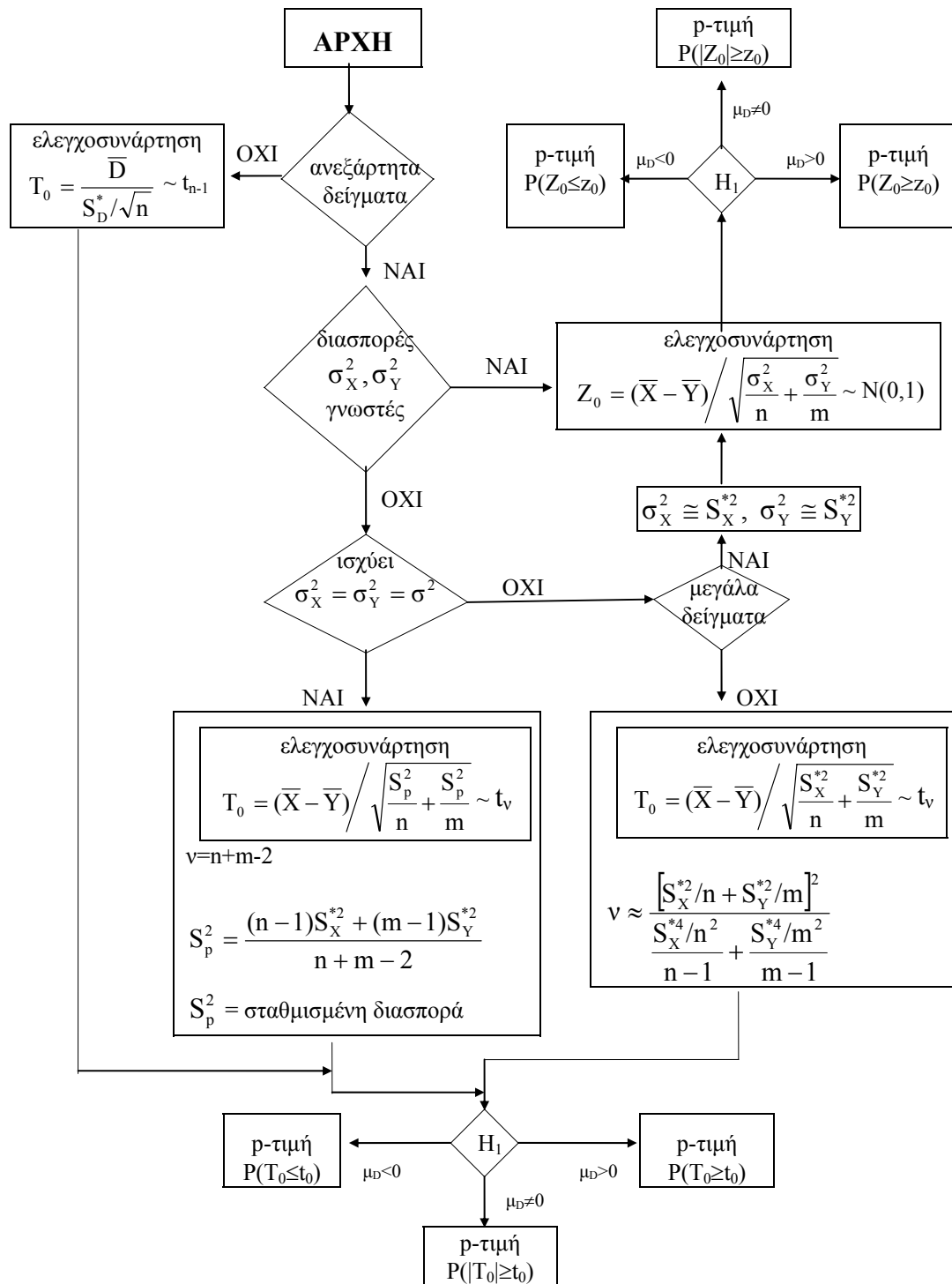
όπου

$$\bar{D} = \sum_{i=1}^n D_i/n \quad \text{και} \quad S_D^{*2} = \frac{\sum_{i=1}^n (D_i - \bar{D})^2}{n-1}$$

Σχηματική Παρουσίαση της Διαδικασίας Ελέγχου Υποθέσεων για την Διαφορά Μέσων Τιμών Κανονικών Πληθυσμών

Η διαδικασία που ακολουθείται για τον έλεγχο υποθέσεων για την διαφορά μέσων τιμών κανονικών πληθυσμών και οι ενέργειες που πρέπει να γίνουν ανάλογα με το εάν χρησιμοποιούνται ανεξάρτητα δείγματα ή ζεύγη παρατηρήσεων, λαμβάνοντας υπόψη και την σχέση των διακυμάνσεων των δύο πληθυσμών, μπορεί να παρασταθεί σχηματικά με το παρακάτω διάγραμμα.

**Έλεγχος της Υπόθεσης $H_0: \mu_D \equiv \mu_X - \mu_Y = 0$
(με την χρήση της p-τιμής)**



Μερικά Παραδείγματα

Παράδειγμα 1ο: Μια εταιρεία ενδιαφέρεται να υιοθετήσει μια νέα πολιτική για τους πωλητές της σύμφωνα με την οποία κάθε πωλητής θα παίρνει κάθε μήνα ένα δώρο ανάλογα με τις πωλήσεις που θα πραγματοποιεί. Η διεύθυνση της εταιρείας, προκειμένου να διερευνήσει τις προσδοκίες των ανδρών και γυναικών που εργάζονται στην εταιρεία επιλέγει δύο τυχαία δείγματα από $n = 40$ άνδρες και $m=40$ γυναίκες από τους οποίους ζητεί να προβλέψουν τις πρόσθετες μηνιαίες αποδοχές που θα έχουν αν υιοθετηθεί το προτεινόμενο σχήμα. Τα δειγματικά δεδομένα έδωσαν τις εξής πληροφορίες:

$$\begin{aligned}\bar{x} &= 31083 & s_X^{*2} &= 2312 & (\text{για τους άνδρες}) \\ \bar{y} &= 29745 & s_Y^{*2} &= 2569 & (\text{για τις γυναίκες})\end{aligned}$$

Παρέχουν τα δεδομένα αυτά ενδείξεις ότι υπάρχει διαφορά στο μέσο αναμενόμενο πρόσθετο εισόδημα μεταξύ ανδρών και γυναικών πωλητών με την υιοθέτηση της νέας πολιτικής ($\alpha = 0.05$);

Λύση: Ο κατάλληλος στατιστικός έλεγχος είναι,

$$H_0 : \mu_X = \mu_Y$$

$$H_1 : \mu_X \neq \mu_Y$$

Δεδομένου ότι τα μεγέθη των δειγμάτων είναι μεγάλα, μπορούμε να χρησιμοποιήσουμε την κανονική προσέγγιση και να έχουμε ως τιμή της στατιστικής συνάρτησης κάτω από τη μηδενική υπόθεση την

$$\begin{aligned}Z_0 &= \frac{\bar{x} - \bar{y}}{\sqrt{s_X^{*2}/n + s_Y^{*2}/m}} \\ &\cong \frac{31083 - 29745}{\sqrt{2312^2/40 + 2569^2/40}} = 2.45\end{aligned}$$

Δοθέντος ότι $Z_{.975} = 1.96$, με βάση τα στοιχεία των δειγμάτων, καταλήγουμε στο συμπέρασμα ότι η μηδενική υπόθεση θα πρέπει να απορριφθεί στο επίπεδο σημαντικότητας 5%.

Εάν θέλουμε να χρησιμοποιήσουμε την διαδικασία με την p-τιμή, θα έχουμε ότι

$$\begin{aligned} p\text{-τιμή} &= P(Z_0 \geq 2.45 \text{ ή } Z_0 \leq -2.45) \\ &= 2 P(Z_0 \geq 2.45) = 0.0142 \end{aligned}$$

Επομένως, η μηδενική υπόθεση θα απορρίπτεται για κάθε επίπεδο σημαντικότητας α μεγαλύτερο του 0.0142.

Παράδειγμα 2ο: Στο παράδειγμα που έχουμε ήδη εξετάσει με την εταιρεία παρασκευής και συσκευασίας ενός προϊόντος, ας υποθέσουμε ότι η εταιρεία αυτή έχει δύο εργοστάσια σε δύο διαφορετικές τοποθεσίες και ενδιαφέρεται να ελέγξει αν υπάρχει διαφορά στην μέση ποσότητα του προϊόντος που περιέχουν οι συσκευασίες που προέρχονται από τα δύο εργοστάσια, έστω A και B. Τυχαία δειγματοληψία με την λήψη δύο ανεξαρτήτων δειγμάτων από τα δύο εργοστάσια δίνει τα αποτελέσματα που ακολουθούν.

Εργοστάσιο A	Εργοστάσιο B
$\bar{x} = 366.35\text{gr}$	$\bar{y} = 369.74\text{ gr}$
$s_x^{*2} = 16.71\text{ gr}$	$s_y^{*2} = 14.20\text{ gr}$
$n = 25$	$m = 20$

(έχουμε ήδη υποθέσει ότι το περιεχόμενο των συσκευασιών ακολουθεί την κανονική κατανομή. Υποθέτουμε επίσης ότι οι διασπορές του βάρους του περιεχομένου των συσκευασιών στα δύο εργοστάσια είναι ίσες)

Η υπόθεση που θέλουμε να ελέγξουμε είναι η,

$$H_0 : \mu_X = \mu_Y$$

$$H_1 : \mu_X \neq \mu_Y$$

Δίνεται ότι $\alpha = 0.01$.

Λύση: Σύμφωνα με τα στοιχεία που έχουμε, θα πρέπει να χρησιμοποιήσουμε την μέθοδο των σταθμισμένων διασπορών. Επομένως,

$$s_p^2 = \frac{24(16.71)^2 + 19(14.20)^2}{25 + 20 - 2}$$

$$= \frac{10532.538}{43} = 244.943$$

Επομένως, η τιμή της στατιστικής συνάρτησης ελέγχου κάτω από την μηδενική υπόθεση θα είναι,

$$t_0 = \frac{366.35 - 369.74}{\sqrt{244.943/25 + 244.943/20}}$$

$$= \frac{-3.39}{4.695} = -0.72$$

Από τους πίνακες, έχουμε ότι,

$$t_{43,0.05} = 2.6951$$

Επομένως, με βάση τα στοιχεία του δείγματος δεν απορρίπτουμε την μηδενική υπόθεση στο $\alpha = 0.01$ επίπεδο σημαντικότητας.

Αν θέλουμε να χρησιμοποιήσουμε την προσέγγιση με την p -τιμή, η πιθανότητα η στατιστική συνάρτηση ελέγχου κάτω από τη μηδενική υπόθεση να πάρει μια τιμή μικρότερη από το -0.72 με 43 βαθμούς ελευθερίας είναι, κατά προσέγγιση, σύμφωνα με τους πίνακες, μεταξύ 0.10 και 0.25. (Στην πραγματικότητα κάτι λιγότερο από 0.25 δοθέντος ότι εμβαδόν 0.25 αντιστοιχεί σε τιμή του t με 43 βαθμούς ελευθερίας ίση με -0.6802).

Δοθέντος ότι έχουμε αμφίπλευρη εναλλακτική υπόθεση, η τιμή p θα είναι μεταξύ 0.20 και 0.50. Επειδή $\alpha = 0.01$ είναι μικρότερο από 0.20, η μηδενική υπόθεση δεν απορρίπτεται στο $\alpha = 0.01$.

Παράδειγμα 3ο: Έστω ότι ένας ερευνητής αγοράς ενδιαφέρεται να μελετήσει την επίδραση που έχει στις πωλήσεις ενός προϊόντος το μέρος στο οποίο το προϊόν αυτό τοποθετείται μέσα στο κατάστημα. Συγκεκριμένα, ενδιαφέρεται να εξετάσει αν υπάρχει διαφορά στις πωλήσεις αν το προϊόν αυτό τοποθετείται κοντά στην έξοδο ή σε ένα

άλλο μέρος του καταστήματος δίπλα σε άλλα παρόμοια προϊόντα. Για να εξετάσει το πρόβλημα αυτό, ο ερευνητής επιλέγει ένα τυχαίο δείγμα 13 καταστημάτων ιδίου μεγέθους από την συγκεκριμένη αλυσίδα καταστημάτων και σε 7 από αυτά τοποθετεί το προϊόν κοντά στην έξοδο ενώ στα υπόλοιπα 6, στο σημείο όπου έχει και άλλα παρόμοια προϊόντα. Τα αποτελέσματα των πωλήσεων, σε αριθμό πακέτων του προϊόντος που πωλούνται ανά εβδομάδα, δίνονται στον πίνακα που ακολουθεί.

ΕΒΔΟΜΑΔΙΑΙΕΣ ΠΩΛΗΣΕΙΣ (σε αριθμό πακέτων)	
Τοποθέτηση στην έξοδο X	Τοποθέτηση στο εσωτερικό Y
107	90
153	83
82	86
158	94
141	89
87	93
119	
$\bar{x} = 121$ $s_X^{*2} = 945$ $n = 7$	$\bar{y} = 89.17$ $s_Y^{*2} = 17.37$ $m = 6$

Λύση: Υποθέτουμε ότι οι πωλήσεις ανά εβδομάδα του προϊόντος αυτού ακολουθούν την κανονική κατανομή. Δοθέντος ότι δεν έχουμε κάποια ένδειξη ότι οι διακυμάνσεις των δύο πληθυσμών είναι ίσες, θα χρησιμοποιήσουμε την μέθοδο των αγνώστων και ανίσων διακυμάνσεων. (Θα χρησιμοποιηθεί το $\alpha = 0.05$).

Η τιμή της στατιστικής συνάρτησης ελέγχου κάτω από την μηδενική υπόθεση θα είναι,

$$t_0 = \frac{121 - 89.17}{\sqrt{945/7 + 17.37/6}} = \frac{31.83}{11.743} = 2.71$$

Οι βαθμοί ελευθερίας στην περίπτωση αυτή θα είναι,

$$v \approx \frac{(945/7 + 17.37/6)^2}{\frac{(945/7)^2}{6} + \frac{(17.37/6)^2}{5}}$$

Από τους πίνακες της κατανομής t έχουμε,

$$t_{6,975} = 2.45$$

Δοθέντος ότι η στατιστική συνάρτηση ελέγχου κάτω από την μηδενική υπόθεση βρίσκεται στην κρίσιμη περιοχή, απορρίπτουμε την μηδενική υπόθεση στο $\alpha = 0.05$.

Λύση με το πακέτο SPSS

Ο έλεγχος υποθέσεων για την διαφορά μέσων τιμών κανονικών πληθυσμών με το στατιστικό πακέτο SPSS γίνεται ως εξής:

- Εισάγουμε τα δεδομένα και των δύο δειγμάτων σε κάποια μεταβλητή π.χ. VAR00001
- Δημιουργούμε μια μεταβλητή (π.χ. VAR00002) η οποία δηλώνει τον πληθυσμό από τον οποίο προέρχεται η κάθε παρατήρηση. Για παράδειγμα, στις παρατηρήσεις από τον πρώτο πληθυσμό, η μεταβλητή αυτή παίρνει την τιμή 1 και, στις παρατηρήσεις από τον δεύτερο πληθυσμό, παίρνει την τιμή 2
- Από την επιλογή Statistics, επιλέγουμε Compare means.
- Επιλέγουμε Independent samples T test
- Στο παράθυρο που εμφανίζεται, επιλέγουμε την μεταβλητή, τα δεδομένα, (VAR00001) και την τοποθετούμε στο πεδίο test variable(s)
- Στο πεδίο Grouping Variable, τοποθετούμε την μεταβλητή που αναφέρεται στον πληθυσμό από τον οποίο προέρχονται οι παρατηρήσεις (VAR00002).
- Πατάμε Define Groups και, στο παράθυρο που εμφανίζεται, γράφουμε την τιμή 1 στο πεδίο Group 1 και την τιμή 2 στο πεδίο Group 2.

- Στο πεδίο Test value, γράφουμε την τιμή για την οποία επιθυμούμε να ελέγξουμε αν ταυτίζεται με την μέση τιμή του πληθυσμού (δηλ. για το συγκεκριμένο παράδειγμα την τιμή 0.5) και πατάμε OK.

Τα αποτελέσματα που δίνει το SPSS είναι τα ακόλουθα

Group Statistics

	SAMPLE	N	Mean	Std. Deviation	Std. Error Mean
DATA	1.00	7	121.0000	30.7409	11.6190
	2.00	6	89.1667	4.1673	1.7013

Independent Samples Test

	Levene's Test for Equality of Variance		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
DATA	15.049	.003	2.501	11	.029	31.8333	12.7275	3.8203	59.8463
Equal variances not assumed			2.711	6.257	.034	31.8333	11.7428	3.3829	60.2838

Στον πρώτο από τους δύο παραπάνω πίνακες, δίνονται τα μεγέθη των δειγμάτων από τους δύο πληθυσμούς (N), οι δύο δειγματικοί μέσοι (Mean), οι δύο δειγματικές τυπικές αποκλίσεις (Std. Deviation) και τα δύο τυπικά σφάλματα του μέσου (Std. Error Mean).

Στον δεύτερο πίνακα, δίνονται αποτελέσματα για την περίπτωση που έχουμε υποθέσει ίσες διακυμάνσεις (Equal variances assumed), καθώς και για την περίπτωση άνισων διακυμάνσεων (Equal variances not assumed). Στην δεύτερη και στην τρίτη στήλη του πίνακα (Levene's Test for Equality of Variances), δίνεται η τιμή της στατιστικής συνάρτησης ελέγχου (F) και η p-τιμή (Sig.) του ελέγχου του Levene για ισότητα διασπορών. Η μηδενική υπόθεση του ελέγχου αυτού είναι ότι οι διασπορές των δύο πληθυσμών είναι ίσες και, συνεπώς, στις περιπτώσεις που αυτή δεν απορρίπτεται (δηλ. στις περιπτώσεις που η p-τιμή είναι μεγαλύτερη κάποιου επιπέδου σημαντικότητας π.χ. 0.05), μπορούμε να συμπεράνουμε ότι έχει νόημα η υπόθεση της ισότητας των διασπορών.

Επίσης, ο δεύτερος πίνακας δίνει (και για τις δύο περιπτώσεις) την τιμή της στατιστικής συνάρτησης ελέγχου (t), τους βαθμούς ελευθερίας (df), την p-τιμή του αμφίπλευρου ελέγχου (Sig. (2-tailed)), την διαφορά των μέσων των δύο δειγμάτων (Mean Difference), το τυπικό σφάλμα της διαφοράς των μέσων τιμών (Std. Error Difference) και ένα 95% διάστημα εμπιστοσύνης για την διαφορά των μέσων τιμών (95% Confidence Interval of the Difference).

Λύση με το πακέτο Minitab

Ο έλεγχος υποθέσεων για την διαφορά μέσων τιμών κανονικών πληθυσμών με το στατιστικό πακέτο Minitab γίνεται ως εξής:

- Στο παράθυρο Data, εισάγουμε τα δεδομένα από τον πρώτο πληθυσμό σε κάποια μεταβλητή π.χ. C1 και τα δεδομένα από τον δεύτερο πληθυσμό σε μια άλλη π.χ. C2 (Πρώτος τρόπος). Εναλλακτικά μπορούμε να εισαγάγουμε τις παρατηρήσεις και των δύο δειγμάτων σε μια μεταβλητή π.χ. C3 και τον πληθυσμό από τον οποίο προέρχεται η κάθε παρατήρηση σε μια άλλη π.χ. C4 (Δεύτερος τρόπος).
- Από την επιλογή Stat, επιλέγουμε Basic Statistics.
- Επιλέγουμε 2-Sample t...
- Αν έχουμε γράψει τα δεδομένα σύμφωνα με τον πρώτο τρόπο, επιλέγουμε Samples in different columns και στα πεδία First και Second επιλέγουμε τα ονόματα των δύο μεταβλητών που έχουμε δημιουργήσει (δηλ. C1 και C2). Εναλλακτικά, αν έχουμε γράψει τα δεδομένα σύμφωνα με τον δεύτερο τρόπο, επιλέγουμε Samples in one column. Στο πεδίο Samples, επιλέγουμε την μεταβλητή με τα δεδομένα (C3) και, στο πεδίο Subscripts, επιλέγουμε την μεταβλητή που δηλώνει τον πληθυσμό από τον οποίο προέρχεται η κάθε παρατήρηση.
- Στο πεδίο Alternative, δηλώνουμε την μορφή της εναλλακτικής υπόθεσης. Όταν αυτή είναι αμφίπλευρη (δηλ. $\mu_1 \neq \mu_2$), επιλέγουμε not equal. Αντίθετα, όταν είναι μονόπλευρη, επιλέγουμε είτε greater than είτε less than (ανάλογα με την μορφή της ανισότητας). Για το παράδειγμα μας, επιλέγουμε not equal.

- Στο πεδίο confidence level, δεν χρειάζεται να αλλάξουμε την τιμή που δίνει το πακέτο, αφού αυτή χρησιμοποιείται αποκλειστικά για τον υπολογισμό του διαστήματος εμπιστοσύνης.
- Τέλος, στο πεδίο Assume equal variances, καθορίζουμε αν θα υποθέσουμε ισότητα των διασπορών. Επειδή στο παράδειγμα μας υποθέτουμε άνισες διασπορές, δεν θα επιλέξουμε το συγκεκριμένο πεδίο.
- Πατάμε OK

Τα αποτελέσματα που δίνει το Minitab είναι τα ακόλουθα:

TWO SAMPLE T FOR C1 VS C2

	N	MEAN	STDEV	SE MEAN
C1	7	121.0	30.7	12
C2	6	89.17	4.17	1.7

95 PCT CI FOR MU C1 - MU C2: (3, 60.6)

TTEST MU C1 = MU C2 (VS NE): T= 2.71 P=0.035 DF= 6

Στην στήλη N, δίνεται ο αριθμός των παρατηρήσεων για τα δύο δείγματα, στην στήλη Mean, δίνονται οι δύο δειγματικοί μέσοι, στην στήλη Stdev, δίνονται οι δύο τυπικές αποκλίσεις και τέλος στην στήλη Se mean, δίνονται τα δύο τυπικά σφάλματα του μέσου. Στην συνέχεια, δίνεται ένα 95% διάστημα εμπιστοσύνης για την διαφορά των μέσων τιμών και στην τελευταία γραμμή, δίνεται η τιμή της στατιστικής συνάρτησης ελέγχου (T), η p-τιμή του ελέγχου (P) και οι βαθμοί ελευθερίας (DF).

Παράδειγμα 4ο: Δύο εργοστάσια A και B παράγουν ένα συγκεκριμένο όμοιο υλικό. Από κάθε εργοστάσιο ένα δείγμα (5 τεμαχίων από το A και 7 τεμαχίων από το B) στέλνονται σε ένα δημόσιο χημικό εργαστήριο. Εκεί μετριέται η αντοχή των κομματιών και προκύπτουν τα ακόλουθα αποτελέσματα:

Εργοστάσιο A: 21,0 21,5 21,7 21,0 22,2

Εργοστάσιο B: 20,3 21,1 21,4 21,1 20,8 20,4 20,8

Παρέχουν τα παραπάνω στατιστικά ισχυρές ενδείξεις ότι τα υλικά από τα δύο εργοστάσια διαφέρουν ως προς την αντοχή; (υποθέστε κανονικές κατανομές).

Λύση με το πακέτο SPSS

- Εισάγουμε τα δεδομένα και τον δύο δειγμάτων σε μια μεταβλητή π.χ. var00001
- Σε μια δεύτερη μεταβλητή, έστω var00002, εισάγουμε τις τιμές 1 και 2 ανάλογα με το αν η αντίστοιχη παρατήρηση της μεταβλητής var00001 προέρχεται από το πρώτο ή το δεύτερο δείγμα
- Από την επιλογή Analyze επιλέγουμε Compare Means και στην συνέχεια Independent-Samples T test
- Στο παράθυρο που εμφανίζεται επιλέγουμε την μεταβλητή με τις παρατηρήσεις (var00001) και την μεταφέρουμε στο πεδίο Test Variable(s). Επίσης επιλέγουμε την μεταβλητή που προσδιορίζει το δείγμα από το οποίο προέρχεται η κάθε παρατήρηση (var00002) και την μεταφέρουμε στο πεδίο Grouping Variable
- Επιλέγουμε Define Groups και δίνουμε την τιμή 1 στο πεδίο Group 1 και την τιμή 2 στο πεδίο Group 2 και επιλέγουμε Continue
- Επιλέγουμε OK

Τα αποτελέσματα που δίνει το πακέτο είναι τα ακόλουθα:

Group Statistics

	VAR00002	N	Mean	Std. Deviation	Std. Error Mean
VAR00001	1,00	5	21,4800	,5070	,2267
	2,00	7	20,8429	,3952	,1494

Independent Samples Test

	Levene's Test for Equality of Variances	t-test for Equality of Means								
		F		t		Sig.		95% Confidence Interval of the Difference		
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
VAR00001	Equal variances assumed	,309	,590	2,455	10	,034	,6371	,2596	,0588	1,2155
	Equal variances not assumed			2,347	7,308	,050	,6371	,2715	,0006	1,2737

Στον πρώτο από τους δύο πίνακες, δίνεται ο αριθμός των παρατηρήσεων κάθε δείγματος (N), καθώς και ο αριθμητικός μέσος (Mean), η τυπική απόκλιση (Std. Deviation) και το τυπικό σφάλμα εκτίμησης του μέσου (Std. Error Mean) για το κάθε δείγμα ξεχωριστά.

Στις δύο πρώτες στήλες του δεύτερου πίνακα, δίνεται η τιμή της στατιστικής συνάρτησης ελέγχου και το παρατηρούμενο επίπεδο σημαντικότητας του ελέγχου του Levene για την ισότητα των διασπορών των πληθυσμών από τους οποίους προέρχονται τα δύο δείγματα. Στις υπόλοιπες στήλες του ίδιου πίνακα, δίνονται τα αποτελέσματα του ελέγχου για την ισότητα των μέσων τιμών των δύο πληθυσμών. Στην πρώτη γραμμή του πίνακα, δίνονται τα αποτελέσματα αν υποθέσουμε ότι οι διασπορές των δύο πληθυσμών είναι ίσες, ενώ, στην δεύτερη γραμμή, δίνονται τα αντίστοιχα αποτελέσματα χωρίς την υπόθεση της ισότητας των διασπορών. Στις επόμενες στήλες, δίνεται η τιμή της στατιστικής συνάρτησης ελέγχου (t), οι βαθμοί ελευθερίας (df), το παρατηρούμενο επίπεδο σημαντικότητας (Sig. 2-tailed), η διαφορά των δύο δειγματικών μέσων (Mean Difference), το τυπικό σφάλμα εκτίμησης της διαφοράς των μέσων δηλ. ο παρονομαστής της στατιστικής συνάρτησης ελέγχου (Std. Error Difference). Στις δύο τελευταίες στήλες, δίνονται τα όρια του 95% διαστήματος εμπιστοσύνης για την διαφορά των μέσων.

Λύση με το πακέτο Minitab

Τα δεδομένα του παραδείγματος μπορούν να εισαχθούν στο πακέτο Minitab με δύο διαφορετικούς τρόπους. Συγκεκριμένα, μπορούν να εισαχθούν είτε σε δύο διαφορετικές στήλες (έστω C1 και C2), είτε σε μια στήλη (έστω C3) δημιουργώντας όμως και μια στήλη (έστω C4) που προσδιορίζει το δείγμα από το οποίο προέρχεται η κάθε παρατήρηση.

- Στην συνέχεια, πρέπει να ακολουθηθεί η παρακάτω διαδικασία:
- Επιλέγουμε Stat, Basic Statistics και 2-Sample t.
- Αν όλα τα δεδομένα έχουν εισαχθεί σε μια στήλη, επιλέγουμε Samples in one column και στην συνέχεια επιλέγουμε την στήλη C3 στο πεδίο Samples και την στήλη C4 στο πεδίο Subscripts.

Αντίθετα, αν τα δεδομένα έχουν εισαχθεί σε δύο διαφορετικές στήλες, επιλέγουμε Samples in different columns. Στο πεδίο First, επιλέγουμε την στήλη C1 και στο πεδίο Second, την στήλη C2.

- Στο πεδίο Alternative, δηλώνουμε την μορφή της εναλλακτικής υπόθεσης. Οι δυνατές επιλογές είναι φυσικά 3 (not equal, less than, greater than). Για το παράδειγμα μας, πρέπει να επιλέξουμε not equal.
- Στο πεδίο Confidence level, δηλώνουμε το επίπεδο εμπιστοσύνης, καθώς το πακέτο δίνει πάντοτε πέρα από τα αποτελέσματα για τον έλεγχο της υπόθεσης και το αντίστοιχο διάστημα εμπιστοσύνης.
- Αν υποθέσουμε ότι οι διασπορές είναι ίσες, τότε κάνουμε κλικ στο τετράγωνο δίπλα από την επιλογή Assume equal variances.
- Επιλέγουμε OK

Στην περίπτωση που έχουμε υποθέσει ίσες διασπορές, το πακέτο δίνει τα ακόλουθα αποτελέσματα:

Two Sample T-Test and Confidence Interval				
Two sample T for C1				
C2	N	Mean	StDev	SE Mean
1	5	21,480	0,507	0,23
2	7	20,843	0,395	0,15

95% CI for mu (1) - mu (2): (0,06; 1,22)
T-Test mu (1) = mu (2) (vs not =): T = 2,45 P = 0,034 DF = 10
Both use Pooled StDev = 0,443

Αντίθετα, αν δεν έχουμε δηλώσει ισότητα διασπορών το Minitab δίνει τα παρακάτω αποτελέσματα:

Two Sample T-Test and Confidence Interval				
Two sample T for C1				
C2	N	Mean	StDev	SE Mean
1	5	21,480	0,507	0,23
2	7	20,843	0,395	0,15

95% CI for mu (1) - mu (2): (-0,01; 1,28)
T-Test mu (1) = mu (2) (vs not =): T = 2,35 P = 0,051 DF = 7

Τα αποτελέσματα που δίνει το πακέτο αποτελούνται από τον μέσο (Mean), την τυπική απόκλιση (StDev) και το τυπικό σφάλμα εκτίμησης της μέσης τιμής (SE Mean) για τα δύο δείγματα, ένα 95% διάστημα εμπιστοσύνης για την διαφορά των δύο μέσων (95% CI for $\mu(1) - \mu(2)$), την τιμή της στατιστικής συνάρτησης ελέγχου (T), το παρατηρούμενο επίπεδο σημαντικότητας (P) και τους βαθμούς ελευθερίας (DF). Στην περίπτωση που έχουμε υποθέσει ισότητα διασπορών, το πακέτο δίνει και την τιμή της σταθμισμένης τυπικής απόκλισης (Pooled StDev).

Παράδειγμα 5ο: Ας υποθέσουμε ότι μας ενδιαφέρει να συγκρίνουμε δύο βαθμολογητές που χρησιμοποιούνται από το Υπουργείο Παιδείας για την διόρθωση γραπτών στις πανελλήνιες εξετάσεις. Μας ενδιαφέρει να δούμε αν οι βαθμολογίες των δύο αυτών βαθμολογητών διαφέρουν στατιστικά σημαντικά μεταξύ τους. Για τον λόγο αυτό, διαλέγουμε ένα τυχαίο δείγμα πέντε γραπτών που έχουν βαθμολογηθεί από τους δύο αυτούς καθηγητές. Οι βαθμολογίες των δύο βαθμολογητών για τα πέντε αυτά γραπτά εμφανίζονται στον πίνακα που ακολουθεί:

Μαθητής	Βαθμολογητής 1 X_i	Βαθμολογητής 2 Y_i	Διαφορά ($D_i=X_i-Y_i$)
1	90	93	-3
2	94	96	-2
3	91	92	-1
4	85	88	-3
5	88	90	-2
Μέσοι	89.6	91.8	-2.2
Τυπικές Αποκλίσεις	3.3615	3.0332	.83666

Είναι φανερό ότι ο σχεδιασμός που επιλέξαμε χρησιμοποιεί τις παρατηρήσεις κατά ζεύγη. Αυτό είναι λογικό αφού με τον σχεδιασμό αυτό αποκλείουμε άλλες διαφορές που θα ήταν δυνατόν να υπεισέλθουν στις βαθμολογίες των δύο καθηγητών, όπως π.χ. διαφορές στις γνώσεις των μαθητών που βαθμολογούνται. (Στο παράδειγμα αυτό, φαίνονται καθαρά τα πλεονεκτήματα που έχει η

επιλογή δειγμάτων κατά ζεύγη έναντι αυτής των ανεξαρτήτων δειγμάτων).

Αν παρατηρήσουμε για λίγο τα δεδομένα, βλέπουμε ότι οι διαφορές βαθμολογιών των δύο καθηγητών δεν είναι ιδιαίτερα μεγάλες. Βλέπουμε, όμως, ότι οι βαθμολογίες του βαθμολογητή 1 στο συγκεκριμένο δείγμα είναι πάντοτε χαμηλότερες από τις αντίστοιχες βαθμολογίες του βαθμολογητή 2. Δοθέντος ότι μας ενδιαφέρει να διαπιστώσουμε αν υπάρχει διαφορά μεταξύ των δύο βαθμολογητών, θα πρέπει να ελέγξουμε την υπόθεση

$$H_0: \mu_D = \mu_X - \mu_Y = 0$$

$$H_1: \mu_D = \mu_X - \mu_Y \neq 0$$

Οι τιμές της διαφοράς $\mu_D = \mu_X - \mu_Y$ που παρατηρήθηκαν με βάση το δείγμα των πέντε γραπτών εμφανίζονται στην τελευταία στήλη του πίνακα (και είναι -3, -2, -1, -3, -2). Για το δείγμα αυτό των διαφορών, έχουμε ότι $\bar{x} - \bar{y} = -2.2$ και $s_D^* = .8366 \cong .837$.

Η τιμή t_0 της στατιστικής συνάρτησης ελέγχου T_0 για την συγκεκριμένη περίπτωση είναι

$$t_0 = \frac{\bar{D}}{s_D^*/\sqrt{n}} = \frac{-2.2}{.837/\sqrt{5}} = \frac{-2.2}{.374} = -5.88$$

Για την τιμή αυτή της στατιστικής συνάρτησης ελέγχου, η p-τιμή βρίσκεται από τους πίνακες της κατανομής t να είναι

$$p\text{-τιμή} = P(|T_0| \geq 5.88) = .0042$$

Παρατηρούμε ότι η p-τιμή του ελέγχου αυτού είναι εξαιρετικά χαμηλή (.0042). Επομένως, ακόμη και σε επίπεδο σημαντικότητας 5%, με βάση τις παρατηρήσεις του δείγματος αυτού η μηδενική υπόθεση θα πρέπει να απορριφθεί, γεγονός που μας οδηγεί στο συμπέρασμα ότι υπάρχει στατιστικά σημαντική διαφορά στον τρόπο βαθμολόγησης των βαθμολογητών.

Λύση με το πακέτο SPSS

- Εισάγουμε τα δεδομένα του πρώτου δείγματος σε μια μεταβλητή (έστω var00001) και εκείνα του δεύτερου σε μια άλλη (έστω var00002)

- Από την επιλογή Analyze, επιλέγουμε Compare Means και στην συνέχεια Paired-Samples T test
- Στο παράθυρο που εμφανίζεται, επιλέγουμε πρώτα την μεταβλητή var00001 και στην συνέχεια την var00002 έτσι ώστε να εμφανισθούν και οι δύο στο σημείο του παραθύρου που έχει τον τίτλο Current Selections.
- Πατάμε με το ποντίκι το βέλος που βρίσκεται στο κέντρο του παραθύρου έτσι ώστε να εμφανισθούν στο πεδίο Paired-Variables οι δύο μεταβλητές.
- Επιλέγουμε OK

Τα αποτελέσματα που δίνει το πακέτο είναι τα ακόλουθα:

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	VAR00001	89,6000	5	3,3615	1,5033
	VAR00002	91,8000	5	3,0332	1,3565

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	VAR00001 & VAR00002	5	,971	,006

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	VAR00001 - VAR00002	-2,2000	,8367	,3742	-3,24	-1,1611	-5,880	4	,004

Στον πρώτο από τους τρεις πίνακες, δίνονται ο μέσος (Mean), ο αριθμός των παρατηρήσεων (N), η τυπική απόκλιση (Std. Deviation) και το τυπικό σφάλμα εκτίμησης της μέσης τιμής (Std. Error Mean) και για τα δύο δείγματα.

Στον δεύτερο πίνακα, δίνεται η τιμή του συντελεστή συσχέτισης (correlation) των δύο δειγμάτων και το παρατηρούμενο επίπεδο σημαντικότητας για τον έλεγχο του ότι η τιμή του συντελεστή συσχέτισης είναι ίση με μηδέν.

Στον τρίτο πίνακα, δίνεται ο μέσος των διαφορών των δύο δειγμάτων (Mean), η τυπική απόκλιση των διαφορών (Std. Deviation) και το τυπικό σφάλμα εκτίμησης της μέσης τιμής των διαφορών (Std. Error Mean). Επίσης δίνονται τα όρια του 95% διαστήματος εμπιστοσύνης για την διαφορά των μέσων (95% Confidence Interval of the Difference), η τιμή της στατιστικής συνάρτησης ελέγχου (t), οι βαθμοί ελευθερίας (df) και το παρατηρούμενο επίπεδο σημαντικότητας (Sig. 2-tailed)

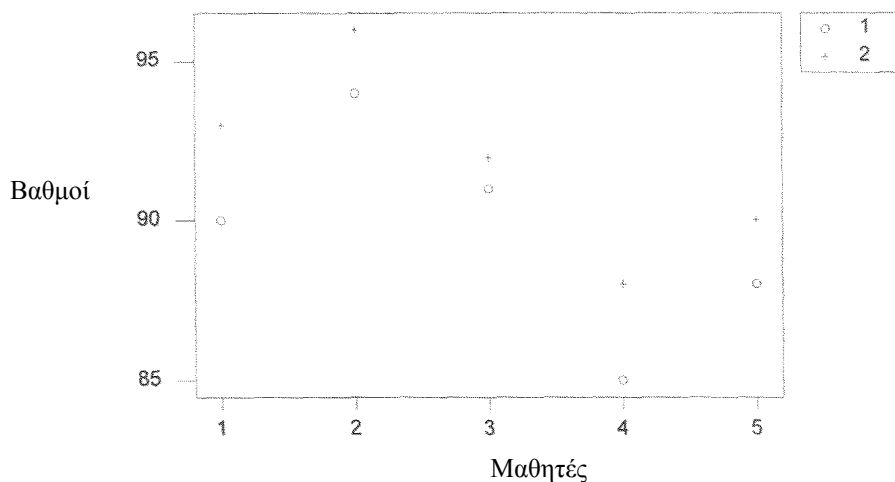
Λύση με το πακέτο Minitab

- Εισάγουμε τα δεδομένα του πρώτου δείγματος σε μια μεταβλητή (έστω C1) και εκείνα του δεύτερου σε μια άλλη (έστω C2)
- Επιλέγουμε Stat, Basic Statistics και Paired t.
- Στο πεδίο First Sample, επιλέγουμε την στήλη C1 και στο πεδίο Second Sample την στήλη C2.
- Επιλέγοντας options, μπορούμε να δηλώσουμε το επίπεδο εμπιστοσύνης για το διάστημα εμπιστοσύνης που δίνει υποχρεωτικά το πακέτο (confidence level), την τιμή με την οποία θέλουμε να ελέγξουμε αν είναι ίση η διαφορά των δύο μέσων τιμών (Test mean) και την μορφή της εναλλακτικής υπόθεσης (Alternative). Για το παράδειγμα μας, πρέπει να επιλέξουμε not equal.
- Επιλέγουμε OK για να επιστρέψουμε στο παράθυρο Paired t
- Επιλέγουμε OK

Το πακέτο δίνει τα παρακάτω αποτελέσματα:

Paired T-Test and Confidence Interval				
Paired T for C1 - C2				
	N	Mean	StDev	SE Mean
C1	5	89,60	3,36	1,50
C2	5	91,80	3,03	1,36
Difference	5	-2,200	0,837	0,374
95% CI for mean difference: (-3,239; -1,161)				
T-Test of mean difference = 0 (vs not = 0): T-Value = -5,88 P-Value = 0,004				

Στον πίνακα, δίνονται ο μέσος (Mean), η τυπική απόκλιση (StDev) και το τυπικό σφάλμα εκτίμησης της μέσης τιμής (SE Mean) για τα δύο δείγματα και τις διαφορές τους, ένα 95% διάστημα εμπιστοσύνης για την διαφορά των μέσων, η τιμή της στατιστικής συνάρτησης ελέγχου (T-Value), και το παρατηρούμενο επίπεδο σημαντικότητας (P-value).



Παρατήρηση: Στο συγκεκριμένο πρόβλημα, όπως και σε άλλα παρόμοια προβλήματα σύγκρισης των μέσων δύο πληθυσμών, μπορεί κανείς να επιλέξει μεταξύ της μεθόδου των ανεξαρτήτων δειγμάτων και της μεθόδου των δειγμάτων κατά ζεύγη. Στο συγκεκριμένο πρόβλημα, αν είχαμε επιλέξει την μέθοδο των ανεξαρτήτων δειγμάτων θα είχαμε διαλέξει ένα τυχαίο δείγμα βαθμολογιών του πρώτου βαθμολογητή και ένα αντίστοιχο τυχαίο δείγμα βαθμολογιών του δεύτερου βαθμολογητή χωρίς να εξετάζουμε αν οι βαθμολογίες αυτές αναφέρονται στους ίδιους μαθητές ή όχι. Με τον τρόπο αυτό όμως, η όποια διαφορά που θα παρατηρούσαμε στις μέσες βαθμολογίες των δύο βαθμολογητών θα μπορούσε να οφείλεται, εκτός από την διαφορά αυτών καθ' εαυτών των βαθμολογητών, σε διαφορές των μαθητών των οποίων τα γραπτά έχουν επιλεγεί στα δείγματα. (Π.χ. οι μαθητές θα μπορούσαν να προέρχονται από διαφορετικό σχολείο, όπως επίσης να είναι μαθητές διαφορετικού επιπέδου γνώσεων ή διαφορετικής προετοιμασίας για τις εξετάσεις). Η

χρησιμοποίηση των δειγμάτων κατά ζεύγη εξαλείφει την επίδραση των παραγόντων αυτών οι οποίοι είναι ενδεχόμενο να συνεισφέρουν στις διαφορές των βαθμολογιών και μας επιτρέπουν να συγκρίνουμε μόνο τους βαθμολογητές αυτούς καθ' εαυτούς.

ΣΥΓΚΡΙΣΗ ΑΝΑΛΟΓΙΩΝ ΔΥΟ ΠΛΗΘΥΣΜΩΝ

Σε πολλά προβλήματα, το ενδιαφέρον του ερευνητή δεν εστιάζεται στην σύγκριση των μέσων δύο πληθυσμών αλλά στην σύγκριση των αναλογιών των στοιχείων των πληθυσμών που έχουν κάποια συγκεκριμένη ιδιότητα. Τέτοιες περιπτώσεις είναι π.χ. η σύγκριση των αναλογιών (των ποσοστών) των φοιτητών που αποτυγχάνουν στις εξετάσεις σε δύο διαφορετικά μαθήματα, των ποσοστών των πολιτών που υποστηρίζουν ένα συγκεκριμένο πολιτικό κόμμα σε δύο διαφορετικές περιοχές της χώρας κ.λ.π.

Στο πλαίσιο του ελέγχου υποθέσεων, αν υποθέσουμε ότι επιλέγουμε δύο τυχαία δείγματα μεγέθους n και m αντίστοιχα από τους δύο πληθυσμούς και συμβολίσουμε με X και Y αντίστοιχα τον αριθμό των στοιχείων σε καθένα από τα δείγματα που έχουν την ιδιότητα την οποία μας ενδιαφέρει να μελετήσουμε και με p_X , p_Y τις αναλογίες (ποσοστά) των στοιχείων των δύο πληθυσμών που έχουν την ιδιότητα αυτή (τις οποίες αναλογίες θέλουμε να συγκρίνουμε), θα έχουμε στο πλαίσιο του στατιστικού ελέγχου υποθέσεων τα εξής:

$$\begin{array}{lll} H_0 : p_X = p_Y & & \\ H_1 : p_X < p_Y & H_1 : p_X \neq p_Y & H_1 : p_X > p_Y \end{array}$$

Δοθέντος ότι τα δείγματα είναι ανεξάρτητα και ότι καθένα από τα στοιχεία των πληθυσμών είτε έχει την ιδιότητα που ενδιαφέρει είτε όχι, θα έχουμε ότι τα X και Y ακολουθούν διωνυμικές κατανομές με παραμέτρους (n, p_X) και (m, p_Y) αντίστοιχα. Δηλαδή,

$$X \sim b(n, p_X) \quad \text{και} \quad Y \sim b(m, p_Y)$$

Για μεγάλα δείγματα, λόγω της προσέγγισης της διωνυμικής με την κανονική κατανομή για τις εκτιμήτριες \hat{p}_X, \hat{p}_Y των p_X, p_Y αντίστοιχα, θα έχουμε,

$$\hat{p}_X = \frac{X}{n} \underset{appr.}{\sim} N\left(p_X, \frac{p_X q_X}{n}\right), \quad q_X = 1 - p_X$$

$$\hat{p}_Y = \frac{Y}{m} \underset{appr.}{\sim} N\left(p_Y, \frac{p_Y q_Y}{m}\right), \quad q_Y = 1 - p_Y$$

Για ανεξάρτητα δείγματα θα έχουμε,

$$\hat{p}_X - \hat{p}_Y = \frac{X}{n} - \frac{Y}{m} \underset{appr.}{\sim} N\left(p_X - p_Y, \frac{p_X q_X}{n} + \frac{p_Y q_Y}{m}\right)$$

Επομένως, η τυποποιημένη στατιστική συνάρτηση θα είναι η,

$$Z = \frac{\hat{p}_X - \hat{p}_Y - (p_X - p_Y)}{\sqrt{\frac{p_X q_X}{n} + \frac{p_Y q_Y}{m}}} \sim N(0, 1)$$

Επειδή p_X, p_Y είναι άγνωστα, προκειμένου να προχωρήσουμε σε συμπερασματολογία, θα πρέπει να αντικαταστήσουμε την τυπική απόκλιση $\sqrt{\frac{p_X q_X}{n} + \frac{p_Y q_Y}{m}}$ της εκτιμήτριας $\hat{p}_X - \hat{p}_Y$ με κάποια εκτιμώμενη από τα δείγματα τιμή.

Κάτω από τη μηδενική υπόθεση, έχουμε, ότι $p_X = p_Y = p$ και επομένως έχουμε δύο εκτιμήτριες τις, $\hat{p}_X = \frac{X}{n}$ και $\hat{p}_Y = \frac{Y}{m}$ που εκτιμούν την κοινή αναλογία p . Είναι, επομένως, χρήσιμο να θεωρήσουμε ως εκτιμήτρια του p , την σταθμισμένη εκτιμήτρια \hat{p}_p που προέρχεται από τα δύο δείγματα.

Έτσι, θα έχουμε

$$\hat{p}_p = \frac{n\hat{p}_X + m\hat{p}_Y}{n + m} = \frac{X + Y}{n + m}$$

Η στατιστική συνάρτηση ελέγχου κάτω από την μηδενική υπόθεση θα είναι επομένως,

$$Z_0 = \frac{\hat{p}_X - \hat{p}_Y}{\sqrt{\hat{p}_p \hat{q}_p \left(\frac{1}{n} + \frac{1}{m} \right)}}$$

ή, ισοδύναμα,

$$Z_0 = \frac{\frac{X}{n} - \frac{Y}{m}}{\sqrt{\frac{X+Y}{n+m} \left(1 - \frac{X+Y}{n+m} \right) \left(\frac{1}{n} + \frac{1}{m} \right)}}$$

Για τις διάφορες εναλλακτικές υποθέσεις θα έχουμε:

1) Αν $H_1 : p_X < p_Y$, απορρίπτουμε την μηδενική υπόθεση αν

$$Z_0 < Z_{\alpha/2} = -Z_{1-\alpha/2}$$

2) Αν $H_1 : p_X \neq p_Y$, απορρίπτουμε την μηδενική υπόθεση αν

$$Z_0 < Z_{\alpha/2} = -Z_{1-\alpha/2} \quad \text{ή} \quad Z_0 > Z_{1-\alpha/2}$$

3) Αν $H_1 : p_X > p_Y$, απορρίπτουμε τη μηδενική υπόθεση αν

$$Z_0 > Z_{1-\alpha/2}$$

Παράδειγμα: Η εταιρεία παρασκευής και συσκευασίας ενός προϊόντος, που είδαμε σε προηγούμενα παραδείγματα, έχει δύο εργοστάσια συσκευασίας σε δύο διαφορετικές περιοχές έστω Α και Β. Στον έλεγχο ποιότητας της συσκευασίας, αν διαπιστωθεί ότι η ποσότητα που τοποθετείται στα κουτιά αποκλίνει πέρα από κάποιο συγκεκριμένο σημείο, τότε η εταιρεία απορρίπτει τα κουτιά αυτά. Έστω ότι ο υπεύθυνος του ελέγχου ποιότητας της εταιρείας ενδιαφέρεται να διαπιστώσει εάν υπάρχει στατιστικά σημαντική διαφορά στην αναλογία των κουτιών που χρειάζεται να απορριφθούν στα δύο εργοστάσια. Προκειμένου να γίνει αυτό, επιλέγονται δύο τυχαία δείγματα από 200 κουτιά από κάθε εργοστάσιο από τα οποία προκύπτει ότι 19 από το πρώτο εργοστάσιο και 21 από το δεύτερο χρειάζεται να απορριφθούν. Να ελεγχθεί η υπόθεση ότι η διαφορά των αναλογιών στα κουτιά που προέρχονται από τα δύο εργοστάσια και που πρέπει να απορριφθούν δεν είναι στατιστικά σημαντική.

Λύση: Έχουμε ότι $n = m = 200$, $X = 19$, $Y = 21$.

Η υπόθεση που πρέπει να ελέγξουμε είναι

$$H_0 : p_X = p_Y$$

$$H_1 : p_X \neq p_Y.$$

Έστω $\alpha = 0.10$.

Από τα στοιχεία των δειγμάτων έχουμε,

$$\frac{X}{n} = \frac{19}{200} = 0.095 \quad \text{και} \quad \frac{Y}{m} = \frac{21}{200} = 0.105$$

επίσης, $\hat{p}_p = \frac{19 + 21}{200} = 0.10$

Επομένως, η τιμή της στατιστικής συνάρτησης ελέγχου είναι η

$$\begin{aligned} Z_0 &= \frac{\hat{p}_X - \hat{p}_Y}{\sqrt{\hat{p}_p \hat{q}_p \left(\frac{1}{n} + \frac{1}{m} \right)}} \\ &= \frac{0.095 - 0.105}{\sqrt{(0.1)(0.9) \left(\frac{1}{200} + \frac{1}{100} \right)}} = -0.33 \end{aligned}$$

Από τους πίνακες της κανονικής κατανομής, έχουμε ότι

$$Z_{1-\alpha/2} = Z_{0.975} = 1.645$$

Η τιμή της στατιστικής συνάρτησης ελέγχου κάτω από την μηδενική υπόθεση δεν βρίσκεται στην περιοχή απόρριψης. Επομένως, καταλήγουμε στο συμπέρασμα ότι, με βάση τα στοιχεία των δειγμάτων, δεν υπάρχει στατιστικά σημαντική διαφορά στο ποσοστό των ελαττωματικών συσκευασιών που πρέπει να αντικατασταθούν στα δύο εργοστάσια.

Σημείωση: Αν χρησιμοποιήσουμε την προσέγγιση με την p-τιμή (παρατηρούμενο επίπεδο σημαντικότητας), έχουμε ότι

$$P(Z < -0.33) = 0.37037$$

Επειδή έχουμε αμφίπλευρη εναλλακτική υπόθεση, η p-τιμή είναι

$$p\text{-τιμή} = 2(0.37037) = 0.7414$$

Επομένως, μόνο για επίπεδο σημαντικότητας μεγαλύτερο του 0.74 θα απορρίψουμε την μηδενική υπόθεση με βάση τα στοιχεία του δείγματος. Η p -τιμή αποτελεί μια ένδειξη ότι, με βάση τα στοιχεία του δείγματος, είναι εξαιρετικά απίθανο να απορριφθεί η μηδενική υπόθεση. (Θα πρέπει δηλαδή να χρησιμοποιηθεί μια πολύ μεγάλη τιμή για το επίπεδο σημαντικότητας ώστε, με βάση τα στοιχεία του δείγματος, να απορριφθεί η μηδενική υπόθεση).