

ΚΕΦΑΛΑΙΟ 14

ΔΙΑΣΤΗΜΑΤΑ ΕΜΠΙΣΤΟΣΥΝΗΣ ΓΙΑ ΤΗΝ ΔΙΑΦΟΡΑ ΜΕΣΩΝ ΤΙΜΩΝ ΚΑΝΟΝΙΚΩΝ ΠΛΗΘΥΣΜΩΝ

Έστω X_1, X_2, \dots, X_n και Y_1, Y_2, \dots, Y_m δύο τυχαία δείγματα μεγέθους n και m αντίστοιχα από δύο ανεξάρτητους κανονικούς πληθυσμούς $N(\mu_X, \sigma_X^2)$ και $N(\mu_Y, \sigma_Y^2)$. Ενδιαφερόμαστε να κατασκευάσουμε ένα διάστημα εμπιστοσύνης για τη διαφορά $\mu_X - \mu_Y$.

Η προσέγγιση που ακολουθείται για την κατασκευή αυτού του διαστήματος εμπιστοσύνης βασίζεται συνήθως σε ανεξάρτητα δείγματα, δηλαδή σε δείγματα που επιλέγονται με τρόπο ώστε οι παρατηρήσεις τους ενός να είναι ανεξάρτητες των παρατηρήσεων του άλλου.

Πολλές φορές όμως, όπως θα δούμε στην συνέχεια, προκειμένου να ελαχιστοποιήσουμε την επίδραση εξωτερικών παραγόντων, επιλέγουμε τα τυχαία δείγματα με τρόπο που οι παρατηρήσεις του ενός “συνδέονται” με τις παρατηρήσεις του άλλου σχηματίζοντας ζεύγη. Αυτό, για παράδειγμα, συμβαίνει στην περίπτωση που έχουμε δείγματα για εισοδήματα συζύγων ή αδελφών ή στις περιπτώσεις που έχουμε να συγκρίνουμε τις πωλήσεις δύο υποκαταστημάτων μιας αλυσίδας καταστημάτων σε δύο διαφορετικές τοποθεσίες.

Στις επόμενες τρεις ενότητες (Α, Β, Γ) εξετάζουμε την περίπτωση των ανεξαρτήτων δειγμάτων και, στην συνέχεια, στην ενότητα που ακολουθεί (Δ), την περίπτωση δειγμάτων που οι παρατηρήσεις τους έχουν επιλεγεί κατά ζεύγη.

Α. Περίπτωση Γνωστών Διακυμάνσεων

(Ανεξάρτητα Δείγματα)

Όπως είναι γνωστό, μια σημειακή εκτιμήτρια για την διαφορά $\mu_X - \mu_Y$ είναι το $\bar{X} - \bar{Y}$ όπου \bar{X} , \bar{Y} είναι οι δειγματικοί μέσοι των δύο δειγμάτων.

Όπως γνωρίζουμε,

$$\bar{X} \sim N(\mu_X, \sigma_X^2/n)$$

και

$$\bar{Y} \sim N(\mu_Y, \sigma_Y^2/m)$$

Επομένως,

$$\bar{X} - \bar{Y} \sim N\left(\mu_X - \mu_Y, \frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}\right)$$

ή ισοδύναμα,

$$Z = \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sqrt{\sigma_X^2/n + \sigma_Y^2/m}} \sim N(0, 1)$$

Συνεπώς, το $100(1-\alpha)\%$ διάστημα εμπιστοσύνης για το $\mu_X - \mu_Y$ είναι το

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

Αυτό είναι ένα ακριβές διάστημα εμπιστοσύνης.

Παρατήρηση: Αν οι πληθυσμοί δεν είναι ακριβώς κανονικοί αλλά τα μεγέθη των δειγμάτων n και m είναι αρκετά μεγάλα, το διάστημα εμπιστοσύνης που προαναφέρθηκε μπορεί, λόγω του κεντρικού οριακού θεωρήματος, να χρησιμοποιηθεί ως ένα κατά προσέγγιση διάστημα εμπιστοσύνης για τη διαφορά $\mu_X - \mu_Y$.

Παράδειγμα: Προκειμένου να ελεγχθεί η ποιότητα δύο ειδών ελαστικών αυτοκινήτων, έγινε ένας έλεγχος σε 100 ελαστικά από κάθε είδος τυχαία επιλεγμένα ($n=m=100$). Ως στοιχείο ποιότητας χρησιμοποιήθηκε ο αριθμός των χιλιομέτρων που τα ελαστικά αυτά χρησιμοποιήθηκαν μέχρις ότου φθάσουν σε ένα συγκεκριμένο σημείο φθοράς. Τα αποτελέσματα αυτά (σε χιλιόμετρα) ήταν ως εξής:

$$\bar{x} = 26400 \text{ km} \quad \bar{y} = 25100 \text{ km}$$

Από προηγούμενη πείρα είναι γνωστό ότι

$$\sigma_X^2 = 1440000 \quad \sigma_Y^2 = 1960000$$

Να κατασκευασθεί ένα 99% διάστημα εμπιστοσύνης για την διαφορά $\mu_X - \mu_Y$.

Λύση: Το ζητούμενο διάστημα είναι

$$\bar{x} - \bar{y} \pm Z_{.995} \sqrt{\sigma_X^2/n + \sigma_Y^2/m}$$

Για το δείγμα μας έχουμε,

$$1300 \pm 2.58 (184) \quad \text{ή}$$

$$(825, 1775)$$

Επομένως, με πιθανότητα .99 η διαφορά της μέσης διάρκειας ζωής μεταξύ των δύο αυτών ειδών ελαστικών έχει μια τιμή στο διάστημα (825, 1775).

B. Περίπτωση Αγνώστων Ίσων Διακυμάνσεων

(Ανεξάρτητα Δείγματα)

Υποθέτουμε και πάλι ότι $X \sim N(\mu_X, \sigma_X^2)$ και $Y \sim N(\mu_Y, \sigma_Y^2)$ και ότι X, Y είναι ανεξάρτητες. Τότε,

$$Z = \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sqrt{\sigma_X^2/n + \sigma_Y^2/m}} \sim N(0, 1)$$

Υποθέτουμε ότι $\sigma_X^2 = \sigma_Y^2 = \sigma^2$.

Έστω S_X^2, S_Y^2 οι διασπορές των δύο δειγμάτων αντίστοιχα (S_X^{*2}, S_Y^{*2} οι αντίστοιχες αμερόληπτες εκτιμήτριες).

Λόγω της ανεξαρτησίας, θα έχουμε ότι

$$V = \frac{nS_X^2}{\sigma^2} + \frac{mS_Y^2}{\sigma^2} \sim \chi_{n+m-2}^2$$

(Οι βαθμοί ελευθερίας της V είναι $n-1+m-1 = n+m-2$).

Δοθέντος επίσης ότι Z και V είναι ανεξάρτητες έχουμε

$$T = \frac{Z}{\sqrt{V/(n+m-2)}} \sim t_{n+m-2}$$

Από τις προηγούμενες σχέσεις προκύπτει ότι,

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

όπου

$$S_p^2 = \frac{nS_X^2 + mS_Y^2}{n + m - 2} = \frac{(n-1)S_X^{*2} + (m-1)S_Y^{*2}}{n + m - 2}$$

δηλαδή,

$$S_p^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2 + \sum_{i=1}^m (Y_i - \bar{Y})^2}{n + m - 2}$$

Επομένως το $100(1-\alpha)\%$ διάστημα εμπιστοσύνης για την διαφορά $\mu_X - \mu_Y$ το στη περίπτωση αυτή είναι,

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

Παρατήρηση: Η διασπορά S_p^2 ονομάζεται *σταθμισμένη διασπορά* (*pooled variance*). Χρησιμοποιείται δε γιατί έχουμε δύο ανεξάρτητες εκτιμήσεις που προέρχονται από δύο διαφορετικά δείγματα για την ίδια ποσότητα (το σ^2). Είναι λοιπόν φυσικό να κάνουμε χρήση και των δύο αυτών εκτιμήσεων λαμβάνοντας όμως υπόψη μας (και δίνοντας την αντίστοιχη βαρύτητα) στην ποιότητα της καθεμιάς από αυτές (δηλαδή στο πόσο ακριβής είναι η καθεμιά από αυτές με βάση το μέγεθος του δείγματος από το οποίο έχει προέλθει).

Παρατήρηση: Είναι δυνατόν να αποδειχθεί, είτε μαθηματικά είτε πειραματικά, ότι η σταθμισμένη εκτιμήτρια S_p^2 της κοινής διασποράς σ^2 είναι μια αμερόληπτη εκτιμήτρια του σ^2 .

Σημείωση: Αν $\sigma_X^2 \neq \sigma_Y^2$ αλλά n, m είναι αρκετά μεγάλα, μπορούμε να αντικαταστήσουμε το σ_X^2 με το $nS_X^2/(n-1)$ και το σ_Y^2 με το

$mS_Y^2/(m-1)$ στην αρχική σχέση (αντίστοιχα το σ_X^2 με το S_X^{*2} και το σ_Y^2 με το S_Y^{*2} δηλαδή, τις αμερόληπτες εκτιμήτριες των σ_X^2 και σ_Y^2 αντίστοιχα). Τότε ένα κατά προσέγγιση $100(1-\alpha)\%$ διάστημα εμπιστοσύνης για την διαφορά $\mu_X - \mu_Y$ είναι το

$$\boxed{\bar{X} - \bar{Y} \pm Z_{1-\alpha/2} \sqrt{S_X^2/(n-1) + S_Y^2/(m-1)}}$$

ή ισοδύναμα, με τις αμερόληπτες εκτιμήτριες των διασπορών,

$$\boxed{\bar{X} - \bar{Y} \pm Z_{1-\alpha/2} \sqrt{S_X^{*2}/n + S_Y^{*2}/m}}$$

Παράδειγμα: Προκειμένου να συγκριθούν οι φοιτητές των τμημάτων Στατιστικής δύο διαφορετικών Πανεπιστημίων, επελέγησαν δύο τυχαία δείγματα από $n=50$ και $m=60$ φοιτητές του ίδιου έτους στους οποίους δόθηκε ένα συγκεκριμένο διαγώνισμα. Τα αποτελέσματα στο διαγώνισμα αυτό (σε βαθμολογία με κλίμακα 0-100) ήταν ως εξής:

$$\begin{aligned} \bar{x} &= 77 & \bar{y} &= 68 \\ S_X^2 &= 64 & S_Y^2 &= 100 \end{aligned}$$

Να κατασκευασθεί ένα 95% διάστημα εμπιστοσύνης για την διαφορά $\mu_X - \mu_Y$ της πραγματικής μέσης απόδοσης των φοιτητών στα δύο αυτά Πανεπιστήμια.

Λύση: Υποθέτουμε ότι $\sigma_X^2 = \sigma_Y^2$ και ότι η απόδοση στα διαγωνίσματα ακολουθεί την κανονική κατανομή.

Για την σταθμισμένη διασπορά έχουμε,

$$S_p^2 = \frac{50 * 64 + 60 * 100}{108} = 85.18$$

Το 95% διάστημα εμπιστοσύνης για τη διαφορά $\mu_X - \mu_Y$ είναι,

$$9 \pm t_{108, 975} \sqrt{85.18/50 + 85.18/60}$$

ή, ισοδύναμα,

$$(5.60, 12.40)$$

(Επειδή στους πίνακες της κατανομής t δεν δίνεται τιμή για 162 βαθμούς ελευθερίας, χρησιμοποιούμε, από τους πίνακες της κανονικής κατανομής, την τιμή $Z_{.975} = 1.96$. Άλλωστε, για τους βαθμούς αυτούς ελευθερίας η κατανομή t προσεγγίζεται σχεδόν πλήρως από την τυποποιημένη κανονική κατανομή).

Σημείωση 1: Μερικές φορές κάποιο από τα άκρα του διαστήματος, ή και τα δύο, μπορεί να είναι αρνητικό πράγμα το οποίο αποτελεί ένδειξη ότι η μέση τιμή του δεύτερου πληθυσμού υπερβαίνει την μέση τιμή του πρώτου πληθυσμού.

Σημείωση 2: Στο προηγούμενο πρόβλημα, δοθέντος ότι τα n και m ήταν μεγάλα, μπορούσαμε να βρούμε ένα κατά προσέγγιση διάστημα εμπιστοσύνης για το $\mu_X - \mu_Y$ χωρίς να υποθέσουμε ότι $\sigma_X^2 = \sigma_Y^2$, χρησιμοποιώντας τον αντίστοιχο τύπο που δώσαμε παραπάνω.

Παρατήρηση: Η χρησιμοποίηση της μεθόδου που αναλύθηκε στην περίπτωση που $\sigma_X^2 = \sigma_Y^2$ απαιτεί την επιβεβαίωση της υπόθεσης αυτής. Όπως θα δούμε αργότερα (στους ελέγχους υποθέσεων), υπάρχει στατιστική μεθοδολογία με την οποία μπορούμε να ελέγξουμε την υπόθεση αυτή.

Παράδειγμα: Σε ένα εργοστάσιο έχει βρεθεί ότι απαιτείται χρόνος περίπου ενός μηνός για να εκπαιδευθεί ένας καινούργιος εργαζόμενος και να φθάσει τη μέγιστη δυνατή απόδοση. Στο εργοστάσιο αυτό έχει προταθεί μια καινούργια μέθοδος εκπαίδευσης για την ελάττωση του χρόνου αυτού. Προκειμένου να ελεγχθεί η αποτελεσματικότητα της μεθόδου αυτής, επιλέγονται δύο ομάδες από 9 εργαζόμενους η καθεμία. Οι εργαζόμενοι αυτοί εκπαιδεύονται για μία περίοδο δύο εβδομάδων, την μιά βδομάδα με χρησιμοποίηση της παλιάς μεθόδου και την άλλη με χρησιμοποίηση της καινούργιας μεθόδου. Στην συνέχεια, ανατίθεται σε καθένα από τους εργαζόμενους που εκπαιδεύτηκαν η συναρμολόγηση μιας συσκευής (εργασία για την οποία εκπαιδεύονται). Οι χρόνοι (σε λεπτά) που απαιτήθηκαν από τον κάθε εργαζόμενο για να συναρμολογήσει το συγκεκριμένο

αντικείμενο μετά από το χρόνο της εκπαίδευσης αυτής δίνονται στον πίνακα που ακολουθεί:

ΧΡΟΝΟΣ ΠΟΥ ΑΠΑΙΤΗΘΗΚΕ (σε λεπτά)	
Συνήθης Μέθοδος Εκπαίδευσης X	Νέα Μέθοδος Εκπαίδευσης Y
32	35
37	31
35	29
28	25
41	34
44	40
35	27
31	32
34	31

Να κατασκευασθεί ένα 95% διάστημα εμπιστοσύνης για την διαφορά μεταξύ των μέσων χρόνων που απαιτούνται για την συναρμολόγηση του συγκεκριμένου προϊόντος μετά από περίοδο εκπαίδευσης δύο εβδομάδων για τη συνήθη και την καινούργια μέθοδο εκπαίδευσης.

Λύση: Ας υποθέσουμε ότι μ_X και μ_Y είναι οι μέσοι χρόνοι που απαιτούνται για τη συναρμολόγηση του προϊόντος μετά από εκπαίδευση με τη συνήθη και τη νέα μέθοδο αντίστοιχα.

Ας υποθέσουμε ότι η διακύμανση στους μέσους χρόνους συναρμολόγησης είναι στην πραγματικότητα συνάρτηση των διαφορών που οφείλονται στις προσωπικότητες των εργαζομένων και ότι οι διακυμάνσεις για τις μετρήσεις από τους δύο πληθυσμούς μπορούν να θεωρηθούν κατά προσέγγιση ίσες.

Από τα δύο δείγματα μπορούν να υπολογισθούν τα εξής στοιχεία:

$$\bar{x} = 35.22 \quad \bar{y} = 31.56 \quad n = m = 9$$

$$\sum_{i=1}^n (x_i - \bar{x})^2 = 195.56$$

$$\sum_{i=1}^m (y_i - \bar{y})^2 = 160.22$$

$$\begin{aligned}
S_p^2 &= \frac{\sum_{i=1}^n (x_i - \bar{x})^2 + \sum_{i=1}^m (y_i - \bar{y})^2}{n + m - 2} \\
&= \frac{195.56 + 160.22}{9 + 9 - 2} \\
&= 22.24
\end{aligned}$$

Επομένως,

$$s_p = \sqrt{22.24} = 4.72$$

Από τους πίνακες της κατανομής t

$$t_{9+9-2, .975} = t_{16, .975} = 2.120$$

Αντικαθιστώντας τις τιμές στον τύπο που δίνει το διάστημα εμπιστοσύνης

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

έχουμε,

$$(35.22 - 31.56) \pm (2.120) (4.72) \sqrt{2/9}$$

ή ισοδύναμα,

$$3.66 \pm 4.72$$

δηλαδή τελικά,

$$(-1.06, 8.38)$$

Επομένως, το εκτιμώμενο 95% διάστημα εμπιστοσύνης για την διαφορά των μέσων χρόνων συναρμολόγησης είναι το (-1.06, 8.38).

Λύση με τη χρήση του πακέτου Minitab

Η κατασκευή του διαστήματος εμπιστοσύνης στο πρόβλημα μπορεί να γίνει από το πακέτο Minitab ως εξής:

- Στο παράθυρο Data εισάγουμε τα δεδομένα σε δύο μεταβλητές π.χ. C1, C2.
- Από την επιλογή Stat επιλέγουμε Basic Statistics.
- Επιλέγουμε 2-Sample t...
- Επιλέγουμε Samples in different columns, στο First τοποθετούμε την μεταβλητή C1 και στο Second την μεταβλητή C2.

- Επιλέγουμε Assume equal variances και OK.

```
MTB > TwoSample 95.0 C1 C2;
SUBC> Alternative 0;
SUBC> Pooled.
```

```
TWOSAMPLE T FOR C1 VS C2
      N  MEAN  STDEV  SE MEAN
C1    9  35.22  4.94    1.6
C2    9  31.56  4.48    1.5
```

95 PCT CI FOR MU C1 - MU C2: (-1.0, 8.4)

TTEST MU C1 = MU C2 (VS NE): T= 1.65 P=0.12 DF= 16

POOLED STDEV = 4.72

Στον πίνακα αποτελεσμάτων, έχουμε πρώτα τα ονόματα των μεταβλητών (C1, C2), μετά τον αριθμό των παρατηρήσεων (N), τον δειγματικό μέσο (Mean), την τυπική απόκλιση (StDev), το τυπικό σφάλμα του μέσου (SE Mean) για κάθε μεταβλητή. Στην συνέχεια, έχουμε το διάστημα εμπιστοσύνης (95 PCT CI FOR MU C1 - MU C2) και τέλος κάποια στοιχεία για τον έλεγχο υπόθεσης.

Σημείωση: Παρατηρούμε ότι το μήκος του διαστήματος που κατασκευάσαμε είναι πολύ μεγάλο και δεν δίνει στην πραγματικότητα πολλές πληροφορίες για ουσιαστικά συμπεράσματα. Σε μια τέτοια περίπτωση, θα ήταν, ίσως, χρήσιμο να αυξηθεί το μέγεθος του δείγματος και να επαναληφθεί η διαδικασία.

Παρατήρηση: Όσον αφορά τις υποθέσεις που κάναμε για την κατασκευή διαστημάτων εμπιστοσύνης στην ενότητα αυτή, θα πρέπει να παρατηρήσουμε ότι μικρές αποκλίσεις από την υπόθεση ότι ο πληθυσμός ακολουθεί την κανονική κατανομή δεν επηρεάζουν σοβαρά τις ιδιότητες των εκτιμητριών ή τον συντελεστή εμπιστοσύνης στο αντίστοιχο διάστημα εμπιστοσύνης. Από το άλλο

μέρος όπως είπαμε, θα πρέπει οι δύο υπό μελέτη πληθυσμοί να έχουν, κατά προσέγγιση τουλάχιστον, ίσες διασπορές.

Όταν οι διασπορές δεν είναι ίσες αλλά ούτε τα μεγέθη των δειγμάτων είναι μεγάλα, τότε δεν ακολουθείται καμιά από τις διαδικασίες που προαναφέρθηκαν. Στην τελευταία αυτή περίπτωση, ακολουθούμε την παρακάτω διαδικασία.

Γ. Περίπτωση Αγνώστων Ανίσων Διακυμάνσεων (Ανεξάρτητα Δείγματα)

Αν έχουμε λόγους να πιστεύουμε ότι οι διακυμάνσεις των δύο πληθυσμών απέχουν πολύ από το να είναι ίσες, θα πρέπει να γίνουν οι εξής αλλαγές στη διαδικασία που εκθέσαμε μέχρι τώρα:

Η σταθμισμένη εκτιμήτρια S_p^2 δεν είναι πιά κατάλληλη και θα πρέπει να χρησιμοποιηθούν οι δειγματικές διασπορές S_X^{*2} και S_Y^{*2} των σ_X^2 και σ_Y^2 αντίστοιχα. Επομένως, το $100(1-\alpha)\%$ διάστημα εμπιστοσύνης για το $\mu_X - \mu_Y$ στη περίπτωση αυτή είναι,

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

όπου οι βαθμοί ελευθερίας v της κατανομής t δίνονται από τον τύπο

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

Φυσικά, το αποτέλεσμα στρογγυλοποιείται στον πλησιέστερο ακέραιο.

Αν τα μεγέθη των δειγμάτων n και m είναι μεγάλα, χρησιμοποιούμε τις εκτιμήτριες S_X^{*2} και S_Y^{*2} (αντίστοιχα τις S_X^2 , S_Y^2) των διασπορών σ_X^2 και σ_Y^2 των δύο πληθυσμών και αξιοποιούμε το γεγονός ότι για μεγάλα δείγματα η στατιστική συνάρτηση

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

ακολουθεί, σύμφωνα με το κεντρικό οριακό θεώρημα, κατά προσέγγιση, την κατανομή $N(0,1)$. Το διάστημα εμπιστοσύνης, επομένως, κατασκευάζεται όπως στην περίπτωση Α. Συγκεκριμένα, το $100(1-\alpha)\%$ διάστημα εμπιστοσύνης για την διαφορά $\mu_X - \mu_Y$ έχει άκρα

$$\bar{X} - \bar{Y} \pm Z_{1-\alpha/2} \sqrt{S_X^2/n + S_Y^2/m}$$

Παράδειγμα: (Συνέχεια του προηγούμενου παραδείγματος).

Αν στο τελευταίο παράδειγμα δεν είμαστε σε θέση να υποθέσουμε ισότητα των διασπορών των δύο πληθυσμών, θα χρησιμοποιούσαμε την τελευταία αυτή μέθοδο για την κατασκευή του διαστήματος εμπιστοσύνης οπότε ο αριθμός v των βαθμών ελευθερίας για την κατανομή t που θα χρησιμοποιούσαμε θα ήταν,

$$v = \frac{\left(\frac{29.44}{9} + \frac{20.03}{9} \right)^2}{\frac{\left(\frac{29.44}{9} \right)^2}{8} + \frac{\left(\frac{20.03}{9} \right)^2}{8}}$$

Επομένως, το 95% διάστημα εμπιστοσύνης για την διαφορά $\mu_X - \mu_Y$ είναι,

$$(35.22 - 31.56) \pm (2.120) \sqrt{24.44 / 9 + 20.03 / 9}$$

ή ισοδύναμα,

$$3.66 \pm (2.120) \sqrt{4.95}$$

επομένως,

$$3.66 \pm 4.71$$

δηλαδή τελικά,

$$(-1.05, 8.37)$$

Παρατηρούμε ότι το διάστημα εμπιστοσύνης που προέκυψε με την μέθοδο αυτή ελάχιστα διαφέρει από εκείνο που είχαμε υπολογίσει προηγουμένως.

Χρησιμοποίηση του στατιστικού πακέτου STATGRAPHICS

Και στην περίπτωση των δύο δειγμάτων μπορούμε να χρησιμοποιήσουμε το στατιστικό πακέτο STATGRAPHICS για να κάνουμε την ανάλυση που μόλις εξηγήσαμε.

Από το βασικό μενού επιλέγουμε το ESTIMATION AND TESTING. Στην συνέχεια, από αυτό επιλέγουμε τη δυνατότητα 2 (TWO SAMPLE ANALYSIS). Μετά την εισαγωγή των στοιχείων στα πεδία SAMPLE 1 και SAMPLE 2, αντίστοιχα με το F6 παίρνουμε τον πίνακα που ακολουθεί.

TWO-SAMPLE ANALYSIS RESULTS

	SAMPLE 1	SAMPLE 2
POOLED		
SAMPLE STATISTICS:NUMBER OF OBS.	9	18
AVERAGE	35.2222	31.5556
VARIANCE	24.4444	20.0278
STD. DEVIATION	4.94413	4.47524
MEDIAN	35	31

DIFFERENCE BETWEEN MEANS = 3.66667

CONF. INTERVAL FOR DIFF. IN MEANS: 95 PERCENT

(EQUAL VARS.) SAMPLE 1-SAMPLE 2 -1.04687 8.38021
16D.F.

(UNEQUAL VARS.)SAMPLE 1-SAMPLE 2 -1.05066 8.38399
15.8D.F.

RATIO OF VARIANCES = 1.22053

CONF. INTERVAL FOR RATIO OF VARIANCES: 0 PERCENT
SAMPLE 1 ÷ SAMPLE 2

HYPOTHESIS TEST FOR H0:DIFF=0 COMPUTED

t STATISTIC = 1.64948

VS ALT: NE SIG. LEVEL = 0.11854

AT ALPHA = .05 SO DO NOT REJECT H0.

Στον πίνακα αυτό, έχουμε τις δειγματικές στατιστικές συναρτήσεις (*sample statistics*) για τα δύο δείγματα (SAMPLE 1, SAMPLE 2) και για το σταθμισμένο δείγμα (POOLED). Για κάθε μια από τις περιπτώσεις αυτές, ο πίνακας μας δίνει τις τιμές του μέσου (AVERAGE), της διασποράς (VARIANCE), της τυπικής απόκλισης (STD. DEVIATION) και της διαμέσου (MEDIAN).

Στην συνέχεια, δίνει τις διαφορές των δειγματικών μέσων (DIFFERENCE BETWEEN MEANS) και το διάστημα εμπιστοσύνης για τη διαφορά των δύο μέσων (CONF. INTERVAL FOR DIFF. IN MEANS) για επίπεδο που επιλέγει ο ερευνητής (στην περίπτωσή μας έχουμε επιλέξει το 95%). Στην συνέχεια, παρέχονται τα σχετικά αποτελέσματα για την περίπτωση που η υπόθεση των ίσων διασπορών ισχύει (EQUAL VARS.) μαζί με τους βαθμούς ελευθερίας, όπως επίσης και για την περίπτωση ανίσων διασπορών (UNEQUAL VARS.) μαζί με τους βαθμούς ελευθερίας, πριν από την στρωγγυλοποίηση, σύμφωνα με τον τύπο που έχουμε δώσει.

Έτσι, για την περίπτωση των ίσων διασπορών δίνει διάστημα εμπιστοσύνης το (-1.04687, 8.38021) που προκύπτει από 16 βαθμούς ελευθερίας ενώ για την περίπτωση ανίσων διασπορών δίνει διάστημα εμπιστοσύνης το (-1.05066, 8.38399) με 15.8 βαθμούς ελευθερίας.

Το υπόλοιπο μέρος του πίνακα αναφέρεται στον λόγο των διασπορών επίσης και στον έλεγχο υποθέσεων για μέσες τιμές που θα συναντήσουμε αργότερα.

Λύση με το πακέτο SPSS

Η κατασκευή του διαγράμματος ελέγχου της κανονικότητας στο πρόβλημα μπορεί να γίνει από το πακέτο SPSS ως εξής:

- Εισάγουμε τα δεδομένα των δύο μεθόδων εκπαίδευσης X και Y σε μια μεταβλητή π.χ. XY.
- Εισάγουμε μια καινούργια μεταβλητή π.χ. GR τις οποίας οι τιμές είναι 1 και 2 που υποδηλώνουν σε ποιά μέθοδο εκπαίδευσης αντιστοιχεί η κάθε παρατήρηση της μεταβλητής XY.
- Από την επιλογή Statistics, επιλέγουμε Compare means.

- Επιλέγουμε Independent-Samples T Test.
- Στο παράθυρο που εμφανίζεται στο πεδίο Test Variables, επιλέγουμε την μεταβλητή XY και στο πεδίο Grouping variable, επιλέγουμε την μεταβλητή GR.
- Επιλέγουμε Define groups.
- Στο παράθυρο που εμφανίζεται, επιλέγουμε Use specified groups και στο Group 1 βάζουμε την τιμή 1 που αντιστοιχεί στην μέθοδο εκπαίδευσης X και Group 2 βάζουμε την τιμή 2 που αντιστοιχεί στην μέθοδο εκπαίδευσης Y και επιλέγουμε continue.
- Στο αρχικό παράθυρο, επιλέγουμε OK.

Group Statistics

	GR	N	Mean	Std. Deviation	Std. Error Mean
XY	1.00	9	35.2222	4.9441	1.6480
	2.00	9	31.5556	4.4752	1.4917

Independent Samples Test

		Levene's Test for quality 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
XY	Equal variances assumed	.061	.807	1.649	16	.119	3.6667	2.2229	-1.0457	8.3790
	Equal variances not assumed			1.649	15.844	.119	3.6667	2.2229	-1.0495	8.3828

Στον πρώτο πίνακα αποτελεσμάτων, φαίνονται τα ονόματα των μεταβλητών (XY, GR), ο αριθμός των παρατηρήσεων (N), ο δειγματικός μέσος (Mean), η τυπική απόκλιση (Std. Deviation) και το τυπικό σφάλμα του μέσου (Std. Error Mean) για κάθε μέθοδο εκπαίδευσης.

Στον δεύτερο πίνακα αποτελεσμάτων, έχουμε δύο σειρές αποτελεσμάτων, πρώτα για την περίπτωση που έχουμε ίσες διακυμάνσεις (παράδειγμα σελίδα 58) και μετά για την περίπτωση που έχουμε άνισες διακυμάνσεις (παράδειγμα σελίδα 60). Στην αρχή, υπάρχει ο έλεγχος του Leven (Leven's Test for Equality of Variances) για το αν μπορούμε να δεχθούμε την υπόθεση ότι οι διακυμάνσεις των

δύο μεταβλητών είναι ίδιες. Βλέπουμε ότι το p-value είναι 0.807, άρα δεχόμαστε την υπόθεση της ισότητας των διακυμάνσεων. Στην συνέχεια, υπάρχουν κάποια στοιχεία που αφορούν τον έλεγχο στατιστικής υπόθεσης (που θα δούμε αργότερα) και το διάστημα εμπιστοσύνης (95% Confidence Interval of the Difference).

Λύση με τη χρήση του πακέτου Minitab

Η κατασκευή του διαστήματος εμπιστοσύνης στο πρόβλημα μπορεί να γίνει από το πακέτο Minitab ως εξής:

- Στο παράθυρο Data, εισάγουμε τα δεδομένα σε δύο μεταβλητές π.χ. C1, C2.
- Από την επιλογή Stat, επιλέγουμε Basic Statistics.
- Επιλέγουμε 2-Sample t...
- Επιλέγουμε Samples in different columns, στο First τοποθετούμε την μεταβλητή C1 και στο Second την μεταβλητή C2 και OK.

```
MTB > TwoSample 95.0 C1 C2;  
SUBC> Alternative 0.
```

```
TWOSAMPLE T FOR C1 VS C2  
      N   MEAN   STDEV  SE MEAN  
C1   9   35.22   4.94    1.6  
C2   9   31.56   4.48    1.5
```

```
95 PCT CI FOR MU C1 - MU C2: (-1.1, 8.4)
```

```
TTEST MU C1 = MU C2 (VS NE): T= 1.65 P=0.12 DF= 15
```

Στον πίνακα αποτελεσμάτων, έχουμε πρώτα τα ονόματα των μεταβλητών (C1, C2), μετά τον αριθμό των παρατηρήσεων (N), τον δειγματικό μέσο (Mean), την τυπική απόκλιση (StDev), το τυπικό σφάλμα του μέσου (SE Mean) για κάθε μεταβλητή. Στην συνέχεια, έχουμε το διάστημα εμπιστοσύνης (95 PCT CI FOR MU C1 - MU C2) και τέλος κάποια στοιχεία για τον έλεγχο υπόθεσης.

Σημείωση 1: Υπάρχει η δυνατότητα τα δεδομένα να εισαχθούν σε μια μεταβλητή και να κατασκευάσουμε μια δεύτερη μεταβλητή που θα δηλώνει σε ποια μέθοδο εκπαίδευσης αντιστοιχεί κάθε παρατήρηση (όπως στον τρόπο εισαγωγής των δεδομένων στο SPSS).

Σημείωση 2: Αν θέλουμε διάστημα εμπιστοσύνης διαφορετικό από 95%, επιλέγουμε Options στο παράθυρο Independent-Samples T Test και στην συνέχεια δηλώνουμε το επιθυμητό επίπεδο εμπιστοσύνης και Continue.

Σημείωση 3: Το πρόβλημα που εξετάσαμε στην τελευταία ενότητα, το οποίο αναφέρεται σε πληθυσμούς με διαφορετικές διασπορές στην βιβλιογραφία χαρακτηρίζεται ως πρόβλημα των Behrens-Fisher από τα ονόματα των επιστημόνων οι οποίοι πρώτοι το αντιμετώπισαν.

Δ. Παρατηρήσεις Κατά Ζεύγη

Στις τρεις προηγούμενες ενότητες, εξετάσαμε μεθόδους κατασκευής διαστημάτων εμπιστοσύνης για την διαφορά των μέσων τιμών δύο πληθυσμών που είναι ανεξάρτητοι με μεθοδολογία που βασίζεται σε επιλογή ανεξαρτήτων τυχαίων δειγμάτων. Υπάρχουν όμως περιπτώσεις που οι παρατηρήσεις του ενός δείγματος δεν είναι ανεξάρτητες από τις παρατηρήσεις στο άλλο δείγμα. Η εξάρτηση μπορεί να εμφανισθεί, είτε διότι οι παρατηρήσεις στα δύο δείγματα έχουν επιλεγεί κατά ζεύγη με βάση κάποιο χαρακτηριστικό, είτε διότι έχουμε επαναλαμβανόμενες παρατηρήσεις πάνω στα ίδια άτομα ή στοιχεία. Και στις δύο αυτές περιπτώσεις, η μεταβλητή που χρησιμοποιείται για την μελέτη του προβλήματος είναι εκείνη που αναφέρεται στη διαφορά των τιμών των παρατηρήσεων και όχι σ' αυτές καθαυτές τις παρατηρήσεις.

Ένα παράδειγμα προβλήματος που εμπίπτει στην πρώτη κατηγορία είναι αυτό που αναφέρεται στην μελέτη της αποτελεσματικότητας ενός νέου φαρμάκου. Για παράδειγμα, έστω ότι μας ενδιαφέρει αν ένα νέο φάρμακο για την χοληστερίνη ελαττώνει τον κίνδυνο καρδιακής προσβολής. Όπως είναι φυσικό, υπάρχουν και άλλοι παράγοντες που επηρεάζουν τον κίνδυνο καρδιακής

προσβολής, όπως η ηλικία, η κληρονομικότητα, το κάπνισμα, η πίεση κ.λ.π. Αν λοιπόν σε μια τέτοια περίπτωση παίρναμε δύο ανεξάρτητα δείγματα, οι παρατηρήσεις θα είχαν μεγάλες τυπικές αποκλίσεις μέρος των οποίων θα οφειλόταν σε άλλους παράγοντες και όχι σ' αυτούς που θα θέλαμε να εξετάσουμε. Ένας τρόπος να αποφύγουμε το πρόβλημα αυτό είναι να πάρουμε ένα τυχαίο δείγμα από άτομα με τα ίδια περίπου χαρακτηριστικά και στην συνέχεια να το χωρίσουμε σε δύο δείγματα, έτσι ώστε σε κάθε άτομο που ανήκει στην ελεγχόμενη ομάδα να αντιστοιχεί ένα άτομο της ομάδας που θα υποστεί τη θεραπεία. Η αντιστοίχιση θα γίνει με βάση την ομάδα ηλικίας, το επίπεδο χοληστερίνης, τις συνήθειες καπνίσματος, την πίεση του αίματος και όλους τους άλλους παράγοντες που είναι δυνατόν να θεωρηθεί ότι επηρεάζουν το ενδεχόμενο καρδιακής προσβολής. Με τον τρόπο αυτό, είναι σαν να έχουμε, κατά προσέγγιση, το ίδιο άτομο ταυτόχρονα και στην ελεγχόμενη ομάδα και στην ομάδα που θα υποστεί τη θεραπεία.

Ως παράδειγμα της δεύτερης κατηγορίας, ας θεωρήσουμε την περίπτωση που μας ενδιαφέρει να ελέγξουμε και να συγκρίνουμε την ποιότητα δύο διαφορετικών ειδών λαδιού αυτοκινήτου. Αν πάρουμε δύο ανεξάρτητα δείγματα και ελέγξουμε την απόδοση της καθεμιάς από τις δύο μάρκες λαδιού, σε καθένα από τα δύο δείγματα είναι αναμενόμενο να έχουμε μεγάλες αποκλίσεις που να οφείλονται σε άλλους παράγοντες (π.χ. κυβισμός αυτοκινήτου, ποιότητα ελαστικών κ.λ.π.) και όχι σε αυτόν που πραγματικά θέλουμε να ελέγξουμε. Σε τέτοιες περιπτώσεις, παίρνουμε ένα μόνο τυχαίο δείγμα από αυτοκίνητα, τοποθετούμε το λάδι της μιας κατασκευάστριας εταιρείας στα αυτοκίνητα αυτά και ελέγχουμε τα χιλιόμετρα που τα αυτοκίνητα αυτά θα διανύσουν μέχρις ότου τα λάδια φθάσουν σε ένα συγκεκριμένο επίπεδο φθοράς. Στην συνέχεια, χρησιμοποιούμε τα ίδια αυτοκίνητα (κατά προτίμηση με τους ίδιους οδηγούς και κάτω από τις ίδιες συνθήκες οδήγησης) και ελέγχουμε την δεύτερη μάρκα λαδιού όσον αφορά τις αποστάσεις που διανύονται μέχρις ότου το λάδι φθάσει στο ίδιο επίπεδο φθοράς όπως και προηγούμενα.

Και στις δύο περιπτώσεις που προαναφέραμε, θα έχουμε δύο σειρές μετρήσεων X_1, X_2, \dots, X_n και Y_1, Y_2, \dots, Y_n .

Δοθέντος όμως ότι τα δείγματα δεν θα είναι ανεξάρτητα, θεωρούμε ότι έχουμε n ζευγάρια παρατηρήσεων (X_i, Y_i) , $i=1,2,\dots,n$. Στην συνέχεια θεωρούμε τις διαφορές των παρατηρήσεων $D_i=X_i - Y_i$, $i=1,2,\dots,n$. Καταλήγουμε δηλαδή τελικά σε ένα πρόβλημα n παρατηρήσεων D_1, D_2, \dots, D_n , οπότε η ανάλυσή μας θα είναι ανάλυση που αναφέρεται ουσιαστικά σε ένα μόνο δείγμα. Για το δείγμα των διαφορών θα έχουμε,

$$\bar{D} = \frac{\sum_{i=1}^n D_i}{n}$$

$$S_D^2 = \frac{\sum_{i=1}^n (D_i - \bar{D})^2}{n}$$

ή, ισοδύναμα,

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

Ακολουθώντας πιά τη γνωστή μεθοδολογία, θα έχουμε ότι το $100(1-\alpha)\%$ διάστημα εμπιστοσύνης για τη διαφορά των μέσων τιμών των δύο πληθυσμών $\mu_X - \mu_Y$ θα είναι το,

$$\bar{D} \pm t_{n-1, 1-\alpha/2} \frac{S_D}{\sqrt{n-1}}$$

ή ισοδύναμα,

$$\bar{D} \pm t_{n-1, 1-\alpha/2} \frac{S_D^*}{\sqrt{n}}$$

Σημείωση: Εξυπακούεται ότι η συμπερασματολογία αυτή ισχύει με την προϋπόθεση ότι οι διαφορές των παρατηρήσεων ακολουθούν την κανονική κατανομή.

Παράδειγμα: Προκειμένου να γίνει σύγκριση της ποιότητας δύο ειδών λαδιού αυτοκινήτου, μια εταιρεία προστασίας καταναλωτών ενδιαφέρεται να κατασκευάσει ένα 95% διάστημα εμπιστοσύνης για τη διαφορά $\mu_D=\mu_X-\mu_Y$ της μέσης κατανάλωσης βενζίνης (μετρούμενης σε χιλιόμετρα/λίτρο) με την χρησιμοποίηση των δύο διαφορετικών

ειδών λαδιού. Για τον λόγο αυτό, χρησιμοποιούνται τέσσερα αυτοκίνητα τα οποία δοκιμάζονται σε απόσταση 1000km, την πρώτη φορά, χρησιμοποιώντας το λάδι μηχανής τύπου A και την δεύτερη φορά, χρησιμοποιώντας το λάδι μηχανής τύπου B. Οι μετρήσεις στις οποίες η εταιρεία κατέληξε δίνονται στον πίνακα που ακολουθεί.

	Αυτοκίνητο			
	1	2	3	4
Λάδι A: X_i	19.77	18.90	20.20	16.29
Λάδι B: Y_i	18.91	18.21	18.84	16.92
$D_i = X_i - Y_i$	0.86	0.69	1.36	-0.33

Από τις παρατηρήσεις αυτές έχουμε:

$$\bar{D} = 0.645$$

$$S_D^{*2} = 0.503$$

Δοθέντος ότι $t_{0.975,3}=3.182$, θα έχουμε ότι το 95% διάστημα εμπιστοσύνης για την διαφορά $\mu_D = \mu_X - \mu_Y$ θα είναι το,

$$0.645 \pm 3.182 \frac{0.709}{2}$$

ή

$$(-0.485, 1.775)$$

Λύση με το πακέτο SPSS

Η κατασκευή του διαστήματος εμπιστοσύνης στο πρόβλημα μπορεί να γίνει από το πακέτο SPSS ως εξής:

- Εισάγουμε τα δεδομένα για το λάδι A στη μεταβλητή π.χ. X.
- Εισάγουμε τα δεδομένα για το λάδι B στη μεταβλητή π.χ. Y.
- Από την επιλογή Statistics, επιλέγουμε Compare means.
- Επιλέγουμε Paired Samples T-test.
- Στο παράθυρο που εμφανίζεται, επιλέγουμε τις μεταβλητές μας (X και Y) τις τοποθετούμε στο πεδίο Paired Variables και επιλέγουμε OK.

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	X	18.7900	4	1.7522	.8761
	Y	18.2200	4	.9221	.4610

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	X & Y	4	.990	.010

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	X - Y	.5700	.8490	.4245	-.7810	1.9210	1.343	3	.272

Στον πρώτο πίνακα αποτελεσμάτων, φαίνονται τα ονόματα των μεταβλητών (X, Y), ο δειγματικός μέσος (Mean), ο αριθμός των παρατηρήσεων (N), η τυπική απόκλιση (Std. Deviation) και το τυπικό σφάλμα του μέσου (Std. Error Mean) για κάθε λάδι.

Στον δεύτερο πίνακα αποτελεσμάτων, έχουμε την συσχέτιση μεταξύ των μεταβλητών X και Y. Στον τρίτο πίνακα, έχουμε κάποια στοιχεία που αφορούν τον έλεγχο στατιστικής υπόθεσης (που θα δούμε αργότερα) και το διάστημα εμπιστοσύνης (95% Confidence Interval of the Difference).

Σημείωση: Αν θέλουμε διάστημα εμπιστοσύνης διαφορετικό από 95%, επιλέγουμε Options στο παράθυρο Paired-samples T Test και στην συνέχεια, δηλώνουμε το επιθυμητό επίπεδο εμπιστοσύνης και Continue.

Λύση με τη χρήση του πακέτου Minitab

Η κατασκευή του διαστήματος εμπιστοσύνης στο πρόβλημα μπορεί να γίνει από το πακέτο Minitab ως εξής:

- Στο παράθυρο Data, εισάγουμε τα δεδομένα σε δυο μεταβλητές π.χ. C1,C2 για κάθε τύπο λαδιού.

- Από την επιλογή Calc, επιλέγουμε Mathematical Expressions.
- Στο πεδίο Variable, τοποθετούμε τη μεταβλητή C3.
- Στο πεδίο expression, γράφουμε C1-C2 και επιλέγουμε OK.
- Από την επιλογή Stat, επιλέγουμε Basic Statistics.
- Επιλέγουμε 1-Sample t...
- Στο παράθυρο που ανοίγει, επιλέγουμε τη μεταβλητή που έχουμε περάσει τη διαφορά των δεδομένων δηλαδή C3 και την τοποθετούμε στο πεδίο Variables.
- Επιλέγουμε διάστημα εμπιστοσύνης (Confidence Interval) και στο πεδίο level επιλέγουμε 95% και OK.

MTB > TInterval 95.0 C3.

	N	MEAN	STDEV	SE MEAN	95.0 PERCENT C.I.
C3	4	0.570	0.849	0.425	(-0.781, 1.921)

Στον πίνακα αποτελεσμάτων, έχουμε πρώτα το όνομα της μεταβλητής (C3), μετά τον αριθμό των παρατηρήσεων (N), τον δειγματικό μέσο (Mean), την τυπική απόκλιση (StDev), το τυπικό σφάλμα του μέσου (SE Mean) και τέλος το διάστημα εμπιστοσύνης (95.0% C.I.).

Σημείωση: Η επιλογή Test mean χρησιμοποιείται για τον έλεγχο υποθέσεων όπως θα δούμε στη συνέχεια.

Παρατήρηση: Η ανάλυση του προβλήματος με παρατηρήσεις κατά ζεύγη στο στατιστικό πακέτο Statgraphics δίνεται με την θεώρηση των διαφορών των παρατηρήσεων των δύο δειγμάτων ως παρατηρήσεις που προέρχονται από ένα δείγμα και την στη συνέχεια ανάλυσή τους με την μέθοδο του ενός δείγματος.

Σχηματική Παρουσίαση της Διαδικασίας Κατασκευής Διαστημάτων Εμπιστοσύνης για την Διαφορά Μέσων Τιμών Κανονικών Πληθυσμών

Ο τρόπος κατασκευής διαστημάτων εμπιστοσύνης για την διαφορά μέσων τιμών κανονικών πληθυσμών και οι ενέργειες που πρέπει να γίνουν ανάλογα με την σχέση των διακυμάνσεων των δύο

αυτών πληθυσμών μπορεί να παρουσιασθεί με το παρακάτω διάγραμμα.

**Διαστήματα Εμπιστοσύνης για την Διαφορά Μέσων Τιμών
Κανονικών Πληθυσμών**

