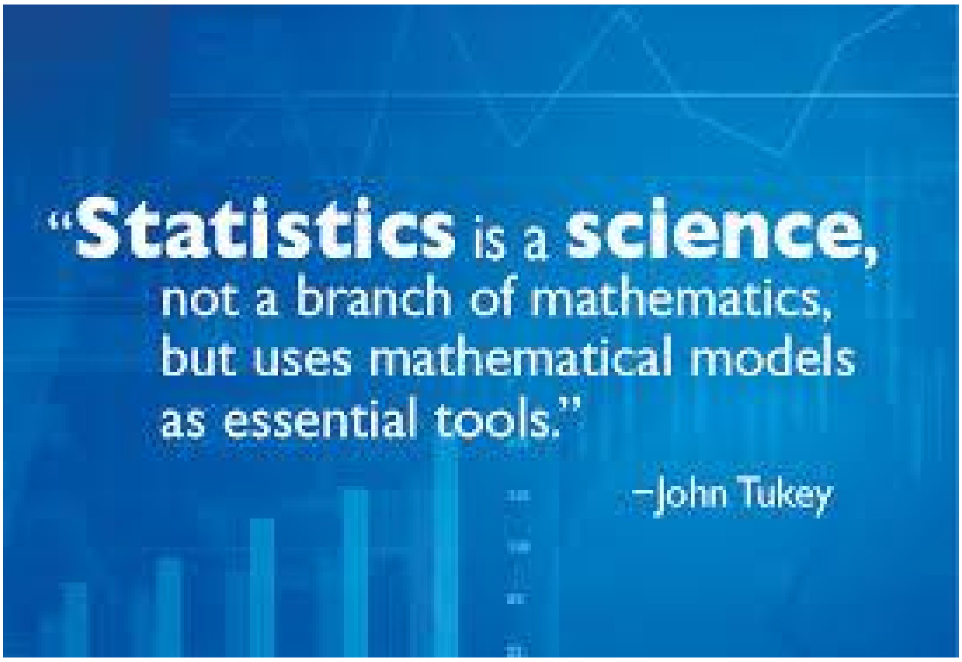
ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ ΔΥΤΙΚΗΣ ΕΛΛΑΔΑΣ
ΣΧΟΛΗ: ΔΙΟΙΚΗΣΗΣ ΚΑΙ ΟΙΚΟΝΟΜΙΑΣ
ΤΜΗΜΑ: ΔΙΟΙΚΗΣΗ ΕΠΙΧΕΙΡΗΣΕΩΝ



ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

«ΠΑΡΑΜΕΤΡΙΚΟΙ ΚΑΙ ΜΗ ΠΑΡΑΜΕΤΡΙΚΟΙ ΕΛΕΓΧΟΙ ΥΠΟΘΕΣΕΩΝ ΚΑΙ ΕΦΑΡΜΟΓΗ ΤΟΥΣ ΣΤΙΣ ΕΠΙΧΕΙΡΗΣΕΙΣ»

ΣΠΟΥΔΑΣΤΕΣ: ΞΕΝΟΣ ΝΙΚΟΛΑΟΣ-ΣΑΡΑΝΤΗΣ ΚΑΙ ΚΡΙΤΣΩΤΑΚΗ ΝΑΤΑΛΙΑ
ΕΠΙΒΛΕΠΟΥΣΑ ΚΑΘΗΓΗΤΡΙΑ: ΠΑΠΑΘΑΝΑΣΟΠΟΥΛΟΥ ΧΡΥΣΑΝΘΗ

Πάτρα-2018

Πρόλογος

Η συγγραφή της πτυχιακής εργασίας πραγματοποιείται στο πλαίσιο του Προπτυχιακού προγράμματος σπουδών του τμήματος Επιχειρηματικού Σχεδιασμού και Πληροφοριακών Συστημάτων του Τεχνολογικού Ιδρύματος Δυτικής Ελλάδας.

Το θέμα που πραγματεύεται η παρούσα εργασία εστιάζεται στους Παραμετρικούς και Μη Παραμετρικούς ελέγχους υποθέσεων και την εφαρμογή τους στις επιχειρήσεις. Ο σκοπός μας, δηλαδή, είναι να παρουσιάσουμε, να περιγράψουμε και να αναλύσουμε με τρόπο απλό και κατανοητό όλα όσα χρειάζεται κανείς να γνωρίζει για την επίλυση προβλημάτων ελέγχου υποθέσεων σε Παραμετρικά ή μη Παραμετρικά δεδομένα.

Συμπερασματικά, σε κάθε κεφάλαιο θα παραθέσουμε την θεωρία της κάθε μεθόδου ξεχωριστά, ως προς την κατανόηση αυτής, καθώς και παραδείγματα ανάλυσης δεδομένων είτε χειρόγραφα, είτε με την βοήθεια των στατιστικών λογισμικών SPSS, Minitab ή Excel, όπου αυτό είναι εφικτό.

Στο σημείο αυτό, θα θέλαμε να ευχαριστήσουμε θερμά την επιβλέποντα καθηγήτριά μας Χρυσάνθη Παπαθανασοπούλου για την συμβολή, την πολύτιμη βοήθεια της και την μετάδοση της αγάπης της προς την Στατική σε εμάς.

Επίσης, θα θέλαμε να ευχαριστήσουμε και τους γονείς μας για την υποστήριξή τους σε κάθε βήμα μας.

Περίληψη

Η παρούσα εργασία καταπιάνεται με την εφαρμογή των παραμετρικών και των μη παραμετρικών μεθόδων στις επιχειρήσεις. Πρόκειται για στατιστικές μεθόδους ελέγχου υποθέσεων, με ποσοτικά, ονομαστικά αλλά και διατακτικά δεδομένα , οι οποίες μας βοηθούν να αναλύσουμε και να εξετάσουμε όλα τα στοιχεία που χρειάζεται, για να καταλήξουμε στην ορθότερη δυνατή απόφαση.

Το ερώτημα είναι πως προέκυψε η ανάγκη για την εύρεση και στην συνέχεια εξέλιξη τέτοιων μεθόδων. Η απάντηση είναι ότι όταν έχουμε δύο υποθέσεις, πολλές φορές είναι δύσκολο έως αδύνατο να επιλέξουμε τη σωστή υπόθεση άμεσα. Μάλιστα η βαρύτητα της απόφασης μας αυξάνεται όταν η απόφαση που πρέπει να λάβουμε επηρεάζει σε μεγάλο βαθμό το μέλλον μας , οικονομικό και προσωπικό, αλλά και τρίτους. Παραδείγματος χάριν εάν χρειάζεται να αποφασίσουμε για την αθωότητα ή την ενοχή ενός ατόμου για κάποιο έγκλημα, την αύξηση του προσωπικού μιας επιχείρησης, την επέκταση της με μια νέα μονάδα σε άλλη πόλη εφόσον προβλέπουμε μελλοντική κερδοφορία εξ αυτής της κίνησης.

Όπως καταλαβαίνουμε η λήψη ορισμένων αποφάσεων είναι πολύ σημαντικό να γίνει σε πλαίσια σιγουριάς και υπευθυνότητας, κάτι το οποίο μας προσφέρουν τα στατιστικά στοιχεία σε μεγάλο βαθμό, γι' αυτό και στην παρούσα εργασία θα δούμε τους τρόπους με τους οποίους μπορούμε να φθάσουμε στα ασφαλέστερα δυνατά συμπεράσματα, χωρίς να στηριζόμαστε σε «άτοπες» προσωπικές εκτιμήσεις και προβλέψεις, κάτι το οποίο συνέβαινε αρκετά στο παρελθόν κυρίως, με καταστροφικές συνέπειες.

Για να κάνουμε όσο το δυνατόν πιο κατανοητές τις εν λόγω μεθόδους μελετήσαμε ενδεδειγμένα στατιστικά βιβλία και συγκεντρώσαμε δεδομένα, στατιστικά στοιχεία, καθώς και πραγματικά προβλήματα που αντιμετωπίζουν επιχειρήσεις. Μέσα απ' όλα αυτά θα επιχειρήσουμε να κατανοήσουμε εξ ολοκλήρου τη χρησιμότητα των παραμετρικών και των μη παραμετρικών μεθόδων σε πραγματικό χρόνο, καθώς και το γιατί η χρήση τους πλέον κρίνεται απαραίτητη. Ακόμη, για την επίλυση των προβλημάτων μας βοήθησαν ιδιαίτερα τα στατιστικά προγράμματα Minitab και SPSS, καθώς και το EXCEL, τα οποία με την μεγάλη ακρίβεια και τα εκτενέστερα αποτελέσματα τους, μας παρείχαν σημαντικές λεπτομέρειες και διευκολύνσεις σε όλη την προσπάθεια μας.

Τα αποτελέσματα των ερευνών μας όπως θα διαπιστώσουμε είναι άκρως ικανοποιητικά, καθώς μέσω της εφαρμογής των εν λόγω μεθόδων, δόθηκαν λύσεις σε προβλήματα τα οποία θα μπορούσαν να θέσουν ακόμη και το οικονομικό τέρμα μιας επιχείρησης με μια λανθασμένη εκτίμηση. Τη σήμερον ημέρα οι περισσότερες αποφάσεις που πρέπει να παρθούν σχετίζονται με μελλοντικές καταστάσεις. Αυτό δείχνει τη σημαντικότητα του να μπορέσουμε να απωλέσουμε όσο περισσότερα στοιχεία για το θέμα που μας αφορά, καθώς και να τα αξιολογήσουμε με τον καλύτερο δυνατό τρόπο ώστε να καταλήξουμε στα ασφαλέστερα δυνατά συμπεράσματα. Οι παραμετρικές και οι μη παραμετρικές μέθοδοι κρίνονται άκρως απαραίτητες γι' αυτόν τον λόγο.

Abstract

This study is dealing with the implementation of parametric and non-parametric methods in enterprises. These are statistical methods of hypothesis testing, with quantitative, nominal and prescriptive data, which help us analyze and examine all the elements we need in order to make the right decision.

The question is how the need has arisen for finding and then developing such methods. The answer is that when we have two assumptions, it is sometimes difficult or even impossible to choose the right case in a short amount of time. Indeed, the gravity of the situation increases when the decision we have to make affects not only our future, financially and personally, to a great extent, but other people as well. For example, deciding on a person's innocence or guilt in court, increasing a company's staff, or expanding it with a new unit to another city if we anticipate this move to bring about profit in the future.

Furthermore as we understand the importance of decision making, it is crucial that we do so in a confident and responsible manner, which is something that statistics can very successfully offer to us. This is why in this paper we will examine all the ways with which we can reach to the safest possible conclusions, without relying on “alleged” personal predictions and forecasts, something that has happened in many occasions, particularly in the past, with disastrous consequences.

In addition, in order to make these methods as comprehensible as possible, we have studied relevant statistical books and collected data, statistics, as well as real business problems. Through this, we will attempt to fully understand the usefulness of parametric and non-parametric methods in real time, as well as why their use is now deemed necessary. Moreover, statistical programs like Minitab, SPSS, and EXCEL, have also helped us solve our problems with great accuracy and their broader results have also provided us with significant details and convenience throughout our efforts.

As you will see, the results of our research are very satisfactory. Through the application of these methods, solutions were provided to problems that could even put an end to the financial aspect of a business due to a wrong estimation. In present time, most of the decisions that need to be made, relate to future situations. This shows the importance of being able to find extensive information on the subject, evaluate said information in most efficient way so we can come up with the safest possible conclusions. Parametric and non-parametric methods are considered essential for this reason.

Εισαγωγή

Παραμετρικοί και μη παραμετρικοί μέθοδοι. Σίγουρα για όσους το αντικείμενο τους δεν αγγίζει το πλαίσιο της στατιστικής αποτελούν άγνωστες παραμέτρους. Τι συμβαίνει όμως με όσους ασχολούνται με τη στατιστική, αλλά και όσους η στατιστική τους επηρεάζει έμμεσα ή άμεσα; Τότε η ομάδα των συγκεκριμένων ανθρώπων οφείλει να γνωρίζει τις ανωτέρω μεθόδους. Οι λόγοι, σαφείς.

Η ορθή εκτίμηση του ελέγχου μιας υπόθεσης για την καλύτερη απόδοση και αποτελεσματικότητα μιας επιχείρησης είναι καίρια, γιατί η υπερεκτίμηση μιας κατάστασης μπορεί να οδηγήσει σε ζημία της επιχείρησης λόγω παραπανίσιων εξόδων και υπεραξίας στην αγορά, ενώ η υποτίμηση μιας κατάστασης μπορεί να οδηγήσει σε οικονομικές απώλειες. Ως αποτέλεσμα αυτού, είναι σημαντική μια ερευνητική προσπάθεια να προλαμβάνει τις επιπτώσεις των τυχών λανθασμένων υποθέσεων.

Το θέμα της εργασίας μας κρίνεται κατά γενική ομολογία ιδιαίτερα σημαντικό, καθώς οι παραμετρικές και οι μη παραμετρικές μέθοδοι, πολλές φορές είναι στη θέση να δώσουν τις καλύτερες δυνατές λύσεις στα ανωτέρω προβλήματα, η κάθε λεπτομέρεια των οποίων είναι καθοριστική για το μέλλον, αλλά και το παρόν μιας επιχείρησης. Η εφαρμογή αυτών των μεθόδων όμως, δεν περιορίζεται στο πλαίσιο των επιχειρήσεων. Για παράδειγμα, ένας καθηγητής έχει τη δυνατότητα να κρίνει με αρκετά μεγάλη ακρίβεια ποιος εκ δύο φοιτητών είναι πιο αποτελεσματικός στις εξετάσεις, με τη βοήθεια των μεθόδων που θα μελετήσουμε, καθώς με τη βοήθεια των ανάλογων λογισμικών μας δίνεται η μεγαλύτερη δυνατή ακρίβεια των αποτελεσμάτων.

Στόχος της παρούσας εργασίας είναι να καταφέρει να κάνει όσο το δυνατόν πιο κατανοητές αυτές τις μεθόδους στο ευρύτερο κοινό, παρουσιάζοντας όχι μόνο τη θεωρητική τους υπόσταση, αλλά και την πρακτική τους εφαρμογή. Για το λόγο αυτό ανατρέξαμε σε βαρυσήμαντα βιβλία της στατιστικής, ώστε να σχεδιάσουμε το θεωρητικό μέρος, και για το πρακτικό μέρος της εργασίας μας, το οποίο είναι ίσως και το πιο σημαντικό, πραγματοποιήσαμε εκτενής έρευνα σε πλήθος επιχειρήσεων, απ’ τις οποίες συλλέξαμε στοιχεία για τη μισθοδοσία, την παραγωγική διαδικασία, και κυρίως από προβλήματα τα οποία απασχολούν τις επιχειρήσεις σε πραγματικό χρόνο, όπου η λάθος αντιμετώπιση και εκτίμηση των οποίων, μπορεί να αποβεί καταστροφική ακόμη και για τη φυσική παρουσία τους. Τέτοια προβλήματα μπορεί να είναι η πρόσληψη προσωπικού, η επιλογή ή μη ενός καινούργιου διευθυντή σε κάποιον τομέα με βάση την αποδοτικότητα του, η επέκταση ή μη της αλυσίδας καταστημάτων της επιχείρησης, ακόμη και η δημιουργία μιας νέας επιχείρησης με βάση την ενδεχόμενη κερδοφορία της.

Η παρούσα πτυχιακή εργασία μας αφορά όλους, αφού συμβάλει στη νέα γνώση, καθώς παρουσιάζεται με τρόπο σαφή και κατανοητό ως προς το ευρύτερο κοινό, δίνοντας τη δυνατότητα να κατανοήσει αλλά και να χρησιμοποιήσει τις συγκεκριμένες μεθόδους για προσωπικό όφελος. Παραδείγματος χάριν, εάν ένας ιδιοκτήτης μιας επιχείρησης σκέφτεται να προσλάβει καινούργιο υπάλληλο, αλλά δεν διαθέτει το κατάλληλο γνωστικό υπόβαθρο για να υπολογίσει κατά πόσο θα ναι επικερδής αυτή η κίνηση, με τη βοήθεια των μεθόδων που παραθέτουμε θα μπορούσε να καταλήξει σε ένα αρκετά ασφαλές συμπέρασμα για την εκτέλεση ή μη της συγκεκριμένης κίνησης.

Η πτυχιακή εργασία αποτελείται από δώδεκα κεφάλαια:

Στο πρώτο κεφάλαιο παρουσιάζονται οι βασικές έννοιες της στατιστικής που χρειάζονται για την κατανόηση των μεθόδων που αναλύονται στην συνέχεια.

Στο δεύτερο κεφάλαιο παρουσιάζονται οι διάφορες συναρτήσεις της κανονικής κατανομής όπου πρέπει να ακολουθεί κάθε δείγμα ενός πληθυσμού για την υλοποίηση του ελέγχου υποθέσεων με παραμετρικές μεθόδους.

Στο τρίτο κεφάλαιο κάνουμε μία εισαγωγή στο τι είναι ο έλεγχος υποθέσεων και πως προσδιορίζεται, το οποίο αποτελεί την βασικότερη έννοια της εργασίας αυτής.

Στο τέταρτο κεφάλαιο παραθέτουμε τον έλεγχο υποθέσεων για τον μέσο ενός δείγματος με γνωστή τυπική απόκλιση.

Στο πέμπτο κεφάλαιο αναλύουμε τις μεθόδους οι οποίες μετρούν το κατά πόσο είναι αξιόπιστο το αποτέλεσμα του ελέγχου υποθέσεων.

Στο έκτο κεφάλαιο θα μάθουμε πως βρίσκουμε το διάστημα τιμών που μπορεί να πάρει ο έλεγχος υποθέσεων.

Στο έβδομο κεφάλαιο γνωρίζουμε την έννοια της διασποράς, η οποία είναι ένας δείκτης μεταβλητότητας με σκοπό την εκτίμηση της παραμέτρου ενός πληθυσμού από το στατιστικό μέγεθος του δείγματος.

Στο όγδοο κεφάλαιο συναντάμε την περίπτωση που τα δεδομένα είναι ονομαστικά. Σε αυτές τις περιπτώσεις πραγματοποιούμε τον έλεγχο υποθέσεων μέσω της αναλογίας ενός πληθυσμού

Στο ένατο κεφάλαιο συναντάμε για πρώτη φορά δύο δείγματα τα οποία πρέπει να συγκρίνουμε. Τα δείγματα αυτά άλλοτε εμφανίζουν κοινά στοιχεία, και άλλοτε διαφέρουν ολοκληρωτικά.

Στο δέκατο κεφάλαιο ερχόμαστε σε επαφή με τη μέθοδος που εξετάζει τυχόν διαφορές ανάμεσα στους μέσους δύο ή περισσότερων πληθυσμών, οι οποίοι ακολουθούν την κανονική κατανομή.

Στο εντέκατο κεφάλαιο θα γνωρίσουμε τον έλεγχο χ^2, ο οποίος είναι ο έλεγχος υποθέσεων με εφαρμογή σε πληθυσμούς ονομαστικών δεδομένων, όπου η τυχαία μεταβλητή μπορεί να πάρει δύο ή περισσότερες τιμές.

Στο δωδέκατο κεφάλαιο θα παρουσιάσουμε την περίπτωση που τα δεδομένα μας είναι διατακτικά, δηλαδή όταν είναι στη μορφή κλάσεων και δεν μπορούν να εξεταστούν μεμονωμένα.

ΠΕΡΙΕΧΟΜΕΝΑ

Κεφάλαιο Ι: Εισαγωγή στην στατιστική

1. Τι είναι η στατιστική:

Στατιστική είναι η επιστήμη που ασχολείται με τη συλλογή δεδομένων, παρουσίαση και εν συνεχεία εξαγωγή συμπερασμάτων παρατηρήσεων που υπόκεινται σε τυχαίες μεταβολές.
Αυτή χωρίζεται στην α) περιγραφική και στην β) επαγωγική στατιστική:
α) Η περιγραφική στατιστική (descriptive statistics) ασχολείται με τις μεθόδους οργάνωσης, σύνοψης και παρουσίασης των δεδομένων με τρόπο εύχρηστο και κατανοητό. Στις μεθόδους τις περιλαμβάνονται γραφήματα και οι αριθμητικοί δείκτες (π.χ. αριθμητικός μέσος), που επιτρέπουν την εξαγωγή συμπερασμάτων και τη λήψη αποφάσεων.
β) Η επαγωγική στατιστική (inferential statistics) είναι ένα σύνολο μεθόδων που επιτρέπουν την προβολή δεικτών από ένα μικρό δείγμα σε ένα ευρύτερο σύνολο πληθυσμού.

2. Ορισμός παραμετρικών και μη παραμετρικών μεθόδων:

Παραμετρικές μέθοδοι: Βάση της υπόθεσης ότι τα δεδομένα έρχονται από κάποια βαθύτερη κατανομή της οποίας η γενική μορφή είναι γνωστή, οι στατιστικές μέθοδοι για αξιολόγηση και έλεγχο αυτής της υπόθεσης είναι τότε βασισμένες στην κατανομή. Ο στόχος μας βρίσκεται στις παραμέτρους και στον έλεγχο της υπόθεσης γύρω από αυτές.
Μη παραμετρικές μέθοδοι: Δεν γίνονται υποθέσεις πάνω στη βαθύτερη κατανομή. Το συμπέρασμα δεν επικεντρώνεται στις ειδικές πληθυσμιακές παραμέτρους και ίσως είναι λιγότερο δυνατές από τις αντίστοιχες παραμετρικές όταν οι υποθέσεις ικανοποιούνται.

3. Βασικές έννοιες:

I. Πληθυσμός:
Ο πληθυσμός (population) μιας στατιστικής έρευνας είναι το σύνολο των στοιχείων που ενδιαφέρουν τον στατιστικό. Συχνά ο πληθυσμός είναι πολύ μεγάλος και απροσδιόριστος. Στη «γλώσσα» της στατιστικής, πληθυσμός δεν σημαίνει απαραίτητα ένα σύνολο ανθρώπων. Για παράδειγμα, ο πληθυσμός μιας έρευνας θα μπορούσε να είναι το σύνολο των διαμέτρων των ρουλεμάν που κατασκευάζονται σε ένα εργοστάσιο. Κάθε περιγραφικό μέτρο ενός πληθυσμού ονομάζεται παράμετρος (parameter). Για παράδειγμα σε μια έρευνα αναψυκτικών που καταναλώνονται από φοιτητές σε μια πανεπιστημιούπολη, η παράμετρος είναι ο μέσος αριθμός αναψυκτικών.

II. Δείγμα:
Το δείγμα (sample) μιας στατιστικής έρευνας είναι το σύνολο των δεδομένων που συλλέγοντα από τον πληθυσμό. Όμως κάθε περιγραφικό μέτρο του δείγματος ονομάζεται στατιστικό μέγεθος (statistic), γιατί τα στατιστικά μεγέθη χρησιμοποιούνται ως βάση για την επαγωγική εκτίμηση των παραμέτρων του πληθυσμού.

III. Επαγωγή:

Στατιστική επαγωγή (statistical inference) είναι η διαδικασία της εκτίμησης, ή πρόβλεψης, μιας παραμέτρου του πληθυσμού με βάση τα δεδομένα ενός δείγματος. Επειδή οι πληθυσμοί είναι σχεδόν πάντοτε πολύ μεγάλοι, η άμεση καταγραφή των δεδομένων για το σύνολο του πληθυσμού είτε είναι ανέφικτη είτε έχει πολύ μεγάλο κόστος. Είναι πολύ ευκολότερο και πιο οικονομικό να ερευνηθεί ένα μικρό δείγμα του πληθυσμού και από τα στατιστικά μεγέθη του δείγματος να γίνουν εκτιμήσεις για τις παραμέτρους του πληθυσμού. Βέβαια, οι εκτιμήσεις αυτές δεν είναι πάντοτε ακριβείς, και για τον λόγο αυτό η στατιστική επαγωγή συνοδεύεται πάντοτε από ένα μέτρο του βαθμού ακριβείας των προβλέψεων. Υπάρχουν δύο δείκτες ακριβείας: η στάθμη εμπιστοσύνης (confidence level) και η στάθμη σημαντικότητας (significance level). Η στάθμη εμπιστοσύνης είναι το ποσοστό επιβεβαίωσης των προβλέψεων στην πράξη. Για παράδειγμα αν σε μια έρευνα η παράμετρος έχει στάθμη εμπιστοσύνης 95% αυτό σημαίνει ότι αν η στατιστική μέτρηση πραγματοποιηθεί σε μεγάλο αριθμό επαναλαμβανόμενων δειγμάτων, τότε το 95% των εκτιμήσεων θα δώσουν το σωστό αποτέλεσμα. Αντίθετα η στάθμη σημαντικότητας είναι το ποσοστό των εκτιμήσεων που θα αποδειχθούν εσφαλμένες. Δηλαδή αν σε μια έρευνα η παράμετρος έχει η στάθμη σημαντικότητας 5%, αυτό σημαίνει ότι αν η στατιστική μέτρηση πραγματοποιηθεί σε μεγάλο αριθμό επαναλαμβανόμενων δειγμάτων, τότε το 5% των εκτιμήσεων θα αποδειχθούν εσφαλμένες.

4. Τύποι δεδομένων και πληροφοριών:

Ο αντικειμενικός στόχος της στατιστικής είναι η εξαγωγή πληροφοριών από δεδομένα. Υπάρχουν διάφοροι τύποι δεδομένων και πληροφοριών:

I. Μεταβλητή είναι ένα χαρακτηριστικό ενός πληθυσμού ή δείγματος. Για παράδειγμα, ο βαθμός στο μάθημα της στατιστικής είναι μια μεταβλητή του πληθυσμού των φοιτητών που έχουν εξεταστεί στο μάθημα.
II. Τιμή μιας μεταβλητής είναι κάθε παρατηρήσιμη κατάσταση αυτής της μεταβλητής. Οι τιμές του βαθμού της στατιστικής είναι αριθμοί μεταξύ 0 έως 20.
III. Δεδομένα είναι το σύνολο των τιμών που έχουν παρατηρηθεί για μια μεταβλητή. Για παράδειγμα, οι βαθμοί 10 φοιτητών στο μάθημα της στατιστικής σε μια εξέταση είναι: 10, 15, 16, 13, 18, 15, 9, 20, 19, 17. Οι δέκα αυτοί αριθμοί είναι το σύνολο των δεδομένων, από το οποίο θα προσπαθήσουμε να εξάγουμε τις πληροφορίες που επιθυμούμε.

Οι περισσότεροι άνθρωποι πιστεύουν ότι τα δεδομένα είναι πάντοτε ένα σύνολο αριθμών. Στην πραγματικότητα υπάρχουν 3ις τύποι δεδομένων:

1. Συνεχή δεδομένα είναι αυτά που εκφράζονται από πραγματικούς αριθμούς, όπως το ύψος, το βάρος, το εισόδημα και η απόσταση. Τα δεδομένα αυτά ονομάζονται επίσης ποσοτικά ή αριθμητικά. Οι τιμές των συνεχών δεδομένων αντιπροσωπεύουν πραγματικά μεγέθη και έτσι επιτρέπονται κάθε είδους υπολογισμοί πάνω στις τιμές αυτές. Στη συνέχεια θα χρησιμοποιείται πάντα ο όρος ποσοτικά.
2. Ονομαστικά δεδομένα είναι αυτά που εκφράζουν μη αριθμητικές ιδιότητες του πληθυσμού, όπως η οικογενειακή κατάσταση: άγαμος, έγγαμος, διαζευγμένος, χήρος. Συχνά τα ονομαστικά δεδομένα αποθηκεύονται με μια μορφή αρίθμησης π.χ. 1=άγαμος, 2=έγγαμος, 3=διαζευγμένος, 4=χήρος. Αυτό όμως γίνεται μόνο για λόγους απλούστευσης και δεν

συνδέεται με τη φύση των τιμών. Στο ίδιο παράδειγμα θα μπορούσε να χρησιμοποιηθεί οποιαδήποτε άλλη αρίθμηση χωρίς να αλλάζει η σημασία των δεδομένων. Τα ονομαστικά δεδομένα λέγονται επίσης και ποσοτικά ή κατηγορικά. Επειδή η αρίθμηση (κωδικοποίηση) των ονομαστικών δεδομένων είναι απόλυτα αυθαίρετη, δεν επιτρέπεται η εκτέλεση υπολογισμών πάνω στις τιμές της αρίθμησης αυτής γιατί οι αριθμοί αυτοί δεν εκφράζονται ως ποσοτικά δεδομένα. Το μόνο που μπορούμε να κάνουμε με αυτά είναι να καταμετρήσουμε τη συχνότητα κάθε τιμής.

3. Διατακτικά (ή διατάξιμα) δεδομένα είναι αυτά που παίρνουν ένα μικρό πλήθος από διακριτές τιμές, όπως τα ονομαστικά δεδομένα, αλλά μεταξύ των τιμών αυτών υπάρχει η σχέση διάταξης. Η διαφορά μεταξύ ονομαστικών και διατακτικών δεδομένων έχει σημασία και δεν μπορεί να αλλάξει. Συνεπώς, τα δεδομένα μπορούν να κωδικοποιηθούν με μια μορφή αρίθμησης, με τον περιορισμό ότι αυτή δεν θα αλλάζει την σειρά των τιμών. Για παράδειγμα μία αποδεκτή αρίθμηση θα ήταν: 6=κακή, 18=μέτρια, 23=καλή, 45=πολύ καλά, 88=άριστη. Το ιδιαίτερο χαρακτηριστικό των διατακτικών δεδομένων είναι η διάταξη. Κατά συνέπεια, οι μόνοι υπολογισμοί που επιτρέπονται είναι αυτοί που αφορούν αυτή τη διάταξη.

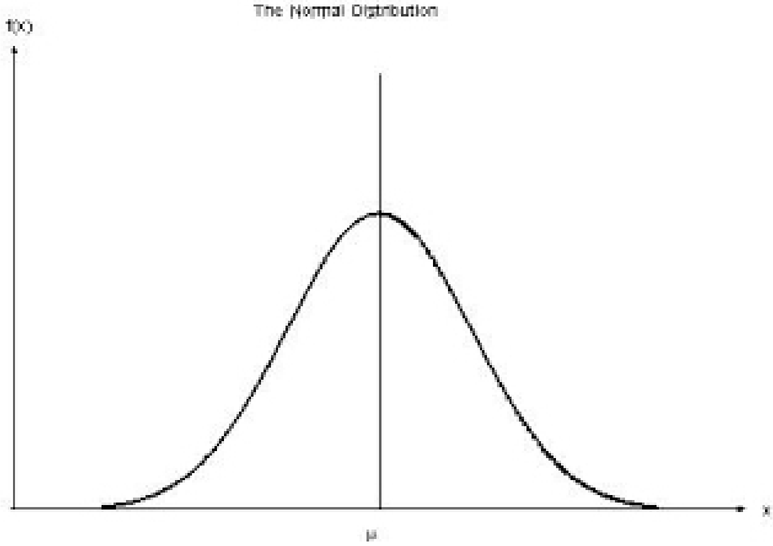
Κεφάλαιο ΙΙ: Συναρτήσεις κατανομής πιθανοτήτων

1. <u>Κανονική κατανομή</u>:

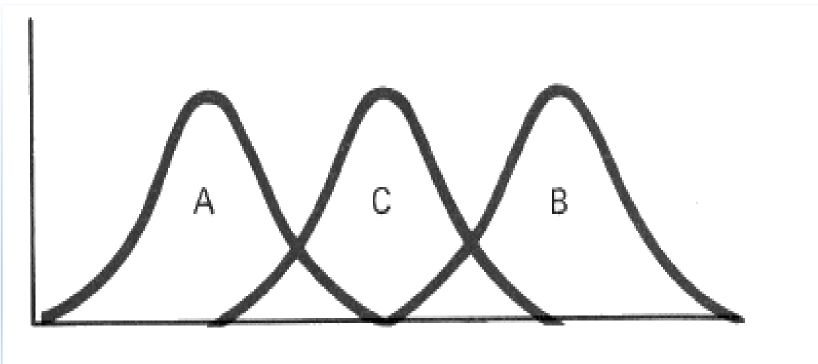
Η σημαντικότερη απ' όλες τις κατανομές πιθανοτήτων είναι η κανονική κατανομή (normal distribution), που αποτελεί την κινητήρια δύναμη της επαγωγικής στατιστικής.

Κανονική συνάρτηση πυκνότητας πιθανοτήτων Η συνάρτηση πυκνότητας πιθανοτήτων μιας κανονικής τυχαίας μεταβλητής (normal random distribution) είναι: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$, για $-\infty < x < +\infty$ όπου e=2,71828... (βάση των φυσικών λογαρίθμων) π=3,14159... (σταθερά της περιφέρειας κύκλου)

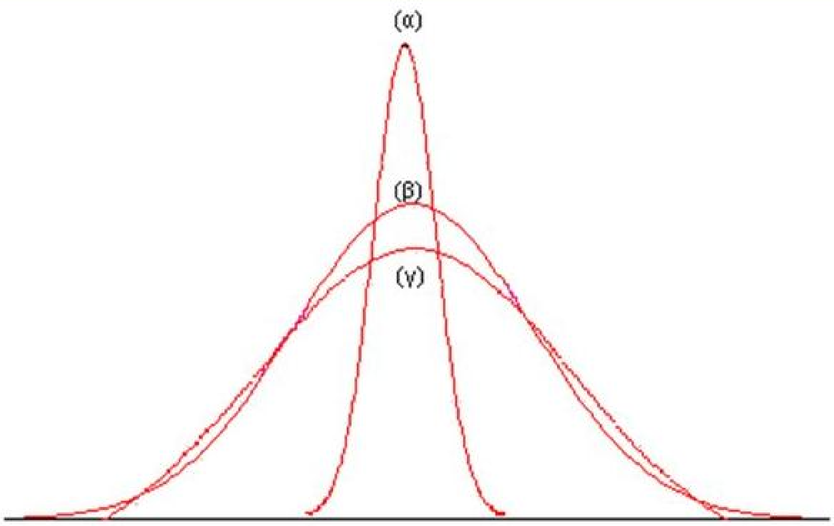
Γραφική παράσταση κανονικής κατανομής:



Παρατηρούμε ότι η μεταβλητή x μπορεί να πάρει όλες τις τιμές από $-\infty$ ως $+\infty$ και ότι η καμπύλη είναι συμμετρική ως προς τον μέσο της. Η καμπύλη έχει το σχήμα καμπάνας. Η συνάρτηση της κανονικής κατανομής εξαρτάται από δύο παραμέτρους, τον αριθμητικό μέσο μ και την τυπική απόκλιση σ της κατανομής. Όπως φαίνεται στην εικόνα παρακάτω, αν αυξηθεί ή μειωθεί ο μέσος μ η καμπύλη μετατοπίζεται δεξιά ή αριστερά αντίστοιχα.



Στην επόμενη εικόνα βλέπουμε ότι αν αυξηθεί η τυπική απόκλιση σ η καμπύλη γίνεται χαμηλότερη και πλατύτερη, ενώ αν μειωθεί η τυπική απόκλιση σ η καμπύλη γίνεται υψηλότερη και πιο στενή.



1.1. Υπολογισμός πιθανοτήτων σε κανονική κατανομή:

Οι στατιστικοί έχουν υπολογίσει την πιθανότητα ορισμένων βαθμολογιών που εμφανίζονται σε κανονική κατανομή με μέση τιμή 0 και τυπική απόκλιση 1. Επομένως, αν έχουμε δεδομένα που έχουν σχήμα κανονικής κατανομής, τότε εάν η μέση τιμή και η τυπική απόκλιση είναι 0 και 1 αντίστοιχα μπορούμε να χρησιμοποιήσουμε τους πίνακες πιθανοτήτων για την κανονική κατανομή για να δούμε πόσο πιθανό είναι να υπάρξει συγκεκριμένη βαθμολογία στα δεδομένα. Το προφανές πρόβλημα είναι ότι όλα τα δεδομένα που συλλέγουμε δεν έχουν μέση τιμή 0 και τυπική απόκλιση 1. Για παράδειγμα, μπορεί να έχουμε ένα σύνολο δεδομένων που να έχει μέσο όρο 567 και τυπική απόκλιση 52,98. Ευτυχώς, κάθε σύνολο δεδομένων μπορεί να μετατραπεί σε σύνολο δεδομένων που έχει μέση τιμή 0 και τυπική απόκλιση 1.

Για να υπολογίσουμε την πιθανότητα η τιμή μιας κανονικής τυχαίας μεταβλητής να πέσει σε ένα συγκεκριμένο διάστημα, πρέπει να υπολογίσουμε το εμβαδόν της επιφάνειας κάτω από την καμπύλη στα όρια του διαστήματος αυτού.

Δυστυχώς η συνάρτηση είναι εξαιρετικά πολύπλοκη και το εμβαδόν δεν μπορεί να υπολογισθεί ούτε με απλά μαθηματικά, ούτε με την βοήθεια ολοκληρωμάτων. Η μοναδική λύση είναι ο προσεγγιστικός υπολογισμός με την βοήθεια πινάκων.

Στην κανονική κατανομή τα πράγματα είναι λίγο απλούστερα, καθώς μπορούμε να ορίσουμε την τυποποιημένη κανονική τυχαία μεταβλητή (standard normal random variable) ως εξής: Πρώτον, για να συγκεντρώσουμε τα δεδομένα γύρω από το μηδέν, λαμβάνουμε κάθε βαθμολογία και αφαιρούμε από αυτό το μέσο όρο. Στη συνέχεια, διαιρούμε την προκύπτουσα βαθμολογία με την τυπική απόκλιση για να εξασφαλίσουμε ότι τα δεδομένα έχουν τυπική απόκλιση. Άρα η εξίσωση είναι:

$$z=\frac{x-\mu}{\sigma}$$

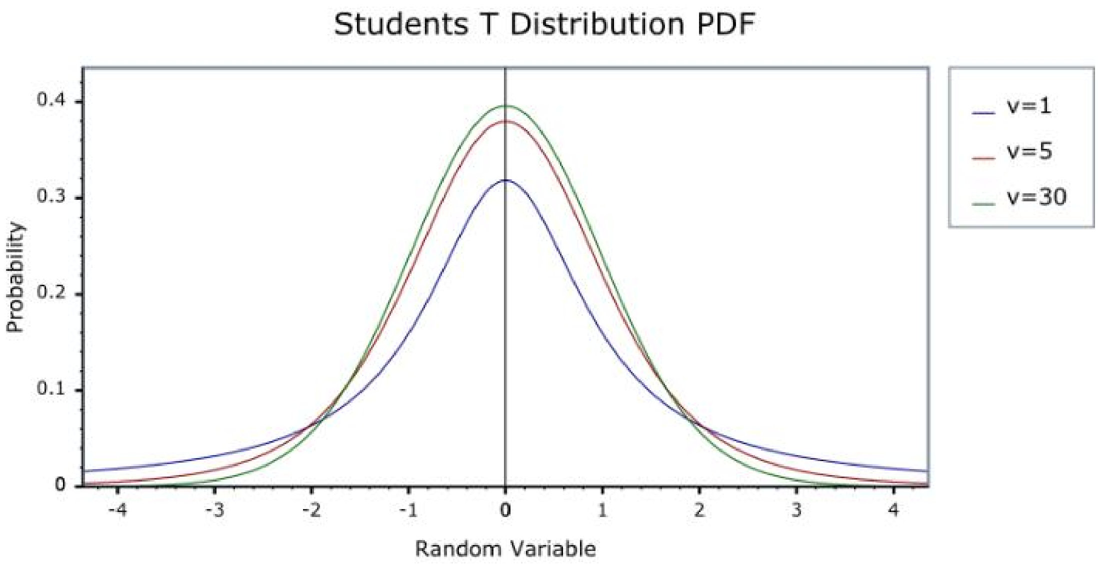
Έτσι κάθε υπολογισμός πιθανότητας για την τυχαία μεταβλητή X ανάγεται σε έναν αντίστοιχο υπολογισμό για την τυχαία μεταβλητή Z, η οποία έχει μέσο ίσο με το μηδέν και τυπική απόκλιση ίση με την μονάδα.

2. Κατανομή student t:

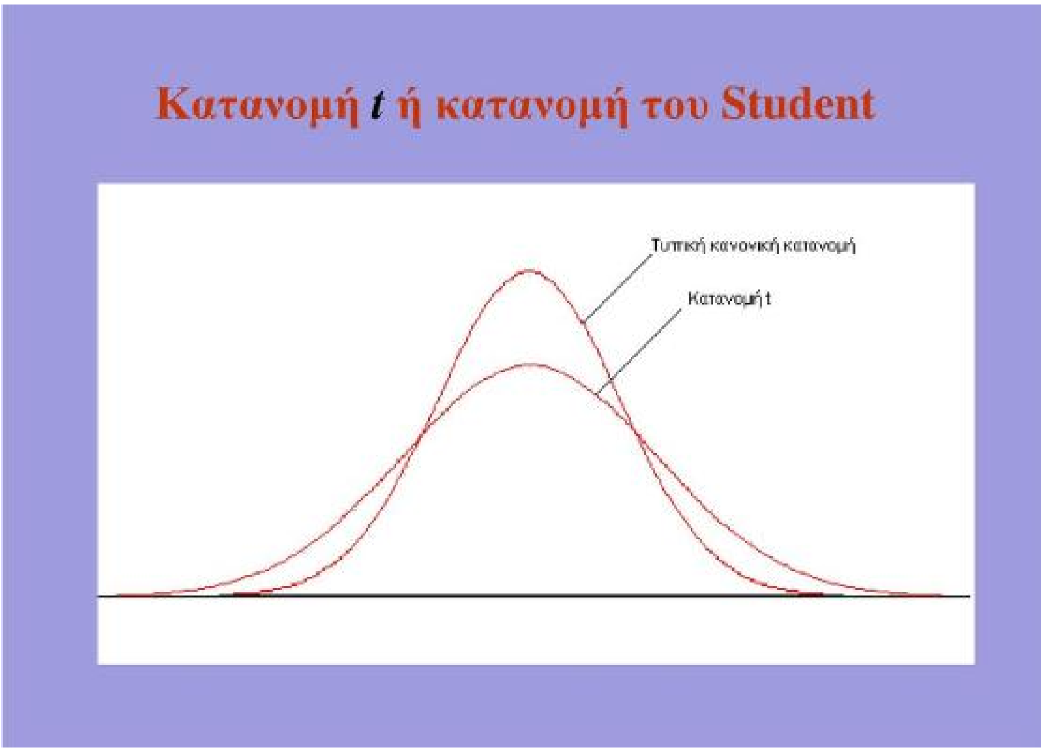
Συνάρτηση πυκνότητας πιθανοτήτων της κατανομής student t
Η συνάρτηση της κατανομής student t (student t distribution) είναι:

$$F(t)= \frac{r(\frac{v-1}{2})}{\sqrt{v\pi}*\Gamma(\frac{v}{2})}*[1 + \frac{t^2}{v}]^{\frac{v+1}{2}}$$

Όπου π=3,14159… είναι η γνωστή σταθερά του μήκους της περιφέρειας του κύκλου, v είναι η παράμετρος της κατανομής που ονομάζεται βαθμοί ελευθερίας (degrees of freedom), και η συνάρτηση Γ ονομάζεται «συνάρτηση Γ» (Gamma function) και είναι επέκταση του παραγοντικού σε μη ακέραιους αριθμούς.
Η κατανομή οφείλει το όνομά της στο ψευδώνυμο «Φοιτητής» (Student) που χρησιμοποιεί ο William S. Gosset για να την δημοσιεύσει το 1908, επειδή ο εργοδότης του δεν του επέτρεπε να δημοσιεύσει με το όνομά του, και στο γεγονός ότι στη δημοσίευση εκείνη η τυχαία μεταβλητή συμβολιζόταν με το γράμμα t.
Ο μέσος και η διασπορά της κατανομής t δίνονται απ' τους τύπους:
E(t)=0
$V(t)= \frac{v}{v-2}$ για v-2
Η καμπύλη της κατανομής student t φαίνεται στην εικόνα που ακολουθεί. Όπως μπορείτε να δείτε έχει αρκετή ομοιότητα με την καμπύλη της τυποποιημένης κατανομής, καθώς είναι και οι δύο συμμετρικές ως προς το μηδέν, ενώ το σχήμα τους έχει πολύ μικρές διαφορές. Για να τις διακρίνουμε, η καμπύλη student t λέγεται ότι έχει σχήμα λόφου (mound shaped) ενώ η καμπύλη της κανονικής κατανομής λέγεται ότι έχει σχήμα καμπάνας (bell shaped).



Οπτικά η διαφορά είναι ότι η καμπύλη της κατανομής student t έχει ελαφρά μικρότερο ύψος και ελαφρά μεγαλύτερο πλάτος από την καμπύλη της τυποποιημένης κανονικής κατανομής. [Η διασπορά της τυποποιημένης κανονικής κατανομής είναι σ^2 = 1, ενώ η διασπορά της κατανομής student t είναι σ^2 = $\frac{v}{v-2}$ που είναι λίγο μεγαλύτερη απ' τη μονάδα.] Οι διαφορές αυτές φαίνονται στην επόμενη εικόνα.



Στην επόμενη εικόνα μπορείτε να δείτε συγκριτικά τις καμπύλες της κατανομής student t για διάφορους βαθμούς ελευθερίας (=v). Παρατηρήστε ότι όσο οι βαθμοί ελευθερίας αυξάνονται η διασπορά της κατανομής student t τείνει στη μονάδα και η καμπύλη τείνει να ταυτιστεί με την καμπύλη της τυποποιημένης κανονικής κατανομής.

Υπολογισμός πιθανοτήτων στην κατανομή student t

Για κάθε τιμή της παραμέτρου v (βαθμοί ελευθερίας) υπάρχει μια διαφορετική κατανομή student t, δηλαδή θα έπρεπε στο παράρτημα Β να υπάρχει ένας διαφορετικός πίνακας πιθανοτήτων. Αυτό όμως δεν είναι πραγματικά εφικτό και έτσι ο υπολογισμός των πιθανοτήτων για την κατανομή student t θα γίνεται με την βοήθεια υπολογιστή, μέσα από το λογισμικό Excel ή Minitab, όπως θα δούμε παρακάτω.

Υπολογισμός τιμών της μεταβλητής για δεδομένη πιθανότητα στην κατανομή student t

Η κατανομή student t έχει ευρύτατη εφαρμογή στην επαγωγική στατιστική, όπου συχνά το ζητούμενο είναι να βρεθεί η κατάλληλη τιμή της τυχαίας μεταβλητής που έχει μια δεδομένη πιθανότητα.

Ο συμβολισμός που χρησιμοποιείται για την κατανομή student t: αν δοθεί μια πιθανότητα Α, η τιμή $t_{A.v}$ είναι η τιμή της μεταβλητής για την οποία ισχύει:

P(t>$t_{A.v}$)= A

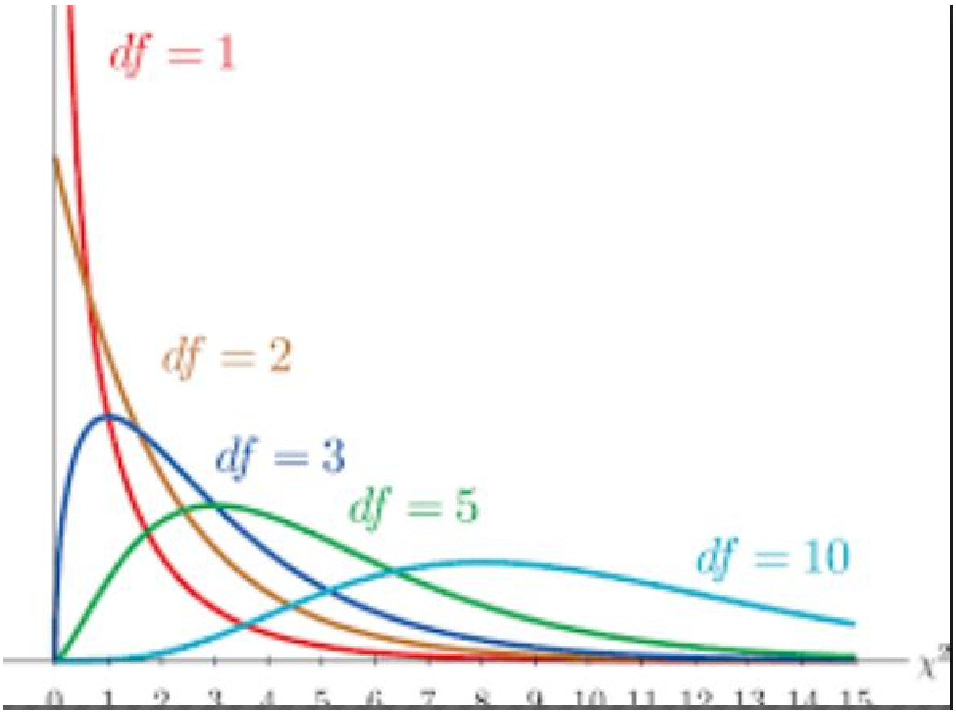
Στην κατανομή student t με παράμετρο v.

3. Κατανομή χ^2:

Μια άλλη πολύ χρήσιμη συνεχής κατανομή πιθανοτήτων είναι αυτή που ονομάζεται χ^2 και προφέρεται «χι τετράγωνο» (chi squared) από το Ελληνικό γράμμα «χ». Η κατανομή ονομάζεται έτσι επειδή η τυχαία μεταβλητή συμβολίζεται ως χ^2.
Η συνάρτηση πυκνότητας πιθανοτήτων της κατανομής χ^2 (chi squared distribution) είναι:

$$f(\chi^2)=\frac{1}{\Gamma(\frac{\nu}{2})} * \frac{1}{2^{\frac{\nu}{2}}} * (\chi^2)^{\frac{\nu}{2}-1} * e^{\frac{\chi^2}{2}}$$

για $\chi^2>0$, όπου e=2,71828… είναι η βάση των φυσικών λογαρίθμων, ν είναι η παράμετρος της κατανομής που ονομάζονται βαθμοί ελευθερίας, και η συνάρτηση Γ ονομάζεται «συνάρτηση Γάμμα» και είναι η επέκταση του παραγοντικού σε μη ακέραιους αριθμούς.
Στην επόμενη εικόνα μπορείτε να δείτε την καμπύλη χ^2. Η τυχαία μεταβλητή χ^2 παίρνει τιμές ίσες ή μεγαλύτερες από το μηδέν και η καμπύλη παρουσιάζει θετική ασσυμετρία.



Η παράμετρος ν (βαθμοί ελευθερίας), όπως και στην περίπτωση της κατανομής student t, επηρεάζει το σχήμα της καμπύλης.
Ο μέσος και η διασπορά της κατανομής χ^2 δίνονται από τους τύπους:
$E(\chi^2)=\nu$
$V(\chi^2)= 2\nu$
Υπολογισμός τιμών της μεταβλητής για δεδομένη πιθανότητα στην κατανομή χ^2
Αν δοθεί μια πιθανότητα Α, η τιμή $\chi^2_{A.\nu}$ είναι η τιμή της μεταβλητής για την οποία ισχύει:
$P(\chi^2 > \chi^2_{A.\nu})=A$
Στην κατανομή χ^2 με παράμετρο ν.

Επειδή η μεταβλητή χ^2 δεν παίρνει αρνητικές τιμές δεν υπάρχει συμμετρία ως προς το μηδέν και το διάστημα στο αριστερό άκρο της καμπύλης που έχει εμβαδόν(=πιθανότητα) Α συμβολίζεται ως $\chi^2_{1-A.\nu}$ και είναι η τιμή της μεταβλητής για την οποία ισχύει:
P(χ^2>$\chi^2_{1-A.\nu}$)=A
Στην κατανομή με παράμετρο ν.

Κεφάλαιο ΙΙΙ: Έλεγχος υποθέσεων, γενική θεωρία

Όταν θέλουμε να μάθουμε το κατά πόσο η τιμή μιας παραμέτρου ενός πληθυσμού είναι αποδεκτή με βάση κάποια ζητούμενα, χρησιμοποιούμε τον **έλεγχο υποθέσεων**.

1. Βασικές έννοιες του ελέγχου υποθέσεων:

Πολλές φορές θέλουμε να ελέγξουμε εάν μια υπόθεση για μια παράμετρο ενός πληθυσμού είναι αληθής. Μια διαδεδομένη μέθοδος είναι ο έλεγχος υποθέσεων, ο οποίος λειτουργεί με βάση δεδομένα και στατιστικά στοιχεία που διαθέτουμε. Στον έλεγχο υποθέσεων υπάρχουν 2 υποθέσεις: Η πρώτη ονομάζεται μηδενική υπόθεση και συμβολίζεται με H_0 και η δεύτερη εναλλακτική υπόθεση ή υπόθεση έρευνας συμβολίζεται με H_1. Ο έλεγχος ξεκινά θεωρώντας ότι η μηδενική υπόθεση είναι αληθής. Στόχος του ελέγχου είναι να καθοριστεί αν υπάρχουν αρκετές αποδείξεις που να στηρίζουν την αλήθεια της εναλλακτικής υπόθεσης.
Οι δυνατές αποφάσεις είναι δύο: Απόρριψη της μηδενικής υπόθεσης ή μη απόρριψη της μηδενικής υπόθεσης.
Ένα παράδειγμα εφαρμογής του ελέγχου υποθέσεων στην καθημερινότητα είναι το εξής: Όταν κάποιος αντιμετωπίζει μια κατηγορία στο δικαστήριο, ο δημόσιος κατήγορος και ο συνήγορος υπεράσπισης παρουσιάζουν αποδείξεις και αναπτύσσουν επιχειρήματα, και τελικά οι ένορκοι αποφασίζουν αν ο κατηγορούμενος είναι ένοχος ή αθώος.
Οι ένορκοι πραγματοποιούν έναν έλεγχο υποθέσεων. Η **μηδενική υπόθεση** είναι:
H_0: Ο κατηγορούμενος είναι αθώος, και η **εναλλακτική υπόθεση** ή **υπόθεση έρευνας** είναι **H_1**: Ο κατηγορούμενος είναι ένοχος
Οι ένορκοι δε γνωρίζουν εξ αρχής ποια από τις δύο υποθέσεις είναι σωστή, έτσι θα πάρουν την απόφαση τους με βάση τα στοιχεία που θα τεθούν στο δικαστήριο.
Κατά τον έλεγχο υποθέσεων, υπάρχουν και δύο δυνατοί τύποι σφάλματος: Το **σφάλμα τύπου Ι** και το **σφάλμα τύπου ΙΙ**:

Σφάλμα τύπου Ι

Σφάλμα τύπου Ι είναι η εσφαλμένη απόρριψη μιας πραγματικής μηδενικής υπόθεσης, δηλαδή η εύρεση ενός αποτελέσματος που δεν υπάρχει. Η πιθανότητα του σφάλματος τύπου Ι ορίζεται ως:
P(σφάλμα τύπου Ι)=**α**
Η πιθανότητα του σφάλματος τύπου Ι ονομάζεται **στάθμη σημαντικότητας**.

Σφάλμα τύπου ΙΙ

Το σφάλμα τύπου ΙΙ συμβαίνει όταν εσφαλμένα δεν απορρίψουμε μια μηδενική υπόθεση, δηλαδή όταν δεν διακρίνουμε ένα αποτέλεσμα που υπάρχει. . Η πιθανότητα του σφάλματος τύπου ΙΙ ορίζεται ως: P(σφάλμα τύπου ΙΙ)=**β**

Στο προηγούμενο παράδειγμα, το σφάλμα τύπου Ι σημαίνει ότι ο κατηγορούμενος ενώ ήταν αθώος καταδικάστηκε, ενώ το σφάλμα τύπου ΙΙ σημαίνει ότι ο κατηγορούμενος ενώ ήταν ένοχος αθωώθηκε.

Όπως βλέπουμε, η πιθανότητα σφάλματος ΙΙ αυξήθηκε. Όσο μειώνεται η στάθμη σημαντικότητας α, μειώνεται και η πιθανότητα σφάλματος τύπου Ι. Οπότε, δεν υπάρχει κάποια στάθμη σημαντικότητας που είναι η καταλληλότερη για όλες τις περιπτώσεις. Το ποια θα επιλέγουμε, εξαρτάται από τα δεδομένα του κάθε προβλήματος.

<u>Προσοχή:</u> Αυτό που δε πρέπει να μπερδεύουμε είναι το δείγμα πληθυσμού με ολόκληρο τον πληθυσμό. Όταν καταλήγουμε στο συμπέρασμα ότι η μηδενική υπόθεση απορρίπτεται, δε μπορούμε να πούμε με βεβαιότητα ότι η εναλλακτική υπόθεση είναι αληθής, διότι ο έλεγχος υπόθεσης δεν πραγματοποιήθηκε σε ολόκληρο τον πληθυσμό. Στην επαγωγική στατιστική μπορούμε να στηρίξουμε μια υπόθεση, αλλά ποτέ να την αποδείξουμε.

Έτσι, αν ο έλεγχος δε βρίσκεται στην περιοχή απόρριψης λέμε ότι δεν υπάρχουν αρκετά στατιστικά στοιχεία που να στηρίζουν ότι η μηδενική υπόθεση είναι ψευδής.

<u>Εν κατακλείδι</u>: Όταν απορρίπτουμε τη μηδενική υπόθεση συμπεραίνουμε ότι υπάρχουν αρκετά στατιστικά στοιχεία τα οποία στηρίζουν την ορθότητα της εναλλακτικής υπόθεσης, ενώ όταν δεν απορρίπτουμε τη μηδενική υπόθεση σημαίνει πως δεν υπάρχουν αρκετά στατιστικά στοιχεία για να γίνει αποδεκτή η εναλλακτική υπόθεση.

Είναι εμφανές ότι και στις δύο περιπτώσεις το αποτέλεσμα επικεντρώνεται στην αλήθεια της εναλλακτικής υπόθεσης. Αυτό οφείλεται στο γεγονός ότι η εναλλακτική υπόθεση είναι εξαρχής ο λόγος που πραγματοποιούμε τον έλεγχο, και γι' αυτό το λόγο η εναλλακτική υπόθεση ονομάζεται και **υπόθεση ελέγχου**.

Κεφάλαιο IV: Έλεγχος υποθέσεων για τον μέσο ενός δείγματος

1. Έλεγχος υποθέσεων όταν η τυπική απόκλιση είναι γνωστή:

Έλεγχος ενός άκρου:

Ho: $\mu = \mu_0$ H_1: $\mu > \mu_0$ (μονόπλευρος έλεγχος)

Ho: $\mu = \mu_0$ H_1: $\mu < \mu_0$ (μονόπλευρος έλεγχος)

Τι συμβαίνει όμως όταν υπάρχει περιοχή απόρριψης και στα δύο άκρα της κατανομής;

Έλεγχος δύο άκρων:

Τα δύο άκρα τα ελέγχουμε όταν έχουμε το σύμβολο $\neq$ για την εναλλακτική υπόθεση, δηλαδή όταν η τιμή στην εναλλακτική υπόθεση πρέπει να είναι είτε μεγαλύτερη είτε μικρότερη από την τιμή με την οποία συγκρίνεται:

H_0: $\mu = \mu_0$ H_1: $\mu \neq \mu_0$ (αμφίπλευρος έλεγχος)

Πριν πάμε να δούμε κάποια παραδείγματα με την εφαρμογή αμφίπλευρου και μονόπλευρου ελέγχου, θα δούμε τι χρειάζεται για να τα επιλύσουμε και να τα ερμηνεύσουμε σωστά στο στατιστικό πρόγραμμα **SPSS**:

1) Το δείγμα πρέπει να έχει επιλεγεί τυχαία απ' τον πληθυσμό.
2) Τα δεδομένα πρέπει να ακολουθούν την κανονική κατανομή. (Για τον έλεγχο κανονικότητας, πηγαίνουμε Analyze→ Descriptivestatistics→Explore)
3) Γνώση του επίπεδου σημαντικότητας α.

Συμπεράσματα απ' τους ελέγχους:

Αμφίπλευρος έλεγχος:

Αν sig>α τότε αποδεχόμαστε την υπόθεση H_0

Αν sig<α τότε απορρίπτουμε την υπόθεση H_0

Μονόπλευρος έλεγχος:

Αν η μέση τιμή του δείγματος ικανοποιεί την ανισότητα της H_1 τότε ισχύει:

Αν$\frac{sig}{2}>\alpha$ τότε αποδεχόμαστε την υπόθεση H_0

Αν$\frac{sig}{2}<\alpha$ τότε απορρίπτουμε την υπόθεση H_0

Αν η μέση τιμή του δείγματος **δεν** ικανοποιεί την ανισότητα της H_1 τότε ισχύει:

Αν 1- $\frac{sig}{2}>\alpha$ τότε αποδεχόμαστε την υπόθεση H_0

Αν 1- $\frac{sig}{2}<\alpha$ τότε απορρίπτουμε την υπόθεση H_0

Ο μονόπλευρος με τον αμφίπλευρο έλεγχο γίνονται με τον ίδιο τρόπο, το μόνο που αλλάζει είναι η ερμηνεία των αποτελεσμάτων. Ας το δούμε στην πράξη μέσα απ' το επόμενο παράδειγμα, το οποίο θα λύσουμε στο SPSS:

<u>Παράδειγμα:</u>

Μια αλυσίδα εστιατορίων ταχείας εξυπηρέτησης εξετάζει την πιθανότητα δημιουργίας ενός νέου εστιατορίου σε μια συγκεκριμένη περιοχή. Με βάση την οικονομική ανάλυση ένα σημείο θεωρείται αποδεκτό για την κατασκευή εστιατορίου μόνο αν ο μέσος αριθμός των διερχομένων πεζών είναι μεγαλύτερος από 100 την ώρα. Για το σκοπό της μελέτης επιλέχθηκε ένα τυχαίο δείγμα 40 διαφορετικών ωρών και καταγράφηκε ο αριθμός των διερχομένων πεζών κάθε ώρα. Αν ο αριθμός αυτός έχει κανονική κατανομή με τυπική απόκλιση 16 την ώρα, μπορούμε να συμπεράνουμε με στάθμη σημαντικότητας 1% ότι η τοποθεσία είναι αποδεκτή για τη δημιουργία του νέου εστιατορίου;

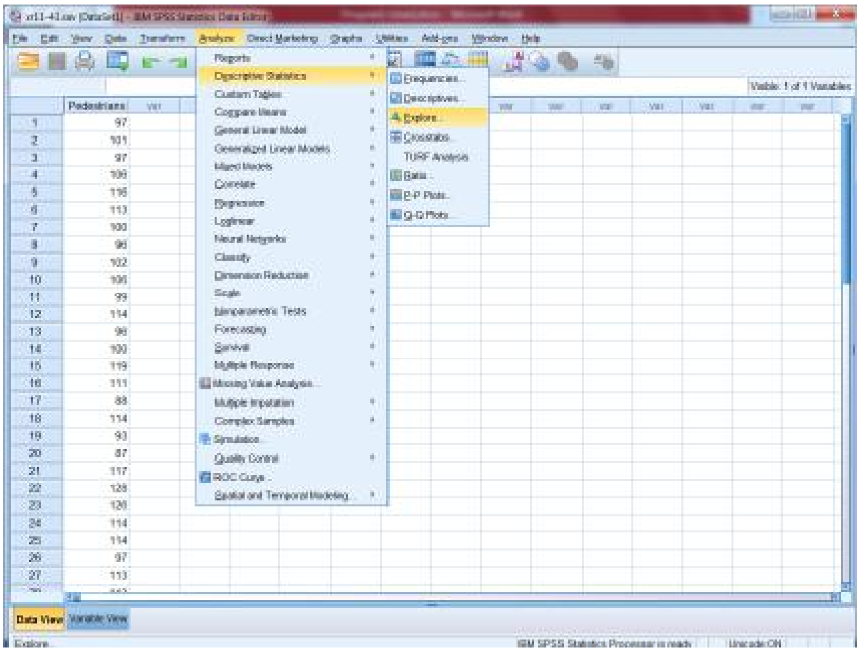
<u>Λύση</u>:

$\mu_0=100$

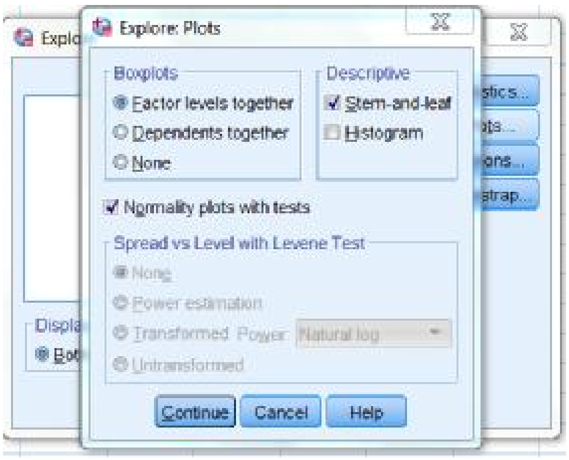
H_0: $\mu=\mu_0$ H1: $\mu\neq\mu_0$

1)Αρχικά φτιάχνουμε μια μεταβλητή «Pedestrians» για τους πεζούς στο Variable View και στο Data View εισάγουμε τα δεδομένα της.

2) Το δείγμα θα πρέπει να ακολουθεί κανονική κατανομή, άρα θα πραγματοποιήσουμε έναν έλεγχο κανονικότητας:
Analyze → Descriptive statistics → explore



Στη συνέχεια περνάω τη μεταβλητή με τις τιμές των πεζών στο κουτί «dependent list» και στην επιλογή «statistics» επιλέγω outliers και percentiles και στην επιλογή «plots» επιλέγω «normality plots with tests» για να μου εμφανιστούν τα απαραίτητα σχεδιαγράμματα :



Το αποτέλεσμα του OUTPUT είναι το εξής:

Tests of Normality

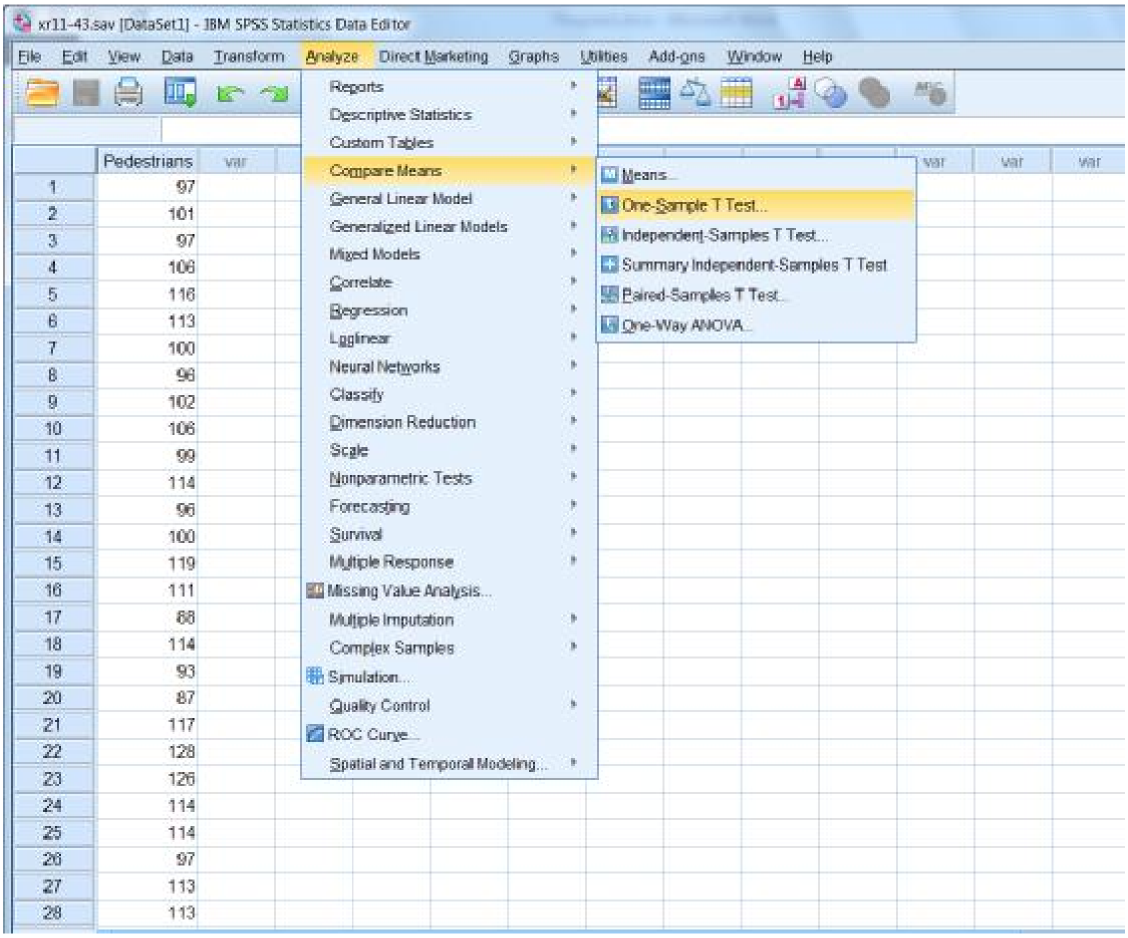
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Pedestrians	,100	40	,200*	,979	40	,664

*. This is a lower bound of the true significance.

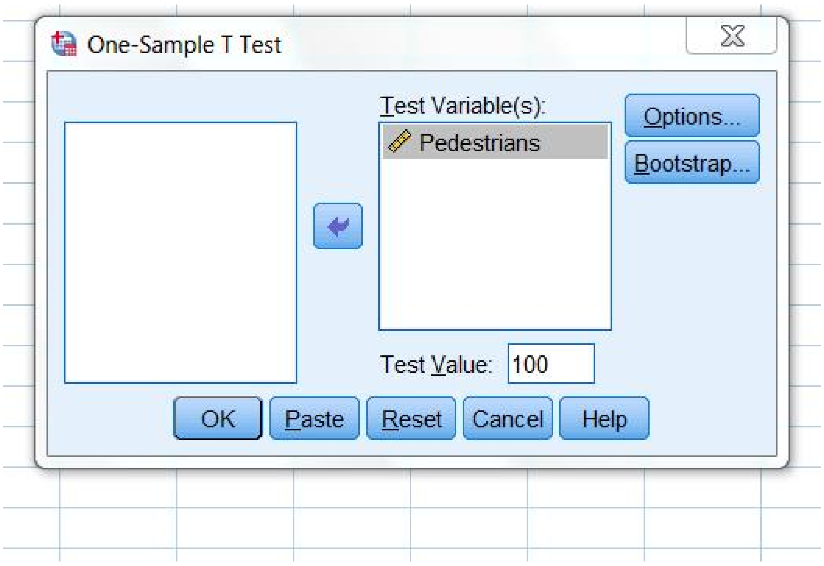
a. Lilliefors Significance Correction

Για να ακολουθεί ο πληθυσμός κανονική κατανομή , πρέπει το sig κάθε κριτηρίου να είναι μεγαλύτερο του α. Εδώ 0,200>0,05 και 0,664>0,05 , άρα αποδεχόμαστε ότι το δείγμα ακολουθεί κανονική κατανομή.

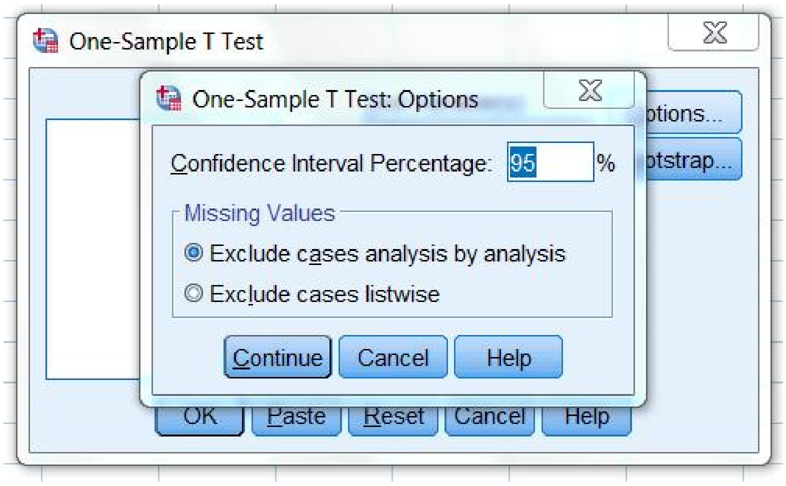
3)Το κριτήριο που θα χρησιμοποιήσουμε για να συγκρίνουμε τον αριθμητικό μέσο μ του δείγματος με την τιμή των 100 πεζών είναι το One sample t-test. Πάμε Analyze →Compare Means→One sample t-test:



→ περνάμε στο test variable τη μεταβλητή με τους πεζούς και στο test_value βάζουμε την τιμή 100 με την οποία θα συγκρίνουμε τη μέση τιμή:



Στη συνέχεια πηγαίνουμε στο κουμπί options και Confidence Interval Percentage όπου γράφουμε το επίπεδο εμπιστοσύνης με το οποίο θα γίνει ο έλεγχος. Στην περίπτωση μας είναι το 99% καθώς έχουμε στάθμη σημαντικότητας 1%.



3) CONTINUE→ OK και πάμε στο OUTPUT που εμφανίζεται:

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Πεζοί	40	105,70	10,837	1,714

Εδώ βλέπουμε ότι οι παρατηρήσεις που έχουμε είναι N=40, και ο αριθμητικός μέσος των παρατηρήσεων του δείγματος είναι μ=**105,70.**

μ=105,7>100 άρα η μέση τιμή του δείγματος ικανοποιεί την ανισότητα της Η1. Πάμε στον επόμενο πίνακα να δούμε την τιμή του sig:

One-Sample Test

	Test Value = 100					
					99% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Πεζοί	3,326	39	,002	5,700	1,06	10,34

Το sig είναι 0,002. Πρέπει να δούμε αν $\frac{sig}{2}>\alpha$ ή αν $\frac{sig}{2}<$ α, 0,002 <0,01 , άρα απορρίπτουμε την υπόθεση H_0.

Ας δούμε και ένα παράδειγμα με αμφίπλευρο έλεγχο:

<u>Παράδειγμα:</u>

Πολλά χιονοδρομικά κέντρα βασίζουν τις οικονομικές τους αναλύσεις στην υπόθεση ότι ο μέσος πελάτης ταξιδεύει για σκι 4 φορές το χρόνο. Για να ελεγχθεί η υπόθεση αυτή επιλέχθηκε ένα τυχαίο δείγμα 63 σκιέρ οι οποίοι ρωτήθηκαν πόσες φορές ταξίδεψαν για σκι τον περασμένο χρόνο. Αν δεχθούμε ότι ο πληθυσμός έχει κανονική κατανομή με τυπική απόκλιση 2, μπορούμε να συμπεράνουμε ότι με στάθμη σημαντικότητας 10% η υπόθεση των οικονομικών αναλύσεων είναι σωστή;

<u>Λύση:</u>

Τα βήματα που ακολουθούμε είναι τα ίδια με την προηγούμενη άσκηση, οπότε πάμε να δούμε τα αποτελέσματα που έχουμε στο OUTPUT, και να τα ερμηνεύσουμε:

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Μέρες_σκι	63	4,84	2,223	,280

Εδώ η μέση τιμή των παρατηρήσεων του δείγματος είναι μ=4,84.

One-Sample Test

	Test Value = 4					
					90% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Μέρες_σκι	3,004	62	,004	,841	,37	1,31

Στον αμφίπλευρο έλεγχο, εξετάζουμε το sig. Εδώ είναι 0,004< 0,1 , οπότε απορρίπτουμε τη μηδενική υπόθεση H_0.

Κεφάλαιο V: Μέθοδοι

Πέρα από τον έλεγχο υποθέσεων, χρειαζόμαστε έναν τρόπο για να ελέγχουμε το πόσο αξιόπιστο είναι το αποτέλεσμα που βρήκαμε. Σε αυτό σημείο, χρησιμοποιούμε τις **στατιστικές μεθόδους**. Οι μέθοδοι είναι τρεις, η **περιοχή απόρριψης, ο τυποποιημένος έλεγχος** και η **τιμή p**(p-value). Ας τις δούμε όμως αναλυτικότερα μέσα από το σημαντικό ερώτημα που δημιουργείται στο επόμενο παράδειγμα:

Ο ιδιοκτήτης ενός γηπέδου γκολφ πιστεύει ότι οι καθυστερήσεις στην ολοκλήρωση των αγώνων περιορίζουν τα έσοδα του και προσπαθώντας να εξερευνήσει την αιτία των καθυστερήσεων επέλεξε ένα τυχαίο δείγμα 10 αγώνων μεταξύ ζευγαριών και κατέγραψε το χρόνο (σε λεπτά) που χρειάστηκαν για να ολοκληρώσουν το γρασίδι στην 18η τρύπα. Οι χρόνοι φαίνονται στον παρακάτω πίνακα:

8	11	5	6	7	8 8	6 6	4 4	8 8	3

Αν είναι γνωστό ότι οι χρόνοι έχουν κανονική κατανομή με τυπική απόκλιση σ=2 λεπτά, μπορούμε να συμπεράνουμε με στάθμη σημαντικότητας 5% ότι ο μέσος χρόνος που απαιτείται για το γρασίδι της 18ης τρύπας είναι πάνω από 6 λεπτά;

Λύση:

Ο ιδιοκτήτης θέλει να ελέγξει εάν ο μέσος χρόνος για την 18η τρύπα είναι μεγαλύτερος από 6 λεπτά. Άρα, η εναλλακτική υπόθεση θα είναι $H_1 : \mu > 6$.
Οπότε η μηδενική υπόθεση θα είναι $H_0 : \mu = 6$

$$\mu = \frac{8+11+5+6+7+8+6+4+8+3}{10} = 6,6$$

Το ερώτημα που δημιουργείται είναι το εξής. Ένας μέσος ίσος με 6,6 , τιμή ούτε πολύ κοντά, ούτε πολύ μακριά από το 6,είναι ικανοποιητική απόδειξη της υπόθεσης ότι ο μέσος χρόνος που απαιτείται είναι πάνω από 6;

1. Περιοχή απόρριψης:

Περιοχή απόρριψης είναι το διάστημα τιμών για το οποίο θεωρούμε ότι αν ο στατιστικός δείκτης του δείγματος βρεθεί εκεί, η μηδενική υπόθεση πρέπει να απορριφθεί.
Η τιμή που αποτελεί το «άκρο», δηλαδή το όριο μεταξύ αποδεκτών τιμών και μη αποδεκτών, ονομάζεται **οριακή τιμή**. Για $\overline{x_L}$ οριακή τιμή, εάν ο μέσος βρεθεί πιο πάνω από $\overline{x_L}$, απορρίπτουμε τη μηδενική υπόθεση. Οπότε περιοχή απόρριψης είναι το σύνολο τιμών: $\bar{x} > \overline{x_L}$
Την οριακή τιμή $\overline{x_L}$, την βρίσκω μέσω του **σφάλματος τύπου I**.

Η πιθανότητα του σφάλματος τύπου I συμβολίζεται με **α** , και είναι: α = P ($\bar{x} > \overline{x_L}$ δεδομένου ότι μ=100). Ο πληθυσμός ακολουθεί την κανονική κατανομή, με μέσο μ και τυπική απόκλιση $\sigma/\sqrt{n}$. Οπότε ανάγουμε την περιοχή απόρριψης στην τυποποιημένη κανονική κατανομή και έχουμε το εξής:

$$P\left(\frac{(\bar{x}-\mu)}{(\sigma/\sqrt{n})} > \frac{(\overline{x_L}-\mu)}{(\sigma/\sqrt{n})}\right) = P\left(Z > \frac{(\overline{x_L}-\mu)}{(\sigma/\sqrt{n})}\right) = \alpha$$

$P(Z>z_{\alpha}) = \alpha$

Άρα από τις δύο παραπάνω ισότητες καταλήγουμε στην εξής ισότητα:

$$\frac{(\overline{x_L}-\mu)}{(\sigma/\sqrt{n})} = z_\alpha$$

Πάμε να βρούμε το z_α .Έχουμε στάθμη σημαντικότητας 5%, άρα:

$z_\alpha = z_{0,05} = 1,645$

Επιστρέφουμε στην ισότητα, και αντικαθιστούμε τα γνωστά δεδομένα. Ο μέσος μ ισούται με 6, η τυπική απόκλιση είναι 2 και το δείγμα είναι 10 τιμές. Άρα:

$\frac{(\overline{x_L} - \mu)}{(\sigma/\sqrt{n})} = z_\alpha \rightarrow \frac{(\overline{x_L} - 6)}{(2/\sqrt{10})} = 1,645 \rightarrow \overline{x_L} = 7,04$

Άρα περιοχή απόρριψης είναι οι τιμές μεγαλύτερες του 7,04.

Η μέση τιμή του δείγματος ίση με 6,6 λεπτά που βρήκαμε προηγουμένως δεν βρίσκεται στην περιοχή απόρριψης, κάτι το οποίο σημαίνει ότι δεν υπάρχουν αρκετά στατιστικά στοιχεία για να συμπεράνουμε ότι ο μέσος χρόνος που απαιτείται για το γρασίδι της 18^{ης} τρύπας είναι πάνω από 6 λεπτά.

2. Τυποποιημένος έλεγχος:

Για την επίλυση του προηγούμενου προβλήματος, χρησιμοποιήσαμε ως έλεγχο το μέσο του δείγματος, και η περιοχή απόρριψης ορίστηκε ως μια οριακή τιμή για το μέσο του δείγματος. Ένας ευκολότερος τρόπος να πραγματοποιήσουμε τον έλεγχο υποθέσεων, είναι ο **τυποποιημένος έλεγχος**. Ο τύπος του είναι ο εξής:

$$\mathbf{z} = \frac{(\bar{\mathbf{x}} - \boldsymbol{\mu})}{(\boldsymbol{\sigma}/\sqrt{\mathbf{n}})}$$

Εδώ, η περιοχή απόρριψης θα ορίζεται ως εξής: $z > z_a$

Ας δούμε τώρα πως θα λυνόταν το τελευταίο παράδειγμα με τον τυποποιημένο έλεγχο:

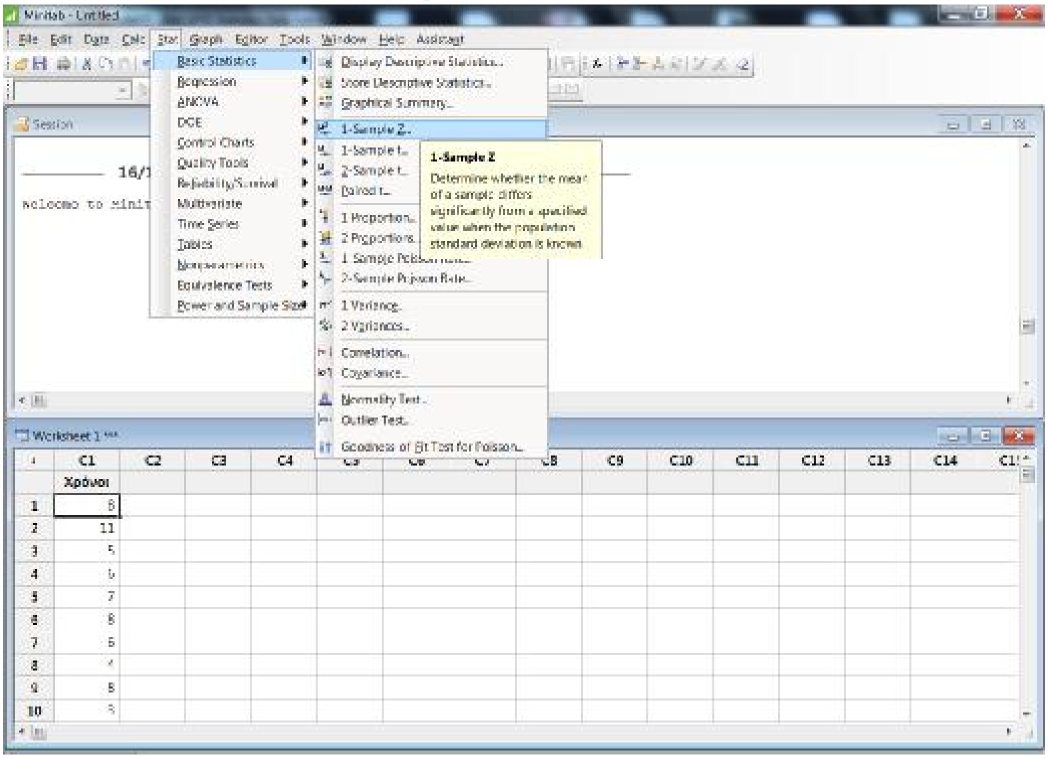
$z = \frac{(\mu - \mu o)}{(\sigma/\sqrt{n})} = \frac{(6,6-6)}{(2/\sqrt{10})} = 0,95$

$z_{a=}\ z_{0,05} = 1.645$

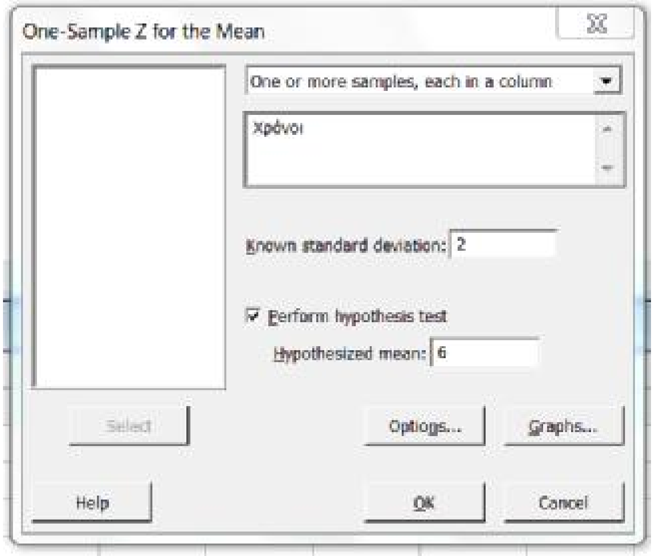
Το Ζ είναι μικρότερο από το Ζα οπότε εφόσον δεν βρίσκεται στην περιοχή απόρριψης αποδεχόμαστε τη μηδενική υπόθεση και καταλήγουμε στο συμπέρασμα ότι δεν υπάρχουν αρκετά στατιστικά στοιχεία για να συμπεράνουμε ότι ο μέσος χρόνος που απαιτείται για την 18^{η} τρύπα είναι μεγαλύτερος από 6 λεπτά.

Στο στατιστικό πρόγραμμα Minitab θα λυνόταν ως εξής:

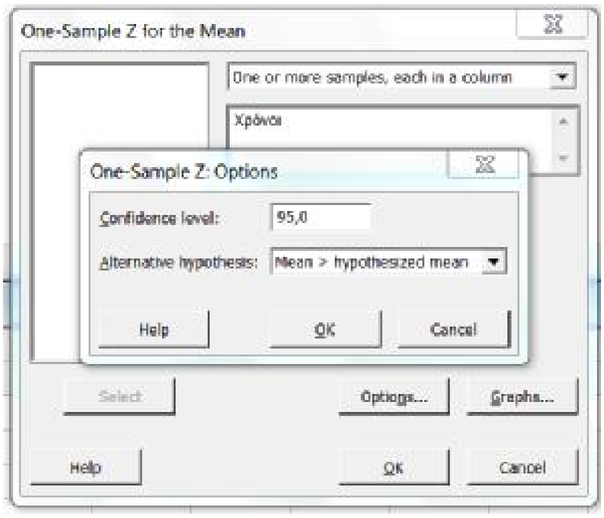
1)Αρχικά θα δημιουργήσουμε μια μεταβλητή με όνομα «Χρόνοι» ,και θα περάσουμε τις τιμές του πίνακα.
2) Πάμε Stat → Basic statistics → 1-Sample z.. :



→Επιλέγω one or more samples, each in a column →Περνάω τη μεταβλητή «Χρόνοι» στο κουτί→ Στο κουτί "Know standard deviation" βάζω την τιμή 2(τυπική απόκλιση) → επιλέγω"perform hypothesis test" → hypothesized mean βάζω την τιμή του μ_0=6:



→Options → confidence level 95, καθώς έχω στάθμη σημαντικότητας: 5% →alternative hypothesis επιλέγω “mean >hypothesized mean“ καθώς θα ελέγξω εάν ο μέσος μ είναι μεγαλύτερος του $μ_0$:



Το OUTPUTμας εμφανίζει τα εξής:

One-Sample Z: χρόνοι

```
Test of μ = 6 vs > 6
The assumed standard deviation = 2


Variable   N   Mean  StDev  SE Mean  95% Lower Bound     Z      P
χρόνοι    10  6,600  2,319    0,632            5,560  0,95  0,171
```

Εδώ βλέπουμε ότι ο μέσος του δείγματος είναι 6,6 και η τιμή του τυποποιημένου ελέγχου είναι z=0,95. $z<z_{\alpha}$=1,645 , άρα συμπεραίνουμε ότι δεν υπάρχουν αρκετά στατιστικά στοιχεία που να απορρίπτουν τη μηδενική υπόθεση.

3. Τιμή p:

Τιμή p είναι η δεσμευμένη πιθανότητα να πάρει ο έλεγχος μια τιμή σαν αυτή που έχει υπολογιστεί από το δείγμα, ή και πιο ακραία, με δεδομένη την αλήθεια της μηδενικής υπόθεσης.

Το ερώτημα που προκύπτει είναι ”πόσο μικρή πρέπει να είναι η τιμή p για να θεωρηθεί απόλυτα σωστή η εναλλακτική υπόθεση; “ Για να απαντηθεί αυτό το ερώτημα, η τιμή p έχει χωριστεί σε συγκεκριμένες **κατηγορίες**:

p<0,01 Εδώ ο έλεγχος είναι στατιστικά πολύ σημαντικός και υπάρχει συντριπτική απόδειξη υπέρ της εναλλακτικής υπόθεσης.
0,01<= p<= 0,05 Οέλεγχος είναι στατιστικά σημαντικός και υπάρχει ισχυρή απόδειξη υπέρ της εναλλακτικής υπόθεσης.
0,05<= p<= 0,10 Ο έλεγχος είναι στατιστικά μη σημαντικός εδώ, και η απόδειξη της εναλλακτική υπόθεσης είναι μικρή.
p>= 0,10 Εδώ δεν υπάρχει καμία απόδειξη υπέρ της εναλλακτικής υπόθεσης.

Ας δούμε πως θα λυνόταν το προηγούμενο πρόβλημα με τους χρόνους του γηπέδου γκολφ μέσω της **τιμής p:**

Λύση:

Για το παράδειγμα μας, τιμή p είναι η πιθανότητα σε έναν πληθυσμό από χρόνους με μέσο χρόνο 6 λεπτά και τυπική απόκλιση 2, ένα δείγμα 10 τιμών να έχει μέσο χρόνο 6,6 λεπτά ή περισσότερα.
Άρα:

$$P(\bar{x} > 6{,}6) = P\frac{(\bar{\mathbf{x}}-\mu)}{(\sigma/n)} > \frac{(\mu-\mu o)}{(\sigma/\sqrt{\mathbf{n}})} = P(\ Z > \frac{(6{,}6-6)}{(2/\sqrt{\mathbf{10}})}) = P(Z>0{,}95)=0{,}001$$

Για να κατανοήσουμε καλύτερα την εφαρμογή της τιμής P, θα επιστρέψουμε λίγο στο προηγούμενο πρόβλημα που επιλύσαμε, με την δημιουργία ή όχι ενός νέου εστιατορίου.
Σε πραγματικό χρόνο, ο υπεύθυνος των εστιατορίων δεν θα λειτουργούσε ακριβώς έτσι. Μετά το αποτέλεσμα του ελέγχου, θα εξέταζε πολλούς ακόμη παράγοντες πριν πάρει την τελική του απόφαση, εκ των οποίων ο ένας είναι η πιθανότητα στατιστικού σφάλματος (σφάλμα τύπου Ι) στην πληροφορία που πρόκειται να χρησιμοποιήσει. Για να είναι πιο αξιόπιστο το αποτέλεσμα του ελέγχου, χρειάζεται να συνοδεύεται από ένα μέτρο της στατιστικής βαρύτητας της πληροφορίας αυτής, ώστε να μπορεί να σταθμιστεί με άλλους παράγοντες που επηρεάζουν την τελική απόφαση, κυρίως οικονομικούς. Το μέτρο αυτό είναι η **τιμή p.**
Η τιμή p είναι η πιθανότητα , οπότε :

$$P(X>100) = P\frac{(\bar{\mathbf{x}}-\mu)}{(\sigma/n)}>$$

$$\frac{(\mu-\mu o)}{(\sigma/\sqrt{\mathbf{n}})} = P(Z>\frac{(105{,}7-100)}{(16/\sqrt{\mathbf{40}})}) = P(Z>2{,}25)=$$

Αυτό μας λέει ότι η πιθανότητα να βρούμε μέσο του δείγματος 100 ή μεγαλύτερο είναι μόλις 0,001,οπότε μιλάμε για ένα αρκετά απίθανο γεγονός, άρα θα απορρίψουμε τη μηδενική υπόθεση. Για να γίνουν πιο κατανοητά τα παραπάνω, παρατηρούμε ότι όσο ο μέσος του δείγματος απομακρύνεται από την τιμή της μηδενικής υπόθεσης, τόσο η τιμή p μειώνεται. Όσο μικρότερη είναι η τιμή p, τόσο μεγαλύτερη είναι η βαρύτητα του αποτελέσματος του ελέγχου, άρα και η απόρριψη της μηδενικής υπόθεσης.

Ας δούμε τώρα μέσα απ' το επόμενο παράδειγμα πως υπολογίζεται η τιμή p με τη βοήθεια του SPSS:

<u>Παράδειγμα:</u>

Τα τελευταία χρόνια λειτουργεί στη Νέα Υόρκη μια τράπεζα με θυρίδα εξυπηρέτησης αυτοκινήτων, η οποία έχει ένα μέσο ρυθμό εξυπηρέτησης 20 ανά ώρα με τυπική απόκλιση 3. Πρόσφατα μια άλλη τράπεζα σε μικρή σχετικά απόσταση δημιούργησε τη δική της θυρίδα εξυπηρέτησης αυτοκινήτων. Προκειμένου να ελέγξει εάν αυτό θα μειώσει τον αριθμό των πελατών, η πρώτη τράπεζα επέλεξε ένα τυχαίο δείγμα 36 διαφορετικών ωρών και κατέγραψε τον αριθμό των αυτοκινήτων που εξυπηρετήθηκαν κάθε ώρα. Οι αριθμοί των αυτοκινήτων κάθε ώρας φαίνονται στον παρακάτω πίνακα:

21	15	17	19	22	19	24	22	20	21
23	21	23	23	18	19	20	19	20	17
18	19	17	14	19	16	18	11	21	20
22	23	23	22	17	15				

Μπορούμε να συμπεράνουμε με στάθμη σημαντικότητας 5% ότι ο μέσος αριθμός των πελατών έχει μειωθεί; Ποια είναι η τιμή p;

<u>Λύση:</u>

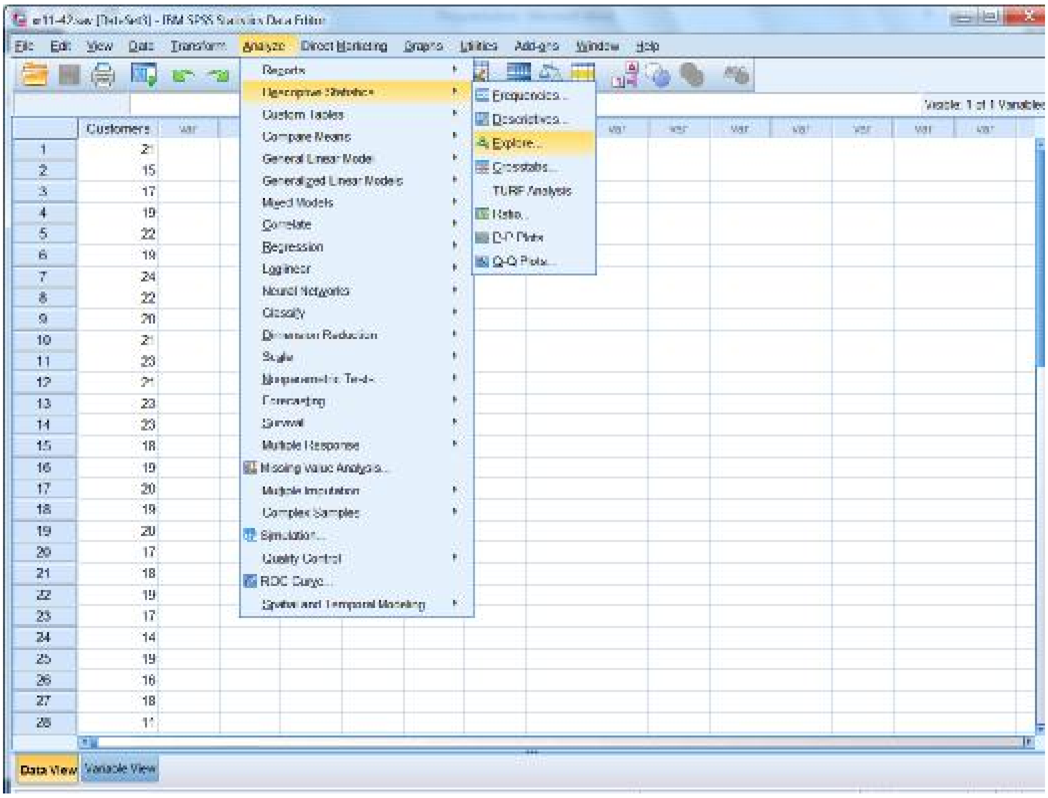
$\mu_0=20$. Ο έλεγχος θα είναι ο εξής:

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

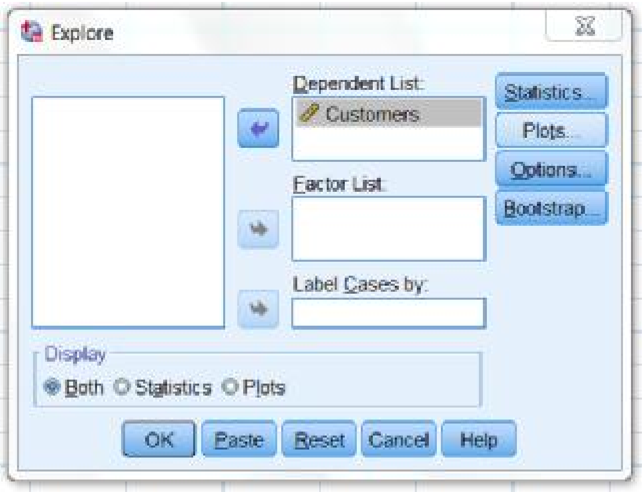
1)Αρχικά φτιάχνουμε τη μεταβλητή που έχουμε στο Variable View και στο Data View εισάγουμε τα δεδομένα της.

2)Πραγματοποιούμε έλεγχο κανονικότητας, για να δούμε εάν τα δεδομένα μας ακολουθούν την κανονική κατανομή:

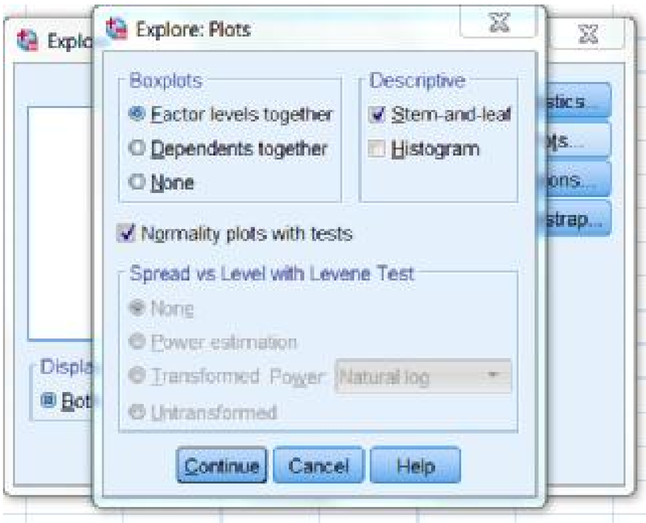
Analyze→Descriptive statistics→explore:



→στο πλαίσιο Dependent list τοποθετώ τη μεταβλητή των αυτοκινήτων:



→ Στην επιλογή Plots επιλέγουμε Normality Plots with tests προκειμένου να μας εμφανίσει τα απαραίτητα test και διαγράμματα για τον έλεγχο κανονικότητας:



→ OK

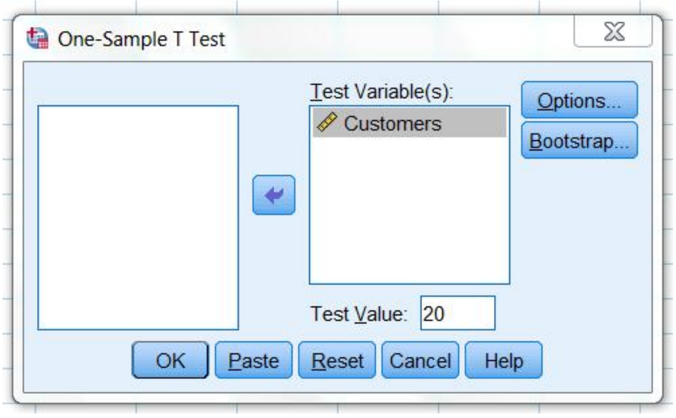
Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Customers	,114	36	,200*	,952	36	,122

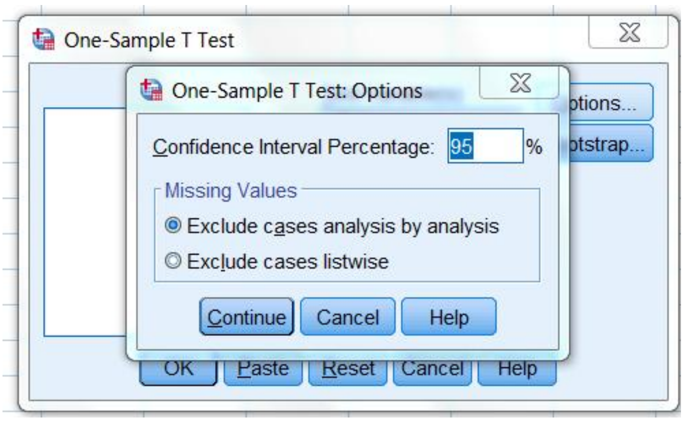
Εδώ έχουμε δύο κριτήρια που εξετάζουν τον έλεγχο κανονικότητας, οπότε θα ελέγξουμε και τα δύο:

Στο πρώτο κριτήριο (Kolmogorov-Smirnov) το sig=0,2> α(0,05) και στο δεύτερο κριτήριο (Shapiro-Wilk) το sig=0,122> α(0,05) ,άρα τα δεδομένα μας ακολουθούν την κανονική κατανομή, οπότε μπορούμε να προχωρήσουμε στο επόμενο βήμα.

3)Το κριτήριο που θα χρησιμοποιήσουμε για να συγκρίνουμε τον αριθμητικό μέσο μ του δείγματος με την τιμή των 20αυτοκινήτων είναι το One sample t-test:
Πάμε Analyze →Compare Means→One sample t-test→ περνάμε στο test variable τη μεταβλητή με τα αυτοκίνητα και στο test_value βάζουμε την τιμή 20 με την οποία θα συγκρίνουμε τη μέση τιμή:



Στη συνέχεια πηγαίνουμε στο κουμπί options και Confidence Interval Percentage όπου γράφουμε το επίπεδο εμπιστοσύνης με το οποίο θα γίνει ο έλεγχος. Στην περίπτωση μας είναι το 95% καθώς έχουμε στάθμη σημαντικότητας 5%:



4) CONTINUE→ OK και πάμε στο OUTPUT που εμφανίζεται:

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Customers	36	19,39	2,950	,492

Ο έλεγχος είναι μονόπλευρος.Απ’ τον πίνακα βλέπουμε ότι μ=19,39 < 20, άρα ικανοποιείται η ανισότητα της Η1 .
Πάμε και στο δεύτερο πίνακα:

One-Sample Test

	Test Value = 20					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Customers	-1,243	35	,222	-,611	-1,61	,39

Το sig=0.222. $\frac{sig}{2}$= 0,111>α(0,05), άρα αποδεχόμαστε την υπόθεση Hο, άρα δεν υπάρχου επαρκή στατιστικά στοιχεία που να δείχνουν ότι ο αριθμός των αυτοκινήτων έχει μειωθεί.
Η τιμή p όταν έχουμε ένα δείγμα, ορίζεται ως $\frac{sig}{2}$. Έτσι στο παράδειγμα μας είναι 0,111

4. Ισχύς ενός στατιστικού ελέγχου:

Ισχύς του ελέγχου, ονομάζεται η ικανότητα του ελέγχου να διακρίνει υπαρκτές σημαντικές διαφορές του δείγματος από την H_0 και έτσι να την απορρίπτει. Εξ αυτού ισούται με 1-β, και ο έλεγχος που για συγκεκριμένα δεδομένα αποδεικνύεται περισσότερες φορές σωστός, λέμε ότι έχει τη **μεγαλύτερη ισχύ**.

Έλεγχος όταν η τυπική απόκλιση είναι άγνωστη

Όλα τα παραπάνω ισχύουν όταν η τυπική απόκλιση σ είναι γνωστή, κάτι που συνήθως δεν υφίσταται. Όταν η τυπική απόκλιση σ του πληθυσμού είναι άγνωστη και ο πληθυσμός έχει κανονική κατανομή, χρησιμοποιούμε την τυπική απόκλιση sτου δείγματος και ο έλεγχος υποθέσεων για τον μέσο μ του πληθυσμού υπολογίζεται από τον τύπο: $\mathbf{t} = \frac{\bar{x}-\mu}{s/n}$.Ακολουθεί την κατανομή **student-t** με v=n-1 βαθμούς ελευθερίας.Τι είναι όμως βαθμός ελευθερίας;
Για να λύσουμε την εξίσωση t, πρέπει να γνωρίζουμε τον μέσο x του δείγματος. Οι κατανομές δειγματοληψίας παράγονται με την επανειλημμένη επιλογή δειγμάτων από τον ίδιο πληθυσμό. Αν έχουμε βρει τον μέσο του δείγματος και θέλουμε να επιλέξουμε ένα δείγμα μεγέθους n, μπορούμε να επιλέξουμε ελεύθερα τα πρώτα n-1 στοιχεία του δείγματος, αλλά το τελευταίο πρέπει να επιλεγεί ώστε ο μέσος να μείνει αμετάβλητος. Για παράδειγμα, αν n=4, x=10, και έχουμε επιλέξει x1=17, x2=5,x3=12 τότε τοx4 δε μπορεί να είναι άλλο νούμερο απ' το 6, για να έχουμε x=10 .Άρα στο παράδειγμα μας μπορέσαμε να επιλέξουμε ελεύθερα τα 3 από τα 4 νούμερα, δηλαδή έχουμε 4-1=3 βαθμούς ελευθερίας.

Ας δούμε τώρα πως γίνεται ο έλεγχος σε πραγματικές συνθήκες εν μέσω του επόμενου παραδείγματος:

<u>Παράδειγμα:</u>

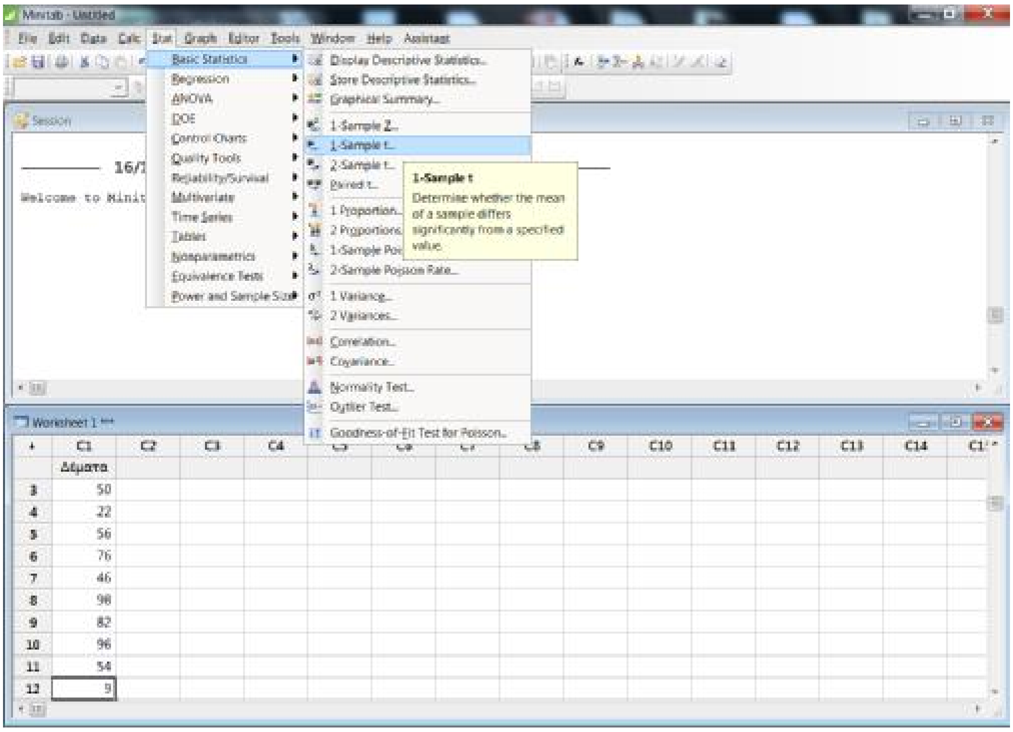
Μια εταιρία διανομής δεμάτων διαφημίζεται ότι ο χρόνος παράδοσης στην ίδια πόλη είναι μικρότερος από 6 ώρες. Για τον έλεγχο του ισχυρισμού αυτού επιλέχθηκε ένα τυχαίο δείγμα 12 δεμάτων με προορισμό στην ίδια πόλη και καταγράφηκε ο χρόνος παράδοσης σε ώρες, όπως φαίνεται στον παρακάτω πίνακα:

3,03	6,33	6,50	5,22	3,56	6,76
7,98	4,82	7,96	4,54	5,09	6,46

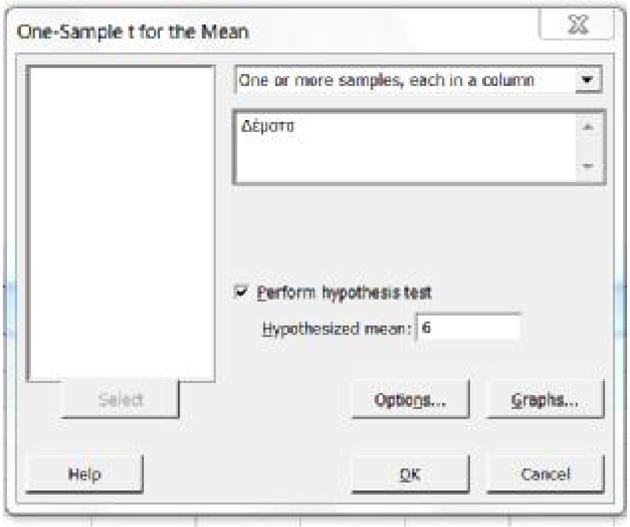
Μπορούμε από τα δεδομένα να συμπεράνουμε με στάθμη σημαντικότητας 5% ότι ο ισχυρισμός είναι αληθής;

<u>Λύση:</u>

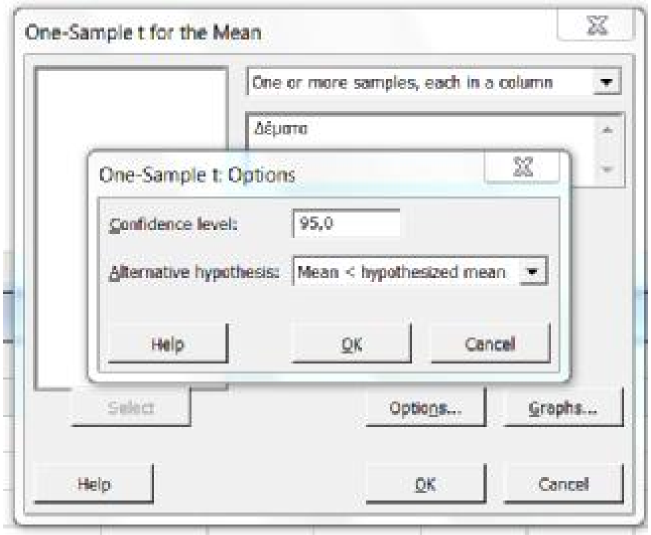
Θα λύσουμε την άσκηση στο στατιστικό πρόγραμμα Minitab:
1)Αρχικά θα δημιουργήσουμε μια μεταβλητή με όνομα «Δέματα» ,και θα περάσουμε τις τιμές του πίνακα.
2) Πάμε Stat → Basic statistics → 1-Sample t…



→Επιλέγω one or more samples, each in a column →Περνάω τη μεταβλητή «Δέματα» στο κουτί→επιλέγω“perform hypothesis test” → hypothesized mean βάζω την τιμή του μ0=6:



→Options → confidence level 95, καθώς έχω στάθμη σημαντικότητας 5% →alternative hypothesis επιλέγω “mean <hypothesized mean” καθώς θα ελέγξω εάν ο μέσος μ είναι μικρότερος του $μ_0$:



Το OUTPUTμας εμφανίζει τα εξής:

```
One-Sample T: Δέματα

Test of μ = 2 vs < 2

Variable   N   Mean  StDev  SE Mean  95% Upper Bound     T      P
Δέματα    12  5,687  1,580    0,456            6,507  8,08  1,000
```

Η τιμή P-value δείχνει πως δεν υπάρχει ισχυρή απόδειξη υπέρ της εναλλακτικής υπόθεσης, άρα δε μπορούμε να συμπεράνουμε ότι ο ισχυρισμός του είναι αληθής.

Κεφάλαιο VI: Εκτιμητής διαστήματος εμπιστοσύνης

1. Εκτιμητής διαστήματος εμπιστοσύνης όταν τυπική απόκλιση γνωστή:

Όταν η τυπική απόκλιση είναι γνωστή, ο πληθυσμός ακολουθεί κανονική κατανομή, και θέλουμε να βρούμε το διάστημα τιμών στο οποίο μπορεί να βρίσκεται ο μέσος του πληθυσμού, μπορούμε να χρησιμοποιήσουμε τον εκτιμητή διαστήματος εμπιστοσύνης, ο οποίος είναι: :$\bar{\boldsymbol{x}} \pm \mathbf{za/2} * \frac{\boldsymbol{\sigma}}{\sqrt{\boldsymbol{n}}}$

Παράδειγμα:

Ένα νέο μεσιτικό γραφείο προσπαθεί να αποφασίσει για το ποιος θα είναι ο ετήσιος μισθός που θα προσφέρει σε στελέχη που θα προσεγγίσει. Ως εκ τούτου έκανε μια έρευνα αγοράς επιλέγοντας ένα τυχαίο δείγμα 350 μεσιτών από όλη τη χώρα και κατέγραψε τις ετήσιες αμοιβές τους. Οι αμοιβές φαίνονται στον παρακάτω πίνακα:

Αμοιβές								
29109	23437	20760	36364	23167	20727	27177	29631	34660
21546	31921	25332	35330	19559	23450	25161	25330	20765
30417	19869	28665	45575	53556	17718	25976	34295	38715
10104	31693	21338	33222	21745	37815	31626	33129	28147
19279	38974	33246	15717	24758	27443	33644	33399	13520
27578	34767	27501	21607	20145	26913	24163	30039	33985
23581	33131	1327	28861	37975	27273	41845	25912	27194
26949	23815	39562	22407	22640	34833	36000	26605	42318
35423	26714	30752	11334	28622	27059	41866	10286	40073
12971	36735	21887	18836	24653	36773	21561	40673	36581
37895	33136	32228	13190	25567	10541	36483	40759	28977
31308	34007	10133	26210	19951	36607	38562	41675	29865
28256	39674	25943	14940	11195	22582	38257	26489	43225
31494	22759	38223	48983	30646	24738	32205	25183	29897

31552	27323	38362	21182	25490	31808	26054	21902	24342
34440	28875	27569	29351	17034	45441	32774	24217	16257
33347	37991	38213	27038	31027	36924	15059	9053	21263
26768	23100	25181	19260	32939	17833	38706	23633	28967
25225	29289	29607	32626	26919	24812	31978	24876	30534
29250	21065	29690	31240	27831	27882	16295	36081	24586
27497	6914	27114	13790	23640	37357	8113	35547	39966
37739	32770	16117	35781	35146	31767	31372	23561	34416
35311	38143	25961	44163	30657	31569	21404	38169	51462
17283	24482	27658	23179	28479	29122	27984	51824	30368
47932	29892	41604	27684	40563	40277	33828	34754	30437
33411	14099	19776	32202	27587	17311	29313	21164	19764
24062	28276	28030	29616	14896	21835	36379	38510	22108
31911	22362	49731	31265	35955	28813	18626	28641	41226
34260	35448	20149	27711	37899	37453	26992	25296	30185
45958	35193	31712	31661	31861	36700	34019	37486	35435
28929	30938	15141	30009	30029	33457	13383	38751	34450
20889	22874	37742	20299	24875	26134	19651	16922	36943
38709	13906	21791	27605	34903	30112	34292	33310	36075
49685	16270	22124	34780	21248	27828	45159	43229	22579
22010	30643	21605	36633	29461	27208	22635	32430	25277
37126	27880	20258	39882	26043	29144	29829	35261	32618
30196	33914	36634	48667	28026	19284	25022	29148	19571
23522	27181	22245	33194	32637	23400	36841	37418	27821
31948	36287	14125	33926	45792	28025	41440	38941	

Αν οι αμοιβές έχουν κανονική κατανομή με τυπική απόκλιση 8.000 δολάρια, μπορούμε να βρούμε με στάθμη σημαντικότητας 5% το διάστημα που κυμαίνεται η μέση τιμή των ετήσιων αμοιβών;

Λύση:

Η στάθμη σημαντικότητας είναι 5%, άρα $Z_{\alpha/2} = Z_{0,025} = 1,96$. Η περιοχή απόρριψης θα είναι: $Z > Z_{\alpha/2}$

Τώρα θα βρούμε τη μέση τιμή του δείγματος:

$\bar{x} = \frac{\sum xi}{n} = \frac{10191832}{350} = 29119,52$

Άρα η τιμή του ελέγχου είναι:

$Z = \bar{x} \pm Z_{a/2} * \frac{\sigma}{\sqrt{n}} = 29119,52 \pm 1,96 * 427,62 = 29119,52 \pm 838,13$

Συνεπώς η μέση τιμή των αμοιβών είναι μεταξύ 29.957,13 και 28.281,39 χιλιάδες ευρώ.

2. <u>Εκτιμητής διαστήματος εμπιστοσύνης όταν τυπική απόκλιση άγνωστη</u>

Όταν η τυπική απόκλιση είναι άγνωστη, ο πληθυσμός ακολουθεί κανονική κατανομή, και χρειαζόμαστε διάστημα τιμών που μπορεί να πάρει ο μέσος μ του πληθυσμού, χρησιμοποιούμε τον εκτιμητή διαστήματος εμπιστοσύνης, ο οποίος δίνεται από τον τύπο:

$$\boldsymbol{\bar{x} \pm t_{\alpha/2} * \frac{s}{\sqrt{n}}}$$

<u>Παράδειγμα:</u>

Στη διάρκεια της τελευταίας δεκαετίας δημιουργήθηκαν στις ΗΠΑ αρκετά ιδρύματα με σκοπό τη βελτίωση προϊόντων και υπηρεσιών, και απομένουν ετήσια βραβεία στις επιχειρήσεις που παρουσιάζουν την υψηλότερη ποιότητα στον τομέα τους. Ένας επενδυτής πιστεύει ότι οι μετοχές των βραβευμένων εταιριών έχουν υψηλότερη απόδοση, και για να επιβεβαιώσει την άποψη αυτή επέλεξε ένα τυχαίο δείγμα 83 εταιριών που βραβεύτηκαν τον τελευταίο χρόνο και κατέγραψε την ετήσια απόδοση των κεφαλαίων τους. Οι αποδόσεις τους φαίνονται στον παρακάτω πίνακα.

16,2	0,58	0,47	11,42	16,34	-1,06	15,52	5,14	15,45
15,34	16,65	27,25	28,26	22,69	22,45	16,82	14,3	21,58
19,55	24,6	25,46	21,45	-0,21	7,23	17,52	24,77	14,15
13,53	-2,27	18,07	21,36	7,76	26,2	29,68	23,33	
6,76	-0,67	2,62	23,8	2,28	21,93	18,02	9,88	
13,9	12,8	13,45	9,01	6,89	21,02	22,37	18,41	
11,72	26,01	9,45	26,58	23,94	17,85	25,32	2,26	
17,63	15,46	12,74	6,51	9,62	7,9	22,78	16,52	
10,97	-1,16	13,6	18,63	13,72	12,54	14,65	5,66	
22,1	7,41	-3,6	24,69	21,39	20,95	21,15	6,39	

Να εκτιμήσετε με στάθμη εμπιστοσύνης 95% τη μέση απόδοση των μετοχών των βραβευμένων εταιριών.

<u>Λύση:</u>

Για να βρούμε το διάστημα τιμών στο οποίο κινείται η μέση απόδοση των μετοχών των εταιριών, θα χρησιμοποιήσουμε τον εκτιμητή διαστήματος εμπιστοσύνης, με τύπο:

$\bar{x} \pm t_{\alpha/2} * \frac{s}{\sqrt{n}}$

Από τα δεδομένα έχουμε:

$\sum x_i$= 1246,43
$\sum x_i^2$=24374,22

Η μέση τιμή $\bar{x}$ είναι: $\bar{x} = \frac{\sum x_i}{n} = \frac{1246,43}{83} = 15,017$

$t_{\alpha/2.n}$= $t_{0025.83} \approx t_{0025.80}$ =1,99

$s^2 = \frac{\sum xi^2 - \frac{\sum xi^2}{n}}{n-1} = \frac{24374,22 - 18717,93}{82} = \frac{5656,29}{82} = 68,98$

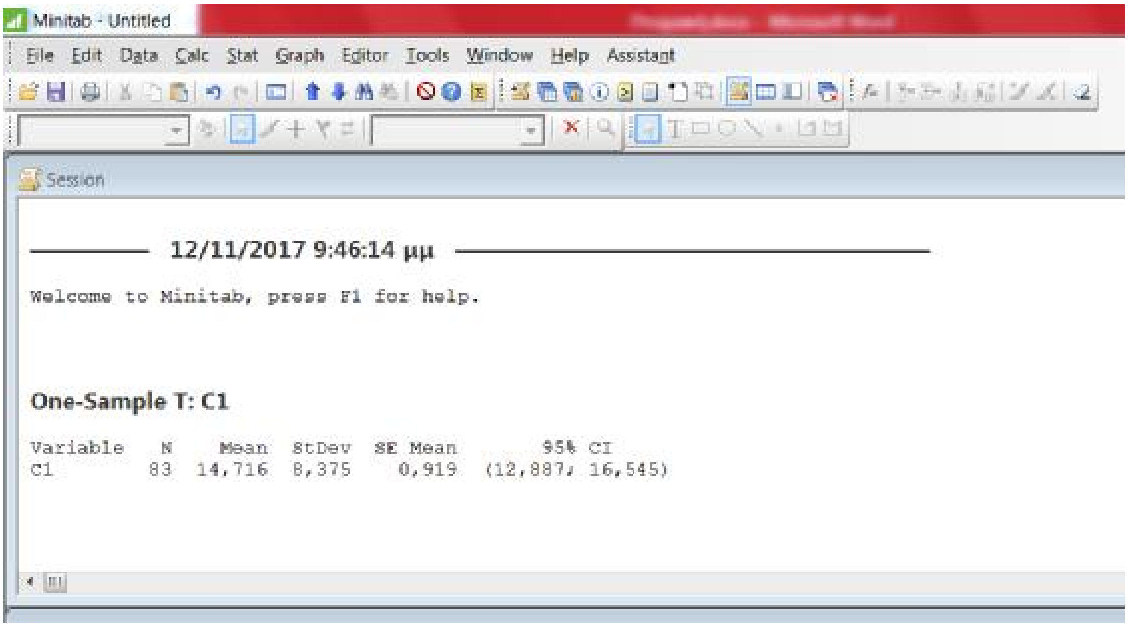
S = $\sqrt{S^2}$ = 8,3

Αντικαθιστώντας τις τιμές στον τύπο, έχουμε το εξής:

$\bar{x} \pm t_{\alpha/2} * \frac{s}{\sqrt{n}} = 15,017 \pm 1,99 \cdot \frac{8,3}{\sqrt{83}} = 15,017 \pm 1,81$

Ή LCL=13,21 και UCL=16,83

Τώρα θα λύσουμε την άσκηση στο στατιστικό πρόγραμμα Minitab:
1)Αρχικά θα δημιουργήσουμε μια μεταβλητή με όνομα «Απόδοση_μετοχών» ,και θα περάσουμε τις τιμές του πίνακα.
2) Πάμε Stat → Basic statistics → 1-Sample t.. →Επιλέγω one or more samples, each in a column →Περνάω τη μεταβλητή «Απόδοση_μετοχών» στο κουτί→options → confidence level 95, καθώς έχω στάθμη σημαντικότητας 5% →alternative hypothesis επιλέγω “mean ≠ hypothesized mean “ καθώς θα ελέγξω το εύρος των τιμών που μπορεί να πάρει η απόδοση των μετοχών.
Το OUTPUTμας εμφανίζει τα εξής:



Σύμφωνα με το αποτέλεσμα, η μέση απόδοση των μετοχών των βραβευμένων εταιριών είναι μεταξύ 12,887 και 16,545 δολαρίων.

Κεφάλαιο VII: Διασπορά ενός πληθυσμού

Για να σχηματίσουμε τον έλεγχο υποθέσεων της διασποράς ενός πληθυσμού, χρειαζόμαστε ένα στατιστικό δείκτη ως εκτιμητή. Ο εκτιμητής πρέπει να πληροί κάποιες προϋποθέσεις: Να είναι **αμερόληπτος εκτιμητής**, δηλαδή η αναμενόμενη τιμή του να ισούται με την τιμή της παραμέτρου του πληθυσμού, και ταυτοχρόνως να είναι **συνεπής εκτιμητής**, δηλαδή η διαφορά μεταξύ της τιμής του και της τιμής της παραμέτρου να μειώνεται όσο αυξάνεται το μέγεθος του πληθυσμού. Θα επιλέξουμε τη διασπορά $\boldsymbol{s^2}$,καθώς είναι ένας αμερόληπτος και συνεπής εκτιμητής της διασποράς του πληθυσμού.
Με δεδομένο ότι έχουμε κανονική κατανομή, το άθροισμα των τετραγώνων των αποκλίσεων θα γίνει:$\sum(xi - x)^2 = (n-1)s^2$. Στη συνέχεια, θα διαιρεθεί με τη διασπορά σ^2 και έτσι θα δημιουργηθεί μια τυχαία μεταβλητή η οποία ακολουθεί την κατανομή χ^2 με ν=n-1 βαθμούς ελευθερίας. Έτσι δημιουργείται ο **έλεγχος $\boldsymbol{\chi^2}$** με τύπο:

$$\frac{(\mathrm{n}-1)s^2}{\sigma^2}$$

Εκτιμητής διαστήματος εμπιστοσύνης:

Το διάστημα εμπιστοσύνης για τη διασπορά του πληθυσμού είναι:

LCL= $\frac{(\mathrm{n}-1)s^2}{\chi\alpha/2^2}$ και UCL= $\frac{(\mathrm{n}-1)s^2}{\chi 1-\alpha/2^2}$

Παράδειγμα:

Στον παρακάτω πίνακα καταγράφεται το ακριβές βάρος για ένα τυχαίο δείγμα συσκευασιών δημητριακών με ονομαστικό περιεχόμενο ένα λίβρα(pound). Να εκτιμήσετε τη διασπορά του πληθυσμού με στάθμη εμπιστοσύνης 90%.

1,05	1,03	0,98	1,00	0,99	0,97	1,01	0,96

Λύση:

Τα δύο άκρα θα είναι τα εξής:

LCL= $\frac{(\mathrm{n}-1)s^2}{\chi\alpha/2^2}$, UCL= $\frac{(\mathrm{n}-1)s^2}{\chi 1-\alpha/2^2}$

$\sum x$= 7.99
$\sum x^2$= 63.8401

$s^2 = \frac{\sum xi^2 - \frac{\sum xi^2}{n}}{n-1} = \frac{63{,}8401 - 7{,}9801}{7} = \frac{55{,}86}{7} = 7{,}98$

$x^2{}_{\alpha/2.\mathrm{n}-1} = x^2{}_{0,1/2.8-1} = x^2{}_{0.05.7} = 15{,}5$
$x^2{}_{1-\alpha/2.\mathrm{n}-1} = x^2{}_{0,95.7} = 2{,}73$

Άρα:

LCL= $\frac{(\mathrm{n}-1)s^2}{\chi\alpha/2^2} = \frac{(8-1)7{,}98}{15{,}5} = 3{,}60$

UCL= $\frac{(\mathrm{n}-1)s^2}{\chi 1-\alpha/2^2} = \frac{(8-1)7{,}98}{2{,}73} = 20{,}46$

Συμπεραίνουμε ότι η διασπορά του βάρους είναι μεταξύ 3,60 και 20,46 με στάθμη εμπιστοσύνης 90%.

Κεφάλαιο VIII: Αναλογία ενός πληθυσμού

Εάν τα δεδομένα που έχουμε δεν είναι συνεχή αλλά ονομαστικά, δεν μπορούμε να εκτιμήσουμε ποσοτικές παραμέτρους, αλλά μόνο συχνότητες και αναλογίες. Δηλαδή σε έναν πληθυσμό μετράμε πόσες φορές εμφανίζεται η κάθε τιμή και βγάζουμε τις συχνότητες, απ' τις οποίες στη συνέχεια βγαίνουν και οι σχετικές συχνότητες, δηλαδή οι **αναλογίες** στον πληθυσμό, οι οποίες συμβολίζονται με $\hat{p}$.Οι ονομαστικές τιμές, μπορούν να πάρουν μόνο 2 τιμές η μια είναι η <<επιτυχία>>, και είναι η τιμή για την οποία ενδιαφερόμαστε, (σε μια δημοσκόπηση <<επιτυχία>> θα είναι η προτίμηση στο πρόσωπο ενός υποψηφίου), και αποτυχία είναι όλες οι υπόλοιπες τιμές. Ο στατιστικός δείκτης έχει τύπο:

$$\hat{p} = \frac{x}{n},$$

όπου x ο αριθμός των επιτυχιών σε ένα δείγμα και n το μέγεθος του δείγματος.

Έλεγχος υποθέσεων:

Ο έλεγχος υποθέσεων στην αναλογία ενός πληθυσμού είναι:

$$z = \frac{\hat{p} - p}{\sqrt{p(1-p)/n}},$$

και ακολουθεί κανονική κατανομή με προϋπόθεση ότι $n*p>5$ και $n*(1-p)>5$.

Διάστημα εμπιστοσύνης:

Το διάστημα εμπιστοσύνης το υπολογίζουμε ως εξής:

$$\hat{p} \pm z_{a/2}\sqrt{p(1-p)/n}$$

με προϋπόθεση ότι $n*p>5$ και $n*(1-p)>5$.

Παράδειγμα:

Την άνοιξη του 1988 ένας εκδότης εφημερίδων κυκλοφόρησε μια νέα εφημερίδα «εθνικής εμβέλειας» στον Καναδά. Το κριτήριο για την οικονομική βιωσιμότητα της νέας εφημερίδας είχε εκτιμηθεί σε ένα ποσοστό μεγαλύτερο από 12% της αγοράς του Τορόντο. Στη διάρκεια της σχεδίασης της εφημερίδας, είχε πραγματοποιηθεί μια έρευνα αγοράς σε ένα τυχαίο δείγμα 400 αναγνωστών στην περιοχή του Τορόντο. Η έρευνα έδινε στους αναγνώστες μια σύντομη περιγραφή των χαρακτηριστικών της εφημερίδας και στη συνέχεια ρωτούσε εάν θα επέλεγαν να γίνουν συνδρομητές μιας εφημερίδας με αυτά τα χαρακτηριστικά αν το κόστος δε ξεπερνούσε τα 20 δολάρια το μήνα. Οι απαντήσεις έχουν κωδικοποιηθεί ως 1=ΟΧΙ και 2=ΝΑΙ, και φαίνονται στον παρακάτω πίνακα:

Απαντήσεις									
1	2	1	1	1	2	2	1	1	1
1	1	1	1	1	1	2	1	2	1
1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1
1	1	1	1	2	1	1	1	2	1
1	1	1	2	1	1	2	1	1	2
1	1	1	1	2	2	1	1	2	2
1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1
1	1	1	2	1	2	1	2	1	2
1	1	2	1	2	1	1	2	1	1
1	1	1	2	1	1	1	1	1	1
1	2	1	2	1	1	2	1	1	1
1	1	1	2	1	1	1	1	1	1
1	1	1	1	1	1	1	2	2	1
1	1	1	1	2	1	1	1	1	1
1	1	1	1	1	2	1	2	1	1
1	1	1	1	1	1	1	1	1	1
1	1	1	1	2	1	1	1	1	2
1	1	1	1	1	1	2	1	1	1
1	2	1	2	1	1	2	2	1	1
1	2	1	1	1	1	1	1	1	
2	1	1	1	1	1	1	2	2	
1	1	1	1	1	2	1	1	1	
1	1	1	1	1	1	1	1	1	
2	1	1	1	1	1	2	1	1	
1	1	1	1	2	1	1	1	1	
1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	
1	1	2	1	1	1	1	1	1	
1	1	1	2	1	2	1	1	1	
1	1	2	2	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	2	
2	1	1	2	1	1	1	1	2	

1	2	1	1	1	1	1	1	1	
1	1	1	1	2	1	1	2	1	
1	2	1	2	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	
1	1	2	1	2	1	1	1	1	
1	1	1	2	1	1	1	1	1	

Μπορούμε από τα δεδομένα της έρευνα να συμπεράνουμε με στάθμη σημαντικότητας 5% αν η εφημερίδα θα είναι οικονομικά βιώσιμη;

Λύση:

Το ζητούμενο είναι να είναι η εφημερίδα οικονομικά βιώσιμη. Για να είναι, πρέπει το μερίδιο αγοράς της να είναι μεγαλύτερο από 12%. Άρα ο έλεγχος υποθέσεων θα είναι:

$$H_o: p = 0{,}12$$

$$H_1: p > 0{,}12$$

Αρχικά θα υπολογίσουμε την αναλογία του πληθυσμού των απαντήσεων των αναγνωστών:

$$\hat{p} = \frac{x}{n} = \frac{65}{400} = 0{,}1625$$

Εν συνεχεία, θα υπολογίσουμε την τιμή του ελέγχου:

$$Z = \frac{\hat{p}-p}{\sqrt{p(1-p)/n}} = \frac{0{,}1625-0{,}12}{\sqrt{0{,}12(1-0{,}12)/400}} = \frac{0{,}0425}{0{,}016} = 2{,}656$$

Οι προϋποθέσεις του ελέγχου, $n*p>5$ και $n*(1-p)>5$ ισχύουν, καθώς:

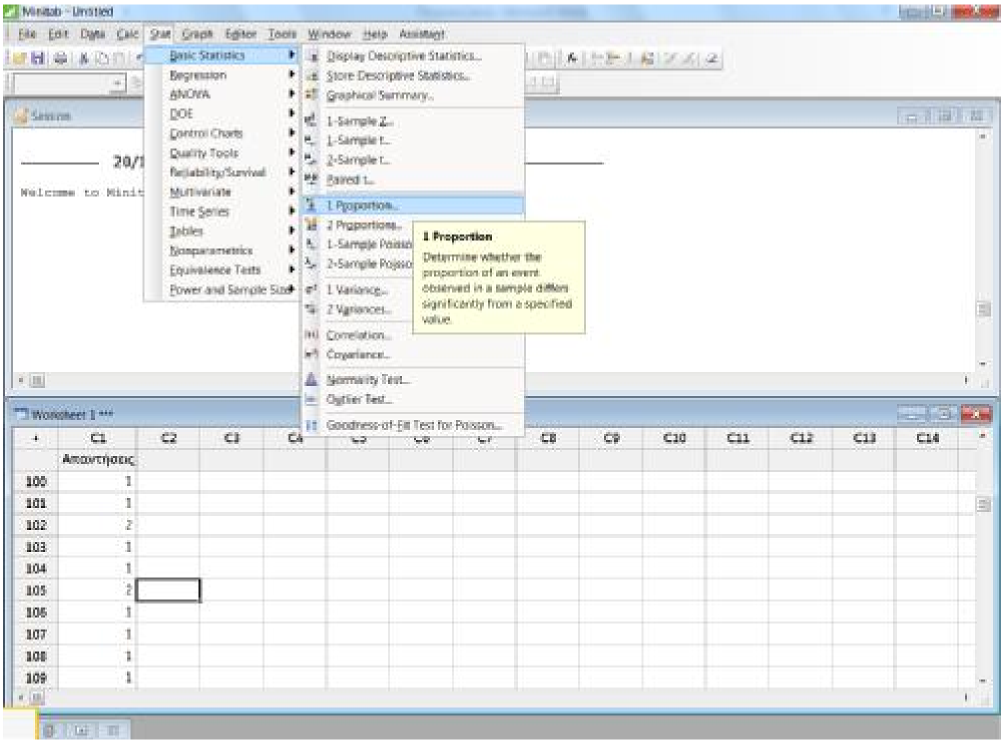
$$400*0{,}12=48 > 5 \text{ και } 400*(1-0{,}12)=352 > 5.$$

Η περιοχή απόρριψης είναι:

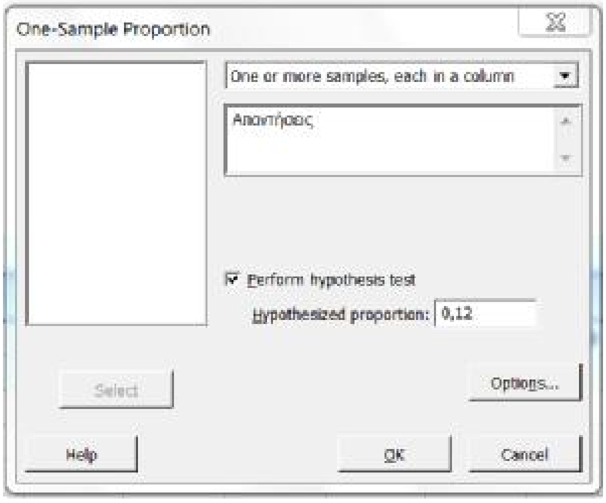
$$z > z_{\alpha} = z_{0,05} = 1{,}645$$

Η τιμή του ελέγχου βρίσκεται στην περιοχή απόρριψης, άρα απορρίπτουμε την μηδενική υπόθεση και συμπεραίνουμε ότι η νέα εφημερίδα θα έχει ένα μερίδιο αγοράς μεγαλύτερο από 12%, άρα θα είναι οικονομικά βιώσιμη.

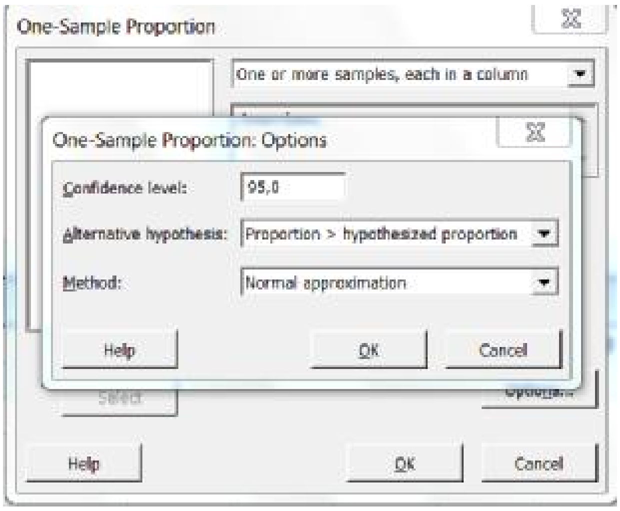
Ας δούμε πως θα λυνόταν στο στατιστικό πρόγραμμα Minitab:
1)Αρχικά θα δημιουργήσουμε μια μεταβλητή με όνομα «Απαντήσεις» και θα περάσουμε τα δεδομένα της.
2)Στη συνέχεια πάμε: Stat → Basic statistics → 1 proportion:



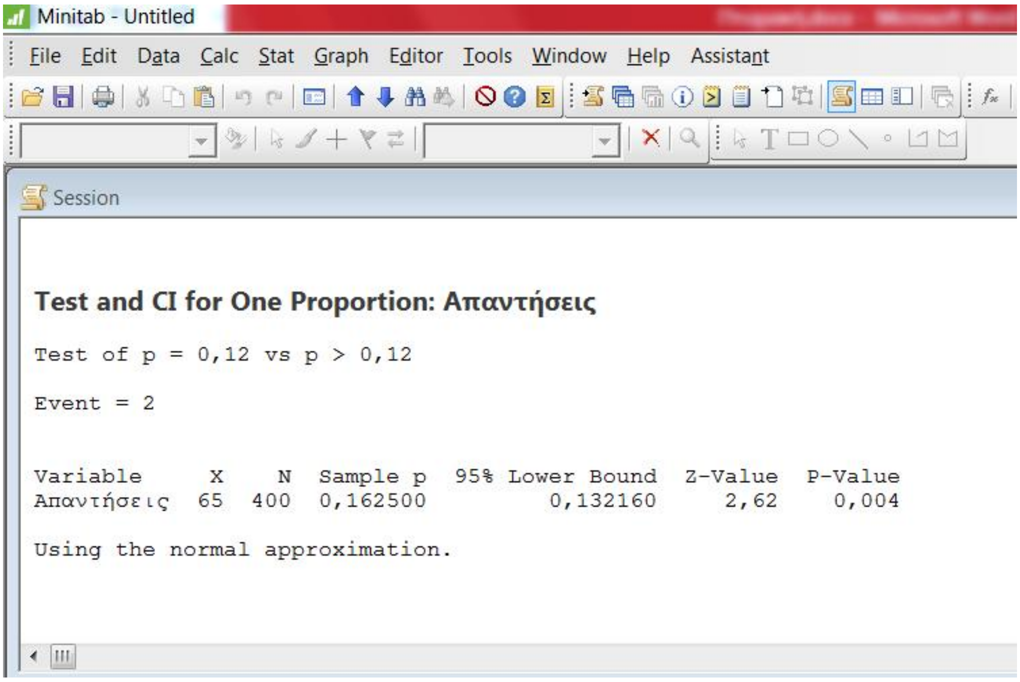
→ επιλέγουμε one or more samples, each in a column και περνάμε τη μεταβλητή «Απαντήσεις»
→ Επιλέγω “perform hypothesis test” και θέτω την τιμή 0,12:



→ options → confidence level: 95% → alternative hypothesis: proportion > hypothesized proportion, διότι για την εναλλακτική υπόθεση θέλουμε p> 0,12. → method “normal approximation” καθώς ακολουθεί την κανονική κατανομή:



3) Πάμε στα αποτελέσματα που εμφανίζει το OUTPUT:



Η τιμή του ελέγχου είναι 2,62 με p=0,004, άρα ο έλεγχος είναι στατιστικά πολύ σημαντικός και υπάρχει συντριπτική απόδειξη υπέρ της εναλλακτικής υπόθεσης, δηλαδή η νέα εφημερίδα θα είναι οικονομικά βιώσιμη.

Κεφάλαιο IX: Έλεγχος υποθέσεων για δύο δείγματα

1. Βασικές έννοιες του ελέγχου υποθέσεων δύο δειγμάτων:

Για να δοκιμάσουμε υποθέσεις πρέπει να μετρήσουμε τις μεταβλητές. Οι μεταβλητές είναι πράγματα που μπορούν να αλλάξουν (ή να ποικίλουν), μπορεί να ποικίλλουν μεταξύ των ατόμων (π.χ. IQ, συμπεριφορά) ή ως προς τοποθετήσεις (π.χ. ανεργία) ή ακόμη και ως προς το χρόνο (π.χ. διάθεση, κέρδος, αριθμός καρκινικών κυττάρων). Οι περισσότερες υποθέσεις μπορούν να εκφραστούν με δύο μεταβλητές: μια προτεινόμενη αιτία και ένα προτεινόμενο αποτέλεσμα. Για παράδειγμα, εάν λάβουμε την επιστημονική δήλωση ότι η «Coca-Cola είναι ένα αποτελεσματικό σπερματοκτόνο» τότε η προτεινόμενη αιτία είναι η «Coca-Cola» και το προτεινόμενο αποτέλεσμα είναι το νεκρό σπέρμα. Τόσο η αιτία όσο και το αποτέλεσμα είναι μεταβλητές: για την αιτία θα μπορούσαμε να διαφοροποιήσουμε τον τύπο του ποτού, και για το αποτέλεσμα, αυτά τα ποτά θα σκοτώσουν διαφορετικές ποσότητες σπέρματος. Το κλειδί για τη δοκιμή τέτοιων δηλώσεων είναι η μέτρηση αυτών των δύο μεταβλητών. Μια μεταβλητή που πιστεύουμε ότι είναι μια αιτία είναι γνωστή ως ανεξάρτητη μεταβλητή (επειδή η αξία της δεν εξαρτάται από άλλες μεταβλητές). Μια μεταβλητή που πιστεύουμε ότι είναι ένα αποτέλεσμα ονομάζεται εξαρτημένη μεταβλητή επειδή η τιμή αυτής της μεταβλητής εξαρτάται από την αιτία (ανεξάρτητη μεταβλητή). Αυτοί οι όροι συνδέονται πολύ στενά με πειραματικές μεθόδους, στις οποίες ο αιτιώδης παράγοντας χειρίζεται πραγματικά τον πειραματιστή. Στην πειραματική εργασία η αιτία ή η ανεξάρτητη μεταβλητή είναι ένας προγνωστικός παράγοντας και η επίδραση ή η εξαρτώμενη μεταβλητή είναι απλώς ένα αποτέλεσμα.

2. Έλεγχος υποθέσεων για τον μέσο δύο δειγμάτων (εξαρτημένα και ανεξάρτητα δείγματα)

1.1 Ανεξάρτητα δείγματα:

Ας δούμε πως συγκρίνουμε τους μέσους 2 πληθυσμών. Επιλέγουμε 2 τυχαία δείγματα από τους πληθυσμούς, τα οποία είναι ανεξάρτητα μεταξύ τους, δηλαδή η επιλογή του ενός δε σχετίζεται με κανένα τρόπο με την επιλογή του άλλου. Τα δύο δείγματα, τα συμβολίζουμε με n_1 και n_2 αντίστοιχα. Αν μ_1 ο μέσος του πληθυσμού 1 και μ_2 ο μέσος του πληθυσμού 2, θα συμβολίσουμε τους μέσους των δύο δειγμάτων με $\bar{x}_1$ και $\bar{x}_2$ αντίστοιχα. Ο καλύτερος εκτιμητής της διαφοράς μ_1 - μ_2 μεταξύ των δύο πληθυσμών, είναι η διαφορά $\boldsymbol{\bar{x}_1}$ - $\boldsymbol{\bar{x}_2}$ των μέσων των αντίστοιχων δειγμάτων.
Η κατανομή δειγματοληψίας της διαφοράς των δύο δειγμάτων $\bar{x}_1$ - $\bar{x}_2$ είναι κανονική εάν οι δύο πληθυσμοί απ' τους οποίους προέρχονται ακολουθούν την κανονική κατανομή, και κατά προσέγγιση κανονική όταν οι δύο πληθυσμοί δεν έχουν κανονική κατανομή αλλά τα δύο δείγματα είναι μεγάλα.

Το τυπικό σφάλμα θα είναι: $\sqrt{\frac{\sigma 1^2}{\mathrm{n1}} + \frac{\sigma 2^2}{n2}}$

Ο έλεγχος υποθέσεων στα ανεξάρτητα δείγματα με τυπική απόκλιση γνωστή διαμορφώνεται ως εξής:

$$z=\frac{(\bar{x}1 - \bar{x}2)-(\mu 1-\mu 2)}{\sqrt{\frac{\sigma 1^2}{n1}+\frac{\sigma 2^2}{n2}}}$$

και ο εκτιμητής διαστήματος εμπιστοσύνης ως:

$$\overline{(x}1 - \bar{x}2) \pm Z\alpha/2\sqrt{\frac{\sigma 1^2}{n1}+\frac{\sigma 2^2}{n2}}$$

Επί του πρακτέος, αυτούς τους τύπους τους χρησιμοποιούμε σπάνια, καθώς σχεδόν πάντα η διασπορά του πληθυσμού είναι άγνωστη. Έτσι, συνήθως χρειάζεται να βρίσκουμε το τυπικό σφάλμα. Ο τρόπος που θα το βρούμε, εξαρτάται απ’ τη σχέση των διασπορών των δύο πληθυσμών. Ας δούμε τις περιπτώσεις:

3. Έλεγχος υποθέσεων για ίσες διασπορές:

Εδώ θα χρησιμοποιήσουμε μια νέα μεταβλητή, την **σταθμισμένη διασποράς** , η οποία συμβολίζεται με sp^2

$$Sp^2=\frac{(n1-1)\,\sigma 1^2+(n2-1)\,\sigma 2^2}{n1+n2-2}$$

$$t=\frac{(\bar{x}1 - \bar{x}2)-(\mu 1-\mu 2)}{\sqrt{sp^2(\frac{1}{n1}+\frac{1}{n2})}}$$

με $v = n_1+n_2-2$ βαθμούς ελευθερίας

Οπότε ο **εκτιμητής διαστήματος εμπιστοσύνης** θα είναι:

$$(\bar{x}1 - \bar{x}2) \pm t\alpha/2\sqrt{Sp^2(\frac{\sigma 1^2}{n1}+\frac{\sigma 2^2}{n2})}$$

με $v = n_1+n_2-2$ βαθμούς ελευθερίας

Προσοχή: Απαραίτητη προϋπόθεση για να ακολουθούν κατανομή student t με $v = n_1+n_2-2$ βαθμούς ελευθερίας ο παραπάνω έλεγχος και ο εκτιμητής διαστήματος εμπιστοσύνης είναι οι δύο πληθυσμοί των δειγμάτων να ακολουθούν **κανονική κατανομή**.

4. Έλεγχος υποθέσεων για άνισες διασπορές:

Όταν οι διασπορές είναι άνισες, δεν μπορούμε να χρησιμοποιήσουμε τη σταθμισμένη διασπορά, οπότε ως εκτιμητής διασποράς κάθε πληθυσμού θα χρησιμοποιηθεί η διασπορά του αντίστοιχου δείγματος. Έτσι ο έλεγχος υποθέσεων θα γίνει:

$$t = \frac{(\overline{x}1 - \bar{x}2) - (\mu 1 - \mu 2)}{\sqrt{(\frac{\sigma 1^2}{n1} + \frac{\sigma 2^2}{n2})}}$$

Με $v = \frac{(\frac{\sigma 1^2}{n1} + \frac{\sigma 2^2}{n2})^2}{\frac{(\frac{\sigma 1^2}{n1})}{n1-1} + \frac{(\frac{\sigma 2^2}{n2})}{n2-1}}$ βαθμούς ελευθερίας

Το διάστημα εμπιστοσύνης διαμορφώνεται ως εξής:

$$(\bar{x}1 - \bar{x}2) \pm t\alpha/2\sqrt{\frac{\sigma 1^2}{n1} + \frac{\sigma 2^2}{n2}}$$

Για $v = \frac{(\frac{\sigma 1^2}{n1} + \frac{\sigma 2^2}{n2})^2}{\frac{(\frac{\sigma 1^2}{n1})}{n1-1} + \frac{(\frac{\sigma 2^2}{n2})}{n2-1}}$ βαθμούς ελευθερίας

5. <u>Έλεγχος για ισότητα διασπορών</u>:

Ένα εύλογο ερώτημα είναι το εξής. Πως μιλάμε για ίσες και άνισες διασπορές, εφόσον είναι άγνωστες; Η απάντηση είναι ότι δε μπορούμε να τις γνωρίζουμε, αλλά μπορούμε να κάνουμε μια σχετική εκτίμηση με βάση τις διασπορές των δύο αντίστοιχων δειγμάτων. Ας δούμε τώρα πως γίνεται ο **έλεγχος** για να συμπεράνουμε αν δύο διασπορές είναι ίσες:

Ho: $\frac{\sigma 1^2}{\sigma 2^2} = 1$, H1: $\frac{\sigma 1^2}{\sigma 2^2} \neq 1$

Καθώς οι διασπορές των αντίστοιχων δειγμάτων είναι αμερόληπτοι εκτιμητές των διασπορών των πληθυσμών, ο έλεγχος θα γίνει:

F= : $\frac{s1^2}{s2^2}$

Για πληθυσμούς που ακολουθούν την κανονική κατανομή, με $v_1 = n_1 - 1$ βαθμούς ελευθερίας για τον αριθμητή και $v_2 = n_2 - 1$ βαθμούς ελευθερίας για τον παρονομαστή. Ο έλεγχος θα ακολουθεί την κατανομή F.

Περιοχή απόρριψης:

Η περιοχή απόρριψης θα βρίσκεται και στα δύο άκρα, καθώς έχουμε αμφίπλευρο έλεγχο, οπότε θα είναι:

$$\mathbf{F > F_{\alpha/2.v1*v2}} \text{ ή } \mathbf{F < F_{1-\alpha/2.v1*v2}}$$

Το ερώτημα είναι, καθώς δεν γνωρίζουμε τις διασπορές, εάν θα χρησιμοποιούμε τους τύπους για ίσες ή για άνισες.

Ο μεγαλύτερος αριθμός βαθμών ελευθερίας οδηγεί σε ισχυρότερους ελέγχους υποθέσεων, δηλαδή υπάρχει μειωμένη πιθανότητα να προκύψει σφάλμα τύπου ΙΙ. Αποδεικνύεται μαθηματικά ότι για τους τύπους ίσων διασπορών έχουμε μεγαλύτερο αριθμό βαθμών ελευθερίας απ’ ότι στους τύπους άνισων διασπορών. Ως εκ τούτου, θα χρησιμοποιούμε τους τύπους για ίσες διασπορές, και τους τύπους για άνισες διασπορές μόνο στην περίπτωση που απορριφθεί η μηδενική υπόθεση στον έλεγχο υποθέσεων.

Παράδειγμα:

Ένας κατασκευαστής κλιματιστικών για αυτοκίνητα σκέφτεται να αλλάξει τον προμηθευτή των συμπυκνωτών που χρησιμοποιεί για τα κλιματιστικά του. Γι’ αυτό το σκοπό συγκρίνει τον τωρινό προμηθευτή Α με έναν άλλο προμηθευτή Β, που προσφέρει τιμές 5% φθηνότερες. Μια ποιοτική ανάλυση έδειξε ότι ο κατασκευαστής πρέπει να αλλάξει προμηθευτή μόνο εάν αποδειχθεί ότι οι συσκευές του προμηθευτή Β έχουν μεγαλύτερη διάρκεια ζωής απ’ αυτές του προμηθευτή Α. Για να ελέγξει την υπόθεση αυτή ο κατασκευαστής εξόπλισε 30 αυτοκίνητα μεσαίας κατηγορίας με συσκευές του προμηθευτή Α και 30 αυτοκίνητα με συσκευές του κατασκευαστή Β και κατέγραψε τη συνολική απόσταση(σε χιλιάδες μίλια) που κάλυψε το κάθε αυτοκίνητο μέχρι να παρουσιάσει βλάβη στο συμπυκνωτή.Ο πίνακας με τις αποστάσεις που διένυσαν τα αυτοκίνητα φαίνεται παρακάτω:

Προμηθευτής Α			Προμηθευτής Β		
156	108	152	109	111	129
146	97	97	75	146	117
93	142	122	131	114	102
152	102	105	131	98	90
80	160	110	129	136	118
111	117	114	147	134	88
107	83	112	78	83	115
118	110	106	124	56	103
115	119	83	86	114	106
125	98	125	127	98	111

Στηρίζουν τα δεδομένα με στάθμη σημαντικότητας 5% την απόφαση για αλλαγή προμηθευτή;

Λύση:

Έστω ότι μ_1 ο μέσος του πληθυσμού των αυτοκινήτων για προμηθευτή Α και μ_2 ο μέσος του πληθυσμού των αυτοκινήτων για προμηθευτή Β. Θα συγκρίνουμε τον πληθυσμό των αυτοκινήτων με το νέο συμπυκνωτή με τον πληθυσμό των αυτοκινήτων με τον παλιό συμπυκνωτή, για να δούμε εάν τα αυτοκίνητα με το νέο συμπυκνωτή κατέγραψαν περισσότερα μίλια. Άρα ο έλεγχος θα είναι:

$$H_0: \mu_2-\mu_1 = 0\ , \quad H_1: \mu_2-\mu_1 > 0$$

Τώρα θα δούμε εάν οι διασπορές είναι ίσες ή άνισες για να δούμε εάν θα χρησιμοποιήσουμε τους τύπους για ίσες ή για άνισες, οπότε θα πραγματοποιήσουμε τον εξής έλεγχο:

$$\text{Ho: } \frac{\sigma 1^{\,2}}{\sigma 2^{\,2}} = 1\ , \quad \text{H1: } \frac{\sigma 1^{\,2}}{\sigma 2^{\,2}} \neq 1$$

Η μέση τιμή $\bar{x}_1$ είναι: $\bar{x}_1 = \frac{\sum \mathbf{x1i}}{n1} = \frac{3465}{30} = 115{,}5$

Η μέση τιμή $\bar{x}_2$ είναι: $\bar{x}_2 = \frac{\sum \mathbf{x2i}}{n2} = \frac{3306}{30} = 110{,}5$

$$s1^2 = \frac{\sum x1i^2 - \frac{\sum x1i^2}{n1}}{n1-1} = \frac{13340{,}25 - 444{,}675}{30} = 429{,}85$$

$$s2^2 = \frac{\sum x2i^2 - \frac{\sum x2i^2}{n2}}{n2-1} = \frac{12210{,}25 - 407{,}008}{30} = 393{,}44$$

Η περιοχή απόρριψης είναι:

$F > F_{\alpha/2.v1*v2}$ ή $F < F_{1-\alpha/2.v1*v2}$ $\rightarrow$ $F > F_{0,025.30*30}$ ή $F < F_{1-0,025.30*30}$

$\rightarrow F > 1{,}84$ ή $F < -{,}084$

Ο έλεγχος είναι:

$F = \frac{s1^{\,2}}{s2^{\,2}} = 1{,}09$

Η τιμή του ελέγχου δεν βρίσκεται στην περιοχή απόρριψης, άρα δεν απορρίπτουμε την μηδενική υπόθεση και θα χρησιμοποιήσουμε τους τύπους των ίσων διασπορών.

Επιστρέφουμε στον αρχικό έλεγχο:

$$H_0: \mu_2-\mu_1 = 0, \quad H_1: \mu_2-\mu_1 > 0\ (?)$$

Ο έλεγχος είναι:

$$t = \frac{\overline{(x}1\ -\ \bar{x}2) - (\mu 1 - \mu 2)}{\sqrt{sp^2\left(\frac{1}{n1}+\frac{1}{n2}\right)}}$$

με $v = n_1+n_2-2$ βαθμούς ελευθερίας

Η περιοχή απόρριψης θα είναι:
$t > t_{\alpha.v} = t_{0,05.30} = 1,697$

$\sigma 1^2 = s1^2 = 429,85$
$\sigma 2^2 = s2^2 = 393,44$

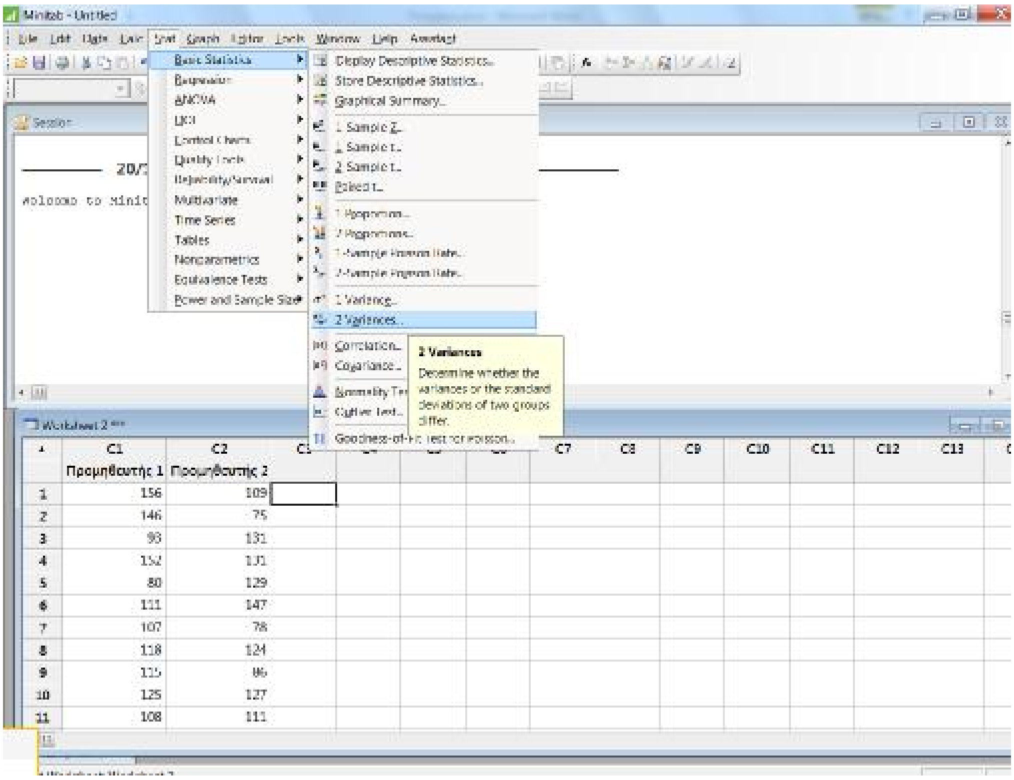
Η σταθμισμένη διασποράς είναι:
$Sp^2 = \frac{(n1-1)\,\sigma 1^{\,2} + (n2-1)\,\sigma 2^{\,2}}{n1+n2-2} = \frac{12465,65 + 11409,76}{58} = 411,645$

Θεωρώντας δεδομένη τη μηδενική υπόθεση έχουμε: $\mu_1=\mu_2$
Άρα: $t = \frac{\overline{(x}1 - \bar{x}2) - (\mu 1 - \mu 2)}{\sqrt{sp^2(\frac{1}{n1}+\frac{1}{n2})}} = \frac{5}{\sqrt{411,645*2/30}} = \frac{5}{\sqrt{411,645*2/30}} = 0,98$

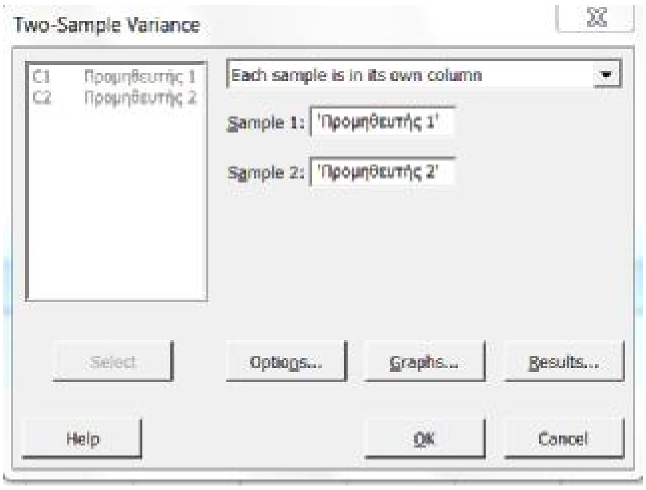
με ν = 30+30-2= 58 βαθμούς ελευθερίας

Ο έλεγχος δε βρίσκεται στην περιοχή απόρριψης, άρα δε μπορούμε να απορρίψουμε τη μηδενική υπόθεση, οπότε δεν μπορεί να στηριχθεί η επιλογή καινούργιου προμηθευτή.

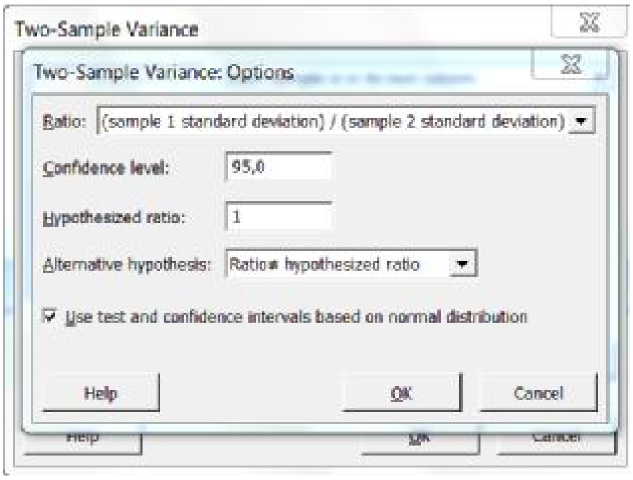
Ας δούμε πως θα λυνόταν στο στατιστικό πρόγραμμα Minitab:
1)Αρχικά θα δημιουργήσουμε δύο μεταβλητές με ονόματα «Προμηθευτής Α» και «Προμηθευτής Β» ,και θα περάσουμε τα δεδομένα τους.
2)Θα ελέγξουμε εάν οι διασπορές είναι ίσες ή άνισες. Για να συγκρίνουμε τις διασπορές πάμε:
Stat → Basic statistics → 2 variances:



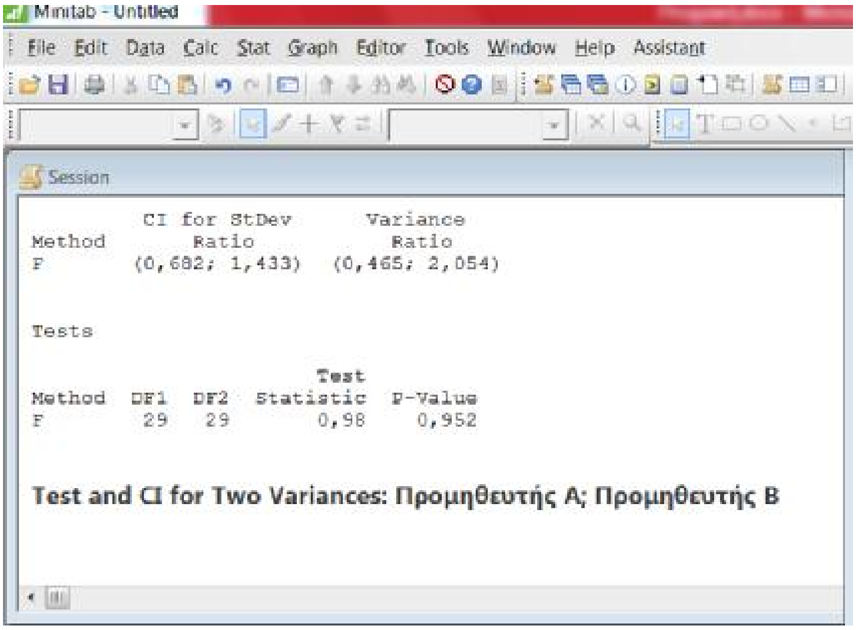
→ Επιλέγω «each sample in its own column» και βάζω όπου sample 1 και sample 2, «Προμηθευτής Α» και «Προμηθευτής Β» αντίστοιχα:



→ options → confidence level 95, καθώς έχω στάθμη σημαντικότητας 5% →hypothesized difference βάζω 1 καθώς εάν οι δύο αναλογίες διαφέρουν ο λόγος τους θα είναι διάφορος του 1, και γι' αυτόν το λόγο επιλέγω alternative hypothesis "ratio ≠ hypothesized ratio. → Τέλος επιλέγω «use test and confidence intervals based on normal distribution", καθώς οι δύο πληθυσμοί ακολουθούν κανονική κατανομή:



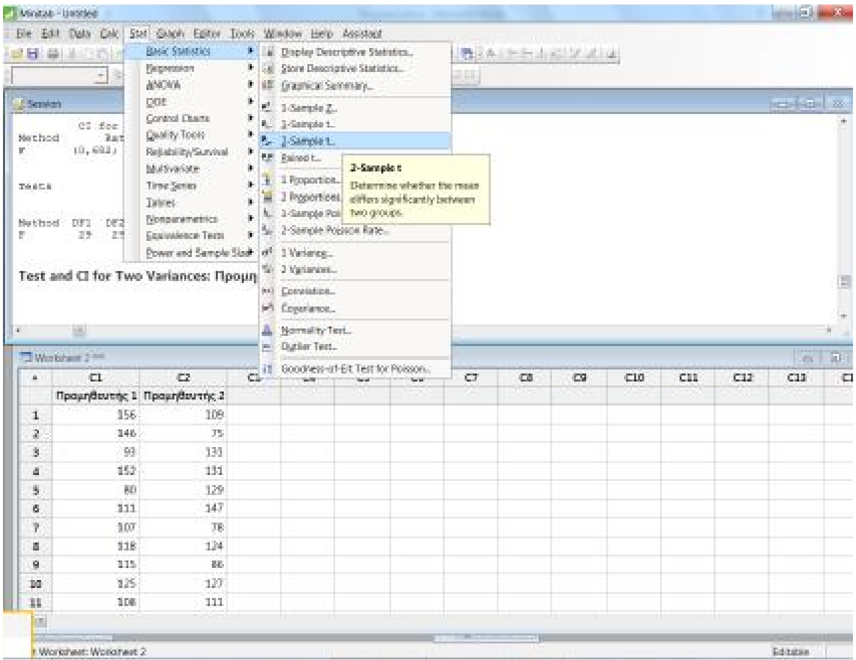
3) Πάμε στο OUTPUT που προκύπτει από τα παραπάνω:



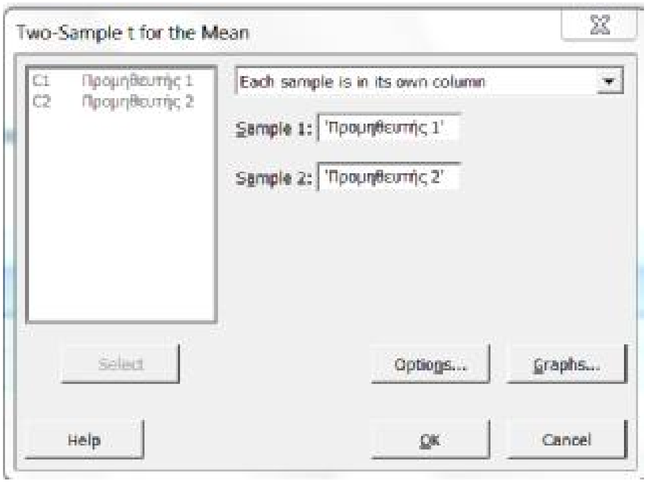
Απ' την τιμή P-Value=0,952 συμπεραίνουμε ότι ο έλεγχος είναι στατιστικά μη σημαντικός, άρα δεν μπορούμε να απορρίψουμε την μηδενική υπόθεση υπέρ της εναλλακτικής. Οι τύποι που θα χρησιμοποιήσουμε στη συνέχεια είναι για ίσες διασπορές.

Εν συνεχεία, θα ελέγξουμε την ισότητα των μέσων των δύο πληθυσμών.

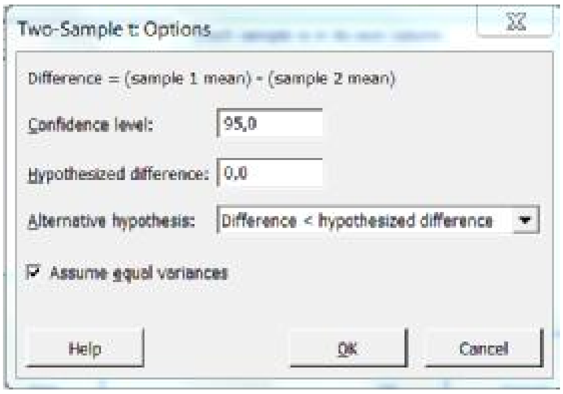
1) Πάμε Stat → Basic statistics → 2-Sample t..:



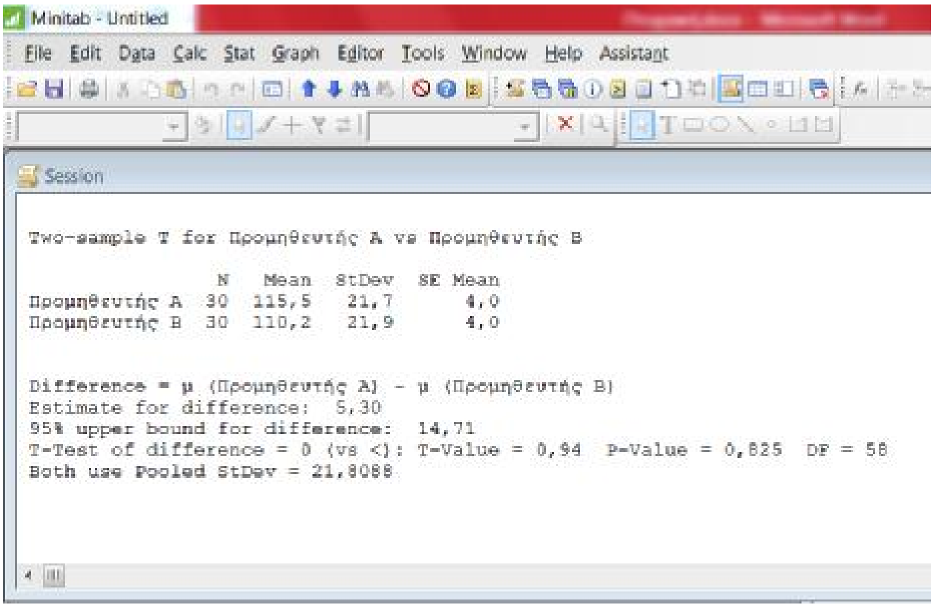
2) →Επιλέγω each sample is in its own column →Περνάω τη μεταβλητή « Προμηθευτής Α » στο κουτί sample 1, και τη μεταβλητή « Προμηθευτής Β » στο κουτί sample 2:



→options → confidence level 95, καθώς έχω στάθμη σημαντικότητας 5% →hypothesized difference βάζω 0, διότι με το 0 θα τα συγκρίνω → alternative hypothesis επιλέγω "difference < hypothesized difference, καθώς εάν σύμφωνα με την εναλλακτική η μέση τιμή του πληθυσμού Β είναι μεγαλύτερη απ' τη μέση τιμή του πληθυσμού Α, μ_1-μ_2 θα πρέπει να μας δώσει αρνητική τιμή → Τέλος επιλέγω «assume equal variances" καθώς πιο πάνω βρήκαμε ότι οι διασπορές είναι ίσες:



2)Το OUTPUTμας εμφανίζει τα εξής:



Η τιμή του ελέγχου είναι 0,94 και η τιμή της P-Value είναι 0,825, άρα ο έλεγχος είναι στατιστικά μη σημαντικός, οπότε δεν μπορούμε να συμπεράνουμε ότι οι συσκευές του προμηθευτή B έχουν μεγαλύτερη διάρκεια ζωής.

6. Σύγκριση δεδομένων κατά ζεύγη:

Ο λόγος συλλογής δεδομένων κατά ζεύγη, είναι να μειωθούν οι εξωγενείς πηγές μεταβλητότητας. Η σύγκριση κατά ζεύγη εμφανίζεται στις εξής περιπτώσεις:

i) Όταν έχουμε μετρήσεις ενός ατόμου σε δύο διαφορετικές χρονικές περιόδους, μεταξύ των οποίων έχουν επέλθει αλλαγές στις μετρήσεις.

ii) Δεδομένα τα οποία είναι ταιριασμένα ένα προς ένα π.χ. α) για φύλο και ηλικία

Γενικά, τα δεδομένα κατά ζεύγη, πάντοτε σχετίζονται μεταξύ τους, οι τιμές των δεδομένων όμως, δεν παίζουν κάποιο ρόλο.
Για να συγκρίνουμε τις τιμές από δύο ομάδες όπου οι παρατηρήσεις είναι κατά ζεύγη ελέγχουμε αν η κατανομή των διαφορών φαίνεται περίπου κανονική. Αν ναι, εφαρμόζουμε **έλεγχο t** κατά ζεύγη.
Ho: η διαφορά των δύο μέσων είναι μηδέν. Αν όχι - εφαρμόζουμε κάποιο μετασχηματισμό σε μια προσπάθεια να μεταποιήσουμε τα δεδομένα ώστε να μπορέσουμε να συνεχίσουμε, ειδάλλως προχωράμε σε μη-παραμετρικές μεθόδους (Π.Χ. Wilcoxon signed ranks test). Ας δούμε μέσα απ’ το επόμενο παράδειγμα πως γίνεται ο στατιστικός έλεγχος:

Παράδειγμα:

Μια μεγάλη επιχείρηση εξετάζει την καθιέρωση ενός προγράμματος φυσικής άσκησης μετά το γεύμα, ώστε να βελτιωθεί η υγεία των εργαζομένων και να μειωθούν οι απουσίες και οι ιατρικές δαπάνες. Έτσι, καθιέρωσε πειραματικά το πρόγραμμα σε ένα τμήμα και κατέγραψε τις ιατρικές δαπάνες (σε χιλιάδες δολάρια) του συγκεκριμένου τμήματος ανά μήνα για ένα χρόνο πριν και ένα χρόνο μετά την εφαρμογή του προγράμματος, όπως φαίνεται στον πίνακα:

	Μήνας											
Δαπάνες	1	2	3	4	5	6	7	8	9	10	11	12
Πριν	68	44	30	58	35	33	52	69	23	69	48	30
Μετά	59	42	20	62	25	30	56	62	25	75	40	26

Μπορούμε από τα δεδομένα να συμπεράνουμε με στάθμη εμπιστοσύνης 95% ότι με την εφαρμογή του νέου προγράμματος μειώθηκαν οι ιατρικές δαπάνες;

Λύση:

Θα ελέγξουμε εάν με την εφαρμογή του νέου προγράμματος μειώθηκαν οι ιατρικές δαπάνες συγκρίνοντας τους μέσους των δύο συνόλων δεδομένων. Έστω ότι μ_1 ο μέσος των δεδομένων απ’ τις δαπάνες πριν την εφαρμογή του προγράμματος, και μ_2 ο μέσος των δεδομένων απ’ τις δαπάνες μετά την εφαρμογή του προγράμματος. Ο έλεγχος υποθέσεων θα είναι:

$$Ho: \mu_1 - \mu_2 = 0, \quad H_1: \mu_1 - \mu_2 > 0$$

Όπως είδαμε προηγουμένως, στον έλεγχο ίσων διασπορών μπορούμε να χρησιμοποιούμε την **σταθμισμένη διασποράς** .

$$Sp^2 = \frac{(n1-1)\,\sigma 1^{\,2} + (n2-1)\,\sigma 2^{\,2}}{n1+n2-2}$$

Ο έλεγχος είναι:

$$t = \frac{(\overline{x}1 - \bar{x}2) - (\mu 1 - \mu 2)}{\sqrt{sp^2(\frac{1}{n1}+\frac{1}{n2})}}$$

Απ' τα δεδομένα έχουμε: $n_1 = n_2 = 12$

Τη μέση τιμή των δύο πληθυσμών θα την βρούμε απ' τους εξής τύπους:

$$\bar{x}_1 = \frac{\sum \mathbf{x1i}}{n1} = \frac{68+44+30+58+35+33+52+69+23+69+48+30}{12} = 46{,}58$$

$$\bar{x}_2 = \frac{\sum \mathbf{x2i}}{n2} = \frac{59+42+20+62+25+30+56+62+25+75+40+26}{12} = 43{,}5$$

Τη διασπορά των δύο πληθυσμών θα την βρούμε ως εξής:

$$s^2{}_1 = \frac{\sum x1i^2 - \frac{\sum x1i^2}{n1}}{n1-1} = \frac{2169{,}7 - 180{,}8}{11} = 180{,}809$$

$$s^2{}_2 = \frac{\sum x2i^2 - \frac{\sum x2i^2}{n2}}{n2-1} = \frac{1892{,}25 - 157{,}687}{11} = 157{,}68$$

Όπως αναφέραμε προηγουμένως, οι έλεγχοι πραγματοποιούνται με βάση τη διασπορά των δειγμάτων, άρα ισχύει : s = σ

Γνωρίζοντας πλέον όλες τις άγνωστες παραμέτρους της εξίσωσης της σταθμισμένης διασποράς, επιλύουμε την εξίσωση:

$$Sp^2 = \frac{(n1-1)\,\sigma 1^{\,2} + (n2-1)\,\sigma 2^{\,2}}{n1+n2-2} = \frac{1988{,}89+1734{,}48}{22} = 169{,}24$$

Με δεδομένη τη μηδενική υπόθεση, έχουμε μ1=μ2.

Άρα ο έλεγχος είναι:

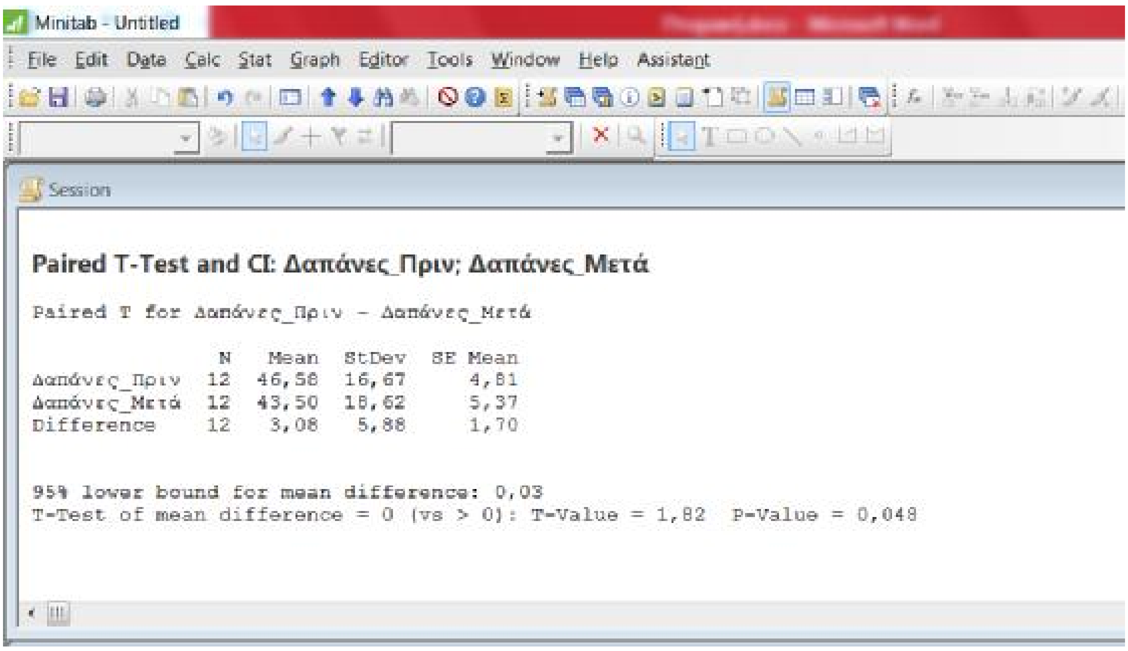
$$t = \frac{(\overline{x}1 - \bar{x}2) - (\mu 1 - \mu 2)}{\sqrt{sp^2(\frac{1}{n1}+\frac{1}{n2})}} = \frac{3{,}08}{\sqrt{\frac{169{,}24}{6}}} = 1.82$$

Η περιοχή απόρριψης είναι: $t > t_{\alpha/2.v} = t_{0.025.12} = 2.179$

Η τιμή t δε βρίσκεται στην περιοχή απόρριψης, άρα μπορούμε να συμπεράνουμε ότι με τη νέα μέθοδο μειώθηκαν οι ιατρικές δαπάνες.

Ας δούμε τώρα πως επιλύεται και στο στατιστικό πρόγραμμα Minitab:

1)Αρχικά θα δημιουργήσουμε δύο μεταβλητές με ονόματα «Δαπάνες_Πριν» και «Δαπάνες_Μετά» ,και θα περάσουμε τα δεδομένα τους
2) Πάμε Stat → Basic statistics → Paired t.. → Επιλέγω each sample is in a column → Περνάω τη μεταβλητή « Δαπάνες_Πριν » στο πρώτο κελί και τη μεταβλητή « Δαπάνες_Μετά » στο δεύτερο → options → confidence level 95, καθώς έχω στάθμη σημαντικότητας 5% → hypothesized difference:0 καθώς η διαφορά των μέσων πρέπει να συγκριθεί με το 0 → επιλέγω “difference > hypothesized difference “ καθώς θέλω να ελέγξω εάν η διαφορά των δύο μέσων ξεπερνάει το 0, δηλαδή εάν ο μέσος του πρώτου δείγματος είναι μεγαλύτερος απ’ το μέσο του δεύτερου.
3) Πάμε στο OUTPUT το οποίο μας εμφανίζει τα εξής:



Η τιμή του ελέγχου είναι 1,82 και η P-Value είναι 0,048, άρα ο έλεγχος είναι στατιστικά σημαντικός και υπάρχει ισχυρή απόδειξη υπέρ της εναλλακτικής υπόθεσης, δηλαδή μειώθηκαν οι ιατρικές δαπάνες.

7. Λόγος δύο διασπορών:

Ας δούμε τώρα, πως μπορούμε να συγκρίνουμε τη μεταβλητότητα δύο πληθυσμών, μέσω της διασποράς τους. Αυτό το κάνουμε συνήθως όταν θέλουμε να συγκρίνουμε τη σταθερότητα δύο διαδικασιών παραγωγής, ή τους επενδυτικούς κινδύνους δύο επενδύσεων
Εδώ θα έχουμε για παράμετρο τον λόγο των διασπορών, διότι η κατανομή δειγματοληψίας υπολογίζει λόγους και όχι διαφορές. Στην περίπτωση μας θα έχουμε για τους δύο πληθυσμούς: $\frac{\sigma 1}{\sigma 2}$. Επειδή όμως έχουμε δει ότι η διασπορά του δείγματος είναι αμερόληπτος εκτιμητής της διασποράς του πληθυσμού, μπορούμε να χρησιμοποιήσουμε το λόγο των διασπορών των δύο δειγμάτων: $\frac{s1}{s2}$
Ο λόγος των δύο μεταβλητών που ακολουθούν την κατανομή χ^2, αν διαιρεθεί με το πλήθος των βαθμών ελευθερίας, ακολουθεί την κατανομή F με βαθμούς ελευθερίας ίσους με τους βαθμούς που έχουν οι δύο κατανομές χ^2. Πριν είδαμε ότι εάν ο πληθυσμός ακολουθεί κανονική κατανομή, η μεταβλητή $\frac{(n-1)\, s^2}{\sigma^2}$ ακολουθεί κατανομή χ^2.
Στην περίπτωση μας, όπου έχουμε 2 πληθυσμούς με κανονική κατανομή, οι δύο μεταβλητές τους θα είναι :
$\frac{(n1-1)s1^2}{\sigma 1^2}$ και $\frac{(n2-1)s2^2}{\sigma 2^2}$

Τώρα εάν διαιρέσουμε κάθε μεταβλητή με τους βαθμούς ελευθερίας της, θα έχουμε: $\frac{\frac{(n1-1)s1^2/\sigma 1^2}{(n1-1)}}{\frac{(n2-1)s2^2/\sigma 2^2}{(n2-1)}}$.

,το οποίο απλοποιείται σε: $\boldsymbol{\frac{s1^2/\sigma 1^2}{s2^2/\sigma 2^2}}$.

Ο λόγος αυτός ακολουθεί την **κατανομή F** με $v_1=n_1-1$ βαθμούς ελευθερίας για τον αριθμητή και $v_2=n_2-1$ βαθμούς ελευθερίας για τον παρονομαστή.

Έλεγχος υποθέσεων: Όταν έχουμε να κάνουμε με λόγο διασπορών, η μηδενική υπόθεση θα είναι πάντα, οι διασπορές να είναι ίσες. Οπότε ο λόγος τους θα πρέπει να ισούται με τη μονάδα.

$$H_0: \frac{\sigma 1^2}{\sigma 2^2} = 1 \qquad H_1: \frac{\sigma 1^2}{\sigma 2^2} \neq 1$$

Ο έλεγχος όπως βρήκαμε πιο πάνω είναι **F** = $\frac{s1^2/\sigma 1^2}{s2^2/\sigma 2^2}$, το οποίο με δεδομένη τη μηδενική υπόθεση γίνεται:

F = $\frac{s1^2}{s2^2}$

Ακολουθώντας την κατανομή F με $v_1=n_1-1$ βαθμούς ελευθερίας για τον αριθμητή και $v_2=n_2-1$ βαθμούς ελευθερίας για τον παρονομαστή.

Η περιοχή απόρριψης είναι:
$F > F_{\alpha/2.\nu1.\nu2}$

Εκτιμητής διαστήματος εμπιστοσύνης:

Ο εκτιμητής διαστήματος εμπιστοσύνης για το λόγο δύο διασπορών είναι:
LCL= $(\frac{s1^2}{s2^2})\frac{1}{F\alpha/2.\nu1.\nu2}$ και UCL= $(\frac{s1^2}{s2^2})\ F_{\alpha/2.\nu1.\nu2}$

Ας δούμε τώρα πως εφαρμόζονται όλα τα παραπάνω για το λόγο δύο διασπορών μέσω του επόμενου παραδείγματος:

<u>Παράδειγμα:</u>

Ο διευθυντής παραγωγής ενός εργοστασίου αντιμετωπίζει προβλήματα σε μια γραμμή παραγωγής, καθώς συχνά δημιουργούνται καθυστερήσεις και μποτιλιαρίσματα. Για το λόγο αυτό σχεδίασε δύο νέες μεθόδους αλλάζοντας τη σειρά των δραστηριοτήτων της γραμμής παραγωγής, και πραγματοποίησε ένα πείραμα καταγράφοντας τους χρόνους ολοκλήρωσης (σε δευτερόλεπτα) της γραμμής παραγωγής με τις δύο μεθόδους, όπως φαίνεται στον παρακάτω πίνακα:

Μέθοδος	Χρόνος				
	8,8	9,6	8,4	9	8,3
1	9,2	9	8,7	8,5	9,4
	9,2	9,4	8,9	9,6	9,7
2	8,4	8,8	8,9	9	9,7

Μπορούμε από τα δεδομένα για επίπεδο αξιοπιστίας 5% να συμπεράνουμε ότι η δεύτερη μέθοδος έχει μικρότερη διασπορά από την πρώτη;

<u>Λύση:</u>

Εάν η δεύτερη μέθοδος έχει μικρότερη διασπορά από την πρώτη, ο λόγος των δύο διασπορών στην εναλλακτική υπόθεση θα πρέπει να είναι μεγαλύτερος της μονάδας. Άρα θα έχουμε:

$$H_0: \frac{\sigma1^2}{\sigma2^2}=1 \quad H_1: \frac{\sigma1^2}{\sigma2^2}>1$$

Για να βρούμε τις δύο διασπορές, αρχικά θα υπολογίσουμε το μέσο των δύο πληθυσμών:

$$\bar{x}_1=\frac{\sum \mathbf{x1i}}{n1}=\frac{88,9}{10}=8,89$$

$$\bar{x}_2=\frac{\sum \mathbf{x2i}}{n2}=\frac{91,6}{10}=9,16$$

Στη συνέχεια θα βρούμε τις διασπορές των δύο δειγμάτων, καθώς είναι αμέριστες εκτιμήτριες των διασπορών των δύο πληθυσμών:

$s^2{}_1 = \frac{\sum x1i^2 - \frac{\sum x1i^2}{n1}}{n1-1}$ = 7903,21 – 790,321 = 7112,889

$s^2{}_2 = \frac{\sum x2i^2 - \frac{\sum x2i^2}{n2}}{n2-1}$ = 8390,56 – 839,056 = 7551,504

Οπότε ο έλεγχος θα είναι:

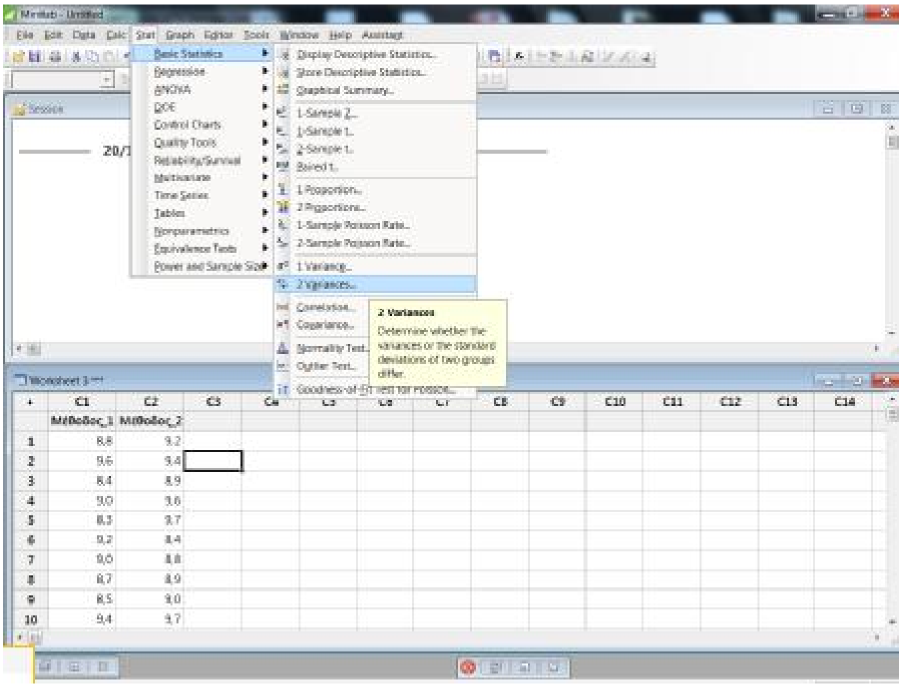
F = $\frac{s1^2}{s2^2} = \frac{7112,889}{7551,504}$ = 0,98

με ν1=9 και ν2=9 βαθμούς ελευθερίας.

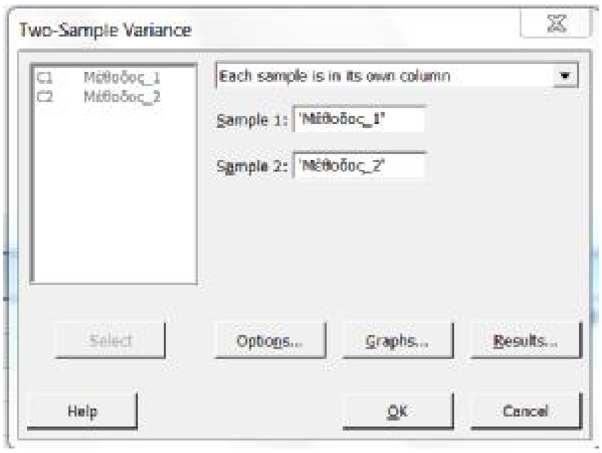
Η περιοχή απόρριψης είναι: F > $F_{\alpha/2.v1.v2} = F_{0,025.10.10} =$ 4,03

Ο έλεγχος δε βρίσκεται στην περιοχή απόρριψης, άρα δεν μπορούμε να απορρίψουμε τη μηδενική υπόθεση και να συμπεράνουμε ότι η διασπορά της δεύτερης μεθόδου είναι μικρότερη από τη διασπορά της πρώτης.

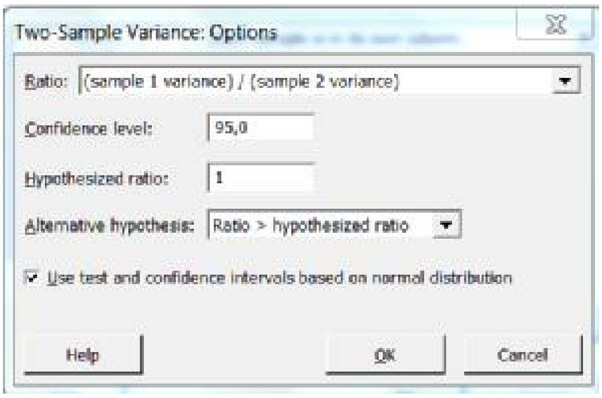
Ας δούμε τώρα πως λύνεται το παραπάνω πρόβλημα στο στατιστικό πρόγραμμα Minitab:
1)Αρχικά θα δημιουργήσουμε δύο μεταβλητές με ονόματα «Μέθοδος_1» και «Μέθοδος_2» ,και θα περάσουμε τα δεδομένα τους.
2) Θα κάνουμε τον έλεγχο υποθέσεων για το λόγο των διασπορών: Πάμε Stat → Basic statistics → 2 Variances:



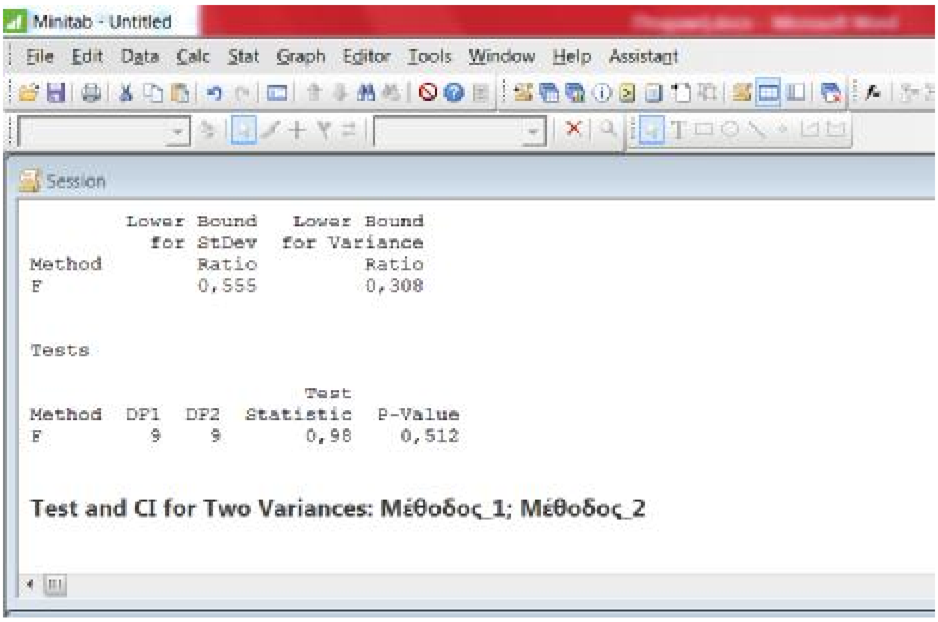
→ Each sample in its own column → Περνάω τη μεταβλητή «Μέθοδος_1» στο κελί sample 1 και τη μεταβλητή «Μέθοδος_2» στο κελί sample 2:



→ options → ratio: sample 1/ sample 2 , confidence level: 95% , hypothesized ratio:1 , alternative hypothesis: ratio > hypothesized ratio. Τέλος επιλέγω use test and confidence intervals based on normal distribution:



Το OUTPUT μας εμφανίζει τα εξής:



Η τιμή της p-value δεν είναι ισχυρή, άρα δεν υπάρχει καμία απόδειξη υπέρ της εναλλακτικής υπόθεσης.

Ας δούμε τώρα πως θα μπορούσαμε να υπολογίσουμε έναν **εκτιμητή διαστήματος εμπιστοσύνης** για το παραπάνω πρόβλημα:
Έστω ότι θέλουμε να εκτιμήσουμε με στάθμη εμπιστοσύνης 95% τον λόγο των διασπορών του παραδείγματος:

Λύση:

Θα υπολογίσουμε τα δύο άκρα του διαστήματος εμπιστοσύνης:

LCL= $(\frac{s1^2}{s2^2})\ \frac{1}{F\alpha/2.v1.v2} = (\frac{7112,889}{7551,504})\ \frac{1}{4,03}$ = 0,233

UCL= $(\frac{s1^2}{s2^2})\ F_{\alpha/2.v1.v2} = (\frac{7112,889}{7551,504})$ 4,03 = 3,788

Άρα ο λόγων των διασπορών των δύο πληθυσμών βρίσκεται μεταξύ 0,233 και 3,788.

8. Έλεγχος ποσοστού για δύο δείγματα:

Πάμε στην περίπτωση που τα δεδομένα των δύο πληθυσμών είναι ονομαστικά. Όταν έχουμε ονομαστικά δεδομένα, υπολογίζουμε πόσες φορές εμφανίζονται στον πληθυσμό, δηλαδή τη **συχνότητα** τους, και την αναλογία τους σε σχέση με τον πληθυσμό, δηλαδή τη **σχετική συχνότητα** τους.
Έστω ότι x_1 ο αριθμός των εμφανίσεων ενός δεδομένου στο δείγμα n_1 ενός πληθυσμού 1 και x_2 ο αριθμός των εμφανίσεων ενός δεδομένου στο δείγμα n_2 ενός πληθυσμού 2. Η **αναλογία** για τον πληθυσμό 1 θα είναι p_1 και για το δείγμα του πληθυσμού $\widehat{p1}$. Αντίστοιχα, η αναλογία για τον πληθυσμό 2 θα είναι p_2 και για το δείγμα του πληθυσμού $\widehat{p2}$. Οι τύποι των αναλογιών των δύο δειγμάτων είναι οι εξής:
$\widehat{p1} = \frac{x1}{n1}$, $\widehat{p2} = \frac{x2}{n2}$

Η διαφορά $\widehat{p1} - \widehat{p2}$ είναι αμερόληπτος εκτιμητής της διαφοράς $p_1 - p_2$.
Για να χρησιμοποιήσουμε τον έλεγχο $\widehat{p1} - \widehat{p2}$, ο οποίος ακολουθεί κατά προσέγγιση κανονική κατανομή είναι απαραίτητο το μέγεθος των δειγμάτων να είναι αρκετά μεγάλο, ώστε να ισχύει:
$n_1 * p_1 \geq 5$, $n_1 * (1 - p_1) \geq 5$
$n_2 * p_2 \geq 5$, $n_2 * (1 - p_2) \geq 5$
Καθώς οι αναλογίες των πληθυσμών είναι άγνωστες, και οι αναλογίες των δειγμάτων είναι αμερόληπτες εκτιμήτριες τους, θα χρησιμοποιούμε αυτές στις παραπάνω ανισότητες:
$n_1 * \widehat{p1} \geq 5$, $n_1 * (1 - \widehat{p1}) \geq 5$
$n_2 * \widehat{p2} \geq 5$, $n_2 * (1 - \widehat{p2}) \geq 5$

Το τυπικό σφάλμα στην περίπτωση αυτή είναι:
$\sigma\widehat{p1} - \widehat{p2} = \sqrt{\frac{p1*(1-p1)}{n1} + \frac{p2*(1-p2)}{n2}}$

Το πρόβλημα είναι ότι οι αναλογίες των δύο πληθυσμών είναι άγνωστες, άρα δεν μπορούμε να επιλύσουμε τον συγκεκριμένο τύπο. Καθώς όμως τα δείγματα είναι αμερόληπτοι εκτιμητές των πληθυσμών, θα βρούμε το τυπικό σφάλμα μέσω αυτών. Το πως θα επιλυθεί εξαρτάται απ' τη μορφή της μηδενικής υπόθεσης του ελέγχου υποθέσεων, η οποία μπορεί να είναι είτε p_1-p_2=0 είτε p_1-$p_2 \neq 0$.
Εάν είναι p_1-p_2=0 , οι δύο αναλογίες είναι ίσες και μπορούμε να χρησιμοποιήσουμε την **εκτιμήτρια αναλογίας:**

Εκτιμήτρια αναλογίας

$$\hat{p} = \frac{\mathrm{x1}+x2}{\mathrm{n1+n2}}$$

Οπότε το τυπικό σφάλμα γίνεται:

$$\sigma\widehat{p1} - \widehat{p2} = \sqrt{\hat{p} * (1-\hat{p})(\frac{1}{\mathrm{n1}} + \frac{1}{\mathrm{n2}})}$$

Ας δούμε πως γίνεται και ο έλεγχος σε αυτές τις δύο περιπτώσεις:
Για Ho: p_1-p_2=0, ο έλεγχος είναι:

$$Z = \frac{(\widehat{p1}-\widehat{p2})-(\mathrm{p1-p2})}{\sqrt{\hat{p}*(1-\hat{p})(\frac{1}{\mathrm{n1}}+\frac{1}{\mathrm{n2}})}}$$

Απ' τη μηδενική υπόθεση έχουμε p_1 - p_2 =0. Άρα:

$$Z = \frac{(\widehat{p1}-\widehat{p2})}{\sqrt{\hat{p}*(1-\hat{p})(\frac{1}{\mathrm{n1}}+\frac{1}{\mathrm{n2}})}}$$

Όταν p_1-$p_2 \neq 0$ για τη μηδενική υπόθεση, ο έλεγχος παραμένει ως έχει, και δεν μπορεί να επιλυθεί καθώς οι αναλογίες p_1 και p_2 είναι άγνωστες.

$$Z = \frac{(\widehat{p1}-\widehat{p2})-(\mathrm{p1-p2})}{\sqrt{\hat{p}*(1-\hat{p})(\frac{1}{\mathrm{n1}}+\frac{1}{\mathrm{n2}})}}$$

Γενικά αυτό δεν θα αποτελέσει πρόβλημα, καθώς συνήθως μας ενδιαφέρει εάν οι αναλογίες των δύο πληθυσμών διαφέρουν, κάτι που απαντιέται μέσω της πρώτης περίπτωσης. Η περίπτωση που οι αναλογίες διαφέρουν, μας ενδιαφέρει όταν θέλουμε να μάθουμε εάν η διαφορά τους ξεπερνά κάποιο μέγεθος. Σε αυτή την περίπτωση, θα χρησιμοποιήσουμε το **διάστημα εμπιστοσύνης** των δειγμάτων τους:

$$(\widehat{p1}-\widehat{p2}) \pm z\alpha/2\sqrt{\frac{\widehat{p1}*(1-\widehat{p1})}{n1} + \frac{\widehat{p2}*(1-\widehat{p2})}{n2}}$$

Με απαραίτητη προϋπόθεση να ισχύουν τα εξής:
$n_1 * \widehat{p1} \geq 5$, $n_1 * (1-\widehat{p1}) \geq 5$

$n_2 * \widehat{p2} \geq 5$, $n_2 * (1 - \widehat{p2}) \geq 5$

Ας δούμε τώρα όμως την εφαρμογή όλων των παραπάνω μέσα απ' το επόμενο πρόβλημα:

Παράδειγμα:

Σύμφωνα με την Καναδική Αντικαρκινική Εταιρία, κάθε χρόνο εμφανίζουν καρκίνο του στήθους 21.000 γυναίκες, από τις οποίες οι 5.000 πεθαίνουν (οι αντίστοιχοι αριθμοί για τις ΗΠΑ είναι τουλάχιστον 10-πλάσιοι). Η βασικότερη μέθοδος αντιμετώπισης είναι η χειρουργική, αλλά σε πολλές γυναίκες ο καρκίνος επανεμφανίζεται. Για το λόγο αυτό, μετά την επέμβαση χορηγείται ένα φάρμακο που ονομάζεται Tamoxifen, αλλά έχει διαπιστωθεί ότι σε μια πενταετία οι όγκοι αναπτύσσουν ανθεκτικότητα και στο συγκεκριμένο φάρμακο. Έτσι αναπτύχθηκε ένα νέο φάρμακό, που ονομάζεται Letrozole, για τη δοκιμή του οποίου επιλέχθηκε ένα δείγμα 5.188 γυναικών μετά από χειρουργική αφαίρεση καρκίνου στο στήθος. Στις μισές χορηγήθηκε το νέο φάρμακο και στις μισές το ψευδοφάρμακο. Η έρευνα προοριζόταν να διαρκέσει 5 χρόνια, όμως μετά από 2,5 χρόνια διαπιστώθηκε ότι 75 από τις γυναίκες που λάμβανα το νέο φάρμακο και 132 απ' αυτές που λάμβαναν το ψευδοφάρμακο είχαν επανεμφάνιση του καρκίνου.
Μπορούμε από τα δεδομένα να συμπεράνουμε με στάθμη σημαντικότητας 1% ότι το νέο φάρμακο είναι αποτελεσματικό;

Λύση:

Θέλουμε να δούμε εάν το νέο φάρμακο είναι πιο αποτελεσματικό. Εάν είναι, η αναλογία του νέου φαρμάκου θα είναι μικρότερη απ' την αναλογία του ψευδοφαρμάκου, καθώς εφόσον είναι πιο αποτελεσματικό, ο αριθμός των γυναικών που το έλαβαν και εμφάνισαν ξανά καρκίνο, θα είναι μικρότερος απ' τον αντίστοιχο αριθμό που έλαβαν το ψευδοφάρμακο. Άρα η εναλλακτική υπόθεση του ελέγχου υποθέσεων θα είναι:

$$H_1 = p_1 - p_2 < 0.$$

Η μηδενική υπόθεση θα είναι:

$$H_0 = p_1 - p_2 = 0.$$

Ο έλεγχος είναι ο εξής:

$$Z = \frac{(\widehat{p1} - \widehat{p2}) - (p1 - p2)}{\sqrt{\hat{p} * (1 - \hat{p}) (\frac{1}{n1} + \frac{1}{n2})}}$$

Απ' τη μηδενική υπόθεση έχουμε $p_1 - p_2 = 0$, άρα ο έλεγχος γίνεται:

$$Z = \frac{(\widehat{p1} - \widehat{p2})}{\sqrt{\hat{p} * (1 - \hat{p})(\frac{1}{\mathrm{n1}} + \frac{1}{\mathrm{n2}})}}$$

Τα δεδομένα μας είναι τα εξής:

$n_1 = n_2 = 2594$
$x_1 = 75$
$x_2 = 132$

Άρα οι αναλογίες των δύο δειγμάτων είναι:

$$\widehat{p1} = \frac{75}{2594} = 0{,}0289$$

$$\widehat{p2} = \frac{132}{2594} = 0{,}050$$

Η εκτιμήτρια αναλογίας είναι:

$$\hat{p} = \frac{\mathrm{x1} + x2}{\mathrm{n1} + \mathrm{n2}} = \frac{207}{5188} = 0{,}0398$$

Με όλα τα απαραίτητα δεδομένα γνωστά, επιστρέφουμε στον έλεγχο:

$$Z = \frac{(\widehat{p1} - \widehat{p2})}{\sqrt{\hat{p} * (1 - \hat{p})(\frac{1}{\mathrm{n1}} + \frac{1}{\mathrm{n2}})}} = \frac{-0{,}0211}{\sqrt{0{,}0398(0{,}6)(7{,}71)}} = \frac{-0{,}0211}{0{,}429} = -0{,}0491$$

Η περιοχή απόρριψης θα είναι: $z > z_{\alpha} = z_{0,01} = 2{,}575$

Η τιμή του ελέγχου δεν βρίσκεται στην περιοχή απόρριψης, οπότε μπορούμε να συμπεράνουμε με στάθμη σημαντικότητας 1% ότι το νέο φάρμακο είναι αποτελεσματικό.

Κεφάλαιο Χ: Ανάλυση διασποράς

Πριν να εξηγήσουμε πώς λειτουργεί η ANOVA, αξίζει να αναφέρουμε γιατί δεν εκτελούμε απλά αρκετούς t-test ελέγχους για να συγκρίνουμε όλους τους συνδυασμούς ομάδων. Φανταστείτε τρεις πειραματικές συνθήκες, όπου μας ενδιαφέρουν οι διαφορές μεταξύ αυτών. Εάν κάναμε δοκιμασίες t-test σε κάθε ζευγάρι ομάδων, τότε θα είχαμε να πραγματοποιήσουμε τρεις ξεχωριστές δοκιμές: μία για να συγκρίνουμε τις ομάδες 1 και 2, μία για να συγκρίνουμε τις ομάδες 1 και 3 και μία για συγκρίνουμε τις ομάδες 2 και 3. Εάν κάθε μία από αυτές τις δοκιμές έχει επίπεδο σημαντικότητας 0,05 τότε για κάθε δοκιμή έχει σημασία η πιθανότητα ψευδούς απόρριψης της μηδενικής υπόθεσης (γνωστή ως σφάλμα τύπου Ι) η οποία είναι μόνο 5%. Επομένως, η πιθανότητα για κανένα σφάλμα Τύπου Ι δεν είναι 0,95 (95%) για κάθε δοκιμή. Αν υποθέσουμε ότι κάθε δοκιμή είναι ανεξάρτητη (ως εκ τούτου, μπορούμε να πολλαπλασιάσουμε τις πιθανότητες) τότε η συνολική πιθανότητα χωρίς σφάλματα Τύπου Ι είναι (0,95) × 3 = 0,95 × 0,95 × 0,95 = 0,857, επειδή η πιθανότητα χωρίς σφάλματα Τύπου Ι είναι 0,95 για κάθε δοκιμή και υπάρχουν τρεις δοκιμές. Δεδομένου ότι η πιθανότητα να μην υπάρχουν σφάλματα Τύπου Ι είναι .857, τότε μπορούμε να υπολογίσουμε την πιθανότητα κάνοντας τουλάχιστον ένα σφάλμα Τύπου Ι αφαιρώντας αυτόν από τον αριθμό 1 (θυμηθείτε ότι η μέγιστη πιθανότητα οποιουδήποτε συμβάντος είναι 1). Έτσι, η πιθανότητα τουλάχιστον ενός σφάλματος Τύπου Ι είναι: 1 – 0,857 = 0,143 ή 14,3%. Ως εκ τούτου, η πιθανότητα εμφάνισης σφάλματος Τύπου Ι αυξήθηκε από 5% σε 14,3%, μια τιμή μεγαλύτερη από το κριτήριο που αποδέχονται οι κοινωνικοί επιστήμονες. Ένα πείραμα με τρεις συνθήκες είναι σχετικά απλό. Αν αυξάναμε τον αριθμό των ομάδων σύγκρισης οι έλεγχοι t-test θα ήταν περισσότεροι και το σφάλμα Τύπου Ι θα ήταν μεγαλύτερο ως ποσοστό.

1. Ανάλυση διασποράς σε ανεξάρτητα δείγματα:

1) Όταν εκτελούμε ένα t-test, ελέγχουμε την υπόθεση ότι τα δύο δείγματα έχουν τον ίδιο μέσο όρο. Ομοίως, η ANOVA μας λέει εάν τρία ή περισσότερα μέσα είναι τα ίδια, επομένως ελέγχει την μηδενική υπόθεση ότι όλα τα μέσα της ομάδας είναι ίσα. Μια ANOVA παράγει μια στατιστική F ή μια αναλογία F, F είναι ο έλεγχος για την σύγκριση των μέσων. Το ANOVA είναι ένα omnibus test, που σημαίνει ότι είναι ένα συνολικό πειραματικό αποτέλεσμα: έτσι, υπάρχουν πράγματα που η ANOVA δεν μπορεί να μας πει. Αν και η ANOVA μας λέει αν ένας πειραματικός χειρισμός ήταν γενικά επιτυχής δεν παρέχει συγκεκριμένες πληροφορίες σχετικά με τις ομάδες που επηρεάστηκαν. Υποθέτοντας ότι το πείραμα διεξήχθη με τρεις διαφορετικές ομάδες, ο λόγος F μας λέει πως τα μέσα αυτών των τριών δειγμάτων δεν είναι ίσα (δηλ. ότι το X1 = X2 = X3 δεν είναι αληθές). Ωστόσο, υπάρχουν διάφοροι τρόποι με τους οποίους οι μέσοι των δειγμάτων μπορούν να διαφέρουν. Η πρώτη δυνατότητα είναι ότι και τα τρία μέσα δειγματοληψίας είναι σημαντικά διαφορετικά (X1 ≠ X2 ≠X3). Μια δεύτερη πιθανότητα είναι ότι τα μέσα της ομάδας 1 και 2 είναι τα ίδια, αλλά η ομάδα 3 έχει διαφορετικό μέσο και από τις δύο άλλες ομάδες (X1 = X2 ≠ X3). Μια άλλη πιθανότητα είναι ότι οι ομάδες 2 και 3 έχουν παρόμοιο μέσο, αλλά η ομάδα 1 έχει διαφορετικό μέσο (X1 ≠ X2 = X3). Τέλος, οι ομάδες 1 και 3 θα μπορούσαν να έχουν

παρόμοιο μέσο, αλλά η ομάδα 2 έχει διαφορετικό μέσο και από τις δύο ($X1 = X3 \neq X2$). Έτσι, σε ένα πείραμα, ο λόγος F μας λέει μόνο ότι ο πειραματικός χειρισμός είχε κάποια επίδραση, αλλά δεν μας λέει συγκεκριμένα το αποτέλεσμα.

2) Αντίθετα, στην περίπτωση που έχουμε να συγκρίνουμε δύο πληθυσμούς, μπορούμε αντί για τον έλεγχο της διαφοράς να εκτελέσουμε μια ανάλυση διασποράς.

Στην ανάλυση διασποράς οι δύο υποθέσεις είναι:

$$H_0: \mu_1=\mu_2$$

$$H_1: \mu_1\neq\mu_2 \text{ (δύο τουλάχιστον μέσοι διαφέρουν)}$$

Άρα η ανάλυση διασποράς μπορεί να ελέγξει μόνο αν οι μέσοι είναι άνισοι (είδαμε ότι ο έλεγχος της διαφοράς δύο μέσων μπορεί <u>και</u> να ελέγξει: $\mu_1>\mu_2$ ή $\mu_1<\mu_2$) και προϋποθέτει ότι οι δύο πληθυσμοί έχουν ίσες διασπορές.

<u>ΠΑΡΑΔΕΙΓΜΑ</u>

Το Viagra είναι ένα σεξουαλικό διεγερτικό (χρησιμοποιείται για τη θεραπεία της ανικανότητας) που έσπασε στη μαύρη αγορά με την πεποίθηση ότι αυτό θα κάνει κάποιον έναν καλύτερο εραστή. Ας υποθέσουμε ότι δοκιμάσαμε αυτή την πεποίθηση με τη λήψη τριών ομάδων συμμετεχόντων και τη χορήγηση μιας ομάδας με ένα εικονικό φάρμακο (placebo) (όπως ένα χάπι ζάχαρης), μία ομάδα με χαμηλή δόση Viagra και μία με υψηλή δόση. Η εξαρτημένη μεταβλητή είναι μία αντικειμενική μέτρηση της λίμπιντο (μετρήθηκε για μία εβδομάδα). Τα δεδομένα είναι τα εξής:

	Placebo	Μικρή δόση	Μεγάλη δόση
	3	5	7
	2	2	4
	1	4	5
	1	2	3
	4	3	6
$\bar{x}$	2,2	3,2	5
s	1,3	1,3	1,58
s^2	1,7	1,7	2,5

Grand Mean = 3.467
Grand SD = 1.767
Grand Variance = 3.124

Για να βρούμε το συνολικό ποσό της διακύμανσης στα δεδομένα μας, υπολογίζουμε τη διαφορά μεταξύ κάθε παρατηρούμενου σημείου δεδομένων ή αλλιώς κάθε δείγματος με τον μεγάλο μέσο ή τον μέσο των πληθυσμών. Στη συνέχεια, υψώνουμε στο τετράγωνο αυτές τις διαφορές και τις προσθέτουμε όλες μαζί για να μας δώσουν το συνολικό άθροισμα τετραγώνων (SST_{otal}):
Επίσης, παρατηρούμε ότι η διακύμανση και τα ποσά των τετραγώνων συσχετίζονται έτσι ώστε η διακύμανση: $S^2 = SS / (N - 1)$, όπου N είναι ο αριθμός των παρατηρήσεων. Επομένως, μπορούμε να υπολογίσουμε το σύνολο των τετραγώνων από τη διακύμανση όλων των παρατηρήσεων (η μεγάλη διακύμανση) με την αναδιάταξη της σχέση :$SS = s^2(N-1)$, n-1 βαθμούς ελευθερίας. Η μεγάλη διακύμανση είναι η διαφορά μεταξύ όλων των δειγμάτων, ανεξάρτητα από την πειραματική κατάσταση από την οποία προέρχονται τα αποτελέσματα.
Επομένως, το SST υπολογίζεται ως εξής:
$SST_{otal} = s^2_{grand}\,(n - 1) = 3{,}124(15 - 1) = 3{,}124 \times 14 = 43{,}74$
Μέχρι τώρα, γνωρίζουμε ότι το συνολικό ποσό της διασποράς εντός των δεδομένων είναι 43,74 μονάδες. Με την ANOVA το μοντέλο βασίζεται στις διαφορές μεταξύ των μέσων των ομάδων και έτσι τα ποσά του μοντέλου των τετραγώνων μας λένε πόσο από το σύνολο της διακύμανσης μπορεί να εξηγηθεί από το γεγονός ότι διαφορετικά σημεία δεδομένων προέρχονται από διαφορετικές ομάδες.

Είδαμε ότι το μοντέλο άθροισμα των τετραγώνων υπολογίζεται λαμβάνοντας τη διαφορά μεταξύ των τιμών που προβλέπονται από το μοντέλο και του μεγάλου μέσου όρου (βλ. Σχήμα 7.4). Στην ANOVA, οι τιμές που προβλέπονται από το μοντέλο είναι τα μέσα των ομάδων. Για τον κάθε συμμετέχοντα η τιμή που προβλέπεται από το μοντέλο είναι ο μέσος όρος για την ομάδα στην οποία το ο συμμετέχων ανήκει. Στο παράδειγμα μας, η προβλεπόμενη τιμή για τους πέντε συμμετέχοντες στην ομάδα του εικονικού φαρμάκου θα είναι 2,2, για τους πέντε συμμετέχοντες στην κατάσταση χαμηλής δόσης που θα είναι 3,2 και για τους πέντε συμμετέχοντες στην κατάσταση υψηλής δόσης θα είναι 5. Το μοντέλο άθροισμα των τετραγώνων απαιτεί να υπολογίσουμε τις διαφορές μεταξύ της προβλεπόμενης τιμής κάθε συμμετέχοντα και του μεγάλου μέσου. Αυτές οι διαφορές στη συνέχεια τετραγωνίζονται και προστίθενται μαζί. Γνωρίζουμε ότι η προβλεπόμενη τιμή για τους συμμετέχοντες σε μία συγκεκριμένη ομάδα είναι ο μέσος όρος αυτής της ομάδας. Επομένως, ο ευκολότερος τρόπος υπολογισμού του SSM είναι:

1) Να υπολογίζουμε τη διαφορά μεταξύ του μέσου όρου κάθε ομάδας και του μεγάλου μέσου όρου.
2) Να τετραγωνίσουμε καθεμία από αυτές τις διαφορές.
3) Να πολλαπλασιάσουμε κάθε αποτέλεσμα με τον αριθμό των συμμετεχόντων στην ομάδα (n_k).
4) Να προσθέσουμε τις τιμές που βρήκαμε για κάθε ομάδα.

Η μαθηματική έκφραση αυτής της διαδικασίας είναι:
$SSM = \sum nk(\overline{x}_k - \overline{x}_{grant})^2$
Χρησιμοποιώντας τα μέσα από τα δεδομένα του παραδείγματός μας, μπορούμε να υπολογίσουμε το SSM ως εξής:
$SSM = 5(2{,}200 - 3{,}467)^2 + 5(3{,}200 - 3{,}467)^2 + 5(5{,}000 - 3{,}467)^2 = 5(-1{,}267)^2 + 5(-0{,}267)^2 + 5(1{,}533)^2 = 8{,}025 + 0{,}355 + 11{,}755 = 20{,}135 \approx 20{,}14$

Για την SSM, οι βαθμοί ελευθερίας θα είναι πάντα μικρότεροι από τον αριθμό των παραμέτρων που υπολογίζουμε. Με λίγα λόγια, αυτή η τιμή θα είναι ο αριθμός των ομάδων μείον ένα (το οποίο το βλέπω ως k - 1). Έτσι, στην περίπτωση των τριών ομάδων οι βαθμοί ελευθερίας θα είναι πάντα 2 (επειδή ο υπολογισμός των αθροισμάτων των τετραγώνων βασίζεται στον μέσο των ομάδων, δύο από τα οποία θα είναι ελεύθερα να μεταβάλλουν τον πληθυσμό αν ο τρίτος κρατηθεί σταθερός).

Μεταβλητότητα εντός των πληθυσμών

Γνωρίζουμε τώρα ότι υπάρχουν 43,74 μονάδες διασποράς που πρέπει να εξηγηθούν στα δεδομένα μας, και αυτό το μοντέλο μπορεί να εξηγήσει 20,14 από αυτές τις μονάδες (σχεδόν το μισό). Το τελικό άθροισμα των τετραγώνων είναι το υπόλοιπο άθροισμα των τετραγώνων (SSE), το οποίο μας λέει πόση από την διασπορά δεν μπορεί να εξηγηθεί από το μοντέλο. Αυτή η τιμή είναι η ποσότητα της μεταβολής που προκαλείται από εξωγενείς παράγοντες, όπως το άτομο διαφορές στο βάρος, τεστοστερόνη ή οτιδήποτε άλλο. Γνωρίζοντας ήδη το SST_{otal} και το SST, ο πιο απλός τρόπος υπολογισμού του SSR είναι να αφαιρέσουμε το SST από το SST_{otal} ($SSR = SST_{otal} - SST$). Ωστόσο αυτό δεν παρέχει πλήρη εικόνα για το τι υπολογίζεται και, φυσικά, αν έχουμε χάσει κάτι στους υπολογισμοί SST ή SST_{otal} (ή και στους δύο!) τότε το SSE θα είναι επίσης λάθος.
Είδαμε ότι το υπολειπόμενο άθροισμα των τετραγώνων είναι η διαφορά μεταξύ του τι το μοντέλο προβλέπει και τι παρατηρήθηκε. Γνωρίζουμε ήδη ότι για ένα συγκεκριμένο συμμετέχοντα, το μοντέλο προβλέπει το μέσο όρο της ομάδας στην οποία ανήκει το άτομο αυτό. Ως εκ τούτου, η SSE υπολογίζεται με την εξέταση της διαφοράς μεταξύ της βαθμολογίας που επιτυγχάνεται από ένα άτομο και του μέσου όρου της ομάδας στην οποία ανήκει το πρόσωπο αυτό.
$SSE = \sum(\bar{x}_{ik} - \bar{x}_k)^2$
Τώρα, το άθροισμα των τετραγώνων για κάθε ομάδα αντιπροσωπεύει το άθροισμα των τετραγωνικών διαφορών μεταξύ της βαθμολογία κάθε συμμετέχοντα σε αυτή την ομάδα και τον μέσο της ομάδας. Ως εκ τούτου, μπορούμε να εκφράσουμε την SSE ως SSE = SSgroup1 + SSgroup2 + SSgroup3 ... και ούτω καθεξής. Δεδομένου ότι γνωρίζουμε τη σχέση μεταξύ της διασποράς και των αθροισμάτων των τετραγώνων, μπορούμε να χρησιμοποιήσουμε τις διασπορές για κάθε ομάδα για να δημιουργήσουμε μια εξίσωση όπως και για το συνολικό άθροισμα των τετραγώνων. Ως εκ τούτου, το SSE μπορεί να εκφραστεί ως:
$SSE = \sum s_k^2(n_k - 1)$
Αυτό σημαίνει, να πάρουμε τη διαφορά από κάθε ομάδα (s^2-k) και να το πολλαπλασιάσουμε κατά ένα μικρότερο από τον αριθμό των ατόμων σε αυτήν την ομάδα (n_k - 1). Όταν το κάνουμε αυτό για κάθε ομάδα, τα προσθέτουμε όλα. Για τα δεδομένα του Viagra, αυτό μας δίνει:
$SSE = s^2_{group1}(n_1 - 1) + s^2_{group2}(n_2 - 1) + s^2_{group3}(n_3 - 1) = (1{,}70)(5 - 1) + (1{,}70)(5 - 1) + (2{,}50)(5 - 1) = (1{,}70 \times 4) + (1{,}70 \times 4) + (2{,}50 \times 4) = 6{,}8 + 6{,}8 + 10 = 23{,}60$
Οι βαθμοί ελευθερίας της SSE είναι N-k: το συνολικό μέγεθος δείγματος, N, μείον τον αριθμό των ομάδων, k.

Το SST μας λέει τη συνολική παραλλαγή που εξηγείται από το μοντέλο παλινδρόμησης (π.χ. η πειραματική επεξεργασία) και η SSE μας λέει τη συνολική διασπορά που οφείλεται σε εξωγενείς παράγοντες. Ωστόσο, επειδή και οι δύο αυτές τιμές είναι αθροισμένες τιμές που θα επηρεαστούν από τον αριθμό των βαθμολογιών που αθροίζονται, για παράδειγμα, η SSM χρησιμοποίησε το άθροισμα μόνο 3 διαφορετικών τιμών (το μέσο της ομάδας) σε σύγκριση με την SSE και την SST_{otal}, στις οποίες χρησιμοποιήσαμε το άθροισμα των 12 και 14 τιμών αντίστοιχα. Για να εξαλείψουμε αυτή την προκατάληψη, μπορούμε να υπολογίσουμε το μέσο άθροισμα των τετραγώνων (γνωστά ως μέσες τιμές τετραγώνων, MS), που είναι απλά το σύνολο τετραγώνων διαιρούμενο με τους βαθμούς ελευθερίας.

Ο λόγος για τον οποίο χωρίζουμε ας προς τους βαθμούς ελευθερίας παρά ως προς τον αριθμό των παραμέτρων που χρησιμοποιούνται για τον υπολογισμό των SS είναι επειδή προσπαθούμε να προεκτείνουμε έναν πληθυσμό και έτσι κάποιες παράμετροι εντός αυτών των πληθυσμών θα κρατηθούν σταθερές. Έτσι, για τα δεδομένα του προβλήματος μας βρίσκουμε τα ακόλουθα τετράγωνα:

$MSM = \frac{SSM}{dfM} = 20{,}135 / 2 = 10{,}067$

$MSR = \frac{SSR}{dfR} = 23{,}60 / 12 = 1{,}967$

Το MSM αντιπροσωπεύει τη μέση ποσότητα παραλλαγής που εξηγείται από το μοντέλο (π.χ. το συστηματική διασπορά), ενώ το MSR είναι ένα μέτρο της μέσης ποσότητας μεταβολής που εξηγείται από τις εξωγενείς μεταβλητές (η μη συστηματική μεταβολή).

Ο λόγος F είναι ένα μέτρο του λόγου της μεταβολής που εξηγείται από το μοντέλο και την διασπορά εξηγούμενη από μη συστηματικούς παράγοντες. Με άλλα λόγια, είναι η αναλογία του πόσο καλό είναι το μοντέλο ενάντια στο πόσο κακό είναι (πόσο λάθος υπάρχει). Μπορεί να υπολογιστεί διαιρώντας τα μοντέλα μέσου τετραγώνου από τα υπολειπόμενα μέσα τετράγωνα.

$F=\frac{MSM}{MSR}$

Ο λόγος F είναι, επομένως, ένα μέτρο της αναλογίας των συστηματικών διασπορών σε μη συστηματικές διασπορές. Στην πειραματική έρευνα, είναι η αναλογία των πειραματικού επιπτώσεων στις μεμονωμένες διαφορές απόδοσης. Ένα ενδιαφέρον σημείο για τον λόγο F είναι αυτό επειδή είναι ο λόγος συστηματικής διασποράς προς μη συστηματική διασπορά, εάν η αξία του είναι μικρότερη από 1, τότε πρέπει, εξ ορισμού, να αντιπροσωπεύει μη σημαντικές επιπτώσεις. Ο λόγος για τον οποίο αυτή η δήλωση είναι αληθής είναι επειδή, αν η αναλογία F είναι μικρότερη από 1 σημαίνει ότι το MSR είναι μεγαλύτερο από το MSM, πράγμα που σημαίνει ότι υπάρχει πιο ασταθής συστηματική διασπορά. Μπορούμε να το σκεφτούμε αυτό από την άποψη των επιπτώσεων των φυσικών διαφορών στην ικανότητα να είναι μεγαλύτερες από τις διαφορές που προκαλούνται από το πείραμα. Σε αυτό το σενάριο, μπορούμε, επομένως, να είμαστε βέβαιοι ότι η πειραματικός μας χειρισμός ήταν ανεπιτυχής (γιατί έχει επιφέρει λιγότερες αλλαγές από ό, τι εάν αφήναμε τους συμμετέχοντες μας μόνο!). Για τα δεδομένα του προβλήματος, ο λόγος F είναι:

$F = \frac{MSM}{MSR} = 10{,}067 / 1{,}967 = 5{,}12$

Αυτή η τιμή είναι μεγαλύτερη από 1, πράγμα που δείχνει ότι ο πειραματικός χειρισμός είχε κάποια επίδραση πέραν των επιπτώσεων των μεμονωμένων διαφορών στις επιδόσεις. Ωστόσο, αυτό δεν μας λέει ακόμη αν η αναλογία F είναι αρκετά μεγάλη ώστε να μην είναι ένα τυχαίο αποτέλεσμα. Για να το ανακαλύψουμε αυτό μπορούμε να συγκρίνουμε την ληφθείσα τιμή του F έναντι της μέγιστης τιμής που θα περίμενε κανείς για να πάρουμε τυχαία αν οι μέσοι της ομάδας είναι ισότιμοι σε μια κατανομή F με τους ίδιους βαθμούς της ελευθερίας, αν η τιμή που αποκτάται υπερβαίνει αυτή την κρίσιμη τιμή μπορούμε να είμαστε σίγουροι ότι αυτό αντανακλά ένα αποτέλεσμα της ανεξάρτητης μας μεταβλητής (επειδή αυτή η τιμή θα ήταν πολύ απίθανη εάν δεν υπήρχε καμία επίδραση στον πληθυσμό). Στην περίπτωση αυτή, με 2 και 12 βαθμούς ελευθερίας οι κρίσιμες τιμές είναι 3,89 (p = .05) και 6,93 (p = 0,01). Η παρατηρούμενη τιμή, 5,12, είναι επομένως σημαντική σε επίπεδο σημαντικότητας 0,05 αλλά δεν είναι σημαντική σε επίπεδο 0,01. Η ακριβής σημαντικότητα που παράχθηκε από το SPSS πρέπει, κατά συνέπεια να πέφτουν κάπου μεταξύ 0,05 και 0,01 (που, παρεμπιπτόντως, το κάνει).

Παράδειγμα

Ένας κατασκευαστής μπρούτζινων γραμματοκιβωτίων και εξωτερικών φωτιστικών δέχθηκε παράπονα ότι τα προϊόντα του σκουριάζουν πρόωρα, και εντόπισε το πρόβλημα στην κακή ποιότητα βαφής. Για να αντικαταστήσει την βαφή που χρησιμοποιούσε επέλεξε πέντε εναλλακτικές βαφές, έβαψε με κάθε μια 25 μεταλλικά γραμματοκιβώτια και κατέγραψε σε πόσες μέρες κάθε γραμματοκιβώτιο εμφάνισε το πρώτο ίχνος σκουριάς. Μπορούμε από τα δεδομένα να συμπεράνουμε με στάθμη σημαντικότητας 1% ότι υπάρχει διαφορά στον χρόνο εμφάνισης σκουριάς των πέντε βαφών; Ποιες είναι οι απαραίτητες προϋποθέσεις για τον υπολογισμό; Μπορούμε να δεχτούμε ότι οι προϋποθέσεις πληρούνται;

Γραμ/τιο 1	Γραμ/τιο 2	Γραμ/τιο 3	Γραμ/τιο 4	Γραμ/τιο 5
133	173	161	160	154
171	182	115	221	180
142	214	137	208	222
154	165	173	249	147
242	168	176	148	159
147	232	138	99	175
171	217	106	182	193
155	220	238	169	209
165	184	129	162	192
156	145	163	158	173
219	126	156	240	188
159	209	140	168	147
179	152	109	121	240
189	139	128	146	183
110	237	121	191	236

198	107	149	222	186
119	266	165	176	168
177	234	121	153	169
185	158	162	238	177
206	154	166	195	128
145	199	190	196	152
81	195	214	240	213
171	133	197	152	135
175	247	130	233	165
165	185	185	138	181

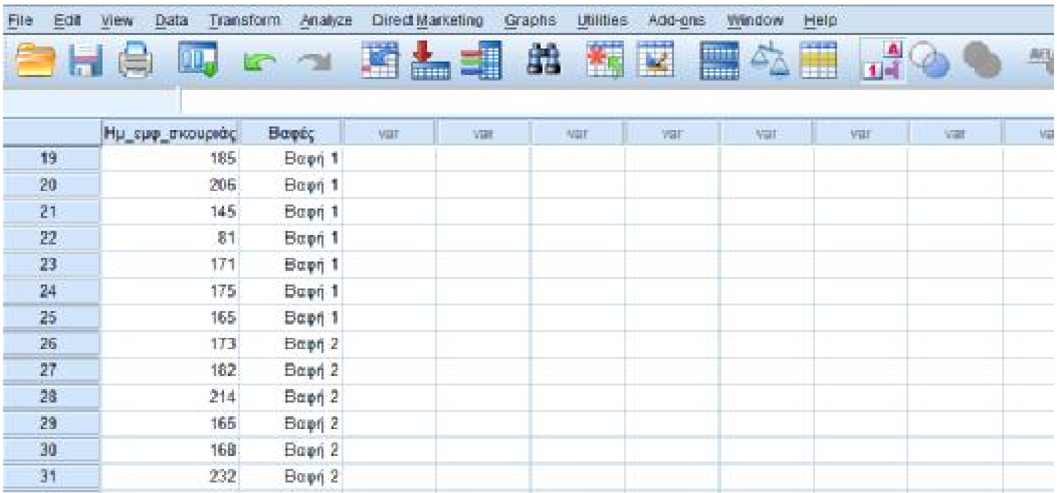
ΛΥΣΗ

Διατύπωση_της μηδενικής και εναλλακτικής υπόθεσης:

$$H_0: \mu_1=\mu_2=\mu_3=\mu_4$$

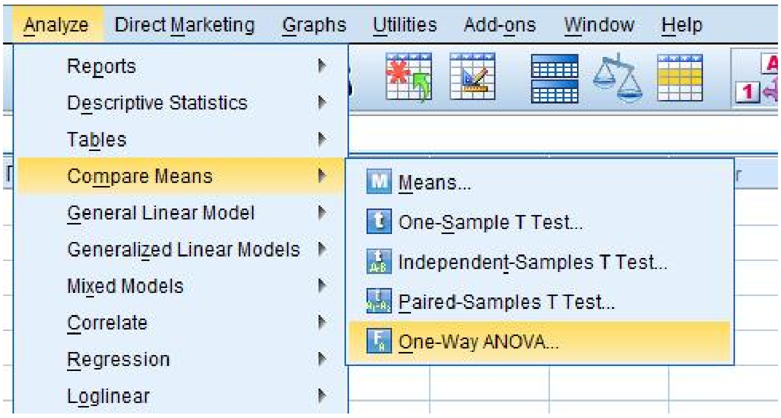
H_1: Δύο τουλάχιστον μέσοι διαφέρουν

Η εισαγωγή δεδομένων στο SPSS γίνεται σε δύο στήλες. Στη πρώτη στήλη εισάγουμε τις μέρες που κάθε γραμματοκιβώτιο εμφάνισε το πρώτο ίχνος σκουριάς (αυτά είναι και τα επίπεδα ή αλλιώς factor levels) και στην δεύτερη εισάγουμε τις βαφές που αντιστοιχούν τα δεδομένα της πρώτης στήλης (αυτός είναι και ο παράγοντας ή αλλιώς factor).

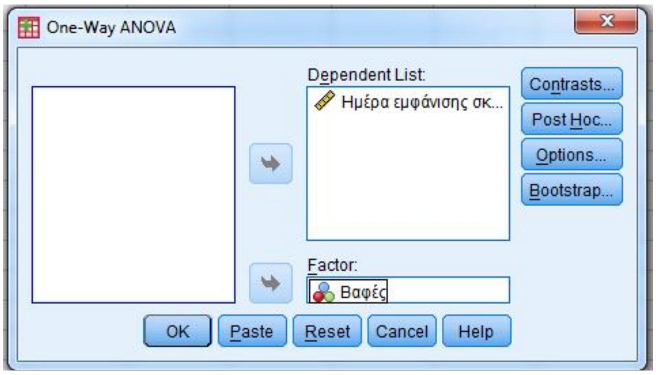


	Ημ_εμφ_σκουριάς	Βαφές
19	185	Βαφή 1
20	206	Βαφή 1
21	145	Βαφή 1
22	81	Βαφή 1
23	171	Βαφή 1
24	175	Βαφή 1
25	165	Βαφή 1
26	173	Βαφή 2
27	162	Βαφή 2
28	214	Βαφή 2
29	165	Βαφή 2
30	168	Βαφή 2
31	232	Βαφή 2

Τα βήματα που ακολουθούμε για την λύση του προβλήματος είναι τα εξής:
(Analyze **à** Compare Means **à** One-Way ANOVA)



Στο παράθυρο (One-Way ANOVA) που ανοίγει στο Dependent List (εξαρτημένη μεταβλητή) εισάγουμε την μεταβλητή «Ημέρα εμφάνισης σκουριάς» και στο factor (παράγοντας) εισάγουμε την μεταβλητή «Βαφές».



Στην επιλογή Post Hoc αλλάζουμε την στάθμη σημαντικότητας σε 0,01 και στην επιλογή options πατάμε την επιλογή Descriptive και Means Plot. Συνεχίζουμε πατώντας ΟΚ για να δούμε τα δεδομένα μας.

ANOVA

Ημέρα εμφάνισης σκουριάς

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	17250,512	4	4312,628	3,319	,013
Within Groups	155941,120	120	1299,509		
Total	173191,632	124			

Η τιμή του ελέγχου είναι F=3,319. Η τιμή p=0.013 (>0,01) είναι μεγαλύτερη απ’ την στάθμη σημαντικότητας, οπότε αποδεχόμαστε την H_0. Συμπεραίνουμε, δηλαδή, ότι δεν υπάρχει διαφορά στον χρόνο εμφάνισης σκουριάς των πέντε βαφών.

2. Πολλαπλές συγκρίσεις:

Αν διαπιστώσουμε διαφορές στους μέσους δύο ή περισσότερων πληθυσμών στην μονοδιάστατη ανάλυση διασποράς, το επόμενο βήμα είναι να εξετάσουμε ποιοι μέσοι διαφέρουν. Φαίνεται απλό, αλλά στην περίπτωση που έχουμε ένα μεγάλο αριθμό πληθυσμών οι υποθέσεις που θα έπρεπε να ελεγχθούν είναι πολλές. Επίσης για τον έλεγχο των υποθέσεων αυτών δεν αρκεί μόνο να γνωρίζουμε τους μέσους των δειγμάτων για αυτό θα χρησιμοποιήσουμε μια ειδική στατιστική μέθοδο, που ονομάζεται: πολλαπλές συγκρίσεις (multiple comparisons)

1.2. Μέθοδος Fisher: ελάχιστη σημαντική διαφορά:

Έχουμε αναφέρει πιο πάνω τον έλεγχο υπόθεσης για την ισότητα των μέσων δύο πληθυσμών με ίσες διασπορές, όπου ο έλεγχος και ο εκτιμητής διαστήματος εμπιστοσύνης δίνονται απ' τους τύπους:

$$t=\frac{(\overline{x_1}-\overline{x_2})-(\mu_1-\mu_2)}{\sqrt{s_p^2(\frac{1}{n_1}+\frac{1}{n_2})}}$$

$$(\bar{x}_1-\bar{x}_2)\pm t_{a/2}\sqrt{s_p^2(\frac{1}{n_1}+\frac{1}{n_2})}$$

με $v=n_1+n_2-2$ βαθμούς ελευθερίας και s_p^2 ο σταθμιμένος εκτιμητής διασποράς (αποτελεί τον καλύετρο αμερόληπτο εκτιμητή της κοινής διασποράς των δύο πληθυσμών με την προϋπόθεση ότι οι δύο πληθυσμοί έχουν ίσες διασπορές).

Ο αμερόληπτος εκτιμητής της κοινής διασποράς των πληθυσμών στην ανάλυση διασποράς ανεξάρτητων δειγμάτων είναι το μέσο τετράγωνο σφάλματος MSE. Επειδή η ποσότητα αυτή βασίζεται στις παρατηρήσεις όλων των πληθυσμών (και όχι μόνο δύο από αυτούς) είναι καλύτερος εκτιμητής και τον αντικαθιστούμε στην θέση του s_p^2. Οπότε και θα αλλάξουμε τους βαθμούς ελευθερίας σε $v=n-k$ (όπου n το συνολικό μέγεθος του δείγματος).

Άρα ο έλεγχος που μας επιτρέπει να καθορίσουμε αν διαφέρουν οι μέσοι δύο πληθυσμών i και j, γίνεται:

$$t=\frac{(\overline{x_i}-\overline{x_j})-(\mu_i-\mu_j)}{\sqrt{MSE(\frac{1}{n_i}+\frac{1}{n_j})}}$$

και ο εκτιμητής διαστήματος εμπιστοσύνης γίνεται:

$$(\bar{x}_i-\bar{x}_j)\pm t_{a/2}\sqrt{MSE(\frac{1}{n_i}+\frac{1}{n_j})}$$

με $v=n-k$ βαθμούς ελευθερίας.

Ορίζουμε ως ελάχιστη σημαντική διαφορά (least significance difference-LSD) την ποσότητα:

LSD = $t_{a/2}\sqrt{MSE(\frac{1}{n_i}+\frac{1}{n_j})}$

Ένας απλός τρόπος για να ελέγξουμε αν οι δύο μέσοι διαφέρουν είναι να συγκρίνουμε τη διαφορά των μέσων των αντιστοίχων δειγμάτων με την ελάχιστη σημαντική διαφορά. Δηλαδή, οι μέσοι μ_i και μ_j διαφέρουν αν:

$|\bar{x}_i-\bar{x}_j|>$ LSD

Αν τα μεγέθη όλων των πληθυσμών είναι ίσα, τότε η ελάχιστη σημαντική διαφορά είναι ίδια για κάθε ζεύγος πληθυσμών.

Έχουμε αναφέρει ότι η μέθοδος σύγκρισης μέσων των πληθυσμών ανά δύο περιέχει αυξημένη πιθανότητα σφάλματος τύπου Ι, δηλαδή αυξημένη πιθανότητα να ανίχνευση διαφορά μεταξύ των μέσων των πληθυσμών ενώ στην πραγματικότητα δεν υπάρχει. Για παράδειγμα, αν χρησιμοποιήσουμε 6 δείγματα του ίδιου πληθυσμού και ελέγξουμε με στάθμη σημαντικότητας 5%, υπάρχει πιθανότητα 54% να συμπεράνουμε ότι κάποιοι από τους μέσους των 6 πληθυσμών διαφέρουν, παρότι πρόκειται για τον ίδιο μέσο. Η στάθμη σημαντικότητας α=5% ονομάζεται ποσοστό σφάλματος τύπου Ι **ανά σύγκριση (comparison-wise)**, ενώ το ποσοστό α_E=54% ονομάζεται ποσοστό σφάλματος τύπου Ι **ανά πείραμα (experiment-wise)**. Τα δύο αυτά ποσοστά συνδέονται με τον τύπο:

$\alpha_E=1-(1-\alpha)^C$

όπου C ο αριθμός των συγκρίσεων των πληθυσμών ανά δυο, δηλαδή:

$C=\frac{k(k-1)}{2}$

Αποδεικνύεται μαθηματικά ότι:

$\alpha_E\leq \alpha\cdot C$

δηλαδή αν θέλουμε να περιορίσουμε την πιθανότητα εμφάνισης ενός τουλάχιστον σφάλματος τύπου Ι σε ένα δεδομένο ποσοστό α_E θα πρέπει να καθορίσουμε:

$\alpha=\frac{a_E}{C}$

Από τις δύο περιπτώσεις αυτές προκύπτει η ακόλουθη μέθοδος.

2.2. Μέθοδος Bonferroni: ελάχιστη σημαντική διαφορά:

Για τη βελτίωση της μεθόδου Fisher ώστε να περιοριστεί η πιθανότητα σφάλματος τύπου Ι, η μέθοδος Bonferroni καθορίζει τη στάθμη σημαντικότητας με βάση τον αριθμό των συγκρίσεων και το επιθυμητό ποσοστό σφάλματος τύπου Ι ανά πείραμα. Για παράδειγμα, αν ο αριθμός των πληθυσμών είναι k=6 και θέλουμε η συνολική πιθανότητα σφάλματος τύπου Ι ανά πείραμα να είναι 5%, βρίσκουμε:

$C=\frac{k(k-1)}{2}=\frac{6(6-1)}{2}=15$

και

$\alpha=\frac{a_E}{C}=\frac{0,05}{15}=0,033$

Παράδειγμα

Το μειονέκτημα της ελάχιστης σημαντικής διαφοράς είναι ότι εμπεριέχει υψηλή πιθανότητα σφάλματος τύπου Ι. Η μέθοδος Bonferroni μειώνει αυτή την πιθανότητα, αλλά αναπόφευκτα αυξάνει την πιθανότητα σφάλματος τύπου ΙΙ, δηλαδή την πιθανότητα να μην ανιχνευτεί διαφορά μεταξύ δύο μέσων ενώ στην πραγματικότητα υπάρχει. Αυτό αντιμετωπίζεται με την επόμενη μέθοδο.

3.2. Μέθοδος Tukey: πολλαπλή σύγκριση:

Όπως και με τις προηγούμενες μεθόδους έτσι και η μέθοδος Tukey υπολογίζει έναν οριακό αριθμό, ώστε αν η απόλυτη τιμή της διαφοράς των μέσων δύο δειγμάτων είναι μεγαλύτερη από τον αριθμό αυτόν μπορούμε να συμπεράνουμε ότι οι μέσοι των αντιστοίχων πληθυσμών διαφέρουν.
Η μέθοδος βασίζεται στην συνάρτηση:

$$q=\frac{\overline{x_{max}}-\overline{x_{min}}}{s/\sqrt{n}}$$

που ονομάζεται «**εύρος student t**» (studentized range) και οφείλει την ονομασία της στο William Gosset, δημιουργό της κατανομής student t, ο οποίος δημοσίευε τις εργασίες του με το ψευδώνυμο «φοιτητής».
$\bar{x}_{max}$ και $\bar{x}_{min}$ είναι ο μέγιστος και ο ελάχιστος από τους μέσους των δειγμάτων. Με βάση την παραπάνω συνάρτηση υπολογίζεται ο **κρίσιμος αριθμός ω**, ως εξής:

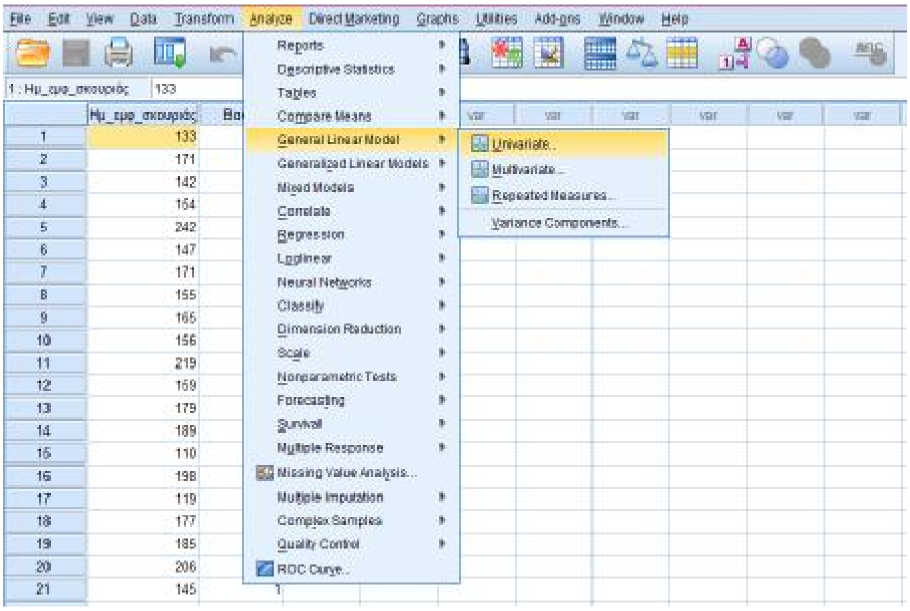
$$\omega=q_{\alpha}(k,v)\sqrt{\frac{MSE}{n_g}}$$

όπου
n = πλήθος παρατηρήσεων ($n=n_1+n_2+...+n_k$)
v = βαθμοί ελευθερίας (v=n-k)
n_g = μέγεθος κάθε δείγματος
α = στάθμη σημαντικότητας
$q_{\alpha}(k,v)$ = κρίσιμη τιμή της συνάρτησης q
Θεωρητικά η μέθοδος αυτή προϋποθέτει πως όλα τα δείγματα έχουν ίδιο μέγεθος, όμως ακόμη και αν τα μεγέθη των δειγμάτων είναι διαφορετικά, αλλά παρόμοια, η ποσότητα n_g μπορεί να υπολογισθεί ως ο **αρμονικός μέσος** των μεγεθών:

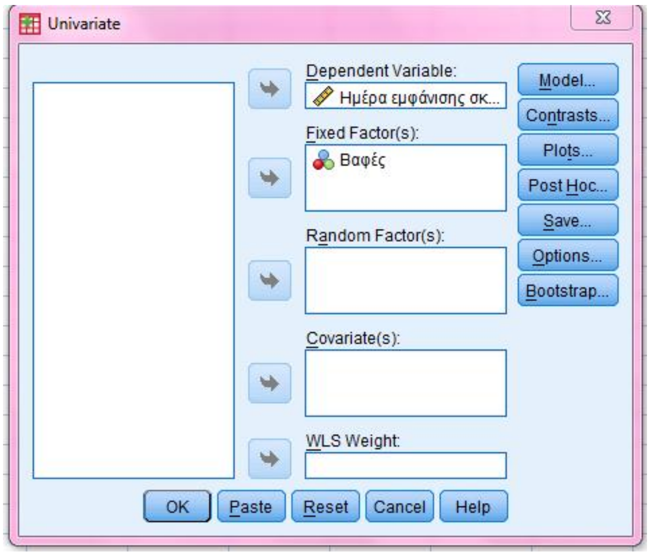
$$n_g=\frac{k}{\frac{1}{n_1}+\frac{1}{n_2}+\cdots+\frac{1}{n_k}}$$

Θα συνεχίσουμε στο προηγούμενο παράδειγμα, στο οποίο θα δούμε αναλυτικά τις τρεις μεθόδους με την χρήση του λογισμικού SPSS.

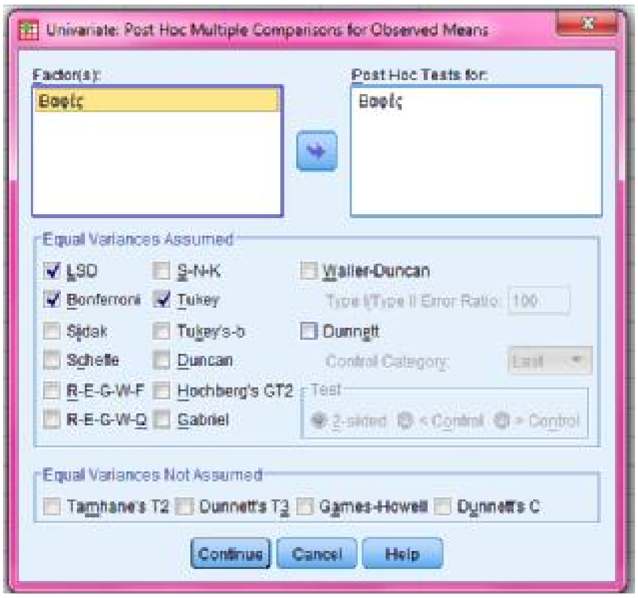
Έχουμε ήδη τα δεδομένα μας καταχωρημένα από πριν, πάμε στην γραμμή εργαλείων και επιλέγουμε Analyze à General Linear Model à Univariate.



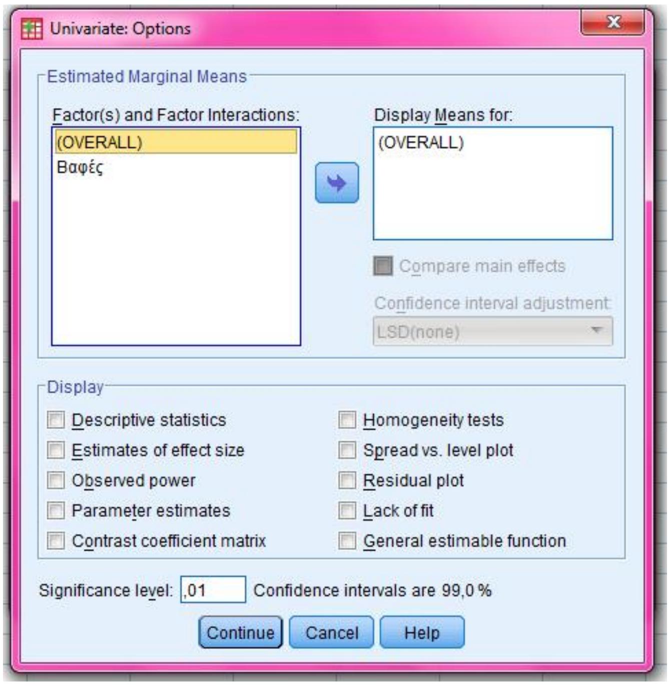
Στο Dependent Variable βάζουμε την μεταβλητή «Ημέρα εμφάνισης σκουριάς» και στο Fixed Factor(s) την μεταβλητή «Βαφές».



Στην επιλογή Post Hoc επιλέγουμε τα LSD, Bonferroni και Tukey ενώ έχουμε ήδη μεταφέρει την μεταβλητή «Βαφές» στο Post Hoc Tests for (στο δεξί πλαίσιο).



Τέλος στην επιλογή options περνάμε το (OVERALL) στο Display Means for και στο Significance Level βάζουμε 0,01 που δινόταν στην εκφώνηση της ασκήσεως.



Τα αποτελέσματα εμφανίζονται στους εξής πίνακες:

Ο πρώτος πίνακας μας δίνει απλά το σύνολο των δεδομένων για κάθε κωδικό ή αλλιώς για κάθε βαφή.

Between-Subjects Factors

		Value Label	N
Βαφές	1	Βαφή 1	25
	2	Βαφή 2	25
	3	Βαφή 3	25
	4	Βαφή 4	25
	5	Βαφή 5	25

Απ' τον δεύτερο πίνακα μας ενδιαφέρει το sig της μεταβλητής «Βαφές» το οποίο είναι 0,013. Αυτή είναι η τιμή p και είναι >0,01 η οποία δεν βρίσκεται στην περιοχή απόρριψης. Αν και μπορούμε να είμαστε σίγουροι ότι οι μέσοι δεν διαφέρουν στις βαφές θα συνεχίσουμε να δούμε αν όντως ισχύει.

Tests of Between-Subjects Effects

Dependent Variable:Ημέρα εμφάνισης σκουριάς στα γραμματοκιβώτια

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	17250,512[a]	4	4312,628	3,319	,013
Intercept	3753591,368	1	3753591,368	2888,468	,000
Βαφές	17250,512	4	4312,628	3,319	,013
Error	155941,120	120	1299,509		
Total	3926783,000	125			
Corrected Total	173191,632	124			

a. R Squared = ,100 (Adjusted R Squared = ,070)

Οι διαφορές δίνονται στον αμέσως επόμενο πίνακα. Το SPSS μας δίνει έναν ενιαίο πίνακα και διαχωρίζει τους ελέγχους με μία γραμμή. Εδώ χωρίσαμε τον πίνακα στα τρία για να εξηγήσουμε ξεχωριστά κάθε μέθοδο για καλύτερη κατανόηση των διαφορών.

LSD	Βαφή 1	Βαφή 2	-21,08	10,196	,041	-47,77	5,61
		Βαφή 3	9,80	10,196	,338	-16,89	36,49
		Βαφή 4	-18,04	10,196	,079	-44,73	8,65
		Βαφή 5	-14,32	10,196	,163	-41,01	12,37
	Βαφή 2	Βαφή 1	21,08	10,196	,041	-5,61	47,77
		Βαφή 3	30,88*	10,196	,003	4,19	57,57
		Βαφή 4	3,04	10,196	,766	-23,65	29,73
		Βαφή 5	6,76	10,196	,509	-19,93	33,45
	Βαφή 3	Βαφή 1	-9,80	10,196	,338	-36,49	16,89
		Βαφή 2	-30,88*	10,196	,003	-57,57	-4,19
		Βαφή 4	-27,84*	10,196	,007	-54,53	-1,15
		Βαφή 5	-24,12	10,196	,020	-50,81	2,57
	Βαφή 4	Βαφή 1	18,04	10,196	,079	-8,65	44,73
		Βαφή 2	-3,04	10,196	,766	-29,73	23,65
		Βαφή 3	27,84*	10,196	,007	1,15	54,53
		Βαφή 5	3,72	10,196	,716	-22,97	30,41
	Βαφή 5	Βαφή 1	14,32	10,196	,163	-12,37	41,01
		Βαφή 2	-6,76	10,196	,509	-33,45	19,93
		Βαφή 3	24,12	10,196	,020	-2,57	50,81
		Βαφή 4	-3,72	10,196	,716	-30,41	22,97

Να αναφέρουμε ότι η μέθοδος Fisher εμφανίζεται στο λογισμικό ως LSD, είναι όμως η μέθοδος που ζητάμε.

Κοιτάμε πάντα την τέταρτη στήλη όπου παρατίθενται τα sig (p-values).

Η στάθμη σημαντικότητας είναι 0,01 οπότε ψάχνουμε να βρούμε τιμές μικρότερης αυτής.

Βλέπουμε διαφορά ως προς τον χρόνο εμφάνισης της σκουριάς μεταξύ Βαφής 2 και 3 με τιμή p 0,003 καθώς και στην Βαφή 3 και 4 με τιμή p 0,007.

Είχαμε αναφέρει ότι το μειονέκτημα της ελάχιστης σημαντικής διαφοράς είναι ότι εμπεριέχει υψηλή πιθανότητα σφάλματος τύπου I. Με την επόμενη μέθοδο του Tukey περιμένουμε να εξαλειφθεί αυτή η πιθανότητα και να μην παρουσιαστούν διαφορές στους μέσους.

	(I) Βαφες	(J) Βαφες	Mean Difference (I-J)	Std. Error	Sig.	99% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	Βαφή 1	Βαφή 2	-21,08	10,196	,241	-55,03	12,87
		Βαφή 3	9,80	10,196	,872	-24,15	43,75
		Βαφή 4	-18,04	10,196	,396	-51,99	15,91
		Βαφή 5	-14,32	10,196	,626	-48,27	19,63
	Βαφή 2	Βαφή 1	21,08	10,196	,241	-12,87	55,03
		Βαφή 3	30,88	10,196	,025	-3,07	64,83
		Βαφή 4	3,04	10,196	,998	-30,91	36,99
		Βαφή 5	6,76	10,196	,964	-27,19	40,71
	Βαφή 3	Βαφή 1	-9,80	10,196	,872	-43,75	24,15
		Βαφή 2	-30,88	10,196	,025	-64,83	3,07
		Βαφή 4	-27,84	10,196	,055	-61,79	6,11
		Βαφή 5	-24,12	10,196	,132	-58,07	9,83
	Βαφή 4	Βαφή 1	18,04	10,196	,396	-15,91	51,99
		Βαφή 2	-3,04	10,196	,998	-36,99	30,91
		Βαφή 3	27,84	10,196	,055	-6,11	61,79
		Βαφή 5	3,72	10,196	,996	-30,23	37,67
	Βαφή 5	Βαφή 1	14,32	10,196	,626	-19,63	48,27
		Βαφή 2	-6,76	10,196	,964	-40,71	27,19
		Βαφή 3	24,12	10,196	,132	-9,83	58,07
		Βαφή 4	-3,72	10,196	,996	-37,67	30,23

Εδώ πραγματικά βλέπουμε τα αποτελέσματα που περιμέναμε εξ αρχής. Οι μέσοι πολύ σωστά δεν παρουσιάζουν διαφορές μεταξύ τους. Παρ' ότι δεν έχει πλέον σημασία θα παρουσιάσουμε και τον πίνακα της μεθόδου Bonferroni.

Όντως βλέπουμε στον πίνακα που ακολουθεί ότι η μία μέθοδος με την σειρά της απαλείφει τα σφάλματα της άλλης και έχουμε μια πιο καθαρή εικόνα ότι όντως δεν υπάρχουν διαφορές στους μέσους. Όλες δηλαδή οι βαφές έχουν την ίδια αντοχή.

[illegible]	Βαφή 1	Βαφή 2	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
		Βαφή 3	9,80	10,196	1,000	-24,60	[illegible]
		Βαφή 4	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
		Βαφή 5	-14,32	10,196	1,000	-40,72	[illegible]
	Βαφή 2	Βαφή 1	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
		Βαφή 3	[illegible]	10,196	[illegible]	[illegible]	[illegible]
		Βαφή 4	3,04	10,196	1,000	[illegible]	[illegible]
		Βαφή 5	[illegible]	10,196	1,000	-27,64	[illegible]
	Βαφή 3	Βαφή 1	-9,80	10,196	1,000	[illegible]	[illegible]
		Βαφή 2	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
		Βαφή 4	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
		Βαφή 5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	Βαφή 4	Βαφή 1	[illegible]	10,196	[illegible]	-16,06	[illegible]
		Βαφή 2	3,04	10,196	1,000	[illegible]	[illegible]
		Βαφή 3	[illegible]	10,196	[illegible]	-6,56	[illegible]
		Βαφή 5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	Βαφή 5	Βαφή 1	[illegible]	10,196	1,000	-20,00	[illegible]
		Βαφή 2	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
		Βαφή 3	24,12	10,196	[illegible]	-10,20	[illegible]
		Βαφή 4	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Based on observed means.
The error term is Mean Square(Error) = 1299,509.
[illegible]

Επιλογή μεθόδου σύγκρισης

Δυστυχώς, ο κάθε στατιστικός πρέπει κάθε φορά να επιλέξει την στατιστική μέθοδο που θέλει να χρησιμοποιήσει συνήθως βάση των εξής κριτηρίων:

- Αν έχουμε εντοπίσει 2-3 πληθυσμούς, από το σύνολο των πληθυσμών που έχουμε, που πρέπει να συγκριθούν πριν από την ανάλυση διασποράς τότε χρησιμοποιούμε την μέθοδο Bonferroni.
- Αν θέλουμε να ελέγξουμε όλες τις δυνατές συγκρίσεις, χρησιμοποιούμε την μέθοδο Tukey.
- Αν ο σκοπός της ανάλυσης είναι να εντοπίσουμε περιοχές που θα πρέπει να ερευνηθούν περισσότερο, ξεκινάμε απ' την μέθοδο Fisher.
- Για τις μεθόδους Fisher και Bonferroni πρέπει να έχει προηγηθεί η ανάλυση διασποράς, κάτι που δεν είναι απαραίτητο για την μέθοδο Tukey.

3. Δισδιάστατη ανάλυση διασποράς (ή σύγκριση κατά ομάδες):

Αναφέραμε προηγουμένως, ότι το κριτήριο με το οποίο διαφοροποιούνται οι πληθυσμοί ονομάζεται **παράγοντας** (**factor**) και κάθε διαφορετική τιμή του παράγοντα ονομάζεται **επίπεδο** (**factor level**). Επίσης το στατιστικό πείραμα πιο πάνω είναι πείραμα ενός παράγοντα.
Σε ένα **πείραμα πολλαπλών παραγόντων** (**multifactor experiment**) οι πληθυσμοί ορίζονται με συνδυασμό δύο ή περισσότερων παραγόντων.
Το σχέδιο αυτό αποτελεί ευθεία γενίκευση του σχεδίου που γνωρίσαμε όταν μιλήσαμε για τη σύγκριση κατά ζεύγη δύο μέσων, μ_1 και μ_2 . Η διαφορά από τη σύγκριση κατά ζεύγη έγκειται μόνο στο ότι αντί να δημιουργούμε έναν αριθμό, έστω b, ζευγών πειραματικών μονάδων επί των οποίων κάνουμε δύο επεμβάσεις, δημιουργούμε έναν αριθμό, b, ομάδων (blocks) αποτελούμενων από $k > 2$ πειραματικές μονάδες η κάθε μια (αντί δύο) γιατί κάνουμε περισσότερες από δύο επεμβάσεις (ή αλλιώς, γιατί ο παράγοντας του οποίου ελέγχουμε την επίδραση στη μεταβλητή απόκρισης, έχει περισσότερες από δύο στάθμες).
Ένας εναλλακτικός σχεδιασμός της σύγκρισης κατά ομάδες είναι η επανάληψη του πειράματος τρεις φορές με το ίδιο δείγμα υποκειμένων (π.χ. εργαζομένων, ανθρώπων, μηχανημάτων), ώστε να συγκριθεί η συμπεριφορά τους κάτω από διαφορετικές συνθήκες. Ο σχεδιασμός αυτός ονομάζεται πείραμα **επαναλαμβανομένων μετρήσεων** (**repeated measures**) και είναι από τεχνική άποψη διαφορετικός από τη σύγκριση κατά ομάδες , ωστόσο η επεξεργασία των δεδομένων γίνεται με τον ίδιο ακριβώς τρόπο και συνεπώς μπορούμε να θεωρήσουμε ότι πρόκειται για την ίδια μέθοδο.
Η σύγκριση των δειγμάτων κατά ομάδες ονομάζεται επίσης **δισδιάστατη** (**two-way)** ανάλυση διασποράς, σε αντίθεση με την επιλογή ανεξάρτητων δειγμάτων που ονομάζεται μονοδιάστατη (one-way).
Αν μια έρευνα περιλαμβάνει όλα τα επίπεδα ενός παράγοντα ονομάζεται ανάλυση διασποράς **σταθερού παράγοντα** (**fixed effects**). Αν περιλαμβάνει μόνο ένα τυχαίο δείγμα των επιπέδων ονομάζεται ανάλυση διασποράς **τυχαίου παράγοντα** (**random effects**).

Ο σκοπός της επιλογής των δειγμάτων κατά ομάδες είναι να εξουδετερωθεί μια πηγή διασποράς στο εσωτερικό των πληθυσμών, ώστε να μπορούν να ανιχνευτούν διαφορές μεταξύ των πληθυσμών
Η συνολική μεταβλητότητα που είχαμε στην μονοδιάστατη ανάλυση διασποράς χωριζόταν σε δύο συνιστώσες:
SS(total)=SST+SSE

Στη δισδιάστατη ανάλυση διασποράς, όπου τα δείγματα επιλέγονται κατά ομάδες, η συνολική μεταβλητότητα θα πρέπει να χωρίζεται σε τρεις συνιστώσες:
SS(total)=SST+SSB+SSE
SSB = μεταβλητότητα εντός των ομάδων (Sums of Squares for Blocks)
Με τον διαχωρισμό αυτό περιορίζεται η μεταβλητότητα εντός των πληθυσμών (SSE) και διευκολύνεται η ανίχνευση της μεταβλητότητας μεταξύ των πληθυσμών (SST).

		Πληθυσμός				Μέσος ομάδας
		1	2	...	K	
Ομάδα	1	X_{11}	X_{12}	...	X_{1k}	$\bar{x}[B]_1$
	2	X_{22}	X_{22}	...	X_{2k}	$\bar{x}[B]_2$
	...	...	...	...	...	...
	b	X_{b1}	X_{b2}	...	x_{bk}	$\bar{x}[B]_k$
Μέσος πληθυσμού		$\bar{x}[T]_1$	$\bar{x}[T]_2$	...	$\bar{x}[T]_k$	

όπου,
$\bar{x}[T]_j$ = Μέσος του πληθυσμού j(j=1,2,...,k)
$\bar{x}[B]_i$ = Μέσος της ομάδας i(i=1,2,...,b)
k = Αριθμός των πληθυσμών
b = Αριθμός των ομάδων
Οι ορισμοί των μεγεθών SS(total) και SST ταυτίζονται με τους ορισμούς που είδαμε στη μονοδιάστατη ανάλυση διασποράς μόνο που εδώ το SSE ισούται με το άθροισμα των μεγεθών SSB+SSE.

Αθροίσματα τετραγώνων στη σύγκριση κατά ομάδες
$SS(total) = \sum_{j=1}^{k}\sum_{i=1}^{b}(x_{ij}-\bar{\bar{x}})^2$
$SST = \sum_{j=1}^{k} b(\bar{x}[T]_j-\bar{\bar{x}})^2$
$SSB = \sum_{i=1}^{b} k(\bar{x}[B]_i-\bar{\bar{x}})^2$
$SSE = \sum_{j=1}^{k}\sum_{i=1}^{b}(x_{ij}-\bar{x}[T]_j-\bar{x}[B]_i+\bar{\bar{x}})^2$

Ο έλεγχος υπολογίζεται με τη βοήθεια των μέσων τετραγώνων, είναι στην ουσία η διαίρεση κάθε αθροίσματος τετραγώνων με τους αντίστοιχους βαθμούς ελευθερίας.

Μέσα τετράγωνα στη σύγκριση κατά ομάδες
$MST = \frac{SST}{k-1}$

$MSB = \frac{SSB}{b-1}$

$MSE = \frac{SSE}{n-k-b+1}$

Ο έλεγχος για την ύπαρξη διαφορών μεταξύ των μέσων των ομάδων είναι:

$F=\frac{MST}{MSE}$

όπου MSB και MSE είναι τα μέσα τετράγωνα που έχουν ήδη υπολογιστεί, και η κατανομή δειγματοληψίας είναι επίσης η κατανομή F με βαθμούς ελευθερίας v_1=k-1 και v_2=n-k-b+1.

Τα αποτελέσματα της δισδιάστατης ανάλυσης διασποράς συνοψίζονται σε ένα πίνακα:

Πηγή μεταβλητότητας	Βαθμοί ελευθερίας	Αθροίσματα τετραγώνων	Μέσα τετράγωνα	Έλεγχος
Χαρακτηριστικά	k-1	SST	MST	F=MST/MSE
Ομάδες	b-1	SSB	MSB	F=MSB/MSE
Σφάλμα	n-k-b+1	SSE	MSE	
Σύνολο	n-1	SS(total)		

ΠΑΡΑΔΕΙΓΜΑ:

Πόσο διαφέρουν οι ειδικότητες των ιατρών ως προς τον χρόνο που αφιερώνουν στην φροντίδα των ασθενών τους; Μια έρευνα επέλεξε ένα τυχαίο δείγμα ιατρών πέντε ειδικοτήτων και κατέγραψε τον αριθμό των ωρών που αφιέρωσε ο καθένας στη φροντίδα των ασθενών του στη διάρκεια της εβδομάδας. Τα δείγματα ομαδοποιήθηκαν με κριτήριο την ηλικία των ιατρών.

Ηλικία	Γενικός Ιατρός	Εσωτερικός Ιατρός	Χειρούργος Ιατρός	Παιδίατρος	Γυναικολόγος
30	54	56,7	55,7	58,3	60,6
31	54,9	61,8	61,5	50,7	65,2
32	55,5	46,6	59,5	56,4	63,2
33	37	44,6	40,8	51,4	37,3
34	52	63,8	44,9	52,5	56,6
35	54,5	57,3	45,2	50,8	56,4
36	46,4	51,6	50	58,5	59
37	42,5	42,2	29,6	42,4	41,5
38	45,3	37	56,6	34,6	55,1
39	45,5	53,3	59,3	49,7	42,5
40	50,4	46,3	44,6	47,1	63,5
41	55,3	67,7	55	56,8	62,8

42	52,1	55,9	52,8	55,5	60,6
43	56,7	65,6	56,1	53,6	57,7
44	44,9	52,2	46,5	43,9	51,2
45	57,4	66,9	57,5	54,3	72,8
46	53,2	55,4	69	53,1	58,2
47	61,7	61	63,4	51,1	73,2
48	57,2	56,8	57	46,9	54,9
49	47,4	48,9	60,9	40	54,3
50	57,3	66,8	62,5	43,4	61,1
51	48,8	53	48,2	54,6	58,1
52	61,9	75	56,1	60,7	65,2
53	47,2	48,6	34,6	37,8	47,8
54	50,1	51,1	55	42,1	58,9
55	57,8	66,5	56	55,9	59,8
56	58,9	58,8	57,9	74,4	70,8
57	64,3	66,3	65,2	49,5	72,8
58	45,9	46,1	68	47,2	58,8
59	57,2	65,4	56,6	59,5	62,4
60	60	66,8	53,2	52,4	68,3
61	61,3	56,9	70,2	58	68
62	58,7	62	65,2	59,3	74,2
63	47,7	37,6	44,9	47,3	57,9
64	53,5	47,1	61,7	53,5	55
65	56,2	57	58,1	54,1	66

Μπορούμε από τα δεδομένα να συμπεράνουμε ότι υπάρχουν διαφορές ανάμεσα στις ειδικότητες ως προς τον χρόνο φροντίδας των ασθενών; Η ομαδοποίηση κατά ηλικία ήταν η καταλληλότερη για την συγκεκριμένη έρευνα;

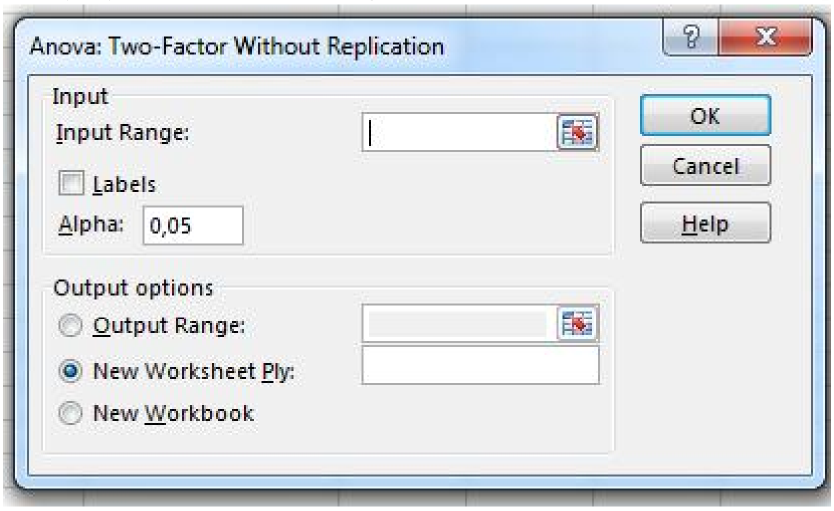
ΛΥΣΗ:

Το πρόβλημα αυτό απαιτεί την σύγκριση 5 πληθυσμών. Η μεταβλητή απόκρισης είναι οι ώρες φροντίδας των ασθενών, ο παράγοντας είναι οι ιατροί, τα επίπεδα του παράγοντα είναι οι πέντε ειδικότητες των ιατρών και οι πέντε πληθυσμοί αντιστοιχούν στα πέντε αυτά επίπεδα.
Η μηδενική και η ενναλακτική υπόθεση είναι:
H_0: $\mu_1=\mu_2=\mu_3=\mu_4=\mu_5$
H_1: Δύο τουλάχιστον μέσοι διαφέρουν
Εισάγουμε τα στοιχεία μας σε στήλες (6 σύνολο) στο EXCEL

F2 | f_x 60,6

	A	B	C	D	E	F
1	Ηλικία	Γενικός Ιατρός	Εσωτερικός Ιατρός	Χειρούργος Ιατρός	Παιδίατρος	Γυναικολόγος
2	30	54	56,7	55,7	58,3	60,6
3	31	54,9	61,8	61,5	50,7	65,2
4	32	55,5	46,6	59,5	56,4	63,2
5	33	37	44,6	40,8	51,4	37,3
6	34	52	63,8	44,9	52,5	56,6
7	35	54,5	57,3	45,2	50,8	56,4
8	36	46,4	51,6	50	58,5	59
9	37	42,5	42,2	29,6	42,4	41,5
10	38	45,3	37	56,6	34,6	55,1
11	39	45,5	53,3	59,3	49,7	42,5
12	40	50,4	46,3	44,6	47,1	63,5
13	41	55,3	67,7	55	56,8	62,8
14	42	52,1	55,9	52,8	55,5	60,6
15	43	56,7	65,6	56,1	53,6	57,7
16	44	44,9	52,2	46,5	43,9	51,2
17	45	57,4	66,9	57,5	54,3	72,8
18	46	53,2	55,4	69	53,1	58,2
19	47	61,7	61	63,4	51,1	73,2

Στην συνέχεια επιλέγουμε Data à Data Analysis- à Anova two-factor without replication



Στο Input επιλέγουμε όλα τα κελιά με τα δεδομένα, το Alpha αντιπροσωπεύει την στάθμη σημαντικότητας, εδώ έχουμε 0,05, και είτε επιλέγουμε New Worksheet Ply για την εμφάνιση των αποτελεσμάτων σε νέο φύλλο εργασίας είτε στην ίδια σελίδα με το Output Range.
Τα αποτελέσματα δίνονται στον πίνακα Anova:

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Rows	7309,684	35	208,8481	6,364696	9,35E-16	1,507334
Columns	1406,39	4	351,5976	10,71502	1,32E-07	2,436317
Error	4593,894	140	32,81353			

Total	13309,97	179

Η ένδειξη Rows (γραμμές) αφορά τις ομάδες με P-value≈ 0 και F=6.36 η ένδειξη Columns (στήλες) αφορά τους πληθυσμούς όπου η τιμή P-value≈ 0 και ο έλεγχος F=10,72. Οπότε από τα δεδομένα μπορούμε να συμπεράνουμε ότι υπάρχουν διαφορές ανάμεσα στις ειδικότητες ως προς τον χρόνο φροντίδας των ασθενών και ότι η ομαδοποίηση κατά την ηλικία ήταν η καταλληλότερη. Άρα αποδεχόμαστε την H_1.

4. Ανάλυση διασποράς δύο παραγόντων:

Ο γενικός όρος που περιγράφει την ανάλυση διασποράς δύο παραγόντων είναι **παραγοντικά πειράματα** (**factorial experiments**). Σε ένα παραγοντικό πείραμα το ζητούμενο είναι αν η **μεταβλητή απόκρισης** (**response variable**) επηρεάζεται από τους διάφορους παράγοντες. Η ανάλυση που θα χρησιμοποιήσουμε είναι **σταθερού παράγοντα** (**fixed effects**) δηλαδή όλα τα επίπεδα (τιμές) των παραγόντων καλύπτονται από το πείραμα. Ένα **πλήρες παραγοντικό πείραμα** (**complete factorial experiment**) είναι ένα πείραμα που καλύπτει όλους τους δυνατούς συνδυασμούς των παραγόντων που εξετάζει. Σε προβλήματα που έχουμε ένα παράγοντα με δύο και ένα παράγοντα με τέσσερα επίπεδα θα πρέπει να χρησιμοποιήσουμε 2x4=8 πληθυσμούς. Αυτό ονομάζεται **πλήρες παραγοντικό πείραμα 2x4**.
Γενικά ο πρώτος παράγοντας (που επιλέγεται αυθαίρετα) συμβολίζεται ως A και το πλήθος των επιπέδων του a, ενώ ο δεύτερος παράγοντας συμβολίζεται ως B και το πλήθος των επιπέδων του b.
Ο αριθμός των παρατηρήσεων για κάθε συνδυασμό (το μέγεθος κάθε πληθυσμού) ονομάζεται **πλήθος αντιγράφων** (**replicates**) και συμβολίζεται ως **r**. Εμείς θα έχουμε να κάνουμε με πληθυσμούς που έχουν το ίδιο μέγεθος (το πλήθος αντιγράφων είναι σταθερό). Αύτο το πείραμα ονομάζεται **ισορροπημένο** (**balanced**).
Εδώ για να εξετάσουμε σε ποιον παράγοντα ή συνδυασμό παραγόντων οφείλονται οι διαφορές μεταξύ των πληθυσμών, θα πρέπει να διαχωρίσουμε την SST σε SS(A), SS(B) και SS(AB).

Παράγοντας Β	Παράγοντας Α							
	1		2		...	a		
1	X_{111} X_{112} ... X_{11r}	$\bar{x}[AB]_{11}$	X_{211} X_{212} ... X_{21r}	$\bar{x}[AB]_{21}$	...	X_{a11} X_{a12} ... X_{a1r}	$\bar{x}[AB]_{a}$	$\bar{x}[B]_1$
2	X_{121} X_{122} ... X_{12r}	$\bar{x}[AB]_{12}$	X_{221} X_{222} ... X_{22r}	$\bar{x}[AB]_{22}$	...	X_{a21} X_{a22} ... X_{a2r}	$\bar{x}[AB]_{a}$	$\bar{x}[B]_2$
...	...	...	...	...	...	...	...	...
b	X_{1b1} X_{1b2} ...	$\bar{x}[AB]_{1b}$	X_{2b1} X_{2b2} ...	$\bar{x}[AB]_{2b}$	....	X_{ab1} X_{ab2} ...	$\bar{x}[AB]_{a}$	$\bar{x}[B]_b$

	X_{1br}		X_{2br}			X_{abr}		
	$\bar{x}[A]_1$		$\bar{x}[A]_2$		…	$\bar{x}[A]_a$		$\bar{\bar{x}}$

Αθροίσματα τετραγώνων στην ανάλυση διαποράς δύο παραγόντων

$SS(total) = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}(x_{ijk}-\bar{\bar{x}})^2$

$SS(A) = rb\sum_{i=1}^{a}(\bar{x}[A]_i-\bar{\bar{x}})^2$

$SS(B) = ra\sum_{i=1}^{b}(\bar{x}[B]_j-\bar{\bar{x}})^2$

$SS(AB) = r\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{x}[AB]_{ij}-\bar{x}[A]_i-\bar{x}[B]_j+\bar{\bar{x}})^2$

$SE = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}(x_{ijk}-\bar{x}[AB]_{ij})^2$

όπου:
x_{ijk} = τιμή k του πληθυσμού ij
$\bar{x}[AB]_{ij}$ = μέσος του πληθυσμού ij
$\bar{x}[A]_i$ = μέσος όλων των πληθυσμών για το επίπεδο i του παράγοντα A
$\bar{x}[B]_j$ = μέσος όλων των πληθυσμών για το επίπεδο j του παράγοντα B
$\bar{\bar{x}}$ = μέσος όλων των πληθυσμών
a = αριθμός επιπέδων του παράγοντα A
b = αριθμός επιπέδων του παράγοντα B
r = αριθμός αντιγράφων

Ο όρος SS(A) εκφράζει την μεταβλητότητα που οφείλεται στον παράγοντα A και υπολογίζεται ως το άθροισμα, για όλα τα επίπεδα του παράγοντα A, των τετραγώνων των διαφορών μεταξύ του μέσου κάθε επιπέδου και του συνολικού μέσου. Όμοια υπολογίζεται ο όρος SS(B), ο οποίος εκφράζει την μεταβλητότητα που οφείλεται στον παράγοντα B.
Ο όρος SS(AB) εκφράζει την μεταβλητότητα που οφείλεται στην αλληλεπίδραση μεταξύ των παραγόντων και υπολογίζεται ως διπλό άθροισμα τετραγώνων των διαφορών μεταξύ του μέσου κάθε πληθυσμού και του συνολικού μέσου από τη μια πλευρά, και των μέσων των αντίστοιχων επιπέδων των δύο παραγόντων από την άλλη (κάθε πληθυσμός αντιπροσωπεύει ένα μοναδικό συνδυασμό επιπέδων των δύο παραγόντων).
Τέλος, ο όρος SSE εκφράζει την μεταβλητότητα εντός των πληθυσμών και υπολογίζεται ως τριπλό άθροισμα τετραγώνων των διαφορών των παρατηρήσεων από τους αντίστοιχους μέσους των πληθυσμών στους οποίους ανήκουν.

Μέσα τετράγωνα στην ανάλυση διασποράς δύο παραγόντων.

$MS(A) = \frac{SS(A)}{a-1}$

$MS(B) = \frac{SS(B)}{b-1}$

$MS(AB) = \frac{SS(AB)}{(a-1)(b-1)}$

$MSE = \frac{SSE}{n-ab}$

ΠΑΡΑΔΕΙΓΜΑ:

Ο πονοκέφαλος είναι από τις πιο συχνές και ταυτόχρονα λιγότερο κατανοητές παθήσεις. Οι περισσότεροι άνθρωποι υποφέρουν μια ή περισσότερες φορές κάθε μήνα, και τα συνηθισμένα παυσίπονα που χορηγούνται χωρίς συνταγή ιατρού είναι συνήθως αρκετά για να ανακουφίσουν από τα συμπτώματα. Υπάρχει όμως και ένα σημαντικό ποσοστό ανθρώπων για τους οποίους τα παυσίπονα δεν είναι αρκετά και ο πονοκέφαλος μπορεί να κάνει τη ζωή τους αφόρητη. Αυτοί αναγκάζονται να καταφεύγουν σε εναλλακτικές θεραπείες, όπως υπνωτικά, υπνωτισμό, βιοανάδραση και βελονισμό, χωρίς ιδιαίτερη επιτυχία. Τα τελευταία χρόνια έχει αναπτυχθεί μια νέα θεραπεία, που περιλαμβάνει ενέσεις τοπικού αναισθητικού στο ινιακό νεύρο (στο πίσω μέρος του λαιμού). Η επικρατούσα δοσολογία είναι μια ένεσα ανά εβδομάδα για 4 εβδομάδες, αλλά έχει προστεθεί και μια πυκνότερη χορήγηση , με μια ένεση ανά 2 ημέρες για ένα σύνολο 4 ενέσεων. Επιπλέον έχουν προστεθεί συνδυασμοί φαρμάκων που κάνουν τη θεραπεία αποτελεσματικότερη.

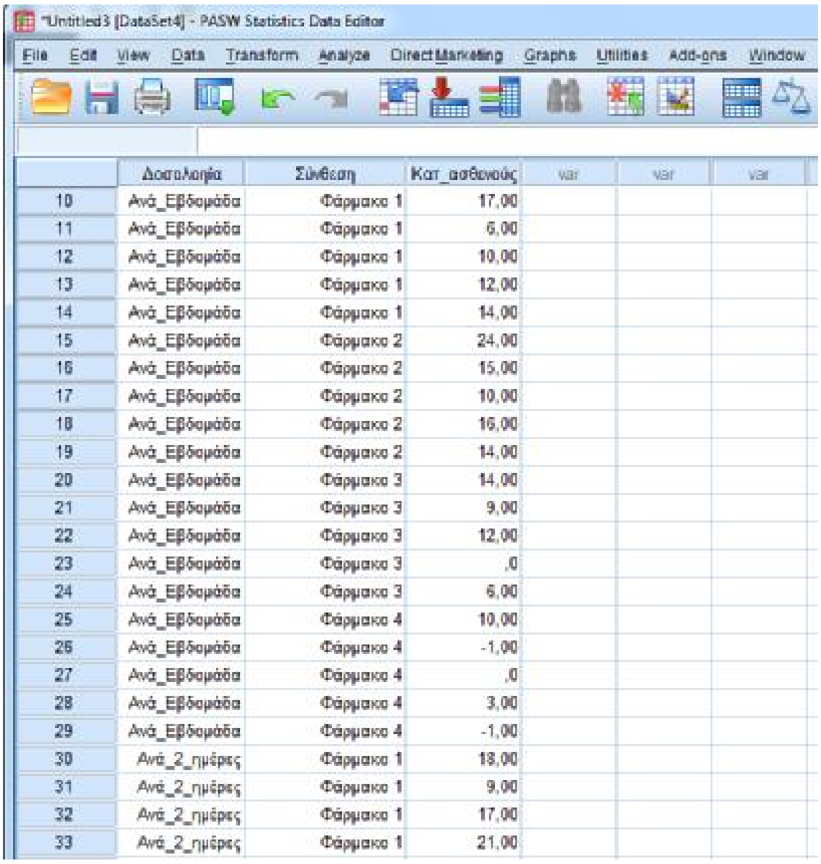
Για τον έλεγχο αυτών των προτάσεων πραγματοποιήθηκε μια πλήρης παραγοντική έρευνα, όπου για κάθε συνδυασμό δοσολογίας και σύνθεσης επιλέχθηκε ένα τυχαίο δείγμα 5 ασθενών, και για κάθε ασθενή καταγράφηκε η συχνότητα, η διάρκεια και η σοβαρότητα των πονοκεφάλων για 30 μέρες μετά την τελευταία ένεση. Από τα δεδομένα αυτά κατασκευάστηκε ένας δείκτης 0-100 για την κατάσταση κάθε ασθενούς πριν και μετά την θεραπεία. Στον πίνακα που ακολουθεί φαίνεται το αποτέλεσμα της θεραπείας ως διαφορά των δύο αυτών δεικτών. Αρνητική διαφορά σημαίνει ότι η κατάσταση του ασθενούς επιδεινώθηκε μετά τη θεραπεία.

Δοσολογία	Σύνθεση			
Ημέρες	Φάρμακο 1	Φάρμακο 2	Φάρμακο 3	Φάρμακο 4
Ανά εβδομάδα	17	24	14	10
Ανά εβδομάδα	6	15	9	-1
Ανά εβδομάδα	10	10	12	0
Ανά εβδομάδα	12	16	0	3
Ανά εβδομάδα	14	14	6	-1
Ανά 2 ημέρες	18	-2	20	-2
Ανά 2 ημέρες	9	0	16	7
Ανά 2 ημέρες	17	17	12	10
Ανά 2 ημέρες	21	2	17	6
Ανά 2 ημέρες	15	6	18	7

Ποιοι είναι οι παράγοντες του πειράματος; Ποια είναι η μεταβλητή απόκρισης; Ποια είναι τα επίπεδα κάθε παράγοντα; Να ελέγξετε με στάθμη σημαντικότητας 5% αν υπάρχουν διαφορές μεταξύ των δοσολογιών, αν υπάρχουν διαφορές μεταξύ των συνθέσεων, και αν υπάρχει αλληλεπίδραση μεταξύ των δυο παραγόντων.

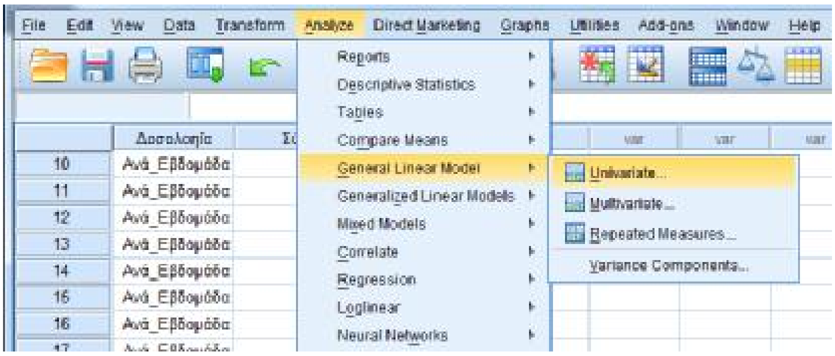
ΛΥΣΗ:

Στο παράδειγμα αυτό οι παράγοντες είναι η σύνθεση και η δοσολογία και η μεταβλητή απόκρισης είναι ο δείκτης βελτίωσης του ασθενούς. Τα επίπεδα της δοσολογίας είναι δύο και της σύνθεσης τέσσερα.
Θα εισάγουμε τώρα τα στοιχεία μας στο SPSS σε τρεις στήλες. Η πρώτη αφορά την δοσολογία (ανά εβδομάδα ή ανά 2 ημέρες), η δεύτερη αφορά την σύνθεση (φάρμακο 1,2,3,4 αντίστοιχα) και η τρίτη την κατάσταση του ασθενούς όπως φαίνεται παρακάτω:

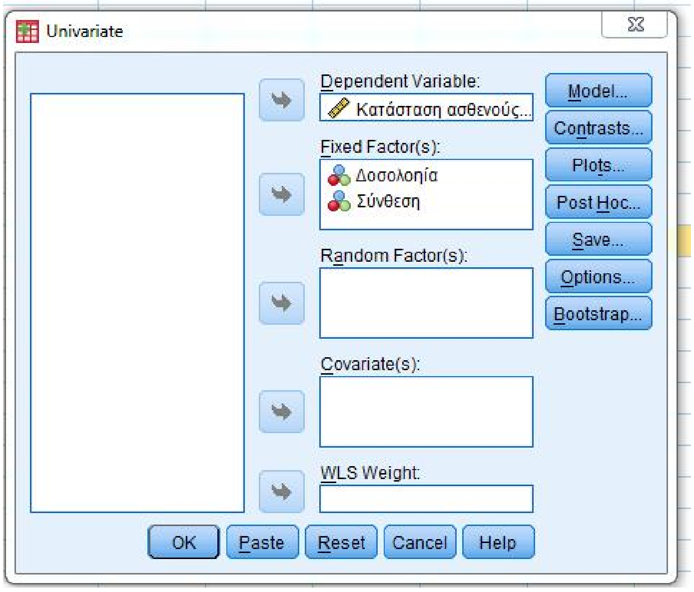


	Δοσολογία	Σύνθεση	Κατ_ασθενούς	var	var	var
10	Ανά_Εβδομάδα	Φάρμακο 1	17,00			
11	Ανά_Εβδομάδα	Φάρμακο 1	6,00			
12	Ανά_Εβδομάδα	Φάρμακο 1	10,00			
13	Ανά_Εβδομάδα	Φάρμακο 1	12,00			
14	Ανά_Εβδομάδα	Φάρμακο 1	14,00			
15	Ανά_Εβδομάδα	Φάρμακο 2	24,00			
16	Ανά_Εβδομάδα	Φάρμακο 2	15,00			
17	Ανά_Εβδομάδα	Φάρμακο 2	10,00			
18	Ανά_Εβδομάδα	Φάρμακο 2	16,00			
19	Ανά_Εβδομάδα	Φάρμακο 2	14,00			
20	Ανά_Εβδομάδα	Φάρμακο 3	14,00			
21	Ανά_Εβδομάδα	Φάρμακο 3	9,00			
22	Ανά_Εβδομάδα	Φάρμακο 3	12,00			
23	Ανά_Εβδομάδα	Φάρμακο 3	,0			
24	Ανά_Εβδομάδα	Φάρμακο 3	6,00			
25	Ανά_Εβδομάδα	Φάρμακο 4	10,00			
26	Ανά_Εβδομάδα	Φάρμακο 4	-1,00			
27	Ανά_Εβδομάδα	Φάρμακο 4	,0			
28	Ανά_Εβδομάδα	Φάρμακο 4	3,00			
29	Ανά_Εβδομάδα	Φάρμακο 4	-1,00			
30	Ανά_2_ημέρες	Φάρμακο 1	18,00			
31	Ανά_2_ημέρες	Φάρμακο 1	9,00			
32	Ανά_2_ημέρες	Φάρμακο 1	17,00			
33	Ανά_2_ημέρες	Φάρμακο 1	21,00			

Μετά την εισαγωγή των στοιχείων πατάμε Analyze **à** General Linear Model **à** Univariate



Στον πίνακα που εμφανίζεται στο Dependent Variable (εξαρτημένη μεταβλητή) εισάγουμε την μεταβλητή «Κατάσταση ασθενούς» και στο Fixed Factors (παράγοντες) εισάγουμε τις μεταβλητές «Δοσολογία» και «Σύνθεση», όπως ακριβώς κάτω. Στην επιλογή options στο significance level βάζουμε την στάθμη σημαντικότητας που θέλουμε, εδώ είναι 0,05.



Τα αποτελέσματα που παίρνουμε είναι τα εξής:

Tests of Between-Subjects Effects

Dependent Variable:Κατάσταση ασθενούς

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1231,602[a]	7	175,943	7,150	,000
Intercept	4577,673	1	4577,673	186,024	,000
Δοσολογία	8,462	1	8,462	,344	,561
Σύνθεση	710,613	3	236,871	9,626	,000
Δοσολογία * Σύνθεση	536,994	3	178,998	7,274	,001

Error	1008,929	41	24,608			
Total	7930,000	49				
Corrected Total	2240,531	48				

a. R Squared = ,550 (Adjusted R Squared = ,473)

Εδώ δεν χρειάζεται να δούμε τους παράγοντες ξεχωριστά αφού υπάρχει αλληλεπίδραση μεταξύ αυτών και φαίνεται απ' τον έλεγχο F=7,27 ($F>F_{a.k-1.n-k}$ => $F>F_{0.05.4-1.49-4}$ => F>2.81) και την p-value=0,001 (<0,05), απορρίπτουμε δηλαδή την H_0.

Κεφάλαιο ΙΧ: Έλεγχος χ^2

1. <u>Έλεγχος καλής προσαρμογής</u>:

Με τον τίτλο «έλεγχος καλής προσαρμογής» (goodness-of-fit) εννοούμε την διαδικασία (ή τις διαδικασίες) εκείνες με τις οποίες μπορούμε να ελέγξουμε αν τα δεδομένα μας ακολουθούν μια συγκεκριμένη κατανομή (π.χ Διωνυμική, Poisson, Κανονική).
Για τον έλεγχο της καλής προσαρμογής υπάρχουν διάφοροι τρόποι. Εδώ θα αναφερθούμε στην πιο γνωστή και εύκολα χρησιμοποιούμενη. Η μέθοδος αυτή είναι γνωστή με το όνομα χ^2.
Ο έλεγχος αυτός εφαρμόζεται σε ένα πολυωνυμικό πείραμα όπου είναι μία γενίκευση του δυωνυμίκου πειράματος και χρησιμοποιείται για την περιγραφή ενός πληθυσμού δεδομένων.
Για το πολυωνυμικό πείραμα:

- Αποτελείται από έναν σταθερό αριθμό, n, δοκιμών.
- Κάθε δοκιμή μπορεί να έχει ένα από τα k ενδεχόμενα, καλούμενα ως κελιά.
- Όλες οι πιθανότητες p_i είναι σταθερές.
- Η συνήθης ιδιότητα των πιθανοτήτων ισχύει: $p_1 + p_2 + ... + p_k = 1$, και
- Κάθε δοκιμή είναι ανεξάρτητη από τις υπόλοιπες δοκιμές.

Ελέγχουμε εάν υπάρχει επαρκή μαρτυρία ώστε να απορρίψουμε ένα προκαθορισμένο σύνολο τιμών για τις p_i. Πιο αναλυτικά η μηδενική υπόθεση είναι:
$H_0 : p_1 = a_1 \;\; p_2 = a_2, ..., p_k = a_k$ (όπου $a_1, a_2, ..., a_k$ είναι οι τιμές που μας ενδιαφέρουν)
Η ερευνητική υπόθεση είναι:
H_1: Τουλάχιστον ένα $p_i = a_i$
Ο έλεγχος πραγματοποιεί την σύγκριση μεταξύ των πραγματικών συχνοτήτων και των αναμενόμενων συχνοτήτων των συμβάντων στα κελιά.
Η αναμενόμενη συχνότητα για κάθε κελί είναι: $e_i = n \cdot p_i$
Ο έλεγχος καλής προσαρμογής δίνεται από τον τύπο:

$\chi^2 = \sum_{i=1}^{k} \frac{(f_i - e_i)^2}{e_i}$

,όπου η κατανομή δειγματοληψίας είναι χ^2 με ν= k-1 βαθμούς ελευθερίας.

Η περιοχή απόρριψης είναι:

$\chi^2 = \chi^2_{\alpha.k-1}$

Αν οι αναμενόμενες και οι παρατηρήσιμες συχνότητες είναι παρόμοιες, ο έλεγχος θα έχει μικρή τιμή. Αυτό εξετάζεται με την βοήθεια της περιοχής απόρριψης και της τιμής p. Η τιμή p πρέπει, δυστυχώς, να υπολογισθεί με την βοήθεια λογισμικού διότι δεν υπάρχει πίνακας που να επιτρέπει απευθείας τον υπολογισμό της.

Τέλος, για να χρησιμοποιήσουμε αυτή την τεχνική, το μέγεθος του δείγματος πρέπει να είναι αρκετά μεγάλο έτσι ώστε η αναμενόμενη τιμή για κάθε κελί είναι 5 ή μεγαλύτερη (δηλαδή $\nu \cdot p_i \geq 5$).

ΠΑΡΑΔΕΙΓΜΑ:

Πολλές επιχειρήσεις ενδιαφέρονται για την ταχύτητα εξόφλησης των λογαριασμών των πελατών που αγοράζουν με πίστωση. Ένα εργαλείο που χρησιμοποιούν είναι η ταξινόμηση των εισπρακτέων λογαριασμών (counts receivable) κατά ηλικία, δηλαδή σύμφωνα με τον χρόνο που έχει περάσει από την τιμολόγηση κάθε αγοράς που παραμένει ανεξόφλητη. Μια μεγάλη επιχείρηση έχει υπολογίσει την κατανομή των εισπρακτέων κατά ηλικία στη διάρκεια 5 ετών ως εξής:

Κωδικός	Ημέρες	Αναλογία
1	0-14	0,72
2	15-29	0,15
3	30-59	0,1
4	60 και άνω	0,03

Λόγω της οικονομικής ύφεσης η επιχείρηση επέλεξε ένα τυχαίο δείγμα 250 εισπρακτέων λογαριασμών και κατέγραψε την ηλικία τους σύμφωνα με την παραπάνω κωδικοποίηση. Να ελέγξετε αν η κατανομή των εισπρακτέων κατά ηλικία έχει αλλάξει.

Δεδομένα:

3	1	1	1	1	2	3	4	1
1	1	3	3	3	4	1	2	3
1	4	1	1	3	1	1	1	3
1	1	1	1	3	1	4	3	3
1	1	1	3	1	1	1	1	1
1	1	1	1	4	1	2	3	1

1	1	1	1	1	4	3	2	1
2	1	3	4	1	1	1	3	2
1	1	1	1	1	1	3	3	2
1	1	1	2	1	1	1	1	1
1	1	1	1	4	1	3	1	1
2	4	4	1	2	1	1	1	1
1	1	1	3	1	1	2	4	1
1	3	2	1	1	1	3	1	1
1	3	3	1	3	1	2	1	1
2	1	2	2	1	1	2	1	2
1	1	1	3	1	3	3	1	2
1	3	3	3	1	1	1	1	1
2	1	1	1	1	2	1	1	1
2	3	1	1	3	1	1	1	1
1	1	3	1	1	1	1	1	3
1	1	3	1	1	1	3	4	1
3	3	1	2	1	1	3	1	2
3	2	1	3	1	1	3	1	1
1	3	1	1	1	1	1	4	1
1	3	3	2	1	3	2	4	1
1	1	4	1	1	1	1	2	
1	1	1	1	1	1	4	1	

ΛΥΣΗ:

Η μηδενική υπόθεση αποτελείται από τις ηλικίες των 250 εισπρακτέων λογαριασμών για τους κωδικούς 1,2,3 και 4.

H_0: p_1=0,72 p_2=0,15 p_3=0,10 p_4=0,03

H_1: Μία τουλάχιστον αναλογία έχει αλλάξει.

Αν η μηδενική υπόθεση είναι αληθής, τότε ο αναμενόμενος αριθμός εισπρακτέων κατά ηλικία του δείγματος θα είναι:

e_1=250·0,72=180

e_2=250·0,15=37,5

e_3=250·0,1=25

e_4=250·0,03=7,5

Ο υπολογισμός του ελέγχου χ^2 παρατίθεται μέσω του πίνακα:

	f_i	e_i	$f_i - e_i$	$\frac{(f_i - e_i)^2}{e_i}$
1	159	180	-21	2,45
2	28	37,5	-9,5	2,41

3	47	25	22	19,36
4	16	7,5	8,5	9,63
Σύνολο	250	250		χ2=33,85

Αν οι αναμενόμενες και οι παρατηρήσιμες συχνότητες είναι παρόμοιες, ο έλεγχος θα έχει μικρή τιμή. Η περιοχή απόρριψης της κατανομής δειγματοληψίας χ^2, για k=4 και α=0,05 είναι:

$\chi^2 > \chi^2_{\alpha.k-1} = \chi^2_{0,05.3} = 7{,}81$

Ο έλεγχος που υπολογίσαμε βρίσκεται στην περιοχή απόρριψης.

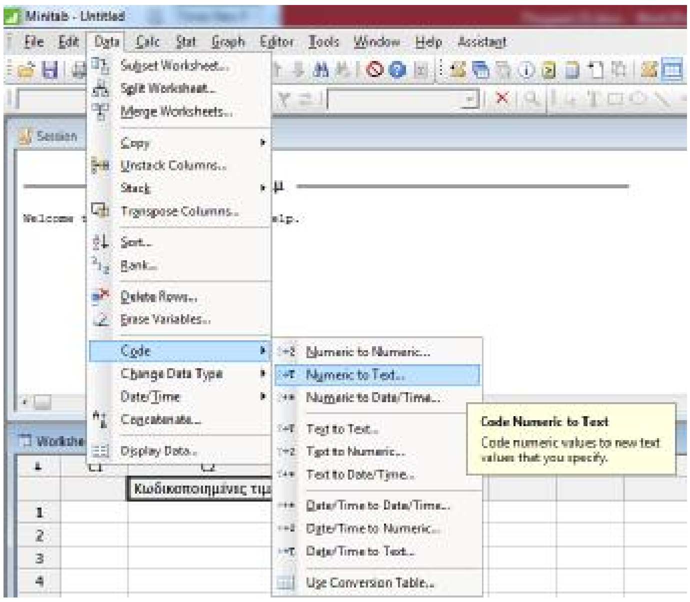
Η τιμή, τώρα, της p υπολογίζεται ως η πιθανότητα:

$P(\chi^2 > 33{,}85)$

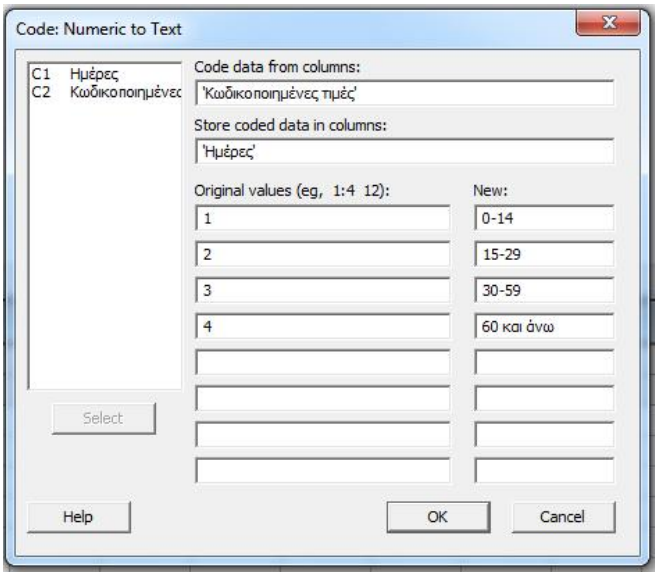
Όπως είπαμε όμως κανένας πίνακας δεν επιτρέπει τον άμεσο υπολογισμό της και χρήζει βοήθεια λογισμικού όπως και θα δούμε στην συνέχεια.

Για την συγκεκριμένη άσκηση έχουμε επιλέξει το λογισμικό Minitab.

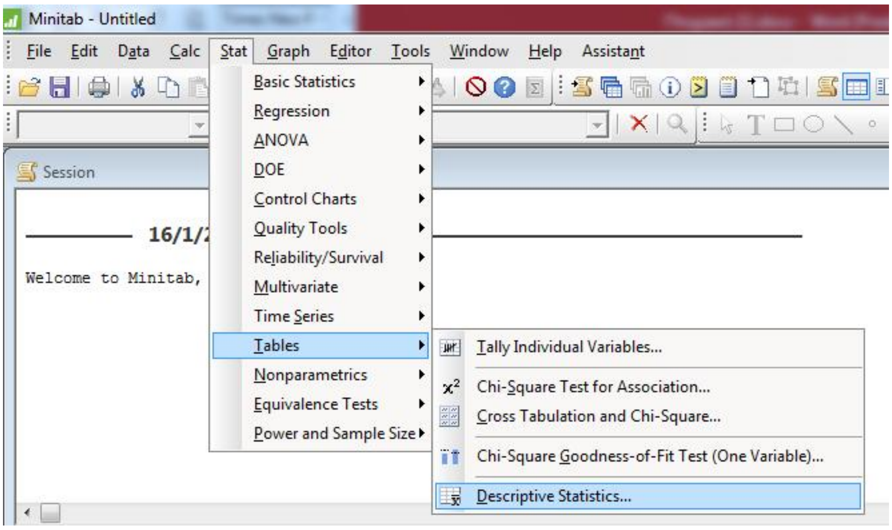
Εισάγουμε τα στοιχεία στο φύλλο εργασίας του Minitab το ένα μετά το άλλο και ονομάζουμε την μεταβλητή «Κωδικοποιημένες τιμές». Δημιουργούμε μία νέα μεταβλητή «Ημέρες» και πάμε στην γραμμή εργαλείων, επιλέγουμε Data **à** Code **à** Numeric to Text



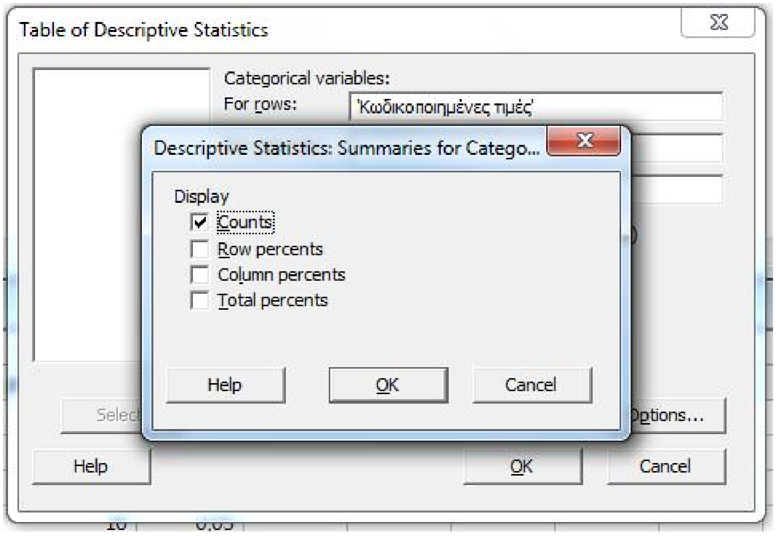
Στο Code data from columns επιλέγουμε την μεταβλητή «Κωδικοποιημένες τιμές» και στο Store coded data in columns την μεταβλητή «Ημέρες». Στο Original values πληκτρολογούμε εμείς τους κωδικούς 1,2,3,4 και στο New πληκτρολογούμε τις ημέρες (0-14, 15-29 κ.ο.κ αντίστοιχα).



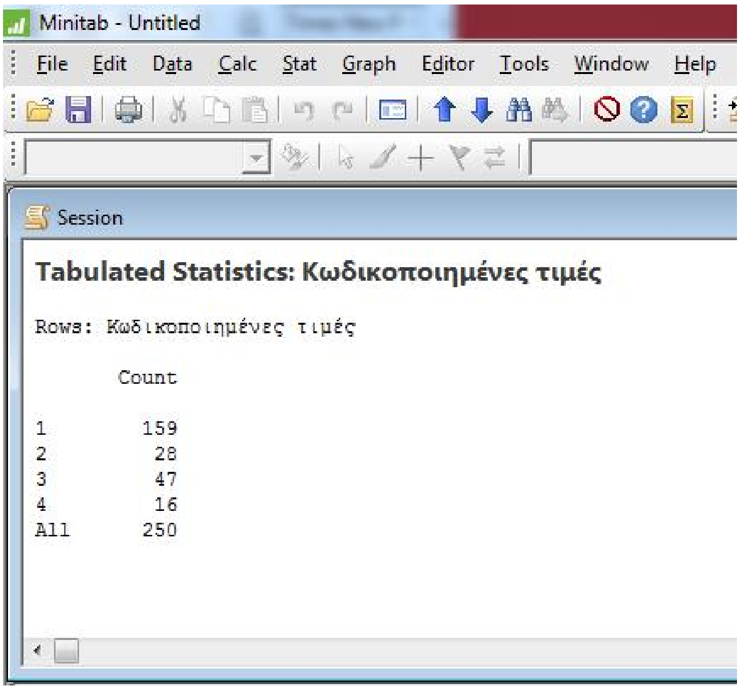
Για να μπορέσουμε στον έλεγχο χ^2 θα πρέπει να βρούμε τις συχνότητες. Πάμε στην γραμμή εργαλείων Stat **à** Tables **à** Descriptive Statistics



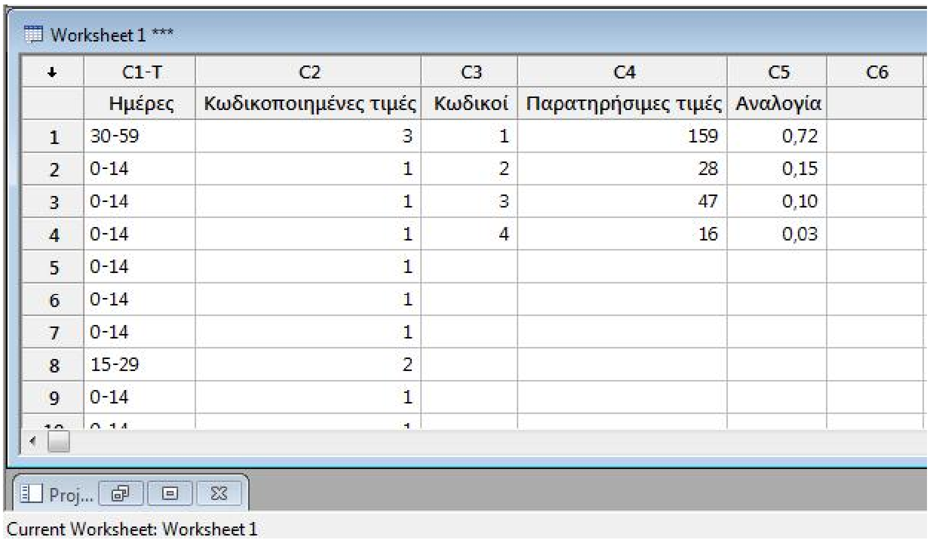
Βάζουμε την μεταβλητή «Κωδικοποιημένες τιμές» στο For rows επιλέγουμε να ανοίξει το Categorical Variables και επιλέγουμε το Counts (για να μας βγάλει τις συχνότητες).



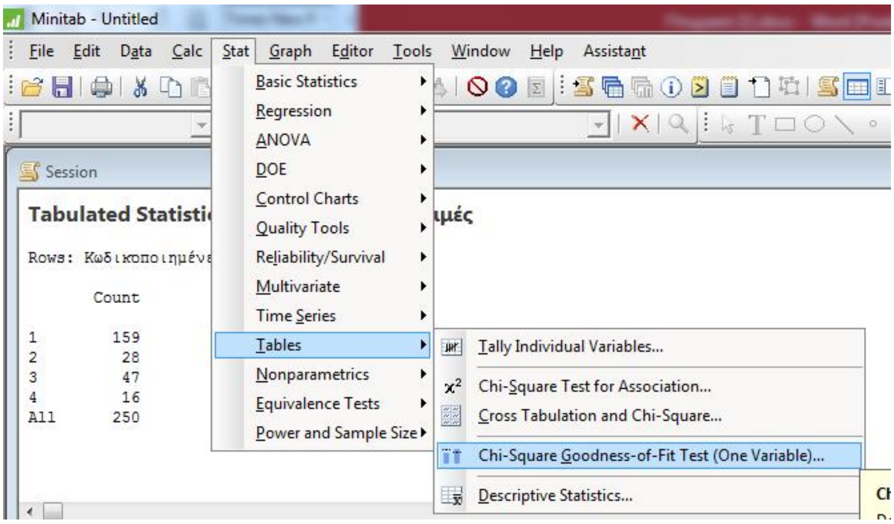
Οι συχνότητες είναι οι εξής:



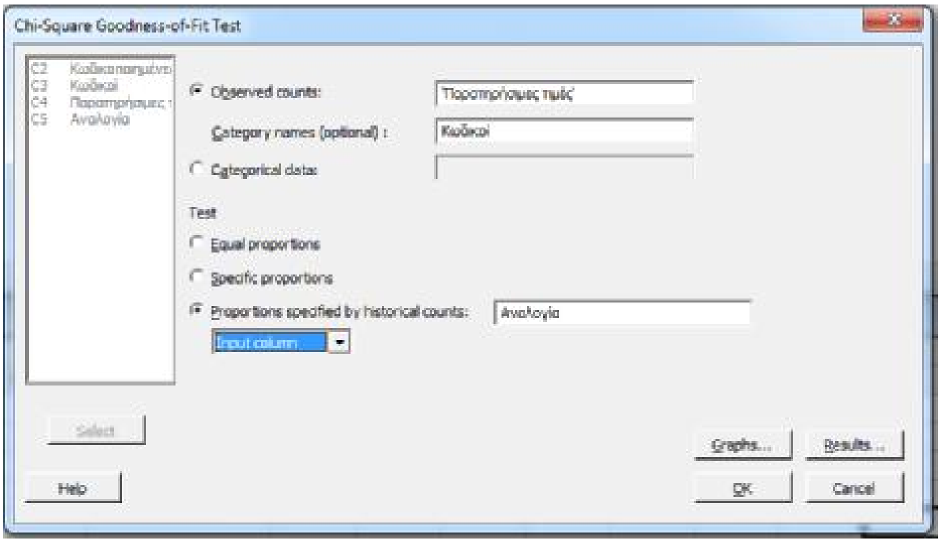
Δημιουργούμε τρεις νέες μεταβλητές: «Κωδικοί» και βάζουμε τους κωδικούς κατά σειρά 1,2,3,4, «Παρατηρήσιμες τιμές» και αντιγράφουμε τις συχνότητες που βρήκαμε και η τελευταία μεταβλητή «Αναλογία» που πληκτρολογούμε τις δοθήσες αναλογίες όπως φαίνεται παρακάτω:



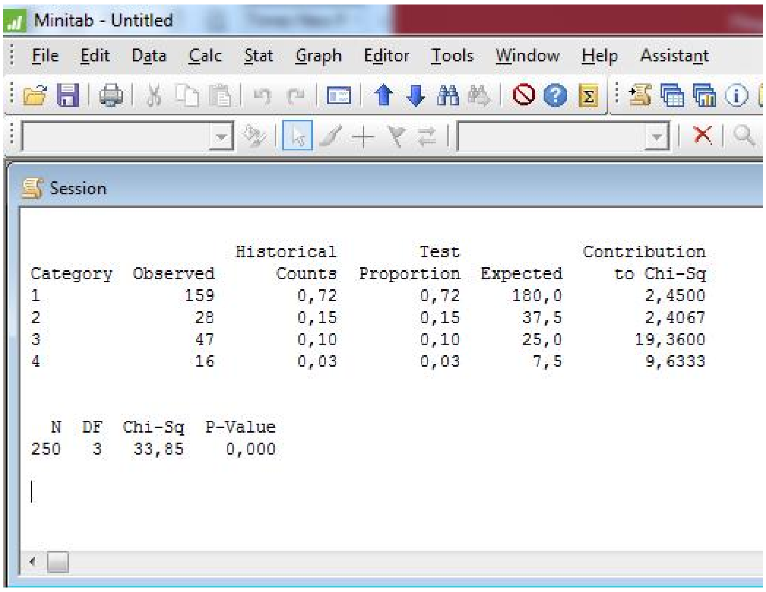
Για να βγάλουμε τα τελικά αποτελέσματα πάμε στην γραμμή εργαλείων Stat **à** Tables **à** Chi-Square Goodness-of-Fit Test (One Variable)



Στο Observed counts βάζουμε την μεταβλητή «Παρατηρήσιμες τιμές», στο Category names βάζουμε την μεταβλητή «Κωδικοί» και στο Proportions specified by historical counts επιλέγουμε την «Αναλογία» όπως και φαίνεται:

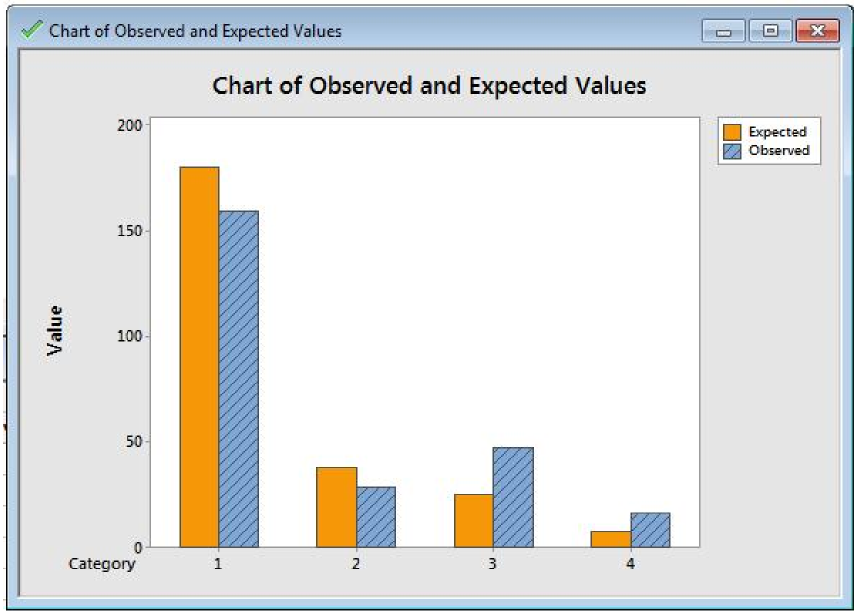


Τα αποτελέσματα είναι τα εξής:



Το χ^2=33,85 (χ^2> $\chi^2_{\alpha.k-1}$ = $\chi^2_{0,05.3}$=7,81 ,όπως είχαμε δει στην χειρόγραφη επίλυση της άσκησης πιο πάνω) και το p=0(<0,05). Άρα το χ^2 και η τιμή p βρίσκονται στην περιοχή απόρριψης. Απορρίπτουμε την μηδενική υπόθεση και μπορούμε να ισχυριστούμε ότι η κατανομή των εισπρακτέων λογαριασμών κατά ηλικία έχει αλλάξει.

Αυτό φαίνεται και στο ραβδόγραμμα για τις παρατηρήσιμες (observed) και τις αναμενόμενες (expected) συχνότητες:



2. <u>Πίνακας συνάφειας</u>:

Το κριτήριο χ^2 είναι γνωστό ως κριτήριο ελέγχου ανεξαρτησίας (chi square test of independence) ή κριτήριο ελέγχου πινάκων συνάφειας (contingency tables). Οι πίνακες συνάφειας ή πίνακες διπλής εισόδου αναπτύσσονται σε (έστω κ) γραμμές και (έστω λ) στήλες. Το κριτήριο εξετάζει αν οι δύο μεταβλητές που απαρτίζουν τον πίνακα διπλής εισόδου είναι ανεξάρτητες ή όχι. Πρόκειται ουσιαστικά για στατιστικό έλεγχο ο οποίος βασίζεται στη χρήση της στατιστικής κατανομής χ^2 σε επίπεδο σημαντικότητας α. Το σκεπτικό του ελέγχου είναι το εξής: Αρχικά υποθέτουμε ότι δεν υπάρχει σχέση μεταξύ των μεταβλητών (δηλαδή είναι ανεξάρτητες οι μεταβλητές μεταξύ τους). Μπορούμε να υπολογίσουμε τότε τις αναμενόμενες συχνότητες (δηλαδή τις συχνότητες που αναμένουμε αν ισχύει η αρχική υπόθεση ότι δεν υπάρχει σχέση). Τις συγκρίνουμε με τις πραγματικές συχνότητες. Αν διαφέρουν πολύ τότε απορρίπτουμε την υπόθεση της ανεξαρτησίας και λέμε ότι υπάρχει σχέση μεταξύ των 2 μεταβλητών.
Γενικά, υπάρχουν δύο τρόποι λύσεων των πινάκων συνάφειας. Ο ένας είναι να θεωρήσουμε ότι ο πίνακας συνάφειας αντιπροσωπεύει δύο μεταβλητές και σκοπός μας είναι να ελέγξουμε αν οι δύο μεταβλητές σχετίζονται μεταξύ τους. Ο άλλος τρόπος είναι να θεωρήσουμε τις ομάδες της μιας μεταβλητής ως πληθυσμούς (ένας πληθυσμός=μία ομάδα) και να ελέγξουμε αν οι αναλογίες διαφέρουν μεταξύ των πληθυσμών, σκοπός μας είναι, δηλαδή, η σύγκριση των πληθυσμών των ονομαστικών δεδομένων. Παρ’ όλα αυτά και οι δύο προσεγγίσεις τέτοιου είδους προβλημάτων καταλήγουν στην ουσία στον ίδιο στατιστικό έλεγχο. Καταλήγοντας στο ότι μπορούμε να επιλέξουμε εμείς όποια προσέγγιση επιθυμούμε.

<u>ΠΑΡΑΔΕΙΓΜΑ</u>:

Η πιθανότητα διαπλοκής μεταξύ ιατρικής έρευνας και φαρμακευτικών εταιριών είναι ένα πολύ σημαντικό θέμα, που προκαλεί μεγάλο ενδιαφέρον. Το έναυσμα ήταν μια μελέτη του 1995, η οποία διαπίστωσε ότι η χρήση της οικογένειας φαρμάκων CCB (Calcium Channel Blockers – ανταγωνιστές ασβεστίου) για την αντιμετώπιση της υπέρτασης αυξάνει τον κίνδυνο καρδιακής προσβολής. Το επιστημονικό περιοδικό New England Journal of Medicine (8-1-1998, σελ. 101) συγκέντρωσε 70 εργασίες που είχαν δημοσιευθεί το 1996-1997 για την οικογένεια φαρμάκων CCB, τις χαρακτήρισε ως 1=ευνοϊκές, 2=ουδέτερες, 3=αρνητικές και στη συνέχεια έλεγξε αν οι συγγραφείς είχαν οικονομικούς δεσμούς με φαρμακευτικές εταιρίες (1=ΝΑΙ, 2=ΟΧΙ). Μπορούμε από τα δεδομένα να συμπεράνουμε ότι τα πορίσματα των δημοσιεύσεων επηρεάζονται από τους οικονομικούς δεσμούς με φαρμακευτικές εταιρίες;

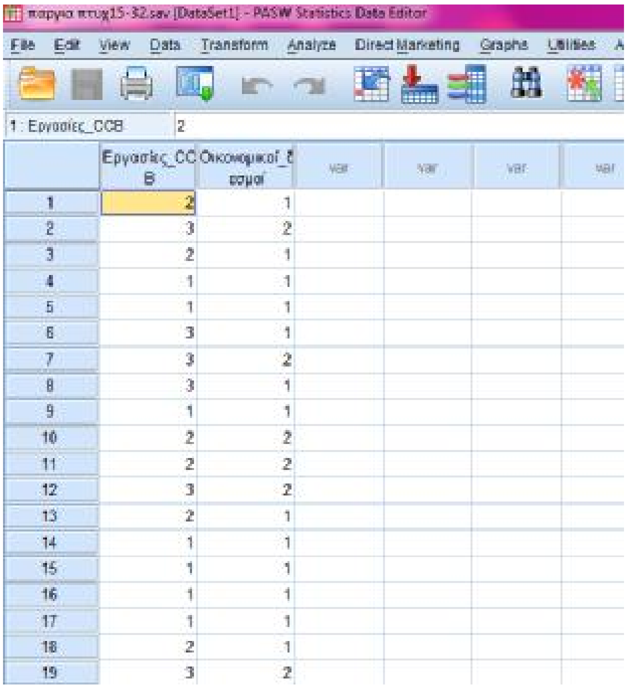
Δεδομένα:

Εργασίες CCB			Οικονομικοί δεσμοί		
2	2	2	1	1	1
3	3	1	2	2	1
2	3		1	2	
1	3		1	2	
1	1		1	1	
3	1		1	1	
3	1		2	1	
3	1		1	1	
1	3		1	2	
2	1		2	2	
2	1		2	1	
3	2		2	2	
2	3		1	2	
1	1		1	1	
1	1		1	1	
1	2		1	1	
1	2		1	2	
2	3		1	2	
3	2		2	2	
1	1		1	1	
1	1		1	1	
1	3		1	1	
3	2		2	2	
1	1		1	1	
1	3		1	1	
3	3		1	1	
3	3		2	2	
1	3		1	2	
1	1		1	1	
3	1		1	1	

2	3		2	1	
2	2		1	1	
1	2		1	1	
1	3		1	1	

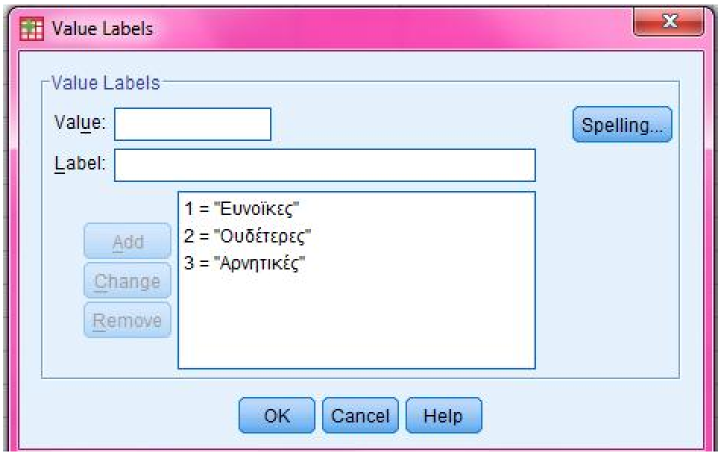
<u>ΛΥΣΗ:</u>

Η επίλυση του προβλήματος θα γίνει στο λογισμικό SPSS.
Εισάγουμε σε δύο στήλες τα δεδομένα, οι μεταβλητές θα ονομαστούν «Εργασίες_CCB» και «Οικονομικοί_δεσμοί» όπως δόθηκαν.

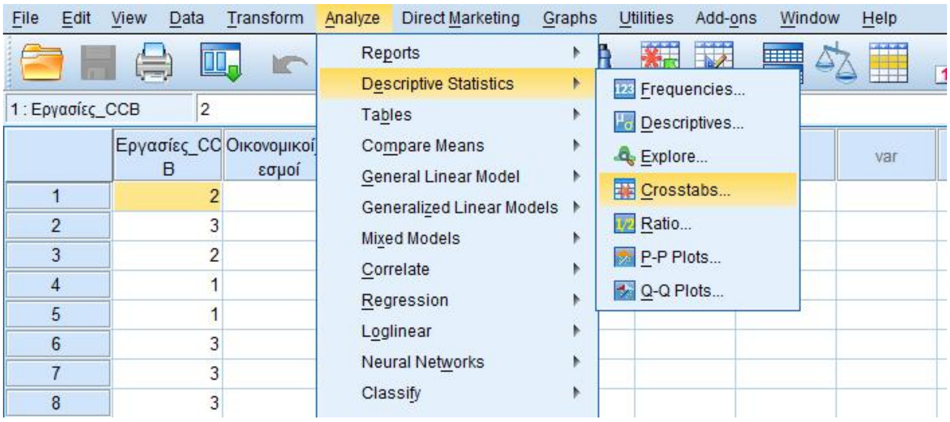


	Εργασίες_CCB	Οικονομικοί_δεσμοί
1	2	1
2	3	2
3	2	1
4	1	1
5	1	1
6	3	1
7	3	2
8	3	1
9	1	1
10	2	2
11	2	2
12	3	2
13	2	1
14	1	1
15	1	1
16	1	1
17	1	1
18	2	1
19	3	2

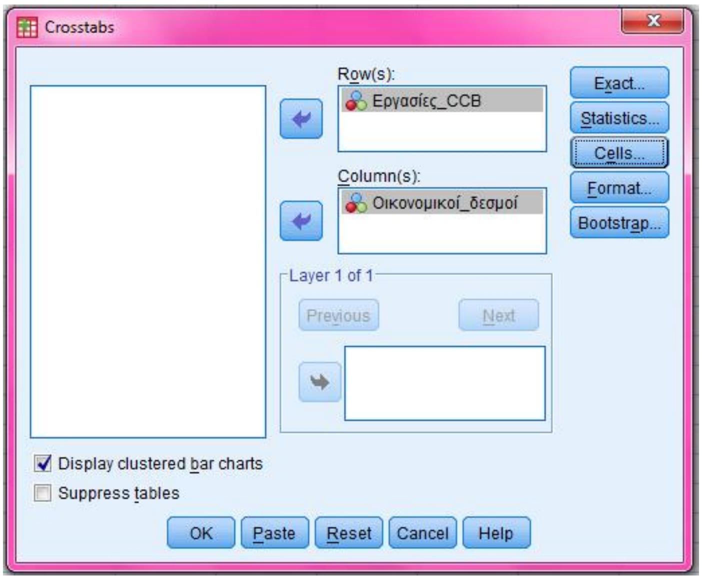
Για την κάθε μεταβλητή περνάμε τους χαρακτηρισμούς που αντιστοιχούν σε κάθε κωδικό. Στην καρτέλα Variable View πατάμε για την κάθε μεταβλητή το Value labels και εισάγουμε τους χαρακτηρισμούς όπως φαίνεται παρακάτω:

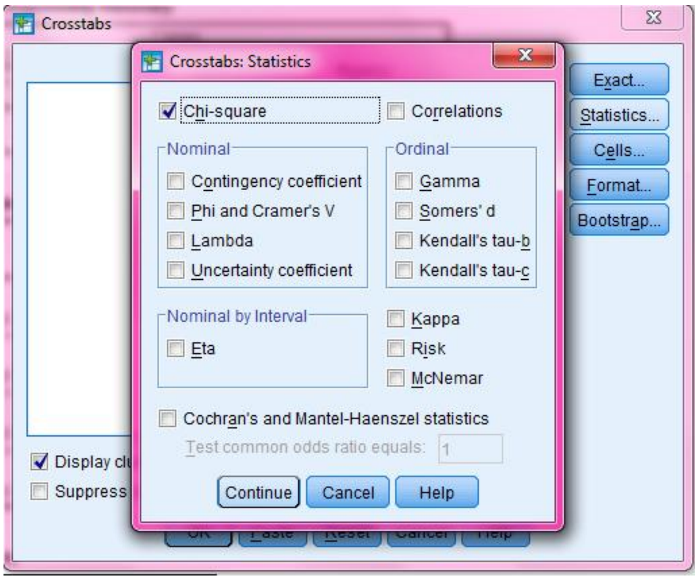


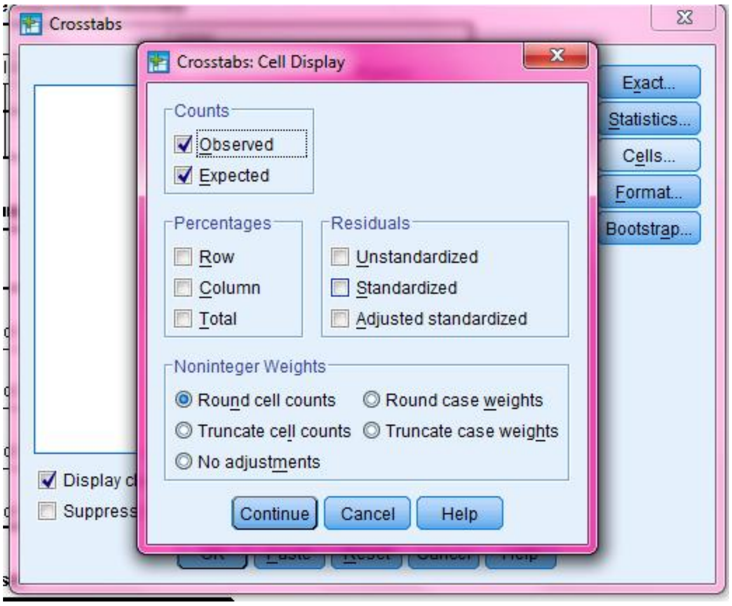
Για τον υπολογισμό του ελέγχου πάμε στην γραμμή εργαλείων Analyze **à** Descriptive Statistics **à** Crosstabs.



Στο Row(s) περνάμε την μεταβλητή «Εργασίες_CCB» και στο Column(s) την μεταβλητή «Οικονομικοί_δεσμοί». Επιλέγουμε το Display clustered bar charts για να εμφανιστεί το ραβδόγραμμα. Στην επιλογή Statistics επιλέγουμε το κριτήριο χ^2 ως Chi-square και στην επιλογή Cells επιλέγουμε το Observed και το Expected για να εμφανιστούν οι παρατηρηθείσες και οι αναμενόμενες συχνότητες, καθώς και το Unstandardized για τη διαφορά παρατηρούμενων – αναμενόμενων συχνοτήτων.


Crosstabs
Row(s):
Εργασίες_CCB
Column(s):
Οικονομικοί_δεσμοί
Layer 1 of 1
Previous
Next
Exact...
Statistics...
Cells...
Format...
Bootstrap...
Display clustered bar charts
Suppress tables
OK
Paste
Reset
Cancel
Help


Crosstabs
Crosstabs: Statistics
Chi-square
Correlations
Nominal
Contingency coefficient
Phi and Cramer's V
Lambda
Uncertainty coefficient
Ordinal
Gamma
Somers' d
Kendall's tau-b
Kendall's tau-c
Nominal by Interval
Eta
Kappa
Risk
McNemar
Cochran's and Mantel-Haenszel statistics
Test common odds ratio equals:
1
Continue
Cancel
Help
Exact...
Statistics...
Cells...
Format...
Bootstrap...



Πατώντας ok εμφανίζονται τα αποτελέσματα στους παρακάτω πίνακες:

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Εργασίες_CCB * Οικονομικοί_δεσμοί	70	100,0%	0	,0%	70	100,0%

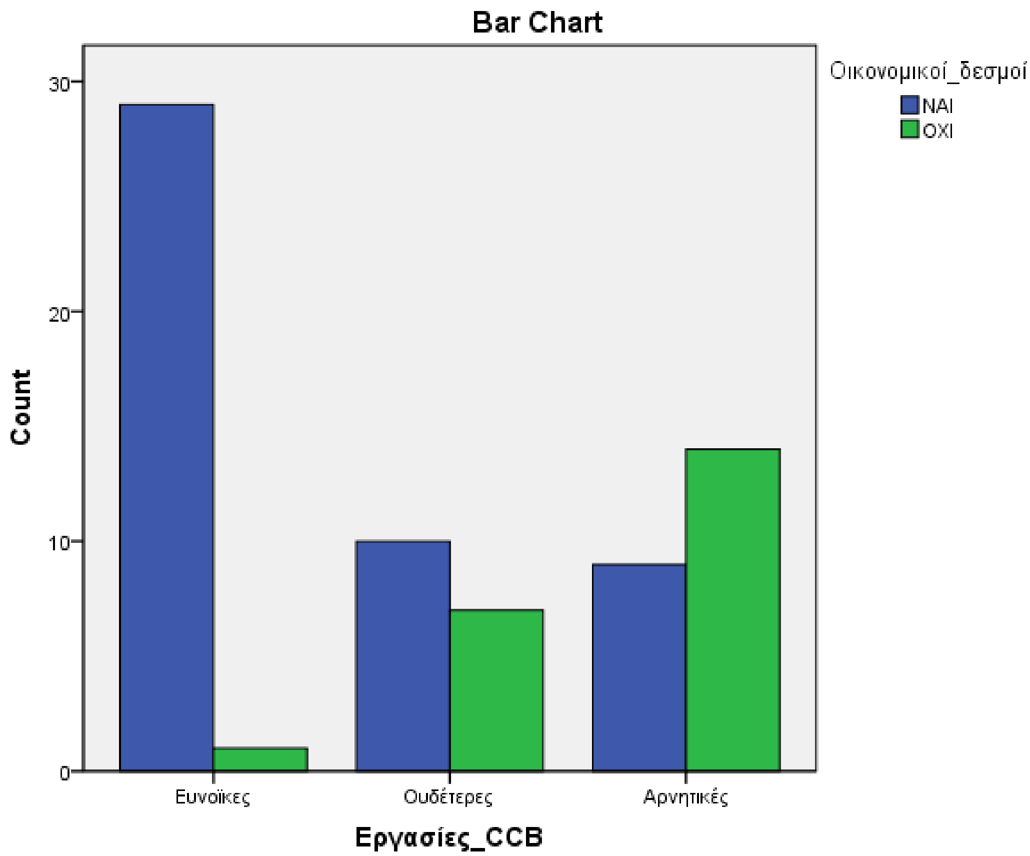
Σε αυτό τον πίνακα μας δίνεται ο ακριβείς αριθμός των δεδομένων. Το missing μας δείχνει τα ελλιπή στοιχεία και σε αυτό το παράδειγμα δεν έχουμε.

Ο πίνακας αυτός μας δείχνει τις παρατηρούμενες (Count) και τις αναμενόμενες (Expected Count) συχνότητες.
Θα εξηγήσουμε την πρώτη γραμμή με τις «Εργασίες CCB» χαρακτηρισμένες ως «ευνοϊκές».
29 φορές οι εργασίες των φαρμάκων CCB που είναι ευνοϊκές έχουν οικονομικούς δεσμούς με τις φαρμακευτικές εταιρίες.
20,6 φορές είναι η αναμενόμενη συχνότητα όπου οι συγγραφείς με ευνοϊκές εργασίες έχουν οικονομικούς δεσμούς με τις φαρμακευτικές εταιρίες.

Εργασίες_CCB * Οικονομικοί_δεσμοί Crosstabulation

			Οικονομικοί_δεσμοί		Total
			NAI	OXI	
Εργασίες_CCB	Ευνοϊκες	Count	29	1	30
		Expected Count	20,6	9,4	30,0
		Residual	8,4	-8,4	
	Ουδέτερες	Count	10	7	17
		Expected Count	11,7	5,3	17,0
		Residual	-1,7	1,7	
	Αρνητικές	Count	9	14	23
		Expected Count	15,8	7,2	23,0
		Residual	-6,8	6,8	
Total		Count	48	22	70
		Expected Count	48,0	22,0	70,0

1 φορά οι ευνοϊκές εργασίες CCB δεν είχαν οικονομικούς δεσμούς με τις φαρμακευτικές εταιρίες.
9,4 φορές είναι η αναμενόμενη συχνότητα όπου οι συγγραφείς με ευνοϊκές εργασίες δεν έχουν οικονομικούς δεσμούς με τις φαρμακευτικές εταιρίες.
Ομοίως και για τα υπόλοιπα.
Ήδη βλέπουμε ότι οι παρατηρούμενες με τις αναμενόμενες τιμές έχουν αρκετή διαφορά είτε θετική είτε αρνητική, το οποίο δίνεται από την τιμή Residual, οπότε μπορούμε να συμπεράνουμε από τώρα ότι η μηδενική υπόθεση είναι αναληθής. Η διαφορά αυτή δίνεται καλύτερα από το ραβδόγραμμα όπου φαίνονται διαγραμματικά οι συχνότητες.



Στον επόμενο πίνακα που μας δίνει το SPSS καταλήγουμε σε σίγουρο αποτέλεσμα αν δεν μπορούμε να κατανοήσουμε τις ήδη υπάρχουσες διαφορές.

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	20,988[a]	2	,000
Likelihood Ratio	24,556	2	,000
Linear-by-Linear Association	20,205	1	,000
N of Valid Cases	70		

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 5,34.

Η βαθμοί ελευθερίας σε πίνακα συνάφειας με r γραμμές και c στήλες, είναι v=(r-1)·(c-1). Στο παράδειγμά μας r=3 και c=2. Άρα:
v = (r-1) ·(c-1) = (3-1) · (2-1) = 2·1 = 2

Με στάθμη σημαντικότητας α=0,05 η περιοχή απόρριψης είναι:
$\chi^2 > \chi^2_{\alpha.\kappa\text{-}1} = \chi^2_{0,05.6} = 12{,}6$
Ο έλεγχος που υπολογίσαμε είναι χ^2=20,988 (>12,6) και η τιμή p=0(>0.05), βρίσκονται και τα δύο στην περιοχή απόρριψης, έτσι απορρίπτουμε την H_0 και μπορούμε να συμπεράνουμε ότι όντως τα πορίσματα των δημοσιεύσεων επηρεάζονται από τους οικονομικούς δεσμούς με φαρμακευτικές εταιρίες, άρα υπάρχει εξάρτηση μεταξύ των δύο μεταβλητών.

Κεφάλαιο XII: Μη παραμετρικές μέθοδοι

Πολλές παραμετρικές στατιστικές μέθοδοι ελέγχου υποθέσεων στηρίζονται στην υπόθεση ότι η μεταβλητότητα των δεδομένων περιγράφεται από μία κατανομή συγκεκριμένης μορφής (για παράδειγμα, κανονική ή κάποια άλλη κατανομή που δεν απέχει πολύ από την κανονική). Έτσι, στους ελέγχους υποθέσεων που αναφέρονται στην μέση τιμή ενός πληθυσμού, γίνεται η υπόθεση ότι ο δειγματικός μέσος κατανέμεται κανονικά. Τέλος, στους ελέγχους υποθέσεων στο πλαίσιο προβλημάτων ανάλυσης διασποράς, υποθέτουμε ότι η εντός των επιδράσεων μεταβλητότητα περιγράφεται από την κανονική κατανομή. Οποτεδήποτε δεν είναι επιτρεπτή η οποιαδήποτε υπόθεση για την μορφή του πληθυσμού, ο ερευνητής στρέφεται στο Κεντρικό Οριακό Θεώρημα και σε εμπειρικές μελέτες για την θεωρητική και εμπειρική υποστήριξη αυτής της "βολικής" υπόθεσης κανονικότητας προκειμένου να χρησιμοποιήσει μια παραμετρική μέθοδο. Στην πράξη, όμως, εμφανίζονται συχνά περιπτώσεις στις οποίες τα δείγματα είναι μικρά και τα δεδομένα κατανέμονται εμφανώς μη κανονικά. Εμφανίζονται επίσης περιπτώσεις, όπου και αν ακόμα υπάρχει κάποια ένδειξη κανονικότητας, ο ερευνητής που ενδεχομένως έχει βαθιά γνώση του πληθυσμού, είναι επιφυλακτικός στο να κάνει μια τέτοια υπόθεση. Οι στατιστικοί επινόησαν αρκετές εναλλακτικές τεχνικές για τον ερευνητή ο οποίος διστάζει να υποθέσει κανονικότητα. Οι τεχνικές αυτές δεν απαιτούν συγκεκριμένες υποθέσεις για την μορφή του πληθυσμού από τον οποίο προέρχονται τα δεδομένα και μπορούν να χρησιμοποιηθούν τόσο για μικρά όσο και για μεγάλα δείγματα. Είναι δηλαδή σχεδιασμένες για να μπορούν να χρησιμοποιηθούν ανεξάρτητα από την κατανομή των δεδομένων. Επομένως, οι τεχνικές αυτές μπορούν να χρησιμοποιηθούν τόσο για κανονικούς πληθυσμούς όσο και για μη κανονικούς πληθυσμούς. Για τον λόγο αυτό, οι τεχνικές αυτές ονομάζονται ελεύθερες κατανομών ή μη παραμετρικές τεχνικές (distribution-free ή non-parametric techniques). Οι μη παραμετρικές τεχνικές είναι πολύ απλές στην χρήση τους εν γένει, και ο σχεδιασμός τους είναι αποτέλεσμα στοιχειωδών θεωρήσεων. Εάν τα δεδομένα ακολουθούν στην πραγματικότητα την κανονική κατανομή, τότε οι μη παραμετρικοί έλεγχοι υποθέσεων δεν είναι το ίδιο ισχυροί όπως οι αντίστοιχοι παραμετρικοί έλεγχοι, οι οποίοι κάνουν χρήση της υπόθεσης της κανονικότητας. Για δεδομένη πιθανότητα σφάλματος τύπου I, οι μη παραμετρικοί έλεγχοι έχουν υψηλότερη πιθανότητα σφάλματος τύπου II. Ένας έλεγχος, ο οποίος αγνοεί πληροφορίες σχετικά με τα δεδομένα, όπως είναι η μορφή της κατανομής τους, δεν αναμένεται να είναι το ίδιο καλός όπως ένας έλεγχος ο οποίος κάνει χρήση αυτής της πληροφορίας. Έτσι, ένας έλεγχος, ο οποίος δεν λαβαίνει υπόψη του ότι τα δεδομένα προέρχονται από μια κανονική κατανομή, δεν αναμένεται να είναι το ίδιο ισχυρός όπως ένας έλεγχος ο οποίος χρησιμοποιεί αυτή την υπόθεση. Από το άλλο μέρος, εάν τα δεδομένα προέρχονται από μη κανονικό πληθυσμό, τότε οι μη παραμετρικοί έλεγχοι έχουν ένα σαφές πλεονέκτημα έναντι των παραμετρικών ελέγχων, οι οποίοι στηρίζονται στην εσφαλμένη υπόθεση της κανονικότητας των δεδομένων. Η σοβαρότητα του σφάλματος και, επομένως, η ακρίβεια των παραμετρικών ελέγχων εξαρτάται, κατά συνέπεια, από το πόσο εσφαλμένη είναι η υπόθεση της κανονικότητας. Επειδή οι μη παραμετρικοί έλεγχοι στηρίζονται σε ελάχιστες υποθέσεις για τους πληθυσμούς από τους οποίους προέρχονται τα δεδομένα και, πάντως, όχι σε υποθέσεις οι οποίες αναφέρονται στην μορφή των πληθυσμών αυτών, είναι εν γένει πολύ ευσταθείς (robust). Συνοψίζοντας, οι μη παραμετρικές μέθοδοι α) αποβλέπουν σε ευρύτερα πεδία εφαρμογής λόγω του ότι οι κατανομές στις οποίες αναφέρονται είναι λιγότερο περιορισμένες από ό,τι στα αντίστοιχα παραμετρικά προβλήματα, β) δεν είναι εξίσου ισχυρές με τις αντίστοιχες παραμετρικές μεθόδους και γ) είναι περισσότερο ευσταθείς επειδή ακριβώς δεν επηρεάζονται από την μορφή της κατανομής των δεδομένων. Παρ' όλα αυτά, οι μη παραμετρικές μέθοδοι συχνά είναι σχεδόν το ίδιο αποτελεσματικές όπως οι παραμετρικές μέθοδοι οι οποίες κάνουν αυστηρές υποθέσεις για τον πληθυσμό από τον οποίο προέρχονται τα δεδομένα. Ένα άλλο σημαντικό πλεονέκτημα των μη παραμετρικών μεθόδων είναι ότι μπορούν να εφαρμοσθούν σε δεδομένα που είναι ταξινομημένα σε κατηγορίες (κατηγορικά δεδομένα) και τα οποία είναι σε κλίμακα διάταξης ή ακόμα και απλώς σε ονομαστική κλίμακα, ενώ οι παραμετρικές μέθοδοι προϋποθέτουν ακριβείς μετρήσεις. Τέλος, θα πρέπει να σημειωθεί ότι οι μη παραμετρικές μέθοδοι μπορούν να θεωρηθούν ως προπαρασκευαστικές για τις παραμετρικές μεθόδους, με την έννοια ότι, η χρησιμοποίηση μιας παραμετρικής μεθόδου, η οποία βασίζεται στην υπόθεση της κανονικότητας, θα πρέπει να έπεται ενός ελέγχου, με μία μη παραμετρική μέθοδο, της υπόθεσης ότι τα δεδομένα έχουν προέλθει από μια κανονική κατανομή.

Εδώ θα εξετάσουμε ορισμένες τεχνικές ανάλυσης δεδομένων, οι οποίες βασίζονται στην τάξη μεγέθους των παρατηρήσεων ενός ζεύγους δειγμάτων και όχι στις παρατηρήσεις αυτές καθαυτές. Οι *βαθμοί (τάξεις μεγέθους, ranks)* των δεδομένων προτιμώνται συχνά από τα δεδομένα για διάφορους λόγους. Ο πρώτος λόγος μπορεί να είναι ότι οι τιμές των παρατηρήσεων δεν έχουν νόημα από μόνες τους παρά μόνο όταν θεωρείται η διάταξή τους σε σχέση με τις υπόλοιπες παρατηρήσεις, οπότε οι τιμές δεν εμπεριέχουν περισσότερες πληροφορίες από ό,τι οι τάξεις μεγέθους τους. Αυτή είναι η φύση των δεδομένων που βρίσκονται σε κλίμακα διάταξης (ordinal scale). Ο δεύτερος λόγος μπορεί να είναι ότι οι τιμές των παρατηρήσεων έχουν μεν νόημα από μόνες τους, αλλά η συνάρτηση κατανομής τους δεν είναι κανονική, οπότε η Θεωρία Πιθανοτήτων δεν προσφέρεται για τον προσδιορισμό της κατανομής της στατιστικής συνάρτησης που χρησιμοποιείται για τον έλεγχο υποθέσεων. Αντίθετα, η Θεωρία Πιθανοτήτων, που απαιτείται για τον έλεγχο υποθέσεων με την χρήση στατιστικών συναρτήσεων που βασίζονται στις τάξεις μεγέθους των παρατηρήσεων, είναι σχετικά απλή και δεν εξαρτάται από την κατανομή από την οποία έχουν προέλθει οι παρατηρήσεις σε πολλές περιπτώσεις.

Παρακάτω θα γίνει η επίλυση τέτοιων προβλημάτων για διάφορες περιπτώσεις δεδομένων με τα τεστ ή τους μη παραμετρικούς ελέγχους Wilcoxon, Kruskal-Wallis, Friedman και Spearman.

1. Τεστ Wilcoxon:

Ο έλεγχος Wilcoxon χρησιμοποιείται για τον έλεγχο υποθέσεων που αναφέρονται σε παραμέτρους κεντρικής τάσης. Εφαρμόζεται στις περιπτώσεις όπου έχουμε ένα μοναδικό δείγμα παρατηρήσεων, όπως επίσης και στις περιπτώσεις όπου το δείγμα μας αποτελείται από ζεύγη παρατηρήσεων, κατάσταση που οδηγεί σε ένα μοναδικό δείγμα, το δείγμα των διαφορών των μελών των αρχικών ζευγών παρατηρήσεων. (Το ζεύγος παρατηρήσεων (n_1, n_2) αποτελεί στην πραγματικότητα μία μοναδική παρατήρηση πάνω σε μία διδιάστατη τυχαία μεταβλητή). Ο έλεγχος του Wilcoxon έχει ως παραμετρικό ανάλογο τον έλεγχο T (Μοιάζει με το *t*-test, μετά από ιεράρχηση των τιμών των δύο μεταβλητών).
Ο έλεγχος, που προτάθηκε από τον Wilcoxon (το 1945), αναπτύχθηκε για τον έλεγχο της υπόθεσης ότι ένα συγκεκριμένο δείγμα προέρχεται από κάποιον πληθυσμό με μία συγκεκριμένη διάμεσο. Όπως επίσης αναφέρθηκε νωρίτερα, ο έλεγχος αυτός μπορεί να χρησιμοποιηθεί και σε περιπτώσεις όπου το δείγμα αποτελείται από ζεύγη παρατηρήσεων, όπως, για παράδειγμα, στην περίπτωση παρατηρήσεων "*πριν*" (n_1) και "*μετά*" (n_2) που έχουν γίνει πάνω σε κάθε ένα από n άτομα με σκοπό να ελεγχθεί αν η δεύτερη τυχαία μεταβλητή στο ζεύγος (n_1, n_2) έχει την ίδια διάμεσο όπως και η πρώτη.
Βασικές προϋποθέσεις των δεδομένων για την επίλυση προβλημάτων με το test Wilcoxon:
α) Δεν έχουμε paired data δηλαδή τους ίδιους αποκρινόμενους για διαφορετικές ερωτήσεις αλλά δύο ομάδες αποκρινόμενων μεγέθους n_1 για το γκρουπ 1 και n_2 για το γκρουπ 2.
β) Τα δύο δείγματα πρέπει να είναι ανεξάρτητα και οι τιμές στην κλίμακα να είναι τακτικές
γ) Να έχουν κατανομές με ίδια κλίση
δ) Οι παρατηρήσεις να είναι μικρότερες από 20 ($n<20$).
Ο έλεγχος υποθέσεων είναι:

H_0: Οι θέσεις των δύο πληθυσμών είναι ίδιες

H_1: Ο πληθυσμός 1 βρίσκεται δεξιά του πληθυσμού 2.

Για να απορρίψουμε τη μηδενική υπόθεση θα πρέπει ο έλεγχος z να βρίσκεται στην περιοχή απόρριψης ($z>z_\alpha$).
Ο έλεγχος που χρησιμοποιείται στο τεστ Wilcoxon είναι το άθροισμα τάξεων (rank sum) των παρατηρήσεων ενός δείγματος σε μια ενιαία διάταξη. Για να τον υπολογίσουμε συγχωνεύουμε τις τιμές των δύο σε έναν ενιαίο πίνακα κατά αύξουσα σειρά. Το άθροισμα των τάξεων για το δείγμα 1 είναι το T_1 και για το δείγμα 2 είναι το T_2, μπορούμε να χρησιμοποιήσουμε ως έλεγχο οποιοδήποτε από τα δύο αθροίσματα. Ο μέσος όρος υπολογίζεται από τον τύπο:

$$E(T) = \frac{n_1 \cdot (n_1+n_2+1)}{2}$$

Η τυπική απόκλιση υπολογίζεται από τον τύπο:

$$\sigma_T = \sqrt{\frac{n_1 \cdot n_2 \cdot (n_1+n_2+1)}{12}}$$

Με βάση τους τύπους αυτούς, μπορεί να χρησιμοποιηθεί ο έλεγχος:

$z = \frac{T-E(T)}{\sigma_T}$, που ακολουθεί την τυποποιημένη κανονική κατανομή.

(Οι τύποι αυτοί μπορούν να χρησιμοποιηθούν για δείγματα με μέγεθος μεγαλύτερο του 10 όπου αποδεικνύεται μαθηματικά ότι προσεγγίζουν την κανονική κατανομή).

Ακολουθεί παράδειγμα για καλύτερη κατανόηση των εννοιών, για την επίλυση αυτού θα χρησιμοποιηθεί το λογισμικό Minitab.

ΠΑΡΑΔΕΙΓΜΑ

Μια έρευνα του Πανεπιστημίου Ohio State University (Πηγή: Working Mother, Απρίλιος 1992) σχετικά με τον τρόπο που οι γυναίκες επηρεάζονται από την εμφάνιση άλλων γυναικών, επέλεξε δύο τυχαία δείγματα 20 γυναικών και τους ζήτησε να βαθμολογήσουν τις επαγγελματικές ικανότητες μιας άλλης γυναίκας σύμφωνα με την παρακάτω κλίμακα:

Κλίμακα	Ερμηνεία
4	Εξαιρετική επαγγελματίας
3	Καλή επαγγελματίας
2	Μέτρια επαγγελματίας
1	Καθόλου επαγγελματίας

Στο πρώτο δείγμα παρουσιάστηκε μια λεπτή γυναίκα με φόρεμα μεγέθους 6 και στο δεύτερο δείγμα παρουσιάστηκε μια εύσωμη με φόρεμα μεγέθους 14. Μπορούμε από τα δεδομένα να συμπεράνουμε ότι η εμφάνιση επηρεάζει την κρίση των γυναικών; (Θεωρείται ότι η στάθμη σημαντικότητας είναι 5%)
Δεδομένα:

Μέγεθος 6	Μέγεθος 14
2	2
3	4
3	3
2	1
3	2
4	3
3	3
3	2
4	3
4	4
2	3
3	4
4	3
4	4
1	4
3	2
4	3
3	1
4	2
2	3

ΛΥΣΗ

Θα θεωρήσουμε ως δείγμα 1 αυτούς που βαθμολόγησαν τις επαγγελματικές ικανότητες της γυναίκας με φόρεμα μεγέθους 6 και ως δείγμα 2 αυτούς που βαθμολόγησαν τις επαγγελματικές ικανότητες της γυναίκας με φόρεμα μεγέθους 12.

Η υπόθεση που πρέπει να ελεγχθεί είναι:

H_0: Οι θέσεις των πληθυσμών είναι ίδιες

H_1: Ο πληθυσμός 1 βρίσκεται δεξιά του πληθυσμού 2.

Για να απορρίψουμε τη μηδενική υπόθεση θα πρέπει ο έλεγχος z να βρίσκεται στην περιοχή απόρριψης: $Z>Z_α = Z_{0,05} = 1,645$

Τώρα θα υπολογίσουμε τον έλεγχο T ιεραρχώντας τα δεδομένα από τις δύο ομάδες από τη μικρότερη στη μεγαλύτερη τιμή, όπως στον πίνακα που ακολουθεί:

α/α	Τιμή	Τάξη 1	Τάξη 2
1	1	2	
2	1		2
3	1		2
4	2	8	
5	2	8	
6	2	8	
7	2	8	
8	2		8
9	2		8
10	2		8
11	2		8
12	2		8
13	3	20,5	
14	3	20,5	
15	3	20,5	
16	3	20,5	
17	3	20,5	
18	3	20,5	
19	3	20,5	
20	3	20,5	
21	3		20,5
22	3		20,5
23	3		20,5
24	3		20,5
25	3		20,5
26	3		20,5
27	3		20,5
28	3		20,5
29	4	34,5	
30	4	34,5	
31	4	34,5	

32	4	34,5	
33	4	34,5	
34	4	34,5	
35	4	34,5	
36	4		34,5
37	4		34,5
38	4		34,5
39	4		34,5
40	4		34,5
	Σύνολο	439,5	344,5

Επειδή έχουμε πολλές ενδείξεις με την ίδια τιμή, αυτές θα βαθμολογηθούν με τον μέσο των τάξεων.
Για την τιμή 1: (1+2+3)/3=2
Για την τιμή 2: (4+5+6+7+8+9+10+11+12)/9= 8
Για την τιμή 3: (13+14+15+16+17+18+19+20+21+22+23+24+25+26+27+28)/16=20,5
Για την τιμή 4: (29+30+31+32+33+34+35+36+37+38+39+40)/12=34,5
Άρα T_1=439,5 και T2=344,5.
Εμείς θα επιλέξουμε ως έλεγχο τον T_1.
Στην συνέχεια υπολογίζουμε τον τυποποιημένο έλεγχο υπολογίζοντας πρώτα τον μέσο όρο και την τυπική απόκλιση.

$$E(T) = \frac{n_1 \cdot (n_1+n_2+1)}{2} = \frac{20 \cdot (10+10+1)}{2} = 410$$

$$\sigma_T = \sqrt{\frac{n_1 \cdot n_2 \cdot (n_1+n_2+1)}{12}} = \sqrt{\frac{20 \cdot 20 \cdot (20+20+1)}{12}} = 39{,}67$$

$$z = \frac{T-E(T)}{\sigma_T} = \frac{439{,}5-410}{39{,}67} = 0{,}7979 \approx 0{,}80$$

Ο έλεγχος $Z<Z_\alpha$ (0,80<1,645), οπότε αποδεχόμαστε την H_0. Η εμφάνιση, δηλαδή, επηρεάζει την κρίση των γυναικών.

2. Προσημασμένο τεστ Wilcoxon:

Τα χαρακτηριστικά που πρέπει να έχει ένα πρόβλημα για την διεξαγωγή του προσημασμένου τεστ Wilcoxon είναι τα εξής:
α) Στόχος είναι η σύγκριση δύο πληθυσμών.
β) Τα δεδομένα είναι είτε διατακτικά, είτε ποσοτικά χωρίς την προϋπόθεση κανονικής κατανομής.
γ) Τα δείγματα έχουν επιλεγεί κατά ζεύγη.

Ο προσημικός έλεγχος μας δίνει τη δυνατότητα, με ένα σχετικά απλό τρόπο και χωρίς ιδιαίτερες παραδοχές, να ελέγξουμε τις διαφορές των τιμών δύο κατά ζεύγη πληθυσμών. Το πρόβλημα με τον προσημικό έλεγχο είναι ότι λαμβάνει υπόψιν μόνο το πρόσημο των διαφορών και όχι το μέγεθος τους, γεγονός που μειώνει σημαντικά την ισχύ του. Το αποτέλεσμα είναι ότι ο προσημικός έλεγχος δεν χρησιμοποιείται συχνά στην πράξη, και οι έλεγχοι των διαφορών δύο συσχετιζόμενων δειγμάτων γίνονται κυρίως με τη δοκιμασία των προσημασμένων θέσεων του Wilcoxon. Ο έλεγχος του Wilcoxon, όπως και όλοι οι έλεγχοι που αναφέρονται σε ζεύγη συσχετιζόμενων τιμών, δεν εξετάζει τις τιμές των δειγμάτων χωριστά, αλλά εξετάζει τις διαφορές των συσχετιζόμενων τιμών με τρόπο ενιαίο. Δεν απαιτεί η κατανομή των διαφορών να είναι κανονική, ενώ λαμβάνει υπόψιν του εκτός του προσήμου των διαφορών και το μέγεθος τους. Η μηδενική υπόθεση του ελέγχου του Wilcoxon είναι ότι η τιμή της διαμέσου των πληθυσμιακών διαφορών είναι ίση με 0.

Για την πραγματοποίηση του ελέγχου του Wilcoxon ξεκινάμε καταρχήν με την επιλογή ενός τυχαίου δείγματος n συσχετιζόμενων τιμών. Για κάθε ζεύγος τιμών υπολογίζουμε τη διαφορά τους και στη συνέχεια, αγνοώντας τα πρόσημα των διαφορών αυτών, τις διατάσσουμε κατ' απόλυτο τιμή από τις μικρότερες ως τις μεγαλύτερες.

Διαφορές ίσες με το 0 απομακρύνονται από την ανάλυση, ενώ το μέγεθος του δείγματος ελαττώνεται κατά 1 μονάδα για κάθε ζεύγος μηδενικής διαφοράς. Διαφορές με την ίδια τιμή λαμβάνουν ως σχετική θέση (rank) κατά τη διάταξη τους τη μέση τιμή των αρχικών θέσεων τους, π.χ. αν η $4^{η}$ και η $5^{η}$ θέση κατά τη διάταξη των διαφορών έχουν την ίδια απόλυτη τιμή, η σχετική θέση των δύο αυτών διαφορών στην τελική διάταξη ορίζεται ως

(4+5)/2=4,5 ενώ η αμέσως μεγαλύτερη από αυτές παίρνει τη θέση 6.

Στις σχετικές θέσεις που προκύπτουν με τη διαδικασία αυτή τίθεται ως πρόσημο το πρόσημο των αρχικών διαφορών των τιμών. Π.χ. αν η διαφορά δύο τιμών είναι ίση με -12 και η θέση που καταλαμβάνει η συγκεκριμένη τιμή στη διάταξη των διαφορών είναι η $3^{η}$, τότε η προσημασμένη θέση της διαφοράς είναι -3.

Τελικό στάδιο κατά την εφαρμογή της διαδικασίας του Wilcoxon είναι ο υπολογισμός του αθροίσματος των θετικών θέσεων T_+ και των αρνητικών θέσεων T_- που έχουν προκύψει. Το μικρότερο κατ' απόλυτο τιμή από τα δύο αυτά αθροίσματα συμβολίζεται με T. Αποδεχόμενοι τη μηδενική υπόθεση ότι η διάμεσος των διαφορών είναι ίση με το 0, αναμένουμε ότι ο αριθμός των προσημασμένων θέσεων με θετικό και αρνητικό πρόσημο να είναι περίπου ο ίδιος και, επιπλέον, το άθροισμα των θετικών θέσεων να είναι περίπου ίσο με το άθροισμα των αρνητικών θέσεων. Η στατιστική συνάρτηση του ελέγχου είναι η ποσότητα T, η οποία εφόσον ισχύει η μηδενική υπόθεση και ο αριθμός n των συσχετιζόμενων τιμών με μη μηδενική διαφορά είναι επαρκώς μεγάλος (n/20), ακολουθεί την κανονική κατανομή με μέση τιμή

$E(T) = \frac{n(n+1)}{4}$

και διακύμανση

$\sigma_T = \sqrt{\frac{n(n+1)(2n+1)}{24}}$

Άρα η τυχαία μεταβλητή

$Z_T = \frac{T-E(T)}{\sigma_T}$

ακολουθεί την τυπική κανονική κατανομή.

Θα τα δούμε αναλυτικότερα στο παράδειγμα που ακολουθεί και με το λογισμικό Minitab.

ΠΑΡΑΔΕΙΓΜΑ

Μια φαρμακευτική εταιρία έχει αναπτύξει ένα νέο υπνωτικό χάπι που δίνεται χωρίς ιατρική συνταγή. Μια έρευνα σχετικά με την αποτελεσματικότητα του φαρμάκου πρέπει να καταγράψει πόσο χρόνο χρειάζεται κάθε μέλος του δείγματος για να κοιμηθεί αφού έχει πάρει το φάρμακο, αλλά επειδή ο χρόνος αυτός διαφέρει από άνθρωπο σε άνθρωπο, οι ερευνητές αποφασίζουν να πραγματοποιήσουν το εξής πείραμα: επιλέγουν ένα τυχαίο δείγμα 100 εθελοντών που υποφέρουν από αϋπνία και στον καθέναν δίνουν σε διαφορετικές μέρες ένα κανονικό υπνωτικό χάπι και ένα ψευδοφάρμακο (χωρίς οι ίδιοι να το γνωρίζουν με τυχαία σειρά), και καταγράφουν τον χρόνο που θα χρειαστεί ο καθένας για να αποκοιμηθεί. Μπορούμε από τα δεδομένα να συμπεράνουμε ότι το νέο υπνωτικό χάπι είναι αποτελεσματικό;
Δεδομένα:

Φάρμακο				Ψευδοφάρμακο			
23	13	16	19	30	13	14	20
19	18	12	20	24	19	16	21
8	16	16	25	6	16	20	26
26	16	18	16	27	17	15	18
34	19	29	20	34	16	33	22
23	19	21	8	27	22	21	12
16	17	10	13	14	17	10	14
10	18	19	16	8	18	22	16
25	10	21	27	26	10	30	27
17	17	19	19	18	22	22	20
23	31	13	19	23	33	15	19
23	40	26	10	23	40	26	12
18	25	24	15	18	26	25	23
18	23	20	8	22	24	19	7
16	14	11	2	13	13	11	4
14	15	24	5	14	15	32	8
20	27	24	20	21	27	22	25
24	18	16	19	25	18	18	19
18	18	14	12	19	26	14	14
12	32	17	18	15	34	18	20
22	22	21	10	23	22	22	14
11	17	12	20	12	19	13	22
13	8	28	14	18	10	28	13
36	4	38	14	36	5	49	14
19	9	25	12	26	12	25	19

ΛΥΣΗ

Η μηδενική και η εναλλακτική υπόθεση εδώ είναι:

H_0: Οι θέσεις των δύο πληθυσμών είναι ίδιες.

H_1: Οι θέσεις των δύο πληθυσμών διαφέρουν.

Για να υπολογίσουμε τον έλεγχο θα σχηματίσουμε τις διαφορές και θα κατασκευάσουμε έναν πίνακα διάταξης, ως εξής:

α/α	Χρόνος		Δ	Τάξη	
	1	2		+	-
1	14	13	1	13,5	
2	20	19	1	13,5	
3	8	7	1	13,5	
4	14	13	1	13,5	
5	26	27	-1		13,5
6	25	26	-1		13,5
7	17	18	-1		13,5
8	20	21	-1		13,5
9	24	25	-1		13,5
10	18	19	-1		13,5
11	22	23	-1		13,5
12	11	12	-1		13,5
13	18	19	-1		13,5
14	16	17	-1		13,5
15	25	26	-1		13,5
16	23	24	-1		13,5
17	4	5	-1		13,5
18	24	25	-1		13,5
19	17	18	-1		13,5
20	21	22	-1		13,5
21	12	13	-1		13,5
22	19	20	-1		13,5
23	20	21	-1		13,5
24	25	26	-1		13,5
25	13	14	-1		13,5
26	19	20	-1		13,5
27	8	6	2	35,5	
28	16	14	2	35,5	
29	10	8	2	35,5	
30	16	14	2	35,5	
31	24	22	2	35,5	

32	31	33	-2		35,5
33	32	34	-2		35,5
34	17	19	-2		35,5
35	8	10	-2		35,5
36	13	15	-2		35,5
37	16	18	-2		35,5
38	16	18	-2		35,5
39	20	22	-2		35,5
40	10	12	-2		35,5
41	2	4	-2		35,5
42	12	14	-2		35,5
43	18	20	-2		35,5
44	20	22	-2		35,5
45	16	13	3	49	
46	19	16	3	49	
47	18	15	3	49	
48	12	15	-3		49
49	19	22	-3		49
50	9	12	-3		49
51	19	22	-3		49
52	19	22	-3		49
53	5	8	-3		49
54	23	27	-4		57
55	18	22	-4		57
56	12	16	-4		57
57	16	20	-4		57
58	29	33	-4		57
59	8	12	-4		57
60	10	14	-4		57
61	19	24	-5		62,5
62	13	18	-5		62,5
63	17	22	-5		62,5
64	20	25	-5		62,5
65	23	30	-7		66
66	19	26	-7		66
67	12	19	-7		66
68	18	26	-8		69
69	24	32	-8		69
70	15	23	-8		69
71	21	30	-9		71
72	38	49	-11		72

Σύνολο	378,5	2249,5

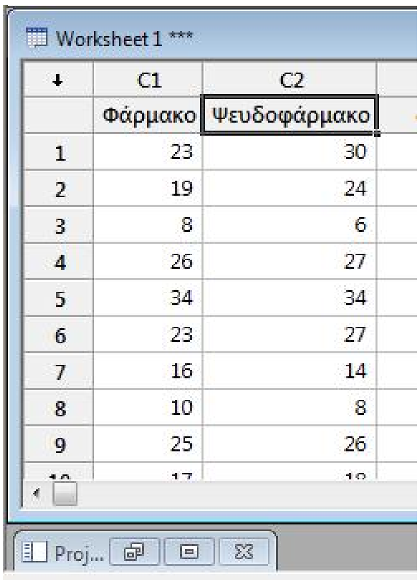
Οι διαφορές κατατάσσονται ως προς την απόλυτη τιμή της διαφοράς και στην συνέχεια αθροίζονται χωριστά οι τάξεις των θετικών και οι τάξεις των αρνητικών διαφορών. Να σημειώσουμε εδώ ότι οι μηδενικές τιμές που βρήκαμε έχουν αφαιρεθεί από τον πίνακα (γενικά τις μηδενικές τιμές τις αφαιρούμε <u>πάντα</u>).
Ο έλεγχος είναι T=378,5.
Στην συνέχεια υπολογίζουμε τον τυποποιημένο έλεγχο υπολογίζοντας πρώτα τον μέσο όρο και την τυπική απόκλιση.

$$E(T) = \frac{n(n+1)}{4} = \frac{72(72+1)}{4} = 1314$$

$$\sigma_T = \sqrt{\frac{n(n+1)(2n+1)}{24}} = \sqrt{\frac{72(72+1)(2\cdot 72+1)}{24}} = 178{,}19$$

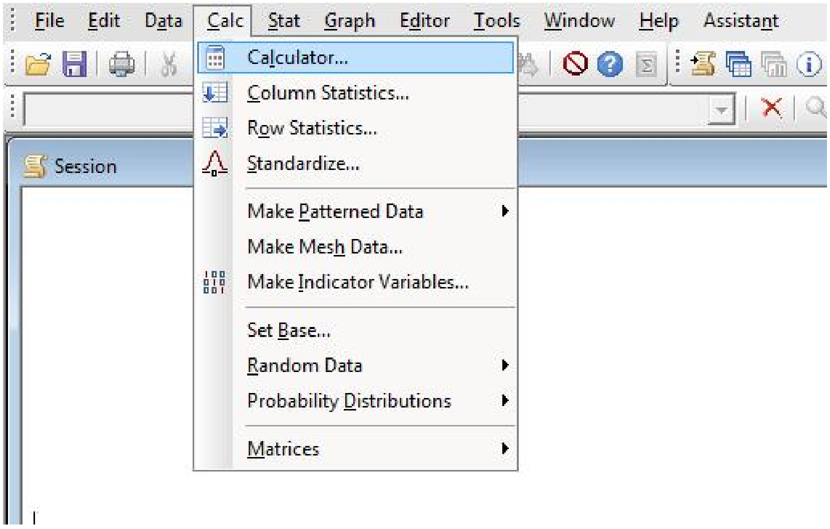
$$Z = \frac{T-E(T)}{\sigma_T} = \frac{378{,}5-1314}{178{,}19} = -5{,}25$$

Η περιοχή απόρριψης είναι:
$Z<-Z_{\alpha/2} = -Z_{0,025} = -1{,}96$
$Z> Z_{\alpha/2} = Z_{0,025} = 1{,}96$
Άρα ο έλεγχος βρίσκεται στην περιοχή απόρριψης, το οποίο μας υποδεικνύει ότι το υπνωτικό χάπι είναι πιο αποτελεσματικό.
Ας λύσουμε τώρα το πρόβλημα και στο περιβάλλον του Minitab.
Πρώτα εισάγουμε τα δεδομένα σε δύο στήλες, σχηματίζοντες τις μεταβλητές «Φάρμακο» και «Ψευδοφάρμακο».

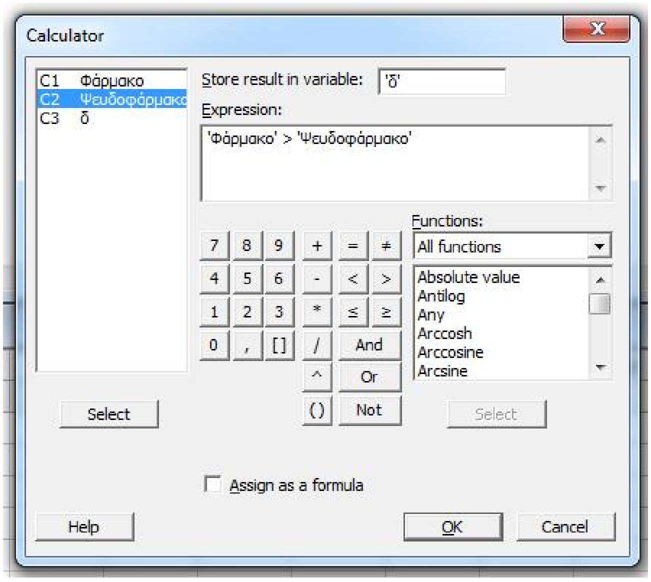

Worksheet 1 ***

↓	C1	C2
	Φάρμακο	Ψευδοφάρμακο
1	23	30
2	19	24
3	8	6
4	26	27
5	34	34
6	23	27
7	16	14
8	10	8
9	25	26

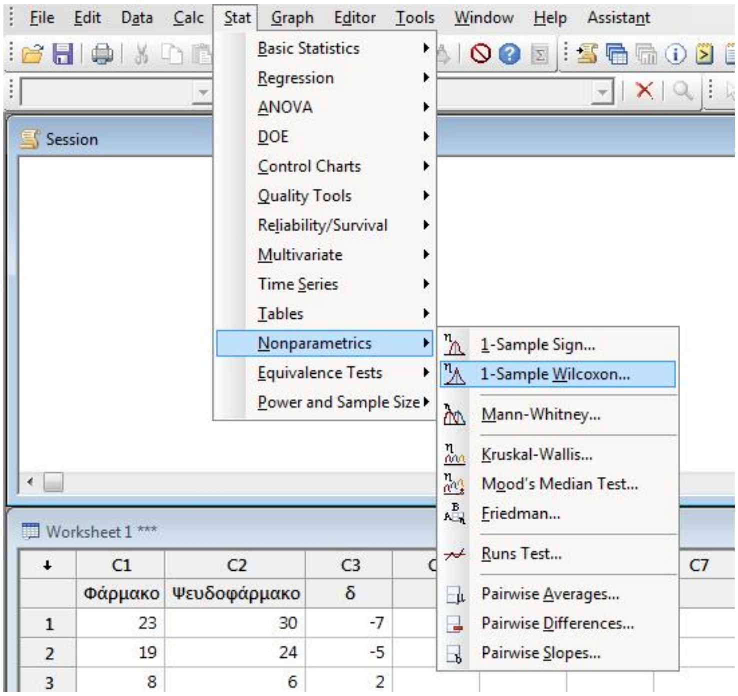
Στην συνέχεια θέλουμε να βρούμε τις διαφορές αυτών. Για την υλοποίηση αυτού, δημιουργούμε μια καινούργια μεταβλητή με όνομα «δ» και στην αριθμομηχανή του Minitab υπολογίζουμε τις διαφορές. Στην γραμμή εργαλείων πάμε Calc **à** Calculator



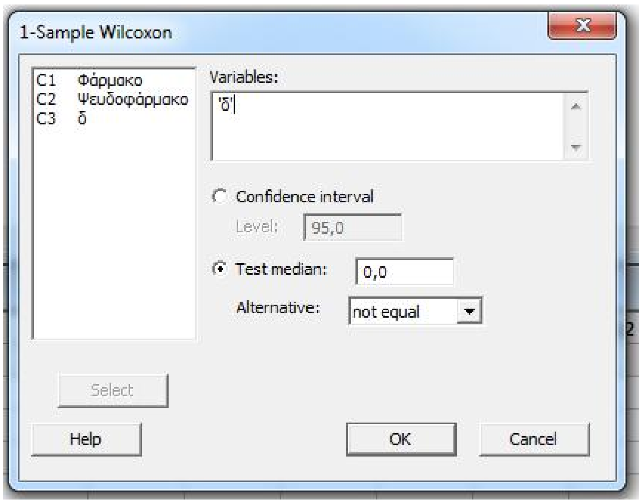
Στον πίνακα που εμφανίζει πατώντας την εντολή, βάζουμε στο Store result in variable την μεταβλητή «δ» και στο Expression περνάμε την μεταβλητή «Φάρμακο», αμέσως μετά το σύμβολο > (μεγαλύτερο) και δίπλα την μεταβλητή «Ψευδοφάρμακο» όπως φαίνεται:



Αφού πατήσουμε OK εμφανίζονται οι διαφορές στην μεταβλητή «δ». Για την υλοποίηση του ελέγχου Wilcoxon πάμε στην γραμμή εργαλείων πατάμε Stat **à** Nonparametrics **à** 1-Sample Wilcoxon:



Στον πίνακα που εμφανίζεται βάζουμε στο Variables τις διαφορές και επιλέγουμε το Test Median αφήνοντας την τιμή 0 ως έχει και στο Alternative επιλέγουμε not equal:



Ας εξηγήσουμε τα αποτελέσματα που εμφανίζονται:

Wilcoxon Signed Rank Test: δ

```
Test of median = 0,000000 versus median ≠ 0,000000

       N for   Wilcoxon          Estimated
    N   Test  Statistic      P      Median
δ  100    72      378,5  0,000      -1,000
```

Το πρώτο Ν που μας δίνεται είναι το σύνολο των δεδομένων, ενώ το N for Test είναι το σύνολο των δεδομένων που θα χρησιμοποιηθούν τελικά στο τεστ ή αλλιώς έχουν αφαιρεθεί οι τιμές με μηδενική διαφορά. Το Wilcoxon Statistic είναι ο έλεγχος Τ και το λογισμικό Minitab μας δίνει και την τιμή p η οποία ισούται με το 0.
Δυστυχώς το Minitab δεν δύναται να μας δώσει την τιμή του τυποποιημένου ελέγχου, ωστόσο έχοντας την τιμή του ελέγχου Τ και το σύνολο των δεδομένων απ' το οποίο έχουν αφαιρεθεί οι μηδενικές τιμές είναι εύκολος ο υπολογισμός του όπως είδαμε στην χειρόγραφη επίλυση του προβλήματος. (Δεν χρειάζεται να συνεχίσουμε σε επίλυση των τύπων γιατί τα δεδομένα είναι ακριβώς ίδια όπως και στην χειρόγραφη επίλυση, η λύση του προβλήματος δεν αλλάζει).

3. Τεστ Kruskal-Wallis:

Ο έλεγχος για δύο ανεξάρτητα δείγματα μπορεί να επεκταθεί για την περίπτωση προβλημάτων που αναφέρονται σε k πληθυσμούς, k>2 όπου και διατύπωσε ο Kruskal-Wallis, (1952). Η πειραματική κατάσταση αντιστοιχεί στην περίπτωση όπου k ανεξάρτητα τυχαία δείγματα είναι διαθέσιμα, ένα από κάθε ένα από k, ενδεχομένως διαφορετικούς, πληθυσμούς και θέλουμε να ελέγξουμε την μηδενική υπόθεση ότι όλοι οι πληθυσμοί είναι ισόνομοι έναντι της εναλλακτικής ότι κάποιοι από τους πληθυσμούς οδηγούν σε παρατηρούμενες τιμές που είναι μεγαλύτερες από αυτές στις οποίες οδηγούν οι άλλοι πληθυσμοί.
Οι υποθέσεις εδώ διατυπώνονται ως εξής:
H_0: Οι συναρτήσεις κατανομής των k πληθυσμών είναι ίσες
H_1: Τουλάχιστον δύο από τους k πληθυσμούς έχουν διαφορετικές μέσες τιμές.
Η μέθοδος βασίζεται στην ίδια λογική των ελέγχων εξετάζοντας αν ένα σύνολο k ανεξάρτητων ομάδων προέρχονται από τον ίδιο πληθυσμό. Σε αντίθεση με τον έλεγχο One-way ANOVA που εξετάζει την ισότητα των μέσων, εδώ χρησιμοποιείται το άθροισμα των βαθμών των θέσεων των παρατηρήσεων. Η μέθοδος είναι μια επέκταση του κριτηρίου Mann-Whitney U για k ανεξάρτητες ομάδες. Κατ' επέκταση η υπόθεση που εξετάζεται είναι αν οι ομάδες έχουν ίσο μέσο βαθμών (δηλ. τυχαία διάταξη) ή κάποια ομάδα διαφοροποιεί τη διάταξη των παρατηρήσεων.
Για την επίλυση προβλημάτων με το τεστ Kruskal-Wallis ακολουθούμε τα ίδια βήματα όπως και στο τεστ Wilcoxon με την διαφορά ότι για κάθε πληθυσμό έχουμε μία τιμή T_k και ο έλεγχος υπολογίζεται από τον τύπο:

$$H = \left[\frac{12}{n(n+1)} \cdot \sum_{j=1}^{k} \frac{T_j^2}{n_j}\right] - 3\cdot(n+1)$$

Η περιοχή απόρριψης δίνεται από τον τύπο:

$H > \chi^2_{\alpha.k-1}$

Στο ακόλουθο παράδειγμα θα αναλύσουμε περαιτέρω το συγκεκριμένο τεστ μέσω του λογισμικού Minitab λόγο του όγκου των δεδομένων είναι εύκολο να γίνουν λάθη στην χειρόγραφη επίλυση.

ΠΑΡΑΔΕΙΓΜΑ

Στη διάρκεια της τελευταίας προεκλογικής εκστρατείας μια εταιρεία δημοσκοπήσεων επέλεξε ένα τυχαίο δείγμα 30 μελών του Δημοκρατικού κόμματος τον Ιανουάριο, 30 τον Φεβρουάριο και 30 τον Μάρτιο. Τα τρία δείγματα απάντησαν στην ερώτηση: πώς κρίνετε τις πιθανότητες να κερδίσει το Δημοκρατικό κόμμα τις προεδρικές εκλογές στην πολιτεία σας: *εξαιρετικές, καλές, μέτριες, κακές*. Οι απαντήσεις κωδικοποιήθηκαν σε μια κλίμακα 4-1. Να ελέγξετε αν η πίστη των μελών του Δημοκρατικού κόμματος για τη νίκη στις εκλογές έχει αλλάξει στη διάρκεια των τριών μηνών με στάθμη σημαντικότητας 5%.

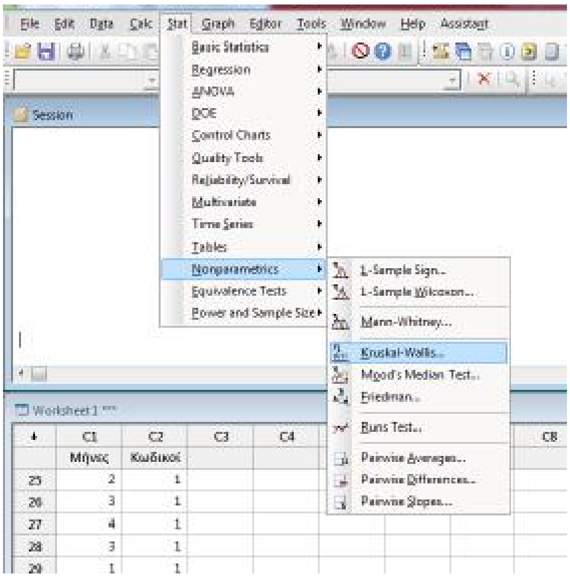
Δεδομένα:

Ιανουάριος		Φεβρουάριος		Μάρτιος	
3	4	2	1	3	1
3	4	4	2	3	1
2	3	3	4	2	1
4	3	1	4	1	4
4	4	2	4	3	2
4	4	1	3	4	4
2	2	3	4	4	2
4	3	4	2	3	4
2	4	2	4	2	1
1	2	3	4	3	4
1	3	4	1	4	1
3	4	4	2	1	1
3	3	2	3	1	2
4	1	3	2	4	3
4	4	2	3	2	3

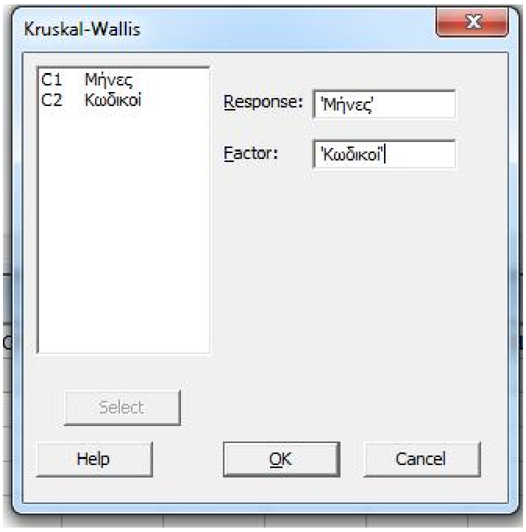
ΛΥΣΗ

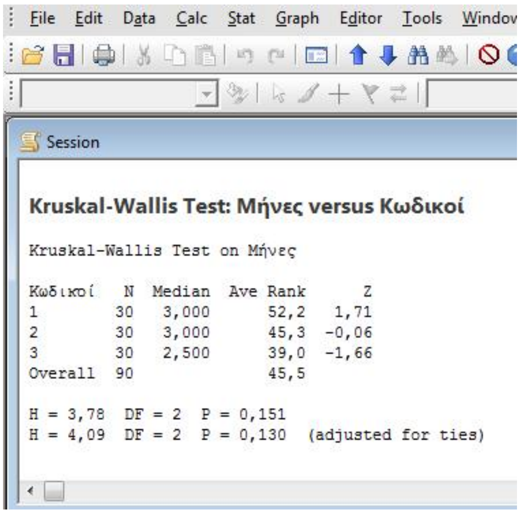
Ανοίγουμε το περιβάλλον του Minitab και εισάγουμε τα δεδομένα σε δύο στήλες. Η πρώτη αποτελείται από τα δεδομένα μας σε στοιβαγμένη μορφή, ονομάζοντας την μεταβλητή ως «Μήνες» και η δεύτερη αποτελεί τον κωδικό του κάθε πληθυσμού (1=Ιανουάριος, 2=Φεβρουάριος, 3=Μάρτιος) και θα ονομάσουμε την μεταβλητή «Κωδικοί» όπως φαίνεται παρακάτω:

Αφού εισάγουμε τα δεδομένα πάμε στην γραμμή εργαλείων και πατάμε Stat à Nonparametrics à Kruskal-Wallis



Στον πίνακα που εμφανίζεται βάζουμε στο πεδίο Response την μεταβλητή «Μήνες» και στο πεδίο Factor την μεταβλητή «Κωδικοί», πατάμε OK για να εμφανιστούν τα αποτελέσματα.





Τα δεδομένα που δίνονται είναι το σύνολο των δεδομένων κάθε μήνα (N), η διάμεσος του κάθε μήνα (Median) και ο τυποποιημένος έλεγχος κάθε μήνα (Z).

Τα αποτελέσματα που μας ενδιαφέρουν είναι ο έλεγχος H=3,78 και η τιμή p=0,151 ενώ το πρόγραμμα μας δίνει και τους βαθμούς ελευθερίας (DF=2).

Η περιοχή απόρριψης είναι:

H>$\chi^2_{\alpha.k-1}$=$\chi^2_{0,05.2}$=5,99 ,όπου α η στάθμη σημαντικότας και k το σύνολο των πληθυσμών.

Βλέπουμε ότι η παραπάνω συνθήκη δεν ισχύει αφού ο έλεγχος είναι μικρότερος, άρα δεν βρίσκεται στην περιοχή απόρριψης.

Καταλήγουμε ότι η πίστη των μελών του Δημοκρατικού κόμματος για τη νίκη στις εκλογές δεν έχει αλλάξει στη διάρκεια των τριών μηνών.

4. Τεστ Friedman:

Είναι μια επέκταση της περίπτωσης που έχουμε δεδομένα ανεξάρτητα σε ζεύγη και εφαρμόζεται όταν ένα άτομο ή περιοχή ελέγχεται τρεις ή περισσότερες φορές.
Αυτός ο έλεγχος είναι ο μη-παραμετρικός ανάλογος της two-way ANOVA. Δεν κάνει παραδοχές για τα δεδομένα (μόνο ότι είναι διακριτά). Είναι κατάλληλος μόνο όταν υπάρχει μια μοναδική παρατήρηση για κάθε συνδυασμό των επιπέδων ενός παράγοντα. Για τις επαναλαμβανόμενες μετρήσεις ένας από τους παράγοντες πρέπει να αντιπροσωπεύει το επίπεδο επανάληψης (π.χ. λεπτά, ημέρες ή μια μέτρηση πριν, κατά και μετά στο πλαίσιο μιας διεργασίας που μελετάμε). Κατόπιν, ο δεύτερος παράγοντας είναι ένας τυπικός παράγοντας όπως περιοχή, είδος, ή τύπος επίδρασης. Η μηδενική υπόθεση είναι ότι οι παρατηρήσεις μέσα στην ίδια ομάδα (επίπεδο του παράγοντα) έχουν την ίδια τιμή διαμέσου (median). Εάν απορριφθεί η μηδενική υπόθεση αυτό σημαίνει ότι τουλάχιστον δύο ομάδες έχουν διαφορετικούς διάμεσους (παρόλο που δεν δείχνει ποιες). Ο έλεγχος Friedman είναι λιγότερο ισχυρός από την ANOVA όταν τα δεδομένα έχουν κανονική κατανομή, αλλά κάνει λιγότερες παραδοχές για τα δεδομένα, και επομένως είναι «ασφαλέστερος».
Ο έλεγχος υπολογίζεται όπως και στα προηγούμενα τεστ (Wilcoxon και Kruskal-Wallis), εδώ όμως δημιουργούμε έναν πίνακα κατάταξης για κάθε ομάδα των πληθυσμών. Στον κάθε πίνακα οι παρατηρήσεις βαθμολογούνται με τον αριθμό της τάξης που καταλαμβάνουν (παρατηρήσεις με την ίδια τιμή βαθμολογούνται με τον αριθμητικό μέσο των τάξεων που καταλαμβάνουν). Έπειτα, υπολογίζουμε για κάθε πληθυσμό το άθροισμα των τάξεων T_1, T_2 . . ., T_k ,όπου k ο αριθμός των πληθυσμών. Αν ο αριθμός των ομάδων είναι b τότε ο έλεγχος υπολογίζεται από τον τύπο:

$$F_r = \left[\frac{12}{b\cdot k\cdot(k+1)}\cdot\sum_{j=1}^{k} T_j^2\right] - 3b(k+1)$$

Αποδεικνύεται μαθηματικά ότι αν k≥5 ή b≥5 τότε ο έλεγχος ακολουθεί την κατανομή χ^2 με k-1 βαθμούς ελευθερίας.
Η περιοχή απόρριψης δίνεται από τον τύπο:
$F_r > \chi^2_{\alpha.k-1}$
Και η τιμή p απ' την πιθανότητα:
$P(\chi^2 > F_r)$
Θα δούμε το τεστ Friedman αναλυτικότερα στο παράδειγμα που ακολουθεί με την βοήθεια του λογισμικού Minitab, λόγο του όγκου των δεδομένων τέτοιου είδους ασκήσεων είναι καλύτερο να λύνονται στον υπολογιστή και όχι χειρόγραφα.

ΠΑΡΑΔΕΙΓΜΑ

Ένας κατασκευαστής αναψυκτικών χρησιμοποιεί την ίδια μυστική συνταγή για περισσότερα από 100 χρόνια, αλλά πρόσφατα λόγω της μείωσης του μεριδίου αγοράς άρχισε να εξετάζει δυο νέες εναλλακτικές συνταγές. Ένα δείγμα 20 καταναλωτών δοκίμασαν την παραδοσιακή συνταγή και τις δύο νέες συνταγές και τις βαθμολόγησαν ως εξής:

Κλίμακα	Ποιότητα αναψυκτικού
1	Απαράδεκτη
2	Κακή
3	Μέτρια
4	Καλή
5	Εξαιρετική

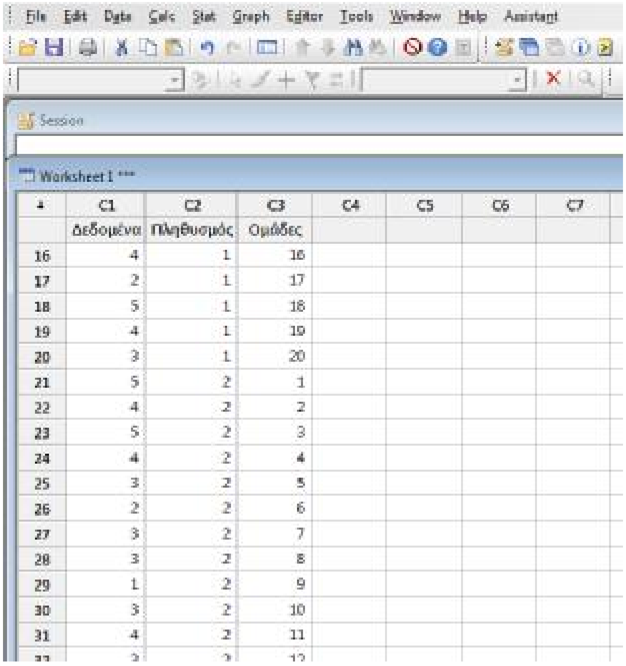
Η εταιρία έχει αποφασίσει ότι δεν πρόκειται να αλλάξει την συνταγή, εκτός αν κάποια από τις νέες συνταγές είναι σημαντικά καλύτερη από την παραδοσιακή. Να ελέγξετε αν υπάρχουν διαφορές ανάμεσα στις τρεις συνταγές.
Τα δεδομένα είναι:

Μυστική Συνταγή	Συνταγή 1	Συνταγή 2
5	5	5
3	4	5
4	5	5
2	4	4
3	3	5
2	2	3
3	3	2
4	3	5
1	1	1
1	3	2
2	4	3
3	3	4
5	3	4
3	2	3
4	3	5
4	4	4
2	4	3
5	3	5
4	5	5
3	5	4

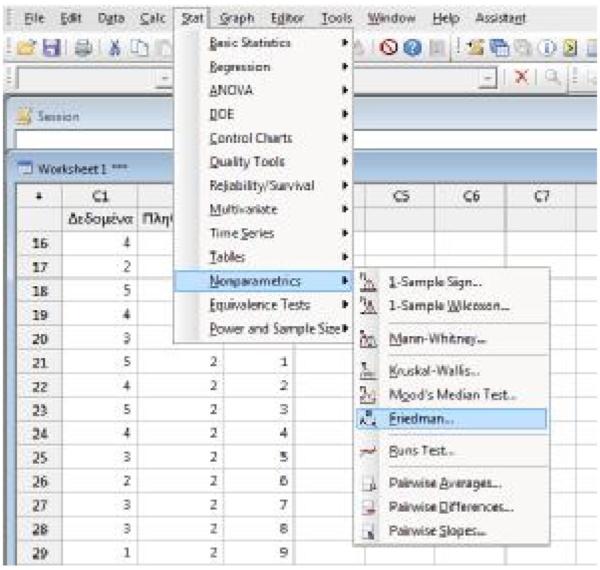
ΛΥΣΗ

Η υπόθεση που θέλουμε να ελεγχθεί είναι:
H_0: Οι θέσεις των τριών πληθυσμών είναι ίδιες.
H_1: Οι θέσεις δύο τουλάχιστον πληθυσμών διαφέρουν.

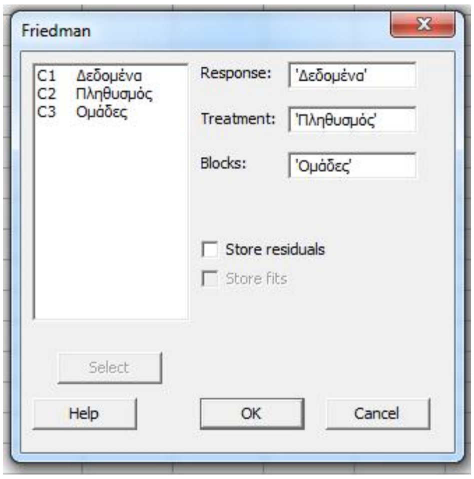
Ανοίγουμε το λογισμικό Minitab και θα δημιουργήσουμε τρεις μεταβλητές. Η πρώτη αποτελείται από όλα τα δεδομένα σε στοιβαγμένη μορφή με όνομα «Δεδομένα», η δεύτερη αποτελείται απ' τους κωδικούς των πληθυσμών με όνομα «Πληθυσμός» και η τρίτη αποτελείται απ' τους κωδικούς των ομάδων με όνομα «Ομάδες» όπως φαίνεται παρακάτω:



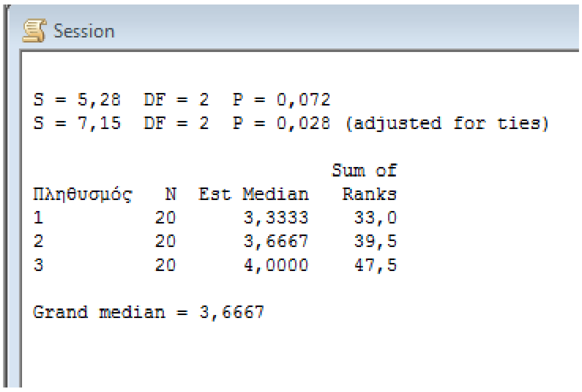
Αφού εισάγουμε τα δεδομένα πάμε στην γραμμή εργαλείων και πατάμε Stat **à** Nonparametrics **à** Friedman...



Στον πίνακα που ανοίγει βάζουμε στο πεδίο Response την μεταβλητή «Δεδομένα», στο πεδίο Treatment εισάγουμε την μεταβλητή «Πληθυσμός» και στο πεδίο Blocks βάζουμε την μεταβλητή «Ομάδες» **à** OK.



Τα αποτελέσματα που εμφανίζονται είναι τα εξής:



Βλέπουμε ότι ο έλεγχος που μας δίνεται ως S στο Minitab είναι F_r=5,28 και η τιμή p=0,072 με 2 βαθμούς ελευθερίας.
Η περιοχή απόρριψης είναι:
$F_r > \chi^2_{\alpha.k-1} = \chi^2_{0{,}05.3-1} = \chi^2_{0{,}05.2}$ = 5,99 (με την βοήθεια του παραρτήματος Β)
Άρα ο έλεγχος F_r δεν βρίσκεται στην περιοχή απόρριψης. Καταλήγοντας στο ότι δεν υπάρχουν διαφορές ανάμεσα στις τρεις συνταγές αναψυκτικού.

5. <u>Συντελεστής συσχέτισης Spearman</u>:

Ο μη-παραμετρικός έλεγχος Spearman είναι κατάλληλος, εφόσον υπάρχουν δύο παρατηρήσεις για κάθε άτομο και οι παρατηρήσεις έχουν μετρηθεί σε μια κλίμακα που μπορεί να μετατραπεί σε μια λογική σειρά κατάταξης. Ο συντελεστής r_s που εκτιμά το βαθμό συσχέτισης κυμαίνεται από -1 (τέλεια αρνητική συσχέτιση) μέχρι το 1 (τέλεια θετική συσχέτιση), και περιγράφει το βαθμό της συσχέτισης ανάμεσα σε δύο μεταβλητές. Η μηδενική και η εναλλακτική υπόθεση εδώ είναι η εξής:
H_0: Δεν υπάρχει συσχέτιση (ο συντελεστής συσχέτισης είναι μηδενικός).

H_1: Υπάρχει συσχέτιση (ο συντελεστής συσχέτισης δεν είναι μηδενικός πληθυσμό) Εδώ ο συντελεστής του Spearman υπολογίζεται με βάση τις τάξεις μεγέθους των δεδομένων. Αυτό είναι και που επιτρέπει την ελευθερία ως προς τη μη ικανοποίηση της κανονικότητας των μεταβλητών. Θα το δούμε αναλυτικότερα μέσα από ένα παράδειγμα με την βοήθεια του λογισμικού SPSS.

ΠΑΡΑΔΕΙΓΜΑ

Πολλοί άνθρωποι υποφέρουν από καούρες στο στομάχι, ένα πρόβλημα που δείχνει να γίνεται εντονότερο με την ηλικία. Μια έρευνα για λογαριασμό μιας φαρμακευτικής εταιρίας επέλεξε ένα τυχαίο δείγμα 325 ενηλίκων και κατέγραψε την ηλικία καθενός, καθώς και τη σοβαρότητα των συμπτωμάτων καούρας που αντιμετωπίζει: ελαφρά, μέτρια, σοβαρά, πολύ σοβαρά (σε κλίμακα 1-4). Μπορούμε από τα δεδομένα να συμπεράνουμε με στάθμη σημαντικότητας 5% ότι τα συμπτώματα γίνονται εντονότερα με την ηλικία;
Τα δεδομένα:

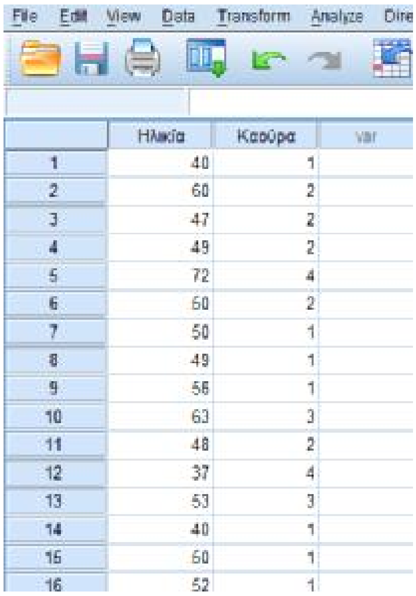
Ηλικία				Καούρα			
40	27	47	59	1	1	4	3
60	37	40	60	2	1	3	2
47	65	71	60	2	1	2	1
49	52	71	74	2	4	1	1
72	56	68	37	4	1	4	1
50	62	50	58	2	2	3	2
50	50	65	62	1	3	3	1
49	65	59	47	1	1	1	1
56	38	57	49	1	3	3	2
63	33	57	50	3	1	1	3
48	47	64	48	2	1	4	1
37	47	65	39	4	1	3	1
53	56	59	64	3	1	1	2
40	49	54	56	1	3	2	1
50	30	70	57	1	1	2	1
52	39	53	47	1	2	2	4
52	65	42	48	4	3	2	2
61	36	54	48	3	2	1	1
51	31	57	66	1	1	4	2
27	50	66	34	3	4	2	1
53	48	51	70	1	3	1	4
60	51	37	50	1	1	4	4
55	61	60	61	2	2	1	1
35	69	68	57	1	2	3	1
39	62	47	48	1	1	1	1
48	70	68		2	1	1	
62	41	52		4	3	1	
32	59	28		1	2	3	
55	75	45		1	2	3	
57	42	51		1	1	3	
55	49	70		2	1	1	

62	53	42		1	1	1	
50	56	57		3	1	3	
57	25	54		3	2	3	
77	28	46		4	2	2	
60	46	32		1	4	1	
57	61	62		2	3	1	
53	39	57		4	2	3	
57	45	44		1	4	3	
80	52	52		4	1	1	
68	72	52		3	1	1	
53	62	43		2	2	2	
66	51	77		3	1	2	
57	51	70		1	1	1	
57	54	57		1	2	1	
60	25	53		2	4	3	
41	69	71		2	3	3	
50	54	41		1	3	2	
47	25	52		1	2	1	
63	53	77		1	1	2	
62	55	48		1	3	2	
67	62	61		1	1	1	
52	49	63		3	1	1	
61	53	56		1	1	3	
51	62	39		3	2	1	
45	58	42		1	1	2	
44	47	74		3	2	1	
39	61	74		1	1	1	
69	38	39		2	2	3	
46	46	32		1	3	4	
48	40	71		2	4	1	
52	71	40		2	1	2	
50	65	68		2	1	1	
60	46	60		3	3	4	
59	64	47		4	1	3	
74	40	54		2	3	1	
41	60	65		1	3	3	
50	47	33		1	2	3	
58	54	58		3	3	3	
53	31	58		1	4	1	
78	71	55		4	1	2	
57	52	41		3	1	2	
45	53	51		1	1	1	
48	66	45		1	1	2	
58	64	44		1	1	1	
59	71	49		1	2	1	
61	46	50		3	4	1	
59	57	43		1	4	2	
48	45	47		1	3	2	
46	70	57		3	1	3	

39	78	48		1	3	1	
67	56	56		2	2	1	
77	31	49		2	1	1	
57	62	70		1	1	1	
79	72	65		2	3	1	
38	43	56		1	1	2	
56	63	54		2	3	1	
58	53	62		2	4	3	
67	64	66		1	1	3	
61	41	61		2	1	3	
54	21	73		1	1	1	
48	65	48		2	1	4	
42	42	37		1	3	1	
36	44	44		1	1	1	
56	49	50		3	1	1	
53	54	63		2	4	1	
47	45	54		2	3	3	
66	48	37		2	1	1	
51	53	77		3	1	3	
63	32	34		3	1	1	

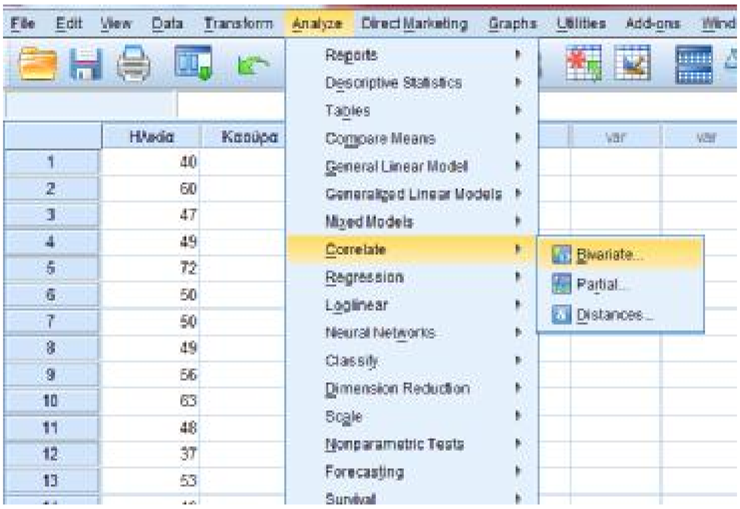
ΛΥΣΗ

Οι έλεγχοι υποθέσεων είναι:
H_0: Δεν υπάρχει συσχέτιση.
H_1: Υπάρχει συσχέτιση.
Εισάγουμε τα δεδομένα μας στο SPSS σε δύο στήλες, η μία απαρτίζεται από τα δεδομένα της ηλικίας και ονομάζουμε την μεταβλητή «Ηλικία» και η άλλη από τα δεδομένα των συμπτωμάτων καούρας και την ονομάζουμε «Καούρα».
Στο φύλλο Variable View θέτουμε και τις δύο μεταβλητές ως Ordinal.

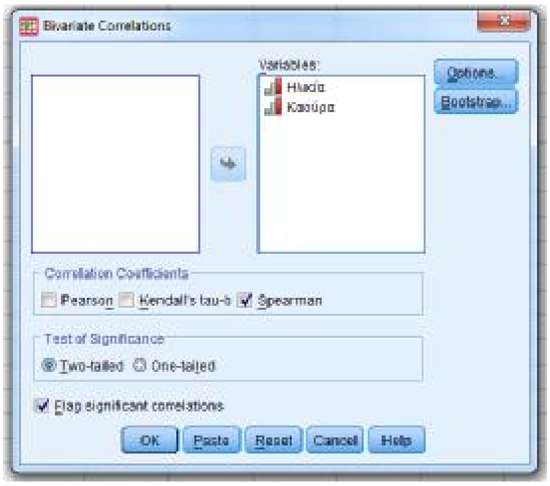


	Ηλικία	Κασύρα
1	40	1
2	60	2
3	47	2
4	49	2
5	72	4
6	50	2
7	50	1
8	49	1
9	56	1
10	63	3
11	48	2
12	37	4
13	53	3
14	40	1
15	60	1
16	52	1

Στην συνέχεια πηγαίνουμε στην γραμμή εργαλείων και πατάμε Analyze **à** Correlate **à** Bivariate...



Στον πίνακα που εμφανίζεται στο Variables περνάμε τις δύο μεταβλητές μας, στο Correlation Coefficients επιλέγουμε τον έλεγχο Spearman, στο Test of Significance επιλέγουμε Two_tailed και OK.



Στο Output εμφανίζεται ο εξής πίνακας:

Correlations

			Ηλικία	Καούρα
Spearman's rho	Ηλικία	Correlation Coefficient	1,000	,030
		Sig. (2-tailed)	.	,587
		N	325	325
	Καούρα	Correlation Coefficient	,030	1,000
		Sig. (2-tailed)	,587	.
		N	325	325

Εδώ μας ενδιαφέρει το Correlation Coefficient και δηλώνεται ως P_s, P_s=0,030 και με βάση το Significance ελέγχουμε αν επαληθεύεται ή όχι η H_0.
Αν P_s > 0,05 δεν μπορούμε να απορρίψουμε την H_0.
Αν P_s < 0,05 απορρίπτουμε την H_0 και δεχόμαστε την H_1.
Στο παράδειγμα ο έλεγχος P_s είναι μικρότερος του 0,05 άρα αποδεχόμαστε την H_1 και μπορούμε να συμπεράνουμε ότι τα συμπτώματα γίνονται εντονότερα με την ηλικία.

Βιβλιογραφία

Ελληνικές βιβλιογραφικές αναφορές

Δημητριάδης Ευ. (2016). «Στατιστική Επιχειρήσεων με εφαρμογές σε SPSS και LISREL», Αθήνα.

Κονόμος Γ. (2009). «Πιθανότητες και Στατιστική», Αθήνα.

Ξεκαλάκη Ευδ. (2001). «Μη-Παραμερική Στατιστική», Αθήνα.

Συμεωνάκη Μαρία (2015). «Στατιστική για όλους με το SPSS», Αθήνα.

Ξένες βιβλιογραφικές αναφορές

Gerald Keller, Joseph L. Rotman (2009). «Στατιστική για οικονομικά & διοίκηση επιχειρήσεων». Εκδόσεις Επίκεντρο, Αθήνα.

Andy Field (2013). «Discovering Statistics Using IBM SPSS Statistics»

Paul Newbold, William Carlson, Betty Thorne (2012). «Statistics for Business and Economics».

Jessica M. Utts (2014). «Seeing Through Statistics».

Ηλεκτρονικές πηγές

http://users.uoi.gr/abatsidis/NonParametricClassNotes2014.pdf

http://opencourses.uom.gr/assets/site/content/courses/72/Notes-SPSS.pdf

https://eclass.upatras.gr/modules/document/file.php/BIO236/GIOKAS/PMS_Data%20Analysis1.pdf

http://nefeli.lib.teicrete.gr/browse/stef/iat/2004/PapachatzakisIoannis/attached-document/2004Papachatzakis.pdf

http://users.uoa.gr/~roussosp/stats/Chapter17.pdf

http://eclass.uth.gr/eclass/modules/document/file.php/MHXC165/9%20Στατιστική%20ανάλυση/TS_SPSS.pdf

https://repository.kallipos.gr/bitstream/11419/5081/1/07_chapter6.pdf

http://users.auth.gr/agpapana/SPSS

http://www.actuar.aegean.gr/notes/Embalotis%20et%20al_%20Stat_Notes.pdf

http://users.auth.gr/agpapana/StatLogistics

https://eclass.uoa.gr/modules/document/file.php/PPP125/%CE%A7%CF%81%CE%B7%CF%83%CF%84%CE%B9%CE%BA%CE%BF%CE%AF%20%CE%BF%CE%B4%CE%B7%CE%B3%CE%BF%CE%AF/%CE%A3%CF%84%CE%B1%CF%84%CE%B9%CF%83%CF%84%CE%B9%CE%BA%CE%AE%20%CE%BC%CE%B5%20%CF%84%CE%BF%20SPSS%3A%20%CE%95%CE%B9%CF%83%CE%B1%CE%B3%CF%89%CE%B3%CE%AE.pdf http://www.math.upatras.gr/~adk/lectures/ida/lab3/slides3.pdf

file:///C:/Users/natalia/Downloads/2_%CE%88%CE%BB%CE%B5%CE%B3%CF%87%CE%BF%CF%82_%CE%A72.pdf

http://www.toposbooks.gr/behavioralstats/samplechapter.pdf

file:///C:/Users/natalia/Downloads/Statistiki%20II%20-%20lab%20(2).pdf

http://opencourses.uom.gr/assets/site/content/courses/72/Notes-SPSS.pdf

http://users.uoi.gr/ktsilidi/Statistical%20tests_KKT.pdf

http://www.statisticshowto.com/probability-and-statistics/chi-square/

http://www2.stat-athens.aueb.gr/~exek/NPar-Statistics/chapter4.pdf

http://apps.fischlerschool.nova.edu/toolbox/StatisticsSPSS.html

http://www.actuar.aegean.gr/notes/Embalotis%20et%20al_%20Stat_Notes.pdf