ΕΓΧΕΙΡΙΔΙΟ ΣΤΑΤΙΣΤΙΚΗΣ

ΕΛΕΓΧΟΣ ΠΡΟΫΠΟΘΕΣΕΩΝ
ΜΕΘΟΔΩΝ ΣΤΑΤΙΣΤΙΚΗΣ ΑΝΑΛΥΣΗΣ
με χρήση SPSS v.16

ΕΙΡΗΝΗ ΜΕΤΑΞΑ
ΠΑΝΑΓΙΩΤΗΣ ΤΖΙΑΧΡΗΣ

ΕΛΛΗΝΙΚΟΣ ΓΕΩΡΓΙΚΟΣ ΟΡΓΑΝΙΣΜΟΣ – ΔΗΜΗΤΡΑ
ΓΕΝΙΚΗ ΔΙΕΥΘΥΝΣΗ ΑΓΡΟΤΙΚΗΣ ΕΡΕΥΝΑΣ

ΙΝΣΤΙΤΟΥΤΟ ΕΔΑΦΟΫΔΑΤΙΚΩΝ ΠΟΡΩΝ (ΙΕΥΠ)
Απρίλιος 2020

Πρόλογος

Το παρόν εγχειρίδιο στατιστικής συγγράφηκε για να δώσει **έμφαση στους ελέγχους**, που πρέπει να υλοποιούνται κατά τη διάρκεια της στατιστικής επεξεργασίας των δεδομένων. Σκοπός του εγχειριδίου είναι αφού κωδικοποιηθούν οι προϋποθέσεις, που πρέπει να ισχύουν για κάθε εργασία, να διευκολυνθούν οι χρήστες στο έργο τους και το αποτέλεσμα της στατιστικής ανάλυσης να είναι όσο το δυνατόν πιο αξιόπιστο. Για διευκόλυνση του χρήστη, η βασική ορολογία της στατιστικής δίνεται και στην αγγλική γλώσσα.
Το στατιστικό πακέτο, που χρησιμοποιείται στο εγχειρίδιο είναι το SPSS v.16.
Τέλος, οι οδηγίες του παρόντος εγχειριδίου δεν έχουν σκοπό να υποκαταστήσουν τη σημασία και την αναγκαιότητα βασικών γνώσεων στατιστικής, αλλά να αποτελέσουν ένα **προβληματισμό** σε όσους, μη ειδικούς επιστήμονες, χρησιμοποιούν πακέτα στατιστικής επεξεργασίας.

Επιστημονική επιμέλεια
Μεταξά Ειρήνη, Γεωλόγος M.Sc., Ειδικός Τεχνικός Επιστήμονας
Τζιαχρής Παναγιώτης, Αγρ. Τοπογράφος M.Sc., PhD, Ερευνητής

Ινστιτούτο Εδαφοϋδατικών Πόρων
57001 Θέρμη Θεσσαλονίκη
τηλ. 2310 472160, 2310 473429
email: i.metaxa@swri.gr
email: p.tziachris@swri.gr

Ευχαριστίες
Ευχαριστούμε τον **κ. Μενεξέ Γεώργιο**, Αναπληρωτή Καθηγητή Βιομετρίας και Γεωργικού Πειραματισμού της Γεωπονικής Σχολής του ΑΠΘ, για τις παρατηρήσεις του κατά την ανάγνωση του παρόντος εγχειριδίου.

Εισαγωγή

Στο πρώτο κεφάλαιο του εγχειριδίου, παρουσιάζεται η σύγκριση μέσων δύο ή περισσότερων πληθυσμών (compare two or several means). Για το σκοπό αυτό γίνεται ανάλυση της παραλλακτικότητας (ANOVA) των δεδομένων με την οποία εξετάζεται η επίδραση, που έχει κάποιος πειραματισμός (factor) στον πληθυσμό (experimental research). Ο πειραματισμός μπορεί να είναι ένας (One-Way ANOVA), δύο (Two-Way ANOVA) ή περισσότεροι.

Στο δεύτερο κεφάλαιο παρουσιάζεται η συσχέτιση μεταξύ μεταβλητών (correlation research) και η εύρεση βέλτιστης γραμμής (regression line) πρόβλεψης της εξαρτημένης μεταβλητής (y) από μία ή περισσότερες (x1, x2, x3,...) ανεξάρτητες μεταβλητές (univariate ANOVA). Στην περίπτωση, που το μοντέλο βασίζεται σε μία ανεξάρτητη μεταβλητή έχουμε την ανάλυση απλής παλινδρόμησης (simple regression analysis), διαφορετικά μιλάμε για ανάλυση πολλαπλής παλινδρόμησης (multiple regression analysis). Αν εξετάζονται περισσότερες από μία εξαρτημένες μεταβλητές, η ανάλυση ονομάζεται MANOVA (multivariate analysis of variance).

Σκοπός του εγχειριδίου είναι να δοθεί έμφαση στις προϋποθέσεις, που πρέπει να ικανοποιούνται για κάθε εργασία στατιστικής ανάλυσης. Η έλλειψη του σχετικού ελέγχου ακόμα και σε επίπεδο διδακτορικών διατριβών ή και επιστημονικών εργασιών δεν είναι σπάνιο φαινόμενο καθώς αυτός αποτελεί μία επιπλέον και επίπονη, μερικές φορές, εργασία. Δυστυχώς, χωρίς τον έλεγχο αυτό, το αποτέλεσμα της στατιστικής επεξεργασίας δεν μπορεί να θεωρηθεί ως αξιόπιστο.

Για τη επίτευξη του σκοπού αυτού, το παρόν εγχειρίδιο χρησιμοποιεί τις απλούστερες περιπτώσεις στατιστικής επεξεργασίας, δηλαδή την τεχνική One-Way ANOVA (μία εξαρτημένη και μία ανεξάρτητη μεταβλητή) και την τεχνική της γραμμικής παλινδρόμησης.

Περιεχόμενα

1. ΣΤΑΤΙΣΤΙΚΟΣ ΕΛΕΓΧΟΣ ΥΠΟΘΕΣΕΩΝ ΓΙΑ ΔΙΑΦΟΡΑ ΜΕΣΩΝ ΤΙΜΩΝ ΔΥΟ Η ΠΕΡΙΣΣΟΤΕΡΩΝ ΠΛΗΘΥΣΜΩΝ (COMPARE TWO OR SEVERAL MEANS) 1
1.1. Θεωρία 1
1.2. Εφαρμογή 2
1.2.1. Παραμετρικά τεστ 2
1.2.1.1 Προϋποθέσεις παραμετρικών τεστ 2
1.2.1.2 Παραμετρικό τεστ για σύγκριση δύο μέσων (t-test) 8
1.2.1.3 Παραμετρικό/α τεστ για σύγκριση περισσότερων (των 2) μέσων 10
1.2.2. Μετασχηματισμός δεδομένων (transformation) 14
1.2.3. Μη-παραμετρικά τεστ 15
1.2.3.1 Προϋποθέσεις μη-παραμετρικών τεστ 15
1.2.3.2 Μη-παραμετρικό/α τεστ για σύγκριση δύο μέσων 15
1.2.3.3 Μη-παραμετρικό τεστ για σύγκριση περισσότερων (των 2) μέσων 16
2. ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ (LINEAR REGRESSION) 17
2.1. Θεωρία 17
2.2. Εφαρμογή 18
2.2.1. Απλή γραμμική παλινδρόμηση (Simple Linear Regression) 18
2.2.2. Απλή μη γραμμική παλινδρόμηση 21
2.2.3. Πολλαπλή γραμμική παλινδρόμηση (Multiple linear regression) 22
ΣΥΜΠΕΡΑΣΜΑΤΑ 25
ΒΙΒΛΙΟΓΡΑΦΙΑ 25

1. ΣΤΑΤΙΣΤΙΚΟΣ ΕΛΕΓΧΟΣ ΥΠΟΘΕΣΕΩΝ ΓΙΑ ΔΙΑΦΟΡΑ ΜΕΣΩΝ ΤΙΜΩΝ ΔΥΟ Η ΠΕΡΙΣΣΟΤΕΡΩΝ ΠΛΗΘΥΣΜΩΝ (COMPARE TWO OR SEVERAL MEANS)

1.1. Θεωρία

Ο **στατιστικός έλεγχος υποθέσεων** για τη σύγκριση των μέσων τιμών δύο ή περισσότερων πληθυσμών εστιάζεται στο εάν η επίδραση συγκεκριμένης παραμέτρου (factor/independent variable) σε δείγματα πληθυσμών είναι/δεν είναι τυχαία και ιδιαίτερα στην διαφορά μεταξύ των δειγμάτων/πληθυσμών (εάν επιζητούμε την απόδειξη της ομοιότητας μεταξύ πληθυσμών τότε κάνουμε equivalence test).

Υπάρχουν **δύο τρόποι συλλογής δεδομένων** για την διερεύνηση της επίδρασης ή μη κάποιας παραμέτρου μεταξύ των δειγμάτων/πληθυσμών. Ο πρώτος τρόπος είναι να συλλεχθούν **ανεξάρτητα δείγματα** (independent samples) δηλαδή δύο ανεξάρτητες μεταξύ τους ομάδες δειγμάτων και σε κάθε ομάδα να εφαρμοστεί διαφορετικός πειραματισμός. Ο δεύτερος τρόπος είναι να συλλεχθούν **εξαρτημένα δείγματα** (dependent/paired/matched samples) και ο διαφορετικός πειραματισμός να εφαρμοστεί στην ίδια ομάδα δείγματος, είτε ταυτόχρονα είτε με επαναληπτικές μετρήσεις σε μεταγενέστερο χρόνο.

Η βασική διαφορά μεταξύ των δύο τρόπων συλλογής είναι η εξής: όταν χρησιμοποιούνται ανεξάρτητα δείγματα στη σύγκριση των μέσων τιμών, συμμετέχουν η **συστηματική παραλλακτικότητα**, που οφείλεται στον πειραματισμό και η **μη συστηματική παραλλακτικότητα**, που οφείλεται σε τυχαία αίτια προερχόμενα κυρίως από την ανεξαρτησία των δειγμάτων. Αντίθετα, όταν στη σύγκριση των μέσων τιμών χρησιμοποιούνται τα ίδια δείγματα, η μη συστηματική παραλλακτικότητα περιορίζεται και γίνεται πιο έκδηλη η επίδραση του πειραματισμού (συστηματική παραλλακτικότητα).

Στον στατιστικό έλεγχο υποθέσεων πάντα υπάρχει μια μηδενική (null hypothesis-H_0) και μια εναλλακτική υπόθεση (alternative hypothesis-H_1).

Ο ερευνητής ορίζει μια μηδενική υπόθεση (συνήθως αυτό που ήδη ισχύει ή είναι γενικά αποδεκτό) και μια εναλλακτική υπόθεση (π.χ. εάν αλλάξει κάτι, αυτό θα έχει επίδραση στον πληθυσμό) και στην συνέχεια υλοποιεί αυτή την επίδραση/αλλαγή μέσω πειράματος. Με την χρήση του στατιστικού ελέγχου υποθέσεων σε δείγματα, αξιολογεί εκ των υστέρων εάν οι διαφορές, που προκύπτουν πριν και μετά το πείραμα είναι τυχαίες ή οφείλονται στην επίδραση του πειράματος.

Το αποτέλεσμα ενός πειράματος πάντα πρέπει να αναφέρεται στην μηδενική υπόθεση και αυτή είναι, που προσπαθεί ένας ερευνητής να απορρίψει (ή αποτυγχάνει να

απορρίψει) και όχι να αποδείξει την εναλλακτική υπόθεση (αποφυγή confirmation bias).

Στόχος στατιστικού έλεγχου υποθέσεων:

απόρριψη

ή

αποτυχία απόρριψης

της Μηδενικής Υπόθεσης (H_0)

1.2. Εφαρμογή

Υπάρχουν δύο είδη δοκιμών (test), που μπορούν να χρησιμοποιηθούν για τη σύγκριση μέσων δύο ή περισσότερων δειγμάτων. Τα παραμετρικά τεστ (t-test, ANOVA), που συγκρίνουν μέσες τιμές και τα μη-παραμετρικά τεστ (Mann-Whitney, Wilcoxon test), που συνήθως χρησιμοποιούν διαμέσους για τη σύγκριση των πληθυσμών.

1.2.1. Παραμετρικά τεστ

1.2.1.1 Προϋποθέσεις παραμετρικών τεστ

Για να εφαρμόσω τα παραμετρικά τεστ πρέπει να ισχύουν τέσσερις προϋποθέσεις:

Α) Είναι ΚΑΝΟΝΙΚΗ Η ΚΑΤΑΝΟΜΗ (normal distribution) των πληθυσμών που θέλω να συγκρίνω;
Β) Οι πληθυσμοί έχουν ΠΑΡΟΜΟΙΑ ΠΑΡΑΛΛΑΚΤΙΚΟΤΗΤΑ (same variance, spread);
Γ) Οι πληθυσμοί αποτελούνται από ΠΟΣΟΤΙΚΕΣ μεταβλητές (quantitative variable or ordinal-continuous scale);
Δ) Οι πληθυσμοί είναι ΑΝΕΞΑΡΤΗΤΟΙ (independence) δηλ., τα δείγματα είναι τυχαία και ανεξάρτητα μεταξύ τους και συνεπώς η τιμή του ενός δεν επηρεάζει την τιμή του άλλου;

Οι προϋποθέσεις Γ και Δ είναι λίγο έως πολύ αυτονόητες για τον ερευνητή. Για τον έλεγχο των προϋποθέσεων Α και Β ακολουθούνται τα παρακάτω βήματα:

Έλεγχος κανονικής κατανομής (προϋπόθεση Α)

Ο έλεγχος της κανονικής κατανομής των δειγμάτων μπορεί να γίνει με δύο τρόπους, ως εξής:

1ος τρόπος: Δημιουργώ τα διαγράμματα των πληθυσμών, που θέλω να συγκρίνω. Τα επικρατέστερα διαγράμματα είναι τα ιστογράμματα (histogram), τα θηκογράμματα (boxplots), κατάλληλα και για λίγα δεδομένα, και τα διαγράμματα πιθανότητας (normal Q-Q plot). Τα διαγράμματα αυτά προκύπτουν εύκολα από τη χρήση γνωστών στατιστικών πακέτων.

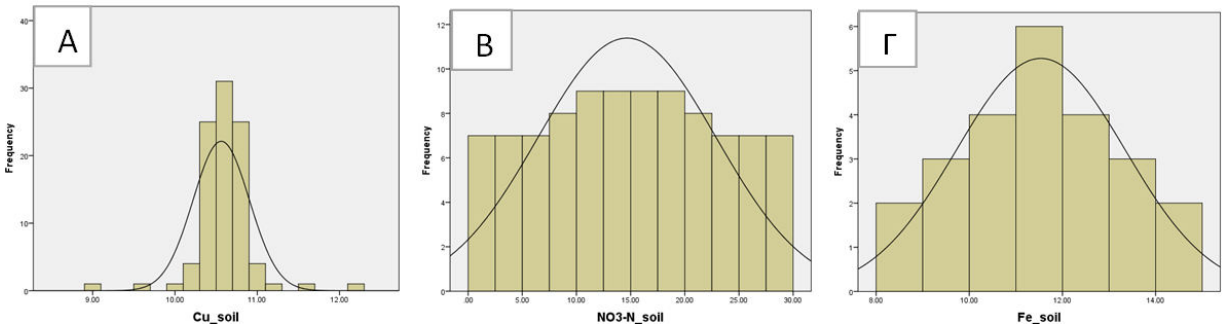
Συμπέρασμα για είδος κατανομής από ιστογράμματα

Υπάρχουν δύο είδη κατανομής: α) η συμμετρική και β) η ασύμμετρη.

α) η συμμετρική κατανομή (skewness=0) παρουσιάζεται με τρεις μορφές, ανάλογα με το βαθμό κύρτωσης (Σχήμα 1.Α,Β,Γ).

- η λεπτόκυρτη (Σχήμα 1.Α), που χαρακτηρίζεται από τη συγκέντρωση των τιμών στο κέντρο της κατανομής (συντ. κύρτωσης >3),
- η πλατύκυρτη (Σχήμα 1.Β), που χαρακτηρίζεται από μικρό βαθμό συγκέντρωσης γύρω από το κέντρο της κατανομής (συντ. κύρτωσης <3) και,
- η μεσόκυρτη (Σχήμα 1.Γ) ή κανονική κατανομή (συντ. κύρτωσης =3) με σχήμα καμπάνας (bell-shape).

ΠΑΡΑΤΗΡΗΣΗ: Η συμμετρική κατανομή δεν είναι απαραίτητα και κανονική.



Σχήμα 1. (Α) Λεπτόκυρτη κατανομή, (Β) Πλατύκυρτη κατανομή, (Γ) Μεσόκυρτη κατανομή

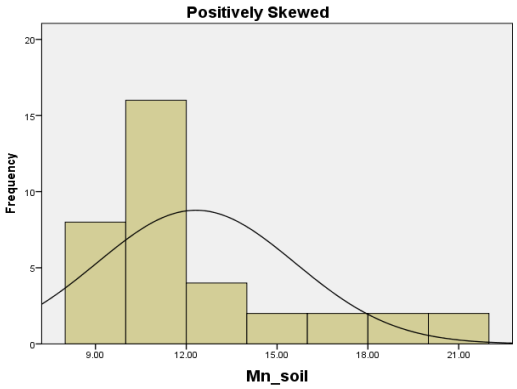
Στο στατιστικό πακέτο SPSS v.16 οι τιμές κύρτωσης δίνονται σε σχέση με το 0. Έτσι, θετικές τιμές κύρτωσης δείχνουν ότι πρόκειται για λεπτόκυρτη κατανομή ενώ αρνητικές τιμές κύρτωσης δείχνουν μία πλατύκυρτη κατανομή. Όσο πιο μακριά είναι ο συντελεστής κύρτωσης από την τιμή 0 τόσο απομακρυνόμαστε από την κανονική κατανομή.

β) στην ασύμμετρη κατανομή (Σχήματα 2,3) υπάρχουν δύο κατηγορίες :

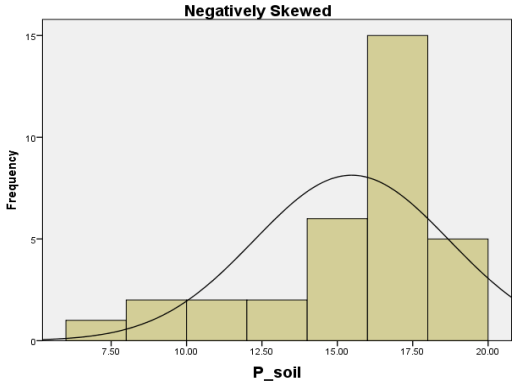
- η ασύμμετρη αριστερά (Σχ. 2, skewness>0, positively skewed) όταν οι μεγάλες συχνότητες συγκεντρώνονται στο κάτω άκρο της κατανομής (που αντιστοιχεί στις χαμηλότερες τιμές της κλίμακας μέτρησης) και ταυτόχρονα οι συχνότητες στο άνω άκρο είναι λιγότερες (εξόγκωση προς τα αριστερά και μεγάλη ουρά προς τα δεξιά).

Στην κατανομή αυτή ισχύει επίσης η σχέση: **μέσος > διάμεσος > επικρατούσα τιμή** και,

- η ασύμμετρη δεξιά (Σχ. 3, skewness<0, negatively skewed), όταν οι μεγάλες συχνότητες συγκεντρώνονται στο πάνω άκρο της κατανομής και ταυτόχρονα οι συχνότητες στο κάτω άκρο είναι λιγότερες (εξόγκωση προς τα δεξιά και μεγάλη ουρά προς τα αριστερά). Στην κατανομή αυτή ισχύει επίσης η σχέση: **επικρατούσα τιμή > διάμεσος > μέσος.**



Σχήμα 2. Ασύμμετρη αριστερά κατανομή



Σχήμα 3. Ασύμμετρη δεξιά κατανομή

Για την κατασκευή ιστογραμμάτων ακολουθώ τα παρακάτω βήματα:

Με χρήση SPSS v.16

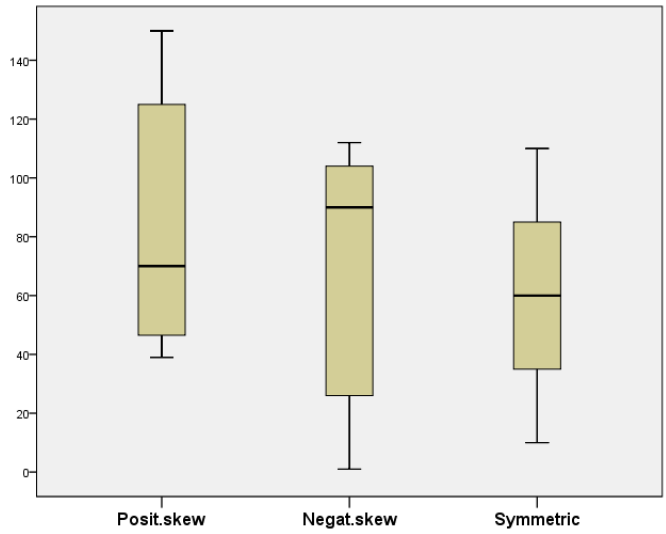
Graphs → Legacy Dialogs → Histogram. Στη συνέχεια μεταφέρω τη μεταβλητή, που θέλω να εξετάσω από το αριστερό πλαίσιο στο δεξιό πλαίσιο, που καλείται *Variable* και τσεκάρω την επιλογή *Display normal curve*.

Συμπέρασμα για είδος κατανομής από boxplots

Στο Σχήμα 4, φαίνονται τα τρία είδη κατανομής, όπως προκύπτουν από τα boxplots. Έτσι, όταν :

- η μέση τιμή βρίσκεται πάνω από τη διάμεσο στο boxplot, η κατανομή είναι ασύμμετρη αριστερά (posit. skewed) και το αντίθετο.
- το πάνω whisker του boxplot είναι μεγαλύτερο από το κάτω, η κατανομή είναι ασύμμετρη αριστερά (posit. skewed) και το αντίθετο.
- η μέση τιμή και η διάμεσος στο boxplot συμπίπτουν και τα whisker (πάνω και κάτω) έχουν ίσο μέγεθος τότε η κατανομή είναι συμμετρική.

Τέλος, στα boxplots απεικονίζονται και τα outliers (εφόσον υπάρχουν), τα οποία συμβάλλουν και αυτά στην ασυμμετρία της κατανομής.



Σχήμα 4. Είδος κατανομής με boxplot διάγραμμα (χωρίς outliers)

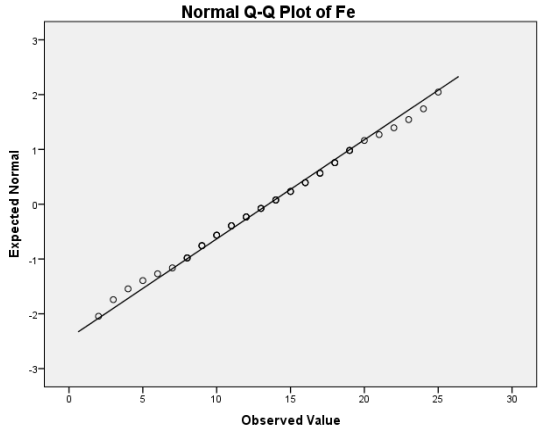
Για την κατασκευή θηκογραμμάτων ακολουθώ τα παρακάτω βήματα:

Με χρήση SPSS v.16

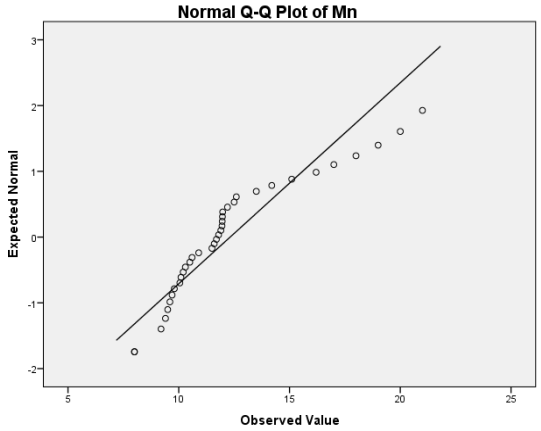
Graphs → Legacy Dialogs → Boxplot. Στη παράθυρο που ανοίγει επιλέγω το *Simple* και τσεκάρω την επιλογή *Summaries of separate variables*. Ακολούθως, κάνω κλικ στο *Define* και μεταφέρω τη μεταβλητή, που θέλω να εξετάσω από το αριστερό πλαίσιο στο δεξιό πλαίσιο, που καλείται *Boxes Represent*.

Συμπέρασμα για είδος κατανομής από Q-Q plots

Τα διαγράμματα αυτά συγκρίνουν τα δεδομένα, που έχουμε με αυτά μίας κανονικής κατανομής, χρησιμοποιώντας ένα τυπικό x-y διάγραμμα. Αν η μεταβλητή μας ακολουθεί κανονική κατανομή (Σχήμα 5α) τα σχηματιζόμενα δεδομένα (dots) πέφτουν πάνω σε μία ευθεία γραμμή (γραμμή κανονικής κατανομής). Αν αποκλίνουν από τη γραμμή είτε προς τα πάνω είτε προς τα κάτω τότε η εξεταζόμενη μεταβλητή αποκλίνει από την κανονικότητα (Σχήμα 5β).



Σχήμα 5α. Κανονική κατανομή με βάση το Q-Q plot



Σχήμα 5β. Μη κανονική κατανομή με βάση το Q-Q plot

Με χρήση SPSS v.16

Analyze → Descriptive Statistics → Explore. Τοποθετώ τις μεταβλητές, που θέλω να εξετάσω στο πλαίσιο με το όνομα *Dependent List* και στη συνέχεια κάνοντας κλικ στο πλαίσιο *plots* τσεκάρω την επιλογή *Normality plots with test*. Από τη διαδικασία προκύπτει ένας πίνακας και διαγράμματα (histogram, boxplot και Q-Q plot), που αποκαλύπτουν την ύπαρξη ή μη κανονικής κατανομής.

2ος τρόπος: Τα παραπάνω σχήματα δείχνουν μία εικόνα της κατανομής των δεδομένων. Συνήθως, στα σχήματα εμφανίζεται κάποια απόκλιση των δεδομένων από την πλήρη κανονικότητα. Αν η απόκλιση αυτή είναι μικρή ή μεγάλη προσδιορίζεται από στατιστικά τεστ (Kolmogorov-Smirnov, Shapiro-Wilk), που εκτελούνται εύκολα με τη χρήση των στατιστικών πακέτων. Τα τεστ αυτά εξετάζουν αν η μηδενική υπόθεση είναι αληθινή (δηλ. ότι οι πληθυσμοί ακολουθούν κανονική κατανομή). Αν το αποτέλεσμα του τεστ δώσει τιμή του $p>0.05$, τότε δεν απορρίπτεται η μηδενική υπόθεση και επομένως ισχύει η κανονικότητα. Αν η τιμή του $p<0.05$, τότε απορρίπτεται η μηδενική υπόθεση, δηλαδή οι πληθυσμοί δεν ακολουθούν την κανονική κατανομή.

Η διαδικασία για την υλοποίηση των παραπάνω βημάτων είναι η εξής:

Με χρήση SPSS v.16

Analyze → Descriptive Statistics → Explore. Τοποθετώ τις μεταβλητές, που θέλω να εξετάσω στο πλαίσιο με το όνομα *Dependent List* και στη συνέχεια κάνοντας κλικ στο πλαίσιο *plots* τσεκάρω την επιλογή *Normality plots with test*. Από τη διαδικασία προκύπτει ένας πίνακας και διαγράμματα (histogram, boxplot και Q-Q plot), που αποκαλύπτουν την ύπαρξη ή μη κανονικής κατανομής.

Συνοψίζοντας, ο έλεγχος της κανονικής κατανομής μπορεί να γίνει είτε με χρήση διαγραμμάτων είτε με στατιστικά τεστ.

Εφαρμόζω δύο βήματα για έλεγχο κανονικής κατανομής:

1) Διαγράμματα (histogram, boxplot, normal Q-Q plot)
2) Στατιστικά τεστ (Kolmogorov-Smirnov, Shapiro-Wilk)

Αν από τα βήματα 1 και 2, διαπιστώσω ότι τα δεδομένα ακολουθούν την κανονική κατανομή, συνεχίζω με τον έλεγχο της προϋπόθεσης Β.

Αν όμως η κατανομή δεν είναι κανονική τότε δεν ισχύει η προϋπόθεση Α για την εφαρμογή του t-test. Σε αυτή τη περίπτωση υπάρχουν δύο επιλογές, που αναλύονται σε επόμενες παραγράφους, όπως :

i) μετασχηματισμός των δεδομένων

ii) χρήση μη-παραμετρικών τεστ.

Έλεγχος παρόμοιας παραλλακτικότητας (προϋπόθεση Β)

Εφόσον τα δεδομένα ακολουθούν την κανονική κατανομή, συνεχίζω με τον έλεγχο της προϋπόθεσης Β. Αν οι τυπικές αποκλίσεις (standard deviation) των πληθυσμών, που θέλω να συγκρίνω είναι κοντά, τότε θεωρώ ότι ισχύει η προϋπόθεση της παρόμοιας παραλλακτικότητας (same variance). Το πόσο κοντά είναι οι τυπικές αποκλίσεις των πληθυσμών μπορεί να ελεγχθεί και με το Levene's test, όπου και εδώ πρέπει η τιμή του $p>0.05$, για να μην απορριφθεί η μηδενική υπόθεση (ότι δηλ. οι τυπ. αποκλίσεις είναι ίσες).

Η διαδικασία για τον έλεγχο της παραλλακτικότητας είναι η εξής:

Με χρήση SPSS v.16

Με το Levene's test, που προκύπτει ταυτόχρονα με τα αποτελέσματα σύγκρισης των μέσων δύο πληθυσμών. Για περισσότερους πληθυσμούς, πρέπει να κάνουμε κλικ στην επιλογή *Homogeneity of variance test* στο πλαίσιο *options* (βλ. 1γi).

Συνοπτικά, ο έλεγχος της παρόμοιας παραλλακτικότητας γίνεται, είτε οπτικά, παρατηρώντας τις αριθμητικές τιμές των τυπικών αποκλίσεων, είτε με στατιστικά τεστ.

Εφαρμόζω δύο βήματα για έλεγχο παρόμοιας παραλλακτικότητας:

1) Τυπικές Αποκλίσεις
2) Στατιστικά τεστ (Levene's test)

Αν από τα βήματα 1 και 2, διαπιστώσω ότι ισχύει η προϋπόθεση της ίσης παραλλακτικότητας συνεχίζω με την εφαρμογή του t-test.

Αν όμως οι παραλλακτικότητες των δειγμάτων διαφέρουν, δεν μπορεί να εφαρμοστεί το t-test. Σε αυτή τη περίπτωση υπάρχουν πάλι οι δύο επιλογές, μετασχηματισμός των δεδομένων και χρήση μη-παραμετρικών τεστ, που αναφέρθηκαν στην προηγουμένη παράγραφο.

Αν μετά τον έλεγχο διαπιστώσω ότι ισχύουν και οι τέσσερις προϋποθέσεις των παραμετρικών τεστ, τότε μόνο προχωράω στην εφαρμογή του t-test.

1.2.1.2 Παραμετρικό τεστ για σύγκριση δύο μέσων (t-test)

Υπάρχουν δύο τεχνικές δοκιμών, t-test, ανάλογα με το είδος των δειγμάτων, που χρησιμοποιούνται στον πειραματισμό. Και στις δύο τεχνικές ελέγχεται εάν η μηδενική υπόθεση είναι αληθινή, που σημαίνει ότι τα εξεταζόμενα δείγματα έχουν ίσες μέσες τιμές, δηλαδή προέρχονται από τον ίδιο πληθυσμό :

i) t-test με **ανεξάρτητα δείγματα** (independent samples), δηλ. όταν δύο διαφορετικές ομάδες πληθυσμών συμμετέχουν στον πειραματισμό

$$\boldsymbol{t} = \frac{\mathbf{x_1} - \mathbf{x_2}}{\sqrt{\frac{\mathbf{s_1}^2}{\mathbf{N_1}} + \frac{\mathbf{s_2}^2}{\mathbf{N_2}}}}$$

όπου x_1, x_2 οι μέσες τιμές των δύο δειγμάτων, s_1 και s_2 οι τυπ. αποκλίσεις των δύο δειγμάτων και N_1, N_2 το μέγεθος του δείγματος κάθε πληθυσμού.

ii) t-test με τα **ίδια δείγματα** (dependent/paired/matched samples), δηλ. όταν μία ομάδα πληθυσμού συμμετέχει στον πειραματισμό

$$\boldsymbol{t = \frac{D}{\frac{s_D}{\sqrt{N}}}}$$

όπου D είναι η διαφορά μεταξύ των μέσων τιμών των δύο δειγμάτων, s_D η τυπ. απόκλιση της διαφοράς των δύο δειγμάτων και N το μέγεθος του δείγματος.

Η διαδικασία σύγκρισης δύο μέσων περιλαμβάνει τα ακόλουθα βήματα:

Με χρήση SPSS v.16

i) για ανεξάρτητα υποκείμενα στον πειραματισμό ακολουθώ το μονοπάτι **Analyze → Compare Means → Independent-Samples T Test.** Όταν ενεργοποιηθεί το παράθυρο αυτό, μεταφέρω την μεταβλητή, που θέλω να εξετάσω από τη λίστα, στο πλαίσιο που καλείται *Test Variable(s)*. Ακολούθως, μεταφέρω την ανεξάρτητη (grouping) μεταβλητή στο πλαίσιο που καλείται *Grouping Variable*. Άμεσα ενεργοποιείται η επιλογή *Define groups*, όπου πρέπει να ορίσω αριθμητικούς κωδικούς (π.χ. 1 και 2), που πρακτικά αντιστοιχούν στους δύο μέσους που θέλω να συγκρίνω. Στα αποτελέσματα, εκτός από τη σύγκριση των δύο μέσων, γίνεται και σύγκριση των διακυμάνσεων (variances) με το Levene's test. Το SPSS δίνει το αποτέλεσμα σε δύο γραμμές, ανάλογα με το αν ισχύει ή όχι η ισότητα της παραλλακτικότητας (Equal variances assumed ή Equal variances not assumed),

ii) για ίδια υποκείμενα στον πειραματισμό ακολουθώ το μονοπάτι **Analyze → Compare Means → Paired-Samples T Test**. Όταν ενεργοποιηθεί το παράθυρο αυτό, μεταφέρω τις μεταβλητές, που θέλω να εξετάσω, από τη λίστα στο πλαίσιο που καλείται *Paired Variables*.

ΥΠΕΝΘΥΜΙΣΗ: Δύο είδη παραλλακτικότητας (variances) υπάρχουν στα δεδομένα. Η συστηματική, που οφείλεται στον πειραματισμό (factor) και η μη συστηματική, που οφείλεται σε άλλα αίτια. Όταν στη σύγκριση των μέσων χρησιμοποιούνται ανεξάρτητα δείγματα συμμετέχουν και τα δύο είδη παραλλακτικότητας. Αντίθετα, όταν στη σύγκριση των μέσων χρησιμοποιούνται τα ίδια δείγματα, η μη συστηματική παραλλακτικότητα περιορίζεται και γίνεται πιο έκδηλη η επίδραση του πειραματισμού (συστηματική παραλλακτικότητα).

Κατά το σχεδιασμό του πειράματος συνιστάται η επιλογή ενός δείγματος/ομάδας (dependent samples), ώστε να απομακρυνθεί η μη συστηματική παραλλακτικότητα, που προκαλείται από τη χρήση των ανεξάρτητων δειγμάτων.

1.2.1.3 Παραμετρικό/α τεστ για σύγκριση περισσότερων (των 2) μέσων

Αν και μόνο αν ισχύουν οι προϋποθέσεις των παραμετρικών τεστ, το ANOVA (analysis of variance/ανάλυση παραλλακτικότητας) εφαρμόζεται για την εύρεση σημαντικής διαφοράς στο σύνολο των εξεταζόμενων δειγμάτων. Ανάλογα με το είδος των δειγμάτων, που έχει συλλεχθεί, υπάρχει και το αντίστοιχο τεστ. Έτσι, για:

i) ανεξάρτητα δείγματα, το ANOVA χρησιμοποιεί το F-test για να ελέγξει την μηδενική υπόθεση (ισότητα μέσων τιμών όλων των δειγμάτων).

Συνοπτικά, η βασική αρχή της τεχνικής αυτής (F-test) είναι (όπως και στο t-test) ο διαχωρισμός της παραλλακτικότητας των δεδομένων σε δύο μέρη:

- την παραλλακτικότητα μεταξύ των εξεταζόμενων ομάδων (between groups), που οφείλεται στον πειραματισμό (SS_T). Υπολογίζεται ως το άθροισμα της διαφοράς (^2) μεταξύ της μέσης τιμής κάθε ομάδας με την ολική μέση τιμή (Μ.Ο. όλων των ομάδων) και,
- την παραλλακτικότητα εντός των ομάδων (within groups), που οφείλεται σε σφάλμα (SS_R). Υπολογίζεται ως το άθροισμα της διαφοράς (^2) μεταξύ κάθε τιμής της ομάδας με τη μέση τιμή της ομάδας.

Το άθροισμα των SS_T και SS_R αποτελεί την ολική παραλλακτικότητα (total variance) των δεδομένων (SST_{ot}). Υπολογίζεται ως το άθροισμα της διαφοράς (^2) μεταξύ κάθε τιμής της ομάδας με την ολική μέση τιμή (Μ.Ο. όλων των ομάδων). Τελικά, το F-test προκύπτει από το πηλίκο των δύο μερών της παραλλακτικότητας των δεδομένων, χρησιμοποιώντας τις μέσες τιμές των διαφορών, ως εξής:

$$\boldsymbol{F = \frac{MS_T}{MS_R}}$$

όπου $\boldsymbol{MS_T = \frac{SS_T}{k-1}}$ και $\boldsymbol{MS_R = \frac{SS_R}{n-k}}$

k-1 και **n-k** είναι οι βαθμοί ελευθερίας (df: degrees of freedom), όπου **k** ο αριθμός των ομάδων και **n** το ολικό μέγεθος του δείγματος (αριθμός μετρήσεων από όλες τις ομάδες).

Όσο μεγαλύτερη τιμή έχει ο αριθμητής στο κλάσμα του F-test τόσο μεγαλύτερη θα είναι η διαφορά, μεταξύ των ομάδων, λόγω του πειραματισμού. Μία μεγάλη λοιπόν τιμή του F-test προειδοποιεί για την ύπαρξη σημαντικής διαφοράς στο σύνολο των εξεταζόμενων πληθυσμών.

Μεγάλη τιμή του F-test μπορεί να προκύψει και από μικρή τιμή του παρονομαστή του κλάσματος, που σημαίνει ότι η παραλλακτικότητα λόγω σφάλματος, μέσα στις ομάδες, είναι μικρή.

Τέλος, για να διαπιστωθεί ποιες ομάδες είναι εκείνες που διαφέρουν μεταξύ τους, υπάρχει μία πληθώρα επιλογής από τεστ, όπως το LSD, Tukey΄s, Dunnett΄s κλπ., γνωστά και ως **post hoc** τεστ. Γενικά, υπάρχουν δύο είδη **post hoc** τεστ, ανάλογα με τον τύπο του σφάλματος, που ελέγχουν:

α) τα συντηρητικά (conservative), που περιορίζουν το σφάλμα τύπου I (απορρίπτει τη μηδενική υπόθεση), και επομένως μπορεί να μην εντοπίσουν σημαντική διαφορά μεταξύ των πληθυσμών, αν και υπάρχει,

β) τα μη συντηρητικά (liberal), που περιορίζουν το σφάλμα τύπου II (δέχεται τη μηδενική υπόθεση) και επομένως μπορεί να βρουν σημαντική διαφορά χωρίς, πράγματι, να υπάρχει.

Το είδος του τεστ, που θα επιλεγεί εξαρτάται από το είδος της έρευνας, που διεξάγεται (αν μας ενδιαφέρει δηλ., να ελέγξουμε το σφάλμα τύπου I ή II και από τα χαρακτηριστικά των εξεταζόμενων πληθυσμών. Συνοπτικά, ισχύουν τα εξής: σε ίσο μέγεθος δειγμάτων (equal sample sizes) και παρόμοιες παραλλακτικότητες **(Equal Variances Assumed)** προτιμώνται τα REGWQ, Tukey και Bonferroni τεστ (το Bonferroni τεστ είναι το πιο συντηρητικό μεταξύ των τριών). Αν οι παραλλακτικότητες είναι παρόμοιες και το μέγεθος των δειγμάτων διαφέρει τότε επιλέγονται το Gabriel's (για μικρές διαφορές στο μέγεθος) ή το Hochberg's GT2 (για μεγαλύτερες διαφορές στο μέγεθος). Αν οι παραλλακτικότητες των πληθυσμών διαφέρουν **(Equal Variances not Assumed)** τότε υπάρχει επιλογή μεταξύ των Tamhane's T2, Dunnett's T3, Games-Howel και Dunnett's C. Τα τεστ αυτά είναι συντηρητικά. Μεταξύ τους, το Games-Howel είναι προτιμότερο (έχει μεγαλύτερη στατιστική ισχύ ακόμη και σε άνισα μεγέθη δειγμάτων). Το Dunnett's τεστ είναι το μόνο συγκρίνει πληθυσμούς με έναν πληθυσμό που ορίζεται ως δείκτης (control). **Προσοχή,** στο σημείο αυτό δεν γνωρίζουμε ακόμη αν ισχύει η ομοιογένεια των παραλλακτικοτήτων, για το λόγο αυτό επιλέγουμε και τις δύο υποθέσεις.

Η διαδικασία σύγκρισης περισσότερων (των δύο) μέσων σε ανεξάρτητα δείγματα, περιλαμβάνει τα ακόλουθα βήματα:

Με χρήση SPSS v.16

Analyze → Compare Means → One Way ANOVA. Τοποθετώ τη μεταβλητή, που θέλω να εξετάσω στο πλαίσιο με το όνομα *Dependent List* και στη συνέχεια μεταφέρω την ανεξάρτητη (grouping) μεταβλητή (ορίζω αριθμητικούς κωδικούς π.χ. 1, 2, 3...n, που πρακτικά αντιστοιχούν στους n μέσους που θέλω να συγκρίνω) στο πλαίσιο που καλείται *Factor*. Στο πεδίο *Post Hoc* επιλέγω τα τεστ, που θέλω να χρησιμοποιήσω για τη σύγκριση (με βάση όσα αναφέρθηκαν στην προηγούμενη παράγραφο). Στο πεδίο options κάνω κλικ στην επιλογή *Homogeneity of variances test* για να ενεργοποιήσω το *Levene's test of Equality of Error Variances*, που παρουσιάζει τα αποτελέσματα του ελέγχου ομοιογένειας των παραλλακτικοτήτων.

ΠΑΡΑΤΗΡΗΣΗ: Τα περισσότερα τεστ είναι αξιόπιστα σε μικρές αποκλίσεις από την κανονική κατανομή. Δεν είναι όμως το ίδιο αξιόπιστα όταν οι παραλλακτικότητες (variances) και το μέγεθος των δειγμάτων (sample sizes) διαφέρουν μεταξύ τους. Η επιλογή του κατάλληλου τεστ σε αυτές τις περιπτώσεις είναι σημαντική.

ii) ίδια ομάδα δείγματος, η ακρίβεια του F-test μειώνεται καθώς δεν ισχύει η προϋπόθεση των ανεξάρτητων πληθυσμών (independent samples). Σε ένα τέτοιο πειραματισμό πρέπει να ισχύει μία αντίστοιχη προϋπόθεση, που καλείται "***sphericity***". Ο όρος αυτός σημαίνει ότι οι παραλλακτικότητες (variances) ανά ζεύγος, στις διάφορες πειραματικές συνθήκες, πρέπει να είναι περίπου ίσες. Δηλ., αν θέλω να συγκρίνω έναν πληθυσμό κάτω από τρεις διαφορετικές συνθήκες A, B, και C θα πρέπει:

παραλλακτικότητα $_{A-B}$ ≈ παραλλακτικότητα $_{A-C}$ ≈ παραλλακτικότητα $_{B-C}$

που πρακτικά αντιστοιχεί στην προϋπόθεση B (των παραμετρικών τεστ), δηλαδή, όταν ο πειραματισμός πραγματοποιείται σε ανεξάρτητους πληθυσμούς.

Ο έλεγχος της ισότητας της παραλλακτικότητας πραγματοποιείται με το λεγόμενο Mauchly's test. Αν το τεστ αυτό είναι σημαντικό ($p<0.05$) τότε η προϋπόθεση της "***sphericity***" παραβιάζεται και οι παραλλακτικότητες διαφέρουν σημαντικά μεταξύ τους. Αντίθετα, οι παραλλακτικότητες θεωρούνται ίσες όταν το τεστ αυτό δεν είναι σημαντικό ($p>0.05$).

Η διαδικασία σύγκρισης για ίδια δείγματα είναι ανάλογη με αυτήν που εφαρμόζεται για τα ανεξάρτητα δείγματα. Αρχικά, δίνεται η τιμή του Mauchly's test για να διαπιστωθεί αν ισχύει η προϋπόθεση της "***sphericity***". Αν ισχύει, τότε η τιμή του F-test θεωρείται ακριβής και εφόσον είναι σημαντική ($p<0.05$) συνεχίζω με την εκτέλεση των post hoc tests για να εντοπίσω σε ποια ζεύγη μετρήσεων υπάρχουν διαφορές ή δεν εκτελώ τα post hoc tests αν η τιμή του F-test είναι στατιστικώς ασήμαντη ($p>0.05$). Αν δεν ισχύει η προϋπόθεση της "***sphericity***" τότε η τιμή του F-test δεν θεωρείται ακριβής

και πρέπει να γίνει κάποια διορθωτική επέμβαση. Τα στατιστικά πακέτα προσφέρουν διάφορα τεστ για το σκοπό αυτό, όπως των Greenhouse-Geisser (conservative) ή Huynh-Feldt (liberal). Με τα τεστ αυτά επαναϋπολογίζονται οι βαθμοί ελευθερίας (df) και με βάση τους νέους βαθμούς, υπολογίζεται εκ νέου, το F-test.

ΠΡΟΣΟΧΗ: Επειδή τα δύο τεστ έχουν διαφορετική στατιστική ισχύ (conservative, liberal) μπορεί να καταλήξουν σε διαφορετικό αποτέλεσμα. Σε αυτή την περίπτωση προτείνεται να ληφθεί υπόψη ο μέσος όρος των δύο τιμών του p.

Η διαδικασία σύγκρισης περισσότερων (των δύο) μέσων στα ίδια δείγματα περιλαμβάνει τα ακόλουθα βήματα:

Με χρήση SPSS v.16

Analyze → General Linear Model → Repeated Measures. Στο παράθυρο που ανοίγει ως *Define Factor(s)*, ονοματίζω τη (repeated measures) μεταβλητή στο πλαίσιο με το όνομα *Within-Subject Factor Name* και τον αριθμό των επαναλαμβανόμενων μετρήσεων στο πλαίσιο *Number of Levels*. Κάνω κλικ στο *Add* και κατόπιν στο *Define* για να ενεργοποιηθεί το τελικό παράθυρο, που περιλαμβάνει τα *Post Hoc* tests για τη σύγκριση των μέσων. Ωστόσο, για την ανάλυση αυτή δεν κάνουμε κλικ στα *Post Hoc* αλλά στο *Options*. Αφού ανοίξει αυτό το παράθυρο, μεταφέρουμε τη μεταβλητή από το αριστερό πλαίσιο *Factor(s) and Factor Interactions* στο δεξιό πλαίσιο με το όνομα *Display Means For*. Άμεσα, ενεργοποιείται η επιλογή *Compare main effects* ώστε να μπορούμε να επιλέξουμε από το πλαίσιο *Confidence interval adjustment* ένα από τα τρία διαθέσιμα post hoc test (LSD, Bonferroni, Sidak). Συνήθως επιλέγουμε Bonferroni. Τα αποτελέσματα της διαδικασίας αυτής περιλαμβάνουν τόσο το Mauchly's test, όσο και τα Greenhouse-Geisser, Huynh-Feldt tests, που ελέγχουν την προϋπόθεση της "***sphericity***", και διορθώνουν τη τιμή του F-test, αντίστοιχα, αν η παραπάνω προϋπόθεση απορριφθεί.

ΠΑΡΑΤΗΡΗΣΗ: Ο πειραματισμός τόσο στα ανεξάρτητα δείγματα (independent samples) όσο και στα ίδια δείγματα (repeated measures) μπορεί να περιλαμβάνει δύο ανεξάρτητες μεταβλητές **(two-way ANOVA)** ή περισσότερες **(factorial ANOVA)**. Στην περίπτωση με τα ανεξάρτητα δείγματα ακολουθούμε το μονοπάτι **Analyze → General Linear Model →Univariate** (όπως και στο One-Way ANOVA, με τη διαφορά ότι τώρα έχω περισσότερες ανεξάρτητες μεταβλητές (factors) τις οποίες και μεταφέρω στο πλαίσιο δεξιά *Fixed Factor(s)*. Κάνοντας κλικ στο πλαίσιο *Model* και συνέχεια στο *Custom* επιλέγουμε *interaction* (αλληλεπίδραση) ή *main effects* (μη αλληλεπίδραση) των ανεξάρτητων μεταβλητών, μεταξύ τους. Στην περίπτωση με τα ίδια δείγματα ακολουθούμε το μονοπάτι **Analyze → General Linear Model → Repeated Measures** (όπως παραπάνω, με τη διαφορά ότι στο πλαίσιο με το όνομα *Within-Subject Factor Name* επιλέγω μία προς μία όλες τις ανεξάρτητες μεταβλητές, ορίζοντας κάθε φορά και

τον αριθμό των επαναλαμβανόμενων μετρήσεων στο πλαίσιο *Number of Levels*). Τέλος, ακόμα πιο σύνθετος πειραματισμός, με ίδια και ανεξάρτητα δείγματα ταυτόχρονα (Mixed Design ANOVA), μπορεί να πραγματοποιηθεί.

Αν μετά τον έλεγχο διαπιστώσω ότι ισχύουν και οι τέσσερις προϋποθέσεις των παραμετρικών τεστ, τότε μόνο προχωράω στην εφαρμογή του ANOVA.

Διαφορετικά μετασχηματίζω τα δεδομένα ή χρησιμοποιώ μη-παραμετρικά τεστ.

1.2.2. Μετασχηματισμός δεδομένων (transformation)

Αν οι πληθυσμοί αποκλίνουν από την κανονική κατανομή (προϋπόθεση Α) ή οι τυπικές αποκλίσεις διαφέρουν σημαντικά μεταξύ τους (προϋπόθεση Β), τότε πρέπει να κάνω μετασχηματισμό των δεδομένων (π.χ. λογαριθμικό, εκθετικό, τετρ. ρίζα κ.ά). Η χρήση τετραγωνικής ρίζας συνήθως φέρει αποτελέσματα. Με τον μετασχηματισμό "διορθώνονται" και οι ακραίες τιμές (outliers, extreme values) και επιτυγχάνεται (δυστυχώς όχι πάντα) το επιθυμητό αποτέλεσμα (προϋποθέσεις Α και Β).

Η διαδικασία για τον μετασχηματισμό των δεδομένων είναι η εξής:

Με χρήση SPSS v.16

ακολουθώ το μονοπάτι **Transform → Compute Variable.** Στο παράθυρο που ανοίγει, πληκτρολογώ το όνομα της νέας μεταβλητής (π.χ. sqrt) στο πλαίσιο *Target Variable*. Στο πλαίσιο δεξιά, που καλείται *Numeric Expression* εισάγεται η αριθμητική εντολή. Η εισαγωγή γίνεται από τη λίστα ***Functions***, που βρίσκεται παρακάτω**.** Δίπλα στο όνομα της μεταβλητής, στη θέση του ερωτηματικού sqrt (?) μεταφέρουμε από το πλαίσιο *Type & Label* την παλιά μεταβλητή τα δεδομένα της οποίας θέλουμε να μετασχηματιστούν.

ΠΑΡΑΤΗΡΗΣΗ : Εναλλακτικά του μετασχηματισμού, μπορεί να γίνει χρήση των μη-παραμετρικών τεστ, όπως προαναφέρθηκε, που εφαρμόζονται κυρίως όταν οι πληθυσμοί αποκλίνουν από την κανονική κατανομή. Η δυνατότητά τους αυτή οφείλεται στο ότι χρησιμοποιούν για το στατιστικό έλεγχο της μηδενικής υπόθεσης είτε κατανομές πληθυσμών είτε τη διάμεσο αντί της μέσης τιμής. Η διάμεσος είναι πιο ευέλικτη από τη μέση τιμή καθώς δεν επηρεάζεται από ακραίες τιμές (outliers, extreme values) ή από ασύμμετρες κατανομές (skewness).

1.2.3. Μη-παραμετρικά τεστ

1.2.3.1 Προϋποθέσεις μη-παραμετρικών τεστ

Ανάλογα με το είδος του δείγματος, που συλλέγεται, πρέπει να ισχύουν οι εξής προϋποθέσεις:

i) για τα **ανεξάρτητα δείγματα**, να έχουν ίδιο σχήμα κατανομής καθώς και παρόμοια παραλλακτικότητα (spread, variance). Όσο αφορά το σχήμα κατανομής, τα boxplots είναι κατάλληλα για μη-παραμετρικά τεστ γιατί απεικονίζουν τη διάμεσο. Τα ιστογράμματα απεικονίζουν καλά το σχήμα της κατανομής αλλά όχι τόσο καλά τη θέση του κέντρου (μέσος ή διάμεσος).

ii) για τα **ίδια δείγματα**, αυτά να αποτελούν ποσοτική μεταβλητή (ordinal ή quantitative) και όχι μεταβλητή κατηγορίας (δηλ. yes or no).

1.2.3.2 Μη-παραμετρικό/α τεστ για σύγκριση δύο μέσων

Όπως και στην περίπτωση των παραμετρικών δεδομένων, υπάρχουν και εδώ δύο τεχνικές σύγκρισης ανάλογα με το είδος των δειγμάτων (subjects), που χρησιμοποιείται στον πειραματισμό:

i) το Mann-Whitney (ή rank sum) test με **ανεξάρτητα δείγματα**, δηλ. όταν δύο διαφορετικές ομάδες συμμετέχουν στον πειραματισμό και,

ii) το Wilcoxon (ή signed-rank) test με τα **ίδια δείγματα**, δηλ. όταν μία ομάδα συμμετέχει στον πειραματισμό.

Συνοπτικά, η βασική φιλοσοφία των μη-παραμετρικών τεστ είναι ότι δεν χρησιμοποιούν τα πραγματικά δεδομένα αλλά μία τεχνική, που λέγεται "**ranked**". Με αυτή την τεχνική, η μικρότερη μέτρηση των πειραματικών δεδομένων παίρνει την τιμή 1, η αμέσως μεγαλύτερη την τιμή 2 και ούτω καθεξής. Το μειονέκτημα της τεχνικής αυτής είναι η πιθανή έλλειψη σημαντικής πληροφορίας από τη μη χρήση των πραγματικών δεδομένων. Αυτό οδηγεί σε μείωση της ισχύος των μη-παραμετρικών τεστ σε σχέση με τα παραμετρικά (δηλ. μπορεί να υπάρχει σημαντική διαφορά στους πληθυσμούς και να μην εντοπιστεί). Η μηδενική υπόθεση στα μη-παραμετρικά τεστ δέχεται ότι οι **διάμεσοι** των δύο πληθυσμών είναι ίσοι. Το αποτέλεσμα απορρίπτει (reject) ή όχι (retain) τη μηδενική υπόθεση.

Η διαδικασία σύγκρισης δύο μέσων περιλαμβάνει τα ακόλουθα βήματα:

Με χρήση SPSS v.16

i) για ανεξάρτητα υποκείμενα στον πειραματισμό ακολουθώ το μονοπάτι **Analyze → Nonparametric Tests → Independent Samples.** Όταν ενεργοποιηθεί το παράθυρο αυτό, επιλέγω το πεδίο *Fields* και μεταφέρω τις μεταβλητές, που θέλω να εξετάσω στο πλαίσιο, που καλείται *Test Fields*. Ακολούθως, μεταφέρω την ανεξάρτητη (grouping) μεταβλητή στο πλαίσιο που καλείται *Groups*. Άμεσα ενεργοποιείται η επιλογή *Define groups*, όπου πρέπει να ορίσω αριθμητικούς κωδικούς (π.χ. 1 και 2), που πρακτικά αντιστοιχούν στους δύο μέσους που θέλω να συγκρίνω. Η επιλογή *Mann-Whitney U* στο πλαίσιο *Test Type* είναι προεπιλεγμένη,

ii) για ίδια υποκείμενα στον πειραματισμό ακολουθώ το μονοπάτι **Analyze → Nonparametric Tests → Related Samples**. Όταν ενεργοποιηθεί το παράθυρο αυτό, επιλέγω το πεδίο *Fields* και μεταφέρω τις μεταβλητές, που θέλω να εξετάσω στο πλαίσιο, που καλείται *Test Fields*.

ΠΑΡΑΤΗΡΗΣΕΙΣ:

- Ο μικρός αριθμός δείγματος δεν ευνοεί την εύρεση σημαντικής διαφοράς.
- Αν οι τυπ. αποκλίσεις είναι μεγάλες σε σχέση με τη μέση τιμή (ή τη διάμεσο), δεν ευνοείται η εύρεση σημαντικής διαφοράς.

1.2.3.3 Μη-παραμετρικό τεστ για σύγκριση περισσότερων (των 2) μέσων

Σε μη-παραμετρικά δεδομένα, στη θέση του F-test, μπορεί να χρησιμοποιηθεί το μη-παραμετρικό τεστ των Kruskal-Wallis (KW) για την εύρεση σημαντικής διαφοράς στο σύνολο των group. Όπως προαναφέρθηκε, οι τιμές των πραγματικών δεδομένων όλων των group κατατάσσονται σε σειρά από τη μικρότερη ως τη μεγαλύτερη παίρνοντας νέες τιμές 1, 2 και ούτω καθεξής (ranked data).

Αφού υπολογιστεί η τιμή του KW (με βάση αριθμητικό τύπο) χρησιμοποιείται η x^2-κατανομή. Αν για την τιμή του KW προκύψει τιμή $p>0.05$, δεν απορρίπτεται η μηδενική υπόθεση δηλ., οι πληθυσμοί δεν διαφέρουν σημαντικά μεταξύ τους. Αν όμως η τιμή του $p<0.05$, τότε το τεστ απορρίπτει την μηδενική υπόθεση στο σύνολο των group. Σε αυτή την περίπτωση ακολουθεί το Wilcoxon rank sum τεστ (ή Mann-Whitney test), που συγκρίνει τις διαμέσους των group (ανά δύο) για να βρει ποια group διαφέρουν μεταξύ τους. Η διαδικασία στο SPSS είναι παρόμοια με αυτή, που περιεγράφηκε στην παράγραφο για σύγκριση δύο μέσων.

ΠΑΡΑΤΗΡΗΣΗ: Εκτός από σύγκριση διαμέσων μπορεί να συγκριθούν οι κατανομές των πληθυσμών. Μάλιστα στο SPSS η επιλογή αυτή (automatically compare distributions across groups) είναι προεπιλεγμένη.

2. ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ (LINEAR REGRESSION)

2.1. Θεωρία

Με την εργασία αυτή επιχειρείται να βρεθεί η βέλτιστη γραμμή (μοντέλο), που συσχετίζει μία εξαρτημένη μεταβλητή (y), με μία (x) ή περισσότερες ανεξάρτητες μεταβλητές (x1, x2, x3,...). Ως βέλτιστη γραμμή θεωρείται εκείνη, που ελαχιστοποιεί το σύνολο της διαφοράς μεταξύ των πραγματικών δεδομένων και των υπολογισμένων από τη γραμμή αυτή. Οι διαφορές μεταξύ των πραγματικών δεδομένων και των υπολογισμένων από τη βέλτιστη γραμμή ονομάζονται σφάλματα (residuals).

Για την εύρεση της βέλτιστης γραμμής χρησιμοποιείται το F-test. Συνοπτικά, το F-test είναι ένα μέτρο, που δείχνει σε ποιο βαθμό το μοντέλο έχει βελτιώσει την πρόβλεψη (σε σχέση με τη χρήση της μέσης τιμής της y μεταβλητής ως μοντέλου πρόβλεψης) συγκρινόμενο πάντα με το σφάλμα, που εξακολουθεί να υπάρχει με τη χρήση του μοντέλου αυτού. Όπως περιεγράφηκε και στην εργασία σύγκρισης των μέσων, το F-test χρησιμοποιεί τις μέσες τιμές των αθροισμάτων των τετραγώνων (mean sums of squares) και υπολογίζεται ως εξής:

$$\boldsymbol{F = \frac{MS_M}{MS_R}}$$

όπου $\boldsymbol{MS_M = \frac{SS_M}{k}}$ και $\boldsymbol{MS_R = \frac{SS_R}{n-k}}$

Ο όρος SS_M (sum of squares model) αντιπροσωπεύει τις διαφορές, που προκύπτουν στην πρόβλεψη της εξαρτημένης μεταβλητής (y), μεταξύ δύο μοντέλων πρόβλεψης, δηλ., της μέσης τιμής της y και της βέλτιστης γραμμής.

Ο όρος SS_R (sum of squares error) αντιπροσωπεύει το σφάλμα, που εξακολουθεί να υπάρχει με τη χρήση του μοντέλου, δηλαδή είναι οι διαφορές μεταξύ των δεδομένων και των προβλεπόμενων από το μοντέλο τιμών.

k και **n-k** είναι οι βαθμοί ελευθερίας (df: degrees of freedom), όπου **k** ο αριθμός των ανεξάρτητων μεταβλητών και **n** ο αριθμός των παρατηρήσεων (δηλ., n-k είναι ο αριθμός των beta συντελεστών συμπεριλαμβανομένης της σταθεράς, στο μοντέλο πρόβλεψης).

Όσο μεγαλύτερη τιμή έχει ο αριθμητής στο κλάσμα του F-test τόσο μεγαλύτερη θα είναι η βελτίωση στην πρόβλεψη της τιμής της y, λόγω του μοντέλου (σε σχέση με τη χρήση της μέσης τιμής της y, ως μοντέλου πρόβλεψης). Όσο μικρότερη τιμή έχει ο παρονομαστής, τόσο μικρότερο είναι το σφάλμα, που εμπεριέχεται στο μοντέλο.

Γενικά, όσο μεγαλύτερη είναι η τιμή του F-test (>1 τουλάχιστον), τόσο καλύτερο είναι το μοντέλο πρόβλεψης.

2.2. Εφαρμογή

2.2.1. Απλή γραμμική παλινδρόμηση (Simple Linear Regression)

Με την εργασία αυτή επιχειρείται να βρεθεί η βέλτιστη γραμμή (μοντέλο), που συσχετίζει την εξαρτημένη μεταβλητή (y) με μία ανεξάρτητη μεταβλητή (x). Ακολουθώ τα εξής βήματα:

i) φτιάχνω ένα διάγραμμα (scatterplot) των δύο μεταβλητών για να δω αν υπάρχει γραμμική σχέση μεταξύ τους και,

ii) βρίσκω το συντελεστή συσχέτισης για να διαπιστώσω αν η σχέση αυτή είναι στατιστικά σημαντική ($p<0.05$).

Ο συνδυασμός των δύο παραπάνω βημάτων είναι απαραίτητος, γιατί: το scatterplot μπορεί να δείξει γραμμική σχέση μεταξύ των δύο μεταβλητών αλλά η σχέση αυτή να μην είναι σημαντική, όπως αποδειχθεί από το συντελεστή συσχέτισης. Από την άλλη, μπορεί ο συντελεστής συσχέτισης να είναι υψηλός (0.9) αλλά οι δύο μεταβλητές να μην έχουν γραμμική σχέση, όπως μπορεί να αποκαλυφτεί από το scatterplot. Η περίπτωση της απλής μη γραμμικής παλινδρόμησης αναφέρεται συνοπτικά στο τέλος της παραγράφου.

Με χρήση SPSS v.16
Graphs → Legacy Dialogs → Scatter/Dot. Στο παράθυρο που ανοίγει (είναι προεπιλεγμένο το simple scatter) κάνω κλικ στο Define, όπου και μεταφέρω την εξαρτημένη μεταβλητή στο Y-Axis και την ανεξάρτητη μεταβλητή στο X-Axis. Για να βρω το συντελεστή συσχέτισης ακολουθώ το μονοπάτι **Analyze → Correlate→ Bivariate**. Στο παράθυρο που ανοίγει, μεταφέρω τις μεταβλητές που θέλω να διερευνήσω τη συσχέτισή τους από το αριστερό πλαίσιο στο δεξιό πλαίσιο που καλείται *Variables*.

ΠΑΡΑΤΗΡΗΣΕΙΣ

- Ο επικρατέστερος συντελεστής συσχέτισης είναι ο Pearson (r) και παίρνει τιμές από -1 έως 1. Αν έχει τιμή κοντά στο μηδέν, σημαίνει ότι δεν υπάρχει γραμμική σχέση μεταξύ των μεταβλητών. Μπορεί, όμως, οι δύο μεταβλητές να έχουν κάποιο άλλο είδος συσχέτισης (μη γραμμικής). Οι συντελεστές συσχέτισης, Spearman's Rho και Kendall's Tau του SPSS χρησιμοποιούνται όταν παραβιάζονται οι προϋποθέσεις των παραμετρικών τεστ. Και οι δύο χρησιμοποιούν "ranked" δεδομένα. Ο συντελεστής Kendall's Tau προτιμάται όταν έχω μικρό αριθμό δεδομένων με μεγάλο αριθμό "tied ranks" δηλ., πολλές ίδιες τιμές "ranks".

- Η μηδενική υπόθεση, που εξετάζεται είναι ότι οι δύο μεταβλητές δεν συσχετίζονται.

ΠΡΟΣΟΧΗ: Η εύρεση της βέλτιστης γραμμής (simple regression analysis) επιτυγχάνεται εύκολα με τη χρήση στατιστικών πακέτων (βλ. παρακάτω). Η εργασία αυτή (μέσω του F-test) όμως, δεν τελειώνει εδώ. Πρέπει να αξιολογηθεί το μοντέλο ελέγχοντας τα εξής:

i) το μοντέλο περιγράφει ικανοποιητικά τα δεδομένα ή επηρεάζεται από κάποιες ακραίες τιμές αυτών (outliers) και,

ii) πρέπει να ισχύουν δύο βασικές προϋποθέσεις (βλ. παρακάτω) πριν δεχτώ ότι το μοντέλο είναι αμερόληπτο (unbiased), που σημαίνει ότι οι πιθανότητες να ισχύει και σε ένα άλλο σετ δεδομένων (πληθυσμό) αυξάνονται:

Έλεγχος αξιολόγησης του μοντέλου

i) *Ακρίβεια μοντέλου και επηρεασμός από ακραίες τιμές*

Εάν τα "standardized residuals" υπερβαίνουν την τιμή ±2.5 σε ποσοστό >1% του δείγματος, θεωρούμε ότι το σφάλμα στο μοντέλο είναι μη αποδεκτό, δηλ., το μοντέλο δεν "ταιριάζει" καλά στο δείγμα. Επιπλέον, εάν τα "standardized residuals" υπερβαίνουν την τιμή ±2.0 σε ποσοστό >5% του δείγματος, θεωρούμε ότι το σφάλμα στο μοντέλο είναι μη αποδεκτό, δηλ., το μοντέλο αποτελεί κακή περιγραφή των δεδομένων. Τα παραπάνω ισχύουν καθώς σε μία κανονική κατανομή, το 95% των τιμών των "standardized residuals" βρίσκεται μεταξύ -2.0 και +2.0 και το 99% πρέπει να κυμαίνεται μεταξύ -2.5 και +2.5. Επομένως, μία τιμή των "standardized residuals" μεγαλύτερη από 3 αποτελεί αιτία προβληματισμού (γιατί τόσο υψηλή τιμή είναι απίθανο να προκύπτει από τύχη).

Δεδομένα των οποίων τα "standardized residuals" δίνουν ακραίες τιμές ελέγχονται για το αν η απομάκρυνσή τους από το δείγμα θα επηρεάσει ή όχι το μοντέλο δηλ., θα μεταβάλλουν τους συντελεστές παλινδρόμησης του μοντέλου. Τα στατιστικά πακέτα παρέχουν τεστ (Cook's, leverage, Mahalanobis distances, Covariance ratio (CVR), Standardized DFBeta), που αξιολογούν τα δεδομένα αυτά. Γενικά, κάποιο δεδομένο με ακραία τιμή "standardized residual" δεν επηρεάζει το μοντέλο, αν η τιμή του Cook's <1, το leverage (κυμαίνεται από 0-1) είναι κοντά στο 0, το Mahalanobis <11 (η τιμή εξαρτάται από το μέγεθος του δείγματος και τον αριθμό των ανεξάρτητων μεταβλητών), το Standardized DFBeta <1.

ii) *Επέκταση μοντέλου*

Οι δύο βασικές προϋποθέσεις επέκτασης του μοντέλου ελέγχονται με βάση τα σφάλματα, που δημιουργούνται από τη γραμμή πρόβλεψης (εκτιμώμενη και πραγματική τιμή της y) και είναι:

Α) Τα σφάλματα να έχουν κανονική κατανομή

B) Ομοσκεδαστικότητα σφαλμάτων, δηλαδή τα σφάλματα να έχουν σταθερή παραλλακτικότητα για κάθε τιμή της μεταβλητής x .

Οι προϋποθέσεις αυτές ελέγχονται εύκολα με διαγράμματα (plots) των στατιστικών πακέτων. Ο έλεγχος της κανονικής κατανομής των σφαλμάτων μπορεί εκτός από το ιστόγραμμα να επιβεβαιωθεί και με το τεστ Kolmogorov-Smirnov.

Το διάγραμμα, που απεικονίζει τον έλεγχο της ομοσκεδαστικότητας (ZRESID/ZPRED) αποκαλύπτει, επιπλέον, και την ύπαρξη ή μη της γραμμικής σχέσης των δύο μεταβλητών (linearity). Αν εφαρμόσουμε ένα γραμμικό μοντέλο σε μεταβλητές που δεν έχουν γραμμική συσχέτιση, αυτό περιορίζει την επέκταση της εφαρμογής του.

Όλες οι παραπάνω εργασίες υλοποιούνται εύκολα με τη χρήση στατιστικών πακέτων, όπως φαίνεται στη συνέχεια:

Με χρήση SPSS v.16

Analyze → Regression → Linear. Στο παράθυρο που ανοίγει, μεταφέρω την εξαρτημένη μεταβλητή στο πλαίσιο με το όνομα *Dependent* και την ανεξάρτητη μεταβλητή στο πλαίσιο με το όνομα *Independent(s)*. Στη συνέχεια, κάνω κλικ στο *Statistics* και τσεκάρω την επιλογή *Casewise diagnostics*, όπου μπορώ να κάνω τον εντοπισμό outliers με την επιλογή *Outliers outside* (± 2 ή 3) *standard deviations*. Κατόπιν, κάνω κλικ στο Save και από το παράθυρο που ανοίγει τσεκάρω από τις διαθέσιμες επιλογές για Predicted Values (Unstandardized, Standardized), για Residuals (Standardized), για Distances (Mahalanobis, Cook's, Leverage values) και για Influence Statistics (Standardized DfBeta(s), Standardized DFFIT, CVR). Ακολούθως, κάνω κλικ στο *Plots* και μεταφέρω τη μεταβλητή ZRESID στον y-άξονα και τη μεταβλητή ZPRED στον x-άξονα. Κατόπιν, τσεκάρω τις επιλογές *Histogram* και *Normal probability plot*. Τέλος, για να εντοπίσω εύκολα, ποιά σημεία είναι τα outliers και αν επηρεάζουν τη δομή του μοντέλου, ακολουθώ το μονοπάτι : **Analyze → Reports → Case Summaries.** Μεταφέρω τις μεταβλητές που με ενδιαφέρουν (Mahalanobis, Cook's, Leverage values, Standardized DfBeta(s), Standardized DFFIT, CVR) δεξιά στο πλαίσιο με το όνομα *Variables*. Κάνω κλικ στην επιλογή *Show case numbers*.

Ακόμα και αν δεν είμαστε βέβαιοι ότι το μοντέλο, που προέκυψε από το δείγμα μας μπορεί να επαναληφθεί και σε ένα άλλο δείγμα, υπάρχουν απλοί τρόποι με τους οποίους μπορούμε να το αξιολογήσουμε (cross-validation). Ο ένας τρόπος είναι μέσω του δείκτη, που καλείται **adjusted R^2**. Ο δείκτης αυτός δίνεται κατά τη διαδικασία εύρεσης του μοντέλου, όπως περιεγράφηκε παραπάνω. Έτσι, ενώ το R^2 δείχνει πόση από τη μεταβλητότητα της Y ερμηνεύεται από το παραγόμενο μοντέλο, που βασίστηκε σε συγκεκριμένο δείγμα, το **adjusted R^2** δείχνει πόση μεταβλητότητα της Y θα

ερμηνευόταν εάν το μοντέλο προερχόταν από τον πληθυσμό από τον οποίο πάρθηκε το παραπάνω δείγμα. Ο δεύτερος τρόπος περιλαμβάνει σπάσιμο των δεδομένων σε δύο ίσα μέρη με υπολογισμό δύο μοντέλων πρόβλεψης και σύγκριση αυτών μεταξύ τους. Το πρόβλημα, με τον τρόπο αυτό, εντοπίζεται συνήθως στο γεγονός ότι σπάνια έχουμε μεγάλο σετ δεδομένων για να πραγματοποιήσουμε αυτού του είδους την ανάλυση.

ΠΑΡΑΤΗΡΗΣΗ: Ο έλεγχος της επέκτασης του μοντέλου πρόβλεψης χρησιμοποιεί τιμές των "standardized residuals".

ΠΡΟΣΟΧΗ: Δεν κάνω επέκταση (extrapolation) του μοντέλου σε δεδομένα, που απέχουν πολύ από τα αρχικές τιμές (ή εκτός λογικών ορίων).

Εφαρμόζω δύο βήματα για
την αξιολόγηση του μοντέλου

1) Εντοπισμός ακραίων τιμών και αντιμετώπισή τους
(με στατιστικά τεστ)
2) Έλεγχος σφαλμάτων για κανονική κατανομή
& ομοσκεδαστικότητα
(με διαγράμματα)

2.2.2. Απλή μη γραμμική παλινδρόμηση

Όπως προαναφέρθηκε, το scatterplot είναι ο καλύτερος τρόπος αποκάλυψης του είδους της σχέσης μεταξύ δύο μεταβλητών. Σε περίπτωση μη γραμμικής σχέσης, οι επικρατέστερες σχέσεις είναι η πολυωνυμική (συνήθως $2^{ου}$ βαθμού) και η εκθετική. Η διαδικασία για την εύρεση μη γραμμικών μοντέλων γίνεται εύκολα με τη χρήση στατιστικών πακέτων, ως εξής:

Με χρήση SPSS v.16

Analyze → Regression → Curve Estimation. Στο παράθυρο που ανοίγει, μεταφέρω την εξαρτημένη μεταβλητή στο πλαίσιο με το όνομα *Dependent* και την ανεξάρτητη μεταβλητή στο πλαίσιο με το όνομα *Independent*. Επιλέγω ένα ή περισσότερα από τα διαθέσιμα μοντέλα, που εμφανίζονται. Τσεκάρω τις επιλογές *Include constant in equation*, *plot models* και *Display ANOVA table*. Στη συνέχεια, κάνω κλικ στο Save και από το παράθυρο που ανοίγει τσεκάρω τις διαθέσιμες επιλογές για Predicted Values, για Residuals και ορίζω το διάστημα εμπιστοσύνης (95%).

Για πιο σύνθετες σχέσεις μεταξύ των δύο μεταβλητών υπάρχει το μονοπάτι: **Analyze → Regression → Nonlinear,** όπου στο πλαίσιο που καλείται *Model Expression* μπορεί να εισαχθεί οποιασδήποτε μορφής σχέση.

ΠΡΟΣΟΧΗ: Τα μη γραμμικά μοντέλα πρέπει να αξιολογηθούν ως προς την δυνατότητα ευρύτερης εφαρμογής στον πληθυσμό. Οι προϋποθέσεις και οι διαδικασίες ελέγχου είναι ανάλογες με αυτές των γραμμικών μοντέλων (χρήση των residual plots και στατιστικών τεστ).

2.2.3. Πολλαπλή γραμμική παλινδρόμηση (Multiple linear regression)

Με την εργασία αυτή επιχειρώ να βρω τη βέλτιστη γραμμή (ή μοντέλο), που συσχετίζει την εξαρτημένη μεταβλητή (y) με δύο η περισσότερες ανεξάρτητες μεταβλητές (x1, x2...). Ακολουθώ τα ίδια βήματα που περιεγράφηκαν στην απλή παλινδρόμηση :

i) φτιάχνω διαγράμματα (scatterplots) μεταξύ της εξαρτημένης μεταβλητής με κάθε μία από τις ανεξάρτητες μεταβλητές για να δω αν υπάρχει γραμμική σχέση μεταξύ τους και,

ii) βρίσκω το συντελεστή συσχέτισης για να διαπιστώσω αν οι σχέσεις αυτές είναι στατιστικά σημαντικές ($p<0.05$).

ΠΡΟΣΟΧΗ: Πριν προχωρήσω στην εύρεση του μοντέλου πρέπει να ελέγξω αν υπάρχει υψηλή συσχέτιση μεταξύ των ανεξάρτητων μεταβλητών ($r>0.7$). Αν π.χ. δύο ανεξάρτητες μεταβλητές συσχετίζονται πολύ μεταξύ τους επιλέγω τη μία από αυτές για τη δημιουργία του μοντέλου. Διαφορετικά δημιουργούνται προβλήματα στο "χτίσιμο" του μοντέλου καθώς θα προκύπτουν πολλές διαφορετικές τιμές για τους συντελεστές (coefficients) των ανεξάρτητων μεταβλητών. Το πρόβλημα αυτό είναι γνωστό ως πολυσυγγραμικότητα (multicollinearity) και το αποτέλεσμα είναι η αστάθεια του μοντέλου και η αδυναμία επέκτασής του σε άλλα δεδομένα.

ΠΑΡΑΤΗΡΗΣΗ: Η διαδικασία εύρεσης του μοντέλου πρόβλεψης στην πολλαπλή παλινδρόμηση είναι ίδια με αυτή της απλής παλινδρόμησης.

Έλεγχος αξιολόγησης του μοντέλου

Η αξιολόγηση μοντέλου πολλαπλής γραμμικής παλινδρόμησης αποτελεί αρκετά σύνθετη διαδικασία και απαιτεί ιδιαίτερη προσοχή και συστηματική προσέγγιση. Έτσι, πέρα των προϋποθέσεων, όπως αυτές αναφέρθηκαν στην περίπτωση της απλής γραμμικής παλινδρόμησης, για τη γενίκευση ενός τέτοιου μοντέλου πρέπει να ισχύουν επιπλέον προϋποθέσεις, όπως η απουσία πολυσυγγραμικότητας, η ανεξαρτησία σφαλμάτων, η απουσία τιμών (σημείων) μόχλευσης, η απουσία σφάλματος εξειδίκευσης κ.ά. Το παρόν εγχειρίδιο επικεντρώνεται σε 4 από τις απαραίτητες προϋποθέσεις:

Α) Τα σφάλματα να έχουν κανονική κατανομή

Β) Ομοσκεδαστικότητα σφαλμάτων, δηλαδή τα σφάλματα να έχουν σταθερή παραλλακτικότητα για κάθε τιμή της μεταβλητής x .

Γ) Απουσία πολυσυγγραμικότητας

Δ) Τα σφάλματα να είναι ανεξάρτητα (δηλ. να μην υπάρχει συσχέτιση μεταξύ των σφαλμάτων).

Οι δύο πρώτες (Α, Β) έχουν αναλυθεί στην περιγραφή της απλής γραμμικής παλινδρόμησης. Οι δύο τελευταίες προϋποθέσεις (Γ, Δ) ελέγχονται με χρήση στατιστικών δεικτών. Διάφορα τεστ, όπως το VIF (μέγιστη τιμή >10 ή μέση τιμή σημαντικά μεγαλύτερη από 1), το Tolerance (τιμή >0.1 σημαντικό πρόβλημα, τιμή >0.2 πιθανό πρόβλημα) ελέγχουν την ύπαρξη ή μη πολυσυγγραμικότητας. Επίσης, με χρήση του Durbin-Watson τεστ ελέγχεται η τέταρτη προϋπόθεση. Όταν ο δείκτης αυτός παίρνει τιμές από -2 έως 2, σημαίνει ότι τα σφάλματα είναι ανεξάρτητα (η ανεξαρτησία μπορεί να φανεί και σε σχετικό διάγραμμα, που καλείται "Residuals versus the Order of the Data", όπου η έλλειψη συγκεκριμένου σχήματος δείχνει την ανεξαρτησία). Τα παραπάνω τεστ προκύπτουν εύκολα με τη χρήση στατιστικών πακέτων, όπως περιγράφεται στη συνέχεια.

Με χρήση SPSS v.16

Analyze → Regression → Linear. Στο παράθυρο που ανοίγει, μεταφέρω την εξαρτημένη μεταβλητή στο πλαίσιο με το όνομα *Dependent* και τις ανεξάρτητες μεταβλητές στο πλαίσιο με το όνομα *Independent(s)*. Στη συνέχεια, κάνω κλικ στο Statistics και τσεκάρω τις επιλογές *Collinearity diagnostics* και *Durbin-Watson*.

ΥΠΕΝΘΥΜΙΣΗ: Ο έλεγχος της επέκτασης του μοντέλου πρόβλεψης χρησιμοποιεί τις τιμές των "standardized residuals" για τις Α, Β προϋποθέσεις, τις τιμές των "residuals" για την Δ προϋπόθεση και τις αρχικές τιμές (πρωτογενή δεδομένα) για την Γ προϋπόθεση**.** Δεν κάνω επέκταση (extrapolation) του μοντέλου σε δεδομένα, που απέχουν πολύ από τα αρχικές τιμές (ή είναι εκτός λογικών ορίων).

Εφαρμόζω δύο βήματα για
την αξιολόγηση του μοντέλου

1) Εντοπισμός ακραίων τιμών και αντιμετώπισή τους (με στατιστικά τεστ)

2) Έλεγχος πολυσυγγραμικότητας και σφαλμάτων για κανονική κατανομή, ομοσκεδαστικότητα & ανεξαρτησία (με διαγράμματα και στατιστικά τεστ)

ΣΥΜΠΕΡΑΣΜΑΤΑ

Η ανάλυση παραλλακτικότητας (ANOVA) χρησιμοποιείται για τη στατιστική επεξεργασία των δεδομένων στις περιπτώσεις, που πρέπει: i) να βρεθεί η επίδραση ενός ή περισσότερων πειραματισμών στα δεδομένα (σύγκριση μέσων) και, ii) να βρεθεί η βέλτιστη γραμμή πρόβλεψης (regression line) μίας εξαρτημένης μεταβλητής από μία ή περισσότερες ανεξάρτητες μεταβλητές. Και για τις δύο αυτές εργασίες χρησιμοποιείται το F-test.

Για την εφαρμογή των τεχνικών t-test, ANOVA (one Way ή περισσότερων) και παλινδρόμησης πρέπει να τηρούνται ορισμένες προϋποθέσεις. Σε κάθε στατιστική επεξεργασία δεδομένων πρέπει να αναφέρονται τα **αποτελέσματα ελέγχου των προϋποθέσεων. Αν δεν γίνει ο έλεγχος των προϋποθέσεων ή αν γίνει και δείξει ότι δεν τηρούνται οι προϋποθέσεις αυτές, τότε το αποτέλεσμα της στατιστικής ανάλυσης είναι είτε έωλο είτε ανακριβές, αντίστοιχα.**

ΒΙΒΛΙΟΓΡΑΦΙΑ

Andy Field, 2000. Discovering Statistics Using SPSS for Windows.

Deborah Rumsey, 2007. Intermediate Statistics FOR DUMMIES.

ΦΥΛΛΟ ΕΛΕΓΧΟΥ

Για σύγκριση μέσων (τσεκάρω αντίστοιχα το κάθε τετραγωνάκι)

Έλεγχος κανονικής κατανομής (διαγράμματα και στατιστικά τεστ)	
Έλεγχος ίσης παραλλακτικότητας (οπτικά και στατιστικά τεστ)	
Ποσοτικές/συνεχείς μεταβλητές	
Ανεξάρτητες μεταβλητές	

Αν δεν ισχύουν οι παραπάνω προϋποθέσεις δεν πρέπει να εφαρμόζονται τα παραμετρικά τεστ (t-test, ANOVA).

Για βέλτιστη γραμμή πρόβλεψης (τσεκάρω αντίστοιχα το κάθε τετραγωνάκι)

i) για απλή γραμμική παλινδρόμηση

Έλεγχος ακραίων τιμών (στατιστικά τεστ)	
σε *Standardized Residuals* Έλεγχος κανονικής κατανομής (διαγράμματα και στατιστικά τεστ)	
σε *Standardized Residuals* Έλεγχος ομοσκεδαστικότητας (διαγράμματα)	

ii) για πολλαπλή γραμμική παλινδρόμηση

Έλεγχος ακραίων τιμών (στατιστικά τεστ)	
σε *Standardized Residuals* Έλεγχος κανονικής κατανομής (διαγράμματα και στατιστικά τεστ)	
σε *Standardized Residuals* Έλεγχος ομοσκεδαστικότητας (διαγράμματα)	
σε πρωτογενή δεδομένα Έλεγχος πολυσυγγραμικότητας (πριν τη δημιουργία του μοντέλου με συντελεστή συσχέτισης και μετά με στατιστικά τεστ)	
σε *Residuals* Έλεγχος ανεξαρτησίας (στατιστικά τεστ και διαγράμματα)	

Αν ισχύουν οι παραπάνω προϋποθέσεις αυξάνονται οι πιθανότητες γενίκευσης του προτεινόμενου μοντέλου.

ΕΛΛΗΝΙΚΟΣ ΓΕΩΡΓΙΚΟΣ ΟΡΓΑΝΙΣΜΟΣ – ΔΗΜΗΤΡΑ
ΓΕΝΙΚΗ ΔΙΕΥΘΥΝΣΗ ΑΓΡΟΤΙΚΗΣ ΕΡΕΥΝΑΣ
Κουρτίδου 56-58 & Νιρβάνα, 11145 Αθήνα
www.elgo.gr

ΙΝΣΤΙΤΟΥΤΟ ΕΔΑΦΟΪΔΑΤΙΚΩΝ ΠΟΡΩΝ (ΙΕΥΠ)
Ελληνική Γεωργική Σχολή 57001 Θέρμη, Θεσσαλονίκη
www.swri.gr