

ΑΤΕΙ ΘΕΣΣΑΛΟΝΙΚΗΣ – ΠΑΡΑΡΤΗΜΑ ΚΑΤΕΡΙΝΗΣ
Λύσεις στις εξετάσεις εργαστηρίου Στατιστικής
ΧΕΙΜΕΡΙΝΟ ΕΞΑΜΗΝΟ - ΑΠΡΙΛΙΟΣ 2007

Όνοματεπώνυμο :
Αριθμός Μητρώου :
Εξάμηνο :
Αρχείο δεδομένων επεξεργασίας :

Σας δίνεται αρχείο το οποίο περιέχει τις βαθμολογίες κάποιων φοιτητών στο μάθημα της Στατιστικής (Θεωρία και Εργαστήριο) καταχωρημένες με τον εξής τρόπο :

gender : φύλο φοιτητή (Κορίτσι 0 ή Αγόρι 1)
origin : προέλευση φοιτητή (Πόλη 0 ή Χωριό 1)
deg_th : Βαθμός στη Θεωρία Στατιστικής
deg_lab : Βαθμός στο εργαστήριο Στατιστικής.

Θέμα 1^ο : Ποια είναι η μέση επίδοση, η τυπική απόκλιση και η διάμεσος της βαθμολογίας στη θεωρία Στατιστικής για το σύνολο των φοιτητών;

Απάντηση

Αρκεί να εκτελεστεί η διαδικασία

Analyze -> Descriptive Statistics -> Explore και να τοποθετηθεί ως **Dependent List** η μεταβλητή **deg_th**. Μετά αρκεί να θυμηθεί ο εξεταζόμενος ότι

Μέση Επίδοση = Mean

Τυπική απόκλιση = Std Variation

Διάμεσος = Median

Ποια είναι η μέση επίδοση για τους φοιτητές οι οποίοι κατάγονται από χωριό και ποια η τυπική απόκλιση και η διάμεσος των βαθμολογιών τους;

Απάντηση

Επιλέγουμε τους φοιτητές από χωριό με την διαδικασία

Data -> Select cases : If condition is satisfied και γράφουμε την συνθήκη **origin = 1**. Η διαδικασία **Explore** όπως προηγουμένως δίνει τα απαραίτητα στατιστικά.

Θέμα 2^ο : Ποιος ο συντελεστής γραμμικής συσχέτισης μεταξύ του βαθμού θεωρίας και του βαθμού εργαστηρίου;

Απάντηση

Analyze -> Correlate -> Bivariate και τοποθετώ στο παράθυρο **Variables** τις μεταβλητές **deg_th** και **deg_lab**. Μετά το **OK** διαβάζουμε τον κατάλληλο αριθμό στον **2x2 πίνακα** που παρουσιάζεται και θυμόμαστε ότι το SPSS δεν εμφανίζει το αρχικό μηδέν σε δεκαδικούς αριθμούς δηλαδή ο αριθμός **0,756** στο SPSS εμφανίζεται ως **,756**.

Είναι αρκετά ισχυρή η γραμμική συσχέτιση των δύο μεταβλητών;

Απάντηση

Κατά γενική παραδοχή, δύο μεταβλητές εμφανίζουν ισχυρή γραμμική συσχέτιση όταν ο συντελεστής γραμμικής συσχέτισης είναι μεγαλύτερος από **0,7**.

Ποια είναι η εξίσωση της ευθείας γραμμικής παλινδρόμησης; (Θεωρείστε τον βαθμό θεωρίας ως ανεξάρτητη και τον βαθμό εργαστηρίου ως εξαρτημένη μεταβλητή).

Απάντηση

Analyze -> Regression -> Linear, και τοποθετώ ως **Dependent** την μεταβλητή **deg_lab** ενώ ως **Independent(s)** την **deg_th**.

Στον πίνακα "**Coefficients**" που εμφανίζεται στο Output και δίπλα στην ονομασία Βαθμός θεωρίας διαβάζω τον συντελεστή της μεταβλητής **deg_th** ενώ ακριβώς από πάνω και δίπλα στην λέξη (Constant) διαβάζω την σταθερά της εξίσωσης της ευθείας. Τα στοιχεία αυτά τα συμπληρώνω στα κατάλληλα κενά.

Βαθμός εργαστηρίου = _____ x βαθμός θεωρίας + _____

Αν κάποιος φοιτητής ο οποίος δεν ανήκει στο δείγμα μας γράψει βαθμό 5 στη θεωρία Στατιστικής τότε ποιος είναι ο βαθμός που ο ίδιος φοιτητής περιμένουμε να γράψει στο εργαστήριο Στατιστικής;

Απάντηση

Απλά τοποθετώ όπου Βαθμός θεωρίας = 5 κάνω τον πολλαπλασιασμό και την πρόσθεση και δίνω την απάντηση.

Θέμα 3^ο : Μπορούμε να ισχυριστούμε ότι τα αγόρια και τα κορίτσια έχουν ίδια μέση επίδοση στο εργαστήριο Στατιστικής;

Απάντηση

Ελέγχουμε την υπόθεση ότι τα αγόρια και τα κορίτσια έχουν ίδια μέση επίδοση στο εργαστήριο Στατιστικής. Τα αγόρια και τα κορίτσια αποτελούν δύο ανεξάρτητες ομάδες συνεπώς και οι βαθμολογίες τους αποτελούν δύο ανεξάρτητα σύνολα δεδομένων. Ο έλεγχος που πρέπει να διεξάγουμε είναι ο **Analyze->Compare Means -> Independent Samples T-test**.

Στο παράθυρο που εμφανίζεται, θέτω **Test Variable(s)** την μεταβλητή **deg_lab** και **Grouping Variable** την μεταβλητή **gender** και μετά το **Define Groups** όπου θέτω τις τιμές **0** και **1** (**gender=0** : κορίτσια και **gender=1** : αγόρια) δίνω **Continue OK**. Πρώτα διαβάζω τον αριθμό στην στήλη (**Sig.**) στην μεγάλη στήλη **Levene's Test for Equality of Variances** όπου διεξάγεται έλεγχος υπόθεσης σχετικά με την ισότητα των διασπορών των δύο ανεξάρτητων πληθυσμών. (Αρχική υπόθεση Ίσες Διασπορές έναντι εναλλακτικής Άνισες Διασπορές) Αν ο αριθμός στην στήλη **Sig.** είναι μικρότερος από **0,05** τότε απορρίπτω την ισότητα των δύο διασπορών ενώ αν είναι μεγαλύτερος ή ίσος τη δέχομαι. Μετά επιλέγω την αντίστοιχη γραμμή “Equal variances assumed” ή “Equal variances not assumed” και διαβάζω τον αριθμό στην στήλη **Sig. (2-tailed)** Αν ο αριθμός στην στήλη **Sig. (2-tailed)** είναι μικρότερος από **0,05** τότε απορρίπτω την ισότητα των δύο μέσων επιδόσεων ενώ αν είναι μεγαλύτερος ή ίσος τη δέχομαι.

Μπορούμε να ισχυριστούμε ότι οι φοιτητές (ανεξαρτήτου φύλου) τα καταφέρνουν εξίσου καλά στη θεωρία όσο και στο εργαστήριο Στατιστικής;

Απάντηση

Ελέγχουμε την υπόθεση ότι όλοι οι φοιτητές έχουν ίδια μέση επίδοση στη θεωρία και στο εργαστήριο Στατιστικής. Οι βαθμολογίες των φοιτητών στα δύο μαθήματα (θεωρία και εργαστήριο) αποτελούν δύο εξαρτημένα σύνολα δεδομένων καθώς προέρχονται από τα ίδια υποκείμενα (ζευγαρωτές παρατηρήσεις). Ο έλεγχος που πρέπει να διεξάγουμε είναι ο

Analyze -> Compare Means -> Paired Samples T-test.

Στο παράθυρο που εμφανίζεται, θέτω **Paired Variables** το ζευγάρι των μεταβλητών **deg_lab** — **deg_th** (η σειρά των μεταβλητών δεν έχει σημασία) και δίνω **OK**. Μετά διαβάζω τον αριθμό στην στήλη **Sig. (2-tailed)** Αν ο αριθμός στην στήλη **Sig. (2-tailed)** είναι μικρότερος από **0,05** τότε απορρίπτω την ισότητα των δύο μέσων επιδόσεων ενώ αν είναι μεγαλύτερος ή ίσος τη δέχομαι.

Θέμα 4^ο : Δημιουργήστε μια νέα μεταβλητή η οποία θα παίρνει τιμή 0 αν ο φοιτητής έγραψε κάτω από τη βάση στη θεωρία Στατιστικής (δηλαδή $deg_th < 5$) και 1 αν έγραψε πάνω από τη βάση ($deg_th \geq 5$). Μπορούμε να ισχυριστούμε ότι η επιτυχία στη θεωρία Στατιστικής εξαρτάται (α) από το φύλο; (β) από την καταγωγή του φοιτητή;

Απάντηση

Εκτελώ την **Transform -> Compute**. Θέτω ως **Target Variable** ένα αποδεκτό από το SPSS όνομα (π.χ. pass) και τοποθετώ τον αριθμό **0** στο παράθυρο **Numeric Expression**:

Μετά πατώ το **If...** και θέτω ως συνθήκη **$deg_th < 5$** αφού πρώτα τοποθετήσω την τελεία στο **Include if case satisfies condition**:

Μετά το **Continue** και το **OK** έχει δημιουργηθεί η μεταβλητή και έχει τοποθετηθεί το 0 στις εγγραφές όπου βρέθηκε βαθμός θεωρίας κάτω από 5. Επαναλαμβάνουμε τη διαδικασία και τοποθετούμε με τον ίδιο τρόπο **pass = 1** στις εγγραφές όπου **$deg_th \geq 5$** . Τη δεύτερη φορά καθώς το SPSS θα αναγνωρίσει την προσπάθεια αλλαγής τιμών μιας υφιστάμενης μεταβλητής θα ρωτήσει ανοίγοντας ένα μικρό παράθυρο **Change Existing Variable?**, στο οποίο δίνουμε **OK**. Τώρα είμαστε σε θέση να κάνουμε τους ελέγχους υποθέσεων. Μελετούμε την πιθανή συσχέτιση δύο ποιοτικών μεταβλητών (επιτυχία, καταγωγή) και (επιτυχία, φύλο) άρα συμπεραίνουμε ότι ο κατάλληλος έλεγχος είναι ο **X^2** . Η διαδικασία είναι η

Analyze -> Descriptive Statistics -> Crosstabs.

Ο έλεγχος **X^2** ελέγχει την αποδοχή της πρότασης

H_0 : Οι δύο μεταβλητές είναι ανεξάρτητες έναντι της

H_1 : όχι η H_0 .

Θέτουμε ως **Row(s)**: τη μεταβλητή **pass** και ως **Column(s)**: τις μεταβλητές **origin** και **gender**. Επιλέγουμε **Statistics** και από το νέο παράθυρο που εμφανίζουμε επιλέγουμε (\checkmark) την επιλογή **Chi-Square** πάνω αριστερά. Μετά δίνουμε **Continue** και **OK**. Στον πίνακα "**Chi-Square Tests**" που εμφανίζεται, στην γραμμή "**Pearson Chi-Square**" διαβάζουμε τον αριθμό στην στήλη **Asymp. Sig. (2-sided)**. Αν αυτός ο αριθμός είναι μικρότερος από 0,1 τότε απορρίπτουμε την υπόθεση H_0 δηλαδή τότε δεχόμαστε ότι υπάρχει σχέση μεταξύ των δύο ποιοτικών μεταβλητών.