

10^ο Φροντιστήριο – Στατιστική Ι



Κεφ. 10 Ο συντελεστής συσχέτισης

Άσκηση 1

Ένας ερευνητής ενδιαφέρεται να μελετήσει την επίδραση του φύλου στη νοητική ικανότητα. Από τα δεδομένα που συγκέντρωσε προσπάθησε να δημιουργήσει ένα διάγραμμα διασποράς. Δεν τα κατάφερε όμως. Γιατί;

Απάντηση:

Για να σχεδιαστεί ένα διάγραμμα διασποράς θα πρέπει και οι δύο μεταβλητές να είναι ποσοτικές και συνεχείς. Στην περίπτωση μας, η μια μεταβλητή (το φύλο) είναι κατηγορική. Συνεπώς δεν μπορεί να γίνει διάγραμμα διασποράς.

Άσκηση 2

Ποιος συντελεστής συσχέτισης είναι πιο υψηλός, το $-0,56$ ή το $0,56$; Γιατί;

Απάντηση:

Είναι ίδιοι!

Το πρόσημο δεν δηλώνει ποιος είναι μεγαλύτερος ή μικρότερος, παρά μόνο την κατεύθυνση (θετική ή αρνητική συσχέτιση)

Άσκηση 3

Ποια είναι η σχέση μεταξύ συνδιακύμανσης και συντελεστή συσχέτισης;

Απάντηση: Ο συντελεστής συσχέτισης είναι το στατιστικό κριτήριο που χρησιμοποιούμε για να διαπιστώσουμε το μέγεθος της συμμετοβολής δύο μεταβλητών και βασίζεται στη συνδιακύμανση. Η συνδιακύμανση είναι στην ουσία ένας αριθμός που δείχνει το βαθμό στον οποίο οι δύο μεταβλητές μεταβάλλονται επηρεάζοντας η μία την άλλη.

Άσκηση 4

Μια ερευνήτρια θέλησε να μελετήσει τη σχέση ανάμεσα στην προσωπικότητα και τη συναισθηματική νοημοσύνη. Για τον σκοπό αυτό χρησιμοποίησε τον συντελεστή συσχέτισης. Από τα αποτελέσματα διαπίστωσε ότι υπήρχε αρνητική συσχέτιση ($r=-0,78$). Έτσι, συμπέρανε ότι υπήρχε χαμηλή συσχέτιση των δύο μεταβλητών. Είναι σωστό το συμπέρασμα που έβγαλε;

Απάντηση:

Δεν είναι σωστό το συμπέρασμα!

Το αρνητικό πρόσημο δεν σημαίνει χαμηλή συσχέτιση.

Μάλιστα, στη συγκεκριμένη περίπτωση, η τιμή $r=-0,78$ σημαίνει υψηλή αρνητική συσχέτιση.

Άσκηση 5

Ένας ερευνητής πραγματοποιεί μία έρευνα στο πλαίσιο της οποίας μελετά κατά πόσο υπάρχει συσχέτιση της στάσης απέναντι στην ισότητα των δύο φύλων και του συντηρητισμού. Από την ερμηνεία του συντελεστή συσχέτισης συμπέρανε ότι ο συντηρητισμός επηρεάζει τη στάση απέναντι στην ισότητα των φύλων. Είναι σωστό το συμπέρασμα που έβγαλε;

Απάντηση:

Το συμπέρασμα δεν είναι σωστό!

Ο συντελεστής συσχέτισης δεν μπορεί να χρησιμοποιείται ως απόδειξη αιτιώδους σχέσης, ανάμεσα σε δύο μεταβλητές, δηλαδή εδώ ότι ο συντηρητισμός επηρεάζει τη στάση απέναντι στην ισότητα των δύο φύλων.

Άσκηση 6

Σε μία πρόσφατη έρευνα διαπιστώθηκε ότι ο συντελεστής συσχέτισης της κούρασης με την ευερεθιστότητα είναι $r=0,43$. Με βάση αυτό το αποτέλεσμα, η ερευνήτρια κατέληξε ότι η κούραση αποτελεί έναν σημαντικό παράγοντα για την πρόκληση της ευερεθιστότητας. Μπορεί αυτό το συμπέρασμα να τεκμηριωθεί από το στατιστικό αποτέλεσμα που προέκυψε;

Απάντηση:

Η λάθος λέξη είναι το “σημαντικός”. Η τιμή 0,43 δείχνει ότι υπάρχει μία συσχέτιση αλλά όχι πολύ μεγάλη. Χρησιμοποιώντας τον δείκτη προσδιορισμού (r^2), βλέπουμε ότι το ποσοστό συνολικής διακύμανσης της ευερεθιστικότητας που οφείλεται στην κούραση είναι μόλις 18,5% ($0,43^2 \times 100 = 18,49\%$)

Θεωρία: Ο r^2 μας δίνει το ποσοστό της διακύμανσης που είναι κοινό στις δύο μεταβλητές ($r^2 : r^2 \times 100$)

Άσκηση 7

Ένας ερευνητής θέλησε να μελετήσει τη σχέση ανάμεσα στη σχολική επίδοση (βαθμοί στο σχολείο) και στον ελεύθερο χρόνο του παιδιού (ώρες κατά τις οποίες δεν διαβάζει). Για τον σκοπό αυτό επέλεξε ένα δείγμα μαθητών της Δ' Δημοτικού από τα τρία σχολεία μιας επαρχιακής πόλης.

Ενώ, λοιπόν, η βιβλιογραφία δείχνει ότι υπάρχει υψηλή αρνητική συσχέτιση των δύο μεταβλητών, η έρευνά του έδειξε αρνητική συσχέτιση αλλά το μέγεθός της ήταν μόλις $r=-0,32$. Πού μπορεί να οφείλεται αυτή η διαφορά ανάμεσα στα αποτελέσματα του ερευνητή και στα αποτελέσματα των άλλων ερευνών;

Απάντηση:

Ο πιο πιθανός λόγος που συνέβη αυτό είναι το γεγονός ότι το δείγμα που χρησιμοποίησε ήταν πολύ ομοιογενές, μια κατάσταση που επηρεάζει το μέγεθος της διακύμανσης. Μάλιστα, όσο μεγαλύτερη είναι η ομοιογένεια τόσο μικρότερος βγαίνει ο συντελεστής συσχέτισης.

Άσκηση 8

Ένας ερευνητής ενδιαφέρεται να μελετήσει αν υπάρχει συσχέτιση της συναισθηματικής νοημοσύνης με την ευζωία. Γι' αυτό τον σκοπό επέλεξε ένα δείγμα 250 ατόμων και τους ζήτησε να συμπληρώσουν δύο ερωτηματολόγια: Ένα για τη συναισθηματική νοημοσύνη και ένα για την ευζωία.

Άσκηση 8

Μετά την αξιολόγηση των ερωτηματολογίων κάθε άτομο είχε μια συνολική βαθμολογία στις δύο αυτές έννοιες που προέκυπτε από το άθροισμα των ερωτήσεων στα αντίστοιχα ερωτηματολόγια. Ποιον συντελεστή συσχέτισης θα χρησιμοποιήσει ο ερευνητής για να δώσει απάντηση στο συγκεκριμένο ερώτημα που τον απασχολεί;

Απάντηση:

Θα χρησιμοποιήσει τον συντελεστή συσχέτισης Pearson r , αφού οι δύο μεταβλητές που μελετά έχουν μετρηθεί σε κλίμακα ίσων διαστημάτων.

Άσκηση 9

Μετά από τη χρήση του κατάλληλου συντελεστή συσχέτισης το αποτέλεσμα ήταν $r=0,78$. Τι συμπέρασμα μπορεί να βγάλει κάποιος από αυτό το αποτέλεσμα αναφορικά με το παραπάνω ερευνητικό πρόβλημα.

Απάντηση:

Μπορούμε να πούμε ότι όσο αυξάνεται η συναισθηματική νοημοσύνη αυξάνεται και η ευζωία. Με άλλα λόγια, τα άτομα που έχουν υψηλή συναισθηματική νοημοσύνη βιώνουν και μεγαλύτερη ικανοποίηση από τη ζωή.

Άσκηση 10

Πόσο ποσοστό της κοινής διακύμανσης των δύο μεταβλητών ερμηνεύει ο συντελεστής συσχέτισης που βρήκε ο ερευνητής στο προηγούμενο παράδειγμα;

Απάντηση:

Θα χρησιμοποιήσουμε τον δείκτη προσδιορισμού (r^2):

$$(-0,78)^2 \times 100 = 61\%$$

Ερμηνεύει λοιπόν το 61%

Άσκηση 11

Ένας ερευνητής θέλησε να διαπιστώσει κατά πόσο υπάρχει συσχέτιση του άγχους με τη σχολική επίδοση. Έδωσε, λοιπόν, ένα ερωτηματολόγιο άγχους σε 250 μαθητές και, ανάλογα με την επίδοσή τους στο τεστ, τους κατέταξε ιεραρχικά ξεκινώντας από το άτομο με το υψηλότερο άγχος. Το ίδιο έκανε και με τους βαθμούς των μαθητών στο τέλος της σχολικής χρονιάς. Ποιον συντελεστή συσχέτισης θα χρησιμοποιήσει ο ερευνητής για να δώσει απάντηση;

Απάντηση:

Θα χρησιμοποιήσει τον συντελεστή συσχέτισης Spearman ρ , αφού και οι δύο μεταβλητές έχουν μετρηθεί σε ιεραρχική κλίμακα.

Άσκηση 13

Κατά την προκαταρκτική διερεύνηση των δεδομένων που συγκεντρώθηκαν σε μία έρευνα διαπιστώθηκε ότι η μορφή της κατανομής τους είναι καμπυλόγραμμη. Τι σημαίνει αυτό για τη συνέχεια της στατιστικής ανάλυσης;

Απάντηση:

Αυτό σημαίνει ότι δεν είναι δυνατή η χρήση των βασικών συντελεστών συσχέτισης (Pearson r και Spearman ρ) καθώς η εφαρμογή του προϋποθέτει τη γραμμική σχέση ανάμεσα στις δύο μεταβλητές που εξετάζονται.

Άσκηση 15

Μία ερευνήτρια θέλει να διαπιστώσει κατά πόσο δύο από τις μεταβλητές που μελετάει σε μία έρευνά της (ύπαρξη παιδιών ή όχι και το εάν η μητέρα εργάζεται ή όχι) συσχετίζονται. Ποιον συντελεστή συσχέτισης θα χρησιμοποιήσει για να δώσει απάντηση στο συγκεκριμένο ερώτημα;

Απάντηση:

Σε αυτή την περίπτωση θα χρησιμοποιήσει τον συντελεστή συσχέτισης ϕ , αφού θέλει να συσχετίσει δύο ποιοτικές μεταβλητές με δύο επίπεδα η κάθε μία:

Υπάρχουν παιδιά-Δεν υπάρχουν παιδιά και Εργαζόμενη-Μη εργαζόμενη.

ΤΕΛΟΣ