**ΚΕΦΑΛΑΙΟ 5 – ΣΤΑΤΙΣΤΙΚΕΣ ΥΠΟΘΕΣΕΙΣ**

**5.1 Γενικά**

Όταν λέμε **Στατιστική** **Υπόθεση**, εννοούμε μια υπόθεση που κάνουμε για τον άγνωστο πληθυσμό που μελετάμε. Π.χ. σε μια έρευνα που γίνεται για την απόδοση ενός φαρμάκου Α, μπορούμε να υποθέσουμε οτι το φάρμακο αυτό καταπολεμά κάποια αρρώστια σε ποσοστό 80% των ασθενών ή οτι είναι καλύτερο απο κάποιο φάρμακο Β, που πιθανόν να κυκλοφορεί στο εμπόριο. Σε μια έρευνα για το ύψος των ανθρώπων μιας συγκεκριμένης φυλής, μπορεί κάποιος να ισχυριστεί οτι το μέσο ύψος τους υπερβαίνει το 1,80 m ή οτι η φυλή Α έχει πιο ψηλά άτομα απο τις φυλές Β, Γ ή Δ. Αυτές οι ερευνητικές υποθέσεις, που κάνει ένας ερευνητής χρειάζονται επιβεβαίωση ή απόρριψη. Η διαδικασία την οποία ακολουθούμε για να επαληθεύσουμε τέτοιους ισχυρισμούς, ονομάζεται **Έλεγχος Υπόθεσης** (ή Στατιστικό τέστ).

Οι Έλεγχοι Υποθέσεων αποτελούν το σημαντικότερο κομμάτι της Στατιστικής αλλά και της έρευνας γενικότερα. Ουσιαστικά οι περισσότερες θεωρίες σήμερα στηρίζονται σε αυτούς. Είναι γεγονός όμως, οτι πολλές φορές γίνεται μια κακή χρήση των στατιστικών τέστ, σε σημείο να ισχυρίζονται κάποιοι οτι τελικά με την Στατιστική μπορεί κανείς να αποδείξει τα πάντα.

Κρίνουμε λοιπόν σκόπιμο να ξεκαθαρίσουμε μερικά πράγματα, για να αποφευχθούν τυχόν παρερμηνείες ή παρεξηγήσεις απο τον αναγνώστη. Και πρώτα απ' όλα, τι σημαίνει "απόδειξη", στην Στατιστική. Στα Μαθηματικά, όταν λέμε να αποδείξουμε μια πρόταση ή ένα θεώρημα, γνωρίζουμε οτι θα ακολουθήσουμε μια επαγωγική διαδικασία, στηριζόμενοι σε κάποια "αξιώματα". Συνεπώς η επαλήθευση ή όχι μιας πρότασης είναι μια διαδικασία καθολικά αληθής, αφού στηρίζεται σε μια αξιωματικά θεμελιωμένη θεωρία. Έτσι πειραματιζόμενοι με κάποια τρίγωνα, ισχυριζόμαστε οτι το άθροισμα των γωνιών του τριγώνου είναι 180 μοίρες. Αυτό μπορεί να διατυπωθεί και να αποδειχθεί με αυστηρά βήματα αφού θέσει κανείς κάποια αξιώματα. Πειράματα όμως με προβλέψιμα αποτελέσματα, όπως το παραπάνω, είναι μάλλον σπάνια, σε σχέση με τα μη προβλέψιμα. Οση γνώση μαθηματικών ή άλλων επιστημών και να έχει κάποιος δεν μπορεί να προβλέψει το φύλο ενός παιδιού πριν να γεννηθεί (ή πριν κάνει υπέρηχο!) ή το ύψος ενός ανθρώπου με ακρίβεια, πριν να τον μετρήσει. Μπορεί βέβαια να ισχυριστεί ότι στα δύο παιδιά που γεννιούνται το ένα είναι κορίτσι, όμως σίγουρα δεν μπορεί να "αποδείξει" τον ισχυρισμό του με έναν αυστηρό (μαθηματικό) τρόπο. Εκείνο που μπορεί να κάνει είναι, έχοντας π.χ. ένα δείγμα απο 500 νεογέννητα παιδιά, να μετρήσει τα κορίτσια και να επιβεβαιώσει ή να απορρίψει τον ισχυρισμό του. Πιο σωστά, μπορεί να ελέγξει αν τα δεδομένα του παρέχουν την ένδειξη οτι ο ισχυρισμός του δεν είναι σωστός ή αν τα δεδομένα του, δεν καταφέρνουν να τον απορρίψουν. Απόδειξη λοιπόν στην Στατιστική είναι απλώς μια ισχυρή ένδειξη.

Στην παραπάνω φράση κρύβεται όλη η ουσία για τους στατιστικούς ελέγχους. Πράγματι, αν ο ερευνητής στα 500 παιδιά έβρισκε 250 κορίτσια και 250 αγόρια, δεν θα είχε κανένα λόγο να αμφιβάλει για την αναλογία 50-50 στα νεογέννητα. Αν πάλι έβρισκε 240 - 260, πάλι το πιθανότερο είναι ότι δεν έχει πολλές ενδείξεις οτι κάτι δεν πάει καλά με την αναλογία 50% - 50% και θα υπέθετε οτι παίρνοντας ένα άλλο δείγμα (1000 ή 2000 νεογέννητα) θα μπορούσε να βρει μια πιο ισορροπημένη αναλογία. Τι γίνεται όμως στην περίπτωση που ανακαλύπτει οτι τα 400 απο τα 500 νεογέννητα είναι κορίτσια; Τότε μπορεί να πει οτι διαθέτει μια ισχυρή ένδειξη οτι η αναλογία πλεον δεν είναι 50% - 50%, αλλά πολύ διαφορετική, έτσι έχει αποδείξει κάτι νέο και με τον τρόπο αυτό συνεισφέρει στην έρευνα με μια σημαντική ανακάλυψη.

Εκείνο που ενδιαφέρει στην περίπτωση ενός νέου ερευνητικού αποτελέσματος, (π.χ. ότι η θεραπεία Α είναι καλύτερη απο τη θεραπεία Β), είναι ο ισχυρισμός μας να είναι πιθανότατα σωστός. Δηλαδή να μην υπάρχει μεγάλη πιθανότητα να βγάλουμε λάθος συμπεράσματα. Στην Στατιστική επιδιώκουμε η πιθανότητα ενός τέτοιου λάθους συμπεράσματος, να μην είναι μεγαλύτερη από 5% (ή 0,05) και αυτό ονομάζεται **Σημαντικότητα του Ελέγχου** (Significance)

**5.2 Περίπτωση A – Διαφορές μεταξύ πληθυσμών, αναφορικά με ένα Ποσοτικό Χαρακτηριστικό (t-test & ANOVA)**

Στην πρώτη περίπτωση, υποθέτουμε πως μελετήσαμε ένα ποσοτικό χαρακτηριστικό (συνήθως συνεχές, αλλά αυτό δεν είναι απαραίτητο). Στα χαρακτηριστικά αυτά αναφερόμαστε συνήθως στη μέση τιμή τους και θεωρούμε ότι η μέση τιμή είναι ένα αντιπροσωπευτικό μέγεθος για το χαρακτηριστικό αυτό. Συνήθως η μέση τιμή βέβαια είναι αντιπροσωπευτική σε συμμετρικές κατανομές (κανονική), ενώ σε περιπτώσεις ασυμμετρίας χρησιμοποιείται η διάμεσος. Για το λόγο αυτό οι διαδικασίες που θα παρουσιαστούν περιληπτικά στην παράγραφο αυτή, κάνουν λόγο για τη μέση τιμή. Αντίστοιχες διαδικασίες που αφορούν μη κανονικές κατανομές, συνεπώς κάνουν λόγο για τη διάμεσο, υπάρχουν, και ονομάζονται Μη Παραμετρικές Δοκιμασίες (non parametric tests). Για τις δοκιμασίες αυτές παραπέμπουμε τον αναγνώστη σε άλλα βοηθήματα.

**Θέτουμε λοιπόν στη συνέχεια το γενικό ερώτημα:**

*«Υπάρχουν διαφορές ανάμεσα στις μέσες τιμές δύο η περισσότερων υποομάδων πληθυσμών?»*

Οι υποομάδες μπορεί να αφορούν π.χ το φύλο (άνδρες/γυναίκες), το κάπνισμα (καπνιστές/μη καπνιστές) την ηλικία μετρημένη με διατάξιμη μεταβλητή (ηλικία άνω των 40 / ηλικία κάτω των 40)

Οι υποομάδες μπορεί να είναι μόνο δύο όπως τα παραπάνω παραδείγματα ή και περισσότερες, π.χ. σύγκριση ανα τόπο κατοικίας (αγροτικές/αστικές/ημιαστικές περιοχές), μορφωτικό επίπεδο (αποφοίτους Δημοτικού/ Γυμνασίου / Λυκείου/ ΑΕΙ) ή την ηλικία μετρημένη με διατάξιμη μεταβλητή (ηλικία κάτω από 30 / ηλικία μεταξύ 31 και 50 / ηλικία από 51 μέχρι 70 / ηλικία από 71 και πάνω)

**Παράδειγμα 1**

Ένας ερευνητής ισχυρίστηκε ότι οι άνδρες είναι πιο ψηλοί από τις γυναίκες. Θέλησε λοιπόν να βρει αν υπάρχει διαφορά στο μέσο ύψος των ανδρών και των γυναικών, για να αποδείξει τον ισχυρισμό του. Ξεκίνησε με δύο δείγματα (ένα με άνδρες και ένα με γυναίκες) και μέτρησε τα ύψη τους. Ας πούμε πως μέτρησε 150 άνδρες και 150 γυναίκες και υποθέτουμε πως βρήκε τις παρακάτω μέσες τιμές στα δείγματα του:

Μέση τιμή ανδρών (στο δείγμα) = 179 cm

Μέση τιμή γυναικών (στο δείγμα) = 175 cm

Ο ερευνητής λοιπόν παρατήρησε μια ΔΙΑΦΟΡΑ 4 εκατοστών στις μέσες τιμές στα δύο δείγματα! Άρα τα μέσα ύψη ανδρών και γυναικών σίγουρα δεν είναι ίδια (ΣΤΑ ΔΕΙΓΜΑΤΑ)

**Το Ερώτημα:**

Η Γενική ερώτηση που θέτουμε στη συνέχεια είναι:

*«Αυτές οι διάφορες που βρήκαμε ανάμεσα στις ομάδες (στην προκειμένη περίπτωση μεταξύ ανδρών και γυναικών), στα δείγματα ήταν τυχαίες ή μήπως ΥΠΑΡΧΟΥΝ πράγματι διαφορές και στους πληθυσμούς;»*

Αν μπορούσαμε δηλαδή να μετρήσουμε ολόκληρους τους πληθυσμούς, από τους οποίους προέρχονται τα δείγματα, θα έβγαινε πάλι διαφορά ή δεν θα έβγαινε; Αν δεν υπάρχει διαφορά στους πληθυσμούς, τότε ότι διαφορά παρατηρήσαμε στα δείγματα, είναι ΤΥΧΑΙΑ.

Στην Στατιστική χρησιμοποιούμε κάποιες διαδικασίες που επιτρέπουν να απαντήσουμε με ένα ΝΑΙ ή ένα ΟΧΙ στην παραπάνω βασική ερώτηση. Οι διαδικασίες αυτές λέγονται **Στατιστικές Δοκιμασίες (Statistical Tests)** και ανάλογα με το πρόβλημα υπάρχει και η κατάλληλη μέθοδος.

**Όταν θέλουμε να συγκρίνουμε τις μέσες τιμές ανάμεσα σε ΔΥΟ υποομάδες χρησιμοποιούμε τον Έλεγχο t (t-test), ενώ για να συγκρίνουμε ΠΕΡΙΣΣΟΤΕΡΕΣ ΑΠΟ ΔΥΟ υποομάδες, χρησιμοποιούμε Ανάλυση Διακύμανσης (Analysis of Variance - ANOVA)**

**H Τιμή P και η Απόφαση**

Πως καταλήγουμε τώρα στην απόφαση μας; Στις σημειώσεις αυτές δεν θα δώσουμε λεπτομέρειες για τις διαδικασίες που αναφέραμε παραπάνω, που κάποιες είναι υπολογιστικά εύκολες, άλλες είναι πιο δύσκολες. Θεωρούμε ότι ο αναγνώστης μπορεί να χρησιμοποιήσει εξειδικευμένα Στατιστικά Προγράμματα (π.χ. EXCEL, SPSS, …) ώστε να γίνουν οι υπολογισμοί, και ο σκοπός μας είναι απλά να μπορεί να κατανοήσει τα αποτελέσματα που δίνουν τα προγράμματα αυτά.

Οι Υπολογιστές μας δίνουν ως αποτέλεσμα, ένα «μαγικό» αριθμό στην κλίμακα από το 0 μέχρι το 1, που μας δίνει την ΠΙΘΑΝΟΤΗΤΑ να είναι οι διαφορές μεταξύ των υποομάδων τυχαίες.

Τον αριθμό αυτό τον ονομάζουμε **Τιμή P (P-value)**

**Σημείωση:** στο SPSS αναφέρεται ως **SIG**.

Αν ο αριθμός αυτός είναι πολύ μικρός, τότε οι διαφορές που παρατηρήθηκαν στα δείγματα δεν μπορεί να είναι τυχαίες, άρα υπάρχουν διαφορές και στους αντίστοιχους πληθυσμούς. Αν ο αριθμός είναι μεγάλος, τότε δεν μπορούμε να συμπεράνουμε πως υπάρχουν διαφορές στους πληθυσμούς, αφού υπάρχει πιθανότητα οι διαφορές στα δείγματα να ήταν τυχαίες. Το όριο ανάμεσα στο «μεγάλο» και στο «μικρό» στην παραπάνω πρόταση είναι το 5% ή το 0,05. Ανακεφαλαιώνοντας

* **Αν το P<0,05 (πιο μικρό από το 5%) τότε οι διαφορές ΔΕΝ είναι τυχαίες, δηλαδή οι διαφορές που βρήκαμε στα δείγματα, επιβεβαιώνονται πιθανότατα και στους πληθυσμούς**
* **Αν το P>0,05 (πιο μεγάλο από το 5%) τότε οι διαφορές μπορεί να είναι τυχαίες, άρα οι διαφορές που βρήκαμε στα δείγματα, ΔΕΝ επιβεβαιώνονται στους πληθυσμούς**

Η πιθανότητα να έχουμε κάνει λάθος στον ισχυρισμό μας είναι ίση με P. Για το λόγο αυτό, όσο πιο μικρό είναι το P, τόσο πιο σίγουροι είμαστε για τις διαφορές που υπάρχουν και πιο έγκυρο είναι το αποτέλεσμα της έρευνας μας.

Αν επανέλθουμε στο παράδειγμα 1, η απάντηση στο ερώτημα αν υπάρχει διαφορά ανάμεσα στα μέσα ύψη ανδρών και γυναικών, δίνεται εφαρμόζοντας τον έλεγχο t, και το αποτέλεσμα (από το EXCEL πχ.) είναι p=0,000012. Προφανώς το p είναι πολύ μικρότερο από το 0,05 άρα οι διαφορά στα ύψη, που παρατηρήθηκε στα δείγματα ΔΕΝ μπορεί να είναι τυχαία και επιβεβαιώνεται και στους πληθυσμούς. Τελικά δηλαδή οι άνδρες είναι πράγματι πιο ψηλοί από τις γυναίκες (κατά μέσο όρο βέβαια!)

**5.3 Περίπτωση B – Συσχετίσεις Ποσοτικών Χαρακτηριστικών**

Στην δεύτερη περίπτωση, υποθέτουμε πως μελετήσαμε ταυτόχρονα δύο ή και περισσότερα ποσοτικά χαρακτηριστικά σε ένα δείγμα, ας πούμε το Χ και το Υ. Θέλουμε να δούμε αν υπάρχει κάποιου είδους σχέση μεταξύ τους.

Σε όλες τις περιπτώσεις που ζητάμε συσχέτιση χαρακτηριστικών, ενδιαφερόμαστε να δώσουμε απάντηση στα ερωτήματα:

1. Υπάρχει σχέση;
2. Τι σχέση είναι αυτή; (π.χ. όταν αυξάνει το ένα, αυξάνει και το άλλο ή αντίστροφα;)
3. Ποιος είναι ο βαθμός αυτής της σχέσης;
4. Ποια είναι η μορφή αυτής της σχέσης;

**Παραδείγματα:**

α) Τι σχέση υπάρχει ανάμεσα στο ύψος και το βάρος των ανθρώπων; όταν αυξάνει το ένα, αυξάνει και το άλλο ή αντίστροφα;

β) Υπάρχει κάποια σχέση ανάμεσα στην βαθμολογία στα Μαθηματικά και στη βαθμολογία στην Ιστορία, για τους μαθητές Λυκείου;

γ) Πως μπορούμε να προβλέψουμε το εισόδημα κάποιου εργαζόμενου, αν είναι γνωστά τα χρόνια προϋπηρεσίας του σε κάποιο αντικείμενο;

Οι μέθοδοι της **Παλινδρόμησης (Regression)** και της **Συσχέτισης (Correlation)** προσπαθούν να δώσουν απαντήσεις στα ερωτήματα αυτά. Το πρόβλημα της σχέσης ανάμεσα σε μεταβλητές είναι αρκετά γενικό και μπορεί να επεκταθεί σε περισσότερες των δύο μεταβλητών. Στις σημειώσεις όμως αυτές θα ασχοληθούμε μόνο με την απλούστερη περίπτωση των μεταβλητών Χ και Υ.

Αν και οι διαδικασίες της Παλινδρόμησης και της Συσχέτισης, ασχολούνται και οι δύο με τις σχέσεις ανάμεσα σε μεταβλητές, διαφέρουν δραστικά στη φύση τους και στον τρόπο με τον οποίο μεταχειρίζονται τις μεταβλητές.

Στην παλινδρόμηση, προσπαθούμε να απαντήσουμε στο ερώτημα 4, από τα παραπάνω, δηλαδή: Πώς αλλάζει η (μέση) τιμή της Υ, όταν αλλάζει η τιμή της Χ; Η Χ ονομάζεται ανεξάρτητη μεταβλητή (independent variable). H μεταβλητή Υ πάντοτε είναι τυχαία, δεν μπορεί να ελεγχθεί από τον ερευνητή και ονομάζεται εξαρτημένη (dependent variable). Αφού κατασκευαστεί η σχέση ανάμεσα στις μεταβλητές, η μεταβλητή Χ μπορεί να χρησιμοποιηθεί για την πρόβλεψη των τιμών της Υ.

Στην συσχέτιση, η οποία δίνει απαντήσεις στα ερωτήματα α,β, δεν υπάρχει διάκριση μεταξύ των μεταβλητών. Θεωρούνται και οι δύο τυχαίες (μη ελεγχόμενες από τον ερευνητή), και δεν διαχωρίζονται σε ανεξάρτητη - εξαρτημένη. Και οι δύο έχουν την ίδια βαρύτητα, στην έρευνά μας και η σχέση τους είναι συμμετρική. Στην περίπτωση αυτή ενδιαφερόμαστε να υπολογίσουμε αριθμητικά, το βαθμό στον οποίο οι μεταβλητές σχετίζονται μεταξύ τους.

Στις παρούσες σημειώσεις θα ασχοληθούμε μόνο με την Απλή Συσχέτιση μεταξύ ΔΥΟ χαρακτηριστικών, η οποία μπορεί να εκφραστεί με έναν αριθμό, που ονομάζεται Συντελεστής Συσχέτισης (Correlation Coefficient). Υπάρχουν κυρίως δύο Συντελεστές Συσχέτισης που χρησιμοποιούνται στα ποσοτικά χαρακτηριστικά. Ο Συντελεστής του Pearson και ο Συντελεστής του Spearman.

Ο **Συντελεστής του Pearson** χρησιμοποιείται κανονικά όταν τα δεδομένα μας είναι συνεχή, αλλά παρουσιάζουν ταυτόχρονα και μια κανονικότητα (συμμετρικότητα δηλαδή). Μπορούμε με ένα ιστόγραμμα ή θηκόγραμμα να ελέγξουμε τη συμμετρία των δεδομένων μας προσεγγιστικά. Ονομάζεται επίσης και Συντελεστής Γραμμικής Συσχέτισης.

Ο **Συντελεστής του Spearman**, χρησιμοποιείται σε όλες τις άλλες περιπτώσεις, καθώς και στην περίπτωση που έχουμε διακριτά δεδομένα, αλλά μπορεί να χρησιμοποιηθεί και στα ποιοτικά διατάξιμα χαρακτηριστικά.

**Παράδειγμα 2**

Θέλουμε να ελέγξουμε αν το ύψος έχει κάποια σχέση με το βάρος σε άτομα 15-18 ετών (μαθητές Λυκείου). Μελετήσαμε 50 μαθητές.

**Διάγραμμα Διάχυσης**

Τα δεδομένα παρουσιάζονται σε ένα γράφημα (στον Χ βάλαμε το ύψος και στο Υ το βάρος). Το διάγραμμα X-Y, ονομάζεται **Διάγραμμα Διάχυσης (Scatter)**



Από το διάγραμμα αυτό, μπορούμε να έχουμε μια πρώτη οπτική επαφή με τα δεδομένα και να βγάλουμε τα πρώτα συμπεράσματα. Στην περίπτωση αυτή φαίνεται ότι υπάρχει μια τάση να έχουμε αυξημένο βάρος σε μαθητές με μεγάλο ύψος και μικρότερο βάρος σε μαθητές με χαμηλό ύψος. Στην περίπτωση αυτή η συσχέτιση είναι ΘΕΤΙΚΗ.

Υπάρχει όμως περίπτωση να έχουμε μια σχέση, όπου σε αυξημένες τιμές ενός χαρακτηριστικού (Χ), αντιστοιχούν μειωμένες τιμές του άλλου (Υ), οπότε το διάγραμμα μπορεί να φαίνεται όπως παρακάτω, και τότε έχουμε ΑΡΝΗΤΙΚΗ συσχέτιση

**Συντελεστής Συσχέτισης του Pearson**

O συντελεστής συσχέτισης του Pearson, συμβολίζεται στο εξής με *r* και παίρνει τιμές ανάμεσα στο -1 και στο +1. Όταν είναι θετικός (δηλαδή από το 0 μέχρι το +1), έχουμε Θετική Συσχέτιση, ενώ όταν είναι αρνητικός (από -1 μέχρι το 0), έχουμε Αρνητική Συσχέτιση.

Όσο πιο κοντά είναι τα δεδομένα μας σε ευθεία γραμμή, τόσο πιο μεγάλος είναι ο συντελεστής (πλησιάζει τον αριθμό 1). Μερικά παραδείγματα βλέπουμε στο παρακάτω σχήμα

(για θετικές συσχετίσεις) και στο παρακάτω για αρνητικές

Γενικά όσο στενεύει η «έλλειψη» που περιλαμβάνει τα δεδομένα μας, τόσο πιο κοντά στον αριθμό 1 είναι ο συντελεστής.

εμπειρικά, μπορούμε να πούμε ότι:

1. Οταν 0<*r*≤0,25, τότε είναι πολύ μικρή συσχέτιση
2. Οταν 0,25≤*r*≤0,50, τότε έχουμε ελαφρά συσχέτιση
3. Οταν 0,50≤*r*≤0,75, η συσχέτιση είναι σχετικά ισχυρή και
4. Οταν 0,75≤*r*≤1, τότε η συσχέτιση είναι πολύ ισχυρή.

*Αντίστοιχα μπορούμε να πούμε και για τις αρνητικές τιμές του r*

**Παρατήρηση:**

O τύπος που μας υπολογίζει το συντελεστή συσχέτισης του Pearson, είναι

$$r=\frac{\sum\_{}^{}\left(X-\overbar{X}\right)(Y-\overbar{Y})}{\sqrt{\sum\_{}^{}(X-\overbar{X})^{2}∙\sum\_{}^{}(Y-\overbar{Y})^{2}∙}}$$

Αν και ο τύπος δεν είναι ιδιαίτερα δύσκολος στους υπολογισμούς, παρόλα αυτά δεν θα τον χρησιμοποιήσουμε στις παρούσες σημειώσεις, όπου μας ενδιαφέρει μόνο η ερμηνεία του και όχι ο υπολογισμός. Αυτός μπορεί να γίνει με οποιοδήποτε πρόγραμμα στατιστικής ανάλυσης.

**Ασυσχέτιστα Χαρακτηριστικά**

όταν ο συντελεστής συσχέτισης είναι ίσος με το 0 (*r*=0), τότε τα χαρακτηριστικά ΔΕΝ έχουν καμία σχέση, και ονομάζονται **ασυσχέτιστα** ή **ανεξάρτητα**. Στο παρακάτω σχήμα βλέπουμε μια περίπτωση ασυσχέτιστων χαρακτηριστικών

Πρεπει κάθε φορά, που υπολογίζουμε κάποιο συντελεστή συσχέτισης στο δείγμα, να μπορούμε να αποκλείσουμε την περίπτωση, σε ολόκληρο τον πληθυσμό να είναι η συσχέτιση μηδενική. Δηλαδή να απαντήσουμε στο βασικό ερευνητικό ερώτημα:

**Το Ερώτημα:**

Το ερώτημα αφορά τα χαρακτηριστικά Χ και Υ, και είναι:

*«Η συσχέτιση που βρήκαμε ανάμεσα στα χαρακτηριστικά στο δείγμα ήταν τυχαία ή μήπως ΥΠΑΡΧΕΙ πράγματι σημαντική συσχέτιση και στον πληθυσμό;»*

Tελικά με το δείγμα μας επιβεβαιώνουμε ότι ΥΠΑΡΧΕΙ πράγματι κάποια σχέση μεταξύ των χαρακτηριστικών και στον πληθυσμό;

**H Τιμή P και η Απόφαση**

Και στην περίπτωση Β, όπως και στην Α, οι Υπολογιστές μας δίνουν ως αποτέλεσμα, ένα «μαγικό» αριθμό στην κλίμακα από το 0 μέχρι το 1, που μας δίνει την ΠΙΘΑΝΟΤΗΤΑ να είναι τα αποτελέσματα της συσχέτισης τυχαία.

Τον αριθμό αυτό τον ονομάζουμε **Τιμή P (P-value)**

Αν ο αριθμός αυτός είναι πολύ μικρός, τότε σ συσχέτιση μεταξύ των Χ και Υ, που παρατηρήθηκε στο δείγμα δεν μπορεί να είναι τυχαία, άρα υπάρχει σημαντική συσχέτιση και στον πληθυσμό. Αν ο αριθμός είναι μεγάλος τότε υπάρχει πιθανότητα η συσχέτιση να ήταν τυχαία. Το όριο ανάμεσα στο «μεγάλο» και στο «μικρό» στην παραπάνω πρόταση είναι το 5% ή το 0,05. Ανακεφαλαιώνοντας

* **Αν το P<0,05 (πιο μικρό από το 5%) τότε η συσχέτιση ανάμεσα στο Χ και στο Υ που βρήκαμε στο δείγμα δεν είναι τυχαία, δηλαδή φαίνεται πως τα Χ και Υ συσχετίζονται και στον πληθυσμό**
* **Αν το P>0,05 (πιο μεγάλο από το 5%) τότε η συσχέτιση ανάμεσα στο Χ και στο Υ που βρήκαμε στο δείγμα μπορεί να θεωρηθεί τυχαία, δηλαδή φαίνεται πως τα Χ και Υ δεν συσχετίζονται στον πληθυσμό**

Η πιθανότητα να έχουμε κάνει λάθος στον ισχυρισμό μας είναι ίση με P. Για το λόγο αυτό, όσο πιο μικρό είναι το P, τόσο πιο σίγουροι είμαστε για την απόφαση μας και πιο έγκυρο είναι το αποτέλεσμα της έρευνας μας.

**Παράδειγμα 2 (συνέχεια - απόφαση)**

Στο παράδειγμα 2, υπολογίσαμε *r* = 0,86 άρα φαίνεται ότι υπάρχει μια θετική συσχέτιση ανάμεσα στο BMI και τη συστολική πίεση (και σύμφωνα με τους εμπειρικούς κανόνες φαίνεται να είναι πολύ ισχυρή). Όμως πόσο σίγουροι είμαστε πως αυτό το εύρημα είναι σημαντικό; Μήπως είναι τυχαίο;

Από ένα στατιστικό πρόγραμμα βρίσκουμε ότι

 Pearson’s Correlation = 0,86

 Sig. = 0,0000… (P-value)

Αφού P<<0,05 άρα υπάρχει μια θετική συσχέτιση ανάμεσα στο ύψος και το βάρος των μαθητών.

Επειδή η Συσχέτιση είναι Θετική, αυτό σημαίνει ότι όσο μεγαλύτερο είναι το ύψος τόσο περισσότερο βάρος περιμένουμε να έχει.

**Πίνακας για τη σημαντικότητα του *r :***

Το P είναι αντιστρόφως ανάλογο με το *r* . Γενικά όσο πιο μεγάλο *r* βρίσκουμε, τόσο το P γίνεται πιο μικρό. Για κάθε συγκεκριμένο μέγεθος δείγματος που παίρνουμε, υπάρχει λοιπόν ένας κρίσιμος αριθμός (***r-κρίσιμο***), που αν το *r* που βρήκαμε εμείς είναι πιο μεγάλο από το *r-κρίσιμο*, τότε είμαστε βέβαιοι ότι και το P θα είναι μικρότερο από το 5%, δηλαδή θα είναι σημαντικό.

Ο παρακάτω πίνακας δίνει τα r-κρίσιμα, για διάφορα μεγέθη δείγματος (n):

|  |  |  |  |
| --- | --- | --- | --- |
| ***n*** | ***r-κρισιμο***  | ***n*** | ***r-κρισιμο*** |
| 10 | 0.632 | 25 | 0,396 |
| 11 | 0.602 | 26 | 0,388 |
| 12 | 0.576 | 27 | 0,381 |
| 13 | 0.553 | 28 | 0,374 |
| 14 | 0.532 | 29 | 0,367 |
| 15 | 0.514 | 30 | 0,361 |
| 16 | 0.497 | 40 | 0,312 |
| 17 | 0.482 | 50 | 0,279 |
| 18 | 0.468 | 60 | 0,254 |
| 19 | 0.456 | 70 | 0,236 |
| 20 | 0.444 | 80 | 0,221 |
| 21 | 0.433 | 90 | 0,209 |
| 22 | 0.423 | 100 | 0,198 |
| 23 | 0.413 | 400 | 0,100 |
| 24 | 0.404 | 500 | 0,089 |

Έτσι για παράδειγμα, αν κάποιος χρησιμοποίησε 50 μετρήσεις και βρήκε *r*=0,358 περιμένει ότι αυτό είναι σημαντικό (δηλ. P<0,05), γιατί το *r-κρίσιμο* για 50 μετρήσεις είναι 0,279 και προφανώς 0,358>0,279. Αν όμως έβρισκε, π.χ. *r*=0,175, δεν θα το θεωρούσε σημαντικό, αφού 0,175<0,279.

**Για τις αρνητικές τιμές ισχύει το ίδιο. Κοιτάμε πάντα την απόλυτη τιμή του *r* που βρήκαμε, αν είναι μεγαλύτερη ή μικρότερη από το *r-κρίσιμο*.**

**Ο συντελεστής συσχέτισης Spearman**

Έχει παρόμοιες ιδιότητες με αυτόν του Pearson, δηλαδή είναι κι αυτός αρνητικός, μηδέν ή θετικός, στην κλίμακα από το -1 μέχρι το +1. Χρησιμοποιείται όταν έχουμε ποιοτικά διατάξιμα χαρακτηριστικά (ή απαντήσεις στην κλίμακα 1-2-3-4-5 – καθόλου, λίγο, μέτρια, πολύ, πάρα πολύ), αλλά μπορεί να χρησιμοποιηθεί και στα ποσοτικά, κυρίως σε περιπτώσεις που δεν έχουμε συμμετρικές κατανομές. Και στην περίπτωση αυτή οι υπολογιστές δίνουν μια τιμή P, με την οποία κρίνουμε τη σημαντικότητα του.

**5.4 Περίπτωση Γ – Συσχετίσεις Ποιοτικών Χαρακτηριστικών**

Αν θέλουμε να ελέγξουμε αν υπάρχει σχέση ανάμεσα σε δύο ΠΟΙΟΤΙΚΑ χαρακτηριστικά, χρησιμοποιούμε τη διαδικασία που ονομάζεται **X2-TEST (Chi-square Test)**

Tα χαρακτηριστικά μπορεί να είναι και τα δύο ονομαστικά ή και τα δύο διατάξιμα ή ένα ονομαστικό κι ένα διατάξιμο.

Όταν σε ένα δείγμα μελετάμε ταυτόχρονα δύο ποιοτικά χαρακτηρηστικά, τα αποτελέσματα παρουσιάζονται σε ένα **πίνακα διπλής εισόδου** ή ένα **πίνακα συνάφειας** (**contingency table**).

Ένας πίνακας συνάφειας για τα χαρακτηριστικά Α και Β είναι της μορφής

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  B A  | 1 | 2 | ... | c |
| 1 | f11 | f12 | ... | f1c |
| 2 | f21 | f22 | ... | f2c |
| ... | ... | ... | ... | ... |
| r | fr1 | fr2 | ... | frc |

Τα f που αναφέρονται στα κελιά του πίνακα, είναι οι συχνότητες που βρήκαμε στο δείγμα και το άθροισμα όλων των συχνοτήτων μας δίνει το *n.*

**Παράδειγμα 3**

Σε ένα ερωτηματολόγιο που δόθηκε σε 140 κοινωνικούς λειτουργούς, που εργαζόταν σε νοσοκομεία μιας ευρωπαϊκής πρωτεύουσας και καταγράφηκε η κλινική που εργαζόταν και η ικανοποίηση τους από τις συνθήκες εργασίας. Η κλινική ήταν το ένα ποιοτικό χαρακτηριστικό (ογκολογική, νευρολογική, καρδιολογική, παιδιατρική) και η ικανοποίηση ήταν το άλλο (καθόλου, λίγο, πολύ). Ο πίνακας συνάφειας έχει τη μορφή

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Ογκολογική** | **Νευρολογική** | **Καρδιολογική** | **Παιδιατρική** |
| **Καθόλου** | 7 | 14 | 12 | 5 |
| **Λίγο** | 22 | 14 | 14 | 12 |
| **Πολύ** | 6 | 9 | 15 | 10 |

Στις στήλες βάλαμε την κλινική και στις γραμμές την ικανοποίηση.

Σε ένα τέτοιο πίνακα, συμπληρώνουμε πάντα μια στήλη και μια γραμμή, που περιλαμβάνουν τα **περιθωριακά αθροίσματα** γραμμών και στηλών, ως εξής

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Ογκολογική** | **Νευρολογική** | **Καρδιολογική** | **Παιδιατρική** | **ΣΥΝΟΛΟ** |
| **Καθόλου** | 7 | 14 | 12 | 5 | **38** |
| **Λίγο** | 22 | 14 | 14 | 12 | **62** |
| **Πολύ** | 6 | 9 | 15 | 10 | **40** |
| **ΣΥΝΟΛΟ** | **35** | **37** | **41** | **27** | **140** |

Τα περιθωριακά αθροίσματα είναι οι συχνότητες που έχουμε καταγράψει σε κάθε μια από τις ερωτήσεις χωριστά, δηλαδή για τον βαθμό ικανοποίησης:

|  |  |  |
| --- | --- | --- |
| **Βαθμός ικανοπ.**  | **Συχνότητα**  | **Σχ.Συχνότητα**  |
| Καθόλου  | 38  | 27,1%  |
| Λίγο  | 62  | 44,3%  |
| Πολύ  | 40  | 28,6%  |
| **ΣΥΝΟΛΟ**  | **140**  | **100,0%**  |

και για την κλινική

|  |  |  |
| --- | --- | --- |
| **Κλινική**  | **Συχνότητα**  | **Σχ.Συχνότητα**  |
| Ογκολογική  | 35  | 25,0%  |
| Νευρολογική | 37  | 26,4%  |
| Καρδιολογική  | 41  | 29,3%  |
| Παιδιατρική  | 27  | 19,3%  |
| **ΣΥΝΟΛΟ**  | **140**  | **100,0%**  |

**Το Σύνθετο Ραβδόγραμμα**.

Όταν έχουμε ένα χαρακτηριστικό, το παρουσιάζουμε με απλό ραβδόγραμμα, όταν όμως έχουμε δύο χαρακτηριστικά (πίνακα συνάφειας), τότε χρησιμοποιούμε **Σύνθετο Ραβδόγραμμα**.

Για το σύνθετο ραβδόγραμμα, επιλέγουμε ένα από τα δύο χαρακτηριστικά και το τοποθετούμε στον άξονα Χ (π.χ. την κλινική). Για να παρουσιάσουμε το δεύτερο χαρακτηριστικό, επιλέγουμε διαφορετικά ορθογώνια με διαφορετικά σχέδια ή χρώματα. Στον άξονα Υ παρουσιάζονται οι συχνότητες.

**Το Ερώτημα**

*Υπάρχει κάποια σχέση ανάμεσα στα δύο ποιοτικά χαρακτηριστικά ή μήπως αυτά είναι ανεξάρτητα;*

To συγκεκριμένο ερώτημα, μπορεί να φαίνεται απλό στη διατύπωση, παρουσιάζει όμως δυσκολίες όταν προσπαθούμε να κατανοήσουμε τι σημαίνει σχέση ανάμεσα σε δύο χαρακτηριστικά και τι σημαίνει ανεξαρτησία. Θα προσπαθήσουμε να βοηθήσουμε τον αναγνώστη, στην κατανόηση αυτή, ακολουθώντας το παράδειγμα 3. Επανερχόμενοι λοιπόν στο παράδειγμα, διατυπώνουμε το ερώτημα:

*«Έχει σχέση ο βαθμός ικανοποίησης με την κλινική που εργάζονται;»*

Θα προσπαθήσουμε στη συνέχεια να κατανοήσουμε Τι σημαίνει «σχέση» ανάμεσα στα δύο αυτά χαρακτηριστικά. Ξεκινώντας ας υπολογίσουμε τα ποσοστά βαθμού ικανοποίησης σε κάθε κλινική χωριστά, στο δείγμα μας

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Ογκολογική** | **Νευρολογική** | **Καρδιολογική** | **Παιδιατρική** |
| **Καθόλου** | 20,0% | 37,8% | 29,3% | 18,5% |
| **Λίγο** | 62,9% | 37,8% | 34,1% | 44,4% |
| **Πολύ** | 17,1% | 24,3% | 36,6% | 37,0% |
| **ΣΥΝΟΛΟ** | **100%** | **100%** | **100%** | **100%** |

Σημείωση: Τα ποσοστά είναι οι σχετικές συχνότητες, δηλαδή για την πρώτη στήλη (Ογκολογική), διαιρούμε όλες τις συχνότητες δια 35 που είναι το σύνολο της ογκολογικής και πολλαπλασιάζουμε με το 100, π.χ.

 7/35\*100 = 20,0 (%) κλπ.

Με τον ίδιο τρόπο συνεχίζουμε και για τις άλλες στήλες (κλινικές).

Όταν τα ποσοστά κατανέμονται με τον ίδιο τρόπο στις τέσσερις κλινικές, ΤΟΤΕ ΔΕΝ υπάρχει σχέση ανάμεσα στο βαθμό ικανοποίησης και την κλινική

Στο παράδειγμα μας όμως έχουμε

Ογκολογική 20% - 63% - 17%

Νευρολογική 38% - 38% - 24%

Καρδιολογική 29% - 34% - 37%

Παιδιατρική 19% - 44% - 37%

Φαίνεται ότι η κατανομή των ποσοστών δεν είναι η ίδια, άρα υπάρχει μια σχέση, στο δείγμα μας.

Για να μην υπάρχει σχέση θα έπρεπε να βρίσκαμε τις ίδιες κατανομές ποσοστών σε κάθε κλινική, π.χ.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Ογκολογική** | **Νευρολογική** | **Καρδιολογική** | **Παιδιατρική** |
| **Καθόλου** | 27% | 27% | 27% | 27% |
| **Λίγο** | 44% | 44% | 44% | 44% |
| **Πολύ** | 29% | 29% | 29% | 29% |
| **ΣΥΝΟΛΟ** | **100%** | **100%** | **100%** | **100%** |

Αυτό δηλαδή θα σήμαινε, ότι σε οποιαδήποτε κλινική και να δουλεύει κανείς, ο βαθμός ικανοποίησης του για την εργασία του είναι ο ίδιος. Δηλ. πάντα το 27 % των νοσηλευτών δεν είναι καθόλου ικανοποιημένοι, το 44% είναι λίγο ικανοποιημένοι και το 29% είναι πολύ ικανοποιημένοι.

Στην περίπτωση όμως που έχουμε διαφορετικά ποσοστά ανα κλινική, τότε έχει σχέση η κλινική που εργάζεται με το βαθμό ικανοποίησης, αφού στην Νευρολογική έχει κανείς πολλές πιθανότητες να μην είναι καθόλου ικανοποιημένος (38%), ενώ στην παιδιατρική οι πιθανότητες αυτές είναι οι μισές (19%). Αντίστοιχα συμβαίνει και στις άλλες κατηγορίες.

Διαφορετικές κατανομές ποσοστών λοιπόν στο δείγμα, σημαίνει ότι υποψιαζόμαστε μια **σχέση** ανάμεσα στα χαρακτηριστικά.

Χρειαζόμαστε στην περίπτωση αυτή επίσης, μια στατιστική διαδικασία (τεστ), ώστε να αποκλείσουμε την πιθανότητα να είναι τα αποτελέσματα του δείγματος, τυχαία. Η μέθοδος που εξετάζει αν δύο ποιοτικά χαρακτηριστικά έχουν σχέση ή όχι, είναι το **τεστ Χ2 (chi-square test)**

**H Τιμή P και η Απόφαση**

Οι Υπολογιστές χρησιμοποιώντας τη διαδικασία Χ2, μας δίνουν ως αποτέλεσμα, ένα «μαγικό» αριθμό στην κλίμακα από το 0 μέχρι το 1, που μας δίνει την ΠΙΘΑΝΟΤΗΤΑ να είναι τυχαία η σχέση που βρήκαμε στο δείγμα.

Τον αριθμό αυτό τον ονομάζουμε **Τιμή P (P-value)**

Αν ο αριθμός αυτός είναι πολύ μικρός, τότε σ σχέση μεταξύ των χαρακτηριστικών Α και Β, που παρατηρήθηκε στο δείγμα δεν μπορεί να είναι τυχαία, άρα υπάρχει σημαντική σχέση και στον πληθυσμό. Αν ο αριθμός είναι μεγάλος τότε υπάρχει πιθανότητα η σχέση να ήταν τυχαία. Το όριο ανάμεσα στο «μεγάλο» και στο «μικρό» στην παραπάνω πρόταση είναι το 5% ή το 0,05. Ανακεφαλαιώνοντας

* **Αν το P<0,05 (πιο μικρό από το 5%) τότε η σχέση ανάμεσα στα ποιοτικά χαρακτηριστικά Α και Β, που βρήκαμε στο δείγμα δεν είναι τυχαία, δηλαδή φαίνεται πως τα Α και Β έχουν κάποια σχέση και στον πληθυσμό**
* **Αν το P>0,05 (πιο μεγάλο από το 5%) τότε η σχέση ανάμεσα στα ποιοτικά χαρακτηριστικά Α και Β, που βρήκαμε στο δείγμα μπορεί να θεωρηθεί τυχαία, δηλαδή φαίνεται πως τα Α και Β δεν έχουν κάποια σχέση στον πληθυσμό**

Η πιθανότητα να έχουμε κάνει λάθος στον ισχυρισμό μας είναι ίση με P. Για το λόγο αυτό, όσο πιο μικρό είναι το P, τόσο πιο σίγουροι είμαστε για την απόφαση μας και πιο έγκυρο είναι το αποτέλεσμα της έρευνας μας.

**Παράδειγμα 3 (συνέχεια - απόφαση)**

Εδώ υπολογίσαμε X2=10,48 και P=0,11.

Αφού **p>0,05** τότε δεν φαίνεται να υπάρχει κάποια σχέση ανάμεσα στο βαθμό ικανοποίησης και στην κλινική που εργάζονται οι κοινωνικοί λειτουργοί. Οι διαφοροποιήσεις σε ποσοστά που βρήκαμε στο δείγμα είναι μάλλον τυχαία.

Δίνουμε στη συνέχεια ένα ακόμη παράδειγμα ώστε ο αναγνώστης να κατανοήσει καλύτερα την παραπάνω διαδικασία

**Παράδειγμα 4**

Σε μια έρευνα που έγινε σε μαθητές Λυκείου, οι οποίοι κατατάχθηκαν αναφορικά με την επίδοση τους στα μαθήματα και στις συνθήκες στέγασης τους, είχαμε τα ακόλουθα αποτελέσματα:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|    | Απουσία στοιχειωδών ανέσεων  | Στοιχειώδεις ανέσεις  | Ανετο αστικό περιβάλλον  | Πολυτελές περιβάλλον  |
| Αριστοι  | 2  | 42  | 75  | 29  |
| Μέτριοι  | 14  | 78  | 70  | 20  |
| Κακοί  | 18  | 50  | 31  | 3  |

Το σύνθετο ραβδόγραμμα και ο πίνακας ποσοστών δίνονται παρακάτω

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|    | Απουσία στοιχειωδών ανέσεων  | Στοιχειώδεις ανέσεις  | Ανετο αστικό περιβάλλον  | Πολυτελές περιβάλλον  |
| Αριστοι  | 5,9 % | 24,7 % | 42,6 % | 55,8 % |
| Μέτριοι  | 41,2 % | 45,9 % | 39,8 % | 38,5 % |
| Κακοί  | 52,9 % | 29,4 % | 17,6 % | 5,8 % |
| **ΣΥΝΟΛΟ** | **100%** | **100%** | **100%** | **100%** |

Από τον πίνακα ποσοστών φαίνεται πως όσοι ζουν σε συνθήκες απουσίας στοιχειωδών ανέσεων, έχουν πολύ λίγες πιθανότητες να είναι άριστοι μαθητές (5,9%), σε αντίθεση με αυτούς που ζουν σε άνετο ή πολυτελές περιβάλλον (42,6% ή 55,8%)

Αφού λοιπόν οι κατανομές των ποσοστών ανα συνθήκες στέγασης, δεν είναι ίδιες, υποψιαζόμαστε ότι υπάρχει κάποια σχέση ανάμεσα στις συνθήκες αυτές και στην επίδοση των μαθητών. Αναλύοντας τα δεδομένα με τη βοήθεια υπολογιστή και με το Χ2 τέστ, βρήκαμε

 X2=48,6 και P=0,0000…

Αφού **p<0,05** τότε φαίνεται ότι υπάρχει σχέση ανάμεσα στις συνθήκες στέγασης και την επίδοση των μαθητών. Όσο πιο άνετες είναι οι συνθήκες τόσο αυξάνει το ποσοστό των άριστων μαθητών και μειώνεται αυτό των μέτριων και κακών.