

## ΑΠΑΡΙΘΜΗΣΗ ΚΑΙ ΤΑΞΙΝΟΜΗΣΗ

Τα δεδομένα συχνά συνιστούν καταμετρήσεις αριθμών στοιχείων με συγκεκριμένα χαρακτηριστικά (ή που ανήκουν σε συγκεκριμένες κατηγορίες), ταξινομημένες σε πίνακες μιας, δύο, τριών ή περισσότερων διαστάσεων. Οι πίνακες αυτοί ονομάζονται συνήθως *πίνακες συναφείας μιας, δύο, τριών ή περισσότερων διαστάσεων (one-, two-, three- or multi - way contingency tables)*. Κάθε διάσταση αντιστοιχεί σε μία ταξινόμηση σε κατηγορίες που αναφέρονται σε ένα χαρακτηριστικό.

### 8.1 ΠΙΝΑΚΕΣ ΣΥΝΑΦΕΙΑΣ (Contingency Tables)

Σύμφωνα με τα παραπάνω, ένας  $r \times c$  πίνακας συναφείας αποτελεί μία παράθεση φυσικών αριθμών (οι οποίοι συνήθως παριστούν συχνότητες εμφάνισης αντικειμένων), ταξινομημένων σε  $r$  γραμμές και  $c$  στήλες, όπως ο πίνακας που ακολουθεί.

Στον πίνακα αυτό,  $O_{ij}$  συμβολίζει τον παρατηρούμενο αριθμό των αντικειμένων που ανήκουν στο  $(i, j)$  κελλί,  $i = 1, 2, \dots, r$ ,  $j = 1, 2, \dots, c$ .

Πίνακες συναφείας αυτής της μορφής χρησιμοποιούνται συνήθως σε σχέση με εφαρμογές για την παρουσίαση δεδομένων που περιέχονται σε  $r$  δείγματα (γραμμές), των οποίων τα στοιχεία

παριστούν μετρήσεις σε ονομαστική κλίμακα τουλάχιστον, για τον έλεγχο της υπόθεσης ότι οι πιθανότητες με τις οποίες ένα τυχαία επιλεγόμενο αντικείμενο θα ανήκει στις κατηγορίες 1, 2, ...,c (στήλες) δεν διαφέρουν από δείγμα σε δείγμα.

	1	2	...	j	...	c
1	$O_{11}$	$O_{12}$	...	$O_{1j}$	...	$O_{1c}$
2	$O_{21}$	$O_{22}$	...	$O_{2j}$	...	$O_{2c}$
...	...	...	...	...	...	...
i	$O_{i1}$	$O_{i2}$	...	$O_{ij}$	...	$O_{ic}$
...	...	...	...	...	...	...
r	$O_{r1}$	$O_{r2}$	...	$O_{rj}$	...	$O_{rc}$

Μία άλλη χρήση του  $r \times c$  πίνακα συναφείας είναι σε σχέση με ένα μοναδικό δείγμα, του οποίου κάθε στοιχείο μπορεί να ταξινομηθεί σε μία από  $r$  διαφορετικές κατηγορίες σύμφωνα με ένα κριτήριο ή χαρακτηριστικό και, ταυτόχρονα, σε μία από  $c$  διαφορετικές κατηγορίες σύμφωνα με ένα άλλο κριτήριο ή χαρακτηριστικό. Στην περίπτωση αυτή, ενδιαφέρει ο έλεγχος της υπόθεσης ότι οι κατηγορίες του ενός κριτηρίου δεν επηρεάζουν σημαντικά τις αναλογίες των αντικειμένων σε κάθε μία από τις κατηγορίες του άλλου κριτηρίου ή χαρακτηριστικού. Η χρήση αυτή των πινάκων συναφείας εξετάσθηκε στο κεφάλαιο 1 για την περίπτωση  $r = c = 2$ , όπου η κατάλληλη

στατιστική τεχνική αποδείχθηκε να είναι μία παραλλαγή του προσημικού ελέγχου, ο έλεγχος McNemar.

Οι όροι κριτήριο ή χαρακτηριστικό χρησιμοποιούνται με την ευρεία έννοιά τους και μπορούν να αναφέρονται επίσης σε καταστάσεις στις οποίες βρίσκονται τα αντικείμενα ενός δείγματος πριν και μετά από μία αγωγή (*treatment*). Στις περιπτώσεις αυτές, η προς έλεγχο μηδενική υπόθεση είναι ότι η αγωγή δεν επηρεάζει σημαντικά τις αναλογίες των αντικειμένων στις κατηγορίες των δύο καταστάσεων.

Ενας άλλος τρόπος για να ελεγχθεί αυτή η ίδια υπόθεση, βέβαια, είναι με την χρησιμοποίηση ανεξάρτητων τυχαίων δειγμάτων από τον υπό εξέταση πληθυσμό πριν και μετά την αγωγή και την σύγκρισή τους στην συνέχεια. Η πρόσθετη μεταβλητότητα, όμως, που εισάγεται από την χρησιμοποίηση των δύο διαφορετικών δειγμάτων είναι ανεπιθύμητη γιατί, όπως είναι γνωστό, τείνει να «*συσκοτίζει*» τις μεταβολές που προκαλούνται στον πληθυσμό από την χρησιμοποιούμενη αγωγή. Παρόλα αυτά, υπάρχουν περιπτώσεις κατά τις οποίες δεν είναι πρακτικό ή εφικτό να χρησιμοποιηθεί το ίδιο δείγμα δύο φορές. Οι περιπτώσεις αυτές αποτελούν τυπικά παραδείγματα της πρώτης μορφής χρήσης των  $r \times c$  πινάκων συναφείας που αναφέρθηκε παραπάνω.

### 8.1.1 Ο $\chi^2$ Έλεγχος για Υπαρξη Διαφορών σε Πιθανότητες ή στις Αναλογίες Εκπροσώπησης $r$ Πληθυσμών σε $c$ Κατηγορίες

Ας υποθέσουμε ότι έχουμε  $r$  αμοιβαία ανεξάρτητα τυχαία δείγματα μεγέθους  $n_1, n_2, \dots, n_r$ , (ένα από κάθε ένα από  $r$

πληθυσμούς), τα στοιχεία των οποίων μπορούν να ταξινομηθούν σε  $c$  κατηγορίες όπως στον πίνακα που ακολουθεί:

		Κατηγορία				
		1	2	...	$c$	Σύνολο
Δείγμα	1	$O_{11}$	$O_{12}$	...	$O_{1c}$	$n_1$
	2	$O_{21}$	$O_{22}$	...	$O_{2c}$	$n_2$
	...	...	...	...	...	...
	$r$	$O_{r1}$	$O_{r2}$	...	$O_{rc}$	$n_r$
Σύνολο	$C_1$	$C_2$	...	$C_c$	$N$	

Εδώ  $O_{ij}$  συμβολίζει τον αριθμό των παρατηρήσεων που προέρχονται από το  $i$  δείγμα και ανήκουν στην κατηγορία  $j$ ,  $i = 1, 2, \dots, r$ ,  $j = 1, 2, \dots, c$ . Επομένως,

$$n_i = O_{i1} + O_{i2} + \dots + O_{ic}, \quad i = 1, 2, \dots, r.$$

Ο αριθμός των παρατηρήσεων που ανήκουν στην  $j$  κατηγορία, συμβολίζεται με  $C_j$ ,  $j = 1, 2, \dots, c$ . Δηλαδή,

$$C_j = O_{1j} + O_{2j} + \dots + O_{rj}, \quad j = 1, 2, \dots, c.$$

Ο συνολικός αριθμός των παρατηρήσεων από όλα τα δείγματα συμβολίζεται με  $N$ , δηλαδή

$$N = n_1 + n_2 + \dots + n_r.$$

Πέρα από την υπόθεση της ανεξαρτησίας μεταξύ των δειγμάτων, απαιτείται να υποθέσουμε ότι κάθε παρατήρηση μπορεί να ταξινομηθεί σε μία ακριβώς από τις  $c$  διαθέσιμες κατηγορίες. Οι υποθέσεις που επιθυμούμε να ελέγξουμε μπορούν να διατυπωθούν με την μορφή:

$H_0$  : Οι πληθυσμοί, από τους οποίους προήλθαν τα τυχαία δείγματα, εκπροσωπούνται σε ίσες αναλογίες στις διάφορες κατηγορίες.

$H_1$  : Τουλάχιστον δύο από τους πληθυσμούς, από τους οποίους προήλθαν τα τυχαία δείγματα, εκπροσωπούνται σε διαφορετικές αναλογίες στις διάφορες κατηγορίες.

Αν με  $p_{ij}$  συμβολίσουμε την πιθανότητα μία τυχαία επιλεγόμενη τιμή από τον  $i$  πληθυσμό,  $i = 1, 2, \dots, r$  να ανήκει στην  $j$  κατηγορία,  $j = 1, 2, \dots, c$ , οι παραπάνω υποθέσεις παίρνουν την μορφή.

$$H_0 : p_{1j} = p_{2j} = \dots = p_{rj}, \quad j = 1, 2, \dots, c$$

$H_1$  :  $p_{ij} \neq p_{kj}$  για κάποια τιμή του  $j$  και για κάποιο ζεύγος τιμών  $(i, k)$ .

Αν η  $H_0$  είναι αληθής, ο αριθμός στοιχείων που περιμένουμε να παρατηρήσουμε στο κελλί  $(i, j)$  είναι

$$E_{ij} = (\text{μέγεθος του } i \text{ δείγματος})$$

× (ποσοστό των συνολικών παρατηρήσεων, που ανήκουν στην  $j$  κατηγορία),

δηλαδή,

$$E_{ij} = n_i C_j / N.$$

Επομένως, τα δεδομένα παρέχουν ενδείξεις υπέρ της μηδενικής υπόθεσης αν οι παρατηρούμενες τιμές  $O_{ij}$  είναι κοντά στις αναμενόμενες τιμές  $E_{ij}$ . Κατά συνέπεια, για τον έλεγχο αυτής της υπόθεσης, απαιτείται μια στατιστική συνάρτηση που θα αντιπροσωπεύει ένα μέτρο της εγγύτητας των παρατηρούμενων

συχνοτήτων και των αναμενόμενων συχνοτήτων (δηλαδή, των  $O_{ij}$  και των  $E_{ij}$ ). Συνήθως, χρησιμοποιείται η στατιστική συνάρτηση

$$T = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}}.$$

Η ακριβής μορφή της κατανομής της  $T$  δεν είναι εύκολο να προσδιορισθεί. Η διαδικασία για τον καθορισμό της ακριβούς μορφής της κατανομής της στατιστικής συνάρτησης  $T$  είναι πολύπλοκη και χρονοβόρα. Αποδεικνύεται, όμως, ότι μπορεί να προσεγγισθεί από την κατανομή  $\chi^2$  με  $(r-1)(c-1)$  βαθμούς ελευθερίας. Δηλαδή,

$$T = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \sim \chi_{(r-1)(c-1)}^2.$$

Είναι φανερό, από τον ορισμό της  $T$ , ότι οι μεγάλες τιμές της αποτελούν ένδειξη εναντίον της μηδενικής υπόθεσης, αφού δείχνουν μεγάλες αποκλίσεις μεταξύ των τιμών που παρατηρούνται και αυτών που αναμένονται να παρατηρηθούν αν η μηδενική υπόθεση είναι αληθής. Επομένως, η κρίσιμη περιοχή του ελέγχου ορίζεται από την ανισότητα

$$T > \chi_{(r-1)(c-1), 1-\alpha}^2,$$

όπου  $\chi_{(r-1)(c-1), 1-\alpha}^2$  συμβολίζει το  $(1-\alpha)$ -ποσοστιαίο σημείο της κατανομής  $\chi_{(r-1)(c-1)}^2$ .

Μία ισοδύναμη έκφραση για την στατιστική συνάρτηση  $T$ , η οποία προσφέρεται για ταχύτερο υπολογισμό της τιμής της, είναι η εξής:

$$T = \sum_{i=1}^r \sum_{j=1}^c \frac{O_{ij}^2}{E_{ij}} - N.$$

Επειδή χρησιμοποιείται η ασυμπτωτική κατανομή της στατιστικής συνάρτησης  $T$ , η θεωρούμενη τιμή του επιπέδου σημαντικότητας αποτελεί μία καλή προσέγγιση της πραγματικής τιμής του επιπέδου σημαντικότητας στην περίπτωση που οι αναμενόμενες συχνότητες  $E_{ij}$ ,  $i = 1, 2, \dots, r$ ,  $j = 1, 2, \dots, c$  έχουν αρκετά μεγάλες τιμές. Αν, όμως, κάποιες από τις αναμενόμενες συχνότητες είναι χαμηλές, η προσέγγιση μπορεί να μην είναι καθόλου ικανοποιητική, ιδιαίτερα μάλιστα εάν ο αριθμός των γραμμών και των στηλών είναι μικρός. Όπως και στην περίπτωση του  $\chi^2$  ελέγχου καλής προσαρμογής, συχνά στην πράξη επανακαθορίζονται οι κλάσεις με την συνένωση κελιών με τρόπο ώστε τα προκύπτοντα κελιά να έχουν αναμενόμενες συχνότητες μεγαλύτερες από 5. Ο επανακαθορισμός των κλάσεων ενός πίνακα συναφείας, όμως, απαιτεί ιδιαίτερη προσοχή γιατί η επιλογή των συνενούμενων κελιών ενδέχεται να μην έχει κανένα νόημα.

### 8.1.2 Ο $\chi^2$ Έλεγχος Ανεξαρτησίας (*The Chi-square Test for Independence*)

Η δεύτερη κατηγορία προβλημάτων, στα οποία οι  $r \times c$  πίνακες συναφείας έχουν εφαρμογή, αναφέρονται στην περίπτωση που έχουμε ένα τυχαίο δείγμα μεγέθους  $N$ , του οποίου κάθε παρατήρηση μπορεί να ταξινομηθεί σύμφωνα με δύο κριτήρια ή χαρακτηριστικά. Υπάρχουν  $r$  κατηγορίες (γραμμές) ως προς το ένα κριτήριο ή χαρακτηριστικό και  $c$  κατηγορίες (στήλες) ως προς το δεύτερο κριτήριο ή χαρακτηριστικό. Κάθε παρατήρηση του δείγματος ταξινομείται σύμφωνα και με τα δύο κριτήρια και, επομένως, περιλαμβάνεται σε ένα συγκεκριμένο κελί του  $r \times c$  πίνακα συναφείας.

		Χαρακτηριστικό Β				Σύνολο
		1	2	...	c	
Χαρακτηριστικό Α	1	$O_{11}$	$O_{12}$	...	$O_{1c}$	$R_1$
	2	$O_{21}$	$O_{22}$	...	$O_{2c}$	$R_2$
	...	...	...	...	...	...
	r	$O_{r1}$	$O_{r2}$	...	$O_{rc}$	$R_r$
	Σύνολο	$C_1$	$C_2$	...	$C_c$	$N$

Ο συνολικός αριθμός των παρατηρήσεων στην  $i$  γραμμή του πίνακα συμβολίζεται με  $R_i$ ,  $i = 1, 2, \dots, r$ . Ο συμβολισμός αυτός, σε αντίθεση με αυτόν της προηγούμενης περίπτωσης, χρησιμοποιείται για να δοθεί έμφαση στο γεγονός ότι τα σύνολα των γραμμών είναι τιμές τυχαίες και όχι γνωστές (δεδομένες), όπως θα ήταν αν οι γραμμές του πίνακα αναφέρονταν σε στοιχεία ανεξαρτήτων δειγμάτων. Ο συνολικός αριθμός των παρατηρήσεων στην  $j$  στήλη του πίνακα συμβολίζεται με  $C_j$ ,  $j = 1, 2, \dots, c$ . Το άθροισμα των τιμών σε όλα τα κελιά του πίνακα ισούται με το μέγεθος  $N$  του τυχαίου δείγματος που έχουμε στην διάθεσή μας. Είναι σαφές ότι και στην συγκεκριμένη περίπτωση, απαιτείται να υποθεθεί ότι κάθε παρατήρηση μπορεί να ταξινομηθεί σε μία ακριβώς από τις  $rc$  διαφορετικές κατηγορίες του πίνακα συναφείας.

Η απαιτούμενη κλίμακα μέτρησης είναι ονομαστική, παρόλο που ανώτερες κλίμακες μπορούν επίσης να χρησιμοποιηθούν. Η υπόθεση που επιθυμούμε να ελέγξουμε είναι ότι οι γραμμές και οι στήλες του πίνακα εκπροσωπούν δύο ανεξάρτητα σχήματα ταξινόμησης. Είναι, δηλαδή, η μηδενική υπόθεση μία υπόθεση ελέγχου



ανεξαρτησίας μεταξύ των χαρακτηριστικών A και B. Συγκεκριμένα, η μηδενική υπόθεση μπορεί να διατυπωθεί ως εξής:

$H_0$  : Το ενδεχόμενο {μία παρατήρηση ανήκει στην  $i$  γραμμή} είναι ανεξάρτητο από το ενδεχόμενο {η ίδια παρατήρηση ανήκει στην  $j$  στήλη}, για κάθε  $i$  και  $j$ .

Εστω

$p_{ij} = P$  (μία τυχαία επιλεγόμενη τιμή από τον πληθυσμό ανήκει στο  $(i, j)$  κελλί),

$p_i = P$  (μία τυχαία επιλεγόμενη τιμή από τον πληθυσμό ανήκει στην  $i$  γραμμή),

$p_j = P$  (μία τυχαία επιλεγόμενη τιμή από τον πληθυσμό ανήκει στην  $j$  στήλη).

Από τον ορισμό της ανεξαρτησίας ενδεχομένων, η μηδενική υπόθεση μπορεί, επομένως, να διατυπωθεί ως εξής:

$$H_0 : p_{ij} = p_i p_j, \quad i = 1, 2, \dots, r, \quad j = 1, 2, \dots, c.$$

Η εναλλακτική υπόθεση  $H_1$  διατυπώνεται ως εξής:

$$H_1 : p_{ij} \neq p_i p_j, \text{ για κάποιες τιμές των } i, j.$$

Εάν η μηδενική υπόθεση είναι αληθής, η αναμενόμενη συχνότητα του  $(i, j)$  κελλιού είναι:

$$E_{ij} = (\text{μέγεθος του τυχαίου δείγματος})$$

× (ποσοστό των παρατηρήσεων που ανήκουν στην  $i$  γραμμή)

× (ποσοστό των παρατηρήσεων που ανήκουν στην  $j$  στήλη),

δηλαδή,

$$E_{ij} = N(R_i/N) \cdot (C_j/N),$$

ή, ισοδύναμα,

$$E_{ij} \equiv \# \text{ παρατηρήσεων που αναμένονται στο } (i, j) \text{ κελλί} = R_i C_j / N.$$

Η στατιστική συνάρτηση για τον έλεγχο των παραπάνω υποθέσεων ορίζεται όπως και προηγουμένως. Δηλαδή,

$$\begin{aligned} T &= \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \\ &= \sum_{i=1}^r \sum_{j=1}^c \frac{O_{ij}^2}{E_{ij}} - N. \end{aligned}$$

Η ακριβής μορφή της κατανομής της στατιστικής συνάρτησης  $T$  διαφέρει από αυτήν της στατιστικής συνάρτησης της προηγούμενης περίπτωσης, αλλά η διαδικασία προσδιορισμού της είναι εξίσου πολύπλοκη και χρονοβόρα. Και στην περίπτωση αυτή, όμως, η κατανομή  $\chi^2$  με  $(r-1)(c-1)$  βαθμούς ελευθερίας προσεγγίζει την ακριβή κατανομή της  $T$  αρκετά ικανοποιητικά. Επομένως, ο κανόνας απόφασης και στην περίπτωση αυτή έχει την μορφή:

Η μηδενική υπόθεση  $H_0$  απορρίπτεται αν η τιμή της στατιστικής συνάρτησης  $T$  υπερβαίνει το  $(1-\alpha)$ -ποσοστιαίο σημείο της  $\chi^2$  κατανομής με  $(r-1)(c-1)$  βαθμούς ελευθερίας, δηλαδή αν

$$T > \chi_{(r-1)(c-1), 1-\alpha}^2.$$

**Παράδειγμα 8.1.1:** Ένας κοινωνικός λειτουργός θέλει να εξετάσει αν η οικογενειακή κατάσταση των ανδρών μιας περιοχής εξαρτάται από τις θρησκευτικές τους πεποιθήσεις. Ένα τυχαίο δείγμα 500 ανδρών που ρωτήθηκαν σχετικά με τα δύο αυτά χαρακτηριστικά, έδωσε τα εξής αποτελέσματα:

Θρησκευτική Πεποίθηση

Οικογενειακή κατάσταση	A	B	Γ	Δ	Ε	
Με διαζύγιο	39	19	12	28	18	116
Ποτέ δεν πήραν διαζύγιο	172	61	44	70	37	384
	211	80	56	98	55	500

Να ελεγχθεί σε επίπεδο σημαντικότητας 1% η υπόθεση ότι τα δύο χαρακτηριστικά είναι ανεξάρτητα.

**Λύση:** Έχουμε ένα τυχαίο δείγμα, του οποίου τα άτομα έχουν ταξινομηθεί σε κατηγορίες (διαβαθμίσεις) ως προς τα χαρακτηριστικά «θρησκευτική πεποίθηση» και «οικογενειακή κατάσταση». Η προς έλεγχο μηδενική υπόθεση είναι η

$$H_0 : \text{οικογενειακή κατάσταση ανεξάρτητη των} \\ \text{θρησκευτικών πεποιθήσεων,}$$

ή, ισοδύναμα,

$$H_0 : p_{ij} = p_i \cdot p_j, \text{ για κάθε } i \text{ και } j,$$

όπου

$$p_{ij} = P(\text{Ένας τυχαία επιλεγόμενος άνδρας ανήκει στην} \\ \text{κατηγορία } (i, j))$$

$$p_i = P(\text{Η οικογενειακή κατάσταση ενός τυχαία επιλεγόμενου} \\ \text{άνδρα είναι } i)$$

και

$$p_j = P(\text{Η θρησκευτική πεποίθηση ενός τυχαία επιλεγόμενου} \\ \text{άνδρα είναι } j)$$

Αν η μηδενική υπόθεση είναι αληθής, οι αναμενόμενες συχνότητες στα κελιά του δοθέντος πίνακα υπολογίζονται από τον τύπο  $E_{ij} = R_i C_j / N$ , για κάθε  $i$  και  $j$ . Οι τιμές αυτών των συχνοτήτων σημειώνονται σε παρένθεση στον πίνακα που ακολουθεί.

### Θρησκευτική Πεποίθηση

Οικογενειακή κατάσταση	A	B	Γ	Δ	E	
Με διαζύγιο	39 (48.952)	19 (18.560)	12 (12.992)	28 (22.736)	18 (12.76)	116
Ποτέ δεν πήραν διαζύγιο	172 (162.048)	61 (61.440)	44 (43.008)	70 (75.204)	37 (42.24)	384
	211	80	56	98	55	500

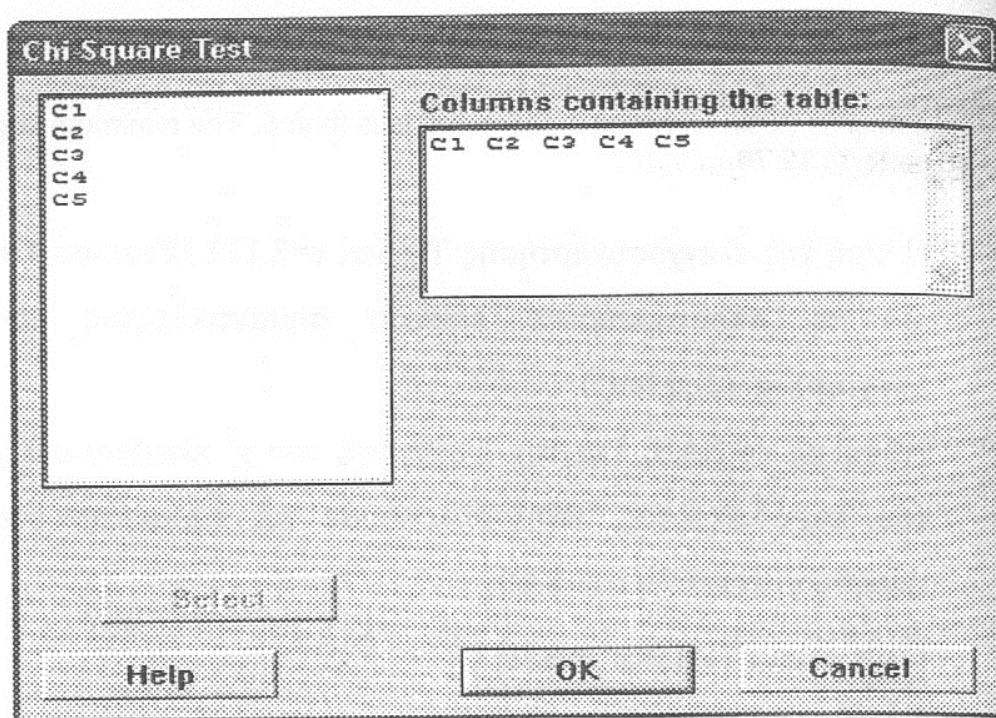
Η κατάλληλη στατιστική συνάρτηση για τον έλεγχο της παραπάνω υπόθεσης δίνεται από τον τύπο:

$$T = \sum_i \sum_j \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

και είναι προφανές ότι η κρίσιμη περιοχή ορίζεται από την ανισότητα  $T > \chi_{4,0.99}^2 \equiv 13.28$ .

Η παρατηρούμενη τιμή της  $T$  είναι  $\tau = 7.135$ . Επομένως, σε επίπεδο σημαντικότητας 1% δε φαίνεται να υπάρχει εξάρτηση μεταξύ των δύο χαρακτηριστικών.

**Λύση με το MINITAB:** Καταχωρίζουμε, σε στήλες του πακέτου, τις στήλες του πίνακα (έστω **c1** έως **c5**) και, στην συνέχεια, επιλέγουμε **Stat, Tables, Chi-Square Test** οδηγούμενοι στο εξής πλαίσιο διαλόγου:



Στο πεδίο **Columns containing the table**, δηλώνουμε τις στήλες του πίνακα συνάφειας (c1, ..., c5) και, πιέζοντας **OK**, οδηγούμεθα στα εξής αποτελέσματα:

Expected counts are printed below observed counts

	C1	C2	C3	C4	C5	Total
1	39	19	12	28	18	116
	48.95	18.56	12.99	22.74	12.76	
2	172	61	44	70	37	384
	162.05	61.44	43.01	75.26	42.24	
Total	211	80	56	98	55	500

Chi-Sq = 2.023 + 0.010 + 0.076 + 1.219 + 2.152 +  
 0.611 + 0.003 + 0.023 + 0.368 + 0.650 = 7.135  
 DF = 4, P-Value = 0.129

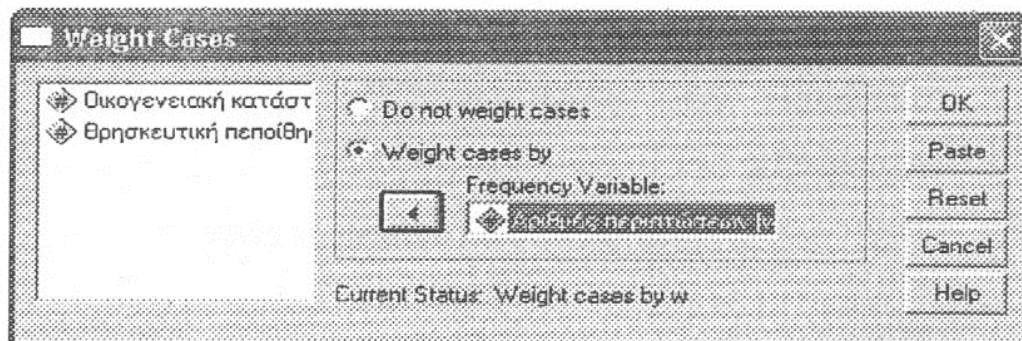
Παρατηρούμε ότι το πρόγραμμα δίνει τις παρατηρούμενες και τις αναμενόμενες συχνότητες κάθε κελλιού, την τιμή της ελεγχοσυνάρτησης T (**chi-sq=7.135**), τους βαθμούς ελευθερίας

(DF=4) και το παρατηρούμενο επίπεδο σημαντικότητας (P-value=0.129).

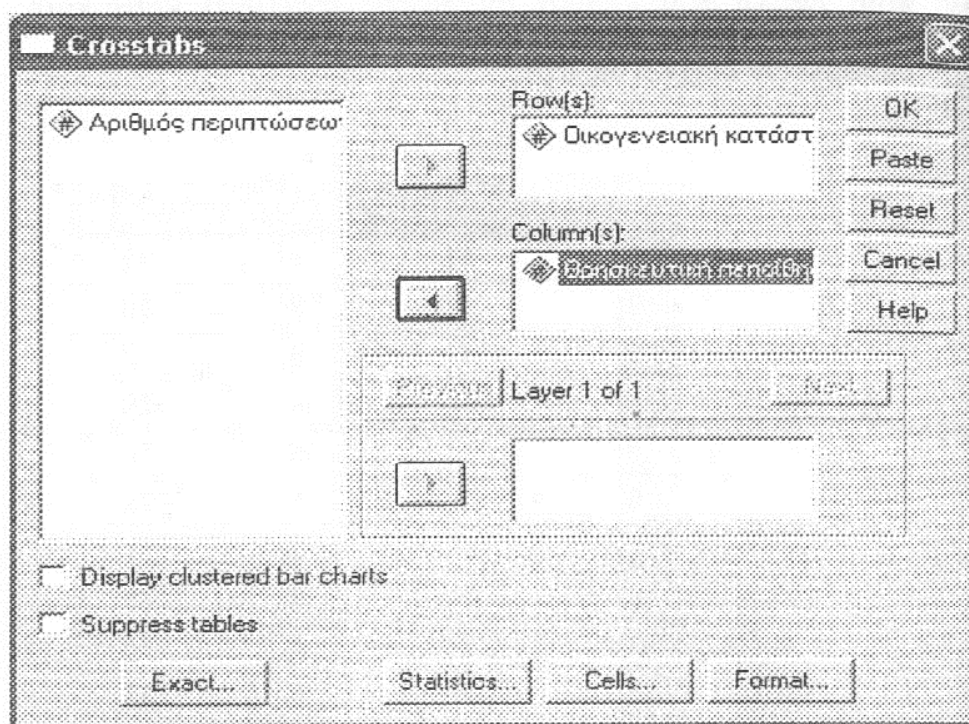
**Λύση με το SPSS:** Σε μία στήλη (με όνομα *w*), καταχωρίζουμε τις τιμές των κελιών του πίνακα και σε δύο στήλες-δείκτες (με ονόματα *row* και *col*) καταχωρίζουμε τιμές που δείχνουν σε ποια γραμμή και ποια στήλη του πίνακα συναφείας αναλογεί κάθε τιμή. Το φύλλο δεδομένων δείχνει ως εξής:

	row	col	w
1	1	1	39
2	1	2	19
3	1	3	12
4	1	4	28
5	1	5	18
6	2	1	172
7	2	2	61
8	2	3	44
9	2	4	70
10	2	5	37

Στην συνέχεια, επιλέγοντας **Data, Weight Cases**, προκύπτει το εξής πλαίσιο διαλόγου:



Επιλέγουμε **Weight cases by** και, στο επίπεδο **Frequency Variable** που ενεργοποιείται, δηλώνουμε την μεταβλητή με τις τιμές των κελιών. Πιέζουμε **OK** και, στην συνέχεια, επιλέγουμε **Analyze, Descriptive Statistics, Crosstabs** οδηγούμενοι στο εξής πλαίσιο διαλόγου:



Στα πεδία **Row(s)** και **Column(s)**, δηλώνουμε προφανώς τις αντίστοιχες μεταβλητές-δείκτες. Οι τιμές των κελιών έχουν δηλωθεί εμμέσως με τη διαδικασία του **weight cases**. Πιέζουμε το πλήκτρο **Statistics** και, στο πλαίσιο διαλόγου που προκύπτει, επιλέγουμε μόνο **Chi-Square**. Στο επανεμφανιζόμενο αρχικό πλαίσιο διαλόγου, πιέζουμε **OK**.

Μέρος των αποτελεσμάτων που παίρνουμε είναι:

#### Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	7.13	4	.129
Likelihood Ratio	6.98	4	.137

Linear-by-Linear Association 6.49 1 .011  
 N of Valid Cases 50

a 0 cells (.0%) have expected count less than 5. The minimum expected count is 12.76.

Η τιμή της ελεγκοσυνάρτησης T είναι  $\tau=7.135$  (**Pearson Chi-square**) και το παρατηρούμενο επίπεδο σημαντικότητας είναι 0.129 (**AsympSig. (2-sided)**).

**Λύση με το SAS:** Για την διεξαγωγή του  $\chi^2$  ελέγχου ανεξαρτησίας του παραδείγματος, πληκτρολογούμε τις παρακάτω εντολές στο παράθυρο εντολών του SAS:

```
data divorced;
input status religion counts;
cards;
1 1 39
1 2 19
1 3 12
1 4 28
1 5 18
2 1 172
2 2 61
2 3 44
2 4 70
2 5 37
;
run;
proc freq;
tables status*religion / chisq;
weight counts;
exact chisq;
run;
```

Το αποτέλεσμα αυτής της διαδικασίας περιέχεται στον πίνακα που ακολουθεί.

The SAS System  
 TABLE OF STATUS BY RELIGION

STATUS	RELIGION					Total
	1	2	3	4	5	
Frequency,						
Percent ,						
Row Pct ,						
Col Pct						
1	39	19	12	28	18	116
	7.80	3.80	2.40	5.60	3.60	23.20
	33.62	16.38	10.34	24.14	15.52	
	18.48	23.75	21.43	28.57	32.73	
2	172	61	44	70	37	384
	34.40	12.20	8.80	14.00	7.40	76.80
	44.79	15.89	11.46	18.23	9.64	
	81.52	76.25	78.57	71.43	67.27	
Total	211	80	56	98	55	500



42.20 16.00 11.20 19.60 11.00 100.00

STATISTICS FOR TABLE OF STATUS BY RELIGION

Statistic	DF	Value	Prob (Asymptotic)	Prob (Exact)
Chi-Square	4	7.135	0.129	0.129
Likelihood Ratio Chi-Square	4	6.985	0.137	0.142
Mantel-Haenszel Chi-Square	1	6.494	0.011	0.011
Phi Coefficient		0.119		
Contingency Coefficient		0.119		
Cramer's V		0.119		
Sample Size = 500				

Το πακέτο παρέχει την τιμή της ελεγχοσυνάρτησης που υπολογίσθηκε ίση με 7.135 (**Chi-Square**). Επίσης, παρέχει προσεγγιστικά το παρατηρούμενο επίπεδο στατιστικής σημαντικότητας (**Prob. Asymptotic**) αλλά και το ακριβές (**Prob. Exact**) που ζητήθηκε με την υποεντολή **exact chisq**. Αξίζει να σημειωθεί ότι το ασυμπτωτικό παρατηρούμενο επίπεδο σημαντικότητας είναι ίσο με το ακριβές λόγω του μεγάλου μεγέθους δείγματος που χρησιμοποιήθηκε.

**Παράδειγμα 8.1.2:** Από δύο ανεξάρτητα τυχαία δείγματα 20 ανδρών και 20 γυναικών ζητήθηκε να δοκιμάσουν μια νέα οδοντόκρεμα και να συγκρίνουν την γεύση της με την γεύση αυτής, που συνήθως χρησιμοποιούσαν. Οι απαντήσεις τους περιλαμβάνονται στον εξής πίνακα:

	Γεύση			
	καμιά διαφορά	χειρότερη	καλύτερη	
άνδρες	4	7	9	20
γυναίκες	1	6	13	20
	5	13	22	N=40

Δείχνουν τα αποτελέσματα αυτά στατιστικά σημαντική διαφορά στις προτιμήσεις ανδρών και γυναικών; ( $\alpha=5\%$ )

**Λύση:** Είναι προφανές ότι η προς έλεγχο μηδενική υπόθεση μπορεί να διατυπωθεί ως εξής:

$H_0$  : οι προτιμήσεις δεν διαφέρουν,

ή, ισοδύναμα,

$H_0$  :  $P_{\text{ανδρ},j} = P_{\text{γυν},j}$ , για κάθε  $j$ .

Κάτω από την μηδενική υπόθεση, οι αναμενόμενες συχνότητες  $E_{ij}$  στα κελιά του παραπάνω πίνακα υπολογίζονται από την σχέση

$$E_{ij} = \frac{n_i C_j}{N}, \text{ για κάθε } i \text{ και } j.$$

Οι τιμές τους σημειώνονται σε παρένθεση στα αντίστοιχα κελιά του πίνακα που ακολουθεί.

	Γεύση			
	καμιά διαφορά	χειρότερη	καλύτερη	
άνδρες	4 (2.5)	7 (6.5)	9 (11)	20
γυναίκες	1 (2.5)	6 (6.5)	13 (11)	20
	5	13	22	N=40

Η κατάλληλη στατιστική συνάρτηση για τον έλεγχο της παραπάνω υπόθεσης δίνεται από την σχέση

$$T = \sum_i \sum_j \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \sim \chi_2^2,$$

οπότε, η κρίσιμη περιοχή του ελέγχου ορίζεται από την ανισότητα

$$T > \chi_{2,0.95}^2 \equiv 5.991.$$

Η παρατηρηθείσα τιμή  $t$  της στατιστικής συνάρτησης  $T$  είναι 2.604. Επομένως, σε επίπεδο σημαντικότητας 5% δεν φαίνεται να διαφέρουν σημαντικά οι προτιμήσεις ανδρών και γυναικών.

Το κρίσιμο επίπεδο του παραπάνω ελέγχου είναι  
 $\hat{\alpha} = \max P(\text{η } T \text{ να έχει μία τιμή το ίδιο ή περισσότερο ακραία από την παρατηρηθείσα της τιμή} | H_0)$

$$= P(T \geq 2.604 | T \sim \chi_2^2) = P(\chi_2^2 \geq 2.604)$$

$$= 1 - P(\chi_2^2 < 2.604) > 0.20.$$

**Παρατήρηση:** Όπως έχει ήδη τονισθεί, στην πράξη, εάν κάποιες από τις αναμενόμενες συχνότητες είναι πολύ χαμηλές, οι κατηγορίες στις οποίες ανήκουν συνενώνονται «κατάλληλα» με άλλες κατηγορίες προκειμένου να προκύψουν κατηγορίες με αναμενόμενες συχνότητες που δεν είναι χαμηλές με στόχο την επίτευξη καλύτερης προσέγγισης του πραγματικού επιπέδου σημαντικότητας του ελέγχου. Συμπύσσοντας, σύμφωνα με αυτή την λογική, τις κατηγορίες «καμία διαφορά» και «χειρότερη» σε μία νέα κατηγορία με το όνομα «όχι καλύτερη» (ώστε, με κάποια έννοια, οι υποθέσεις να διατηρούν το νόημά τους) και επαναλαμβάνοντας τον έλεγχο, η τιμή της  $T$  που προκύπτει είναι  $t=1.616$  με αντίστοιχο κρίσιμο επίπεδο  $\hat{\alpha}=0.204$ .

**Λύση με το MINITAB:** Διεξάγοντας τον έλεγχο ανεξαρτησίας  $\chi^2$  με την βοήθεια του MINITAB κατά τα ήδη γνωστά από το προηγούμενο παράδειγμα, παίρνουμε τα εξής αποτελέσματα:

Expected counts are printed below observed counts

	C1	C2	C3	Total
1	4	7	9	20
	2.50	6.50	11.00	

	2	1	6	13	20
		2.50	6.50	11.00	
Total		5	13	22	40

Chi-Sq = 0.900 + 0.038 + 0.364 +  
0.900 + 0.038 + 0.364 = 2.604  
DF = 2, P-Value = 0.272  
2 cells with expected counts less than 5.0

Το πρόγραμμα μας προειδοποιεί ότι 2 από τα 6 κελιά του πίνακα έχουν αναμενόμενη συχνότητα μικρότερη του 5. Αυτά αντιστοιχούν στα άτομα που δεν βρήκαν καμία διαφορά στην γεύση μεταξύ νέας οδοντόκρεμας και αυτής που χρησιμοποιούν. Υιοθετώντας σύμπτυξη της κατηγορίας με τις χαμηλές αναμενόμενες συχνότητες (κατηγορία «καμία διαφορά») με την κατηγορία «χειρότερη» σε μία νέα κατηγορία με όνομα «όχι καλύτερη» και επαναλαμβάνοντας τον έλεγχο, παίρνουμε τα εξής αποτελέσματα:

$$\text{Chi-Sq} = 0.444 + 0.364 + 0.444 + 0.364 = 1.616$$

$$\text{DF} = 1, \text{ P-Value} = 0.204$$

Το κρίσιμο επίπεδο του ελέγχου τώρα προσεγγίζεται από την τιμή 0.204.

**Λύση με το SPSS:** Εργαζόμενοι όπως στο προηγούμενο παράδειγμα για την εφαρμογή του ελέγχου  $\chi^2$ , παίρνουμε τα εξής αποτελέσματα:

#### Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	2.604 <sup>a</sup>	2	.272
Likelihood Ratio	2.736	2	.255
Linear-by-Linear Association	2.416	1	.120
N of Valid Cases	40		

a 2 cells (33.3%) have expected count less than 5. The minimum expected count is 2.50.

Το πρόγραμμα μας προειδοποιεί ότι 2 από τα 6 κελιά του πίνακα έχουν αναμενόμενη συχνότητα μικρότερη του 5. Αυτά αντιστοιχούν στα άτομα που δεν βρήκαν καμία διαφορά στην γεύση μεταξύ νέας οδοντόκρεμας και αυτής που χρησιμοποιούν. Υιοθετώντας σύμπτυξη της κατηγορίας με τις χαμηλές αναμενόμενες συχνότητες (κατηγορία «καμία διαφορά» με την κατηγορία «χειρότερη» σε μία νέα κατηγορία με όνομα «όχι καλύτερη» και επαναλαμβάνοντας τον έλεγχο, παίρνουμε τα εξής αποτελέσματα:

#### Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	1.616 <sup>b</sup>	1	.204		
Continuity Correction <sup>a</sup>	.909	1	.340		
Likelihood Ratio	1.628	1	.202		
Fisher's Exact Test				.341	.170
Linear-by-Linear Association	1.576	1	.209		
N of Valid Cases	40				

a Computed only for a 2x2 table

b 0 cells (.0%) have expected count less than 5. The minimum expected count is 9.00.

Βλέπουμε ότι το παρατηρούμενο επίπεδο σημαντικότητας προσεγγίζεται τώρα από την τιμή 0.204 (πεδίο **Asymp. Sig. (2-sided)**).

**Λύση με το SAS:** Για την διεξαγωγή του ελέγχου  $\chi^2$  με το SAS, εργαζόμαστε με τον ήδη γνωστό από το προηγούμενο παράδειγμα τρόπο: Πληκτρολογούμε στο παράθυρο εντολών τις παρακάτω εντολές

```
data taste;
input gender taste counts;
cards;
1 1 4
1 2 7
1 3 9
2 1 1
2 2 6
2 3 13
;
run;
proc freq;
tables gender*taste / chisq;
```

```
weight counts;
exact chisq;
run;
```

Τα αποτελέσματα που παίρνουμε δίνονται στον πίνακα που ακολουθεί.

The SAS System  
TABLE OF GENDER BY TASTE

GENDER	TASTE			Total
Frequency,	1,	2,	3,	
1,	4,	7,	9,	20
Percent	10.00	17.50	22.50	50.00
Row Pct	20.00	35.00	45.00	
Col Pct	80.00	53.85	40.91	
2,	1,	6,	13,	20
Percent	2.50	15.00	32.50	50.00
Row Pct	5.00	30.00	65.00	
Col Pct	20.00	46.15	59.09	
Total	5	13	22	40
	12.50	32.50	55.00	100.00

STATISTICS FOR TABLE OF GENDER BY TASTE

Statistic	DF	Value	Prob (Asymptotic)	Prob (Exact)
Chi-Square	2	2.604	0.272	0.303
Likelihood Ratio Chi-Square	2	2.736	0.255	0.303
Mantel-Haenszel Chi-Square	1	2.416	0.120	0.182
Phi Coefficient		0.255		
Contingency Coefficient		0.247		
Cramer's V		0.255		

Sample size = 40

WARNING: 33% of the cells have expected counts less than 5.  
(Asymptotic) Chi-Square may not be a valid test.

Η τιμή της ελεγχοσυνάρτησης υπολογίσθηκε ίση με 2.604 και το ακριβές παρατηρούμενο επίπεδο στατιστικής σημαντικότητας ίσο με 0.303.

Σημειώνεται ότι η ασυμπτωτική τιμή του παρατηρούμενου επιπέδου στατιστικής σημαντικότητας δεν είναι ίση με την ακριβή τιμή και η προσέγγιση της κατανομής της T από την  $\chi^2$  ενδέχεται να μην ισχύει λόγω των χαμηλών συχνοτήτων κάποιων κελιών του πίνακα συναφείας. Αυτή ακριβώς την προειδοποίηση δίνει το πακέτο.

### 8.1.3 Ο $\chi^2$ Ελεγχος με Γνωστά Αθροίσματα Γραμμών και Στηλών (The Chi-square Test with Fixed Marginal Totals)

Η ενότητα αυτή αναφέρεται σε μία τρίτη κατηγορία προβλημάτων εφαρμογής των πινάκων συναφείας, τα σύνολα των γραμμών αλλά και των στηλών είναι δεδομένα (γνωστά).

Τα δεδομένα συνοψίζονται σε ένα  $r \times c$  πίνακα συναφείας, όπως και στις δύο προηγούμενες περιπτώσεις με την διαφορά ότι τα σύνολα των γραμμών και τα σύνολα των στηλών δεν είναι τυχαίες τιμές, αλλά είναι σταθερές που συμβολίζονται με  $n_i$ ,  $i = 1, 2, \dots, r$  και  $n_j$ ,  $j = 1, 2, \dots, c$ , αντίστοιχα. Ο συνολικός αριθμός των παρατηρήσεων είναι  $N$ .

	1	2	...	c	Σύνολο
1	$O_{11}$	$O_{12}$	...	$O_{1c}$	$n_{1.}$
2	$O_{21}$	$O_{22}$	...	$O_{2c}$	$n_{2.}$
...	...	...	...	...	...
r	$O_{r1}$	$O_{r2}$	...	$O_{rc}$	$n_{r.}$
Σύνολο	$n_{.1}$	$n_{.2}$	...	$n_{.c}$	$N$

Και στην περίπτωση αυτή, οι παρατηρήσεις προέρχονται από ένα μόνο τυχαίο δείγμα και κάθε παρατήρηση μπορεί να ταξινομηθεί σε ένα μόνο κελί του πίνακα.

Οι υποθέσεις που ενδιαφερόμαστε να ελέγξουμε στην περίπτωση αυτή, μπορεί να έχουν οποιαδήποτε από τις μορφές των υποθέσεων που αναφέρονται στις δύο προηγούμενες περιπτώσεις με την πρόσθετη προϋπόθεση ότι τα σύνολα των γραμμών και των

στηλών έχουν σταθερές τιμές. Εναλλακτικά, οι υποθέσεις μπορούν να διατυπώνονται με κατάλληλες τροποποιήσεις των υποθέσεων των προηγούμενων δύο κατηγοριών προβλημάτων, ώστε να ανταποκρίνονται στην συγκεκριμένη πειραματική περίπτωση. Συνήθως, οι υποθέσεις αποτελούν παραλλαγές της υπόθεσης της ανεξαρτησίας της περίπτωσης που εξετάστηκε στην προηγούμενη ενότητα με τις απαραίτητες τροποποιήσεις που υπαγορεύονται από το συγκεκριμένο πείραμα. (Μία συνήθης πρόσθετη «απαίτηση» είναι η κατασκευή ενός πίνακα συναφείας με ίσες αναμενόμενες συχνότητες στα κελιά του). Έτσι, όπως θα δούμε στο πλαίσιο του επόμενου παραδείγματος, ο έλεγχος αυτής της περίπτωσης μπορεί να χρησιμοποιηθεί για τον έλεγχο της υπόθεσης

$H_0$ : Οι τυχαίες μεταβλητές  $X$  και  $Y$  είναι ανεξάρτητες με βάση ένα δείγμα ανεξαρτήτων παρατηρήσεων  $(X_1, Y_1), (X_2, Y_2), \dots, (X_N, Y_N)$  πάνω στην διδιάστατη μεταβλητή  $(X, Y)$ .

Η εναλλακτική υπόθεση είναι η άρνηση της μηδενικής, δηλαδή

$H_1$ : Οι τυχαίες μεταβλητές  $X$  και  $Y$  είναι εξαρτημένες.

Αν η μηδενική υπόθεση είναι αληθής, οι αναμενόμενες συχνότητες στα κελιά του πίνακα δίνονται από την σχέση

$$E_{ij} = n_i n_j / N, \quad i = 1, 2, \dots, r \text{ και } j = 1, 2, \dots, c.$$

Και πάλι, η προφανής επιλογή στατιστικής συνάρτησης είναι η στατιστική συνάρτηση

$$T = \sum_{i=1}^r \sum_{j=1}^c (O_{ij} - E_{ij})^2 / E_{ij}$$

και η κρίσιμη περιοχή του ελέγχου θα αντιστοιχεί σε μεγάλες τιμές αυτής.



Η ακριβής μορφή της κατανομής της στατιστικής συνάρτησης  $T$  μπορεί να προκύψει ευκολότερα από ό,τι στις δύο προηγούμενες περιπτώσεις εφαρμογών λόγω του ότι τα σύνολα των γραμμών και των στηλών του πίνακα συναφείας είναι σταθερά. Παρόλα αυτά, η διαδικασία προσδιορισμού της κατανομής της στατιστικής συνάρτησης  $T$  εξακολουθεί να είναι πολύπλοκη και χρονοβόρα. Η  $\chi^2$  κατανομή και στην περίπτωση αυτή προσφέρει μία ικανοποιητική προσέγγιση της κατανομής της  $T$ . Συγκεκριμένα,

$$T = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \sim \chi_{(r-1)(c-1)}^2.$$

Επομένως, η κρίσιμη περιοχή του ελέγχου των παραπάνω υποθέσεων ορίζεται από την ανισότητα

$$T > \chi_{(r-1)(c-1), 1-\alpha}^2.$$

Όσο αφορά το επίπεδο σημαντικότητας  $\alpha$  του ελέγχου, ισχύουν τα σχόλια τα οποία διατυπώθηκαν στις προηγούμενες δύο περιπτώσεις εφαρμογών.

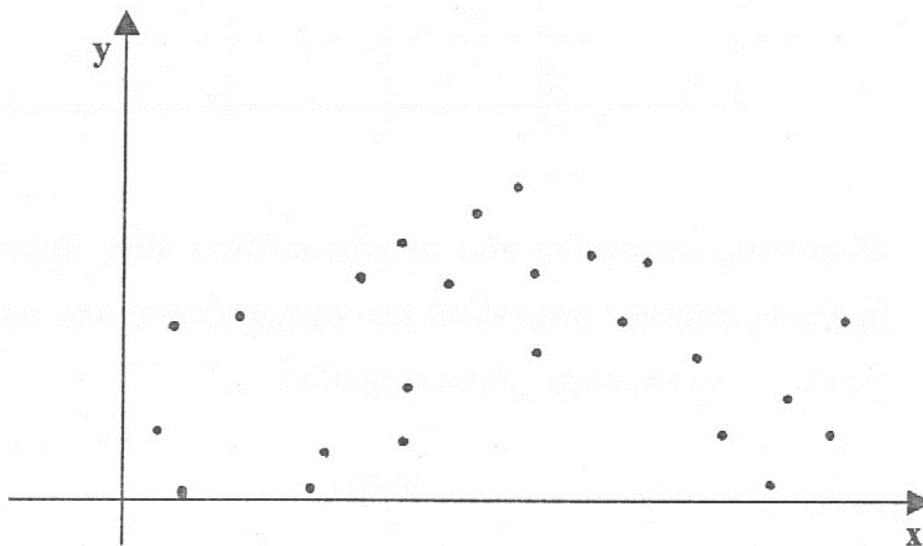
**Σημείωση:** Όπως αναφέρθηκε στα προηγούμενα, συχνά απαιτείται η κατασκευή πινάκων συναφείας έτσι ώστε οι αναμενόμενες συχνότητες στα προκύπτοντα κελιά να είναι ίσες. Μεταξύ αυτών των περιπτώσεων περιλαμβάνονται και αυτές στις οποίες περιμένει κανείς ότι οι αναμενόμενες συχνότητες θα έχουν χαμηλές τιμές, πράγμα το οποίο θα έχει επίδραση στον βαθμό που θα είναι ικανοποιητική η προσέγγιση της κατανομής της στατιστικής συνάρτησης  $T$  από την κατανομή  $\chi^2$ . Η επίδραση αυτή «αίρεται» στην περίπτωση που οι αναμενόμενες συχνότητες είναι περίπου ίσες.

**Παράδειγμα 8.1.3:** Ας θεωρήσουμε το παρακάτω διάγραμμα 24 σημείων, τα οποία παριστούν ανεξάρτητες παρατηρήσεις  $(X_1, Y_1), (X_2, Y_2), \dots, (X_{24}, Y_{24})$  πάνω στην διδιάστατη τυχαία μεταβλητή  $(X, Y)$ . Η x-συνιστώσα κάθε σημείου παριστά την παρατηρηθείσα τιμή της μεταβλητής  $X$  και η y-συνιστώσα την παρατηρηθείσα τιμή της τυχαίας μεταβλητής  $Y$  σε κάθε μία από τις παρατηρήσεις πάνω στην  $(X, Y)$ . Ας υποθέσουμε ότι τα παρατηρηθέντα ζεύγη  $(X_i, Y_i)$ ,  $i = 1, 2, \dots, 24$  είναι αμοιβαία ανεξάρτητα. Να ελεγχθούν οι υποθέσεις

$H_0$ : Οι μεταβλητές  $X$  και  $Y$  είναι ανεξάρτητες

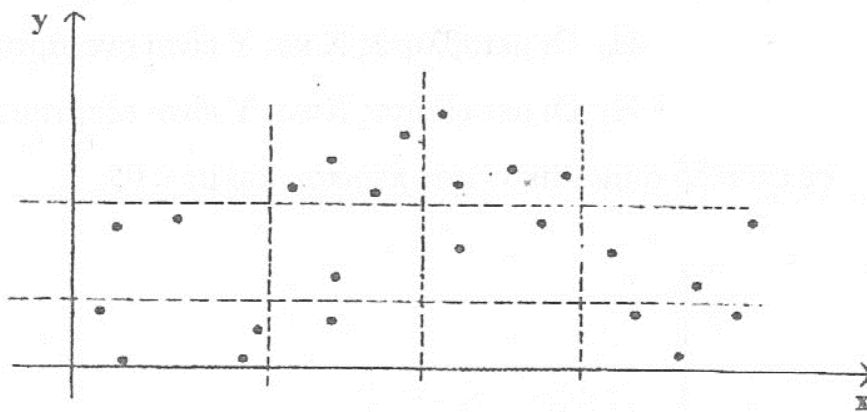
$H_1$ : Οι μεταβλητές  $X$  και  $Y$  είναι εξαρτημένες

σε επίπεδο σημαντικότητας περίπου ίσο με 0.05.



**Λύση:** Προκειμένου να σχηματίσουμε έναν πίνακα συναφείας μπορούμε να χωρίσουμε τα σημεία του διαγράμματος σε γραμμές και στήλες με την βοήθεια διακεκομμένων γραμμών παράλληλων και κάθετων στον άξονα των  $x$ . Οι αριθμοί των σημείων που θα εμφανίζονται στα κελιά τα οποία θα σχηματισθούν, θα αποτελούν τις

παρατηρούμενες συχνότητες εμφάνισης των παρατηρήσεων πάνω στην μεταβλητή (X, Y). Προκειμένου να σχηματισθεί ένας πίνακας συναφείας τέτοιος ώστε όλες οι αναμενόμενες τιμές  $E_{ij}$  να είναι ίσες, χωρίζουμε τα σημεία του διαγράμματος σε 3 γραμμές των 8 σημείων και σε 4 στήλες των 6 σημείων όπως στο διάγραμμα που ακολουθεί. Αυτός ο τρόπος χωρισμού οδηγεί σε ίσα αθροίσματα γραμμών και ίσα αθροίσματα στηλών.



Ο πίνακας συναφείας που προκύπτει δίνει τους παρατηρούμενους αριθμούς σημείων στα κελιά που σχηματίζονται από τις γραμμές και τις στήλες του προηγούμενου σχήματος.

	Στήλη				
Γραμμή	1	2	3	4	Σύνολο
1	0	4	4	0	8
2	2	1	2	3	8
3	4	1	0	3	8
Σύνολο	6	6	6	6	24

Η στατιστική συνάρτηση για τον έλεγχο των υποθέσεων του προβλήματος αυτού δίνεται από την σχέση

$$T = \sum_{i=1}^3 \sum_{j=1}^4 (O_{ij} - E_{ij})^2 / E_{ij}$$

και η κρίσιμη περιοχή μεγέθους περίπου ίσου με 0.05 ορίζεται από την ανισότητα

$$T > \chi_{6,0.95}^2 \equiv 12.59.$$

Προφανώς  $E_{ij} = (6)(8)/24 = 2$ , για κάθε  $i$  και  $j$ . Επομένως, η παρατηρούμενη τιμή  $\tau$  της στατιστικής συνάρτησης  $T$  είναι

$$\tau = \sum_{i=1}^3 \sum_{j=1}^4 \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = \sum_{i=1}^3 \sum_{j=1}^4 \frac{(O_{ij} - 2)^2}{2} = 14.$$

Επειδή η τιμή  $\tau$  υπερβαίνει την κρίσιμη τιμή, η μηδενική υπόθεση απορρίπτεται σε επίπεδο σημαντικότητας 0.05 και συμπεραίνουμε ότι η υπόθεση της ανεξαρτησίας των μεταβλητών  $X$  και  $Y$  δεν είναι εύλογη. Το κρίσιμο επίπεδο του ελέγχου είναι:

$$\hat{\alpha} = P(T \geq 14 | H_0) = P(\chi_6^2 \geq 14) = 0.03.$$

**Λύση με το MINITAB:** Κατασκευάζουμε έναν πίνακα συνάφειας όπως στην αναλυτική λύση του παραδείγματος και, στην συνέχεια, διεξάγουμε τον έλεγχο  $\chi^2$  με το MINITAB κατά τον ήδη γνωστό τρόπο. Τα αποτελέσματα που παίρνουμε είναι τα εξής:

```
Chi-Sq = 2.000 + 2.000 + 2.000 + 2.000 +
          0.000 + 0.500 + 0.000 + 0.500 +
          2.000 + 0.500 + 2.000 + 0.500 = 14.000
DF = 6, P-Value = 0.030
12 cells with expected counts less than 5.0
```

Το κρίσιμο επίπεδο είναι 0.03.

Παρατηρούμε ότι το πακέτο μας προειδοποιεί ότι 12, δηλαδή, όλα τα κελλιά του πίνακα συναφείας έχουν αναμενόμενη συχνότητα μικρότερη του 5.

**Λύση με το SPSS:** Κατασκευάζουμε έναν πίνακα συνάφειας όπως στην αναλυτική λύση του παραδείγματος και, στην συνέχεια, διεξάγουμε τον έλεγχο  $\chi^2$  με το SPSS κατά τον ήδη γνωστό τρόπο. Τα αποτελέσματα που παίρνουμε είναι τα εξής:

#### Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	14.00	6	.030
Likelihood Ratio	18.72	6	.005
Linear-by-Linear Association	.19	1	.662
N of Valid Cases	2		

a 12 cells (100.0%) have expected count less than 5. The minimum expected count is 2.00.

Το κρίσιμο επίπεδο είναι 0.03.

Παρατηρούμε ότι το πακέτο μας προειδοποιεί ότι όλα τα κελλιά του πίνακα συναφείας έχουν αναμενόμενη συχνότητα μικρότερη του 5.

**Λύση με το SAS:** Εργαζόμενοι κατά τα ήδη γνωστά, πληκτρολογούμε τις εξής εντολές στο παράθυρο εντολών:

```
data graph;
input rows columns counts;
cards;
1 1 0
1 2 4
1 3 4
1 4 0
2 1 2
2 2 1
2 3 2
2 4 3
3 1 4
3 2 1
3 3 0
3 4 3
;
run;
proc freq;
tables rows*columns / chisq;
weight counts;
exact chisq;
run;
```

Τα αποτελέσματα που παίρνουμε, περιέχονται στον πίνακα που ακολουθεί.

The SAS System  
TABLE OF ROWS BY COLUMNS

ROWS	COLUMNS				Total
Frequency,	1,	2,	3,	4,	
Percent ,					
Row Pct ,					
Col Pct	1,	2,	3,	4,	Total
1	0	4	4	0	8
	0.00	16.67	16.67	0.00	33.33
	0.00	50.00	50.00	0.00	
	0.00	66.67	66.67	0.00	
2	2	1	2	3	8
	8.33	4.17	8.33	12.50	33.33
	25.00	12.50	25.00	37.50	
	33.33	16.67	33.33	50.00	
3	4	1	0	3	8
	16.67	4.17	0.00	12.50	33.33
	50.00	12.50	0.00	37.50	
	66.67	16.67	0.00	50.00	
Total	6	6	6	6	24
	25.00	25.00	25.00	25.00	100.00

STATISTICS FOR TABLE OF ROWS BY COLUMNS

Statistic	DF	Value	Prob (Asymptotic)	Prob (Exact)
Chi-Square	6	14.000	0.030	0.033
Likelihood Ratio Chi-Square	6	18.729	0.005	0.015
Mantel-Haenszel Chi-Square	1	0.192	0.662	0.747
Phi Coefficient		0.764		
Contingency Coefficient		0.607		
Cramer's V		0.540		
Sample Size = 24				

WARNING: 100% of the cells have expected counts less than 5.  
(Asymptotic) Chi-Square may not be a valid test

Η τιμή της ελεγχοσυνάρτησης υπολογίσθηκε ίση με 14, ενώ το ακριβές παρατηρούμενο επίπεδο στατιστικής σημαντικότητας είναι ίσο με 0.033.

Παρατηρούμε ότι το πακέτο μας προειδοποιεί ότι όλα τα κελιά του πίνακα συναφείας έχουν αναμενόμενη συχνότητα μικρότερη του 5 και ότι η προσέγγιση της κατανομής της ελεγχοσυνάρτησης από την  $\chi^2$  ενδέχεται να μην είναι ικανοποιητική.

## 8.2 2×2 ΠΙΝΑΚΕΣ ΣΥΝΑΦΕΙΑΣ – ΜΙΑ ΕΝΔΙΑΦΕΡΟΥΣΑ ΕΙΔΙΚΗ ΠΕΡΙΠΤΩΣΗ

Οι  $r \times c$  πίνακες συναφείας που εξετάστηκαν στην προηγούμενη ενότητα, αποτελούν εν γένει μία παράθεση φυσικών αριθμών ταξινομημένων σε  $r$  γραμμές και  $c$  στήλες που, έτσι, οδηγούν σε  $rc$  κατηγορίες που εκπροσωπούνται από τα κελιά των πινάκων. Οι περιπτώσεις των πινάκων όπου  $r=2$  και  $c=2$ , δηλαδή οι  $2 \times 2$  πίνακες συναφείας, αποτελούν μία ενδιαφέρουσα ειδική περίπτωση των  $r \times c$  πινάκων συναφείας, ιδιαίτερα στην περιοχή των βιοϊατρικών προβλημάτων. Για τον λόγο αυτό εξετάζονται ως μία ξεχωριστή κατηγορία στην ενότητα αυτή.

Κατ' αναλογία με τους  $r \times c$  πίνακες συναφείας, μία εφαρμογή των  $2 \times 2$  πινάκων συναφείας προκύπτει όταν επιθυμούμε να αναλύσουμε τα στοιχεία δύο ανεξαρτήτων δειγμάτων από δύο διαφορετικούς πληθυσμούς για να εξετάσουμε κατά πόσο οι δύο πληθυσμοί εκπροσωπούνται με ίσα ή με διαφορετικά ποσοστά στοιχείων σε μία συγκεκριμένη κατηγορία. (Ειδικότερα, δύο τυχαία δείγματα επιλέγονται, ένα από κάθε πληθυσμό, για να ελεγχθεί η μηδενική υπόθεση ότι η πιθανότητα κάποιου συγκεκριμένου ενδεχομένου  $A$  είναι η ίδια για τους δύο πληθυσμούς).

Μία άλλη εφαρμογή των  $2 \times 2$  πινάκων συναφείας προκύπτει όταν  $N$  αντικείμενα (ή πρόσωπα), επιλεγέντα με τυχαίο τρόπο από κάποιο πληθυσμό, ταξινομούνται σε μία από δύο κατηγορίες πριν από την εφαρμογή ενός προγράμματος ή μίας αγωγής (ή πριν από την εμφάνιση κάποιας κατάστασης ή χαρακτηριστικού). Μετά την εφαρμογή του προγράμματος ή της αγωγής (ή μετά την εμφάνιση της κατάστασης ή χαρακτηριστικού), τα  $N$  αντικείμενα (ή πρόσωπα) επανεξετάζονται και

ταξινομούνται εκ νέου στις δύο κατηγορίες. Το ερώτημα που επιθυμούμε να εξετάσουμε διατυπώνεται ως εξής: « *Η εφαρμογή του προγράμματος ή της αγωγής (ή η εμφάνιση της κατάστασης) μεταβάλλει σημαντικά το ποσοστό των μονάδων ή προσώπων σε κάθε μία από τις δύο κατηγορίες;* ».

Ένας άλλος τρόπος για τον έλεγχο της ίδιας υπόθεσης είναι να χρησιμοποιηθούν εκτιμήσεις των ποσοστών των μονάδων του πληθυσμού στις διάφορες κατηγορίες βασισμένες σε συσχετισμένες παρατηρήσεις, δηλαδή ζεύγη παρατηρήσεων.

Οι έλεγχοι υποθέσεων που βασίζονται σε ένα και μοναδικό δείγμα ή σε δύο συσχετισμένα δείγματα (δηλαδή σε ένα δείγμα ζευγών παρατηρήσεων), είναι ιδιαίτερα σημαντικοί, κυρίως στην περιοχή των βιοϊατρικών προβλημάτων, όπου τυπικό επιδημιολογικό εργαλείο αποτελούν οι αναδρομικές έρευνες μεταξύ μιας ομάδας ατόμων που παρουσιάζουν την συγκεκριμένη διαταραχή ή πάθηση και μιας ομάδας ατόμων που δεν την παρουσιάζουν, ανεξάρτητα από εάν είχαν εκτεθεί στον παράγοντα κινδύνου. (Στην βιοϊατρική ορολογία, τα άτομα που παρουσιάζουν την πάθηση ονομάζονται *περιπτώσεις (cases)*, ενώ τα μη ασθενούντα άτομα της ομάδας σύγκρισης ονομάζονται *μάρτυρες (controls)*). Ας υποθέσουμε ότι ένας ερευνητής επιθυμεί να εξετάσει κατά πόσο υπάρχει συσχέτιση μεταξύ ενός χαρακτηριστικού A (για παράδειγμα, ενός παράγοντα κινδύνου, όπως η χρήση αντισυλληπτικών χαπιών) και ενός χαρακτηριστικού B (για παράδειγμα, της εμφάνισης ενός προβλήματος ή μιας διαταραχής της υγείας, όπως η θρομβοεμβολή). Επειδή η εμφάνιση νέων περιστατικών της πάθησης ενδέχεται να είναι χαμηλή, θα χρειαζόταν μία εξαιρετικά μεγάλη αναδρομική μελέτη για τον εντοπισμό επαρκούς αριθμού περιπτώσεων. Μία στρατηγική θα ήταν να ξεκινήσει κανείς με τον αριθμό των περιπτώσεων



που έχουν παρατηρηθεί. Τότε, το πρόβλημα θα ήταν να βρεθούν κατάλληλοι μάρτυρες (μονάδες σύγκρισης) για τις περιπτώσεις. Στην περίπτωση των μελετών που στηρίζονται σε συσχετισμένα δείγματα, ορίζεται για κάθε περίπτωση ένας μάρτυρας (μία μονάδα σύγκρισης). Ο μάρτυρας, ο οποίος δεν έχει την πάθηση, πρέπει να επιλεγεί ώστε να «ταυτίζεται» με την περίπτωση ως προς όλα εκείνα τα χαρακτηριστικά στοιχεία που κρίνονται ότι μπορούν να επηρεάζουν ή να συνδέονται με την εμφάνιση της πάθησης, ενδεχομένως με εξαίρεση αυτό της έκθεσης στον παράγοντα κινδύνου (π.χ., ηλικία, βάρος, φύλο, άλλες παθήσεις κ.λ.π.). Ο έλεγχος που χρησιμοποιείται για την ανάλυση δεδομένων πινάκων τέτοιας μορφής είναι γνωστός ως έλεγχος McNemar.

Στην περίπτωση που τα δεδομένα προέρχονται από ένα και μοναδικό δείγμα και είναι ταξινομημένα σε πίνακα συναφείας με σταθερά αθροίσματα γραμμών και στηλών, ο χρησιμοποιούμενος έλεγχος είναι γνωστός ως έλεγχος Fisher (Fisher's exact test).

### **8.2.1 Έλεγχος $\chi^2$ για Διαφορές σε Πιθανότητες ή σε Αναλογίες (περίπτωση ανεξαρτήτων δειγμάτων)**

Ένα τυχαίο δείγμα  $n_1$  παρατηρήσεων επιλέγεται από ένα πληθυσμό και κάθε παρατήρηση ταξινομείται σε μία από δύο κατηγορίες (κατηγορία 1 ή κατηγορία 2). Ένα δεύτερο τυχαίο δείγμα  $n_2$  παρατηρήσεων επιλέγεται από έναν άλλο πληθυσμό (ή από τον πρώτο πληθυσμό μετά από την εφαρμογή κάποιας αγωγής) και κάθε παρατήρησή του επίσης ταξινομείται στις κατηγορίες 1 και 2. Προκύπτει, επομένως, ο εξής  $2 \times 2$  πίνακας συναφείας

	Κατηγορία 1	Κατηγορία 2	Σύνολο
Πληθυσμός 1	$O_{11}$	$O_{12}$	$n_1$
Πληθυσμός 2	$O_{21}$	$O_{22}$	$n_2$
Σύνολο	$C_1$	$C_2$	$N = n_1 + n_2$

Εστω ότι η πιθανότητα με την οποία ένα τυχαία επιλεγόμενο στοιχείο από τον πληθυσμό 1 ανήκει στην κατηγορία 1 είναι  $p_1$  και ότι η αντίστοιχη πιθανότητα για τον πληθυσμό 2 είναι  $p_2$ . Οι υποθέσεις που ενδιαφερόμαστε να ελέγξουμε έχουν μία από τις εξής μορφές:

**A. (Αμφίπλευρος Έλεγχος)**

$$H_0: p_1 = p_2$$

$$H_0: p_1 \neq p_2$$

**B. (Μονόπλευρος Έλεγχος)**

$$H_0: p_1 \leq p_2$$

$$H_0: p_1 > p_2$$

**Γ. (Μονόπλευρος Έλεγχος)**

$$H_0: p_1 \geq p_2$$

$$H_0: p_1 < p_2$$

Όπως και στην περίπτωση των  $r \times c$  πινάκων, αν η  $H_0$  είναι αληθής, ο αριθμός των στοιχείων που περιμένουμε να παρατηρήσουμε στο κελλί  $(i, j)$  του πίνακα είναι

$$E_{ij} = n_i C_j / N.$$

Και πάλι, ως μέτρο της εγγύτητας μεταξύ παρατηρούμενων και αναμενόμενων συχνοτήτων προκειμένου να εξετασθεί αν τα δεδομένα παρέχουν ενδείξεις υπέρ της μηδενικής υπόθεσης, χρησιμοποιείται η στατιστική συνάρτηση

$$T = \sum_{i=1}^2 \sum_{j=1}^2 \frac{(O_{ij} - E_{ij})^2}{E_{ij}},$$

η οποία μπορεί να γραφεί με την μορφή

$$T = \frac{N(O_{11}O_{22} - O_{12}O_{21})^2}{n_1 n_2 C_1 C_2}.$$

Για τους μονόπλευρους ελέγχους χρησιμοποιείται συνήθως η στατιστική συνάρτηση

$$\begin{aligned} T_1 &= \sqrt{T} \\ &= \frac{\sqrt{N}(O_{11}O_{22} - O_{12}O_{21})}{\sqrt{n_1 n_2 C_1 C_2}}. \end{aligned}$$

Η ακριβής μορφή της κατανομής της στατιστικής συνάρτησης  $T$ , αντίστοιχα της  $T_1$ , δεν είναι εύκολο να προσδιορισθεί λόγω του μεγάλου αριθμού των διαφορετικών συνδυασμών των δυνατών τιμών για τις παρατηρούμενες συχνότητες  $O_{11}$ ,  $O_{12}$ ,  $O_{21}$  και  $O_{22}$ . (Η διαδικασία προσδιορισμού της κατανομής της στατιστικής συνάρτησης  $T$  περιγράφεται στην συνέχεια). Χρησιμοποιούνται, επομένως, προσεγγίσεις των κατανομών αυτών. Έτσι, σύμφωνα με τα όσα ισχύουν για την περίπτωση των  $r \times c$  πινάκων, η κατανομή της στατιστικής συνάρτησης  $T$  μπορεί να προσεγγισθεί από την κατανομή  $\chi^2$  με 1 βαθμό ελευθερίας. Δηλαδή,

$$T \sim \chi_1^2.$$

Όσο αφορά την κατανομή της στατιστικής συνάρτησης  $T_1$ , αποδεικνύεται ότι μπορεί να προσεγγισθεί από την τυποποιημένη κανονική κατανομή, δηλαδή

$$T_1 \sim N(0, 1).$$

Επομένως, ο κανόνας απόφασης έχει την εξής μορφή:

#### **A. (Αμφίπλευρος Έλεγχος)**

Η μηδενική υπόθεση  $H_0$  απορρίπτεται σε επίπεδο σημαντικότητας περίπου ίσο με  $\alpha$  αν  $T > \chi_{1, 1-\alpha}^2$ , όπου  $\chi_{1, 1-\alpha}^2$  συμβολίζει το  $(1-\alpha)$ -ποσοστιαίο σημείο της κατανομής  $\chi_1^2$ .

#### **B. (Μονόπλευρος Έλεγχος)**

Η μηδενική υπόθεση  $H_0$  απορρίπτεται σε επίπεδο σημαντικότητας περίπου ίσο με  $\alpha$  αν  $T_1 > z_{1-\alpha}$ , όπου  $z_{1-\alpha}$  είναι το  $(1-\alpha)$ -ποσοστιαίο σημείο της τυποποιημένης κανονικής κατανομής.

#### **Γ. (Μονόπλευρος Έλεγχος)**

Η μηδενική υπόθεση  $H_0$  απορρίπτεται σε επίπεδο σημαντικότητας περίπου ίσο με  $\alpha$  αν  $T_1 < z_\alpha$ , όπου  $z_\alpha$  είναι το  $\alpha$ -ποσοστιαίο σημείο της τυποποιημένης κανονικής κατανομής.

**Παράδειγμα 8.2.1:** Σε μία στρατιωτική σχολή τοποθετήθηκε ένα νέο σύστημα φωτισμού στους θαλάμους των μαθητευομένων. Προκειμένου να ελεγχθεί ο ισχυρισμός ότι το νέο σύστημα φωτισμού προκαλούσε κόπωση

στα μάτια που οδηγούσε σε κακή όραση, επελέγησαν δύο ανεξάρτητα τυχαία δείγματα 825 και 816 αποφοιτησάντων από την σχολή, την χρονιά που προηγήθηκε της εγκατάστασης του νέου συστήματος φωτισμού (πρώτο δείγμα) και την χρονιά των πρώτων αποφοίτων μετά την εγκατάσταση του νέου φωτισμού (δείγμα 2). Τα αποτελέσματα συνοψίζονται στον πίνακα που ακολουθεί.

	Καλή όραση	Κακή όραση	
<b>Παλιός Φωτισμός</b>	$O_{11} = 714$	$O_{12} = 111$	$n_1 = 825$
<b>Νέος Φωτισμός</b>	$O_{21} = 662$	$O_{22} = 154$	$n_2 = 816$
<b>Σύνολα</b>	1376	265	$N = 1641$

Να ελεγχθεί ο ισχυρισμός σε επίπεδο σημαντικότητας περίπου ίσο με  $\alpha=0.05$ .

**Λύση:** Έστω  $p_1$  η πιθανότητα ότι ένας τυχαία επιλεγόμενος απόφοιτος είχε καλή όραση κάτω από το παλιό σύστημα φωτισμού και  $p_2$  η αντίστοιχη πιθανότητα κάτω από το νέο σύστημα φωτισμού. Οι υποθέσεις που έχει έννοια να ελεγχθούν μπορούν να γραφούν με την μορφή:

$$H_0: p_1 \leq p_2$$

$$H_1: p_1 > p_2.$$

Η κατάλληλη στατιστική συνάρτηση για τον έλεγχο των παραπάνω υποθέσεων είναι η  $T_1$  και είναι προφανές ότι η κρίσιμη περιοχή μεγέθους 0.05 του ελέγχου ορίζεται από την ανισότητα

$$T_1 > z_{0.95} = 1.645 .$$

Η παρατηρούμενη τιμή της  $T_1$  είναι

$$\tau_1 = \frac{\sqrt{1641} [(714)(154) - (111)(662)]}{\sqrt{(825)(816)(1376)(265)}} = 2.982 .$$

Η τιμή αυτή βρίσκεται μέσα στην κρίσιμη περιοχή μεγέθους 0.05 και, επομένως, η μηδενική υπόθεση απορρίπτεται σε επίπεδο σημαντικότητας 5%.

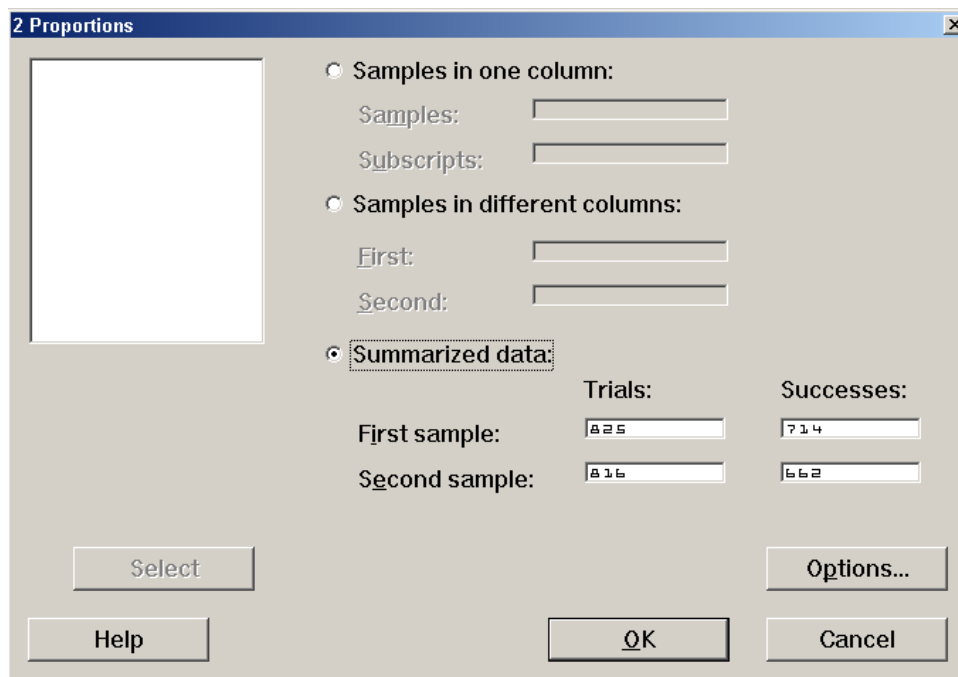
Το κρίσιμο επίπεδο του ελέγχου είναι

$$\hat{\alpha} = P(T_1 \geq 2.982 | H_0) = 1 - \Phi(2.982) = 0.002 .$$

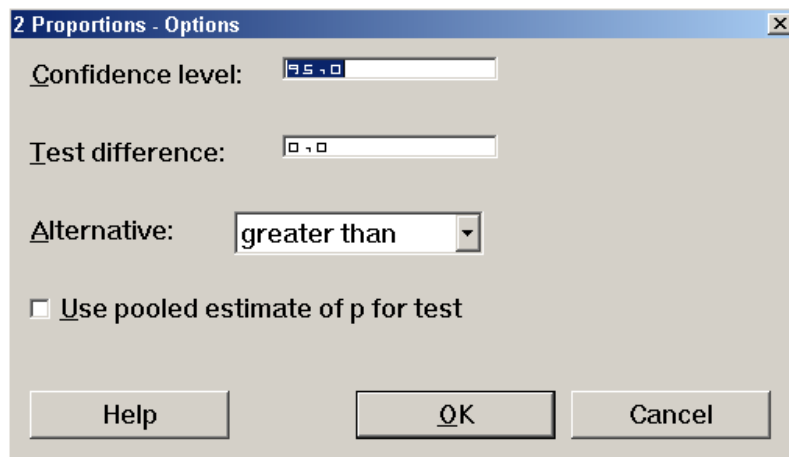
Είναι εύλογο, επομένως, να συμπεράνει κανείς ότι οι πληθυσμοί των δύο τάξεων αποφοίτων διαφέρουν στατιστικά σημαντικά όσο αφορά τα ποσοστά αυτών που έχουν κακή όραση σύμφωνα με τον ισχυρισμό. Δηλαδή, ο πληθυσμός των αποφοίτων πριν από την εγκατάσταση του νέου συστήματος φωτισμού εμφανίζεται να έχει καλύτερη όραση από αυτή του πληθυσμού των αποφοίτων με το νέο σύστημα φωτισμού.

**Παρατήρηση:** Θα πρέπει να σημειωθεί ότι το ερώτημα «κατά πόσο η κακή όραση είναι αποτέλεσμα της τοποθέτησης του νέου συστήματος φωτισμού» δεν έχει απαντηθεί. Όμως, έχει τεκμηριωθεί μία σχέση μεταξύ της χαμηλής όρασης και του νέου φωτισμού.

**Λύση με το MINITAB:** Το MINITAB παρέχει τον παραπάνω έλεγχο μέσω της επιλογής **Stat, Basic Statistics, 2 Proportions**, η οποία οδηγεί στο εξής πλαίσιο διαλόγου:



Παρατηρούμε ότι το πρόγραμμα παρέχει διάφορες επιλογές όσο αφορά τον τρόπο εισαγωγής των δεδομένων. Μεγαλύτερη ευκολία στους υπολογισμούς προσφέρει η επιλογή **Summarized data**, με την οποία πρέπει να δηλώσουμε τον αριθμό δοκιμών (πεδίο **Trials**) και τον αριθμό επιτυχιών (πεδίο **Successes**) για κάθε ένα από τους πληθυσμούς. Ως πρώτο δείγμα (πεδίο **First sample**), δηλώνουμε το δείγμα των αποφοιτησάντων πριν εγκατασταθεί το νέο σύστημα φωτισμού και ως δεύτερο (πεδίο **Second sample**), δηλώνουμε το δείγμα των αποφοιτησάντων μετά την εγκατάσταση του νέου συστήματος φωτισμού. Στην συνέχεια, πιέζοντας **Options**, οδηγούμεθα στο εξής πλαίσιο διαλόγου:



Στο πλαίσιο αυτό, δηλώνουμε το επίπεδο σημαντικότητας (πεδίο **Confidence level**), την τιμή της διαφοράς  $p_1 - p_2$  που υπονοεί η  $H_0$  (πεδίο **Test difference**) και την εναλλακτική υπόθεση. Η εναλλακτική που απαιτεί το MINITAB πρέπει να δηλώνει τι είναι η  $p_1$  (που αναλογεί στο πρώτο δείγμα) σε σχέση με την  $p_2$  (που αναλογεί στο δεύτερο). Εμείς εδώ, επιλέγουμε **greater than**. Πιέζοντας **OK** και πάλι **OK** στο επανεμφανιζόμενο αρχικό πλαίσιο, και παίρνουμε το εξής αποτέλεσμα:

Test and Confidence Interval for Two Proportions

Sample	X	N	Sample p
1	714	825	0.865455
2	662	816	0.811275

Estimate for  $p(1) - p(2)$ : 0.0541800  
 95% CI for  $p(1) - p(2)$ : (0.0186416; 0.0897184)  
 Test for  $p(1) - p(2) = 0$  (vs  $> 0$ ): Z = 2.99 P-Value = 0.001

Παρατηρούμε ότι η τιμή της ελεγχοσυνάρτησης  $T_1$  είναι 2.99 (**Z**) και το παρατηρούμενο επίπεδο σημαντικότητας είναι μόλις 0.001.



**Λύση με το SPSS:** Το SPSS δεν παρέχει μονόπλευρους ελέγχους της παραπάνω μορφής.

**Λύση με το SAS:** Το SAS δεν παρέχει μονόπλευρους ελέγχους της παραπάνω μορφής.

### Προσδιορισμός της ακριβούς κατανομής της στατιστικής συνάρτησης T

Η ακριβής μορφή της κατανομής της στατιστικής συνάρτησης T όταν η μηδενική υπόθεση  $H_0: p_1 = p_2 = p$  είναι σωστή, μπορεί να προσδιορισθεί ως εξής:

Για το δείγμα που προέρχεται από τον πληθυσμό 1, η πιθανότητα ότι ακριβώς  $x_1$  στοιχεία ανήκουν στην κατηγορία 1 και  $n_1 - x_1$  στοιχεία ανήκουν στην κατηγορία 2 δίνεται από την σχέση

$$P \left( \begin{array}{cc} \text{Κατηγορία 1} & \text{Κατηγορία 2} \\ \text{Πληθυσμός 1} & [x_1 \quad n_1 - x_1] \end{array} \right) = \binom{n_1}{x_1} p^{x_1} (1-p)^{n_1-x_1} .$$

Με όμοιο τρόπο, η πιθανότητα ότι το δείγμα που προέρχεται από τον πληθυσμό 2 θα έχει ακριβώς  $x_2$  στοιχεία στην κατηγορία 1 και  $n_2 - x_2$  στοιχεία στην κατηγορία 2 δίνεται από την σχέση:

$$P \left( \begin{array}{cc} \text{Κατηγορία 1} & \text{Κατηγορία 2} \\ \text{Πληθυσμός 2} & [x_2 \quad n_2 - x_2] \end{array} \right) = \binom{n_2}{x_2} p^{x_2} (1-p)^{n_2-x_2} .$$

Επειδή τα δύο δείγματα είναι ανεξάρτητα, η από κοινού πιθανότητα των δύο ενδεχομένων προκύπτει ίση με

$$P \left( \begin{array}{cc} \text{Κατηγορία 1} & \text{Κατηγορία 2} \\ \text{Πληθυσμός 1} & [x_1 \quad n_1 - x_1] \\ \text{Πληθυσμός 2} & [x_2 \quad n_2 - x_2] \end{array} \right) = \binom{n_1}{x_1} \binom{n_2}{x_2} p^{x_1+x_2} (1-p)^{N-x_1-x_2}$$

Στην απλή περίπτωση όπου  $n_1=2$  και  $n_2=2$ , υπάρχουν εννέα διαφορετικά σημεία στον δειγματικό χώρο, αντιστοιχούντα στις εννέα δυνατές ταξινομήσεις σε πίνακες συναφείας οι οποίες παρατίθενται στην συνέχεια.

**Πίνακας Δυνατών ταξινομήσεων των στοιχείων δύο ανεξάρτητων δειγμάτων μεγέθους  $n_1=n_2=2$  σε  $2 \times 2$  πίνακα συναφείας**

Δυνατές Ταξινομήσεις	Πιθανότητα αν η $H_0$ είναι αληθής		T	
	$p = 1/2$	$p = 1$		
$\begin{bmatrix} 2 & 0 \\ 2 & 0 \end{bmatrix}$	$p^4$	1/16	1	Δεν ορίζεται
$\begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}$	$2p^3(1-p)$	1/8	0	4/3
$\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$	$p^2(1-p)^2$	1/16	0	4
$\begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}$	$2p^3(1-p)$	1/8	0	4/3
$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$	$4p^2(1-p)^2$	1/4	0	0
$\begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$	$2p(1-p)^3$	1/8	0	4/3
$\begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}$	$p^2(1-p)^2$	1/16	0	4
$\begin{bmatrix} 0 & 2 \\ 1 & 1 \end{bmatrix}$	$2p(1-p)^3$	1/8	0	4/3
$\begin{bmatrix} 0 & 2 \\ 0 & 2 \end{bmatrix}$	$(1-p)^4$	1/16	0	Δεν ορίζεται

Οι τιμές της στατιστικής συνάρτησης  $T$ , οι οποίες δεν ορίζονται αντιστοιχούν στις περιπτώσεις της απροσδιόριστης μορφής  $0/0$ . Επειδή, όμως, τα δύο ενδεχόμενα τα οποία οδηγούν σε μη οριζόμενες τιμές της  $T$  είναι ενδεικτικά του ότι η  $H_0$  είναι αληθής, όπως ακριβώς και το πέμπτο ενδεχόμενο είναι ενδεικτικό ότι η  $H_0$  είναι αληθής, η τιμή της  $T$  μπορεί αυθαίρετα να ορισθεί ίση με  $0$  για το πρώτο και για το τελευταίο ενδεχόμενο, κατ' αναλογία με την περίπτωση που αντιστοιχεί στο πέμπτο ενδεχόμενο. Τότε, η στατιστική συνάρτηση  $T$  έχει την εξής κατανομή πιθανότητας:

$p = 1/2$	$p = 1$
$P(T = 0) = 3/8$	$P(T = 0) = 1$
$P(T = 4/3) = 1/2$	
$P(T = 4) = 1/8$	

Με τον ίδιο τρόπο, για οποιαδήποτε μεγέθη δειγμάτων  $n_1$  και  $n_2$ , η ακριβής μορφή της κατανομής πιθανότητας της στατιστικής συνάρτησης  $T$  μπορεί να προσδιορισθεί ορίζοντας κατάλληλα τις «μη οριζόμενες» τιμές της  $T$ .

**Παρατήρηση:** Η συνάρτηση πιθανότητας της στατιστικής συνάρτησης  $T$ , όμως, δεν ορίζεται μονοσήμαντα ακόμη και στην περίπτωση που η  $H_0$  υποτίθεται ότι είναι αληθής, αλλά εξαρτάται από την παράμετρο  $p$ . Άρα, η μηδενική υπόθεση του παραπάνω ελέγχου είναι σύνθετη. Μπορεί να αποδειχθεί, ότι το μέγεθος της κρίσιμης περιοχής  $\alpha$  μεγιστοποιείται όταν

$p=1/2$ . Επομένως, το επίπεδο σημαντικότητας  $\alpha$  της κατανομής της στατιστικής συνάρτησης  $T$  μπορεί να προσδιορισθεί στην απλή περίπτωση που έχουμε θεωρήσει παραπάνω θέτοντας  $p=1/2$ . Αν η κρίσιμη περιοχή αντιστοιχεί στην μέγιστη τιμή της στατιστικής συνάρτησης  $T$  (δηλαδή,  $\tau=4$ ), τότε  $\alpha=0.125$ .

Όπως αναφέρθηκε στα προηγούμενα, η ασυμπτωτική κατανομή της στατιστικής συνάρτησης  $T$  είναι η  $\chi^2$  με 1 βαθμό ελευθερίας. (Για την απόδειξη του αποτελέσματος αυτού ο ενδιαφερόμενος αναγνώστης παραπέμπεται στο βιβλίο του Cramér (1946)).

**Παρατήρηση:** Προκειμένου να αντισταθμισθεί εν μέρει η ανακρίβεια που εισάγεται από την χρήση μίας συνεχούς συνάρτησης κατανομής (της  $\chi^2$ ) για την προσέγγιση της διακριτής συνάρτησης κατανομής της στατιστικής συνάρτησης  $T$ , ο Yates (1934) πρότεινε την λεγόμενη *διόρθωση συνεχείας* (*correction for continuity*). Η κατά Yates τροποποιημένη μορφή της στατιστικής συνάρτησης  $T$  έχει την μορφή

$$T' = \frac{N[|O_{11}O_{22} - O_{12}O_{21}| - (N/2)]^2}{n_1 n_2 C_1 C_2}.$$

Η διόρθωση συνίσταται στην μείωση της απόλυτης τιμής της διαφοράς  $O_{11}O_{22} - O_{12}O_{21}$  κατά την ποσότητα  $N/2$  πριν από την ύψωσή της στο τετράγωνο. Όμως, πολλοί στατιστικοί (Pearson (1947), Plackett (1964), Grizzle (1967), Conover (1974)) θεωρούν ότι η διόρθωση αυτή οδηγεί σε μία πολύ συντηρητική τιμή για το μέτρο της εγγύτητας των παρατηρούμενων και αναμενόμενων τιμών και αντιτίθενται στην εφαρμογή της γιατί οι τιμές της στατιστικής συνάρτησης  $T$  βρίσκονται εγγύτερα σ'

αυτές μιας  $\chi^2$  μεταβλητής με 1 βαθμό ελευθερίας από τις τιμές της στατιστικής συνάρτησης  $T'$ .

### 8.2.2 Ο $\chi^2$ Έλεγχος Ανεξαρτησίας (Περίπτωση Ενός και Μοναδικού Δείγματος)

Όπως και στην περίπτωση των  $r \times c$  πινάκων, οι υποθέσεις που θέλουμε να ελέγξουμε αναφέρονται στην ανεξαρτησία των ενδεχομένων *{μία παρατήρηση ανήκει στην  $i$  γραμμή}* και *{η ίδια παρατήρηση ανήκει στην  $j$  στήλη}*.

Από τον ορισμό της ανεξαρτησίας ενδεχομένων, οι προς έλεγχο υποθέσεις μπορούν, προφανώς, να διατυπωθούν ως εξής:

$$H_0: p_{ij} = p_i \cdot p_j, \quad i = 1, 2 \quad j = 1, 2$$

$$H_1: p_{ij} \neq p_i \cdot p_j, \quad \text{για κάποιες τιμές των } i, j,$$

όπου  $p_i$  είναι η πιθανότητα εμφάνισης του πρώτου ενδεχομένου,  $p_j$  είναι η πιθανότητα εμφάνισης του δεύτερου ενδεχομένου και  $p_{ij}$  είναι η από κοινού πιθανότητα των δύο ενδεχομένων. Επομένως, εάν η μηδενική υπόθεση είναι αληθής, η αναμενόμενη συχνότητα στο  $(i, j)$  κελλί προκύπτει από την σχέση

$$E_{ij} = N(R_i/N) \cdot (C_j/N) = R_i C_j / N.$$

Η στατιστική συνάρτηση για τον έλεγχο των παραπάνω υποθέσεων ορίζεται όπως και προηγουμένως από την σχέση

$$\begin{aligned}
T &= \sum_{i=1}^2 \sum_{j=1}^2 \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \\
&= \sum_{i=1}^2 \sum_{j=1}^2 \frac{O_{ij}^2}{E_{ij}} - N \\
&= \frac{N(O_{11}O_{22} - O_{12}O_{21})^2}{n_1 n_2 C_1 C_2},
\end{aligned}$$

η οποία ταυτίζεται με την στατιστική συνάρτηση  $T$  της προηγούμενης περίπτωσης.

Όπως και στην προηγούμενη περίπτωση, η ακριβής μορφή της κατανομής της στατιστικής συνάρτησης  $T$  δεν είναι εύκολο να προσδιορισθεί. (Η διαδικασία προσδιορισμού της κατανομής της στατιστικής συνάρτησης  $T$  περιγράφεται στην συνέχεια). Για τον λόγο αυτό, συνήθως χρησιμοποιείται η προσέγγισή της από την κατανομή  $\chi_1^2$ . Δηλαδή,

$$T \sim \chi_1^2.$$

Η κρίσιμη περιοχή του ελέγχου των παραπάνω υποθέσεων, επομένως, βρίσκεται στην δεξιά ουρά της κατανομής  $\chi^2$  με 1 βαθμό ελευθερίας και προσδιορίζεται από την ανισότητα

$$T > \chi_{1, 1-\alpha}^2.$$

**Παράδειγμα 8.2.2:** 160 πιλότοι, υποψήφιοι για κάποιες θέσεις σε μία αεροπορική εταιρεία, υποβάλλονται σε ένα ψυχολογικό τεστ σύμφωνα με το οποίο ταξινομούνται ως εσωστρεφείς ή εξωστρεφείς και σε ένα τεστ δεξιότητας στο οποίο βαθμολογούνται ως επιτυχόντες ή αποτυχόντες. Τα στοιχεία του πίνακα που ακολουθεί συνοψίζουν τ' αποτελέσματα.

### Ψυχολογικό τεστ

Τεστ δεξιότητας	Εσωστρεφείς	Εξωστρεφείς	Σύνολα
Επιτυχόντες	13	73	86
Αποτυχόντες	17	57	74
Σύνολα	30	130	160

Θα ήταν εύλογο να θεωρήσει κανείς ότι τα δεδομένα υπαινίσσονται ύπαρξη συσχέτισης μεταξύ του τύπου της προσωπικότητας και των δεξιοτήτων των πιλότων;

**Λύση:** Η προς έλεγχο μηδενική υπόθεση είναι η

$H_0$ : Η δεξιότητα ενός πιλότου δεν σχετίζεται με τον τύπο της προσωπικότητάς του.

Η τιμή της στατιστικής συνάρτησης για τα δεδομένα του παραπάνω πίνακα προκύπτει ίση με

$$\tau = \frac{160((13)(57) - (73)(17))^2}{(86)(74)(30)(130)} = 1.61.$$

Το κρίσιμο επίπεδο του ελέγχου είναι ίσο με

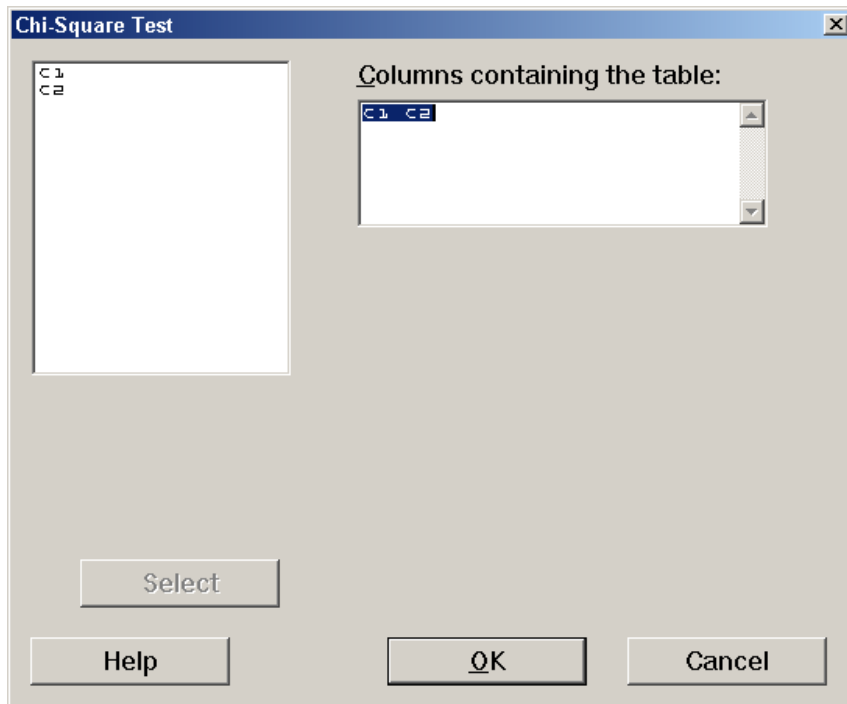
$$\hat{\alpha} = P(T \geq 1.61 | H_0) = P(\chi_1^2 \geq 1.61) \cong 1 - 0.78 = 0.22.$$

Επομένως, η υπόθεση ότι η δεξιότητα ενός πιλότου δεν σχετίζεται με τον τύπο της προσωπικότητάς του μπορεί να θεωρηθεί μία εύλογη υπόθεση.

**Λύση με το MINITAB:** Στο MINITAB, ο έλεγχος  $\chi^2$  απαιτεί να καταχωρισθεί κάθε στήλη του πίνακα συναφείας σε μία μεταβλητή. Για το συγκεκριμένο παράδειγμα, το παράθυρο δεδομένων είναι:

	C1	C2
	13	73
	17	57

Επιλέγοντας **Stat, Tables, Chi-Square Test**, προκύπτει το εξής πλαίσιο διαλόγου:



Στο πεδίο **Columns containing the table**, δηλώνουμε τις μεταβλητές που περιέχουν τις στήλες του πίνακα συναφείας και, στην συνέχεια, πιέζοντας **OK**, παίρνουμε τα εξής αποτελέσματα:



### Chi-Square Test

Expected counts are printed below observed counts

	C1	C2	Total
1	13 16.12	73 69.87	86
2	17 13.88	57 60.13	74
Total	30	130	160

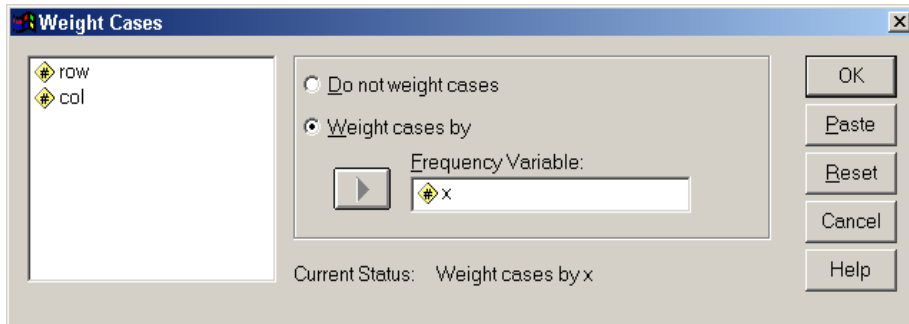
$$\text{Chi-Sq} = 0.606 + 0.140 + 0.704 + 0.162 = 1.612$$
$$\text{DF} = 1, \text{ P-Value} = 0.204$$

Το πρόγραμμα μας δίνει τον πίνακα συναφείας με τις παρατηρούμενες και τις αναμενόμενες (κάτω από την υπόθεση της ανεξαρτησίας) τιμές στα κελιά του, καθώς και την τιμή της ελεγχοσυνάρτησης T και την ασυμπτωτική τιμή του κρίσιμου επιπέδου.

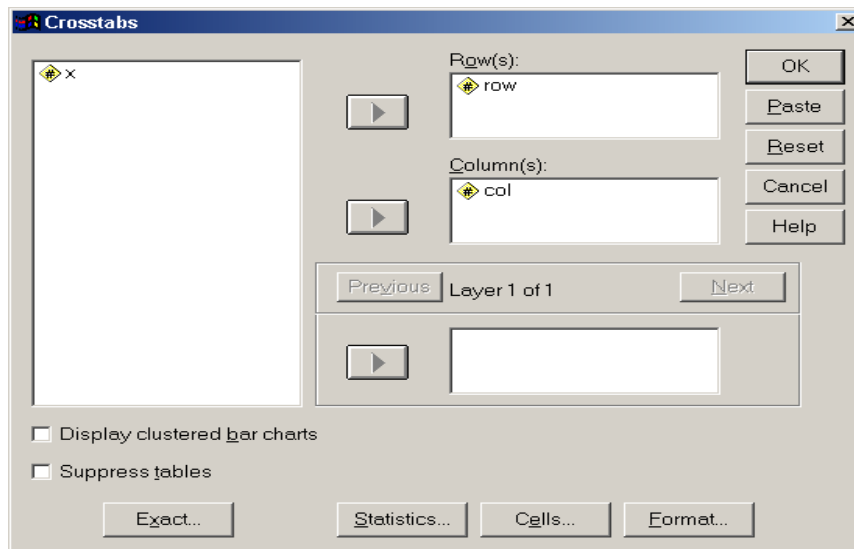
**Λύση με το SPSS:** Για την διεξαγωγή του ελέγχου ανεξαρτησίας  $\chi^2$  με το SPSS, τα δεδομένα πρέπει να εισαχθούν σε ειδική μορφή. Συγκεκριμένα, σε μία στήλη (με όνομα **x**), εισάγουμε τις τιμές των κελιών του 2×2 πίνακα δηλώνοντας ταυτόχρονα, σε δύο άλλες στήλες (έστω **row** και **col**), τον αριθμό της γραμμής και τον αριθμό της στήλης του πίνακα συναφείας στις οποίες ανήκει κάθε ένα από αυτά τα κελιά. Για το δικό μας παράδειγμα, το παράθυρο δεδομένων θα δείχνει ως εξής:

	row	col	x
1	1	1	13
2	1	2	73
3	2	1	17
4	2	2	57

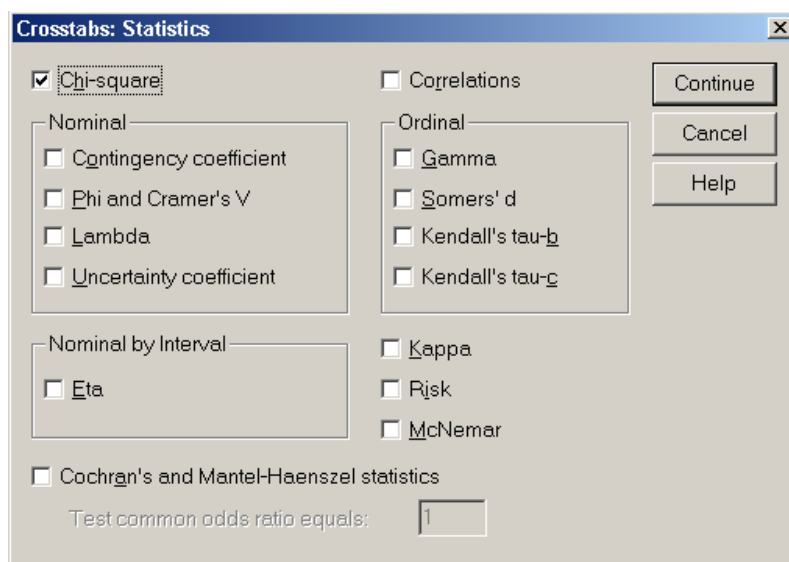
Στην συνέχεια, επιλέγοντας **Data, Weight Cases**, προκύπτει το εξής πλαίσιο διαλόγου:



Επιλέγουμε **Weight cases by** και δηλώνουμε το όνομα της μεταβλητής που περιέχει τις τιμές των κελιών (**x**). Πιέζουμε **OK** και, εν συνεχεία, επιλέγουμε **Analyze, Descriptive Statistics, Crosstabs** οδηγούμενοι στο εξής πλαίσιο:



Στα πεδία **Row(s)** και **Column(s)**, δηλώνουμε τις μεταβλητές που δείχνουν τις γραμμές και τις στήλες αντίστοιχα, στις οποίες ανήκουν οι παρατηρήσεις (**row** και **col**) όπως στην παραπάνω εικόνα. (Το πρόγραμμα δίνει στον χρήστη την επιλογή να ζητήσει την μη εμφάνιση του πίνακα συναφείας στην οθόνη, επιλέγοντας **Suppress tables**). Πιέζοντας **Statistics**, εμφανίζεται το πλαίσιο:



Στο πλαίσιο αυτό, επιλέγουμε **Chi-square**. Πιέζουμε **OK** και ξανά **OK** στο επανεμφανιζόμενο αρχικό πλαίσιο και παίρνουμε τα εξής αποτελέσματα:

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square <sup>a</sup>	1.612 <sup>b</sup>	1	.204	.228	.143	
Continuity Correction	1.137	1	.286			
Likelihood Ratio	1.608	1	.205	.228	.143	

Fisher's Exact Test				.228	.143	
Linear-by-Linear Association	1.602 <sup>c</sup>	1	.206	.228	.143	.073
N of Valid Cases	160					

a Computed only for a 2x2 table

b 0 cells (.0%) have expected count less than 5. The minimum expected count is 13.88.

c The standardized statistic is -1.266.

Η τιμή  $\tau$  της ελεγχουσυνάρτησης  $T$  που μας ενδιαφέρει, είναι 1.612 (**Pearson Chi-square**) με αντίστοιχο παρατηρούμενο επίπεδο σημαντικότητας 0.204 (πεδίο **Asymp. Sig. (2-sided)**).

Στον παραπάνω πίνακα, παρατηρούμε ότι το SPSS δίνει επίσης την τιμή του μονόπλευρου κρίσιμου επιπέδου χωρίς όμως να δίνει και το πρόσημο της ελεγχουσυνάρτησης  $T_1$ . Επομένως, στην περίπτωση μονόπλευρων ελέγχων (όπως στο παράδειγμα 8.2.1), απαιτείται προσοχή.

**Λύση με το SAS:** Στο παράθυρο εντολών εισάγουμε τις εξής εντολές:

```
data pilots;
input result psych counts;
cards;
1 1 13
1 2 73
2 1 17
2 2 57
;
run;
proc freq;
tables result*psych / chisq;
weight counts;
exact chisq;
run;
```

Στις εντολές εισαγωγής των δεδομένων, η μεταβλητή **counts** χρησιμοποιείται ως στάθμιση. Για παράδειγμα, αντί να καταχωρίσουμε 13 φορές την τιμή του κελιού των εσωτερικών επιτυχόντων, δηλαδή τον κωδικό **1 1** του οποίου το πρώτο ψηφίο δηλώνει την γραμμή του πίνακα διπλής εισόδου, και το δεύτερο την στήλη του κελιού αυτού, εισάγουμε την τιμή αυτή (δηλαδή, τον κωδικό) μόνο μια φορά, δηλώνοντας ταυτόχρονα

την αντίστοιχη συχνότητα (την ονομάσαμε **counts**). Στην συνέχεια, στην διαδικασία του SAS **proc freq**, η οποία παρέχει αναλύσεις πινάκων συναφείας, δηλώνουμε **weight counts** που χρησιμοποιείται για την στάθμιση των δεδομένων. Η εκτέλεση των εντολών αυτών δίνει τον παρακάτω πίνακα αποτελεσμάτων.

TABLE OF RESULT BY PSYCH

RESULT	PSYCH		Total
Frequency,	1,	2,	
Percent			
Row Pct			
Col Pct			
1	13	73	86
	8.13	45.63	53.75
	15.12	84.88	
	43.33	56.15	
2	17	57	74
	10.63	35.63	46.25
	22.97	77.03	
	56.67	43.85	
Total	30	130	160
	18.75	81.25	100.00

STATISTICS FOR TABLE OF RESULT BY PSYCH

Statistic	DF	Value	Prob (Asymptotic)	Prob (Exact)
Chi-Square	1	1.612	0.204	0.228
Likelihood Ratio Chi-Square	1	1.608	0.205	0.228
Continuity Adj. Chi-Square	1	1.137	0.286	
Mantel-Haenszel Chi-Square	1	1.602	0.206	0.228
Fisher's Exact Test (Left)				0.143
(Right)				0.929
(2-Tail)				0.228
Phi Coefficient		-0.100		
Contingency Coefficient		0.100		
Cramer's V		-0.100		
Sample Size =	160			

Το πεδίο **Chi-Square** περιέχει την τιμή της ελεγχοσυνάρτησης, ενώ τα πεδία **Prob(Asymptotic)** και **Prob, (Exact)** περιέχουν την ασυμπτωτική και την ακριβή τιμή του κρίσιμου επιπέδου αντίστοιχα.

### Προσδιορισμός της ακριβούς κατανομής της στατιστικής συνάρτησης T

Η ακριβής μορφή της κατανομής της στατιστικής συνάρτησης T μπορεί να προσδιορισθεί με τρόπο ανάλογο αυτού της προηγούμενης περίπτωσης  $2 \times 2$  πινάκων συναφείας.

Ας θεωρήσουμε και πάλι την απλή περίπτωση  $N=4$ . Έστω  $p_{ij}$  η πιθανότητα ότι μία παρατήρηση ανήκει στην  $i$  γραμμή και  $j$  στήλη. (Η πιθανότητα αυτή δεν είναι η ίδια με την πιθανότητα  $p_{ij}$  της προηγούμενης περίπτωσης. Εδώ το άθροισμα των  $p_{ij}$  σε όλα τα κελιά του πίνακα ( $\sum_{i,j} p_{ij} = 1$ ) είναι ίσο με 1. Στην προηγούμενη περίπτωση, οι τιμές  $p_{ij}$  σε κάθε γραμμή άθροισαν στην μονάδα ( $\sum_i p_{ij} = 1$ ). Επομένως, η πιθανότητα του ενδεχομένου

	Στήλη 1	Στήλη 2	
Γραμμή 1	a	b	
Γραμμή 2	c	d	
			N

υπολογίζεται με βάση την πολυωνυμική κατανομή ίση με

$$\frac{N!}{a!b!c!d!} (p_{11})^a (p_{12})^b (p_{21})^c (p_{22})^d$$

γιατί ο αριθμός των τρόπων με τους οποίους N αντικείμενα μπορούν να ταξινομηθούν στα παραπάνω κελιά ισούται με τον πολυωνυμικό συντελεστή  $N!/a!b!c!d!$  και κάθε αποτέλεσμα έχει πιθανότητα

$$(p_{11})^a (p_{12})^b (p_{21})^c (p_{22})^d.$$

Μπορεί να αποδειχθεί ότι το μέγιστο μέγεθος της κρίσιμης περιοχής, η οποία βρίσκεται στην δεξιά ουρά της κατανομής της στατιστικής

συνάρτησης  $T$ , όταν η μηδενική υπόθεση  $H_0$  είναι αληθής, αντιστοιχεί στην περίπτωση

$$p_{ij} = 1/4, \quad i = 1, 2, \quad j = 1, 2.$$

Επομένως, το επίπεδο σημαντικότητας  $\alpha$  προσδιορίζεται από την σχέση

$$P\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \frac{N!}{a!b!c!d!} \left(\frac{1}{4}\right)^N,$$

για κάθε δυνατή ταξινόμηση των  $N$  στοιχείων. Για  $N=4$  υπάρχουν 35 διαφορετικές τέτοιες ταξινομήσεις (πίνακες συναφείας), οι οποίες παρατίθενται στην συνέχεια.

**Πίνακας Δυνατών ταξινομήσεων των στοιχείων ενός τυχαίου δείγματος  
μεγέθους  $N=4$  σε  $2 \times 2$  πίνακα συναφείας**

$T=0$		$T=4/9$		$T=4/3$		$T=4$	
Ενδεχόμενο	Πιθανότητα	Ενδεχόμενο	Πιθανότητα	Ενδεχόμενο	Πιθανότητα	Ενδεχόμενο	Πιθανότητα
$\begin{pmatrix} 4 & 0 \\ 0 & 0 \end{pmatrix}$	$(1/4)^4$	$\begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}$	$12(1/4)^4$	$\begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}$	$12(1/4)^4$	$\begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$	$4(1/4)^4$
$\begin{pmatrix} 0 & 4 \\ 0 & 0 \end{pmatrix}$	$(1/4)^4$	$\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$	$12(1/4)^4$	$\begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix}$	$12(1/4)^4$	$\begin{pmatrix} 0 & 3 \\ 1 & 0 \end{pmatrix}$	$4(1/4)^4$
$\begin{pmatrix} 0 & 0 \\ 0 & 4 \end{pmatrix}$	$(1/4)^4$	$\begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}$	$12(1/4)^4$	$\begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}$	$12(1/4)^4$	$\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$	$4(1/4)^4$
$\begin{pmatrix} 0 & 0 \\ 4 & 0 \end{pmatrix}$	$(1/4)^4$	$\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$	$12(1/4)^4$	$\begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix}$	$12(1/4)^4$	$\begin{pmatrix} 0 & 1 \\ 3 & 0 \end{pmatrix}$	$4(1/4)^4$
$\begin{pmatrix} 3 & 1 \\ 0 & 0 \end{pmatrix}$	$4(1/4)^4$	Σύνολο =	$48/256$	$\begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}$	$12(1/4)^4$	$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$	$6(1/4)^4$
$\begin{pmatrix} 0 & 3 \\ 0 & 1 \end{pmatrix}$	$4(1/4)^4$			$\begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix}$	$12(1/4)^4$	$\begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}$	$6(1/4)^4$
						Σύνολο =	$28/256$

(συνεχίζεται)

**Πίνακας Δυνατών ταξινομήσεων των στοιχείων ενός τυχαίου δείγματος  
μεγέθους N=4 σε 2×2 πίνακα συναφείας**

(συνέχεια)

<u>T=0</u>		<u>T=4/9</u>		<u>T=4/3</u>		<u>T=4</u>	
Ενδεχόμενο	Πιθανότητα	Ενδεχόμενο	Πιθανότητα	Ενδεχόμενο	Πιθανότητα	Ενδεχόμενο	Πιθανότητα
$\begin{pmatrix} 0 & 0 \\ 1 & 3 \end{pmatrix}$	$4(1/4)^4$			$\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$	$12(1/4)^4$		
$\begin{pmatrix} 1 & 0 \\ 3 & 0 \end{pmatrix}$				$\begin{pmatrix} 0 & 1 \\ 2 & 1 \end{pmatrix}$	$12(1/4)^4$		
$\begin{pmatrix} 1 & 3 \\ 0 & 0 \end{pmatrix}$	$4(1/4)^4$			Σύνολο =	96/256		
$\begin{pmatrix} 0 & 1 \\ 0 & 3 \end{pmatrix}$	$4(1/4)^4$						
$\begin{pmatrix} 0 & 0 \\ 3 & 1 \end{pmatrix}$	$4(1/4)^4$						
$\begin{pmatrix} 3 & 0 \\ 1 & 0 \end{pmatrix}$	$4(1/4)^4$						
$\begin{pmatrix} 2 & 2 \\ 0 & 0 \end{pmatrix}$	$6(1/4)^4$						
$\begin{pmatrix} 0 & 2 \\ 0 & 2 \end{pmatrix}$	$6(1/4)^4$						
$\begin{pmatrix} 0 & 0 \\ 2 & 2 \end{pmatrix}$	$6(1/4)^4$						
$\begin{pmatrix} 2 & 0 \\ 2 & 0 \end{pmatrix}$	$6(1/4)^4$						
$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$	$24(1/4)^4$						
Σύνολο =	84/256						



Όπως και προηγουμένως, ορίζουμε την απροσδιόριστη τιμή 0/0 ως ίση με 0. Επομένως, η ακριβής κατανομή της στατιστικής συνάρτησης  $T$  όταν όλες οι τιμές  $p_{ij}$  είναι ίσες με  $\frac{1}{4}$  είναι η

$$P(T=0)=84/256= 0.33$$

$$P(T=4/9)=48/256= 0.19$$

$$P(T=4/3)=96/256= 0.37$$

$$P(T=4)=28/256= 0.11$$

Αν η κρίσιμη περιοχή αντιστοιχεί στην μέγιστη τιμή,  $\tau=4$ , της στατιστικής συνάρτησης  $T$ , τότε  $\alpha=0.11$ .

Σύγκριση της διαδικασίας προσδιορισμού της ακριβούς μορφής της κατανομής της στατιστικής συνάρτησης  $T$  στην περίπτωση ενός μόνο δείγματος (παρούσα περίπτωση), όπου μόνο η τιμή  $N$  είναι γνωστή, με την διαδικασία προσδιορισμού της κατανομής της αντίστοιχης στατιστικής συνάρτησης  $T$  της προηγούμενης περίπτωσης, όπου και τα αθροίσματα γραμμών είναι επίσης γνωστά, δείχνει ότι, στην προκειμένη περίπτωση, η κατανομή της στατιστικής συνάρτησης  $T$  είναι πολύ πιο δύσκολο να προσδιορισθεί λόγω του πολύ μεγαλύτερου αριθμού δυνατών ταξινομήσεων.

### **8.2.3 Ο $\chi^2$ Έλεγχος Ανεξαρτησίας - Ο Έλεγχος McNemar (Περίπτωση Δύο Συσχετισμένων Δειγμάτων)**

Όπως ήδη αναφέρθηκε, η υπόθεση ότι η εμφάνιση ενός χαρακτηριστικού  $A$  δεν επηρεάζεται από ένα άλλο χαρακτηριστικό  $B$  που ελέγχει ο έλεγχος ανεξαρτησίας με ένα και μοναδικό δείγμα, μπορεί εναλλακτικά να ελεγχθεί με βάση δύο συσχετισμένα δείγματα, ένα από τον πληθυσμό των μονάδων που έχουν το χαρακτηριστικό  $A$  και ένα από τον

πληθυσμό των μονάδων που δεν έχουν το χαρακτηριστικό A. Τα μέλη των ζευγών παρατηρήσεων πρέπει να «ταιριάζουν» ως προς όλα εκείνα τα στοιχεία που κρίνονται ότι μπορούν να συνδέονται με το χαρακτηριστικό A ενδεχομένως με εξαίρεση το υπό μελέτη χαρακτηριστικό B. Τα αποτελέσματα μιας μελέτης που βασίζεται σε συσχετισμένα δείγματα συνοψίζονται μέσω των συσχετισμένων ζευγών παρατηρήσεων όπως στον πίνακα που ακολουθεί.

Κατάσταση Μονάδων με το χαρακτηριστικό A ως προς το χαρακτηριστικό B	Κατάσταση Μονάδων χωρίς το χαρακτηριστικό A ως προς το χαρακτηριστικό B	
	Ναι (B)	Όχι (B)
Ναι (B)	a	b
Όχι (B)	c	d

Κάθε ένα από τα μέλη ενός ζεύγους παρατηρήσεων έχει ή δεν έχει το χαρακτηριστικό A και ταυτόχρονα, φυσικά, έχει ή δεν έχει το χαρακτηριστικό B.

Τα δεδομένα τέτοιων μελετών είναι, προφανώς, σε ονομαστική κλίμακα με δύο κατηγορίες «Ναι» και «Όχι» ή 0 και 1, αντίστοιχα. Επομένως, τα στοιχεία του παρακάτω πίνακα αποτελούν συχνότητα εμφάνισης των τιμών ενός δείγματος ανεξάρτητων παρατηρήσεων πάνω στο ζεύγος μεταβλητών (X, Y), όπου X, αντίστοιχα Y, αναφέρεται στην κατάσταση μιας μονάδας με το χαρακτηριστικό A, αντίστοιχα μιας μονάδας χωρίς το χαρακτηριστικό A, όσο αφορά την παρουσία του χαρακτηριστικού

B. En γένει, δηλαδή, τα δεδομένα αποτελούνται από  $n'$  ανεξάρτητες παρατηρήσεις  $(X_i, Y_i)$ ,  $i = 1, 2, \dots, n'$  πάνω σε ένα ζεύγος δίτιμων μεταβλητών  $(X, Y)$  ταξινομημένων σε έναν  $2 \times 2$  πίνακα συναφείας της μορφής:

		<b>Ταξινόμηση των τιμών των <math>Y_i</math></b>	
<b>Ταξινόμηση των τιμών των <math>X_i</math></b>		<b>0</b>	<b>1</b>
<b>0</b>		a (αριθμός (0, 0))	b (αριθμός (0, 1))
<b>1</b>		c (αριθμός (1, 0))	d (αριθμός (1, 1))

Ζεύγη με την ίδια κατάσταση όσο αφορά του χαρακτηριστικού B τόσο για το μέλος με το χαρακτηριστικό A όσο και για το μέλος χωρίς το χαρακτηριστικό A (δηλαδή ζεύγη της μορφής (1, 1) και (0, 0)) ονομάζονται *εναρμονισμένα (concordant)*. Τα ζεύγη που αντιστοιχούν σε διαφορετικές καταστάσεις ως προς το χαρακτηριστικό B (δηλαδή ζεύγη της μορφής (1, 0) και (0, 1)) ονομάζονται *μη εναρμονισμένα (discordant)*.

Οι υποθέσεις που ενδιαφερόμαστε να ελέγξουμε στις περιπτώσεις αυτές έχουν την μορφή:

$H_0$ : Δεν υπάρχει συσχέτιση μεταξύ των χαρακτηριστικών A και B

$H_1$ : Υπάρχει συσχέτιση μεταξύ των χαρακτηριστικών A και B

Κάτω από την μηδενική υπόθεση ότι δεν υπάρχει συσχέτιση μεταξύ των χαρακτηριστικών A και B, σε κάθε μη εναρμονισμένο ζεύγος το

μέλος με το χαρακτηριστικό A και το μέλος χωρίς το χαρακτηριστικό A έχουν ίσες πιθανότητες να έχουν το χαρακτηριστικό B. Επομένως, μόνο τα ζεύγη της μορφής (0, 1) και (1, 0) είναι ουσιαστικής σημασίας, γιατί η οποιαδήποτε απόκλιση από την μηδενική υπόθεση συνδέεται προφανώς με διαφορές στην αναλογία εμφάνισης αυτών των ζευγών. Ως εκ τούτου, στην πράξη, τα ζεύγη της μορφής (1, 1) και (0, 0) αγνοούνται και η ανάλυση στηρίζεται στα απομένοντα n ζεύγη όπου

$$n = (\text{αριθμός } (0, 1) \text{ ζευγών} + (\text{αριθμός } (1, 0) \text{ ζευγών}).$$

Είναι προφανές, ότι οι προς έλεγχο υποθέσεις γράφονται με την μορφή:

$$H_0 : P(X = 0, Y = 1) = P(X = 1, Y = 0)$$

$$H_1 : P(X = 0, Y = 1) \neq P(X = 1, Y = 0).$$

Οι υποθέσεις αυτές παίρνουν την ισοδύναμη μορφή (με την προσθήκη του όρου  $P(X = 0, Y = 0)$  και στα δύο μέλη της εξίσωσης που εμφανίζεται στην  $H_0$ ) ως εξής:

$$H_0 : P(X = 0) = P(Y = 0)$$

$$H_1 : P(X = 0) \neq P(Y = 0).$$

Οι παραπάνω υποθέσεις είναι επίσης ισοδύναμες με τις υποθέσεις:

$$H_0 : P(X = 1) = P(Y = 1)$$

$$H_1 : P(X = 1) \neq P(Y = 1).$$

Θέτοντας

$$p_1 = P(0, 1)$$

$$p_2 = P(1, 0),$$

οι προς έλεγχο υποθέσεις γράφονται με την μορφή

$$H_0 : p_1 = p_2$$

$$H_1 : p_1 \neq p_2,$$

ή, ισοδύναμα, με την μορφή

$$H_0 : p_1 = 1/2$$

$$H_1 : p_1 \neq 1/2.$$

Η προφανής επιλογή στατιστικής συνάρτησης για τον έλεγχο των παραπάνω υποθέσεων είναι η

$$T = \text{αριθμός των } (0, 1) \text{ ζευγών,}$$

της οποίας η κατανομή είναι η διωνυμική με παραμέτρους  $n$  και  $p=p_1=1/2$ ,

κάτω από την μηδενική υπόθεση. (Συμβολικά,  $T \stackrel{H_0}{\sim}$  διωνυμική ( $n, p=1/2$ )). Η

διωνυμική αυτή κατανομή είναι απόλυτα συμμετρική (αφού  $p=1/2$ ) και μπορεί να προσεγγισθεί για  $n > 20$  από την κανονική κατανομή με μέση τιμή

$\mu=n(1/2)$  και διασπορά  $\sigma^2= n(1/2)(1/2)$ . (Συμβολικά,  $T \stackrel{H_0}{\sim} N(n/2, n/4)$ ).

Επομένως, στην θέση της επιλεγείσας ελεγχουσυνάρτησης, μπορεί να χρησιμοποιηθεί η τυποποιημένη μορφή της

$$T' = \frac{T - n(1/2)}{\sqrt{n(1/2)(1/2)}} = \frac{b - n(1/2)}{(1/2)\sqrt{n}}.$$

Επειδή  $n=b+c$ , η παραπάνω σχέση γράφεται με την μορφή

$$T' = \frac{b - [(b+c)/2]}{\sqrt{b+c/2}} = \frac{b-c}{\sqrt{b+c}}.$$

Η κρίσιμη περιοχή του ελέγχου μπορεί, επομένως, να προσδιορισθεί χρησιμοποιώντας τις ουρές της τυποποιημένης κανονικής κατανομής και ορίζεται από την ανισότητα  $T' > z_{1-\alpha}$ .

Στην πράξη, συχνά, αντί της στατιστικής συνάρτησης  $T'$ , χρησιμοποιείται η στατιστική συνάρτηση

$$T_1 = (T')^2 = \frac{(b-c)^2}{b+c},$$

η οποία προφανώς ακολουθεί την  $\chi^2$  κατανομή με 1 βαθμό ελευθερίας. Η κρίσιμη περιοχή μεγέθους  $\alpha$ , επομένως, ορίζεται από την ανισότητα

$$T_1 > \chi_{1,1-\alpha}^2.$$

**Παράδειγμα 8.2.3:** Τα στοιχεία του πίνακα που ακολουθεί αναφέρονται σε μία αναδρομική μελέτη περιπτώσεων-μαρτύρων με συσχετισμένα δείγματα των Sartwell et al. (1969) για την μελέτη της σχέσης μεταξύ θρομβοεμβολής και χρήσης αντισυλληπτικών χαπιών. Οι περιπτώσεις ήταν 175 γυναίκες στην αναπαραγωγική ηλικία (15-44), οι οποίες πήραν εξιτήριο από 43 νοσοκομεία σε 5 πόλεις των ΗΠΑ μετά από εκδήλωση ιδιοπαθούς θρομβοφλεβίτιδας, πνευμονικής εμβολής ή εγκεφαλικής θρόμβωσης. Σε κάθε ασθενούντα (περίπτωση) αντιστοιχίστηκε ένας μάρτυρας, η επιλογή του οποίου έγινε με τρόπο που να ικανοποιούνται οι προϋποθέσεις για την «ταύτιση» του με τον ασθενούντα ως προς διάφορους σχετικούς με την έρευνα παράγοντες. Συγκεκριμένα, οι μάρτυρες αντιστοιχίστηκαν με περιπτώσεις που είχαν τα ίδια ή παρόμοια χαρακτηριστικά όσο αφορά το

νοσοκομείο στο οποίο νοσηλεύθηκαν, τον τόπο μόνιμης κατοικίας, τον χρόνο εισαγωγής στο νοσοκομείο, την φυλή στην οποία ανήκαν, την ηλικία, την οικογενειακή κατάσταση, το ιατρικό ιστορικό τους (ενδεχόμενες κυήσεις) και το είδος της περίθαλψης (ιδιωτική, ημι-ιδιωτική). Ειδικότερα, οι μάρτυρες ήταν γυναίκες που νοσηλεύθηκαν στο ίδιο νοσοκομείο κατά την διάρκεια του ίδιου εξάμηνου χρονικού διαστήματος. Τα δεδομένα όσο αφορά την χρήση αντισυλληπτικών χαπιών συνοψίζονται στον πίνακα που ακολουθεί.

Περιπτώσεις	Μάρτυρες	
	Ναι	Όχι
Ναι	10	57
Όχι	13	95

Το ερώτημα που ενδιαφέρει να εξετασθεί είναι: *Είναι οι περιπτώσεις περισσότερο πιθανές από τους μάρτυρες να έχουν χρησιμοποιήσει αντισυλληπτικά χάπια;*

**Λύση:** Αν οι καταστάσεις «Ναι» και «Όχι» αντιστοιχισθούν με τις καταστάσεις «0» και «1», οι προς έλεγχο υποθέσεις μπορούν να γραφούν με την μορφή

$$H_0 : p_1 = p_2$$

$$H_1 : p_1 \neq p_2,$$

ή, ισοδύναμα, την μορφή

$$H_0 : p_1 = 1/2$$

$$H_1 : p_1 \neq 1/2,$$

όπου

$$p_1 = P(0, 1)$$

$$p_2 = P(1, 0).$$

Σύμφωνα με τα προηγούμενα, η κατάλληλη ελεγχοσυνάρτηση για τον έλεγχο των παραπάνω υποθέσεων είναι η

$$T_1 = (T')^2 = \frac{(b-c)^2}{b+c},$$

η οποία ακολουθεί την  $\chi^2$  κατανομή με 1 βαθμό ελευθερίας. Η κρίσιμη περιοχή μεγέθους  $\alpha$ , επομένως, ορίζεται από την ανισότητα

$$T_1 > \chi_{1,1-\alpha}^2.$$

Με βάση τον παραπάνω πίνακα, η παρατηρούμενη τιμή της στατιστικής ελεγχοσυνάρτησης είναι

$$\begin{aligned} \tau_1 &= \frac{(b-c)^2}{b+c} \\ &= \frac{(57-13)^2}{57+13} \\ &= 27.66. \end{aligned}$$

Το κρίσιμο επίπεδο του ελέγχου είναι ίσο με

$$\hat{\alpha} = P(T_1 \geq 27.66 \mid H_0) = P(\chi_1^2 \geq 27.66) < 0.0005.$$

Επομένως, η υπόθεση ότι δεν υπάρχει συσχέτιση μεταξύ του παράγοντα κινδύνου (χρήση αντισυλληπτικών χαπιών) και εμφάνισης θρομβοεμβολής δεν μπορεί να θεωρηθεί εύλογη.



**Λύση με το MINITAB:** Εδώ πρέπει να διεξάγουμε έλεγχο McNemar. Έχουμε ξαναπεί ότι το MINITAB δεν δίνει αυτούσιο τον έλεγχο αυτό. Είναι όμως στην ουσία διωνυμικός έλεγχος. Το στατιστικό  $T_1$  είναι το τετράγωνο του στατιστικού  $Z$  που δίνει ο διωνυμικός έλεγχος όταν χρησιμοποιεί την κανονική προσέγγιση της διωνυμικής κατανομής. Κάνοντας διωνυμικό έλεγχο (70 δοκιμές, 57 επιτυχίες) παίρνουμε:

Test and Confidence Interval for One Proportion

Test of  $p = 0,5$  vs  $p \text{ not} = 0,5$

Sample	X	N	Sample p	95,0 % CI	Exact P-Value
1	57	70	0,814286	(0,703385; 0,897239)	0,000

Το στατιστικό  $Z$  δεν μας δίνεται αλλά μας δίνεται το παρατηρούμενο επίπεδο σημαντικότητας. Επειδή ο έλεγχος που κάνουμε είναι αμφίπλευρος το παρατηρούμενο επίπεδο σημαντικότητας είναι το ίδιο είτε χρησιμοποιούμε το  $Z$  είτε το  $T_1$ . Άρα έχουμε πολύ μικρό παρατηρούμενο επίπεδο σημαντικότητας και η μηδενική υπόθεση πρέπει να απορριφθεί.

**Λύση με το SPSS:** Το SPSS διαθέτει τον έλεγχο McNemar και έχουμε ξαναδεί πως διεξάγεται. Έτσι για το συγκεκριμένο παράδειγμα παίρνουμε τα αποτελέσματα:

ROW \* COL Crosstabulation  
Count

		COL		Total
		1	2	
ROW	1	10	57	67
	2	13	95	108
Total		23	152	175

#### Chi-Square Tests

	Value	Exact Sig. (2-sided) <sup>a</sup>
McNemar Test		,000
N of Valid Cases	175	

a Binomial distribution used.

Βλέπουμε και πάλι ότι το παρατηρούμενο επίπεδο σημαντικότητας είναι πολύ μικρό και η μηδενική υπόθεση πρέπει να απορριφθεί.

**Λύση με το SAS:** Στο παράθυρο των εντολών πληκτρολογούμε τις εξής εντολές:

```
data women;
input case control counts;
cards;
1 1 10
1 2 57
2 1 13
2 2 95
;
run;
proc freq;
tables case*control / chi sq;
weight counts;
exact mcnem;
run;
```

Η διεξαγωγή του ελέγχου McNemar απαιτεί την προσθήκη της υποεντολής **exact mcnem** στην διαδικασία **proc freq**, η οποία παρέχει την ασυμπτωτική αλλά και την ακριβή τιμή του κρίσιμου επιπέδου.

Τα αποτελέσματα που προκύπτουν περιέχονται στον παρακάτω πίνακα.

TABLE OF CASE BY CONTROL

CASE	CONTROL		Total
Frequency,	1,	2,	
Percent ,			
Row Pct ,			
Col Pct ,			
1	10	57	67
	5.71	32.57	38.29
	14.93	85.07	
	43.48	37.50	
2	13	95	108
	7.43	54.29	61.71
	12.04	87.96	
	56.52	62.50	
Total	23	152	175
	13.14	86.86	100.00

STATISTICS FOR TABLE OF CASE BY CONTROL

Statistic	DF	Value	Prob
Chi-Square	1	0.302	0.583
Likelihood Ratio Chi-Square	1	0.298	0.585
Continuity Adj. Chi-Square	1	0.102	0.749
Mantel-Haenszel Chi-Square	1	0.300	0.584
Fisher's Exact Test (Left)			0.784
(Right)			0.370
(2-Tail)			0.648
Phi Coefficient		0.042	
Contingency Coefficient		0.042	
Cramer's V		0.042	

McNemar's Test

Statistic = 27.657 DF = 1 Prob (Asymptotic) = 0.001  
 Prob (Exact) = 1.03E-07

Simple Kappa Coefficient

Kappa = 0.033 ASE = 0.061 95% Confidence Bounds  
 -0.087 0.153

Sample Size = 175

Στο πεδίο **McNemar's Test** περιέχεται η τιμή της ελεγχουσυνάρτησης  $T_1$ . Τα πεδία **Prob(Asymptotic)** και **Prob(Exact)** περιέχουν την ασυμπτωτική και την ακριβή τιμή του κρίσιμου επιπέδου, αντίστοιχα.

### 8.3 ΕΛΕΓΧΟΣ ΥΠΟΘΕΣΕΩΝ ΓΙΑ ΤΗΝ ΔΙΑΜΕΣΟ c ΠΛΗΘΥΣΜΩΝ

Ο έλεγχος της διαμέσου αυτής της ενότητας αναφέρεται στις περιπτώσεις που ενδιαφερόμαστε να ελέγξουμε κατά πόσο c πληθυσμοί έχουν την ίδια διάμεσο με βάση c ανεξάρτητα δείγματα που επιλέγονται, ένα από κάθε πληθυσμό. Στην πραγματικότητα, ο έλεγχος της διαμέσου αποτελεί μία ειδική περίπτωση εφαρμογής του ελέγχου  $\chi^2$  για την περίπτωση πινάκων συναφείας με σταθερά σύνολα γραμμών και στηλών που εξετάστηκε στην προηγούμενη ενότητα. Παρά το ότι αποτελεί ειδική περίπτωση αυτού του ελέγχου, ο έλεγχος της διαμέσου εξετάζεται ξεχωριστά λόγω της μεγάλης χρησιμότητάς του σε σχέση με τις εφαρμογές. Αξίζει να σημειωθεί στο σημείο αυτό, ότι για  $c = 2$ , ο έλεγχος αυτός αποτελεί το ανάλογο του προσημικού ελέγχου για τις διαμέσους δύο πληθυσμών που αναπτύχθηκε στην ενότητα 2.2, αλλά για την περίπτωση ανεξαρτήτων δειγμάτων.

Οι υποθέσεις τις οποίες ενδιαφερόμαστε να ελέγξουμε έχουν την μορφή

$H_0$  : Οι c πληθυσμοί έχουν την ίδια διάμεσο

$H_1$  : τουλάχιστον 2 από τους πληθυσμούς έχουν διαφορετικές διαμέσους.

Για τον έλεγχο των υποθέσεων αυτής της μορφής επιλέγονται c αμοιβαία ανεξάρτητα τυχαία δείγματα, ένα από κάθε πληθυσμό, μεγέθους  $n_1, n_2, \dots, n_c$ . Ας θεωρήσουμε το δείγμα το οποίο προκύπτει από την συνένωση των c δειγμάτων και ας θεωρήσουμε ότι m

παριστάνει την διάμεσο των  $N = \sum_{i=1}^c n_i$  στοιχείων του. Τότε, αν η

μηδενική υπόθεση είναι αληθής, ισχύει ότι  $P(X_{ij} \leq m) = p$  για κάθε  $i$ ,  $j$  όπου  $X_{ij}$  συμβολίζει την  $j$  τυχούσα παρατήρηση του  $i$  τυχαίου δείγματος,  $i = 1, 2, \dots, c$ . Δηλαδή, κάτω από την μηδενική υπόθεση, το ποσοστό των παρατηρήσεων από κάθε δείγμα που δεν υπερβαίνει την διάμεσο του συνολικού δείγματος είναι το ίδιο για κάθε επιμέρους δείγμα. Αυτό είναι άμεση συνέπεια του γεγονότος ότι αν όλοι οι πληθυσμοί έχουν την ίδια διάμεσο, τότε όλοι οι πληθυσμοί έχουν την ίδια πιθανότητα μία τιμή τους να είναι  $>$  ή  $\leq$  από την διάμεσο του συνολικού πληθυσμού.

Αν, λοιπόν, με  $O_{i1}$  συμβολίσουμε τον αριθμό των παρατηρήσεων του  $i$  επιμέρους δείγματος που η τιμή τους υπερβαίνει την τιμή  $m$  της διαμέσου του συνολικού δείγματος και με  $O_{i2}$  συμβολίσουμε τον αριθμό των παρατηρήσεων του  $i$  δείγματος που οι τιμές τους είναι  $\leq$  από την τιμή  $m$  της διαμέσου του συνολικού δείγματος, δηλαδή,

$$O_{i1} = \# \text{ των παρατηρήσεων του } i \text{ δείγματος που είναι } > m$$

$$O_{i2} = \# \text{ των παρατηρήσεων του } i \text{ δείγματος που είναι } \leq m,$$

μπορούμε να κατασκευάσουμε τον εξής  $2 \times c$  πίνακα συναφείας:

		Δείγμα					
		1	2	3	...	r	Σύνολο
# παρατηρήσεων $> m$	<table style="border-collapse: collapse; width: 100%;"> <tr> <td style="border: 1px solid black; padding: 5px;"><math>O_{11}</math></td> <td style="border: 1px solid black; padding: 5px;"><math>O_{12}</math></td> <td style="border: 1px solid black; padding: 5px;"><math>O_{13}</math></td> <td style="border: 1px solid black; padding: 5px;">...</td> <td style="border: 1px solid black; padding: 5px;"><math>O_{1c}</math></td> </tr> </table>	$O_{11}$	$O_{12}$	$O_{13}$	...	$O_{1c}$	$R_1$
$O_{11}$	$O_{12}$	$O_{13}$	...	$O_{1c}$			
# παρατηρήσεων $\leq m$	<table style="border-collapse: collapse; width: 100%;"> <tr> <td style="border: 1px solid black; padding: 5px;"><math>O_{21}</math></td> <td style="border: 1px solid black; padding: 5px;"><math>O_{22}</math></td> <td style="border: 1px solid black; padding: 5px;"><math>O_{23}</math></td> <td style="border: 1px solid black; padding: 5px;">...</td> <td style="border: 1px solid black; padding: 5px;"><math>O_{2c}</math></td> </tr> </table>	$O_{21}$	$O_{22}$	$O_{23}$	...	$O_{2c}$	$R_2$
$O_{21}$	$O_{22}$	$O_{23}$	...	$O_{2c}$			
Σύνολο	<table style="border-collapse: collapse; width: 100%;"> <tr> <td style="border: 1px solid black; padding: 5px;"><math>n_1</math></td> <td style="border: 1px solid black; padding: 5px;"><math>n_2</math></td> <td style="border: 1px solid black; padding: 5px;"><math>n_3</math></td> <td style="border: 1px solid black; padding: 5px;"></td> <td style="border: 1px solid black; padding: 5px;"><math>n_c</math></td> </tr> </table>	$n_1$	$n_2$	$n_3$		$n_c$	$N = \sum_{i=1}^c n_i$
$n_1$	$n_2$	$n_3$		$n_c$			

Εστω

$E_{i1} = \#$  παρατηρήσεων του  $i$  δείγματος που είναι  $> m$

$E_{i2} = \#$  παρατηρήσεων του  $i$  δείγματος που είναι  $\leq m$ .

Τότε, η προφανής επιλογή στατιστικής συνάρτησης για τον έλεγχο των παραπάνω υποθέσεων είναι η στατιστική συνάρτηση

$$T = \sum_{i=1}^2 \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}},$$

όπου οι αναμενόμενες συχνότητες  $E_{ij}$  δίνονται από την σχέση

$$E_{ij} = \frac{R_i n_j}{N}.$$

Ο προσδιορισμός της ακριβούς μορφής της κατανομής της στατιστικής συνάρτησης  $T$  είναι δυνατός, αλλά η διαδικασία που απαιτείται είναι πολύπλοκη και χρονοβόρα. Αποδεικνύεται, όμως, ότι η κατανομή της στατιστικής συνάρτησης  $T$  μπορεί να προσεγγισθεί ικανοποιητικά από την κατανομή  $\chi^2$  με  $(c-1)$  βαθμούς ελευθερίας. Δηλαδή,

$$T \sim \chi_{c-1}^2.$$

Επομένως, η κρίσιμη περιοχή του ελέγχου μεγέθους  $\alpha$  προσδιορίζεται από την ανισότητα

$$T \sim \chi_{c-1, 1-\alpha}^2.$$

**Παράδειγμα 8.3.1:** Ας θεωρήσουμε τα δεδομένα που αναλύθηκαν με τον έλεγχο Kruskal-Wallis για την σύγκριση τεσσάρων μεθόδων παραγωγής καλαμποκιού. Υπενθυμίζεται ότι τα δεδομένα αυτά που για διευκόλυνση του αναγνώστη περιέχονται στον πίνακα που ακολουθεί αναφέρονται στην παραγωγή καλαμποκιού ανά στρέμμα.

Μέθοδος	Παραγωγή									
1	83	91	94	89	89	96	91	92	90	
2	91	90	81	83	84	83	88	91	89	84
3	101	100	91	93	96	95	94			
4	78	82	81	77	79	81	80	81		

Να ελεγχθεί κατά πόσον οι τέσσερις μέθοδοι διαφέρουν σε αποτελεσματικότητα, θεωρώντας ότι διαφορές στις διαμέσους των πληθυσμών παραγωγής μπορούν να ερμηνευθούν ως διαφορές στους βαθμούς αποτελεσματικότητας των μεθόδων που χρησιμοποιήθηκαν.

**Λύση:** Είναι προφανές ότι οι προς έλεγχο υποθέσεις έχουν την μορφή

$H_0$ : Όλες οι μέθοδοι οδηγούν στην ίδια διάμεσο παραγωγή

$H_1$ : Τουλάχιστον δύο από τις μεθόδους οδηγούν σε

διαφορετικές τιμές της διαμέσου παραγωγής.

Το δείγμα που προκύπτει αν συνδυάσουμε τα 4 δείγματα σε ένα ενιαίο δείγμα μεγέθους  $N = n_1 + n_2 + n_3 + n_4 = 34$  και, στην συνέχεια, διατάξουμε τις παρατηρήσεις του προκύπτοντος δείγματος κατά αύξουσα σειρά μεγέθους είναι το εξής:

77 78 79 80 81 81 81 81 82 83 83 83 84 84 88 89 89  
89 90 90 91 91 91 91 91 92 93 94 94 95 96 96 100 101

Η διάμεσος του δείγματος αυτού είναι η τιμή 89. Οι αριθμοί των παρατηρήσεων από κάθε ένα από τα 4 επιμέρους δείγματα που υπερβαίνουν ή είναι  $\leq$  από την διάμεσο του συνολικού δείγματος συνοψίζονται στον πίνακα συναφείας που ακολουθεί (1η γραμμή):

		Μέθοδος				Σύνολο
		1	2	3	4	
Συχνότητα παρατηρήσεων που είναι	>89	6 4.23	3 4.71	7 3.29	0 3.77	16
	≤89	3 4.77	7 5.29	0 3.71	8 4.24	18
Σύνολο		9	10	7	8	N=34

Η στατιστική συνάρτηση για τον έλεγχο αυτών των υποθέσεων δίνεται από την σχέση

$$T = \sum_{i=1}^2 \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

και είναι προφανές ότι κάτω από την μηδενική υπόθεση ακολουθεί την κατανομή  $\chi^2$  με 3 βαθμούς ελευθερίας, δηλαδή  $T \sim \chi_{3,0.95}^2$ . Επομένως, η κρίσιμη περιοχή μεγέθους 0.05 για τον έλεγχο των παραπάνω υποθέσεων ορίζεται από την ανισότητα  $T > \chi_{3,0.95}^2$ . Οι αναμενόμενες συχνότητες των παρατηρήσεων από κάθε ένα από τα 4 δείγματα που υπερβαίνουν ή δεν υπερβαίνουν την τιμή 89 της διαμέσου του συνολικού πληθυσμού υπολογίζονται από την σχέση

$$E_{ij} = \frac{R_i n_j}{N}, \quad i = 1, 2, \quad j = 1, 2, 3, 4$$

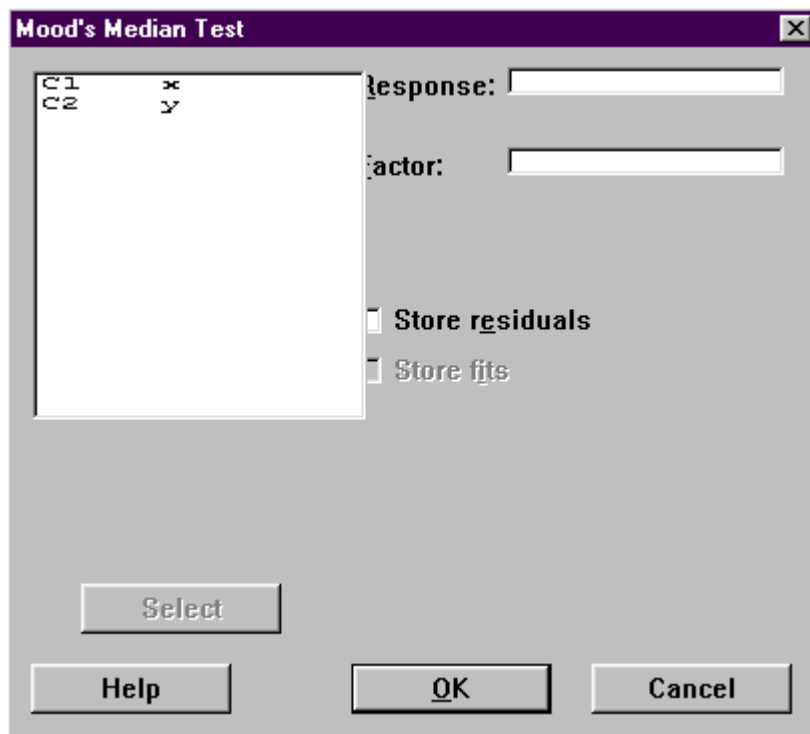
και δίνονται στην δεύτερη γραμμή των κελιών του παραπάνω πίνακα.

Η παρατηρούμενη τιμή  $t$  της στατιστικής συνάρτησης  $T$  είναι 17.57. Επομένως, σε επίπεδο σημαντικότητας 0.05, η μηδενική υπόθεση δεν απορρίπτεται, δηλαδή δεν είναι εύλογο να θεωρηθεί ότι η παραγωγή (και, επομένως, η αποτελεσματικότητα των μεθόδων που



χρησιμοποιήθηκαν) διαφέρει στατιστικά σημαντικά από μέθοδο σε μέθοδο.

**Λύση με το MINITAB:** Καταχωρίζουμε σε μία μεταβλητή (έστω  $x$ ) όλες τις μετρήσεις. Σε μία δεύτερη μεταβλητή (έστω  $y$ ), καταχωρίζουμε τιμές-δείκτες που δηλώνουν σε ποιο πληθυσμό (σε ποια μέθοδο στο συγκεκριμένο παράδειγμα) αντιστοιχεί κάθε παρατήρηση (1 για την πρώτη μέθοδο, 2 για τη δεύτερη, κ.ο.κ.). Στην συνέχεια, επιλέγοντας **Stat, Nonparametrics, Mood's Median Test**, οδηγούμεθα στο εξής πλαίσιο διαλόγου:



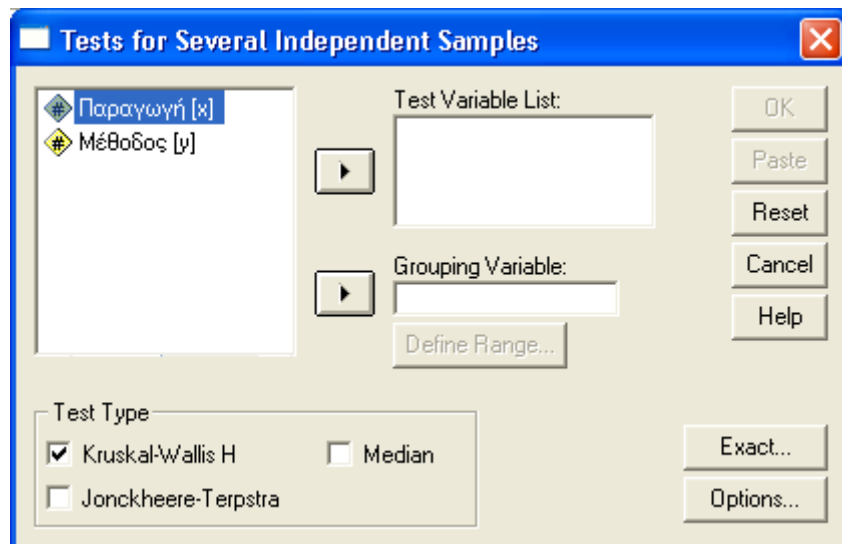
Στο πεδίο **Response**, δηλώνουμε την μεταβλητή  $x$  που περιέχει τις μετρήσεις και, στο πεδίο **Factor**, δηλώνουμε την μεταβλητή  $y$  που

περιέχει τους δείκτες. Πιέζουμε **OK** και οδηγούμεθα στα εξής αποτελέσματα:

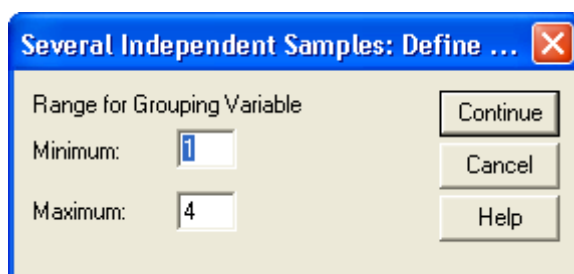
Mood median test for x  
Chi-Square = 17.54    DF = 3    P = 0.001

Παρατηρούμε ότι το κρίσιμο επίπεδο είναι εξαιρετικά χαμηλό και συνεπώς η μηδενική υπόθεση ότι οι τέσσερις μέθοδοι έχουν την ίδια αποτελεσματικότητα δεν μπορεί να θεωρηθεί εύλογη.

**Λύση με το SPSS:** Καταχωρίζουμε σε μία μεταβλητή (έστω **x**) όλες τις μετρήσεις. Σε μία δεύτερη μεταβλητή (έστω **y**), καταχωρίζουμε τιμές-δείκτες που δηλώνουν σε ποιο πληθυσμό (σε ποια μέθοδο στο συγκεκριμένο παράδειγμα) αντιστοιχεί κάθε παρατήρηση (1 για την πρώτη μέθοδο, 2 για τη δεύτερη, κ.ο.κ.). Στην συνέχεια, επιλέγοντας **Analyze, Nonparametric Tests, K Independent Samples**, οδηγούμεθα στο εξής πλαίσιο διαλόγου:



Στο πεδίο **Test Variable List**, δηλώνουμε την μεταβλητή **x** που περιέχει τις μετρήσεις, στο πεδίο **Grouping Variable**, δηλώνουμε την μεταβλητή **y** που δείχνει σε ποιο πληθυσμό αντιστοιχεί κάθε παρατήρηση. Στο πεδίο **Test Type**, ακυρώνουμε την επιλογή **Kruskal-Wallis H** και επιλέγουμε **Median**. Πιέζοντας το πλήκτρο **Define Range**, μπορούμε να ορίσουμε όσες από τις τιμές της μεταβλητής-δείκτη **y** αντιστοιχούν στους πληθυσμούς που επιθυμούμε να συγκρίνουμε. Συγκεκριμένα, οδηγούμεθα στο εξής πλαίσιο:



Στο πλαίσιο αυτό, δηλώνουμε την μικρότερη και την μέγιστη τιμή της **y** που θα ληφθεί υπόψη. (Θα μπορούσαμε δηλαδή να εξαιρέσουμε κάποιες μεθόδους από την ανάλυση). Πιέζοντας **Continue** και **OK** στο επανεμφανιζόμενο αρχικό πλαίσιο, παίρνουμε τα εξής αποτελέσματα:

Frequencies

		Μέθοδος			
		1	2	3	4
Παραγωγή	> Median	6	3	7	0
	<= Median	3	7	0	8

Test Statistics<sup>b</sup>

	Παραγωγή
N	34
Median	89.00
Chi-Square	17.543 <sup>a</sup>
df	3
Asymp. Sig.	.001

- a 7 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 3.3.
- b Grouping Variable: Μέθοδος

Όπως μπορούμε να δούμε, η τιμή της ελεγχουσυνάρτησης T είναι  $\tau=17.543$  και το παρατηρούμενο επίπεδο σημαντικότητας είναι πάρα πολύ μικρό. Συνεπώς, η μηδενική υπόθεση ότι όλες οι μέθοδοι παραγωγής έχουν την ίδια αποτελεσματικότητα δεν μπορεί να θεωρηθεί ως εύλογη.

**Λύση με το SAS:** Για την επίλυση στο SAS πληκτρολογούμε στο παράθυρο εντολών τα παρακάτω:

```
data corn;
input rows columns counts;
cards;
1 1 6
1 2 3
1 3 7
1 4 0
2 1 3
2 2 7
2 3 0
2 4 8
;
run;
proc freq;
tables rows*columns / chisq;
weight counts;
exact chisq;
run;
```

Τα αποτελέσματα δίδονται στον πίνακα που ακολουθεί.

The SAS System

TABLE OF ROWS BY COLUMNS

ROWS	COLUMNS				Total
Frequency,	1,	2,	3,	4,	
Percent					
Row Pct					
Col Pct					
1	6	3	7	0	16
	17.65	8.82	20.59	0.00	47.06
	37.50	18.75	43.75	0.00	
	66.67	30.00	100.00	0.00	
2	3	7	0	8	18
	8.82	20.59	0.00	23.53	52.94
	16.67	38.89	0.00	44.44	
	33.33	70.00	0.00	100.00	
Total	9	10	7	8	34

26.47    29.41    20.59    23.53    100.00  
 STATISTICS FOR TABLE OF ROWS BY COLUMNS

Statistic	DF	Value	Prob (Asymptotic)	Prob (Exact)
Chi-Square	3	17.543	0.001	1.75E-04
Likelihood Ratio Chi-Square	3	23.342	0.001	9.72E-05
Mantel-Haenszel Chi-Square	1	2.881	0.090	0.097
Phi Coefficient		0.718		
Contingency Coefficient		0.583		
Cramer's V		0.718		

Sample Size = 34  
 WARNING: 88% of the cells have expected counts less than 5.  
 (Asymptotic) Chi-Square may not be a valid test.

Η τιμή της ελεγχουσυνάρτησης υπολογίσθηκε ίση με 17.543 ενώ το ακριβές παρατηρούμενο επίπεδο στατιστικής σημαντικότητας είναι ίσο με 0.000175 που οδηγεί σε απόρριψη της αρχικής υπόθεσης σε  $\alpha=5\%$ .

**Παρατήρηση:** Αν ο έλεγχος της διαμέσου οδηγήσει σε απόρριψη της μηδενικής υπόθεσης ότι οι πληθυσμοί έχουν την ίδια διάμεσο και ο ερευνητής επιθυμεί να προβεί σε παραπέρα διερεύνηση των δειγμάτων με σκοπό να προσδιορίσει τίνων πληθυσμών οι διάμεσοι διαφέρουν, ο έλεγχος της διαμέσου μπορεί να επαναλαμβάνεται για οποιοδήποτε υποσύνολο δύο ή περισσότερων πληθυσμών, μέχρις ότου εντοπισθούν οι διαφορές. Θα πρέπει να τονισθεί βέβαια, ότι αυτή η διαδικασία εντοπισμού των διαφορών με την επανάληψη του ίδιου ελέγχου σε υποσύνολα πληθυσμών διαστρεβλώνει το πραγματικό επίπεδο σημαντικότητας όλων των ελέγχων, με εξαίρεση βεβαίως το επίπεδο σημαντικότητας του πρώτου. Η χρησιμοποίηση, επομένως, τέτοιων επαναλαμβανόμενων διαδικασιών ελέγχου θα πρέπει να χρησιμοποιείται με προσοχή και μόνο ως ένα αντικειμενικό κριτήριο διάκρισης μεταξύ των διαφόρων πληθυσμών και, σε καμία περίπτωση, τα αποτελέσματά τους δεν μπορούν να ερμηνευθούν με τον ίδιο τρόπο

μες τον οποίο ερμηνεύονται τα αποτελέσματα του πρώτου ελέγχου με βάση το σύνολο των διαθέσιμων δειγμάτων.

**Παρατήρηση:** Στο παράδειγμα που αναφέρθηκε παραπάνω, τα δεδομένα προέκυψαν στο πλαίσιο ενός πειράματος γνωστού ως *πλήρως τυχαιοποιημένου πειράματος*, το οποίο υποθέτει ότι οι διαφορετικές μέθοδοι παραγωγής εφαρμόστηκαν στις διαφορετικές εκτάσεις με τυχαίο τρόπο. Η συνήθης παραμετρική μέθοδος για την ανάλυση τέτοιων δεδομένων είναι γνωστή ως *ανάλυση διασποράς κατά ένα κριτήριο*.

**Παρατήρηση:** Ο έλεγχος της διαμέσου μπορεί να τροποποιηθεί και να χρησιμοποιηθεί, ως *έλεγχος ποσοστιαίων σημείων* για τον έλεγχο της μηδενικής υπόθεσης  $H_0$ : οι  $c$  ανεξάρτητοι πληθυσμοί έχουν το ίδιο  $p$ -ποσοστιαίο σημείο, για οποιαδήποτε τιμή του  $p$  επιλεγεί. Αυτό μπορεί να επιτευχθεί με την διάκριση των τιμών των δειγμάτων σε τιμές που υπερβαίνουν την τιμή του συγκεκριμένου  $p$ -ποσοστιαίου σημείου του συνενωμένου δείγματος και σε τιμές που είναι το πολύ ίσες με αυτή. Η διεξαγωγή του ελέγχου παραμένει η ίδια.

**Παρατήρηση:** Ο έλεγχος της διαμέσου μπορεί επίσης να χρησιμοποιηθεί στις περιπτώσεις για τις οποίες ο έλεγχος Mann-Whitney είναι κατάλληλος. Η ασυμπτωτική σχετική αποτελεσματικότητα του ελέγχου Mann-Whitney συγκρινόμενου με τον έλεγχο της διαμέσου είναι 1.5 για κανονικούς πληθυσμούς, 3 για ομοιόμορφους πληθυσμούς, αλλά μόνο 0.75 στην περίπτωση εκθετικών πληθυσμών.

Σημειώνεται ότι, για μικρά δείγματα, ο έλεγχος Mann-Whitney ενδέχεται να έχει μεγαλύτερη ισχύ από τον έλεγχο της διαμέσου στην περίπτωση των εκθετικών πληθυσμών. Από το άλλο μέρος, ο έλεγχος της διαμέσου δεν προϋποθέτει οι δύο πληθυσμοί να είναι ίδιοι κάτω από την μηδενική υπόθεση. Αυτό που απαιτεί μόνο είναι οι πληθυσμοί να έχουν την ίδια διάμεσο. Επομένως, ο έλεγχος της διαμέσου μπορεί να χρησιμοποιηθεί στις περιπτώσεις στις οποίες ο έλεγχος Mann-Whitney δεν μπορεί να εφαρμοσθεί.

## ΑΣΚΗΣΕΙΣ

1. Δείγματα από τέσσερα είδη κεραμικής ύλης εκτεθέντα σε υψηλή θερμοκρασία εμφάνισαν αποτελέσματα τριών κατηγοριών, όπως προκύπτει από τον παρακάτω πίνακα.

Είδη κεραμικής ύλης	A	B	Γ	Δ	Σύνολο
Κατεστράφησαν τελείως	22	27	34	31	114
Παρουσίασαν ελαφρές ζημιές	48	56	41	52	197
Παρέμειναν άθικτα	30	17	25	17	89
Σύνολο	100	100	100	100	400

Μπορείτε να ισχυρισθείτε ότι οι κατηγορίες είναι ανεξάρτητες των ειδών;

(Οικ. Παν/μιο Αθηνών – Εξετ. Σεπτ. 1998)

2. 60 φοιτητές χωρίστηκαν σε δύο ομάδες των 30 και διδάχθηκαν πώς να γράφουν ένα πρόγραμμα στον Η/Υ. Η μία ομάδα διδάχθηκε με την

συνηθισμένη μέθοδο, ενώ η άλλη με μία νέα (πειραματική) μέθοδο. Στο τέλος, κάθε φοιτητής έγραψε ένα πρόγραμμα. Τα αποτελέσματα της αξιολόγησης των προγραμμάτων συνοψίζονται στον παρακάτω πίνακα.

Μέθοδος	Πρόγραμμα	
	σωστό	λάθος
Συνηθισμένη μέθοδος	23	7
Πειραματική μέθοδος	27	3

Μπορεί να συμπεράνει κανείς από τα αποτελέσματα ότι η πειραματική μέθοδος είναι ανώτερη;

3. Ένα τυχαίο δείγμα 196 πτυχιούχων ενός μεγάλου πανεπιστημίου οι οποίοι είχαν εισαχθεί στο παν/μιο κατά την διάρκεια τριών διαδοχικών ετών ταξινομήθηκαν σύμφωνα με διάφορα τμήματα του πανεπιστημίου (όπως αυτή καθορίζεται από τα μόρια εισαγωγής) και σύμφωνα με την σειρά με την οποία απεφοίτησαν από αυτά (όπως αυτή καθορίζεται από τους βαθμούς των πτυχιούχων της περιόδου κατά την οποία απεφοίτησαν) ως εξής:

		Σειρά αποφοίτησης			
		1	2	3	άλλη
Σειρά επιτυχίας στα	1-4	15	14	15	52
διάφορα τμήματα	5-9	9	10	9	72

Θα μπορούσε να συμπεράνει κανείς ότι η σειρά αποφοίτησης εξαρτάται από την σειρά επιτυχίας στο παν/μιο;



4. Τα στοιχεία που περιέχονται στον πίνακα που ακολουθεί αναφέρονται σε μια μελέτη που έγινε με βάση δύο ανεξάρτητα δείγματα παιδιών με ή χωρίς ιστορικό βρογχίτιδας στην νηπιακή τους ηλικία προκειμένου να συγκριθούν οι αναλογίες, στις οποίες τα παιδιά των δύο αυτών κατηγοριών παρουσιάζουν αναπνευστικά προβλήματα αργότερα στην παιδική ηλικία. Το πρώτο δείγμα αποτελείται από 273 παιδιά που είχαν βρογχίτιδα κατά την νηπιακή ηλικία, από τα οποία 26 αναφέρθηκαν ως έχοντα σύμπτωμα επίμονου βήχα κατά την διάρκεια της ημέρας ή της νύκτας γύρω στην ηλικία των 12-14 ετών. Το δεύτερο δείγμα αποτελείται από 1046 παιδιά τα οποία δεν είχαν παρουσιάσει βρογχίτιδα κατά τη νηπιακή ηλικία και από τα οποία 44 αναφέρθηκαν ως παρουσιάζοντα το ίδιο σύμπτωμα κατά την ίδια ηλικία.

Σύμπτωμα Βήχα	Ιστορικό Βρογχίτιδας		Σύνολο
	Ναι	Όχι	
Ναι	26	44	70
Όχι	247	1002	1249
Σύνολο	273	1046	1319

Δείχνουν τα στοιχεία αυτά στατιστικά σημαντική διαφορά στις αναλογίες των αντίστοιχων πληθυσμών παιδιών; ( $\alpha=5\%$ )

5. Ο πίνακας που ακολουθεί συνοψίζει το αποτέλεσμα μίας έρευνας με αντικείμενο την μελέτη του κατά πόσο ψυχολογικοί παράγοντες που συνδέονται με το είδος της στέγης κάτω από την οποία ζουν έγκυοι γυναίκες μπορεί να θεωρηθούν ότι περιλαμβάνονται μεταξύ των παραγόντων που οδηγούν σε πρόωρο τοκετό.

Είδος Κατοικίας	Είδος Τοκετού		Σύνολο
	Πρόωρος	Κανονικός	

<u>Ιδιόκτητη</u>	50	849	899
Εργατική	29	229	258
Ενοικιασμένη	11	164	175
Συγκατοικεί με γονείς	6	66	72
Άλλο	3	36	39
<u>Σύνολο</u>	<u>99</u>	<u>1344</u>	<u>1443</u>

Να ελεγχθεί, σε επίπεδο σημαντικότητας 1%, κατά πόσον ο χρόνος τοκετού επηρεάζεται από το είδος της στέγης κάτω από την οποία ζουν έγκυοι γυναίκες.

6. Ο παρακάτω πίνακας παρουσιάζει το αποτέλεσμα ταξινόμησης των παρατηρήσεων ενός τυχαίου δείγματος 94 φοιτητών ενός παν/μίου ως προς την σχολή στην οποία φοιτούν και ως προς το εάν απεφοίτησαν από γυμνάσιο της Αθήνας ή όχι.

	Φυσ/κή	Φιλος.	Νομική	Λοιπές Σχολές
Γυμνάσιο Αθήνας	16	14	13	13
Γυμνάσιο άλλης πόλης	14	6	10	8

Να ελεγχθεί αν η επιτυχία σε μία σχολή εξαρτάται από τον τόπο γυμνασιακών σπουδών. ( $\alpha=5\%$ )

(Παν/μιο Κρήτης, Ηράκλειο – Εξ.ετ. Ιανουαρίου 1985)

7. Στον παρακάτω πίνακα 80, άλογα ιπποδρόμου έχουν ταξινομηθεί ως προς τον αριθμό της λωρίδας εκκίνησης και ως προς τον αριθμό της λωρίδας τερματισμού.

Λωρίδα Τερματισμού

		1	2	3	Άλλη ( $\neq 1, 2, 3$ )
Λωρίδα	1-4	8	6	8	16
Εκκίνησης	5-9	3	6	5	28

Εξαρτάται η θέση του αλόγου κατά τον τερματισμό από την θέση κατά την εκκίνηση;

(Παν/μιο Κρήτης, Ηράκλειο – Εξ. Ιουνίου 1985)

**8.** 214 ασθενείς με καρδιολογικά προβλήματα έλαβαν μέρος σε μια έρευνα για να ελεγχθεί η επίδραση της ασπιρίνης στην επιβίωση από καρδιακό επεισόδιο. Σε 189 ασθενείς χορηγήτο ψευδοφάρμακο και σε 104 ασθενείς χορηγήτο ασπιρίνη για ένα μεγάλο χρονικό διάστημα, στο τέλος του οποίου κατεγράφη ο αριθμός των θανάτων από καρδιακό επεισόδιο.

	Καρδιακή προσβολή	
	Θάνατος	Όχι θάνατος
Ψευδοφάρμακο	18	171
Ασπιρίνη	5	99

**α)** Να υπολογισθεί η πιθανότητα θανάτου δοθέντος ότι στον ασθενή έχει χορηγηθεί ασπιρίνη και η πιθανότητα θανάτου δοθέντος ότι στον ασθενή έχει χορηγηθεί ψευδοφάρμακο. Να συγκριθούν οι δύο πιθανότητες.

**β)** Να ελεγχθεί η υπόθεση της ανεξαρτησίας μεταξύ θανάτου από καρδιακή προσβολή και λήψης φαρμάκου και να συγκριθούν τα αποτελέσματα των δύο ελέγχων?

**9.** Σε δυο ομάδες ψυχασθενών ερευνήθηκαν οι τάσεις αυτοκτονίας. Να ελεγχθεί η υπόθεση της ανεξαρτησίας μεταξύ του τύπου ψυχασθένειας

και της τάσης για αυτοκτονία με τον έλεγχο Fisher. Ποια η ερμηνεία του αποτελέσματος του ελέγχου και ποιο το αποτέλεσμα του κλασσικού ελέγχου  $\chi^2$ ;

	Ψυχωτικοί	Νευρωτικοί
Τάσεις αυτοκτονίας	2	6
Όχι τάσεις αυτοκτονίας	18	14

10. Ο παρακάτω πίνακας συνοψίζει δεδομένα που συνελέγεσαν για τον έλεγχο της υπόθεσης ότι η κατανάλωση αλκοόλ (σε γρ./ημέρα) συνδέεται με την κίρρωση του ήπατος και αναφέρονται σε 268 ασθενείς και 923 μη ασθενείς.

	Επίπεδο κατανάλωσης (γρ./ημέρα)							
	0-19	20-39	40-59	60-79	80-99	100-119	>=120	
<b>Κίρρωση</b>	5	21	30	41	50	45	76	268
<b>Όχι κίρρωση</b>	222	228	225	112	64	33	39	923
	227	249	255	153	114	78	115	1191

Να ελεγχθεί η υπόθεση ότι η κατανάλωση αλκοόλ δεν συνδέεται με την κίρρωση του ήπατος.

11. Χρησιμοποιώντας τα δεδομένα των διαλέξεων (σχέση ψεύδους με ηλικία σε νεαρά αγόρια) ελέγξτε εάν η χρήση των καταλοίπων είναι ικανή να μας δώσει ένδειξη για ύπαρξη γραμμικής τάσης η οποία δε φαίνεται να υπάρχει με βάση τον υπολογισμό του  $\chi^2$  για ανεξαρτησία.

	Ηλικία					
	5-7	8-9	10-11	12-13	14-15	
<b>Ψεύτες</b>	6	18	19	27	25	95
<b>Όχι ψεύτες</b>	15	31	31	32	19	128

	21	49	50	59	44	223
--	----	----	----	----	----	-----

**12.** Για να δοκιμασθεί μια νέα θεραπεία αλλεργικού άσθματος, οι ασθενείς υποβάλλονται πριν από την εφαρμογή της σε ένα περιβάλλον με πολλούς αλλεργιογόνους παράγοντες και καταγράφεται ο αριθμός των ασθενών που αντέχουν πάνω από μια ώρα σε ένα τέτοιο περιβάλλον. Μετά την εφαρμογή της θεραπείας, επαναλαμβάνεται το ίδιο πείραμα. Τα δεδομένα συνοψίζονται στον πίνακα που ακολουθεί. Φαίνεται να μπορεί να υποστηριχθεί ο ισχυρισμός ότι υπάρχει σχέση μεταξύ θεραπείας και αντοχής σε αλλεργιογόνο περιβάλλον;

		Μετά την θεραπεία	
		Αντοχή	Μη αντοχή
Πριν την θεραπεία	Αντοχή	101	22
	Μη αντοχή	61	34

**13.** Οι απαντήσεις τριάντα φοιτητών και τριάντα φοιτητριών, τυχαία επιλεγμένων, στο ερώτημα: "Θα προσπαθούσε να διεκδικήσει ένα μεταπτυχιακό τίτλο μία μετρίας απόδοσης πτυχιούχος;" δίνονται στον παρακάτω πίνακα

<i>Απάντηση στην ερώτηση</i>	<i>Φοιτητές</i>	<i>Φοιτήτριες</i>
<i>ΝΑΙ</i>	9	15
<i>ΟΧΙ</i>	12	2
<i>Δ/Γ</i>	9	13
<i>Σύνολο</i>	30	30

Να ελέγξετε αν υπάρχουν σημαντικές διαφορές στις αντιδράσεις των δύο φύλων ( $\alpha=5\%$ ).

(Οικ. Παν/μιο Αθηνών – Εξετ. Σεπτ. 1997)

14. Με βάση τα δεδομένα του παρακάτω πίνακα και σε επίπεδο σημαντικότητας  $\alpha=5\%$ , να ελέγξετε κατά πόσον η ικανότητα των φοιτητών στο μάθημα της Στατιστικής είναι ανεξάρτητο της ικανότητάς τους στα Μαθηματικά.

		Ικανότητα στην Στατιστική		
		Καλώς	Λ. Καλώς	Άριστα
Ικανότητα στα Μαθηματικά	Καλώς	43	24	15
	Λ. Καλώς	31	46	37
	Άριστα	12	20	49

(Οικ. Παν/μιο Αθηνών – Εξ. Φεβρ. 1995)

15. Εστω ότι ο επόμενος πίνακας διπλής εισόδου δίνει τις συχνότητες επιτυχίας 390 εφήβων μιας περιοχής (οι οποίοι πέτυχαν στα ΑΕΙ) σύμφωνα με την οικονομική τους τάξη και την κατεύθυνση σπουδών που επέλεξαν.

Κατεύθυνση	Οικονομική Τάξη				
	I & II	III	IV	V	Σύνολο
Θετική	23	40	16	2	81
Θεωρητική	11	75	107	14	207
Οικονομική	1	31	60	10	102
Σύνολο	35	146	183	26	390

Μπορείτε να ισχυρισθείτε, σε επίπεδο στατιστικής σημαντικότητας 1%, ότι η αναλογία των επιτυχόντων στις τρεις διαφορετικές κατευθύνσεις είναι η ίδια για όλες τις οικονομικές τάξεις;

(Οικ. Παν/μιο Αθηνών – Εξ. Φεβρ. 1992)

16. Τα παρακάτω δεδομένα συνελέγησαν για να μελετηθεί αν υπάρχει συσχέτιση μεταξύ της επιβίωσης κατά την παιδική ηλικία και της προγεννητικής φροντίδας που παρείχαν δυο κλινικές. Ζητείται ο έλεγχος  $\chi^2$  για την πλήρη ανεξαρτησία. Να εξηγήσετε πώς θα μπορούσατε να προχωρήσετε με την ανάλυση των παραπάνω στοιχείων προκειμένου να εξετάσετε το πρόβλημα αυτό. Ποιά είναι τα συμπεράσματά σας;

	Παιδική επιβίωση			
	Θάνατος		Επιβίωση	
	Λιγότερο	Περισσότερο	Λιγότερο	Περισσότερο
Προγεννητική φροντίδα				
Κλινική Α	3	4	176	293
Κλινική Β	17	2	197	23

17. Να εξετασθεί αν τα τυχαία δείγματα

6 3 1 14 8 9 12 1 3 2 1 7

6 2 1 1 3 1 2 4 8 1 2

προέρχονται από πληθυσμούς με την ίδια διάμεσο. ( $\alpha=5\%$ )

(Παν/μιο Κρήτης, Ηράκλειο – Εξετ. Ιαν. 1985)

18. Employee integrity is important to the success of all businesses. Suppose an industrial security firm wants to conduct a study of criminal cases involving stolen company money in which employees have been found guilty. Among the data they record are the employee's salary (wages) and the amount of money stolen from the company for 400 recent cases. Does this information provide evidence of a relationship between employee income and amount stolen? Use  $\alpha = .05$ .

AMOUNT STOLEN (THOUSANDS OF DOLLARS)

		Under 5	5-10	11-20	Over 20
Income of Employee (Thousands of Dollars)	Under 15	46	39	17	5
	15-25	78	79	61	19
	Over 25	5	14	25	12

Before you give an answer

- a) State  $H_0$  and  $H_1$ .
- b) State the assumptions that you make.

*(University of Missouri, Columbia – Ežet. Mařov 1982)*

**19.** In a comparison on visual acuity of deaf and hearing children, eye movement rates were taken on ten deaf and ten hearing children. A clinical psychologist believes that deaf children have greater visual acuity than hearing children. Test the psychologist's claim using the data given in the table. (The larger a child's eye movement rate, the more visual acuity the child possesses).

#### Eye Movement Rates

Deaf Children: 2.75 3.14 3.23 2.30 2.64 1.95 2.17 2.45 1.83 2.23

Hearing Children: 1.15 1.65 1.43 1.83 1.75 1.23 2.03 1.64 1.96 1.37

Before you give an answer

- c) State  $H_0$  and  $H_1$ .
- d) State the assumptions that you make.

*(University of Iowa, Iowa City - Ežet. Mařtiov 1983)*

**20.** An experiment was conducted to determine whether a test designed to identify a particular form of mental illness could be easily interpreted



with little psychological training. Thirty judges were selected to review the results of 100 tests, half of which were given to disturbed patients and half to normal people. Of hospital, ten were staff members of a mental hospital, ten were trainees at the hospital and ten were undergraduate psychology majors. The results below give the percentages of the 100 tests correctly classified by each judge. Do the data provide sufficient evidence to indicate that the three types of judges differ in their ability to identify the mentally disturbed patients? Test using  $\alpha = .05$ .

<u>Staff:</u>	78.2	79.4	85.2	93.4	90.1	76.2	86.4	88.2	84.3	81.6
<u>Trainees:</u>	80.4	75.1	72.0	68.3	75.2	69.3	81.2	76.4	72.5	76.9
<u>Undergraduates:</u>	65.2	70.3	74.2	78.1	68.3	74.4	80.1	73.6	75.7	73.2

- a) State  $H_0$  and  $H_1$
- b) State the assumptions that you make.

*(Trinity College, University of Dublin – Ežet. Mařov 1980)*