

ΚΕΦΑΛΑΙΟ 2

ΜΕΡΙΚΟΙ ΕΛΕΓΧΟΙ ΥΠΟΘΕΣΕΩΝ ΒΑΣΙΣΜΕΝΟΙ ΣΤΗΝ ΔΙΩΝΥΜΙΚΗ ΚΑΤΑΝΟΜΗ

Όπως είναι ήδη γνωστό, η διωνυμική κατανομή χρησιμοποιείται για να περιγράψει πιθανότητες που σχετίζονται με τον αριθμό των επιτυχιών σε μία σειρά n ανεξαρτήτων δοκιμών. Κάθε μία από n ανεξάρτητες δοκιμές μπορεί να καταλήξει σε *επιτυχία* με πιθανότητα p , ή *αποτυχία* με πιθανότητα $q=1-p$. Η διωνυμική κατανομή περιγράφει την πιθανότητα να παρατηρηθούν ακριβώς k επιτυχίες.

Στην εφαρμοσμένη έρευνα, πολλές πειραματικές καταστάσεις μπορούν να περιγραφούν με το διωνυμικό μοντέλο. Αρκετοί πελάτες μπαίνουν σε ένα κατάστημα και, ανεξάρτητα ο ένας από τον άλλο, αποφασίζουν να αγοράσουν ή όχι ένα συγκεκριμένο προϊόν. Σε αρκετά πειραματόζωα χορηγείται ένα συγκεκριμένο φαρμακευτικό προϊόν, οπότε κάθε ένα από αυτά θεραπεύεται ή δεν θεραπεύεται. Παραδείγματα αυτής της μορφής μπορούν να βρεθούν σχεδόν σε κάθε επιστήμη. Τα δεδομένα που λαμβάνονται σε αυτές τις περιπτώσεις μπορούν να αναλυθούν με την χρήση ορισμένων από τις απλούστερες στατιστικές μεθόδους που είναι γνωστές, αυτές που βασίζονται στην διωνυμική κατανομή. Στο κεφάλαιο αυτό παρουσιάζονται μερικές από τις διαθέσιμες αυτές μεθόδους.

2.1 Ο ΔΙΩΝΥΜΙΚΟΣ ΕΛΕΓΧΟΣ

Ο διωνυμικός έλεγχος που περιγράφεται στα επόμενα, μπορεί να προσαρμοσθεί για τον έλεγχο σχεδόν οποιασδήποτε υπόθεσης και με δεδομένα οποιασδήποτε μορφής η οποία να επιδέχεται στατιστική ανάλυση. Σε μερικές περιπτώσεις, ο διωνυμικός έλεγχος είναι ο πιο ισχυρός έλεγχος. Στις περιπτώσεις αυτές, ο διωνυμικός έλεγχος διεκδικείται τόσο από την Παραμετρική όσο και από την Μη Παραμετρική Στατιστική. Σε άλλες περιπτώσεις, υπάρχουν περισσότερο ισχυροί έλεγχοι. Στις περιπτώσεις αυτές, ο διωνυμικός έλεγχος διεκδικείται μόνο από την μη Παραμετρική Στατιστική. Παρ' όλα αυτά, ακόμα και στις περιπτώσεις όπου υπάρχουν περισσότερο ισχυροί έλεγχοι, ο διωνυμικός έλεγχος συχνά προτιμάται γιατί είναι απλός στη χρήση του, απλός στην ερμηνεία του και μερικές φορές,

αρκετά ισχυρός για να απορρίπτει μία μηδενική υπόθεση όταν αυτή πρέπει να απορριφθεί.

Το δείγμα αποτελείται από τα αποτελέσματα n ανεξάρτητων δοκιμών. Κάθε αποτέλεσμα ανήκει είτε στην *κατηγορία 1* (*επιτυχία*) είτε στην *κατηγορία 2* (*αποτυχία*), αλλά όχι και στις δύο κατηγορίες. Ο αριθμός των παρατηρήσεων στην κατηγορία 1 είναι O_1 και ο αριθμός των παρατηρήσεων στην κατηγορία 2 είναι $O_2 = n - O_1$. Οι n δοκιμές είναι αμοιβαία ανεξάρτητες και κάθε μία από αυτές έχει πιθανότητα p να οδηγήσει σε ενδεχόμενο της κατηγορίας 1, όπου η πιθανότητα p είναι η ίδια για κάθε μία από τις n δοκιμές.

Έστω p_0 μία συγκεκριμένη τιμή στο διάστημα $[0,1]$, δηλαδή, $0 \leq p_0 \leq 1$. Οι υποθέσεις που ενδέχεται να ενδιαφερόμαστε να ελέγξουμε μπορούν να έχουν μία από τις εξής τρεις μορφές.

A. (Αμφίπλευρος έλεγχος)

$$H_0: p = p_0$$

$$H_1: p \neq p_0$$

B. (Μονόπλευρος έλεγχος)

$$H_0: p \leq p_0$$

$$H_1: p > p_0$$

Γ. (Μονόπλευρος έλεγχος)

$$H_0: p \geq p_0$$

$$H_1: p < p_0 .$$

Επειδή ενδιαφερόμαστε και στις τρεις αυτές περιπτώσεις για την πιθανότητα του ενδεχομένου *επιτυχία*, η προφανής επιλογή στατιστικής συνάρτησης T για τον έλεγχο των υποθέσεων αυτών είναι ο αριθμός των φορών που το ενδεχόμενο *επιτυχία* εμφανίστηκε στις n ανεξάρτητες δοκιμές. Δηλαδή,

$$T = O_1$$

= αριθμός των επιτυχιών στις n ανεξάρτητες δοκιμές.

Αν η ακολουθία X_1, X_2, \dots, X_n συμβολίζει τα αποτελέσματα των n δοκιμών, όπου

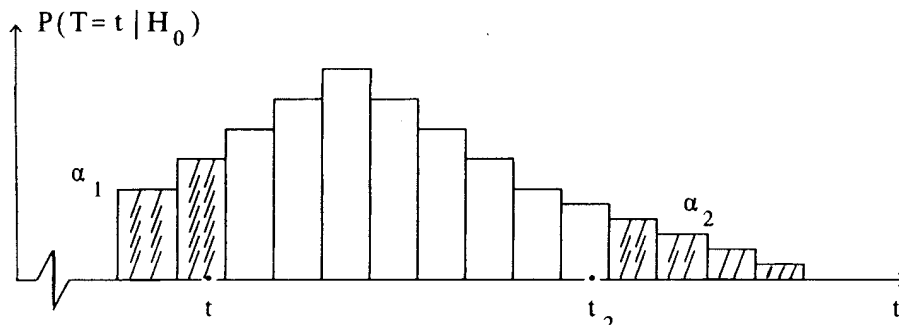
$$X_i = \begin{cases} 1, & \text{αν η } i \text{ δοκιμή οδηγεί σε επιτυχία} \\ 0, & \text{διαφορετικά,} \end{cases}$$

τότε

$$T = \sum_{i=1}^n X_i .$$

Προφανώς, η στατιστική συνάρτηση T ακολουθεί την διωνυμική κατανομή με παραμέτρους n και p . (Συμβολικά, $T \sim \text{διωνυμική}(n, p)$). Επειδή, επομένως, η στατιστική συνάρτηση T είναι διακριτή τυχαία μεταβλητή, ο έλεγχος των υποθέσεων των περιπτώσεων A , B και Γ σπάνια θα μπορεί να γίνει σε επίπεδο σημαντικότητας ακριβώς ίσο με το επιθυμητό επίπεδο α .

A. (Αμφίπλευρος έλεγχος): Η κρίσιμη περιοχή μεγέθους α για τον έλεγχο των υποθέσεων της περίπτωσης αυτής αντιστοιχεί στις δύο ουρές της διωνυμικής κατανομής με παραμέτρους p_0 και n , όπου το μέγεθος της κάτω ουράς είναι α_1 , το μέγεθος της πάνω ουράς είναι α_2 και $\alpha_1 + \alpha_2 = \alpha$ (σχήμα 2.1.1).



Σχήμα 2.1.1

Κρίσιμη περιοχή του ελέγχου των υποθέσεων

$H_0: p=p_0, H_1: p \neq p_0$

Πράγματι, από την μορφή της εναλλακτικής υπόθεσης, είναι προφανές ότι οι πολύ μεγάλες τιμές και οι πολύ μικρές τιμές της στατιστικής συνάρτησης T θα αποτελούν ένδειξη εναντίον της ισχύος της μηδενικής υπόθεσης H_0 . Αν, επομένως, t_1 και t_2 , $t_1 < t_2$, συμβολίζουν τις κρίσιμες τιμές της στατιστικής συνάρτησης T , τότε η κρίσιμη περιοχή του ελέγχου έχει την μορφή $(-\infty, t_1] \cup (t_2, +\infty)$. Δηλαδή, τιμές της στατιστικής συνάρτησης T που δεν υπερβαίνουν την τιμή t_1 ($T \leq t_1$) ή που υπερβαίνουν την τιμή t_2 ($T > t_2$) θα συνηγορούν υπέρ της απόρριψης της μηδενικής υπόθεσης H_0 . Οι κρίσιμες τιμές t_1 και t_2 θα καθορισθούν έτσι ώστε

$$\begin{aligned} \alpha &= \max P(H_0 | H_0) \\ &= P[(\{T \leq t_1\} \cup \{T > t_2\}) | n, p=p_0] \\ &= P(T \leq t_1 | p=p_0) + P(T > t_2 | n, p=p_0). \end{aligned}$$

Οι τιμές των α_1 και α_2 πρέπει να είναι περίπου ίσες. Όπως ήδη αναφέρθηκε, επειδή η στατιστική συνάρτηση T είναι διακριτή, το μέγεθος του ελέγχου σπάνια θα είναι ακριβώς ίσο με α . Τότε,

χρησιμοποιώντας τον πίνακα 1 του παραρτήματος για τις συγκεκριμένες τιμές των p_0 και n , προσδιορίζουμε τις τιμές t_1 και t_2 έτσι ώστε

$$P(T \leq t_1 | n, p=p_0) = \alpha_1 \cong \alpha/2$$

και

$$P(T > t_2 | n, p=p_0) = \alpha_2 \cong \alpha/2.$$

Η τελευταία σχέση γράφεται ισοδύναμα

$$P(T \leq t_1 | n, p=p_0) = 1 - \alpha_2 \cong 1 - \alpha/2.$$

Ο κανόνας απόφασης είναι ο εξής:

Η υπόθεση H_0 απορρίπτεται σε επίπεδο σημαντικότητας να μην χωρίζει $\alpha_1 + \alpha_2 \cong \alpha$ όταν η τιμή της στατιστικής συνάρτησης T είναι μικρότερη ή ίση από την τιμή t_1 ($T \leq t_1$), ή όταν η τιμή της στατιστικής συνάρτησης T υπερβαίνει την τιμή t_2 ($T > t_2$), όπου t_1, t_2 προσδιορίζονται από τις σχέσεις

$$P(T \leq t_1 | n, p=p_0) = \alpha_1 \cong \alpha/2$$

και

$$P(T > t_2 | n, p=p_0) = \alpha_2 \cong \alpha/2.$$

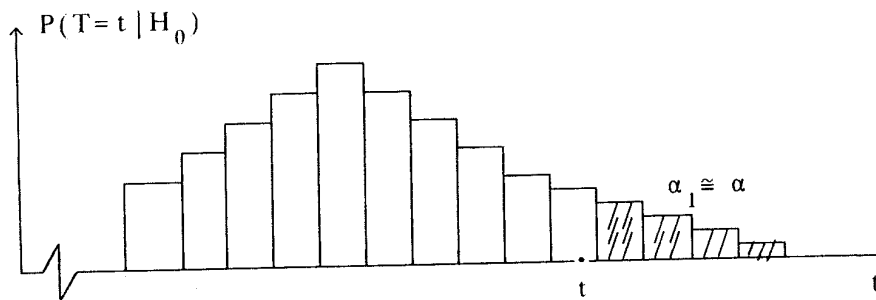
B. (Μονόπλευρος έλεγχος): Από την μορφή της εναλλακτικής υπόθεσης της περίπτωσης αυτής, είναι προφανές ότι μεγάλες τιμές της στατιστικής συνάρτησης T θα αποτελούν ένδειξη ότι η μηδενική υπόθεση H_0 δεν είναι αληθής. Τότε, η κρίσιμη περιοχή του ελέγχου θα βρίσκεται στην δεξιά ουρά της κατανομής, και θα έχει μέγεθος $\alpha_1 \cong \alpha$ (σχήμα 2.1.2). Δηλαδή, η κρίσιμη περιοχή θα αποτελείται από όλες τις τιμές της στατιστικής συνάρτησης T οι οποίες υπερβαίνουν την τιμή t ,

όπου η τιμή t προσδιορίζεται από τον πίνακα 1 του παραρτήματος, για τις συγκεκριμένες τιμές p_0 και n και έτσι ώστε

$$P(T > t \mid n, p=p_0) = \alpha_1 \cong \alpha,$$

ή, ισοδύναμα

$$P(T \leq t \mid n, p=p_0) = 1 - \alpha_2 \cong 1 - \alpha.$$



Σχήμα 2.1.2

Κρίσιμη περιοχή του ελέγχου των υποθέσεων $H_0: p \leq p_0$, $H_1: p > p_0$

Επομένως, ο κανόνας απόφασης είναι ο εξής:

Η μηδενική υπόθεση H_0 απορρίπτεται σε επίπεδο σημαντικότητας $\alpha_1 \cong \alpha$, αν η τιμή της στατιστικής συνάρτησης T υπερβαίνει την τιμή t ($T > t$), όπου η τιμή t προσδιορίζεται από την σχέση

$$P(T \leq t \mid n, p=p_0) = 1 - \alpha_2 \cong 1 - \alpha.$$

Γ. (Μονόπλευρος έλεγχος): Για τον έλεγχο των υποθέσεων της περίπτωσης αυτής, παρατηρούμε ότι οι μικρές τιμές της στατιστικής συνάρτησης T θα συνηγορούν υπέρ της απόρριψης της μηδενικής υπόθεσης. Επομένως, η κρίσιμη περιοχή μεγέθους $\alpha_1 \cong \alpha$ θα βρίσκεται στην αριστερή ουρά της κατανομής (σχήμα 2.1.3). Δηλαδή, η κρίσιμη περιοχή αποτελείται από όλες τις τιμές της στατιστικής

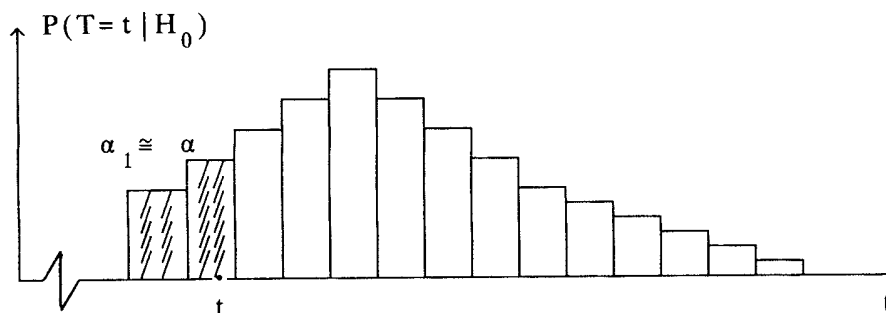
συνάρτησης T οι οποίες δεν υπερβαίνουν την τιμή t , όπου η τιμή t προσδιορίζεται από τον πίνακα 1 του παραρτήματος για τις συγκεκριμένες τιμές p_0 και n έτσι ώστε

$$P(T \leq t \mid n, p=p_0) = \alpha_1 \cong \alpha.$$

Επομένως, ο κανόνας απόρριψης είναι ο εξής:

Η μηδενική υπόθεση H_0 απορρίπτεται σε επίπεδο σημαντικότητας $\alpha_1 \cong \alpha$, αν η τιμή της στατιστικής συνάρτησης T είναι μικρότερη ή ίση από την τιμή t ($T \leq t$), όπου η τιμή t προσδιορίζεται από την σχέση

$$P(T \leq t \mid n, p=p_0) = \alpha_1 \cong \alpha.$$



Σχήμα 2.1.3

Κρισιμή περιοχή του ελέγχου των υποθέσεων $H_0: p \geq p_0$, $H_1: p < p_0$

Σημείωση: Τα αποτελέσματα των ελέγχων υποθέσεων είναι περισσότερο χρήσιμα, αν μαζί με αυτά προσδιορίζεται και η τιμή του *κρισίμου επιπέδου* (*critical level*) $\hat{\alpha}$. (Ο συνήθης συμβολισμός για το κρίσιμο επίπεδο είναι p (*p-value*). Για να αποφευχθούν όμως

ασάφειες που ενδέχεται να προκύψουν, επειδή το σύμβολο p χρησιμοποιείται πολύ συχνά στην μη παραμετρική στατιστική ως παράμετρος της διωνυμικής κατανομής, θα χρησιμοποιείται στην συνέχεια το σύμβολο $\hat{\alpha}$). Όπως είναι γνωστό, το κρίσιμο επίπεδο είναι το ελάχιστο επίπεδο σημαντικότητας στο οποίο η μηδενική υπόθεση θα απορριπτόταν για τις δοθείσες παρατηρήσεις. Το κρίσιμο επίπεδο, που είναι επίσης γνωστό ως *παρατηρούμενο επίπεδο σημαντικότητας* (*observed level of significance*) ή *τιμή πιθανότητας* (*probability value, p-value*), είναι, επομένως, η μέγιστη πιθανότητα να παρατηρηθεί μια τιμή της στατιστικής συνάρτησης T το ίδιο ακραία ή περισσότερο ακραία από την τιμή που παρατηρήθηκε, δοθέντος ότι η μηδενική υπόθεση είναι αληθής. Δηλαδή,

$$\hat{\alpha} = \max P(\text{να παρατηρηθεί μια τιμή της } T \text{ το ίδιο ακραία ή πιο ακραία από την παρατηρηθείσα τιμή της } | H_0).$$

Συγκεκριμένα, αν τ παριστάνει την παρατηρηθείσα τιμή της στατιστικής συνάρτησης T , τότε το κρίσιμο επίπεδο του ελέγχου της περίπτωσης B , το οποίο αντιστοιχεί στην τιμή τ της στατιστικής συνάρτησης T ορίζεται από την σχέση

$$\begin{aligned} \hat{\alpha} &= \max P(T \geq \tau | n, p \leq p_0) \\ &= P(T \geq \tau | n, p = p_0). \end{aligned}$$

Με όμοιο τρόπο, το κρίσιμο επίπεδο που αντιστοιχεί στην τιμή τ της στατιστικής συνάρτησης T , η οποία χρησιμοποιείται για τον έλεγχο των υποθέσεων της περίπτωσης Γ , ορίζεται από την σχέση

$$\begin{aligned} \hat{\alpha} &= \max P(T \leq \tau | n, p \geq p_0) \\ &= P(T \leq \tau | n, p = p_0). \end{aligned}$$

Το κρίσιμο επίπεδο $\hat{\alpha}$ που αντιστοιχεί στην τιμή τ της στατιστικής συνάρτησης T , η οποία χρησιμοποιείται για τον έλεγχο των υποθέσεων της περίπτωσης A , μπορεί να προσδιορισθεί αν παρατηρήσουμε ότι η περιοχή αποδοχής του ελέγχου είναι κάποια περιοχή η οποία εκτείνεται σχεδόν συμμετρικά γύρω από την τιμή np_0 . Τότε, στην περίπτωση που η παρατηρηθείσα τιμή τ της στατιστικής συνάρτησης T είναι μικρότερη ή ίση από την τιμή np_0 , δηλαδή $\tau \leq np_0$, το ήμισυ της τιμής του κρίσιμου επιπέδου $\hat{\alpha}$ προκύπτει από τον υπολογισμό της πιθανότητας $P(T \leq \tau | n, p=p_0)$. Δηλαδή, $\hat{\alpha}/2 = P(T \leq \tau | n, p=p_0)$. Ακολουθώντας την ίδια επιχειρηματολογία, θα έχουμε, στην περίπτωση που $\tau > np_0$, ότι $\hat{\alpha}/2 = P(T \geq \tau | n, p=p_0)$. Επομένως, στην περίπτωση A , το κρίσιμο επίπεδο που αντιστοιχεί στην τιμή τ της στατιστικής συνάρτησης T ορίζεται από την σχέση

$$\hat{\alpha} = \begin{cases} 2 P(T \leq \tau | n, p = p_0), & \text{αν } \tau \leq np_0 \\ 2 P(T \geq \tau | n, p = p_0), & \text{αν } \tau > np_0. \end{cases}$$

Από τον ορισμό του κρίσιμου επιπέδου $\hat{\alpha}$ ως του ελάχιστου επιπέδου σημαντικότητας, στο οποίο η μηδενική υπόθεση H_0 θα απορριπτόταν για τις δοθείσες παρατηρήσεις, συνεπάγεται ότι οι ενδείξεις από τα δεδομένα θα συνηγορούν υπέρ της απόρριψης της μηδενικής υπόθεσης H_0 σε κάθε επίπεδο σημαντικότητας μεγαλύτερο ή ίσο από το κρίσιμο επίπεδο. Δηλαδή, η μηδενική υπόθεση H_0 απορρίπτεται σε επίπεδο σημαντικότητας α , αν $\alpha \geq \hat{\alpha}$.

Παράδειγμα 2.1.1: Πέντε από τους δεκατρείς θανάτους εργατών που ζούσαν κοντά σε ένα εργοστάσιο πυρηνικής ενέργειας ήταν το αποτέλεσμα καρκίνου. Με βάση αυτό το δεδομένο, μπορεί να

υποστηριχθεί ο ισχυρισμός ότι το 25% όλων των θανάτων που συνέβησαν την χρονιά αυτή σε εργάτες που ζούσαν κοντά σε εργαστάσια πυρηνικής ενέργειας μπορούν να θεωρηθούν ως αποτέλεσμα καρκίνου; ($\alpha=0.05$)

Λύση: Θεωρούμε τις τυχαίες μεταβλητές X_1, X_2, \dots, X_{13} , όπου

$$X_i = \begin{cases} 1, & \text{αν ο } i \text{ θάνατος οφείλεται σε καρκίνο} \\ 0, & \text{διαφορετικά,} \end{cases} \quad i=1,2,\dots,13$$

Τότε, η πιθανότητα $p = P(X_i=1)$, $i=1, 2, \dots, 13$ παριστάνει το ποσοστό των θανάτων οι οποίοι οφείλονται σε καρκίνο. Ο ισχυρισμός μπορεί να διατυπωθεί με την μορφή $p=0.25$ και να ελεγχθεί μέσω των υποθέσεων

$$H_0: p = 0.25$$

$$H_1: p \neq 0.25.$$

Αν θεωρήσουμε ως δοκιμή τον θάνατο και ως επιτυχία τον θάνατο που οφείλεται σε καρκίνο, τότε η κατάλληλη στατιστική συνάρτηση T για τον έλεγχο των παραπάνω υποθέσεων είναι ο αριθμός των επιτυχιών στις 13 δοκιμές, δηλαδή

$$T = \sum_{i=1}^{13} X_i.$$

Προφανώς, η στατιστική συνάρτηση T ακολουθεί την διωνυμική κατανομή με παραμέτρους $n=13$ και $p=P(X_i=1)$. Τότε, κάτω από την μηδενική υπόθεση, η στατιστική συνάρτηση T ακολουθεί την διωνυμική κατανομή με παραμέτρους $n=13$ και $p=0.25$. Επομένως, η υπόθεση H_0 θα απορριφθεί σε επίπεδο σημαντικότητας περίπου ίσο με $\alpha=0.05$, αν $T \leq t_1$ ή $T > t_2$, όπου οι κρίσιμες τιμές t_1, t_2 ορίζονται από τις σχέσεις

$$P(T \leq t_1 \mid n=13, p=0.25) = \alpha_1 \cong 2.5\%$$

και

$$P(T > t_2 \mid n=13, p=0.25) = \alpha_2 \cong 2.5\%$$

ή, ισοδύναμα

$$P(T \leq t_2 \mid n=13, p=0.25) = 1 - \alpha_2 \cong 97.5\%.$$

Από τον πίνακα 1 του παραρτήματος, έχουμε για τις δοθείσες τιμές των παραμέτρων ($n=13$ και $p=0.25$), ότι

$$P(T \leq 0 \mid n=13, p=0.25) = 0.0238 \cong 0.025$$

και

$$P(T \leq 6 \mid n=13, p=0.25) = 0.9757 \cong 0.975.$$

Αν θεωρήσουμε, επομένως, ως α_1 την τιμή 0.0238 και ως $1 - \alpha_2$ την τιμή 0.9757, οπότε

$$\begin{aligned} \alpha_1 + \alpha_2 &= 0.0238 + (1 - 0.9757) = 0.0238 + 0.0243 \\ &= 0.0238 + 0.0243 \\ &= 0.0481 \\ &\cong 0.05, \end{aligned}$$

τότε οι τιμές 0 και 6 είναι οι κρίσιμες τιμές του ελέγχου μεγέθους $\alpha_1 + \alpha_2 = 0.0481$. Δηλαδή, $t_1 = 0$ και $t_2 = 6$. Επομένως, ο κανόνας απόφασης παίρνει την εξής μορφή:

Η μηδενική υπόθεση H_0 απορρίπτεται σε επίπεδο σημαντικότητας 0.0481 αν $T \leq 0$ ή $T > 6$.

Η παρατηρηθείσα, όμως, τιμή της στατιστικής συνάρτησης T είναι ίση με 5. Δηλαδή, $\tau = 5$. Κατά συνέπεια, σε επίπεδο σημαντικότητας ίσο με 0.0481 δεν απορρίπτεται η μηδενική υπόθεση, δηλαδή, σε επίπεδο σημαντικότητας ίσο με 0.0481 τα δεδομένα δεν συνηγορούν υπέρ της απόρριψης του ισχυρισμού.

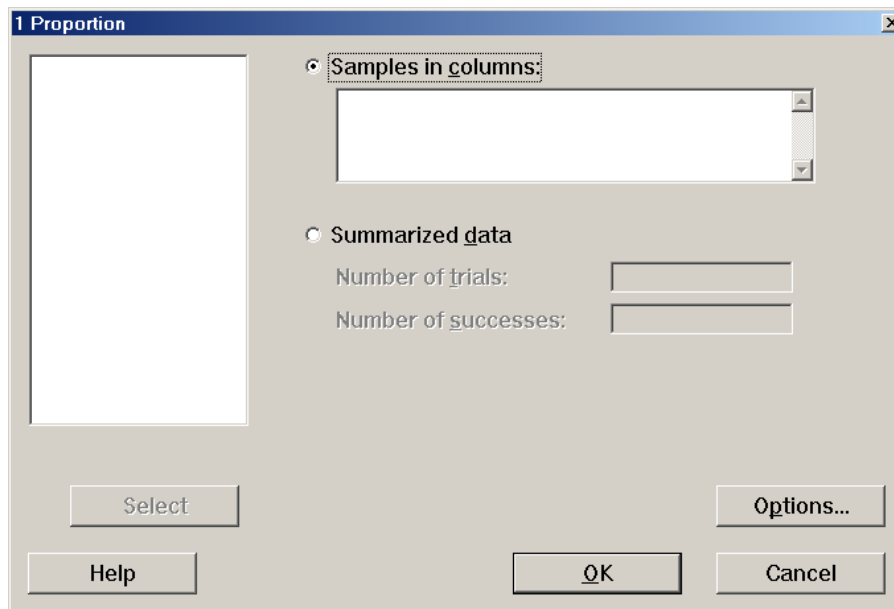
Το κρίσιμο επίπεδο $\hat{\alpha}$ του ελέγχου, το οποίο αντιστοιχεί στην τιμή $\tau=5$ της στατιστικής συνάρτησης T μπορεί να προσδιορισθεί, αν

παρατηρήσουμε ότι η περιοχή αποδοχής του ελέγχου είναι μια περιοχή γύρω από την τιμή $np_0 = (13)(0.25) = 3.25$. Επειδή η παρατηρηθείσα τιμή $\tau=5$ είναι μεγαλύτερη από την τιμή $np_0 = 3.25$, το ήμισυ της τιμής $\hat{\alpha}$ του κρίσιμου επιπέδου ισούται με την πιθανότητα

$$\begin{aligned} P(T \geq 5 \mid n=13, p=0.25) &= 1 - P(T \leq 4 \mid n=13, p=0.25) \\ &= 1 - 0.7940 \\ &= 0.206. \end{aligned}$$

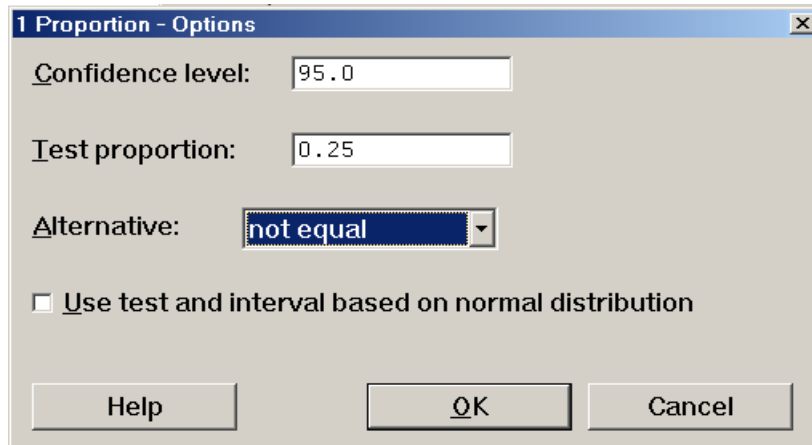
Επομένως, $\hat{\alpha} = 2(0.206) = 0.412$, οπότε η μηδενική υπόθεση H_0 θα απορρίπτεται σε επίπεδο σημαντικότητας α , αν $\alpha \geq 0.412$.

Λύση με το MINITAB: Ο διωνυμικός έλεγχος είναι εξαιρετικά εύκολος με το MINITAB 12.0. Δεν χρειάζεται να καταχωρίσουμε το δείγμα μας σε κάποια μεταβλητή του προγράμματος. Απλώς, επιλέγουμε κατά σειρά **Stat, Basic Statistics, 1 Proportion** από το βασικό μενού και προκύπτει το παρακάτω πλαίσιο διαλόγου:



Στο πεδίο **Summarized data**, μπορούμε να δηλώσουμε απ' ευθείας τον αριθμό δοκιμών (πεδίο **Number of trials**) και τον αριθμό επιτυχιών

(πεδίο **Number of successes**). Μετά, πιέζουμε το πλήκτρο **Options** και εμφανίζεται το εξής πλαίσιο:



The screenshot shows the '1 Proportion - Options' dialog box in Minitab. It contains the following fields and options:

- Confidence level:** 95.0
- Test proportion:** 0.25
- Alternative:** not equal (selected from a dropdown menu)
- Use test and interval based on normal distribution**

Buttons at the bottom: Help, OK, Cancel.

Στο πεδίο **Confidence level**, δηλώνουμε την τιμή του $1-\alpha$, όπου α το επίπεδο σημαντικότητας του ελέγχου. Για το παράδειγμα μας δηλαδή δίνουμε 0.95. Στο πεδίο **Test proportion**, δίνουμε την τιμή του p_0 . Στην περίπτωση μας, δίνουμε 0.25. Τέλος, στο πεδίο **Alternative**, δηλώνουμε την κατεύθυνση της εναλλακτικής υπόθεσης. Για αμφίπλευρο έλεγχο, δηλώνουμε **not equal**. Πιέζοντας δύο φορές **OK**, παίρνουμε τα εξής αποτελέσματα:

Test and Confidence Interval for One Proportion

Test of $p = 0.25$ vs $p \text{ not } = 0.25$

Exact					
Sample	X	N	Sample p	95.0 % CI	P-Value
1	5	13	0.384615	(0.138579, 0.684222)	0.333

Το MINITAB μας δίνει εκτίμηση του ποσοστού επιτυχίας, 95% διάστημα εμπιστοσύνης και την τιμή του κρίσιμου επιπέδου (P-value). Παρατηρούμε ότι, για τα δεδομένα του παραδείγματός μας, η τιμή του κρίσιμου επιπέδου είναι κατά πολύ μεγαλύτερη του 0.05 και συνεπώς η μηδενική υπόθεση φαίνεται να είναι μία εύλογη υπόθεση. Το ίδιο

συμπέρασμα προκύπτει από το γεγονός ότι το 95% διάστημα εμπιστοσύνης περιέχει την τιμή 0.25.

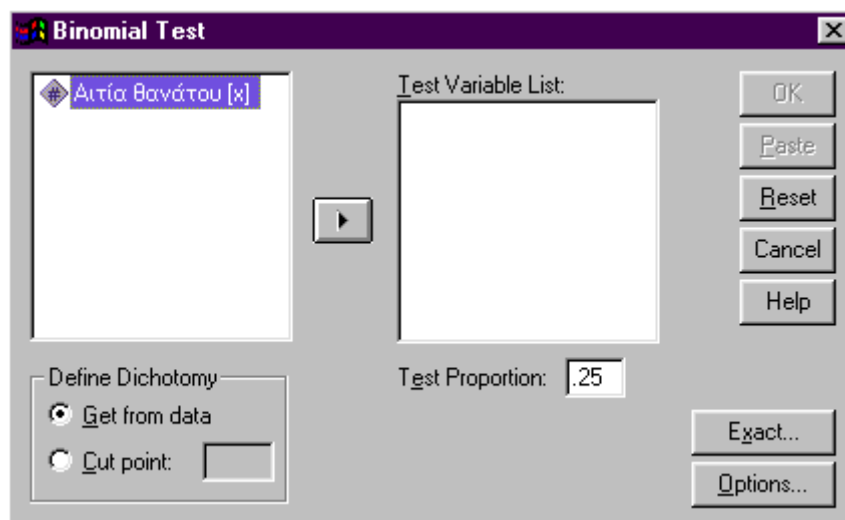
Σημείωση: Παρατηρούμε ότι η τιμή του κρίσιμου επιπέδου (p-value), που το MINITAB δίνει για το αμφίπλευρο έλεγχο, διαφέρει πολύ από αυτό που υπολογίσθηκε στην αναλυτική λύση. Αυτό συμβαίνει γιατί το MINITAB χρησιμοποιεί διαφορετικό τρόπο υπολογισμού. Στην περίπτωση του παραδείγματός μας έχει κάνει το εξής: Υπολόγισε την πιθανότητα τουλάχιστον 5 επιτυχιών σε 13 δοκιμές ίση με 0.206. Θεωρώντας ότι η τιμή αυτή είναι μικρότερη ή ίση του μισού κρίσιμου επιπέδου, αναζήτησε την μεγαλύτερη τιμή k για την οποία η πιθανότητα να έχουμε k ή λιγότερες επιτυχίες σε 13 δοκιμές είναι μικρότερη ή ίση του 0.206. Η τιμή αυτή είναι $k=1$ και η αντίστοιχη πιθανότητα είναι ίση με 0.1267. Κατά συνέπεια, το κρίσιμο επίπεδο είναι $0.206+0.1267=0.333$.


Σημείωση: Παρατηρούμε ότι, στο δεύτερο πλαίσιο διαλόγου, υπάρχει και η επιλογή **Use test and interval based on normal distribution**, η οποία δίνει την δυνατότητα χρησιμοποίησης της κανονικής προσέγγισης διωνυμικής κατανομής της ελεγχοσυνάρτησης.

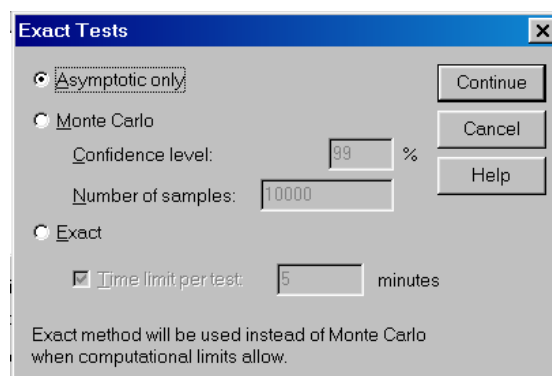
Σημείωση: Αν θέλουμε να καταχωρίσουμε το δείγμα μας σε μία μεταβλητή, οι τιμές του πρέπει να είναι κωδικοποιημένες (π.χ. 0 και 1). Το MINITAB θεωρεί ότι η μεγαλύτερη από τις δύο κωδικές τιμές δηλώνει την επιτυχία (π.χ. αν έχουμε τιμές -2 και -1 , η τιμή -1 θα θεωρηθεί επιτυχία).

Λύση με το SPSS: Για να κάνουμε διωνυμικό έλεγχο με το SPSS 10.1, καταχωρίζουμε το δείγμα μας σε μια nominal μεταβλητή. Κάθε παρατήρηση του δείγματος καταχωρίζεται ως κωδικός αριθμός. Έχουμε δύο κωδικούς αριθμούς, έναν για να συμβολίζουμε την «επιτυχία» και έναν για την «αποτυχία». Οι παρατηρήσεις μπορούν να καταχωρισθούν με όποια σειρά θέλουμε. Θα πρέπει να τονισθεί, ότι το SPSS θεωρεί ως «επιτυχία» την πρώτη καταχώριση που συναντά στη μεταβλητή.

Στο συγκεκριμένο παράδειγμα, συμβολίζουμε την «επιτυχία» (θάνατος από καρκίνο) με 0 και την «αποτυχία» (θάνατος από άλλη αιτία) με 1. Δημιουργούμε μια μεταβλητή με όνομα **x** και «ετικέτα» **Αιτία θανάτου**. Καταχωρίζουμε τις τιμές 0,0,0,0,0,1,1,1,1,1,1,1 στην μεταβλητή **x**. Η σειρά δεν παίζει ρόλο εφόσον έχουμε 5 μηδενικά, 8 μονάδες και η πρώτη τιμή είναι 0. Το ίδιο θα ίσχυε αν είχαμε 5 μονάδες, 8 μηδενικά με την πρώτη τιμή ίση με 1. Για να διεξαγάγουμε διωνυμικό έλεγχο, επιλέγουμε **Analyze, Nonparametric Tests, Binomial** και οδηγούμεθα στο παρακάτω πλαίσιο διαλόγου:



Επιλέγουμε την μεταβλητή x και την «στέλνουμε» στο πεδίο **Test Variable List** πιέζοντας το . Στο πεδίο **Define Dichotomy**, επιλέγουμε **Get from data** αφού η μεταβλητή περιέχει μόνο δύο δυνατές τιμές. Στο πεδίο **Test Proportion**, δηλώνουμε την τιμή p_0 που αντιστοιχεί στην μηδενική υπόθεση, δηλαδή 0.25. Στην συνέχεια, πιέζοντας **Exact**, οδηγούμεθα στο εξής πλαίσιο διαλόγου:



Παρατηρούμε ότι έχουμε την επιλογή να διεξάγουμε τον έλεγχο χρησιμοποιώντας την ασυμπτωτική προσέγγιση της διωνυμικής κατανομής από την κανονική (επιλογή **Asymptotic only**) ή χρησιμοποιώντας την διωνυμική κατανομή (επιλογή **Exact**). Στο παράδειγμα μας, το δείγμα είναι μικρό. Για τον λόγο αυτό, επιλέγουμε **Exact**. Υπάρχει επίσης η επιλογή **Time limit per test**. Αυτή χρησιμεύει σε περιπτώσεις που ο υπολογισμός του κρίσιμου επιπέδου (p -value) είναι χρονοβόρος όταν η μορφή της κατανομής της στατιστικής συνάρτησης ελέγχου είναι πολύπλοκη. Στη δική μας περίπτωση, ο υπολογισμός της p -τιμής δεν απαιτεί πολύ χρόνο, οπότε αφήνουμε την διάρκεια στα 5 λεπτά. Πιέζοντας κατά σειρά **Continue** και **OK**, παίρνουμε τα παρακάτω αποτελέσματα:

Binomial Test

		Category	N	Observed	Test Prop.	Exact Sig.	Point
--	--	----------	---	----------	------------	------------	-------

				Prop.		(1-tailed)	Probability
Αιτία θανάτου	Group 1	Καρκίνος	5	.38	.25	.206	.126
	Group 2	Άλλη	8	.62			
	Total		13	1.00			

Παρατηρούμε ότι το ποσοστό «επιτυχιών» στο δείγμα μας είναι 0.38 (στήλη **Observed Proportion**). Το κρίσιμο επίπεδο για μονόπλευρο έλεγχο (υποσημείωση **1-tailed** στην στήλη **Exact Significance**) είναι 0.206. Η τιμή αυτή αντιστοιχεί στην μέγιστη πιθανότητα του ενδεχομένου $T \geq 5$ κάτω από την H_0 . Για τον αμφίπλευρο έλεγχο του παραδείγματός μας, χρειάζεται να διπλασιάσουμε την τιμή αυτή, οπότε η τιμή του κρίσιμου επιπέδου προκύπτει ίση με 0.412.

Σημείωση: Ανάλογα με το αν η τιμή **Observed Proportion** είναι μεγαλύτερη ή μικρότερη του p_0 (στήλη **Test Proportion**), το SPSS υπολογίζει το κρίσιμο επίπεδο για τον αντίστοιχης κατεύθυνσης μονόπλευρο έλεγχο. Η στήλη **Point Probability** εμφανίζεται μόνο αν έχουμε επιλέξει Exact και δίνει (στην περίπτωση μας) την πιθανότητα 5 «επιτυχιών» σε 13 «δοκιμές» με πιθανότητα «επιτυχίας» ίση με 0.25.

Λύση με το SAS: Η έκδοση του SAS που χρησιμοποιείται εδώ, δεν διεξάγει το διωνυμικό έλεγχο. Παρέχει όμως ένα ασυμπτωτικό έλεγχο.

Στο παράθυρο εντολών του SAS πληκτρολογούμε τις εντολές:

```
data cancer;
input x @@;
cards;
1 1 1 1 1 2 2 2 2 2 2 2 2
;
run;

proc freq;
tables x / testp=(0.25 0.75);
run;
```

Η εντολή **proc freq**; μας δίνει τον πίνακα συχνοτήτων της υπό εξέταση μεταβλητής ενώ η υποεντολή **testp=(0.25 0.75)**; διεξάγει τον έλεγχο περί ισότητας των ποσοστών που αντιστοιχούν σε κάθε κατηγορία με τις τιμές που δίδονται (στο συγκεκριμένο παράδειγμα εξετάζεται αν το ποσοστό θανάτων από τη συγκεκριμένη αιτία στον υπό εξέταση πληθυσμό είναι 0.25). Θα πρέπει να σημειωθεί ότι, στην παρένθεση της υποεντολής, εισάγονται τόσες τιμές όσες είναι οι κατηγορίες της μεταβλητής που εξετάζουμε, ενώ το άθροισμα των τιμών αυτών θα πρέπει να ισούται με 1.

Επίσης, ένας εναλλακτικός τρόπος εισαγωγής των δεδομένων στην περίπτωση αυτή θα ήταν να χρησιμοποιήσουμε μεταβλητή στάθμισης. Έτσι, αντί των εντολών που χρησιμοποιήθηκαν παραπάνω, μπορούμε να πληκτρολογήσουμε τις εξής εντολές (το αποτέλεσμα είναι ταυτόσημο).

```
data cancer;
x=1; f=5; output;
x=2; f=8; output;
run;
proc freq;
weight f;
tables x / testp=(0.25 0.75);
run;
```

Προφανώς, ο δεύτερος αυτός τρόπος μας διευκολύνει ιδιαίτερα αν ο αριθμός των δεδομένων είναι μεγάλος.

Το αποτέλεσμα του ελέγχου δίδεται στον πίνακα που ακολουθεί.

The SAS System					
Cumulati ve				Test	Cumul ati ve
Percent	X	Frequency	Percent	Percent	Frequency
38.5	1	5	38.5	25.0	5
100.0	2	8	61.5	75.0	13

Chi-Square Test for Specified Proportions		
Statistic = 1.256	DF = 1	Prob = 0.262

WARNING: 50% of the cells have expected counts less than 5. Chi-Square may not be a valid test.

Ο έλεγχος χ^2 που το πακέτο χρησιμοποιεί ως προσέγγιση του διωνυμικού δεν δίνει αξιόπιστα αποτελέσματα όταν οι αναμενόμενες συχνότητες των δεδομένων είναι μικρότερες του 5. Στο συγκεκριμένο παράδειγμα, η αναμενόμενη συχνότητα των θανάτων από καρκίνο είναι 3.25 (αναμενόμενο ποσοστό 0.25 επί μέγεθος δείγματος 13). Έτσι, το πακέτο διευκρινίζει ότι ο έλεγχος ενδέχεται να μην είναι αξιόπιστος.

Το πεδίο **statistic** παρέχει την τιμή της ελεγχουσυνάρτησης χ^2 , η οποία για το συγκεκριμένο παράδειγμα είναι 1.256, το πεδίο **DF** δίνει τους βαθμούς ελευθερίας, ενώ το πεδίο **Prob** δίνει την τιμή του κρίσιμου επιπέδου που είναι 0.262. Όπως τονίστηκε και προηγουμένως, παρατηρούμε ότι το παρατηρούμενο επίπεδο στατιστικής σημαντικότητας του ασυμπτωτικού ελέγχου απέχει αρκετά από το αντίστοιχο παρατηρούμενο επίπεδο στατιστικής σημαντικότητας του διωνυμικού ελέγχου που είναι 0.412. Βέβαια, ακόμη και με τη χρήση του ασυμπτωτικού ελέγχου δεν θα απορρίπταμε την αρχική υπόθεση ότι το πραγματικό ποσοστό στον υπό εξέταση πληθυσμό μπορεί να είναι 0.25, σε κάθε σύνηθες επίπεδο σημαντικότητας α .

Επίσης, θα πρέπει να σημειωθεί ότι νεότερες εκδόσεις του SAS (έκδοση 7 και μεταγενέστερες) διεξάγουν τον διωνυμικό έλεγχο χωρίς να χρησιμοποιούν προσέγγιση. Αυτό, στο παράδειγμα μας θα υπολογιζόταν με τις εντολές

```
data cancer;  
  x=1; f=5; output;  
  x=2; f=8; output;  
run;
```

```

proc freq;
  weight f;
  table x / binomial(p=.25);
run;

```

Παράδειγμα 2.1.2: Για να μελετήσει τα αποτελέσματα της κόπωσης, ένας ερευνητής δίδαξε 18 παιδιά δύο διαφορετικούς τρόπους να δένουν τον ίδιο κόμπο. Τα μισά από τα παιδιά (που επελέγησαν τυχαία από την ομάδα των 18 παιδιών) διδάχθηκαν πρώτα τη μέθοδο Α και τα υπόλοιπα παιδιά διδάχθηκαν πρώτα τη μέθοδο Β. Στο τέλος μιας κουραστικής σχολικής μέρας και αμέσως μετά από ένα διαγώνισμα στα Μαθηματικά, ζητήθηκε από κάθε ένα από τα 18 παιδιά να δέσουν τον κόμπο που έμαθαν να δένουν. Η πρόβλεψη του ερευνητή ήταν ότι η κόπωση θα προκαλούσε *παλινδρόμηση*, δηλαδή τα παιδιά θα παλινδρομούσαν στην πρώτη μέθοδο δεσίματος του κόμπου που έμαθαν. Κάθε παιδί ταξινομήθηκε σε μία από δύο κατηγορίες, ανάλογα με το εάν χρησιμοποίησε την μέθοδο που έμαθε πρώτα ή την μέθοδο που έμαθε τελευταία όταν του ζητήθηκε να δέσει τον κόμπο. Το αποτέλεσμα του πειράματος ήταν ότι όλα τα παιδιά εκτός από δύο χρησιμοποίησαν τη μέθοδο που έμαθαν πρώτα. Είναι εύλογη η πρόβλεψη του ερευνητή; ($\alpha=0.05$).

Λύση: Έστω p η πιθανότητα που έχει ένα παιδί να χρησιμοποιήσει την μέθοδο που διδάχθηκε πρώτα. Είναι φανερό ότι, για κάθε παιδί, υπάρχουν δύο δυνατές περιπτώσεις (δυνατά αποτελέσματα) κατά το δέσιμο του κόμπου (δοκιμή): Να χρησιμοποιηθεί η πρώτη μέθοδος που το παιδί διδάχθηκε (επιτυχία) και να χρησιμοποιηθεί η δεύτερη μέθοδος που το παιδί διδάχθηκε (αποτυχία). Τα δεδομένα του πειράματος, επομένως, μπορούν να παρασταθούν με την ακολουθία των τυχαίων μεταβλητών X_1, X_2, \dots, X_{18} , όπου, προφανώς,

$$X_i = \begin{cases} 1, & \text{αν το παιδί χρησιμοποίησε την μέθοδο που διδάχθηκε πρώτα} \\ 0, & \text{διαφορετικά} \end{cases}$$

και $P(X_i = 1) = p$, $i = 1, 2, \dots, 18$. Ο ισχυρισμός, επομένως, του ερευνητή μπορεί να διατυπωθεί με την μορφή $p > 1 - p$ ή, ισοδύναμα, $p > 0.5$ (η θεωρία της παλινδρόμησης είναι ορθή). Μία εύλογη υπόθεση που θα αντιτίθετο στην λογική της πρόβλεψης του ερευνητή, θα ήταν ότι η κόπωση δεν έχει καμία επίδραση στις πιθανότητες με τις οποίες το παιδί διαλέγει τη μία ή την άλλη μέθοδο ($p = 1 - p$ ή, ισοδύναμα, $p = 0.5$).

Επομένως, ένα κατάλληλο ζεύγος υποθέσεων για τον έλεγχο της υπόθεσης του ερευνητή είναι το εξής:

$$H_0 : p = 0.5$$

$$H_1 : p > 0.5.$$

Η κατάλληλη στατιστική συνάρτηση για τον έλεγχο των παραπάνω υποθέσεων είναι ο αριθμός των παιδιών που διάλεξαν να χρησιμοποιήσουν την πρώτη μέθοδο που διδάχθηκαν, δηλαδή

$$T = \sum_{i=1}^{18} X_i.$$

Η κατανομή της στατιστικής αυτής συνάρτησης είναι προφανώς η διωνυμική με παραμέτρους $n=18$ και p . Επομένως, κάτω από τη μηδενική υπόθεση, η στατιστική συνάρτηση T ακολουθεί την διωνυμική κατανομή με παραμέτρους $n=18$ και $p=0.5$.

Από την μορφή της εναλλακτικής υπόθεσης, είναι σαφές ότι οι μεγάλες τιμές της στατιστικής συνάρτησης T θα συνηγορούν υπέρ της απόρριψης της μηδενικής υπόθεσης.

Τότε, η υπόθεση H_0 θα απορρίπτεται σε επίπεδο σημαντικότητας περίπου ίσο με την τιμή $\alpha=0.05$, αν $T>t$, όπου η κρίσιμη τιμή t ορίζεται από τη σχέση

$$P(T > t \mid n = 18, p = 0.5) = \alpha_1 \cong \alpha \equiv 0.05,$$

ή, ισοδύναμα

$$P(T \leq t \mid n = 18, p = 0.5) = 1 - \alpha_1 \cong 1 - \alpha \equiv 0.95.$$

Από τον πίνακα 1 του παραρτήματος, για τις δοθείσες τιμές των παραμέτρων n και p , παρατηρούμε ότι

$$P(T \leq 12 \mid n = 18, p = 0.5) = 0.9519.$$

Αν, λοιπόν, θέσουμε $1 - \alpha_1 = 0.9519$ ή, ισοδύναμα, $\alpha_1 = 0.0481$, η τιμή $\tau = 12$ αποτελεί την κρίσιμη τιμή ελέγχου των υποθέσεών μας που έχει μέγεθος $\alpha_1 = 0.0481 \cong \alpha = 0.05$. Επομένως, ο κανόνας απόφασης διαμορφώνεται ως εξής:

Η υπόθεση H_0 απορρίπτεται σε επίπεδο σημαντικότητας $\alpha_1 = 0.0481$, αν η $\tau = 16$. Η τιμή της στατιστικής συνάρτησης T η οποία παρατηρήθηκε, είναι η $\tau = 16$. Η τιμή αυτή υπερβαίνει την κρίσιμη τιμή και, επομένως, η μηδενική υπόθεση δεν είναι εύλογη σε επίπεδο σημαντικότητας $\alpha_1 = 0.0481$. Επομένως, με πιθανότητα σφάλματος τύπου I ίση με 0.0481 οι ενδείξεις από το δείγμα υποστηρίζουν την θεωρία του ερευνητή.

Όπως εύκολα μπορούμε να διαπιστώσουμε από τον πίνακα 1 του παραρτήματος, η τιμή του κρίσιμου επιπέδου είναι πάρα πολύ μικρή. Πράγματι,

$$\begin{aligned}
\hat{\alpha} &= \max P(T \geq \tau \mid H_0) \\
&= P(T \geq 16 \mid n = 18, p = 0.5) \\
&= 1 - P(T \leq 15 \mid n = 18, p = 0.5) \\
&= 1 - 0.9993 \\
&= 0.0007 < 0.1\%.
\end{aligned}$$

Επομένως, με βάση τα δεδομένα του προβλήματος, θα οδηγούμεθα σε απόρριψη της μηδενικής υπόθεσης σε κάθε επίπεδο σημαντικότητας α για το οποίο ισχύει ότι $\alpha \geq 0.0007$.

Λύση με το MINITAB: Το παράδειγμα επιλύεται με διωνυμικό έλεγχο πολύ εύκολα. Έχουμε 16 «επιτυχίες» σε 18 δοκιμές με εναλλακτική υπόθεση ότι το ποσοστό «επιτυχίας» είναι μεγαλύτερο του 0.5. Τα αποτελέσματα του ελέγχου είναι τα εξής:

Test and Confidence Interval for One Proportion

Test of $p = 0.5$ vs $p > 0.5$

Sample	X	N	Sample p	95.0 % CI	Exact P-Value
1	16	18	0.888889	(0.652880, 0.986249)	0.001

Η τιμή του κρίσιμου επιπέδου είναι πολύ μικρή και συνεπώς η H_0 δεν μπορεί να θεωρηθεί εύλογη.

Λύση με το SPSS: Για να επιλύσουμε με το SPSS το παράδειγμα αυτό, πρέπει να δημιουργήσουμε μία μεταβλητή (έστω x) που θα περιέχει 16 μονάδες («επιτυχίες») και δύο μηδενικά («αποτυχίες»). Η πρώτη τιμή της μεταβλητής πρέπει να είναι 1 για να «καταλάβει» το SPSS ότι αυτή είναι η κωδική τιμή της «επιτυχίας». Τα αποτελέσματα του ελέγχου είναι:

Binomial Test

	Category	N	Observed	Test Prop.	Exact Sig.	Point

				Prop.		(2-tailed)	Probability
X	Group 1	1	16	.89	.50	.001	.001
	Group 2	0	2	.11			
	Total		18	1.00			

Το πρόγραμμα έχει διεξαγάγει αμφίπλευρο έλεγχο του οποίου το κρίσιμο επίπεδο είναι 0.001. Για τον μονόπλευρο έλεγχο του παραδείγματος, το κρίσιμο επίπεδο προκύπτει, με διαίρεση της τιμής αυτής με 2, ίσο με 0.0005.

Λύση με το SAS: Και σε αυτό το παράδειγμα, το SAS θα χρησιμοποιήσει τον ασυμπτωτικό έλεγχο αντί του διωνυμικού. Στο παράθυρο εντολών πληκτρολογούμε

```
data regress;
input x @@;
cards;
1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
;
run;

proc freq;
tables x /testp=(0.5 0.5);
run;
```

Το αποτέλεσμα περιέχεται στον πίνακα που ακολουθεί.

```
The SAS System

      Test Cumulative Cumulative
      X  Frequency Percent Percent Frequency Percent
-----
1     2         11.1     50.0     2         11.1
2    16         88.9     50.0    18        100.0

Chi-Square Test for Specified Proportions
-----
Statistic = 10.889    DF = 1    Prob = 0.001
```

Θα πρέπει να δοθεί προσοχή στο ότι η τιμή του κρίσιμου επιπέδου, που παρέχεται στο πεδίο **Prob**, αντιστοιχεί σε αμφίπλευρη εναλλακτική υπόθεση. Για τον μονόπλευρο έλεγχο του παραδείγματος μας θα

προκύψει με διαίρεση τις τιμές αυτής με 2. Έτσι, βάσει της πολύ χαμηλής τιμής του παρατηρούμενου επιπέδου στατιστικής σημαντικότητας απορρίπτουμε την αρχική υπόθεση σε κάθε σύνηθες επίπεδο σημαντικότητας α .

Παρατήρηση: Όπως είναι γνωστό, όταν η παράμετρος n ενός διωνυμικού πληθυσμού έχει μία αρκετά μεγάλη τιμή ($n > 20$) και η κατανομή πιθανότητάς του είναι περίπου συμμετρική ($np \geq 5$ και $n(1-p) \geq 5$), η διωνυμική κατανομή που περιγράφει τον πληθυσμό αυτό μπορεί να προσεγγισθεί ικανοποιητικά από μία κανονική κατανομή με μέση τιμή και διασπορά την μέση τιμή και την διασπορά αντίστοιχα της διωνυμικής κατανομής. Δηλαδή, στην περίπτωση αυτή, η διωνυμική κατανομή με παραμέτρους n και p προσεγγίζεται από μία κανονική κατανομή με παραμέτρους $\mu = np$ και $\sigma^2 = np(1-p)$. Τότε, στα προβλήματα ελέγχου υποθέσεων, η κρίσιμη περιοχή ενός ελέγχου μεγέθους περίπου ίσου με α μπορεί να προσδιορισθεί κατά προσέγγιση με βάση την συνάρτηση κατανομής της κανονικής κατανομής. Συγκεκριμένα, αν p_0 είναι η μηδενική τιμή της παραμέτρου p της διωνυμικής κατανομής (δηλαδή η τιμή της p κάτω από την μηδενική υπόθεση H_0), αποδεικνύεται ότι το κρίσιμο σημείο t ενός ελέγχου μεγέθους περίπου ίσου με α δίνεται από την σχέση

$$t = np_0 + z_r \sqrt{np_0(1-p_0)},$$

όπου z_r παριστάνει το κατάλληλο r -ποσοστιαίο σημείο της τυποποιημένης κανονικής κατανομής ($r = \alpha/2, \alpha, 1-\alpha/2, 1-\alpha$).

Παράδειγμα 2.1.3: Σύμφωνα με τον απλό νόμο της κληρονομικότητας του Mendel, η διασταύρωση μεταξύ φυτών δύο συγκεκριμένων γενοτύπων, ενδέχεται να οδηγήσει σε *απογόνους*, οι οποίοι κατά το

1/4 είναι *νάνοι* και κατά τα 3/4 είναι *κανονικοί*. Σε ένα πείραμα, για να προσδιορισθεί κατά πόσον η υπόθεση του απλού νόμου της κληρονομικότητας του Mendel είναι εύλογη σε μία συγκεκριμένη περίπτωση, μία διασταύρωση οδήγησε σε *απογόνους* φυτά από τα οποία τα 243 ήταν *νάνοι* και τα 682 ήταν *κανονικά*.. Πώς θα ερμηνεύατε τα αποτελέσματα του πειράματος αυτού σε επίπεδο σημαντικότητας περίπου ίσο με $\alpha=0.05$;

(Παν/μιο Κρήτης, Εξετάσεις Δεκεμβρίου 1984)

Λύση: Αν θεωρήσουμε ως επιτυχία ένα φυτό-*απόγονο* που είναι *κανονικό*, και συμβολίσουμε την πιθανότητα *επιτυχίας* με p , η μηδενική υπόθεση του απλού νόμου της κληρονομικότητας του Mendel είναι ισοδύναμη με την υπόθεση

$$H_0: p = 3/4.$$

Η εναλλακτική που μας ενδιαφέρει έχει την μορφή

$$H_1: p \neq 3/4.$$

Ο αριθμός των *δοκιμών* εδώ είναι $n=243+682=925$. Επομένως, η στατιστική συνάρτηση T , που ορίζεται ως ο αριθμός των επιτυχιών στις 925 δοκιμές, ακολουθεί την διωνυμική κατανομή με παραμέτρους $n=925$ και p . Δηλαδή, κάτω από την μηδενική υπόθεση, η κατανομή της στατιστικής συνάρτησης T είναι η διωνυμική με παραμέτρους $n=925$, $p=p_0=0.75$. Είναι, δηλαδή, η κατανομή της στατιστικής συνάρτησης T κάτω από την μηδενική υπόθεση περίπου συμμετρική αφού $np_0 = 693.75 \geq 5$ και $n(1-p_0)=231.25 \geq 5$. Επομένως, οι κρίσιμες τιμές t_1 , t_2 του ελέγχου των παραπάνω υποθέσεων μεγέθους περίπου ίσου με α , προσδιορίζονται από τις σχέσεις

$$\alpha/2 = P(T \leq t_1 \mid n = 925, p = p_0 = 0.75)$$

$$\cong P(Z \leq \frac{t_1 - \mu}{\sigma} \mid \mu = np_0 = 693.75, \sigma^2 = np_0(1 - p_0) = 231.25)$$

και

$$1 - \alpha/2 = P(T \leq t_2 \mid n = 925, p = p_0 = 0.75)$$

$$\cong P\left(Z \leq \frac{t_2 - \mu}{\sigma} \mid \mu = np_0 = 693.75, \sigma^2 = np_0(1 - p_0) = 231.25\right),$$

αντίστοιχα, όπου Z παριστάνει μια τυποποιημένη κανονική τυχαία μεταβλητή. Από τις σχέσεις αυτές έπεται ότι

$$\frac{t_1 - \mu}{\sigma} = z_{\alpha/2} \quad \text{και} \quad \frac{t_2 - \mu}{\sigma} = z_{1-\alpha/2},$$

δηλαδή,

$$\begin{aligned} t_1 &= \sigma z_{\alpha/2} + \mu \\ &= np_0(1 - p_0)z_{\alpha/2} + np_0 \\ &= 231.25z_{0.025} + 693.75 \end{aligned}$$

και

$$\begin{aligned} t_2 &= \sigma z_{1-\alpha/2} + \mu \\ &= np_0(1 - p_0)z_{1-\alpha/2} + np_0 \\ &= 231.25z_{0.975} + 693.75 \end{aligned}$$

Από τον πίνακα 2 της τυποποιημένης κανονικής κατανομής, προκύπτει ότι

$$z_{1-\alpha/2} = z_{0.975} = 1.96$$

$$z_{\alpha/2} = z_{0.025} = -z_{0.975} = -1.96.$$

Επομένως, οι προσεγγίσεις των κρίσιμων τιμών του ελέγχου των υποθέσεών μας προκύπτουν να είναι οι τιμές $t_1 = 667.94$ και $t_2 = 719.56$, αντίστοιχα. Η τιμή της στατιστικής συνάρτησης T , η οποία παρατηρήθηκε στο πείραμα αυτό, είναι $\tau = 682$. Επομένως, η μηδενική υπόθεση δεν απορρίπτεται σε επίπεδο σημαντικότητας 0.05. Αρα, με πιθανότητα σφάλματος τύπου I ίση με 0.05, τα αποτελέσματα

του πειράματος μπορούν να θεωρηθούν ως ενδείξεις υπέρ της απλής υπόθεσης κληρονομικότητας του Mendel.

Το κρίσιμο επίπεδο $\hat{\alpha}$ του ελέγχου μπορεί να προσδιορισθεί κατά προσέγγιση από την σχέση που ακολουθεί, λαβαίνοντας υπόψη ότι η παρατηρηθείσα τιμή της στατιστικής συνάρτησης T είναι $\tau = 682 \leq 693.75 = np_0$. Συγκεκριμένα,

$$\begin{aligned} \hat{\alpha}/2 &= P(T \leq \tau \mid n, p = p_0) \\ &= P(T \leq 682 \mid n = 925, p = 0.75) \\ &\approx P\left(Z < \frac{682 - \mu}{\sigma} \mid \mu = 693.75, \sigma^2 = 231.25\right) \\ &= \Phi(-0.8922) \\ &\approx 0.186. \end{aligned}$$

Δηλαδή, για την απόρριψη της μηδενικής υπόθεσης, θα χρειαζόταν ένας έλεγχος μεγέθους τουλάχιστον ίσου με 0.372. Θα χρειαζόταν, δηλαδή, ένα επίπεδο σημαντικότητας τουλάχιστον ίσο με 0.372 για να απορριφθεί η μηδενική υπόθεση ($\hat{\alpha} = 2(0.186) = 0.372$). Κατά συνέπεια, τα δεδομένα παρέχουν ισχυρές ενδείξεις υπέρ της μηδενικής υπόθεσης.

Λύση με το MINITAB: Τα αποτελέσματα που δίνει το MINITAB με χρήση της κανονικής προσέγγισης της διωνυμικής κατανομής περιέχονται στον εξής πίνακα:

Test and Confidence Interval for One Proportion

Test of $p = 0.75$ vs $p \text{ not } = 0.75$

Sample	X	N	Sample p	95.0 % CI	Z-Value	P-Value
1	682	925	0.737297	(0.708936, 0.765659)	-0.89	0.372

Λύση με το SPSS: Για να διεξαγάγουμε διωνυμικό έλεγχο, πρέπει να καταχωρίσουμε σε μια μεταβλητή 682 μηδενικά και 243 άσσους με την πρώτη τιμή της μεταβλητής να είναι 0. Αυτό γίνεται εύκολα με **Copy**,

Paste. Αν υποθέσουμε ότι η μεταβλητή ονομάζεται x και ζητήσουμε διωνυμικό έλεγχο με χρήση της διωνυμικής κατανομής (επιλογή **Exact**), οδηγούμεθα στα εξής αποτελέσματα:

Binomial Test

		Category	N	Observed Prop.	Test Prop.	Asymp. Sig. (1-tailed)	Exact Sig. (1-tailed)
X	Group 1	Κανονικός	682	.74	.75	.196 ^{a,b}	.196
	Group 2	Νάνος	243	.26			
	Total		925	1.00			

a Alternative hypothesis states that the proportion of cases in the first group < .75.

b Based on Z Approximation.

Παρατηρούμε ότι το SPSS από μόνο του έχει υπολογίσει και την ακριβή αλλά και την ασυμπτωτική τιμή του κρίσιμου επιπέδου (p -value) επειδή έχουμε μεγάλο δείγμα. Όπως είναι αναμενόμενο, οι δύο τιμές συμφωνούν. Για τον αμφίπλευρο έλεγχο του παραδείγματός μας, η τιμή του κρίσιμου επιπέδου προκύπτει, με διπλασιασμό της τιμής 0.196, ίση με 0.392.

Σημείωση: Για τον υπολογισμό της τιμής της ελεγχουσυνάρτησης Z , το SPSS προσέθεσε στον αριθμητή την τιμή 0.5 (διόρθωση συνέχειας του Yates).

Λύση με το SAS: Εισάγουμε τα δεδομένα σε αρχείο με όνομα **mendel.prn** σε μία στήλη που περιέχει 236 τιμές ίσες με 1 και 682 τιμές ίσες με 2 (μπορεί να χρησιμοποιηθεί οποιαδήποτε κατάληξη αρχείου, αρκεί να δηλωθεί το πλήρες όνομα και στην εντολή **infile** του SAS). Οι εντολές που χρειάζονται για την διεξαγωγή του ασυμπτωτικού ελέγχου είναι

```
data mendel;
infile 'c:\mendel.prn';
input x;
```

```

run;
proc print;
run;

proc freq;
tables x /testp=(0.25 0.75);
run;

```

Βέβαια, αντί να καταχωρίσουμε τα δεδομένα σε κάποιο αρχείο, θα μπορούσαμε, εναλλακτικά, να χρησιμοποιήσουμε μια μεταβλητή στάθμισης (βλ. Παράδειγμα 2.1.1). Στη περίπτωση αυτή, όπου το πλήθος των δεδομένων είναι μεγάλο, ο τρόπος αυτός θα μας διευκόλυνε σημαντικά.

Το αποτέλεσμα περιέχεται στον πίνακα που ακολουθεί

The SAS System					
X	Frequency	Test Percent	Cumulative Percent	Cumulative Frequency	Percent
1	243	26.3	25.0	243	26.3
2	682	73.7	75.0	925	100.0

Chi-Square Test for Specified Proportions

 Statistic = 0.796 DF = 1 Prob = 0.372

Παρατηρούμε ότι το κρίσιμο επίπεδο του ασυμπτωτικού αυτού ελέγχου (0.372) είναι το ίδιο με εκείνο που προέκυψε παραπάνω με βάση την κανονική προσέγγιση της κατανομής της ελεγχοσυνάρτησης. Αυτό συμβαίνει επειδή το μέγεθος του δείγματος είναι αρκετά μεγάλο. Κατά συνέπεια, η μηδενική υπόθεση μπορεί να θεωρηθεί εύλογη στα συνήθη επίπεδα σημαντικότητας.

2.2 Ο ΠΡΟΣΗΜΙΚΟΣ ΕΛΕΓΧΟΣ Ή ΕΛΕΓΧΟΣ ΠΡΟΣΗΜΩΝ (*The Sign Test*)

Ο προσημικός έλεγχος ή έλεγχος προσήμων (sign test) είναι ο αρχαιότερος από όλους τους μη παραμετρικούς ελέγχους. Χρησιμοποιήθηκε για πρώτη φορά το 1710. Στην πραγματικότητα, ο προσημικός έλεγχος δεν είναι παρά ο διωνυμικός έλεγχος, όταν $p_0 = 1/2$. Παρά το γεγονός αυτό, ο προσημικός έλεγχος αποτελεί μια παραλλαγή του διωνυμικού ελέγχου, η οποία απαιτεί ξεχωριστή μελέτη λόγω του ενδιαφέροντος που παρουσιάζει η ευέλικτη μορφή της που οφείλεται στο γεγονός ότι $p_0 = 1/2$.

Ο προσημικός έλεγχος χρησιμοποιείται κυρίως για τον έλεγχο της υπόθεσης ότι οι τιμές μιας από τις τυχαίες μεταβλητές του ζεύγους (X, Y) τείνουν να είναι μεγαλύτερες από τις τιμές της άλλης. Χρησιμοποιείται επίσης για τον έλεγχο ύπαρξης τάσης σε μια ακολουθία μετρήσεων σε κλίμακα διάταξης (ordinal measurements), όπως επίσης και ως έλεγχος ύπαρξης συσχέτισης. Σε πολλές από τις περιπτώσεις όπου μπορεί να χρησιμοποιηθεί ο προσημικός έλεγχος, υπάρχουν διαθέσιμοι περισσότερο ισχυροί μη παραμετρικοί έλεγχοι για το ίδιο μοντέλο. Παρ' όλα αυτά, ο προσημικός έλεγχος προτιμάται λόγω της απλότητάς του.

Τα δεδομένα αποτελούνται από n ' παρατηρήσεις πάνω σε ένα διδιάστατο τυχαίο δείγμα $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$. Ο τρόπος πραγματοποίησης των παρατηρήσεων αυτών θα πρέπει να υπαινίσσεται ότι η μεταβλητή X_i ($i = 1, 2, \dots, n$) δεν είναι ανεξάρτητη από την μεταβλητή Y_i ($i = 1, 2, \dots, n$) που αντιστοιχεί σ' αυτήν στο ζεύγος (X_i, Y_i) . Στην αντίθετη περίπτωση, η υπόθεση ότι οι τιμές της μίας μεταβλητής τείνουν να είναι μεγαλύτερες από τις

τιμές της άλλης στο ζεύγος (X, Y) , μπορεί να ελεγχθεί με ένα ισχυρότερο έλεγχο, ο οποίος αναπτύσσεται στα επόμενα και είναι γνωστός στην βιβλιογραφία ως *έλεγχος των Mann και Whitney (Mann - Whitney test)*.

Σε κάθε ζεύγος (X_i, Y_i) γίνεται μια εσωτερική σύγκριση και το ζεύγος αυτό ταξινομείται ως "+" ή "συν" ζεύγος ("+" ή "*plus*" pair), εάν $X_i < Y_i$, ως "-" ή "πλην" ζεύγος ("-" ή "*minus*" pair), αν $X_i > Y_i$, ή ως "0" ή "ισοβαθμούν" ζεύγος ("0" ή "*tied*" pair), αν $X_i = Y_i$, $i = 1, 2, \dots, n$. Επομένως, η κλίμακα μέτρησης των δεδομένων αρκεί να είναι κλίμακα διάταξης (ordinal scale).

Αν οι τιμές της μεταβλητής X τείνουν να είναι μεγαλύτερες από τις τιμές της μεταβλητής Y , τότε τα ζεύγη (X_i, Y_i) είναι *εσωτερικά συνεπή*, με την έννοια ότι αν $P(+)>P(-)$ για κάποιο ζεύγος (X_i, Y_i) , τότε $P(+)>P(-)$, για όλα τα ζεύγη. Η εσωτερική συνέπεια ορίζεται με ανάλογο τρόπο και στην περίπτωση κατά την οποία οι τιμές της μεταβλητής Y τείνουν να είναι μεγαλύτερες από τις τιμές της μεταβλητής X (οπότε οι αντίστοιχες πιθανότητες των "+" και "-" ζευγών ικανοποιούν την ανισότητα $P(+)<P(-)$), όπως επίσης και στην περίπτωση που δεν υπάρχει τάση στις τιμές της μιας μεταβλητής να υπερβαίνουν τις τιμές της άλλης (οπότε οι πιθανότητες των "+" και "-" ζευγών είναι ίσες, δηλαδή, $P(+)=P(-)$). Επομένως, οι υποθέσεις τις οποίες ενδέχεται να ενδιαφερόμαστε να ελέγξουμε έχουν τη μορφή:

A. (Αμφίπλευρος έλεγχος):

$$H_0: P(+)=P(-)$$

$$H_1: P(+)\neq P(-)$$

B. (Μονόπλευρος έλεγχος):

$$H_0: P(+)\leq P(-)$$

$$H_1: P(+)> P(-)$$

Γ. (Μονόπλευρος έλεγχος):

$$H_0: P(+)\geq P(-)$$

$$H_1: P(+)< P(-)$$

Η μορφή της μηδενικής υπόθεσης στην περίπτωση Α υπαινίσσεται ότι οι μεταβλητές X και Y έχουν την ίδια παράμετρο θέσης. Επίσης, στην περίπτωση Β η μηδενική υπόθεση μπορεί να θεωρηθεί ενδεικτική του ότι οι τιμές της μεταβλητής X τείνουν να είναι μεγαλύτερες από τις τιμές της μεταβλητής Y, αφού η H_0 δηλώνει ότι το ενδεχόμενο οι τιμές της μεταβλητής X να υπερβαίνουν τις τιμές της μεταβλητής Y είναι περισσότερο πιθανό από το ενδεχόμενο οι τιμές της μεταβλητής Y να υπερβαίνουν τις τιμές της μεταβλητής X. Αντίστοιχα, στην περίπτωση Γ, η μηδενική υπόθεση δηλώνει ότι η μεταβλητή X έχει την τάση να παίρνει τιμές μικρότερες από τις τιμές της μεταβλητής Y, αφού θεωρεί ότι το ενδεχόμενο οι τιμές της Y να υπερβαίνουν τις τιμές της X είναι περισσότερο πιθανό από το ενδεχόμενο οι τιμές της X να υπερβαίνουν τις τιμές της Y. Επομένως, ο προσημικός έλεγχος θα μπορούσε να χρησιμοποιηθεί και για τον έλεγχο υποθέσεων της μορφής:

Α. (Αμφίπλευρος έλεγχος):

$$H_0: E(X) = E(Y)$$

$$H_1: E(X) \neq E(Y)$$

Β. (Μονόπλευρος έλεγχος):

$$H_0: E(X) \geq E(Y)$$

$$H_1: E(X) < E(Y)$$

Γ. (Μονόπλευρος έλεγχος):

$$H_0: E(X) \leq E(Y)$$

$$H_1: E(X) > E(Y)$$

Για τους ίδιους λόγους, ο προσημικός έλεγχος μπορεί να χρησιμοποιηθεί ως έλεγχος υποθέσεων για τις διαμέσους των μεταβλητών X και Y .

Είναι προφανές, ότι μια κατάλληλη μορφή στατιστικής συνάρτησης για τον έλεγχο των παραπάνω υποθέσεων είναι η

$$T = \text{αριθμός των "+" ζευγών.}$$

Δηλαδή, η στατιστική συνάρτηση T συμβολίζει τον αριθμό των ζευγών (X_i, Y_i) στα οποία η μεταβλητή X_i είναι μικρότερη από τη μεταβλητή Y_i . Αγνοώντας τα "0" ενδεχόμενα, και περιοριζόμενοι σε n ζεύγη, όπου

$$n = (\text{αριθμός των "+" ζευγών}) + (\text{αριθμός των "-" ζευγών}),$$

μπορούμε εύκολα να δούμε ότι το πρόβλημα του ελέγχου των παραπάνω υποθέσεων τοποθετείται στο πλαίσιο του διωνυμικού μοντέλου.

Πράγματι, κάθε ζεύγος από τα απομένοντα n ζεύγη παρατηρήσεων (X_i, Y_i) θεωρείται ως μια δοκιμή, η οποία μπορεί να οδηγήσει σε ένα από δύο δυνατά αποτελέσματα:

επιτυχία, "+" ($\{X_i < Y_i\}$) με πιθανότητα $P(+)$

αποτυχία, "-" ($\{X_i > Y_i\}$) με πιθανότητα $P(-)$

Κατά συνέπεια, η ελεγχοσυνάρτηση T ακολουθεί την διωνυμική κατανομή με παραμέτρους n και $p = P(+)$. Συμβολικά,

$T \sim$ διωνυμική ($n, p = P(+)$).

Επομένως, οι υποθέσεις των περιπτώσεων A, B και Γ γράφονται ισοδύναμα με την εξής μορφή:

A. (Αμφίπλευρος έλεγχος):

$H_0: P(+) = 1/2$

$H_1: P(+) \neq 1/2$

B. (Μονόπλευρος έλεγχος):

$H_0: P(+) \leq 1/2$

$H_1: P(+) > 1/2$

Γ. (Μονόπλευρος έλεγχος):

$H_0: P(+) \geq 1/2$

$H_1: P(+) < 1/2$

Είναι προφανές ότι για τον έλεγχο των παραπάνω υποθέσεων μπορεί να χρησιμοποιηθεί ο διωνυμικός έλεγχος, αφού κάτω από τη μηδενική υπόθεση, η στατιστική συνάρτηση T ακολουθεί την διωνυμική κατανομή με παραμέτρους n και $p=P(+)=1/2$. Συμβολικά,

$T \sim$ διωνυμική ($n, p=P(+)=1/2$).

Τότε, σε επίπεδο σημαντικότητας περίπου ίσο με α , ο κανόνας απόφασης έχει την μορφή:

A. Η H_0 απορρίπτεται αν $T \leq t_1$ ή $T > t_2$, όπου t_1, t_2 ορίζονται έτσι

ώστε $P(T \leq t_1 | n, p=1/2) \cong \alpha/2$ και $P(T > t_2 | n, p=1/2) \cong \alpha/2$.

B. Η H_0 απορρίπτεται αν $T > t$, όπου t ορίζεται έτσι ώστε

$P(T > t | n, p=1/2) \cong \alpha$.

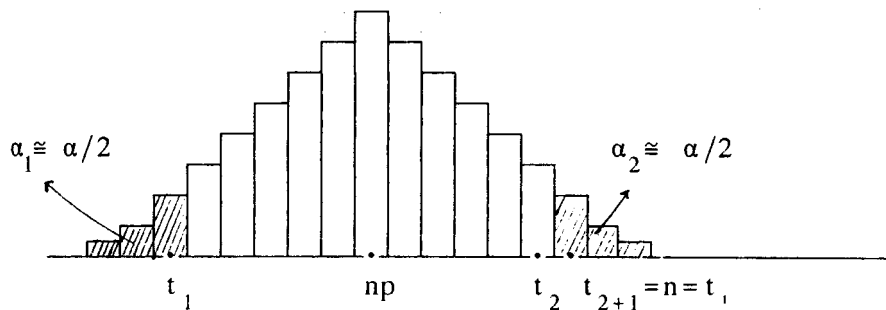
Γ. Η H_0 απορρίπτεται αν $T \leq t$, όπου t ορίζεται έτσι ώστε

$$P(T \leq t | n, p=1/2) \cong \alpha.$$

Λόγω συμμετρίας ($p=1/2$), ισχύει για την κρίσιμη περιοχή της περίπτωσης Α ότι

$$\begin{aligned} \alpha/2 &\cong P(T \leq t_1 | n, p=1/2) \\ &= P(T > t_2 | n, p=1/2) \\ &= P(T \geq t_2 + 1 | n, p=1/2) \\ &= P(T \geq n - t_1 | n, p=1/2). \end{aligned}$$

Η αλήθεια της παραπάνω σχέσης επιβεβαιώνεται και από το σχήμα 2.2.1.



Σχήμα 2.2.1.

Κατά συνέπεια, η κρίσιμη περιοχή μεγέθους κατά προσέγγιση ίσου με α του ελέγχου των υποθέσεων Α, Β και Γ έχει την μορφή:

Α. $T \leq t$ και $T \geq n - t$, όπου η τιμή t προσδιορίζεται από τον πίνακα 1 του παραρτήματος, έτσι ώστε $P(T \leq t | n, p=1/2) \cong \alpha/2$.

B. $T \geq n-t$, όπου t ορίζεται έτσι ώστε $P(T \leq t | n, p=1/2) \cong \alpha$.

Γ. $T \leq t$, όπου t ορίζεται έτσι ώστε $P(T \leq t | n, p=1/2) \cong \alpha$.

Παράδειγμα 2.2.1: Έξι άτομα υποβλήθηκαν σε μια δίαιτα για να χάσουν βάρος. Τα βάρη των ανθρώπων πριν (X_i) και μετά (Y_i) την δίαιτα ($i = 1, 2, \dots, 6$) είναι τα εξής :

X_i : 174 191 188 182 201 188

Y_i : 165 186 183 178 203 181

Παρέχουν τα δεδομένα ενδείξεις ότι η δίαιτα ήταν αποτελεσματική;

Λύση: Αν η δίαιτα ήταν αποτελεσματική, τότε θα πρέπει $P(+)$ < $P(-)$, όπου ως "+" ενδεχόμενο ορίζεται το ενδεχόμενο $\{X < Y\}$. Επομένως, οι υποθέσεις που θα μπορούσαν να δώσουν απάντηση στο ερώτημά μας μπορούν να γραφούν με τη μορφή

$H_0: P(+)$ \geq $P(-)$ (μη αποτελεσματική δίαιτα)

$H_1: P(+)$ < $P(-)$ (μη αποτελεσματική δίαιτα).

Οι υποθέσεις αυτές είναι ισοδύναμες με τις υποθέσεις:

$H_0: P(+)$ \geq 1/2

$H_1: P(+)$ > 1/2.

Είναι προφανές ότι, κάτω από την μηδενική υπόθεση, η στατιστική συνάρτηση T ακολουθεί την διωνυμική κατανομή με παραμέτρους $n=6$, $p=1/2$. Από τον πίνακα που ακολουθεί, προκύπτει ότι η τιμή της στατιστικής συνάρτησης $T =$ αριθμός των "+" ζευγών, είναι ίση με 1, δηλαδή $\tau=1$.

$X_i :$	174	191	188	182	201	188
$Y_i :$	165	186	183	178	203	181
Πρόσημο ζεύγους ($X_i Y_i$) :	-	-	-	-	+	-

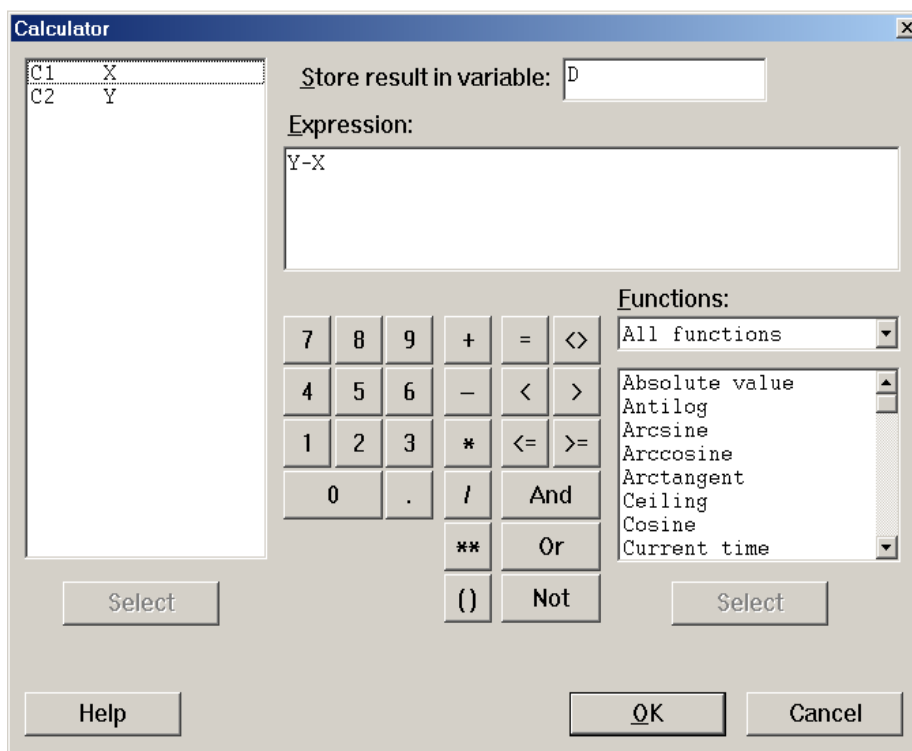
Επιπλέον, $n' = n = 6$. Από τον πίνακα 1 του παραρτήματος, προκύπτει ότι το κρίσιμο επίπεδο του ελέγχου είναι 10.94%. Πράγματι,

$$\hat{\alpha} = P(T \leq 1 | n=6, p=1/2) = 0.1094.$$

Κατά συνέπεια, η μηδενική υπόθεση H_0 θα απορρίπτεται σε κάθε επίπεδο σημαντικότητας α , αν $\alpha \geq 0.1094$.

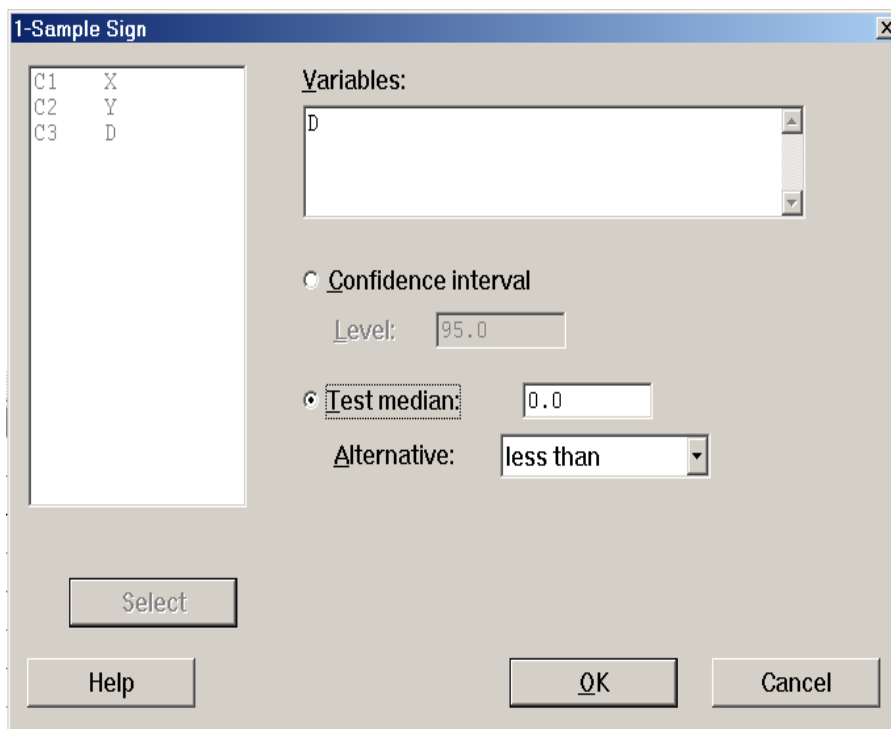
Λύση με το MINITAB: Η διεξαγωγή του προσημικού ελέγχου για ζεύγη παρατηρήσεων με το MINITAB ανάγεται σε προσημικό έλεγχο για ένα δείγμα. Συγκεκριμένα, για τον έλεγχο αυτό, μπορούμε να θεωρήσουμε το δείγμα των διαφορών των μελών των ζευγών παρατηρήσεων $D_1=Y_1-X_1, D_2=Y_2-X_2, \dots, D_n=Y_n-X_n$ και σ' αυτές να αντιστοιχίσουμε τα πρόσημά τους (+ στις θετικές (περίπτωση $X_i < Y_i$) και - στις αρνητικές (περίπτωση $X_i > Y_i$)).

Καταχωρίζουμε τα δύο δείγματα στις μεταβλητές **X** (στήλη C1) και **Y** (στήλη C2) αντίστοιχα και πιέζοντας **Calc, Calculator** στο βασικό μενού προκύπτει το εξής πλαίσιο διαλόγου:



Στο πεδίο **Store result in variable** δηλώνουμε το όνομα της μεταβλητής που θα καταχωρισθούν οι διαφορές D_i . Στο πεδίο **Expression** απλά πληκτρολογούμε την μαθηματική έκφραση που θα οδηγήσει στις διαφορές D_i πιέζοντας **OK**, οι διαφορές αυτές αποθηκεύονται στην στήλη D.

Πιέζοντας **Stat, Nonparametrics, 1-Sample sign** στο βασικό μενού προκύπτει το παρακάτω πλαίσιο διαλόγου:



Στο πεδίο **Variables**, δηλώνουμε την μεταβλητή που περιέχει τις διαφορές. Αυτό γίνεται με το να την επιλέξουμε από το πεδίο στα αριστερά και να πιάσουμε **Select**. Το MINITAB χειρίζεται τον προσημικό έλεγχο σαν έλεγχο για την διάμεσο της διαφοράς των δύο τυχαίων μεταβλητών. Συνεπώς στο παράδειγμα αυτό η μηδενική υπόθεση ισοδυναμεί με την υπόθεση ότι η διάμεσος των διαφορών $Y - X$ είναι ≥ 0 και η εναλλακτική με την υπόθεση ότι η διάμεσος των διαφορών είναι < 0 . Στο MINITAB δεν υπάρχει τρόπος να ξεχωρίσουμε μηδενικές υποθέσεις που δηλώνουν ισότητα της διαμέσου με μια συγκεκριμένη τιμή από μηδενικές υποθέσεις που δηλώνουν ανισότητα. Απλά στο πεδίο **Test median** δηλώνουμε την τιμή της διαμέσου που εμφανίζεται στην μηδενική υπόθεση, δηλαδή 0. Στην εναλλακτική υπόθεση δηλώνουμε **less than** που αντιστοιχεί στην υπόθεση ότι η

δίαιτα είναι αποτελεσματική (δηλαδή ότι η Y τείνει να έχει μικρότερες τιμές από την X). Τα αποτελέσματα του ελέγχου είναι:

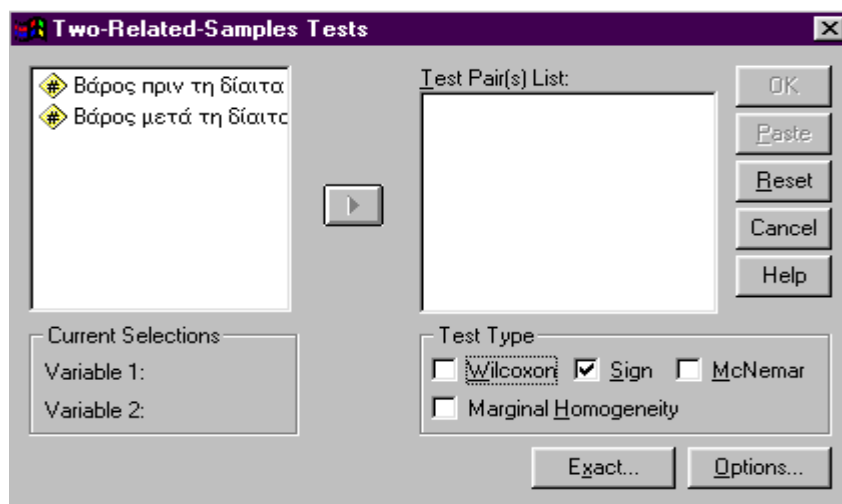
Sign Test for Median


Sign test of median = 0.00000 versus < 0.00000

	N	Below	Equal	Above	P	Median
D	6	5	0	1	0.1094	-5.000

Παρατηρούμε ότι έχουμε 5 αρνητικές διαφορές και 1 θετική. Η τιμή του κρίσιμου επιπέδου είναι 0.1094 και συνεπώς η μηδενική υπόθεση δεν απορρίπτεται στα συνήθη επίπεδα σημαντικότητας.

Λύση με το SPSS: Για την διεξαγωγή του προσημικού ελέγχου, πρέπει τα δύο δείγματα να καταχωρισθούν σε δύο μεταβλητές από τις οποίες η πρώτη θα αντιστοιχεί στην X και η δεύτερη στην Y μεταβλητή. Ας υποθέσουμε ότι καταχωρίζουμε τα βάρη πριν τη δίαιτα στην μεταβλητή x (με label «Βάρος πριν τη δίαιτα») και τα βάρη μετά τη δίαιτα στην μεταβλητή y (με label «Βάρος μετά τη δίαιτα»). Επιλέγοντας κατά σειρά **Analyze, Nonparametric Tests, 2 Related Samples** στο βασικό μενού οδηγούμεθα στο παρακάτω πλαίσιο διαλόγου:



Επιλέγουμε τις μεταβλητές **x** και **y** και βλέπουμε ότι στο πεδίο **Current Selections** εμφανίζονται οι ετικέτες τους. Πιέζοντας το πλήκτρο , αυτές μεταφέρονται στο πεδίο **Test Pair(s) List**. Στην συνέχεια, στο πεδίο **Test Type**, επιλέγουμε **Sign** (για προσημικό έλεγχο). Επειδή το δείγμα μας είναι μικρό (6 ζεύγη παρατηρήσεων) ζητάμε **Exact** έλεγχο. Τα αποτελέσματα που δίνει ο έλεγχος είναι τα εξής:

Frequencies

		N
Βάρος μετά τη δίαιτα - Βάρος πριν τη δίαιτα	Negative Differences ^a	5
	Positive Differences ^b	1
	Ties ^c	0
	Total	6

a Βάρος μετά τη δίαιτα < Βάρος πριν τη δίαιτα

b Βάρος μετά τη δίαιτα > Βάρος πριν τη δίαιτα

c Βάρος πριν τη δίαιτα = Βάρος μετά τη δίαιτα

Test Statistics^b

	Βάρος μετά τη δίαιτα - Βάρος πριν τη δίαιτα
Exact Sig. (2-tailed)	.219 ^a
Exact Sig. (1-tailed)	.109
Point Probability	.094

a Binomial distribution used.

b Sign Test

Παρατηρούμε ότι έχουμε 5 αρνητικές διαφορές (Βάρος μετά < Βάρος πριν) και 1 θετική. Δεν έχουμε καμμία ισοβαθμία τιμών (πεδίο **Ties**). Το SPSS μας επεξηγεί στις υποσημειώσεις του πίνακα **Frequencies** ποιες διαφορές θεωρεί αρνητικές και ποιες θετικές.

Στον πίνακα **Test Statistics** παρέχονται οι τιμές του κρίσιμου επιπέδου για μονόπλευρο και για αμφίπλευρο έλεγχο. Στο συγκεκριμένο παράδειγμα, έχουμε μονόπλευρο έλεγχο και επομένως η τιμή του κρίσιμου επιπέδου είναι 0.109. Το SPSS μας πληροφορεί, σε υποσημείωση, ότι η τιμή του κρίσιμου επιπέδου υπολογίσθηκε με βάση την διωνυμική κατανομή της στατιστικής συνάρτησης ελέγχου.

Λύση με το SAS: Εισάγουμε στο παράθυρο εντολών τις παρακάτω εντολές.

```
data data221;
input x y @@;
dif=y-x;
cards;
174 165 191 186 188 183 182 178 201 203 188 181
;
run;
proc univariate;
var dif;
run;
```

Με την εντολή **dif=y-x**; υπολογίζεται η διαφορά των δύο μεταβλητών. Η εντολή **proc univariate**; χρησιμοποιείται για τον υπολογισμό περιγραφικών μέτρων μιας μεταβλητής καθώς και για την διεξαγωγή μιας σειράς ελέγχων που βασίζονται στην κατανομή της μεταβλητής αυτής, η οποία συμπεριλαμβάνει τον προσημικό έλεγχο.

Τα αποτελέσματα περιέχονται στον παρακάτω πίνακα.

The SAS System
Univariate Procedure

Variabl e=DI F

Quantiles(Def=5)		Moments					
N	6	Sum Wgts	6	100% Max	2	99%	2
Mean	-4.66667	Sum	-28	75% Q3	-4	95%	2
Std Dev	3.723797	Variance	13.86667	50% Med	-5	90%	2
Skewness	1.176167	Kurtosis	2.333349	25% Q1	-7	10%	-9
USS	200	CSS	69.33333	0% Min	-9	5%	-9
CV	-79.7957	Std Mean	1.520234			1%	-9
T: Mean=0	-3.0697	Pr> T	0.0278			Range	11
Num ^= 0	6	Num > 0	1			Q3-Q1	3
M(Sign)	-2	Pr>= M	0.2188			Mode	-5
Sgn Rank	-9.5	Pr>= S	0.0625				

Extremes			
Lowest	Obs	Highest	Obs
-9(1)	-7(6)
-7(6)	-5(2)
-5(3)	-5(3)
-5(2)	-4(4)
-4(4)	2(5)

Το πακέτο δίνει όλα τα διαθέσιμα περιγραφικά μέτρα μαζί με κάποιους ελέγχους που αφορούν την υπό εξέταση μεταβλητή. Το αποτέλεσμα που αναφέρεται στον προσημικό έλεγχο, περιέχεται στα πεδία

$$M(\text{Sign}) \quad -2 \quad Pr \geq |M| \quad 0.2188$$

Ας σημειωθεί ότι το πακέτο υπολογίζει την τιμή του κρίσιμου επιπέδου που αντιστοιχεί σε αμφίπλευρο έλεγχο και παρέχεται από το πεδίο ($Pr \geq |M|0.2188$). Για τα δεδομένα του παραδείγματος, η τιμή αυτή είναι 0.2188. Η τιμή του κρίσιμου επιπέδου που αντιστοιχεί στον μονόπλευρο έλεγχο του παραδείγματος προκύπτει με διαίρεση της τιμής αυτής με 2. Δηλαδή, $\hat{\alpha} = 0.2188/2 = 0.1094$.

Παράδειγμα 2.2.2: Ο χρόνος αντίδρασης 28 υπαλλήλων γραφείου μετρήθηκε πριν και μετά το μεσημεριανό φαγητό. Για 22 υπαλλήλους, ο χρόνος αυτός ήταν μεγαλύτερος πριν από το φαγητό, ενώ, για 2 υπαλλήλους, δεν βρέθηκε διαφορά. Είναι ο χρόνος αντίδρασης μετά το φαγητό σημαντικά μεγαλύτερος από τον χρόνο αντίδρασης πριν από το φαγητό;

Λύση: Ορίζουμε ως «+» ενδεχόμενο το ενδεχόμενο {χρόνος αντίδρασης πριν από το φαγητό < από τον χρόνο αντίδρασης μετά το φαγητό}, ως «-» ενδεχόμενο το ενδεχόμενο {χρόνος αντίδρασης πριν από το φαγητό > από τον χρόνο αντίδρασης μετά το φαγητό} και ως «0» ενδεχόμενο το ενδεχόμενο {χρόνος αντίδρασης πριν από το φαγητό = χρόνος αντίδρασης μετά το φαγητό}.

Οι υποθέσεις που ενδιαφερόμαστε να ελέγξουμε έχουν την μορφή

$$H_0 : P(+) = 1/2$$

$$H_1 : P(+) > 1/2$$

Η προφανής επιλογή στατιστικής συνάρτησης για τον έλεγχο των υποθέσεων αυτών, είναι η στατιστική συνάρτηση $T = \# \text{ των } +$.

Προφανώς, $n' = 28$ και $n = n' - (\# \text{ των } 0) = 28 - 2 = 26$. Επομένως, $T \sim \text{διων} (n = 26, p = P(+))$ και, κατά συνέπεια, κάτω από την H_0 ,

$T \sim \text{διων} (n = 26, p = 1/2)$. Άρα, η κρίσιμη περιοχή ορίζεται από την ανισότητα $T \geq n - t$, όπου t ορίζεται έτσι ώστε

$$P(T \geq n - t \mid n = 26, p = 1/2) \cong \alpha.$$

Η τιμή του κρίσιμου επιπέδου του ελέγχου είναι:

$$\begin{aligned} \hat{\alpha} &= \max P(T \geq 22 \mid n = 26, p \leq 1/2) = P(T \geq 22 \mid n = 26, p = 1/2) \\ &\cong P(X_{N(\mu, \sigma^2)} \geq 22 \mid \mu = 26/2, \sigma^2 = 26/4) = P\left(Z \geq \frac{22 - 13}{\sqrt{13/2}}\right) \\ &= 1 - \Phi(3.53) \cong 0.0002 \end{aligned}$$

Η τιμή αυτή είναι πολύ χαμηλή. Επομένως, δεν είναι εύλογο να υποθέσουμε ότι ο χρόνος αντίδρασης μετά το φαγητό είναι μεγαλύτερος από τον χρόνο αντίδρασης πριν από το φαγητό.

Λύση με το MINITAB: ο προσημικός έλεγχος του παραδείγματος αυτού αντιστοιχεί σε διωνυμικό έλεγχο, όπου η ελεγχοσυνάρτηση T παριστάνει τον αριθμό των επιτυχιών σε $n=26$ δοκιμές και της οποίας η κατανομή κάτω από την H_0 είναι η διωνυμική με $n=26$ και $p=p_0=0.5$. Ο έλεγχος αυτός μπορεί να εφαρμοσθεί μέσω του MINITAB από την επιλογή **Stat, Basic Statistics, 1 Proportion** στο κυρίως μενού. Επειδή το πακέτο μας δίνει την δυνατότητα εφαρμογής του ακριβούς ελέγχου, δεν επιλέγουμε την χρησιμοποίηση κανονικής προσέγγισης. Τα αποτελέσματα που παρέχονται είναι:

Test of p = 0.5 vs p > 0.5					
Sample	X	N	Sample p	95.0 % CI	Exact P-Value
1	22	26	0.846154	(0.651321, 0.956437)	0.000

Όπως φαίνεται από το πεδίο **p-value**, η τιμή του κρίσιμου επιπέδου είναι τόσο μικρή, ώστε από το πρόγραμμα να αναφέρεται ως 0. Συνεπώς, η μηδενική υπόθεση ότι ο χρόνος αντίδρασης μετά το

μεσημεριανό φαγητό είναι μεγαλύτερος από τον χρόνο αντίδρασης πριν από αυτό δεν φαίνεται να είναι μια εύλογη υπόθεση.

Χρήση της κανονικής προσέγγισης δίνει τα ακόλουθα

Test of $p = 0.5$ vs $p > 0.5$

Sample	X	N	Sample p	95,0 % CI	Z-Value	P-Value
1	22	26	0.846154	0.707469; 0.984839)	3.53	0.000

NOTE * The normal approximation may be inaccurate for small samples.

Λύση με το SPSS: ο προσημικός έλεγχος του παραδείγματος αυτού ισοδυναμεί με διωνυμικό με 26 δοκιμές, 22 επιτυχίες και $p_0=0.5$. Διεξάγοντας τον με SPSS και με κανονική προσέγγιση της διωνυμικής κατανομής και χωρίς αυτήν λαμβάνουμε τα αποτελέσματα:

Binomial Test

	Category	N	Observed Prop.	Test Prop.	Asymp. Sig. (2-tailed)	Exact Sig. (2-tailed)
A	Group 1	1	22	.85	.50	.001 ^a
	Group 2	0	4	.15		
	Total		26	1.00		

a Based on Z Approximation.

Παρατηρούμε ότι, στην περίπτωση αμφίπλευρου ελέγχου με βάση τόσο την ακριβή κατανομή της ελεγχοσυνάρτησης όσο και την κανονική προσέγγιση, το κρίσιμο επίπεδο είναι 0.001. Η τιμή του κρίσιμου επιπέδου στην περίπτωση του μονόπλευρου ελέγχου, προκύπτει με διαίρεση της τιμής 0.001 με 2.

Λύση με το SAS: Το πακέτο δεν παρέχει άμεσα τον διωνυμικό έλεγχο. Μπορούμε όμως να διεξαγάγουμε τον έλεγχο που βασίζεται στην κανονική προσέγγιση της κατανομής της ελεγχοσυνάρτησης χρησιμοποιώντας τις εντολές:

```
data reaction;
x=1; f=22; output;
x=2; f=4; output;
run;
proc freq;
```

```
weight f;
tables x / testp=(0.5 0.5);
run;
```

Το αποτέλεσμα του ελέγχου δίνεται στον πίνακα που ακολουθεί

The SAS System					
X	Frequency	Percent	Test Percent	Cumul ative Frequency	Cumul ative Percent
1	22	84.6	50.0	22	84.6
2	4	15.4	50.0	26	100.0

Chi-Square Test for Specified Proportions

Statistic = 12.462 DF = 1 Prob = 0.001

Το αμφίπλευρο παρατηρούμενο επίπεδο στατιστικής σημαντικότητας δίνεται στο πεδίο **Prob = 0.001**. Διαίρεση αυτού με το 2 δίνει το μονόπλευρο παρατηρούμενο επίπεδο σημαντικότητας. Κατά συνέπεια, η μηδενική υπόθεση δεν μπορεί να θεωρηθεί εύλογη στα συνήθη επίπεδα σημαντικότητας.

2.3 ΠΑΡΑΛΛΑΓΕΣ ΤΟΥ ΠΡΟΣΗΜΙΚΟΥ ΕΛΕΓΧΟΥ

2.3.1 Ο Έλεγχος McNemar για την Σημαντικότητα της Αλλαγής μιας Κατάστασης (*McNemar Test for Significance of Changes*)

Τα δεδομένα αποτελούνται από παρατηρήσεις πάνω σε n' ανεξάρτητες διδιάστατες τυχαίες μεταβλητές (X_i, Y_i) , ($i = 1, 2, \dots, n'$). Η κλίμακα μέτρησης των X_i και Y_i είναι ονομαστική (nominal scale) με δύο κατηγορίες, τις οποίες μπορούμε να ονομάσουμε "0" και "1". Δηλαδή, οι δυνατές τιμές του ζεύγους (X_i, Y_i) είναι (0, 0), (0, 1), (1, 0) και (1, 1). Το ερώτημα που συχνά απασχολεί σε σχέση με δεδομένα αυτής της μορφής αφορά το κατά πόσο μπορούμε να ανιχνεύσουμε μια διαφορά μεταξύ της πιθανότητας πραγματοποίησης

του ενδεχομένου (0, 1) και της πιθανότητας πραγματοποίησης του ενδεχομένου (1, 0). Ερωτήματα αυτής της μορφής αναφέρονται όταν η μεταβλητή X_i στο ζεύγος (X_i, Y_i) συμβολίζει την κατάσταση της i μονάδας ενός πληθυσμού πριν από κάποιο πείραμα, ενώ η μεταβλητή Y_i συμβολίζει την κατάστασή της μετά το πείραμα.

Τα ζεύγη των παρατηρήσεων είναι συνήθως ταξινομημένα στις κατηγορίες (0, 0), (0, 1), (1, 0) και (1, 1) ως εξής:

		Y	
		0	1
X	0	(0, 0)	(0, 1)
	1	(1, 0)	(1, 1)

Ο έλεγχος της σημαντικότητας μιας αλλαγής θα βασίζεται στην συχνότητα εμφάνισης των ζευγών (0, 1) και (1, 0). Επομένως, οι υποθέσεις που ενδέχεται να ενδιαφερόμαστε να ελέγξουμε έχουν τη μορφή:

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

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

Είναι προφανές ότι τα ζεύγη (0, 0) και (1, 1) (δηλαδή ζεύγη (X_i, Y_i) για τα οποία ισχύει $\{X_i=Y_i\}$) υπονοούν ότι δεν πραγματοποιήθηκε καμία αλλαγή και, επομένως, μπορούν να αγνοηθούν. Ας θεωρήσουμε, λοιπόν, τα απομένοντα n ζεύγη, όπου

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

Ας θεωρήσουμε επίσης το ζεύγος (0, 1) ως επιτυχία ($\{X < Y\} = "+"$ ζεύγος) και το ζεύγος (1, 0) ως αποτυχία ($\{X > Y\} = "-"$ ζεύγος). Τότε, η στατιστική συνάρτηση

$T =$ αριθμός των επιτυχιών

ακολουθεί την διωνυμική κατανομή με παραμέτρους n και $p=1/2$ ($T \sim$ διωνυμική ($n, p=1/2$)). Είναι προφανές, επομένως, ότι οι υποθέσεις που θέλουμε να ελέγξουμε γράφονται με την μορφή:

$$H_0: P(+)=P(-)$$

$$H_1: P(+)\neq P(-),$$

ή, ισοδύναμα

$$H_0: P(+)=1/2$$

$$H_1: P(+)\neq 1/2,$$

οι οποίες μπορούν να ελεγχθούν σύμφωνα με την προηγούμενη ενότητα.

Παράδειγμα 2.3.1: Από 135 άτομα, που επελέγησαν τυχαία, ζητήθηκε να εκφράσουν γνώμη για την εξωτερική πολιτική της χώρας τους. 43 άτομα ήταν εναντίον της ακολουθούμενης εξωτερικής πολιτικής. Μετά από κάποιο διάστημα, τα άτομα αυτά ρωτήθηκαν ξανά. 37 άτομα ήταν εναντίον της ακολουθούμενης εξωτερικής πολιτικής, από τα οποία 30 ήταν άτομα που αρχικά ήταν υπέρ. Θα μπορούσε να συμπεράνει κανείς ότι η μεταβολή στον αριθμό των ατόμων που ήταν εναντίον της ακολουθούμενης εξωτερικής πολιτικής της χώρας τους είναι σημαντική;

Λύση: Ας συμβολίσουμε με X και Y την γνώμη ενός ατόμου την πρώτη φορά που ρωτήθηκε και την δεύτερη φορά που ρωτήθηκε, αντίστοιχα. Τότε, αντιστοιχίζοντας την τιμή 0 στο ενδεχόμενο αρνητικής γνώμης και την τιμή 1 στο ενδεχόμενο θετικής γνώμης, τα 135 άτομα μπορούν να ταξινομηθούν σύμφωνα με τον πίνακα που ακολουθεί.

		Y		
		0	1	
X	0	7	36(+)	43
	1	30(-)	62	92
		37	98	135

Προφανώς, $n = 36+30 = 66$. Τότε, ο έλεγχος που θα οδηγήσει στην απάντηση του ερωτήματος θα αναφέρεται στις υποθέσεις

$$H_0: P(0, 1) = 1/2$$

$$H_1: P(0, 1) \neq 1/2,$$

με βάση τα $n=66$ ενδεχόμενα της μορφής $(1, 0)$ και $(0, 1)$. Η προφανής επιλογή στατιστικής συνάρτησης είναι ο αριθμός T των $(0, 1)$ ζευγών (αριθμός των "+"). Παρατηρούμε ότι, κάτω από την μηδενική υπόθεση, $T \sim$ διωνυμική ($n=66, p=1/2$).

Επειδή, $n=66 > 20$ και λόγω συμμετρίας της κατανομής της στατιστικής συνάρτησης T κάτω από τη μηδενική υπόθεση ($p=1/2$), ισχύει κατά προσέγγιση ότι $T \sim N(\mu=np=n/2, \sigma^2=np(1-p)=n/4)$.

Επομένως, η κρίσιμη περιοχή του ελέγχου ορίζεται από τις ανισότητες $T \leq t$ και $T \geq n-t$, όπου η τιμή t ορίζεται από την σχέση $P(T \leq t | n, p=1/2) \cong \alpha/2$. Δηλαδή, η τιμή t προσδιορίζεται από την σχέση

$$\alpha/2 \cong P(T \leq t | n, p=1/2) \cong P\left[\frac{T - n/2}{\sqrt{n/4}} < \frac{t - n/2}{\sqrt{n/4}}\right] = P\left[Z < \frac{t - n/2}{\sqrt{n/4}}\right],$$

όπου Z συμβολίζει μια τυποποιημένη κανονική μεταβλητή.

Από την παραπάνω σχέση προκύπτει ότι

$$\frac{t - n/2}{\sqrt{n/4}} = Z_{\alpha/2} = Z_{1-\alpha/2}$$

ή, ισοδύναμα

$$t = n/2 - z_{1-\alpha/2} \sqrt{n/4} .$$

Αν, επομένως, επιθυμούμε να ελέγξουμε τις παραπάνω υποθέσεις σε επίπεδο σημαντικότητας $\alpha=0.10$, οπότε $z_{1-\alpha/2} = z_{0.95} = 1.645$ (πίνακας 2 του παραρτήματος), η κρίσιμη περιοχή θα έχει την μορφή:

$$\{T \leq t\} \cup \{T \geq n-t\},$$

όπου

$$\begin{aligned} t &= 66/2 - 1.645 \sqrt{66/4} \\ &= 26.32 = 27. \end{aligned}$$

Δηλαδή, η κρίσιμη περιοχή μεγέθους $\alpha=0.10$ ορίζεται από τις ανισότητες $T \leq 27$ και $T \geq 66-27 = 39$.

Η παρατηρούμενη, όμως, τιμή της στατιστικής συνάρτησης T ($\tau = 36$) δεν βρίσκεται στην κρίσιμη περιοχή. Επομένως, σε επίπεδο σημαντικότητας 10%, η υπόθεση H_0 δεν απορρίπτεται, δηλαδή η μεταβολή στον αριθμό των ατόμων που ήταν εναντίον της ακολουθούμενης εξωτερικής πολιτικής της χώρας τους δεν μπορεί να θεωρηθεί σημαντική σε επίπεδο $\alpha=0.10$.

Το κρίσιμο επίπεδο του ελέγχου υπολογίζεται με βάση τον πίνακα 2 του παραρτήματος.

$$\begin{aligned} \hat{\alpha} &= 2P(T \geq 36 \mid n=66, p=1/2) \\ &\cong 2P(X_N \geq 36 \mid \mu=33, \sigma^2=16.5) \\ &= 2P\left(Z \geq \frac{36-33}{16.5}\right) \\ &= 2(1 - P(Z < 0.74)) = 0.4592 \end{aligned}$$

(Η μεταβλητή X_N στην δεύτερη ισότητα συμβολίζει μια κανονική τυχαία μεταβλητή).

Λύση με το MINITAB: Επειδή ο έλεγχος McNemar ως παραλλαγή του διωνυμικού ελέγχου είναι στην ουσία διωνυμικός έλεγχος, δεν δίνεται χωριστά στο MINITAB. Έτσι, εκτελείται από το πακέτο, όπως ο διωνυμικός έλεγχος. Στο παράδειγμα μας, έχουμε 66 δοκιμές και 36 επιτυχίες με πιθανότητα επιτυχίας $p_0=0.5$. Τ' αποτελέσματα του ελέγχου είναι τα εξής:

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	Z-Value	P-Value
1	36	66	0.545455	(0.425327, 0.665582)	0.74	0.460

Η τιμή του κρίσιμου επιπέδου είναι 0.46 και, κατά συνέπεια η H_0 μπορεί να θεωρηθεί ως εύλογη.

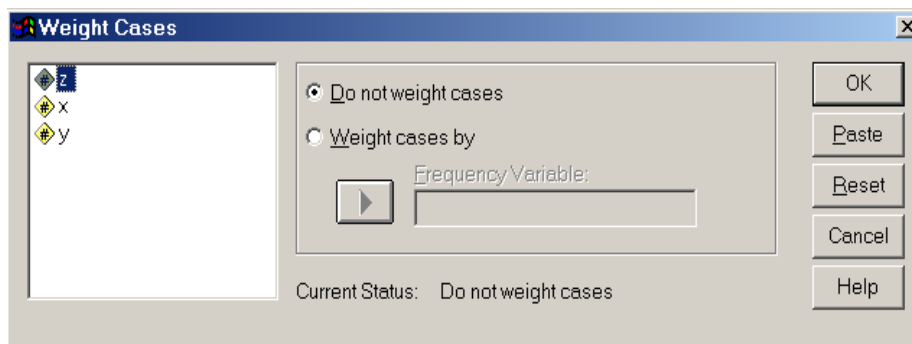
Λύση με το SPSS: Για την διεξαγωγή του ελέγχου McNemar με το SPSS, τα δεδομένα πρέπει να εισαχθούν σε ειδική μορφή. Σε μια μεταβλητή (έστω z), καταχωρίζονται οι συχνότητες του πίνακα στον οποίο είναι ταξινομημένα τα δεδομένα ενώ, σε δύο άλλες (έστω x και y), χρειάζεται να καταχωρισθούν τιμές-δείκτες που δηλώνουν σε ποια γραμμή και στήλη αναλογεί κάθε καταχώριση της μεταβλητής z . Για το συγκεκριμένο παράδειγμα, το παράθυρο δεδομένων θα έχει την μορφή:

	z	x	y	
1	7	1	1	
2	36	1	2	
3	30	2	1	
4	62	2	2	
5				

Στην γραμμή 1 και στήλη 2, δηλαδή, αναλογούν 36 μέλη του δείγματος.

Σημείωση: Αντί των τιμών 1 και 2 των μεταβλητών x και y , θα μπορούσαμε να είχαμε χρησιμοποιήσει οποιαδήποτε άλλα ζεύγη τιμών, αφού η κλίμακα μέτρησης των x και y είναι ονομαστική.

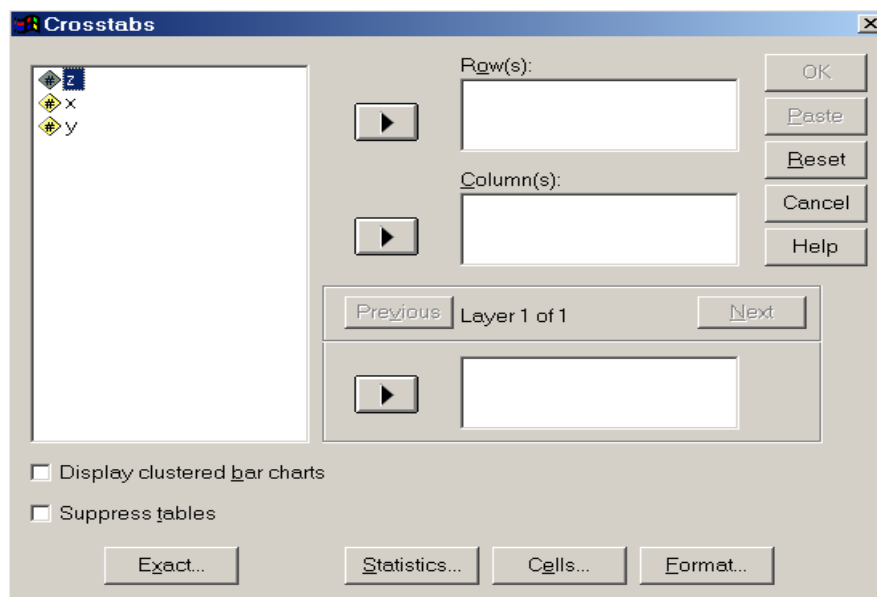
Για την διεξαγωγή του ελέγχου McNemar πρέπει πρώτα να σταθμίσουμε τις παρατηρήσεις με συγκεκριμένο τρόπο. Στο κυρίως μενού, επιλέγουμε **Data, Weight Cases** και οδηγούμεθα στο παρακάτω πλαίσιο διαλόγου.



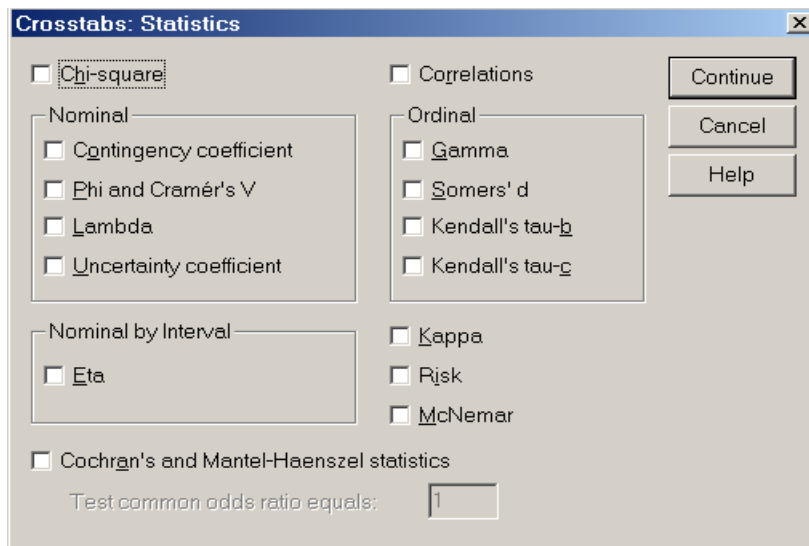
Επιλέγουμε **Weight cases by** και, στο πεδίο **Frequency Variable**, δηλώνουμε z (το όνομα της μεταβλητής με τις συχνότητες του πίνακα ταξινόμησης) και πιέζουμε **OK**.

Σημείωση: Όταν θα έχουμε ολοκληρώσει την ανάλυση (στο τέλος του παραδείγματος) πρέπει να επανέλθουμε εδώ και να επιλέξουμε **Do not weight cases**.

Στην συνέχεια, επιλέγουμε **Analyze, Descriptive Statistics, Crosstabs** στο βασικό μενού και προκύπτει το παρακάτω πλαίσιο διαλόγου:



Στο πεδίο **Row(s)**, δίνουμε **x** και στο πεδίο **Column(s)** δίνουμε **y**. Επειδή το μέγεθος του δείγματος είναι μεγάλο, χρησιμοποιούμε την κανονική προσέγγιση της διωνυμικής κατανομής της ελεγχουσυνάρτησης. Πιέζοντας το πλήκτρο **Statistics**, προκύπτει το επόμενο πλαίσιο διαλόγου:



Επιλέγοντας **McNemar** και πιέζοντας **Continue**, επιστρέφουμε στο προηγούμενο πλαίσιο από όπου παίρνουμε τα αποτελέσματα:

Case Processing Summary

	Cases		Missing	Total	
	N	Percent		N	Percent
X * Y	135	100.0%	0	.0%	135 100.0%

X * Y Crosstabulation
Count

		Y		Total
		1	2	
X	1	7	36	43
	2	30	62	92
Total		37	98	135

Chi-Square Tests

	Value	Exact Sig. (2-sided)
McNemar Test		.539 ^a
N of Valid Cases	135	

a Binomial distribution used.

Παρατηρούμε ότι, παρά το ότι δεν ζητήσαμε να χρησιμοποιηθεί η ακριβής κατανομή της στατιστικής συνάρτησης ελέγχου, το πακέτο παρέχει την τιμή του κρίσιμου επιπέδου για αμφίπλευρο έλεγχο. **Σημείωση:** Ο έλεγχος αυτός μπορεί να γίνει και από το πλαίσιο διαλόγου που χρησιμοποιούμε για τον προσημικό έλεγχο. Σε αυτή την περίπτωση, όμως, πρέπει να δημιουργήσουμε δύο μεταβλητές που έχουν τόσες τιμές όσες και οι παρατηρήσεις του δείγματος, όλες ίσες με 0 ή 1 ανάλογα με τις τιμές των αντιστοιχούντων σ' αυτές τις παρατηρήσεις του δείγματος. Ο τρόπος που δείξαμε προτιμάται γιατί είναι οικονομικότερος.

Λύση με το SAS: Εισάγουμε τα δεδομένα σε αρχείο με όνομα "examp231.txt", το οποίο θα περιέχει δύο στήλες: μια ως δείκτη της γραμμής του πίνακα διπλής εισόδου και μια ως δείκτη της στήλης.

Έτσι, το αρχείο θα περιέχει 7 γραμμές με τους κωδικούς «0 0», 36 γραμμές με τους κωδικούς «0 1», 30 γραμμές με τους κωδικούς «1 0» και 62 γραμμές με τους κωδικούς «1 1». Στην συνέχεια, πληκτρολογούμε τις εντολές

```
data data231;
infile 'c:\examp231.txt';
input x y;
run;
proc print;
run;

proc freq;
tables x*y /agree;
run;
```

Η εντολή **proc freq; tables x*y;** μας οδηγεί στον πίνακα διπλής εισόδου για τις μεταβλητές **x** και **y**, ενώ η υποεντολή **/agree;** μας οδηγεί στον πίνακα που ακολουθεί, όπου, μεταξύ άλλων αποτελεσμάτων, παρέχεται και το αποτέλεσμα εφαρμογής του ελέγχου McNemar.

```
The SAS System
TABLE OF X BY Y
```

X	Y		Total
	0	1	
0	7	36	43
	5.19	26.67	31.85
	16.28	83.72	
	18.92	36.73	
1	30	62	92
	22.22	45.93	68.15
	32.61	67.39	
	81.08	63.27	
Total	37	98	135
	27.41	72.59	100.00

```
STATISTICS FOR TABLE OF X BY Y

McNemar's Test
-----
Statistic = 0.545      DF = 1      Prob = 0.460
```

Παρατήρηση: Με την προσθήκη και στα δύο μέλη της ισότητας στην H_0 της πιθανότητας $P(X=1, Y=1)$, αυτή παίρνει την μορφή $H_0: P(X=1)=P(Y=1)$, που είναι ευκολότερο να ερμηνευθεί στο πλαίσιο πρακτικών εφαρμογών. Για τον λόγο αυτό, ο έλεγχος McNemar συχνά χρησιμοποιείται και για τον έλεγχο των υποθέσεων

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

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

ή, ισοδύναμα, των υποθέσεων

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

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

2.3.2 Ο Έλεγχος των Cox και Stuart για την Υπαρξη Τάσης σε μια Ακολουθία Παρατηρήσεων (Cox and Stuart Test for Trend)

Μια άλλη παραλλαγή του προσημικού ελέγχου προτάθηκε από τους Cox και Stuart το 1955 για τον έλεγχο ύπαρξης τάσης σε μια ακολουθία παρατηρήσεων. Μια ακολουθία παρατηρήσεων λέγεται ότι έχει τάση (*trend*), εάν οι τελευταίοι όροι της ακολουθίας τείνουν να έχουν μεγαλύτερες τιμές από τους πρώτους όρους της ακολουθίας (*αυξητική τάση, upwards trend*) ή μικρότερες τιμές από τις τιμές των πρώτων όρων της ακολουθίας (*πτωτική τάση, downwards trend*). Ο έλεγχος αυτός βασίζεται στην δημιουργία ζευγών παρατηρήσεων, των οποίων το πρώτο μέλος προέρχεται από τους πρώτους όρους της ακολουθίας και το δεύτερο μέρος από τους τελευταίους όρους της

ακολουθίας. Εάν υπάρχει τάση, το ένα μέλος κάθε ζεύγους θα έχει την τάση να έχει υψηλότερη ή χαμηλότερη τιμή από το άλλο. Από το άλλο μέρος, αν δεν υπάρχει τάση και η ακολουθία των παρατηρήσεων στην πραγματικότητα αποτελείται από παρατηρήσεις πάνω σε ανεξάρτητες και ισόνομες τυχαίες μεταβλητές, τότε οι τιμές ενός μέλους του ζεύγους δεν θα τείνουν να είναι μεγαλύτερες ή μικρότερες από τις τιμές του άλλου μέλους.

Τα δεδομένα αποτελούνται από n' παρατηρήσεις πάνω σε μια ακολουθία τυχαίων μεταβλητών $X_1, X_2, \dots, X_{n'}$, των οποίων οι δείκτες υποδεικνύουν την σειρά με την οποία οι τυχαίες μεταβλητές παρατηρήθηκαν. Για τον έλεγχο της υπόθεσης ότι οι όροι της ακολουθίας παρουσιάζουν τάση, συνδυάζουμε τις τυχαίες μεταβλητές σε ζεύγη της μορφής $(X_1, X_{1+c}), (X_2, X_{2+c}), \dots, (X_{n'-c}, X_{n'})$, όπου

$$c = \begin{cases} n'/2, & \text{αν } n' \text{ άρτιος} \\ (n'+1)/2, & \text{αν } n' \text{ περιττός} \end{cases}$$

(Δηλαδή, η μεσαία τυχαία μεταβλητή αγνοείται με αυτό τον τρόπο συνδυασμού των μεταβλητών, αν η τιμή του n' είναι περιττός αριθμός). Στην συνέχεια, σε κάθε ζεύγος (X_i, X_{i+c}) , ($i = 1, 2, \dots, n'-c$) αντιστοιχίζεται ένα "+" εάν $X_i < X_{i+c}$, ένα "-" εάν $X_i > X_{i+c}$ ή ένα "0" αν $X_i = X_{i+c}$. Έστω n ο αριθμός των "+" και "-" ζευγών.

Οι υποθέσεις που συνήθως ενδιαφερόμαστε να ελέγξουμε έχουν τη μορφή:

A. (Αμφίπλευρος έλεγχος)

H_0 : Δεν υπάρχει τάση

H_1 : Υπάρχει τάση

B. (Μονόπλευρος έλεγχος)

H_0 : Δεν υπάρχει αυξητική τάση

H_1 : Υπάρχει αυξητική τάση

Γ. (Μονόπλευρος έλεγχος)

H_0 : Δεν υπάρχει πτωτική τάση

H_1 : Υπάρχει πτωτική τάση.

Η μορφή της μηδενικής υπόθεσης στην περίπτωση Α υπαινίσσεται ότι η σειρά εμφάνισης ή καταγραφής των παρατηρήσεων X_1, X_2, \dots, X_n πάνω στην τυχαία μεταβλητή X είναι τυχαία, αφού, σύμφωνα με την H_0 , η κατανομή των παρατηρήσεων $X_i, i = 1, 2, \dots, n$ δεν συνδέεται με την χρονική στιγμή ή την σειρά κατά την οποία αυτές ελήφθησαν. Η εναλλακτική υπόθεση αντίθετα, υπαινίσσεται ότι η κατανομή των $X_i, i = 1, 2, \dots, n$ συνδέεται με τον χρόνο με τρόπο ώστε, με την πάροδο του χρόνου οι τιμές της ακολουθίας παρατηρήσεων τείνουν να είναι μεγαλύτερες ή μικρότερες. Ο έλεγχος αυτός μπορεί, επομένως, να χρησιμοποιηθεί για την ανίχνευση οποιασδήποτε μορφής μη τυχαίου σχήματος στην σειρά πραγματοποίησης των παρατηρήσεων X_1, X_2, \dots, X_n , όπως ημιτονοειδή κύμανση ή άλλης μορφής περιοδικότητα. Οι παρατηρήσεις ανακατατάσσονται ώστε οι μικρότερες (σύμφωνα με το προς έλεγχο σχήμα) να είναι πιο κοντά στην αρχή της ακολουθίας και οι μεγαλύτερες προς το τέλος.

Προφανώς, η ύπαρξη αυξητικής ή πτωτικής τάσης στην προκύπτουσα ακολουθία αποτελεί ένδειξη παρουσίας μη τυχαίου σχήματος στην αρχική ακολουθία παρατηρήσεων.

Είναι προφανές, ότι οι παραπάνω υποθέσεις μπορούν να γραφούν με την μορφή:

Α. (Αμφίπλευρος έλεγχος):

H_0 : $P(+)=P(-)$

$$H_1: P(+)\neq P(-)$$

B. (Μονόπλευρος έλεγχος):

$$H_0: P(+)\leq P(-)$$

$$H_1: P(+)>P(-)$$

Γ. (Μονόπλευρος έλεγχος):

$$H_0: P(+)\geq P(-)$$

$$H_1: P(+)<P(-)$$

Η στατιστική συνάρτηση T, όπως και στον προσημικό έλεγχο, ορίζεται ως ο αριθμός των "+" ζευγών (ζευγών, δηλαδή, όπου $X_{i+c} > X_i$) και ο έλεγχος διεξάγεται όπως περιγράφηκε στην παράγραφο 2.3.

Παράδειγμα 2.3.2: Σε κάποια πόλη, ο ρυθμός των θανάτων ανά 100.000 κατοίκους που οφείλεται σε τροχαία ατυχήματα παρουσίασε τις εξής τιμές για τα τελευταία 15 χρόνια:

17.3 17.9 18.4 18.1 18.3 19.6 18.6
19.2 17.7 20.0 19.0 18.8 19.3 20.2 19.9

Είναι βάσιμο να υποθέσει κανείς ότι ο ρυθμός θανάτων, που οφείλεται σε τροχαία ατυχήματα, τείνει να αυξάνει; ($\alpha=0.05$).

Λύση: Είναι προφανές ότι οι υποθέσεις, των οποίων ο έλεγχος θα οδηγήσει στην απάντηση του ερωτήματος, έχουν την μορφή:

$$H_0: \text{Δεν υπάρχει αυξητική τάση}$$

$$H_1: \text{Υπάρχει αυξητική τάση.}$$

Από τα δεδομένα, προκύπτει ότι $n'=15$, οπότε $c = (15+1)/2 = 8$.

Αγνοώντας, επομένως, την όγδοη παρατήρηση στην σειρά με την οποία εμφανίζονται οι παρατηρήσεις, σχηματίζουμε τα ζεύγη παρατηρήσεων:

(X_i, X_{i+8})	Πρόσημο ζεύγους
(17.3, 17.7)	+
(17.9, 20.0)	+
(18.4, 19.0)	+
(18.1, 18.8)	+
(18.3, 19.3)	+
(19.6, 20.2)	+
(18.6, 19.9)	+

Στον παραπάνω πίνακα, έχει θεωρηθεί ως "+" ζεύγος το ζεύγος (X_i, X_{i+c}) , για το οποίο $X_i < X_{i+c}$. Προκύπτει, επομένως, ότι οι αρχικές υποθέσεις είναι ισοδύναμες με τις υποθέσεις:

$$H_0: P(+) \leq 1/2$$

$$H_1: P(-) > 1/2.$$

Ορίζοντας, κατά συνέπεια, ως στατιστική συνάρτηση T τον αριθμό των "+" ζευγών και παρατηρώντας ότι κάτω από την μηδενική υπόθεση η κατανομή της T είναι η διωνυμική με παραμέτρους $n=7$ και $p=1/2$, έχουμε ότι η κρίσιμη περιοχή του ελέγχου μεγέθους περίπου ίσου με α ορίζεται από την ανισότητα $T \geq n-t$, όπου t ορίζεται έτσι ώστε

$$P(T \geq n-t \mid n=7, p=1/2) = \alpha_1 \cong \alpha=0.05$$

ή, ισοδύναμα, έτσι ώστε

$$P(T \leq t \mid n=7, p=1/2) = \alpha_1 \cong 0.05.$$

Από τον πίνακα της διωνυμικής κατανομής, προκύπτει ότι για $\alpha_1 = 0.0625$, $t=1$. Επομένως, ο κανόνας απόφασης διατυπώνεται ως εξής:

Η υπόθεση H_0 απορρίπτεται σε επίπεδο σημαντικότητας $\alpha=0.0625$, αν $T \leq 1$ ή αν $T \geq n-1=7-1=6$. Επειδή η παρατηρηθείσα τιμή της στατιστικής συνάρτησης βρίσκεται μέσα στην κρίσιμη περιοχή ($\tau=7$), η μηδενική υπόθεση απορρίπτεται σε επίπεδο σημαντικότητας $\alpha=0.0625$.

Το κρίσιμο επίπεδο $\hat{\alpha}$ έχει μια πολύ χαμηλή τιμή. Πράγματι, από τον πίνακα 1 του παραρτήματος, προκύπτει ότι

$$\hat{\alpha} = P(T \geq 7 \mid n=7, p=1/2) = 1 - P(T \leq 6 \mid n=7, p=1/2) = 0.0078.$$

Επομένως, η μηδενική υπόθεση θα απορρίπτεται σε κάθε επίπεδο σημαντικότητας α του οποίου η τιμή είναι μεγαλύτερη ή ίση από την τιμή 0.78%.

Λύση με το MINITAB: Ο έλεγχος Cox-Stuart για την ύπαρξη τάσης δεν είναι απ' ευθείας διαθέσιμος από το MINITAB. Κατά συνέπεια, σχηματίζουμε τα ζεύγη παρατηρήσεων της μορφής (X_i, X_{i+c}) και αποθηκεύουμε τις τιμές των μελών των ζευγών αυτών σε δύο μεταβλητές. Στην συνέχεια, θεωρούμε την διαφορά τους (δεύτερο μέλος - πρώτο μέλος) και επιλέγουμε **1-Sample Sign**. Τα αποτελέσματα του ελέγχου είναι:

Sign Test for Median

Sign test of median = 0.00000 versus > 0.00000

	N	Below	Equal	Above	P	Median
D	7	0	0	7	0.0078	0.7000

Αξίζει να σημειωθεί ότι ως εναλλακτική υπόθεση θεωρήσαμε την υπόθεση ότι η διάμεσος του πληθυσμού των διαφορών είναι θετική. Η

τιμή του κρίσιμου επιπέδου, όπως παρέχεται στήλη **P**, είναι πολύ χαμηλή. Επομένως, με βάση τα δεδομένα, η υπόθεση ύπαρξης ανοδικής τάσης δεν φαίνεται να είναι εύλογη.

Λύση με το SPSS: Επειδή, όπως έχει ήδη παρατηρηθεί, το SPSS δεν διαθέτει απ' ευθείας τον έλεγχο Cox-Stuart για την ύπαρξη τάσης, πρέπει να καταχωρίσουμε τα δεδομένα μας στην κατάλληλη μορφή για προσημικό έλεγχο. Θα χρησιμοποιήσουμε, δηλαδή, δύο μεταβλητές, μία για το πρώτο μισό του δείγματος και μία για το δεύτερο μισό. Στο παράδειγμα μας έχουμε 15 τιμές. Κατά συνέπεια, η μεσαία παρατήρηση αγνοείται. Στις μεταβλητές **x** και **y** καταχωρίζουμε τις πρώτες και τελευταίες 7 τιμές του δείγματος αντίστοιχα. Κατόπιν, εκτελούμε προσημικό έλεγχο, ζητώντας **exact** έλεγχο (αφού το δείγμα είναι μικρό), με τα εξής αποτελέσματα:

Frequencies

		N
Y - X	Negative Differences ^a	0
	Positive Differences ^b	7
	Ties ^c	0
	Total	7

a Y < X

b Y > X

c Y = X

Test Statistics^b

	Y - X
Exact Sig. (2-tailed)	.016 ^a
Exact Sig. (1-tailed)	.008
Point Probability	.008

a Binomial distribution used.

b Sign Test

Λύση με το SAS: Προκειμένου να διεξαγάγουμε τον έλεγχο με το SAS, καταχωρίζουμε σε δύο μεταβλητές τα ζεύγη παρατηρήσεων και διεξάγουμε προσημικό έλεγχο.

Στο παράθυρο εντολών πληκτρολογούμε


```

data cox_stuart;
input x y @@;
di f=y-x;
cards;
17.3 17.7 17.9 20 18.4 19 18.1 18.8 18.3 19.3 19.6 20.2 18.6 19.9
;
run;
proc univariate;
var di f;
run;

```

Τα αποτελέσματα του ελέγχου είναι τα εξής:

M(Si gn) 3.5 Pr>=|M| 0.0156

όπου στο πεδίο **Pr>=|M|** δίνεται το αμφίπλευρο παρατηρούμενο επίπεδο στατιστικής σημαντικότητας. Διαίρεση αυτού με 2, παρέχει το ζητούμενο κρίσιμο επίπεδο $\hat{\alpha} = 0.0078$.

Παράδειγμα 2.3.3: Τα τελευταία 34 χρόνια, το μέσο ύψος των ανδρών μιας περιοχής (σε εκατοστά) έχει παρουσιάσει τις τιμές που περιέχονται στον πίνακα που ακολουθεί.

170.8 (X ₁)	171.8 (X ₇)	173.0 (X ₁₃)	173.0 (X ₁₉)	173.5 (X ₂₅)	173.0 (X ₃₁)
171.5 (X ₂)	172.3 (X ₈)	173.0 (X ₁₄)	172.0 (X ₂₀)	173.3 (X ₂₆)	173.0 (X ₃₂)
171.0 (X ₃)	172.5 (X ₉)	172.3 (X ₁₅)	171.8 (X ₂₁)	173.3 (X ₂₇)	172.8 (X ₃₃)
170.3 (X ₄)	172.0 (X ₁₀)	171.5 (X ₁₆)	173.8 (X ₂₂)	173.8 (X ₂₈)	174.8 (X ₃₄)
171.0 (X ₅)	172.5 (X ₁₁)	171.5 (X ₁₇)	171.8 (X ₂₃)	173.8 (X ₂₉)	
170.5 (X ₆)	171.5 (X ₁₂)	172.0 (X ₁₈)	172.0 (X ₂₄)	172.5 (X ₃₀)	

Μπορεί να υποστηριχθεί ότι τα δεδομένα δείχνουν την ύπαρξη μιας τάσης αύξησης του μέσου ύψους των ανδρών της περιοχής αυτής;

Λύση: Προφανώς, οι υποθέσεις που ενδιαφερόμαστε να ελέγξουμε έχουν την μορφή

H_0 : Δεν υπάρχει αυξητική τάση

H_1 : Υπάρχει αυξητική τάση

Από τα δεδομένα, προκύπτει ότι $n' = 34$, οπότε $c = 34/2 = 17$ και μπορούμε να σχηματίσουμε $n' = 17$ ζεύγη παρατηρήσεων ως εξής

$$\begin{array}{c} \underline{(X_i, X_{i+17})} \\ (X_1, X_{18}) \\ (X_2, X_{19}) \\ (X_3, X_{20}) \\ \vdots \\ (X_{17}, X_{34}) \end{array}$$

Ορίζουμε ως «+» ενδεχόμενο το ενδεχόμενο $\{X_i < X_{i+17}\}$. Τότε, οι αρχικές υποθέσεις είναι ισοδύναμες με τις υποθέσεις

$$\begin{array}{l} H_0 : P(+) \leq 1/2 \\ H_1 : P(+) > 1/2. \end{array}$$

Η προφανής επιλογή ελεγχουσυνάρτησης για τον έλεγχο των υποθέσεων αυτών είναι η στατιστική συνάρτηση $T = \#$ των +.

Από την εσωτερική σύγκριση των μελών των ζευγών παρατηρήσεων που σχηματίσαμε, προκύπτει ο εξής πίνακας

X_i	170.8	171.5	171.0	170.3	171.0	170.5	171.8	172.3	172.5
X_{i+17}	172.0	173.0	172.0	171.8	173.8	171.8	172.0	173.5	173.3
πρόσημο	+	+	+	+	+	+	+	+	+
X_i	172.0	172.5	171.5	173.0	173.0	172.3	171.5	171.5	
X_{i+17}	173.3	173.8	173.8	172.5	173.0	173.0	172.8	174.8	
πρόσημο	+	+	+	-	0	+	+	+	

Προφανώς, $n' = 17$ και $n = 17 - \#$ των «0» ζευγών $\equiv 17-1=16$.

Επομένως, $T \sim \text{διων. } (n = 16, p = P(+))$. Άρα, κάτω από την H_0 , $T \sim \text{διων. } (n = 16, p = 1/2)$ και η κρίσιμη περιοχή του ελέγχου ορίζεται από την ανισότητα $T > t$, όπου t προσδιορίζεται από την σχέση

$$P(T > t \mid n = 16, p \leq 1/2) \cong \alpha.$$

Η τιμή του κρίσιμου επιπέδου του ελέγχου είναι

$$\begin{aligned} \hat{\alpha} &= \max P(T \geq 15 \mid n = 16, p \leq 1/2) \\ &= P(T \geq 15 \mid n = 16, p = 1/2) \\ &= 1 - P(T \leq 14 \mid n = 16, p = 1/2) \\ &= 1 - 0.9997 \\ &= 0.0003 \end{aligned}$$

Η τιμή αυτή είναι πολύ χαμηλή. Επομένως, ο ισχυρισμός ότι υπάρχει αυξητική τάση στο ύψος των ανδρών της περιοχής αυτής δεν είναι εύλογος.

Λύση με το MINITAB: Ο έλεγχος Cox-Stuart για την ύπαρξη τάσης δεν είναι απ' ευθείας διαθέσιμος από το MINITAB. Κατά συνέπεια, σχηματίζουμε τα ζεύγη παρατηρήσεων της μορφής (X_i, X_{i+c}) και αποθηκεύουμε τις τιμές των μελών των ζευγών αυτών σε δύο μεταβλητές. Στην συνέχεια, θεωρούμε την διαφορά τους (δεύτερο μέλος – πρώτο μέλος), την αποθηκεύουμε στην μεταβλητή **d** και επιλέγουμε **1-Sample Sign**. Τα αποτελέσματα του ελέγχου είναι:

Sign Test for Median

Sign test of median = 0.00000 versus > 0.00000

	N	Below	Equal	Above	P	Median
d	17	1	1	15	0.0003	1.300

Αξίζει να σημειωθεί ότι ως εναλλακτική υπόθεση, θεωρήσαμε την υπόθεση ότι η διάμεσος του πληθυσμού των διαφορών είναι θετική. Η τιμή του κρίσιμου επιπέδου, όπως παρέχεται από το πεδίο **P**, είναι πολύ

χαμηλή. Επομένως, με βάση τα δεδομένα, η υπόθεση ύπαρξης ανοδικής τάσης δεν φαίνεται να είναι εύλογη.

Λύση με το SPSS: Επειδή, όπως έχει ήδη παρατηρηθεί, το SPSS δεν διαθέτει απ' ευθείας τον έλεγχο Cox-Stuart για την ύπαρξη τάσης, καταχωρίζουμε τα δύο μισά του δείγματος σε δύο μεταβλητές (έστω **X1** για το πρώτο μισό και **X2** για το δεύτερο) και κάνουμε προσημικό έλεγχο. Επιλέγοντας κατά τα γνωστά, **Analyze, Nonparametric Tests, 2 Related Samples**, αφού δηλώσουμε τις δύο μεταβλητές, επιλέγουμε προσημικό έλεγχο ζητώντας να χρησιμοποιηθεί η ακριβής κατανομή της ελεγχοσυνάρτησης T. Τα αποτελέσματα που προκύπτουν είναι τα εξής:

Frequencies

		N
X2 - X1	Negative Differences ^a	1
	Positive Differences ^b	15
	Ties ^c	1
	Total	17

a X2 < X1

b X2 > X1

c X1 = X2

Test Statistics^b

	X2 - X1
Exact Sig. (2-tailed)	.001 ^a
Exact Sig. (1-tailed)	.000
Point Probability	.000

a Binomial distribution used.

b Sign Test

Παρατηρούμε ότι η τιμή του κρίσιμου επιπέδου του μονόπλευρου ελέγχου είναι πολύ μικρό (το SPSS το αναφέρει ως 0) και, κατά συνέπεια συμπεραίνουμε ότι η υπόθεση ότι δεν υπάρχει αυξητική τάση στις παρατηρήσεις δεν είναι εύλογη.

Λύση με το SAS: Προκειμένου να διεξαγάγουμε τον έλεγχο, καταχωρίζουμε σε δύο μεταβλητές τα ζεύγη παρατηρήσεων και κάνουμε προσημικό έλεγχο.

Στο παράθυρο εντολών πληκτρολογούμε

```
data cox_stuart;
input x y @@;
di f=y-x;
cards;
170.8 172.0 171.5 173.0 171.0 172.0 170.3 171.8 171.0 173.8 170.5 171.8
171.8 172.0
172.3 173.5 172.5 173.3 172.0 173.3 172.5 173.8 171.5 173.8 173.0 172.5
173.0 173.0
172.3 173.0 171.5 172.8 171.5 174.8
;
run;
proc univariate;
var di f;
run;
```

Τα αποτελέσματα που προκύπτουν είναι τα εξής:

M(Si gn) 7 Pr>=|M| 0.0005

Οπου, στο πεδίο **Pr>=|M|**, δίνεται το αμφίπλευρο παρατηρούμενο επίπεδο στατιστικής σημαντικότητας. Διάρθρωση αυτού με 2, δίνει το ζητούμενο κρίσιμο επίπεδο $\hat{\alpha} = 0.00025$.

2.3.3 Ο Προσημικός Έλεγχος για τον Έλεγχο Ύπαρξης Συσχέτισης (*The Sign Test for Correlation*)

Οι προηγούμενες ενότητες αναφέρονται σε δύο μόνο από τους πολλούς τρόπους με τους οποίους ο προσημικός έλεγχος μπορεί να εφαρμοσθεί για να ελέγξει διάφορες μορφές υποθέσεων. Μία ακόμη εφαρμογή του ελέγχου αυτού αναφέρεται στη χρησιμοποίησή του για τον έλεγχο ύπαρξης *sysxatishw (correlation)*, δηλαδή, για τον έλεγχο της υπόθεσης ότι υψηλές τιμές μιας τυχαίας μεταβλητής τείνουν να αντιστοιχούν σε υψηλές τιμές μιας άλλης τυχαίας μεταβλητής, ενώ χαμηλές τιμές της πρώτης τείνουν να αντιστοιχούν

σε χαμηλές τιμές της δεύτερης (*θετική συσχέτιση, positive correlation*) ή της υπόθεσης ότι υψηλές τιμές μιας τυχαίας μεταβλητής τείνουν να αντιστοιχούν σε χαμηλές τιμές μιας άλλης τυχαίας μεταβλητής, και χαμηλές τιμές της πρώτης τείνουν να αντιστοιχούν σε υψηλές τιμές της δεύτερης (*αρνητική συσχέτιση, negative correlation*). Η εφαρμογή αυτή του προσημικού ελέγχου εξετάστηκε από τους Cox και Stuart. Για το λόγο αυτό, ο προσημικός έλεγχος στην περίπτωση αυτή είναι γνωστός ως *έλεγχος των Cox και Stuart*.

Για την εφαρμογή του ελέγχου, τα ζεύγη των παρατηρήσεων ανακατατάσσονται, έτσι ώστε το ένα μέλος του ζεύγους (είτε το πρώτο είτε το δεύτερο) να διατάσσεται κατά αύξουσα σειρά μεγέθους των τιμών του. Εάν πράγματι υπάρχει συσχέτιση, το δεύτερο μέλος του ζεύγους θα επιδεικνύει μια αυξητική τάση, εάν η συσχέτιση είναι θετική, ενώ θα επιδεικνύει μία πτωτική τάση, εάν η συσχέτιση είναι αρνητική. Ο έλεγχος των Cox και Stuart μπορεί, στην συνέχεια, να χρησιμοποιηθεί για την ακολουθία των τιμών, η οποία σχηματίζεται από την προκύπτουσα ακολουθία τιμών του δευτέρου μέλους του ζεύγους.

Παράδειγμα 2.3.4: Εστω X ο αριθμός των παιχνιδιών που ένας παίκτης του basketball έπαιξε εντός έδρας σε μια συγκεκριμένη χρονιά. Εστω, επίσης, Y η μέση βαθμολογία του παίκτη για την χρονιά αυτή. Με βάση το τυχαίο δείγμα των παρατηρήσεων πάνω στο ζεύγος (X, Y) , που περιέχει ο πίνακας που ακολουθεί, θα μπορούσε να συμπεράνει κανείς ότι υπάρχει στατιστικά σημαντική γραμμική συσχέτιση μεταξύ των μεταβλητών X και Y ; ($\alpha=0.05$).

1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993
------	------	------	------	------	------	------	------	------	------	------	------

X_i	7	14	17	15	9	19	16	17	22	17	13	10
Y_i	.212	.232	.234	.210	.201	.256	.261	.247	.255	.241	.238	.235

Λύση: Οι προς έλεγχο υποθέσεις έχουν την μορφή:

H_0 : Δεν υπάρχει γραμμική συσχέτιση μεταξύ X και Y

H_1 : Υπάρχει γραμμική συσχέτιση μεταξύ X και Y .

Κατατάσσοντας τα ζεύγη παρατηρήσεων, έτσι ώστε οι τιμές της μεταβλητής Y να διατάσσονται κατά αύξουσα σειρά, προκύπτει ο πίνακας που ακολουθεί:

$Y^{(i)}$	X_i^*
0.201	4
0.210	15
0.212	7
0.232	14
0.234	17
0.235	10
0.238	13
0.241	17
0.247	14
0.255	22
0.256	19
0.261	16

Στον παραπάνω πίνακα, $Y^{(i)}$ συμβολίζει την i παρατήρηση στο διατεταγμένο δείγμα $Y^{(1)}, Y^{(2)}, \dots, Y^{(n)}$ των τιμών της Y , ενώ X_i^* ,

$i = 1, 2, \dots, n$, συμβολίζει την προκύπτουσα μετάθεση της ακολουθίας των τιμών της μεταβλητής X . Τότε, οι αρχικές υποθέσεις είναι ισοδύναμες με τις υποθέσεις:

H_0 : Δεν υπάρχει τάση στις τιμές της X

H_1 : Υπάρχει τάση στις τιμές της X .

Επειδή το πλήθος των παρατηρήσεων είναι άρτιο ($n=6$), σχηματίζουμε τα ζεύγη (X_i^*, X_{i+c}^*) , όπου $c=n/2=6/2=3$ με την βοήθεια του πίνακα που ακολουθεί.

X_i^*	X_{i+c}^*	Πρόσημο ζεύγους (X_i^*, X_{i+c}^*)
9	13	+
15	17	+
7	17	+
14	22	+
17	19	+
10	16	+

Τοποθετώντας τώρα το πρόβλημα στο πλαίσιο του προσημικού ελέγχου, παρατηρούμε ότι οι προς έλεγχο υποθέσεις παίρνουν την μορφή:

$$\left\{ \begin{array}{l} H_0 : P(+) = P(-) \\ H_1 : P(+) \neq P(-) \end{array} \right\} \Leftrightarrow \left\{ \begin{array}{l} H_0 : P(+) = 1/2 \\ H_1 : P(+) \neq 1/2 \end{array} \right\}$$

Ορίζοντας ως στατιστική συνάρτηση T , τον αριθμό των "+" ζευγών παρατηρήσεων, η κρίσιμη περιοχή του ελέγχου των παραπάνω υποθέσεων ορίζεται από τις ανισότητες $T \leq t$ και $T \geq n-t$, όπου t ορίζεται έτσι ώστε

$$P(T \leq t \mid n=6, p=1/2) \cong \alpha/2.$$

Η τιμή t προσδιορίζεται με τη βοήθεια του πίνακα 1 του παραρτήματος, αφού κάτω από την υπόθεση H_0 η κατανομή της στατιστικής συνάρτησης T είναι διωνυμική με παραμέτρους $n=6$ και $P(+)=1/2$, δηλαδή,

$$T \sim \text{διωνυμική} (n=6, p=P(+)=1/2).$$

Το κρίσιμο επίπεδο του ελέγχου $\hat{\alpha}$ είναι

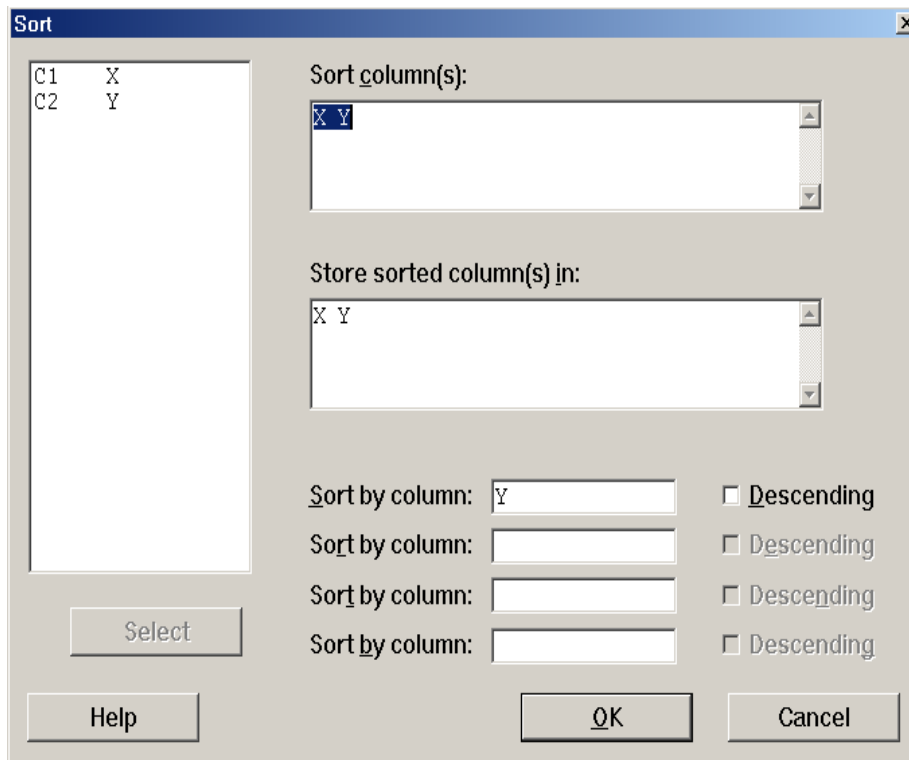
$$\hat{\alpha} = 2P(T \geq 6 \mid n=6, p=1/2) = 0.0312$$

(αφού $t=6 > np=3$). Επομένως, η μηδενική υπόθεση απορρίπτεται σε επίπεδο σημαντικότητας $\alpha=0.05$, αφού $\alpha = 0.050 > .0312 = \hat{\alpha}$. Δηλαδή, σε επίπεδο σημαντικότητας 5% τα δεδομένα παρέχουν ενδείξεις ύπαρξης γραμμικής συσχέτισης μεταξύ του αριθμού των παιχνιδιών που ο συγκεκριμένος παίκτης παίζει εντός έδρας και της μέσης βαθμολογίας του.

Λύση με το MINITAB: Ο έλεγχος Cox-Stuart για την ύπαρξη συσχέτισης δεν είναι επιλέξιμος απ' ευθείας από το MINITAB. Μπορούμε όμως να τον διεξαγάγουμε ως προσημικό έλεγχο για ένα πληθυσμό αρκεί να μετασχηματίσουμε κατάλληλα τα δεδομένα μας.

Οι διαδικασίες μετασχηματισμού είναι οι εξής: Καταχωρίζουμε τα δύο δείγματα σε δύο μεταβλητές (έστω X και Y). Κατόπιν, ταξινομούμε κατά αύξουσα τάξη μεγέθους την μία από τις δύο μεταβλητές (αυτήν με τις λιγότερες ισοβαθμίες τιμών), οπότε οι τιμές της άλλης «μετακινούνται» αντίστοιχα. Τέλος χωρίζουμε στα δύο τη δεύτερη μεταβλητή, καταχωρίζουμε τα δύο μισά σε δύο νέες μεταβλητές (έστω $Y1$ και $Y2$) και δημιουργούμε την μεταβλητή $D=Y2-Y1$. Θεωρώντας τις τιμές της ως ένα δείγμα, διεξάγουμε προσημικό έλεγχο κατά τα ήδη γνωστά.

Το μόνο νέο σημείο αναφέρεται στον τρόπο με τον οποίο ταξινομούμε μια μεταβλητή με ταυτόχρονη μετακίνηση των τιμών των άλλων διαθέσιμων μεταβλητών. Θεωρώντας τις μεταβλητές στις οποίες έχουμε καταχωρίσει τα δείγματα και επιλέγοντας **Manip, Sort** στο βασικό μενού, εμφανίζεται το ακόλουθο πλαίσιο διαλόγου:



Στο πεδίο **Sort column(s)**, δηλώνουμε τις μεταβλητές των οποίων οι τιμές θα «μετακινηθούν» (συμπεριλαμβανομένης αυτής που θα ταξινομηθεί). Στο πεδίο **Store sorted column(s) in**, δηλώνουμε πού θα καταχωρισθούν οι προκύπτουσες μεταβλητές με τα στοιχεία των αρχικών μεταβλητών να έχουν μετακινηθεί στις νέες θέσεις τους. Επειδή οι αρχικές μεταβλητές δεν χρησιμεύουν στην συνέχεια, ζητάμε από το πρόγραμμα να καταχωρίσει τις «νέες» μεταβλητές στην θέση των παλιών. Τέλος, παρατηρούμε ότι υπάρχουν αρκετά (τέσσερα)

πεδία **Sort by column**. Αυτό σημαίνει ότι το SPSS δίνει την δυνατότητα ταξινόμησης των δεδομένων ως προς περισσότερες από μία μεταβλητές. Εμείς εδώ χρειαζόμαστε ταξινόμηση μόνο ως προς την μεταβλητή **Y** την οποία και δηλώνουμε. Μη επιλέγοντας **Descending**, η ταξινόμηση θα γίνει κατά αύξουσα τάξη μεγέθους.

Συνεχίζοντας με τους υπόλοιπους χειρισμούς που απαιτούνται για την διεξαγωγή του προσημικού ελέγχου, προκύπτουν τα εξής αποτελέσματα:

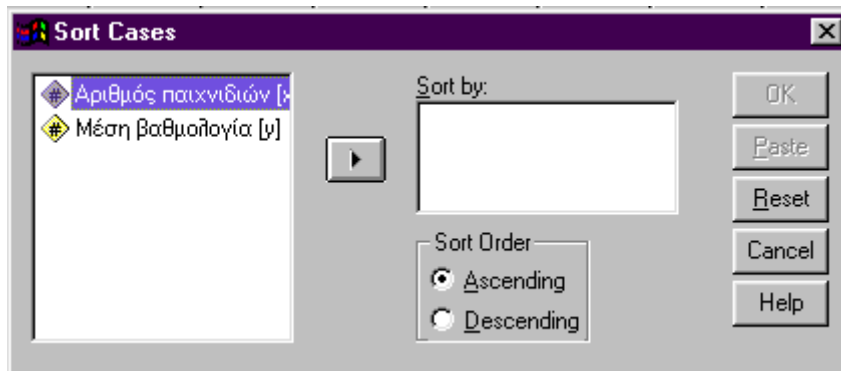
Sign Test for Median

Sign test of median = 0.00000 versus not = 0.00000

	N	Below	Equal	Above	P	Median
D	6	0	0	6	0.0312	5.000

Λύση με το SPSS: Με το SPSS δεν είναι δυνατή η απ' ευθείας διεξαγωγή του ελέγχου Cox-Stuart για την ύπαρξη συσχέτισης. Όπως και στην περίπτωση του MINITAB, μπορούμε να κάνουμε προσημικό έλεγχο, αφού καταχωρίσουμε τα δεδομένα με ειδική μορφή. Συγκεκριμένα, καταχωρίζουμε τα δύο δείγματα σε δύο μεταβλητές (έστω **x** και **y**). Κατόπιν, ταξινομούμε κατά αύξουσα τάξη μεγέθους τις τιμές της μεταβλητής **y**. Με την ταξινόμηση, αλλάζουν θέση και οι τιμές της **x** «ακολουθώντας» η κάθε μία το «ταίρι» της στο ζεύγος τιμών από το οποίο προήλθε. Το προκύπτον δείγμα τιμών το οποίο τώρα αποτελεί την μεταβλητή **x**, το χωρίζουμε στην μέση και τα δύο μισά του καταχωρίζουμε σε δύο νέες μεταβλητές (έστω **x1** και **x2**). Χρησιμοποιώντας τις δύο αυτές μεταβλητές διεξάγουμε προσημικό έλεγχο.

Για να ταξινομήσουμε την y και ταυτόχρονα να αλλάξουν θέση οι τιμές της x δηλώνουμε **Data, Sort Cases** στο βασικό μενού. Αυτό μας οδηγεί στο εξής πλαίσιο διαλόγου:



Στο πεδίο **Sort by**, δηλώνουμε y και, στο πεδίο **Sort Order**, επιλέγουμε **Ascending** για ταξινόμηση κατά αύξουσα τάξη μεγέθους. Πιέζοντας **OK**, το παράθυρο δεδομένων παίρνει την μορφή:

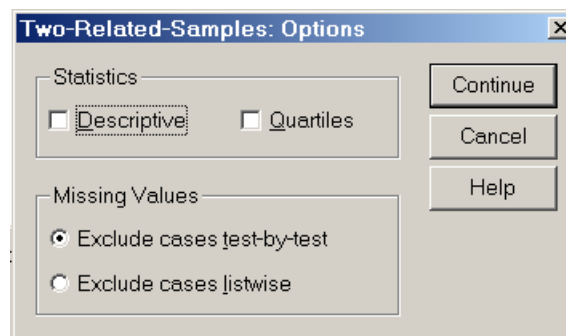
	x	y
1	9	,201
2	15	,210
3	7	,212
4	14	,232
5	17	,234
6	10	,235
7	13	,238
8	17	,241
9	17	,247
10	22	,255
11	19	,256
12	16	,261

Παρατηρούμε ότι τα ζεύγη τιμών διατηρούνται, αλλά τώρα οι τιμές της y είναι σε αύξουσα σειρά. Απομένει τώρα να καταχωρίσουμε τις πρώτες έξι τιμές της x σε μια μεταβλητή και τις υπόλοιπες έξι σε μία

άλλη μεταβλητή. Αυτό γίνεται εύκολα με **Copy, Paste**, οπότε η τελική μορφή του παραθύρου δεδομένων είναι:

	x	y	x1	x2
1	9	,201	9,00	13,00
2	15	,210	15,00	17,00
3	7	,212	7,00	17,00
4	14	,232	14,00	22,00
5	17	,234	17,00	19,00
6	10	,235	10,00	16,00
7	13	,238	.	.
8	17	,241	.	.
9	17	,247	.	.
10	22	,255	.	.
11	19	,256	.	.
12	16	,261	.	.
13				

Πριν προχωρήσουμε στην διεξαγωγή του προσημικού ελέγχου με βάση τις μεταβλητές **x1, x2**, είναι χρήσιμο να παρατηρηθεί ότι το SPSS θεωρεί ότι αυτές οι δύο μεταβλητές έχουν 6 ελλείπουσες τιμές (missing values). Για τον λόγο αυτό, στο παράθυρο διαλόγου του προσημικού ελέγχου, επιλέγουμε **options** και προκύπτει το παρακάτω πλαίσιο διαλόγου:



Στο πλαίσιο **Missing Values**, επιλέγουμε **Exclude cases test-by-test**. Με αυτήν την εντολή, το SPSS εξετάζει τις μεταβλητές που μετέχουν

στον έλεγχο που διεξάγει και εξαιρεί όσες παρατηρήσεις του δείγματος έχουν ελλείπουσες τιμές (missing value) έστω και σε μία από αυτές τις μεταβλητές. Με την επιλογή **exclude cases listwise**, το SPSS εξαιρεί από τους ελέγχους τις ελλείπουσες τιμές κάποιας μεταβλητής ακόμα κι αν αυτή δεν μετέχει στον έλεγχο. Στο παράδειγμά μας, η συγκεκριμένη επιλογή θα είχε αποτέλεσμα, σε ένα έλεγχο με βάση τις μεταβλητές των x και y , να αγνοηθούν οι τελευταίες έξι παρατηρήσεις τους, παρά το ότι δεν είναι ελλείπουσες (οι μεταβλητές x και y δεν έχουν ελλείπουσες τιμές).

Διεξάγοντας προσημικό έλεγχο, προκύπτουν τα εξής αποτελέσματα:

Frequencies

		N
X2 - X1	Negative Differences ^a	0
	Positive Differences ^b	6
	Ties ^c	0
	Total	6

a $X2 < X1$

b $X2 > X1$

c $X1 = X2$

Test Statistics^b

	X2 - X1
Exact Sig. (2-tailed)	,031 ^a
Exact Sig. (1-tailed)	,016
Point Probability	,016

a Binomial distribution used.

b Sign Test

Παρατήρηση: Η ανακατάταξη των ζευγών παρατηρήσεων θα πρέπει να γίνεται έτσι ώστε να διατάσσονται κατά αύξουσα σειρά μεγέθους οι τιμές εκείνου του μέλους του ζεύγους στο οποίο παρατηρούνται οι λιγότερες περιπτώσεις ισότητας τιμών (*ισοβαθμίες, ties*). Έχοντας διαλέξει το μέλος του ζεύγους, ως προς το οποίο θα γίνει η ανακατάταξη, διαλέγουμε εκείνη την διάταξη των τιμών που είναι

λιγότερο πιθανή να οδηγήσει σε απόρριψη της H_0 . Ας θεωρήσουμε, για παράδειγμα, τα ζεύγη τιμών

X: 5 3 6 3 2 3 2 1 4

Y: 7 2 5 1 1 6 3 4 8

Οι προς έλεγχο υποθέσεις έχουν τη μορφή:

H_0 : Δεν υπάρχει θετική γραμμική συσχέτιση

H_1 : Υπάρχει θετική γραμμική συσχέτιση.

Παρατηρούμε ότι η τυχαία μεταβλητή Y έχει τις λιγότερες περιπτώσεις ισότητας τιμών της. Επομένως, ανακατατάσσουμε τα ζεύγη τιμών, έτσι ώστε να διατάσσονται κατά αύξουσα τάξη μεγέθους οι τιμές αυτής της μεταβλητής:

$Y^{(i)}$	X_i^*	
1	;	(3 ή 2)
1	;	(2 ή 3)
2	3	
3	2	
4	1	
5	6	
6	3	
7	5	
8	4	

Στην προκύπτουσα μετάθεση των τιμών της X οι δύο πρώτοι όροι θα είναι 3, 2 αντίστοιχα, γιατί τα ζευγάρια (1, 3) και (1, 2) είναι λιγότερο πιθανά να οδηγήσουν σε απόρριψη της H_0 .

2.4 ΔΙΑΣΤΗΜΑΤΑ ΕΜΠΙΣΤΟΣΥΝΗΣ ΓΙΑ ΜΙΑ ΠΙΘΑΝΟΤΗΤΑ

Η μέθοδος για τον προσδιορισμό ενός διαστήματος εμπιστοσύνης για την άγνωστη πιθανότητα $p=P(A)$ ενός ενδεχομένου A συνδέεται στενά με τον διωνυμικό έλεγχο.

Ένα δείγμα παρατηρήσεων πάνω σε n ανεξάρτητες δοκιμές εξετάζεται και καταγράφεται η συχνότητα T με την οποία το συγκεκριμένο ενδεχόμενο εμφανίζεται. Οι n δοκιμές είναι αμοιβαία ανεξάρτητες και η πιθανότητα p του εν λόγω ενδεχομένου παραμένει σταθερή από δοκιμή σε δοκιμή. Είναι προφανές, ότι η κατανομή της τυχαίας μεταβλητής T είναι η διωνυμική με παραμέτρους n και p . Ένα διάστημα εμπιστοσύνης για την πιθανότητα p είναι προφανές ότι θα αποτελείται από όλες τις τιμές του p_0 οι οποίες είναι τέτοιες ώστε τα δεδομένα του δείγματος θα οδηγούσαν σε μη απόρριψη της μηδενικής υπόθεσης

$$H_0 : p = p_0$$

έναντι της εναλλακτικής

$$H_1 : p \neq p_0.$$

Ειδικότερα, εάν επιθυμούμε την κατασκευή ενός $100(1-\alpha)\%$ διαστήματος εμπιστοσύνης, παρατηρούμε το δείγμα των αποτελεσμάτων των n δοκιμών, καταγράφουμε την τιμή t της μεταβλητής T και ρωτάμε:

«Για την δοθείσα τιμή t της μεταβλητής T , ποιες είναι οι τιμές εκείνες που θα μπορούσε να πάρει το p_0 στην υπόθεση

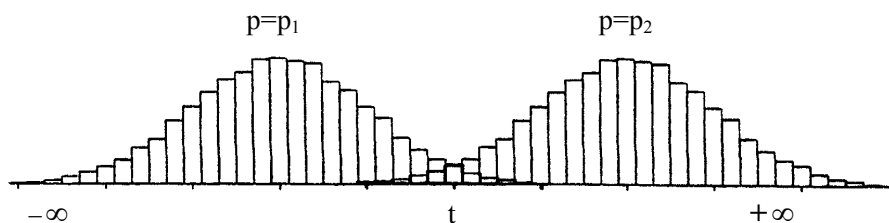
$$H_0 : p = p_0$$

ώστε ένας αμφίδρομος διωνυμικός έλεγχος σε επίπεδο σημαντικότητας α να μην οδηγήσει σε απόρριψη της μηδενικής υπόθεσης; »

Το σύνολο αυτών των τιμών αποτελούν το $100(1-\alpha)\%$ διάστημα εμπιστοσύνης. Οι τιμές του p_0 που οδηγούν σε απόρριψη της μηδενικής υπόθεσης H_0 δεν θα ανήκουν στο διάστημα εμπιστοσύνης. Επομένως, αν (p_1, p_2) είναι ένα $100(1-\alpha)\%$ διάστημα εμπιστοσύνης για την παράμετρο p , τότε οι προτάσεις « H_0 σε επίπεδο σημαντικότητας α » και « $p_0 \notin (p_1, p_2)$ » είναι ισοδύναμες.

Καθορισμός των Άκρων p_1, p_2 του $100(1-\alpha)\%$ Διαστήματος Εμπιστοσύνης

Τα άκρα p_1, p_2 του $100(1-\alpha)\%$ διαστήματος εμπιστοσύνης καθορίζονται έτσι ώστε η παρατηρούμενη τιμή t της μεταβλητής T να αποτελεί ταυτόχρονα ακραία τιμή των διωνυμικών μεταβλητών $T_1 \sim \text{διωνυμική}(n, p_1)$ και $T_2 \sim \text{διωνυμική}(n, p_2)$, δηλαδή $P(T_1 \geq t | n, p = p_1) \cong \alpha/2$ και $P(T_2 \geq t | n, p = p_2) \cong \alpha/2$.



Σχήμα 2.4.1

Επομένως, τα άκρα p_1 και p_2 του διαστήματος εμπιστοσύνης ορίζονται έτσι ώστε

$$P(T \leq t | n, p = p_2) \cong \alpha/2$$

και

$$P(T \geq t \mid n, p = p_1) \cong 1 - \alpha/2$$

ή, ισοδύναμα,

$$P(T \leq t-1 \mid n, p_1) \cong \alpha/2 .$$

Επομένως, ο κανόνας απόρριψης του παραπάνω ελέγχου σε επίπεδο σημαντικότητας α μπορεί να διατυπωθεί μέσω του $100(1-\alpha)\%$ διαστήματος εμπιστοσύνης (p_1, p_2) ως εξής:

Η υπόθεση H_0 απορρίπτεται σε επίπεδο σημαντικότητας α αν $p_0 \notin (p_1, p_2)$.

Παράδειγμα 2.4.1: Ένα τυχαίο δείγμα 10 τέως καπνιστών έδειξε ότι, ένα χρόνο μετά, 6 άτομα ξανάρχισαν το κάπνισμα. Να κατασκευασθεί ένα 99% διάστημα εμπιστοσύνης για το πραγματικό ποσοστό των τέως καπνιστών που ξαναρχίζουν το κάπνισμα.

Λύση: Θεωρούμε ως επιτυχία την επιστροφή στο κάπνισμα. Η παρατηρηθείσα τιμή της στατιστικής συνάρτησης $T =$ αριθμός τέως καπνιστών που ξανάρχισαν το κάπνισμα είναι $t=6$. Έστω p το πραγματικό ποσοστό των τέως καπνιστών, που ξαναρχίζουν το κάπνισμα. Τα άκρα p_1, p_2 του 99% διαστήματος εμπιστοσύνης για το p θα καθορισθούν από τις σχέσεις:

$$P(T \leq 6 \mid n = 10, p = p_2) \cong 0.005$$

και

$$P(T \leq 5 \mid n = 10, p = p_1) \cong 0.995$$

Από τον πίνακα της διωνυμικής κατανομής (πίνακας 1 του παραρτήματος) για $n=10$ και για την τιμή $t=6$ έχουμε ότι η τιμή p_2 του p

που αντιστοιχεί στην τιμή 0.005 της πιθανότητας $P(T \leq 6 | n = 10, p = p_2)$ βρίσκεται στο διάστημα (0.90, 0.95).

Συγκεκριμένα,

p :	0.90	p_2	0.95
$P(T \leq 6 n = 10, p)$:	0.0128	0.005	0.001

Επομένως, $p_2 = 0.90 + \Delta p$, όπου η τιμή του Δp προκύπτει από την σχέση

$$\frac{\Delta p}{0.95 - 0.9} = \frac{0.0128 - 0.005}{0.0128 - 0.001}$$

ίση με $\Delta p = 0.033$. Επομένως, $p_2 = 0.90 + 0.033 = 0.933$.

Με τον ίδιο τρόπο, προκύπτει ότι $p_1 = 0.189$. Κατά συνέπεια, το 99% διάστημα εμπιστοσύνης για την πιθανότητα p είναι το διάστημα (0.189, 0.933).

Αν στο προηγούμενο παράδειγμα μας ενδιέφερε να ελέγξουμε τις υποθέσεις:

$$H_0: p = 0.5$$

$$H_1: p \neq 0.5,$$

τότε, επειδή $p_0 \in (0.189, 0.933)$, η H_0 δεν θα απορριπτόταν σε επίπεδο σημαντικότητας 1%.

Λύση με το MINITAB: Στο MINITAB η κατασκευή διαστημάτων εμπιστοσύνης για ποσοστά γίνεται μέσω της διεξαγωγής διωνυμικού ελέγχου κατά τα ήδη γνωστά. Στο συγκεκριμένο παράδειγμα, όπου έχουμε 10 δοκιμές και 6 «επιτυχίες» (επανέναρξη καπνίσματος) το MINITAB δίνει τα εξής αποτελέσματα:

Test of $p = 0.5$ vs $p \text{ not} = 0.5$

Exact					
Sample	X	N	Sample p	99.0 % CI	P-Value
1	6	10	0.600000	(0.190916; 0.923232)	0.754

Άρα το 99% διάστημα εμπιστοσύνης για το ποσοστό αυτών που ξαναρχίζουν το κάπνισμα είναι (0.191, 0.923).

Λύση με το SPSS: Το SPSS δεν δίνει διαστήματα εμπιστοσύνης για το ποσοστό επιτυχίας μιας διωνυμικής κατανομής χωρίς την χρήση της κανονικής προσέγγισης. Πέραν τούτου, δεν επιτρέπει την «αντίστροφη» διαδικασία που ακολουθήθηκε παραπάνω για την εύρεση των ορίων εμπιστοσύνης p_1 και p_2 . Κατά συνέπεια, το SPSS δεν προσφέρεται για λύση παρόμοιων προβλημάτων.

Λύση με το SAS: Προκειμένου να υπολογίσουμε το ζητούμενο διάστημα εμπιστοσύνης χρησιμοποιώντας την συγκεκριμένη έκδοση του SAS, απαιτείται η εισαγωγή των παρατηρούμενων συχνοτήτων (6 άτομα που άρχισαν ξανά το κάπνισμα και 4 που το απέφυγαν) στην πρώτη γραμμή ενός πίνακα διπλής εισόδου, στην δεύτερη γραμμή του οποίου μπορούν να εισαχθούν οποιεσδήποτε θετικές ακέραιες τιμές. Αυτό επιτρέπει τον έμμεσο υπολογισμό ασυμπτωτικών και ακριβών διαστημάτων εμπιστοσύνης για ποσοστά ως εξής:

Στο παράθυρο εντολών πληκτρολογούμε:

```
data smoke;
input row col count;
cards;
1 1 6
1 2 4
2 1 1
2 2 1
;
run;
proc freq;
tables row*col / ri skdi ff alpha=0.01;
weight count;
run;
```

Με την υποεντολή **riskdiff**, ζητείται ο υπολογισμός των διαστημάτων εμπιστοσύνης και, με την υποεντολή **alpha=0.01**, καθορίζεται το επιθυμητό επίπεδο σημαντικότητας.

Οι παραπάνω εντολές οδηγούν στα εξής αποτελέσματα:

The SAS System

TABLE OF ROW BY COL

ROW	COL		Total
Frequency,	1,	2,	
Percent			
Row Pct			
Col Pct	1,	2,	Total
1	6	4	10
	50.00	33.33	83.33
	60.00	40.00	
	85.71	80.00	
2	1	1	2
	8.33	8.33	16.67
	50.00	50.00	
	14.29	20.00	
Total	7	5	12
	58.33	41.67	100.00

STATISTICS FOR TABLE OF ROW BY COL

Column 1 Risk Estimates

Ri sk	99% Confi dence ASE	99% Confi dence BOUNDS (Asymptoti c)	99% Confi dence BOUNDS (Exact)
Row 1	0.600	0.155	0.201 0.999 0.191 0.923
Row 2	0.500	0.354	0.000 1.000 0.003 0.997
Total	0.583	0.142	0.217 0.950 0.209 0.897
Di fference (Row 1 - Row 2)	0.100	0.386	-0.894 1.000

Column 2 Ri sk Estimates

Ri sk	99% Confi dence ASE	99% Confi dence BOUNDS (Asymptoti c)	99% Confi dence BOUNDS (Exact)
Row 1	0.400	0.155	0.001 0.799 0.077 0.809
Row 2	0.500	0.354	0.000 1.000 0.003 0.997
Total	0.417	0.142	0.050 0.783 0.103 0.791
Di fference (Row 1 - Row 2)	-0.100	0.386	-1.000 0.894

Sampl e Si ze = 12

Παρατηρούμε ότι δίνονται δύο πίνακες με διαστήματα εμπιστοσύνης. Ο πρώτος αφορά στην πρώτη στήλη του πίνακα (άτομα που καπνίζουν, βάσει

της εισαγωγής των δεδομένων) και περιέχει το διάστημα που ζητάμε. Από αυτόν τον πίνακα, μας ενδιαφέρει το πεδίο **Row 1**, δηλαδή η πρώτη γραμμή του πίνακα. (Η δεύτερη γραμμή περιέχει στοιχεία που δεν αφορούν τα δεδομένα μας, αλλά χρησιμοποιήθηκαν για τον έμμεσο υπολογισμό του διαστήματος εμπιστοσύνης). Συνεπώς, το ζητούμενο ακριβές διάστημα εμπιστοσύνης είναι το διάστημα (0.191, 0.923).

Παρατήρηση: Στην περίπτωση κατά την οποία το μέγεθος n του δείγματος είναι αρκετά μεγάλο ($n > 20$) και η κατανομή πιθανότητας της στατιστικής συνάρτησης T είναι περίπου συμμετρική ($np \geq 5$) και ($n(1-p) \geq 5$), το $100(1-\alpha)\%$ διάστημα εμπιστοσύνης μπορεί να προσδιορισθεί κατά προσέγγιση με βάση την συνάρτηση κατανομής της κανονικής κατανομής με παραμέτρους $\mu = np$ και $\sigma^2 = np(1-p)$, η οποία προσεγγίζει την διωνυμική κατανομή της T . Ισχύει, δηλαδή, ότι

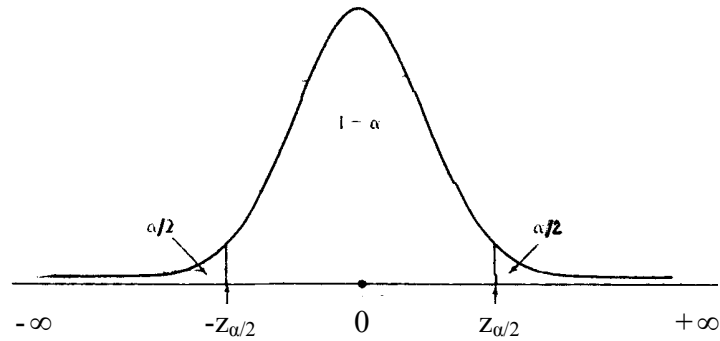
$$\frac{T - np}{\sqrt{np(1-p)}} \sim N(0,1).$$

Η τελευταία σχέση οδηγεί στο συμπέρασμα ότι

$$P(-z_{1-\alpha/2} < \frac{T - np}{\sqrt{np(1-p)}} < z_{1-\alpha/2}) = 1 - \alpha,$$

ή, ισοδύναμα,

$$P\left(\frac{T}{n} - z_{1-\alpha/2} \sqrt{\frac{p(1-p)}{n}} < p < \frac{T}{n} + z_{1-\alpha/2} \sqrt{\frac{p(1-p)}{n}}\right) = 1 - \alpha.$$



Σχήμα 2.4.2

Τα άκρα του διαστήματος εμπιστοσύνης περιέχουν την άγνωστη παράμετρο p . Είναι βέβαια δυνατόν να γραφεί η ανισότητα με τρόπο ώστε τα άκρα της να είναι ανεξάρτητα του p . Στην πράξη, όμως, επιτυγχάνουμε μία ικανοποιητική προσέγγιση αν αντικαταστήσουμε την παράμετρο p με την εκτιμήτριά της $\frac{T}{n}$. Δηλαδή, το $100(1-\alpha)\%$ διάστημα εμπιστοσύνης για την παράμετρο p έχει άκρα:

$$\frac{T}{n} \pm z_{1-\alpha/2} \sqrt{\frac{T/n (1 - T/n)}{n}}.$$

Παράδειγμα 2.4.2: Αν στο παράδειγμα με τους καπνιστές είχαμε ότι στους $n=100$ τέως καπνιστές $\tau=60$ ξανάρχισαν να καπνίζουν, το 99% διάστημα εμπιστοσύνης θα είχε άκρα

$$\frac{60}{100} + z_{0.995} \sqrt{\frac{60/100 (1 - 60/100)}{100}}.$$

Δηλαδή, το 99% διάστημα εμπιστοσύνης θα ήταν το $(0.474, 0.726)$.

Λύση με το MINITAB: Όπως παρατηρήθηκε στο προηγούμενο παράδειγμα, το MINITAB υπολογίζει διάστημα εμπιστοσύνης για ένα ποσοστό μέσω της διεξαγωγής κατάλληλου διωνυμικού ελέγχου. Εργαζόμενοι, όπως στο προηγούμενο παράδειγμα, αλλά ζητώντας από το πρόγραμμα να χρησιμοποιήσει την κανονική προσέγγιση της διωνυμικής κατανομής (αφού έχουμε μεγάλο δείγμα), οδηγούμεθα στα εξής αποτελέσματα:

Test of $p = 0.5$ vs $p \text{ not} = 0.5$

Sample	X	N	Sample p	99.0 % CI	Z-Value	P-Value
1	60	100	0.600000	(0.473811; 0.726189)	2.00	0.046

Κατά συνέπεια, το 99% διάστημα εμπιστοσύνης είναι (0.474 , 0.726).

Λύση με το SPSS: Για τον υπολογισμό διαστήματος εμπιστοσύνης με το SPSS, εργαζόμαστε ως εξής: Καταχωρίζουμε σε μία στήλη τα αποτελέσματα των δοκιμών. Συγκεκριμένα, καταχωρίζουμε σε μία μεταβλητή, έστω x , 60 φορές την τιμή 1 και κατόπιν 40 φορές την τιμή 0. Υπενθυμίζεται ότι η τιμή που καταλαμβάνει το πρώτο κελλί της στήλης θεωρείται ως «επιτυχία» από το πρόγραμμα. Για τον λόγο αυτό, πρώτα καταχωρίζουμε την τιμή 1. Προκειμένου να υπολογίσουμε το ζητούμενο διάστημα εμπιστοσύνης για την μελετώμενη παράμετρο p , θα ζητήσουμε από το SPSS ένα διάστημα εμπιστοσύνης για την μέση τιμή του δίτιμου πληθυσμού από τον οποίο προέρχεται το δείγμα. Αυτό επιτυγχάνεται με την επιλογή **Analyze, Descriptive Statistics, Explore** κατά τα ήδη γνωστά. Στον πίνακα αποτελεσμάτων, που δίνει το SPSS, περιλαμβάνονται και τα εξής:

Descriptives

			Statistic	Std. Error
X	Mean		.60	,049
	99% Confidence Interval for Mean	Lower Bound	.47	
		Upper Bound	.73	

Δηλαδή, το 99% διάστημα εμπιστοσύνης του ποσοστού των ατόμων που ξαναρχίζουν το κάπνισμα μέσα σε ένα χρόνο από την στιγμή που το κόβουν είναι (0.47 , 0.73).

Λύση με το SAS: Όπως και στο προηγούμενο παράδειγμα, για την κατασκευή του ζητούμενου διαστήματος, χρησιμοποιούμε τις εξής εντολές:

```
data smoke;
input row col count;
cards;
1 1 60
1 2 40
2 1 1
2 2 1
;
run;
proc freq;
table row*col / riskdiff alpha=0.01;
weight count;
run;
```

Το διάστημα που προκύπτει περιέχεται στους εξής πίνακες αποτελεσμάτων:

The SAS System
TABLE OF ROW BY COL

ROW	COL		Total
	1	2	
Frequency	60	40	100
Percent	58.82	39.22	98.04
Row Pct	60.00	40.00	
Col Pct	98.36	97.56	
	1	2	
Frequency	1	1	2
Percent	0.98	0.98	1.96
Row Pct	50.00	50.00	
Col Pct			

	1.64	2.44	
Total	61	41	102
	59.80	40.20	100.00

STATISTICS FOR TABLE OF ROW BY COL

		Column 1 Risk Estimates				
Risk	ASE	99% Confidence Bounds (Asymptotic)		99% Confidence Bounds (Exact)		

Row 1	0.600	0.049	0.474	0.726	0.466	0.724
Row 2	0.500	0.354	0.000	1.000	0.003	0.997
Total	0.598	0.049	0.473	0.723	0.466	0.721
Difference (Row 1 - Row 2)	0.100	0.357	-0.819	1.000		

		Column 2 Risk Estimates				
Risk	ASE	99% Confidence Bounds (Asymptotic)		99% Confidence Bounds (Exact)		

Row 1	0.400	0.049	0.274	0.526	0.276	0.534
Row 2	0.500	0.354	0.000	1.000	0.003	0.997
Total	0.402	0.049	0.277	0.527	0.279	0.534
Difference (Row 1 - Row 2)	-0.100	0.357	-1.000	0.819		

Sample Size = 102

Το ασυμπτωτικό 99% διάστημα εμπιστοσύνης είναι (0.474, 0.726). Το SAS, παρά το μεγάλο μέγεθος δείγματος, υπολογίζει και το ακριβές διάστημα εμπιστοσύνης (0.466, 0.724). Όπως μπορούμε να δούμε, τα άκρα του δεν διαφέρουν πολύ από αυτά του ασυμπτωτικού.

Σημείωση: Ο έλεγχος υποθέσεων για την μέση τιμή μιας διωνυμικής κατανομής με παραμέτρους N και p είναι ισοδύναμος με τον έλεγχο των υποθέσεων για την παράμετρο p. Αυτό είναι συνέπεια του γεγονότος ότι

$$\mu = E(X) = N p,$$

από το οποίο συνάγεται ότι

$$\left\{ \begin{matrix} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{matrix} \right\} \Leftrightarrow \left\{ \begin{matrix} H_0 : Np = Np_0 \\ H_1 : Np \neq Np_0 \end{matrix} \right\} \Leftrightarrow \left\{ \begin{matrix} H_0 : p = p_0 \\ H_1 : p \neq p_0 \end{matrix} \right\}.$$

2.5 ΕΛΕΓΧΟΣ ΠΟΣΟΣΤΙΑΙΩΝ ΣΗΜΕΙΩΝ ΜΙΑΣ ΚΑΤΑΝΟΜΗΣ

(*The Quantile Test*)

Ο διωνυμικός έλεγχος μπορεί να χρησιμοποιηθεί για τον έλεγχο υποθέσεων αναφερομένων στα ποσοστιαία σημεία μίας τυχαίας μεταβλητής. Στην περίπτωση αυτή ονομάζεται *Έλεγχος Ποσοστιαίων Σημείων* (*Quantile Test*). Για παράδειγμα, ενδέχεται να ενδιαφερόμαστε να εξετάσουμε αν ένα τυχαίο δείγμα τιμών μιας τυχαίας μεταβλητής X παρέχει ενδείξεις υπέρ της υπόθεσης ότι η διάμεσος της X είναι μεγαλύτερη από κάποια τιμή. Η κλίμακα μέτρησης είναι συνήθως τουλάχιστον κλίμακα διάταξης, παρά το ότι για τον διωνυμικό έλεγχο απαιτείται η (ασθενέστερη) ονομαστική κλίμακα.

Έστω X_1, X_2, \dots, X_n τυχαίο δείγμα παρατηρήσεων πάνω στην τυχαία μεταβλητή X και x_p το p -ποσοστιαίο σημείο της κατανομής της.

Υπενθυμίζεται ότι το p -ποσοστιαίο σημείο μίας κατανομής ορίζεται ως το σημείο x_p με την ιδιότητα

$$\begin{cases} P(X > x_p) \leq 1-p \text{ και } P(X < x_p) \leq p & (\text{διακριτή περίπτωση}) \\ P(X < x_p) = p & (\text{συνεχής περίπτωση}) \end{cases}$$

Έστω ότι x^* και p^* είναι κάποιες συγκεκριμένες τιμές, $0 < p^* < 1$. Οι υποθέσεις που ενδέχεται να ενδιαφερόμαστε να ελέγξουμε μπορούν να έχουν μία από τις εξής τρεις μορφές.

A. (Αμφίπλευρος Έλεγχος)

$$H_0 : x_{p^*} = x^*$$

$$H_1 : x_{p^*} \neq x^*$$

Β. (Μονόπλευρος Έλεγχος)

$$H_0 : x_{p^*} \geq x^*$$

$$H_1 : x_{p^*} < x^*$$

Γ. (Μονόπλευρος Έλεγχος)

$$H_0 : x_{p^*} \leq x^*$$

$$H_1 : x_{p^*} > x^*$$

Και στις τρεις περιπτώσεις, το δείγμα X_1, X_2, \dots, X_n των παρατηρήσεων πάνω στην μεταβλητή X μπορεί να θεωρηθεί ως μία ακολουθία αποτελεσμάτων n ανεξαρτήτων δοκιμών. Κάθε αποτέλεσμα ανήκει στην κατηγορία $\{X < x^*\}$ (επιτυχία) με πιθανότητα $p = P(X < x^*)$ ή στην κατηγορία $\{X > x^*\}$ (αποτυχία) με πιθανότητα $1 - p = P(X > x^*)$.

Α. (Αμφίπλευρος Έλεγχος)

Συνεχής περίπτωση

Από τον ορισμό του p -ποσοστιαίου σημείου, προκύπτει ότι η μηδενική υπόθεση $H_0 : x_{p^*} = x^*$ είναι ισοδύναμη με την υπόθεση $H_0 : P(X < x^*) = p^*$.

Επομένως, ένα τυχαίο δείγμα μεγέθους n θα περιέχει T επιτυχίες κάθε μία από τις οποίες έχει πιθανότητα $p = P(X < x)$. Κατά συνέπεια, η μηδενική υπόθεση $H_0 : x_{p^*} = x^*$ είναι ισοδύναμη με την υπόθεση $H_0 : p = p^*$. Άρα ο έλεγχος των υποθέσεων της μορφής A είναι

ισοδύναμος με έναν διωνυμικό έλεγχο για την παράμετρο p μιας διωνυμικής κατανομής με παραμέτρους n, p .

Παράδειγμα 2.5.1: Πρόκειται να κατασκευασθεί ένα καινούργιο αυτοκίνητο, που να χωράει άνετα ψηλούς οδηγούς εκτός από το 5% των πιο ψηλών οδηγών. Παλιότερες μελέτες δείχνουν ότι το 95ο ποσοστιαίο σημείο της κατανομής του ύψους των οδηγών της περιοχής ήταν (σε εκατοστά). Για να ελεγχθεί αν αυτό ισχύει ακόμα, επελέγη ένα τυχαίο δείγμα μεγέθους 100 και βρέθηκε ότι τα ύψη των 12 πιο ψηλών ατόμων ήταν τα εξής:

181.5	175	178.3	176.3
177.0	190.0	175.3	181.3
177.8	176.5	179.8	182.0

Είναι εύλογο να χρησιμοποιηθεί η τιμή 175.8 ως το 95ο ποσοστιαίο σημείο;

Λύση: Οι προς έλεγχο υποθέσεις είναι

$$\left. \begin{array}{l} H_0 : x_{0.95} = 175.8 \\ H_1 : x_{0.95} \neq 175.8 \end{array} \right\} \Leftrightarrow \left\{ \begin{array}{l} H_0 : P(X < 175.8) = 0.95 \\ H_1 : P(X < 175.8) \neq 0.95 \end{array} \right.$$

Για τον έλεγχο των υποθέσεων αυτών, η προφανής επιλογή στατιστικής συνάρτησης είναι η $T = \#$ των $X_i < 175.8$, η οποία ακολουθεί την διωνυμική κατανομή με παραμέτρους $n=100$ και $p=P(X<175.8)$. (Συμβολικά, $T \sim$ διων ($n=100, p= P(X<175.8)$)). Επομένως, κάτω από την H_0 , $T \sim$ διων ($n = 100, p= P(X<175.8)= 0.95$). Άρα, η κρίσιμη περιοχή του

ελέγχου μεγέθους α ορίζεται από τις ανισότητες $T \leq t_1$ και $T > t_2$, όπου t_1 και t_2 ορίζονται από τις σχέσεις

$$P(T \leq t_1 \mid n = 100, p = 0.95) \cong \alpha/2$$

και

$$P(T > t_2 \mid n = 100, p = 0.95) \cong \alpha/2.$$

Επειδή το μέγεθος του δείγματος είναι μεγάλο ($n=100$) και $np \geq 5$, $n(1-p) \geq 5$, μπορεί να χρησιμοποιηθεί η προσέγγιση της διωνυμικής κατανομής με $n=100$ και $p=0.95$ από την κανονική κατανομή $N(\mu = np = 95, \sigma^2 = np(1-p) = 4.75)$. Τότε,

$$\begin{aligned} \frac{\alpha}{2} &\cong P(T \leq t_1 \mid n = 100, p = 0.95) \\ &\cong P(X_{N(\mu, \sigma^2)} \leq t_1 \mid \mu = 95, \sigma^2 = 4.75) \\ &= P\left(Z \leq \frac{t_1 - 95}{\sqrt{4.75}}\right). \end{aligned}$$

$$\frac{t_1 - 95}{\sqrt{4.75}} = z_{\alpha/2} = -z_{1-\alpha/2}$$

και

$$\frac{t_2 - 95}{\sqrt{4.75}} = z_{1-\alpha/2}.$$

Η κρίσιμη περιοχή του ελέγχου μεγέθους α ορίζεται, επομένως, από τις ανισότητες $t \leq 95 - 2.18 z_{1-\alpha/2}$ και $t > 95 + 2.18 z_{1-\alpha/2}$

Η παρατηρούμενη τιμή της T είναι 90 ($=100-12+2$). Έτσι, η κρίσιμη περιοχή μεγέθους 0.05, που ορίζεται από τις σχέσεις $t \leq 90.7$ και $t > 99.2$, συνεπάγεται απόρριψη της H_0 σε επίπεδο σημαντικότητας 5%.

Το κρίσιμο επίπεδο του ελέγχου είναι

$$\begin{aligned}\hat{\alpha} &= 2P(T \leq 90 \mid n = 100, p = 0.95) \\ &= 2P\left(Z \leq \frac{90 - 95}{4.75}\right) \\ &= 2P(Z \leq -2.5) \\ &= 2(1 - P(Z < 2.5)) \\ &= 2(1 - 0.9938) \\ &= 0.0124.\end{aligned}$$

Επομένως, οι ενδείξεις που παρέχει το δείγμα δεν υποστηρίζουν την υπόθεση ότι το 0.95-ποσοστιαίο σημείο της κατανομής του ύψους είναι η τιμή 175.8.

Λύση με το MINITAB: Το παράδειγμα λύνεται με διωνυμικό έλεγχο. Στο γνωστό πλαίσιο διαλόγου που προκύπτει αν επιλέξουμε **Stat, Basic Statistics, 1 Proportion**, δηλώνουμε ότι έχουμε 100 δοκιμές, 90 επιτυχίες. Πιέζουμε **Options** και, στο πλαίσιο διαλόγου που προκύπτει, δηλώνουμε επίπεδο εμπιστοσύνης 95% (είναι το $1-\alpha$), **Test proportion** 0.95 (είναι το p_0), **Alternative not equal** και ζητάμε να χρησιμοποιηθεί η κανονική προσέγγιση της διωνυμικής κατανομής. Το αποτέλεσμα του ελέγχου είναι:

Test of $p = 0.95$ vs $p \text{ not} = 0.95$

Sample	X	N	Sample p	95.0 % CI	Z-Value	P-Value
1	90	100	0.900000	0.841201, 0.958799)	-2.29	0.022

Παρατηρούμε ότι η τιμή της ελεγχοσυνάρτησης Z είναι στην περιοχή απόρριψης (< -1.96) και ότι το κρίσιμο επίπεδο είναι 0.022, δηλαδή μικρότερο του 0.05.

Λύση με το SPSS: Για να κάνουμε διωνυμικό έλεγχο, καταχωρίζουμε το δείγμα μας σε μια nominal μεταβλητή (έστω **a**). Δηλώνουμε πρώτα 90 επιτυχίες με τον κωδικό 1 και κατόπιν 10 αποτυχίες με τον κωδικό 0. Στην συνέχεια, επιλέγουμε **Analyze, Nonparametric Tests, Binomial** και, μέσα από το γνωστό πλαίσιο διαλόγου, ζητάμε διωνυμικό έλεγχο για $p_0=0.95$ με κανονική προσέγγιση της διωνυμικής κατανομής. Τα αποτελέσματα του ελέγχου είναι:

Binomial Test

	Category	N	Observed Prop.	Test Prop.	Asymp. Sig. (1-tailed)	
A	Group 1	1	90	.90	.95	.019 ^{a,b}
	Group 2	0	10	.10		
	Total		100	1.00		

a Alternative hypothesis states that the proportion of cases in the first group < 95 .

b Based on Z Approximation.

Η τιμή του κρίσιμου επιπέδου που αντιστοιχεί σε μονόπλευρο έλεγχο είναι 0.019. Επομένως, για τον αμφίπλευρο έλεγχο του παραδείγματος, το κρίσιμο επίπεδο είναι το διπλάσιο της τιμής αυτής, δηλαδή 0.038.

Λύση με το SAS: Καταχωρίζουμε σε μια μεταβλητή τις 90 επιτυχίες που αντιστοιχούν σε ισάριθμα άτομα του δείγματος με ύψος κάτω του 175.8 και τις 10 αποτυχίες που αντιστοιχούν σε άτομα του δείγματος με ύψος άνω του 175.8. Στην συνέχεια, διεξάγουμε τον ασυμπτωτικό έλεγχο που παρέχει η συγκεκριμένη έκδοση του SAS, αντί του διωνυμικού. (Τα αποτελέσματα αναμένονται να είναι ικανοποιητικά λόγω του μεγάλου μεγέθους δείγματος. Έτσι, χρησιμοποιώντας τις εντολές:


```

data height;
x=1; f=90; output;
x=2; f=10; output;
run;
proc freq;
weight f;
tables x / testp=(0.95 0.05);
run;

```

το πακέτο μας δίνει τα εξής αποτελέσματα:

```

The SAS System

X      Frequency    Percent      Test      Cumulative  Cumulative
          Frequency    Percent      Percent    Frequency    Percent
-----
1          90         90.0         95.0         90         90.0
2          10         10.0          5.0         100        100.0

Chi-Square Test for Specified Proportions
-----
Statistic = 5.263      DF = 1      Prob = 0.022

```

Η τιμή της ελεγχοσυνάρτησης υπολογίσθηκε ίση με 5.263 και αντιστοιχεί σε παρατηρούμενο επίπεδο στατιστικής σημαντικότητας 0.022.

Διακριτή περίπτωση

Από τον ορισμό του p -ποσοστιαίου σημείου της κατανομής της μεταβλητής X , προκύπτει ότι οι προς έλεγχο υποθέσεις

$$H_0 : x_{p^*} = x^*$$

$$H_1 : x_{p^*} \neq x^*$$

είναι ισοδύναμες με τις υποθέσεις

$$H_0 : P(X \leq x^*) \geq p^* \text{ και } P(X < x^*) \leq p^*$$

$$H_1 : P(X \leq x^*) < p^* \text{ ή } P(X < x^*) > p^* .$$

Για τον έλεγχο των υποθέσεων αυτής της μορφής θεωρούμε δύο στατιστικές συναρτήσεις:

$$T_1 = \# \text{ παρατηρήσεων } X_i \text{ που είναι } \leq x^*$$

και

$$T_2 = \# \text{ παρατηρήσεων } X_i \text{ που είναι } < x^* .$$

Επομένως, η κρίσιμη περιοχή του ελέγχου, αντιστοιχεί σε τιμές της T_2 που είναι πολύ μεγάλες και σε τιμές της στατιστικής συνάρτησης T_1 που είναι πολύ μικρές. Αυτό, γιατί οι μεγάλες τιμές της T_2 είναι ενδεχομένως ενδεικτικές του ότι η πιθανότητα του ενδεχομένου $\{X < x^*\}$ υπερβαίνει την τιμή p^* . Αντίστοιχα, οι μικρές τιμές της T_1 είναι ενδεχομένως ενδεικτικές του ότι η πιθανότητα του ενδεχομένου $\{X \leq x^*\}$ είναι μικρότερη από την τιμή p^* . Δεδομένου ότι, όπως είναι προφανές, οι στατιστικές συναρτήσεις T_1 και T_2 ακολουθούν κάτω από την μηδενική υπόθεση την διωνυμική κατανομή με παραμέτρους n, p^* , ο κανόνας απόρριψης έχει την εξής μορφή:

Η υπόθεση H_0 απορρίπτεται σε επίπεδο σημαντικότητας $\alpha_1 + \alpha_2 \cong \alpha$ αν $T_1 \leq t_1$ ή αν $T_2 > t_2$, όπου οι τιμές t_1 και t_2 προσδιορίζονται από τον πίνακα της διωνυμικής κατανομής (πίνακας 1 του παραρτήματος) έτσι ώστε

$$P(T_1 \leq t_1 \mid n, p = p^*) = \alpha_1 \cong \alpha/2$$

και

$$P(T_2 > t_2 \mid n, p = p^*) = \alpha_2 \cong \alpha/2 .$$

Παράδειγμα 2.5.2: Ένα δείγμα 20 τελειοφοίτων μαθητών Γυμνασίου έδειξε τα εξής αποτελέσματα σ' ένα διαγώνισμα με άριστα το 200.

142 134 98 119 131
 103 154 122 93 137
 86 119 161 144 158
 165 81 117 128 103

Να ελεγχθεί σε επίπεδο σημαντικότητας $\alpha \cong 0.05$ η υπόθεση ότι η διάμεσος της κατανομής των βαθμών είναι η τιμή 103.

Λύση: Στην πραγματικότητα, θέλουμε να ελέγξουμε τις υποθέσεις:

$$H_1 : x_{0.5} = 103$$

$$H_1 : x_{0.5} \neq 103$$

ή, ισοδύναμα,

$$H_0 : P(X \leq 103) \geq 0.5 \text{ και } P(X < 103) \leq 0.5$$

$$H_1 : P(X \leq 103) < 0.5 \text{ ή } P(X < 103) > 0.5.$$

Τα κρίσιμα σημεία t_1 και t_2 θα καθορισθούν έτσι ώστε

$$P(T_1 \leq t_1 \mid n = 20, p = 0.5) = \alpha_1 \cong 2.5\%$$

και

$$P(T_2 > t_2 \mid n = 20, p = 0.5) = \alpha_2 \cong 2.5\%.$$

Από τον πίνακα της διωνυμικής κατανομής (πίνακας 1 του παραρτήματος) έχουμε $t_1 = 5$ αν $\alpha_1 = 0.0207$ και $t_2 = 14$ αν $\alpha_2 = 0.0207$.

Από τα δεδομένα, προκύπτει ότι οι παρατηρηθείσες τιμές των στατιστικών συναρτήσεων T_1 και T_2 είναι $t_1 = 6$ και $t_2 = 4$, αντίστοιχα. Επομένως, σε επίπεδο σημαντικότητας 0.0414, δεν απορρίπτεται η μηδενική υπόθεση H_0 .

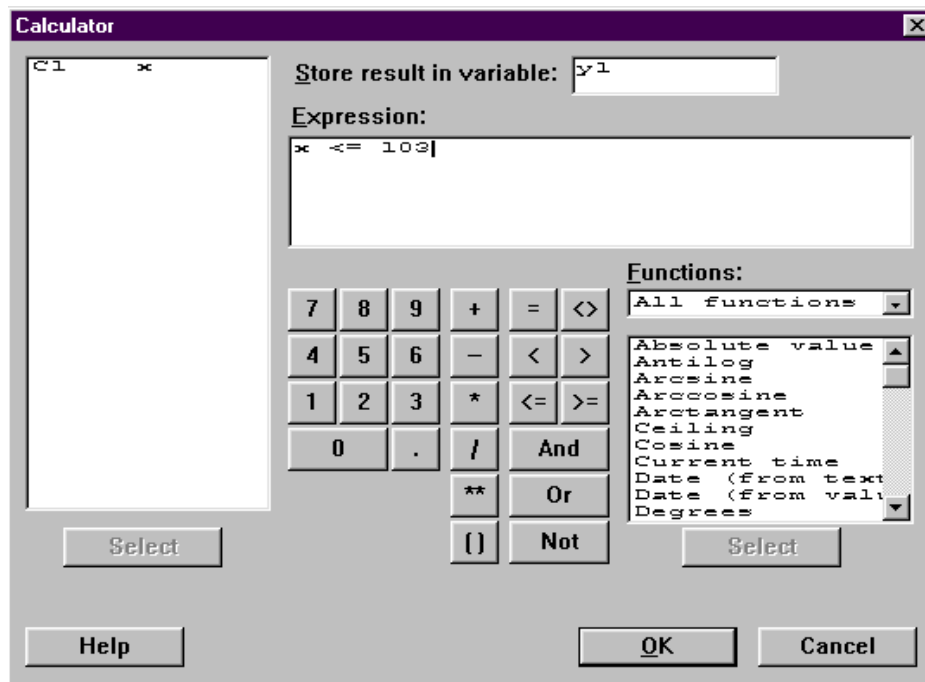
Επειδή έχουμε δύο ελεγχosυναρτήσεις, ο αμφίπλευρος έλεγχος ισοδυναμεί με τον συνδυασμό των μονόπλευρων ελέγχων των περιπτώσεων Β και Γ. Επομένως, το κρίσιμο επίπεδο $\hat{\alpha}$ του ελέγχου ορίζεται από την σχέση: $\hat{\alpha} = 2 \min \{\hat{\alpha}_1, \hat{\alpha}_2\}$, όπου

$$\begin{aligned}\hat{\alpha}_1 &= P(\eta T_1 \text{ έχει μία τιμή ίση ή πιο ακραία από την παρατηρηθείσα } \tau_1 | H_0) \\ &= P(T_1 \leq 6 | n = 20, p = 0.5) = 0.0577\end{aligned}$$

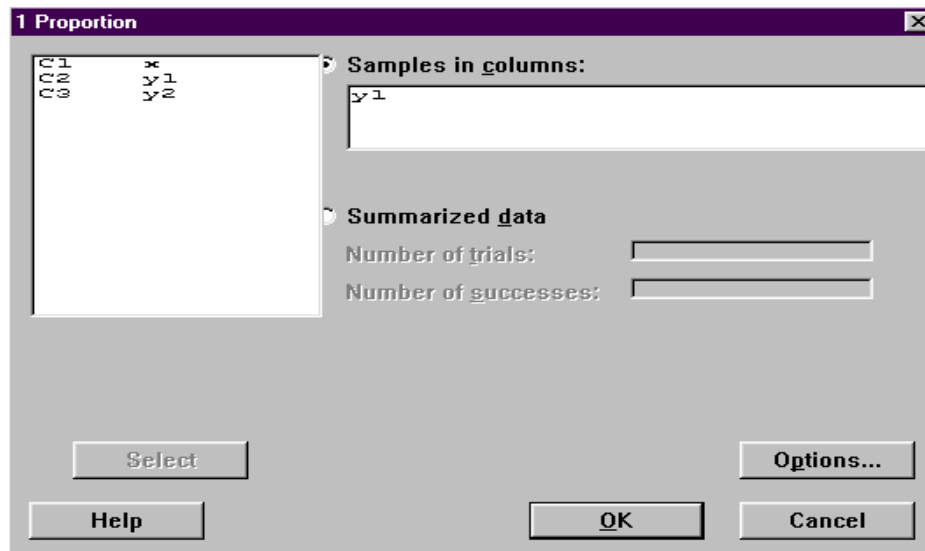
$$\begin{aligned}\hat{\alpha}_2 &= P(\eta T_2 \text{ έχει μία τιμή ίση ή πιο ακραία από την παρατηρηθείσα } \tau_2 | H_0) . \\ &= P(T_2 \geq 4 | n = 20, p = 0.5) = 1 - 0.0059 = 0.9941 .\end{aligned}$$

Άρα $\hat{\alpha} = 2(0.0577) = 0.1154$.

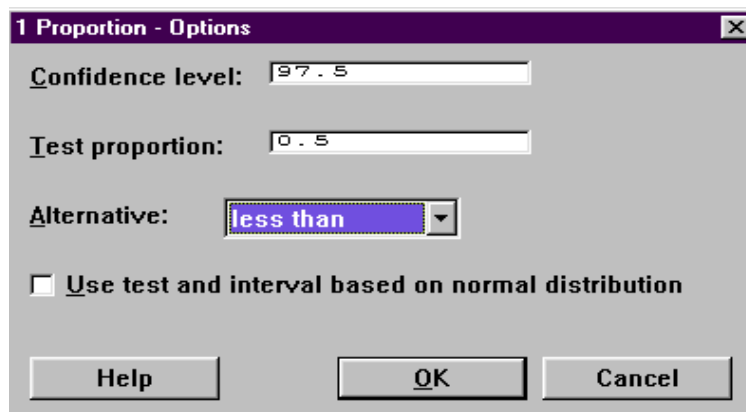
Λύση με το MINITAB: Για να λύσουμε το παράδειγμα με το MINITAB θα διεξαγάγουμε χωριστό διωνυμικό έλεγχο για το κάθε ζεύγος υποθέσεων. Για τον λόγο αυτό, καταχωρίζουμε σε μία μεταβλητή με το όνομα **x** τις τιμές του δείγματος και στην συνέχεια, δημιουργούμε δύο ψευδομεταβλητές **y1** και **y2**. Η **y1** παίρνει την τιμή 1 αν η **x** είναι μικρότερη ή ίση του 103 ενώ η **y2** παίρνει την τιμή 1 αν η **x** είναι μικρότερη του 103. Οι δύο αυτές μεταβλητές παίρνουν ταυτόχρονα την τιμή 0 ή την τιμή 1 οποτεδήποτε η **x** έχει τιμή $\neq 103$. Για την δημιουργία μιας ψευδομεταβλητής εργαζόμαστε ως εξής: Επιλέγουμε **Calc, Calculator** στο βασικό μενού και στο πλαίσιο διαλόγου που εμφανίζεται προσδιορίζουμε τις τιμές και το πεδίο ορισμού της ψευδομεταβλητής.



Συγκεκριμένα, στο πεδίο **Store result in variable**, δηλώνουμε τη μεταβλητή, στην οποία θα καταχωρισθεί η ψευδομεταβλητή (π.χ. **y1**), ενώ, στο πεδίο **Expression**, δηλώνουμε την λογική συνθήκη που πρέπει να ισχύει ώστε η ψευδομεταβλητή να πάρει την τιμή 1. Ανάλογη διαδικασία ακολουθούμε για την δημιουργία της ψευδομεταβλητής **y2**. Στην συνέχεια, διεξάγουμε τους δύο μονόπλευρους διωνυμικούς ελέγχους. Πρέπει να προσέξουμε να δηλώσουμε επίπεδο σημαντικότητας ίσο με 0.025, αφού συνδυασμός των δύο ελέγχων θα δώσει τον έλεγχο που θέλουμε με $\alpha=0.05$. Για το πρώτο ζεύγος υποθέσεων (με $H_1 : P(X \leq 103) < 0.5$), επιλέγουμε από το βασικό μενού **Stat, Basic Statistics, 1 Proportion** και οδηγούμεθα στο εξής πλαίσιο διαλόγου:



Στο πεδίο **Samples in columns**, δηλώνουμε την ψευδομεταβλητή που αντιστοιχεί στο συγκεκριμένο ζεύγος υποθέσεων. Πιέζοντας **Options**, εμφανίζεται το πλαίσιο:



Στο πλαίσιο αυτό, δηλώνουμε το επίπεδο σημαντικότητας, την τιμή της παραμέτρου p και την κατεύθυνση της εναλλακτικής υπόθεσης και πιέζουμε **OK**. Πιέζοντας πάλι **OK** στο πλαίσιο διαλόγου που προκύπτει, οδηγούμεθα στα εξής αποτελέσματα:

Test of $p = 0.5$ vs $p < 0.5$
Success = 1

Exact Variable	X	N	Sample p	97.5 % CI	P-Value
y1	6	20	0.300000	(0.102255; 0.573647)	0.058

Για το δεύτερο ζεύγος υποθέσεων (με $H_1: P(X < 103) > 0.5$) τα αποτελέσματα του ελέγχου είναι:

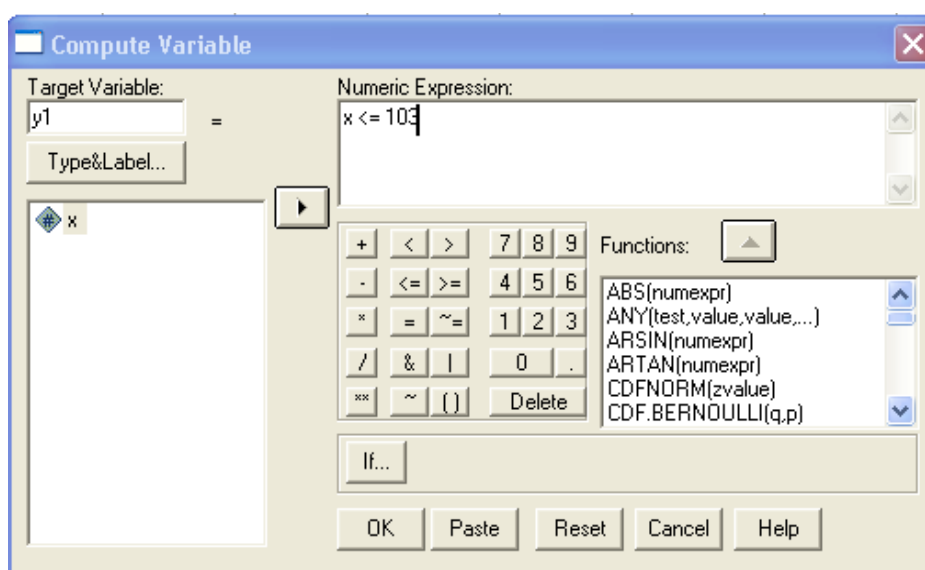
Test of $p = 0.5$ vs $p > 0.5$
Success = 1

Exact Variable	X	N	Sample p	97.5 % CI	P-Value
y2	4	20	0.200000	(0.046559; 0.468610)	0.999

Βλέπουμε ότι το παρατηρούμενο επίπεδο σημαντικότητας του συνολικού ελέγχου είναι $\hat{\alpha} = 2 \min \{\hat{\alpha}_1, \hat{\alpha}_2\} = 2 \cdot 0.058 = 0.116$. Κατά συνέπεια, σε επίπεδο σημαντικότητας $\alpha = 0.05$ η μηδενική υπόθεση πως η διάμεσος της κατανομής των βαθμών είναι ίση με 103 δεν απορρίπτεται.

Λύση με το SPSS: Για να λύσουμε το παράδειγμα με το SPSS θα διεξαγάγουμε χωριστό διωνυμικό έλεγχο για το κάθε ζεύγος υποθέσεων. Για τον λόγο αυτό, καταχωρίζουμε σε μία μεταβλητή με το όνομα x τις τιμές του δείγματος και, στην συνέχεια, δημιουργούμε δύο ψευδομεταβλητές $y1$ και $y2$. Η $y1$ παίρνει την τιμή 1 αν η x είναι μικρότερη ή ίση του 103 ενώ η $y2$ παίρνει την τιμή 1 αν η x είναι

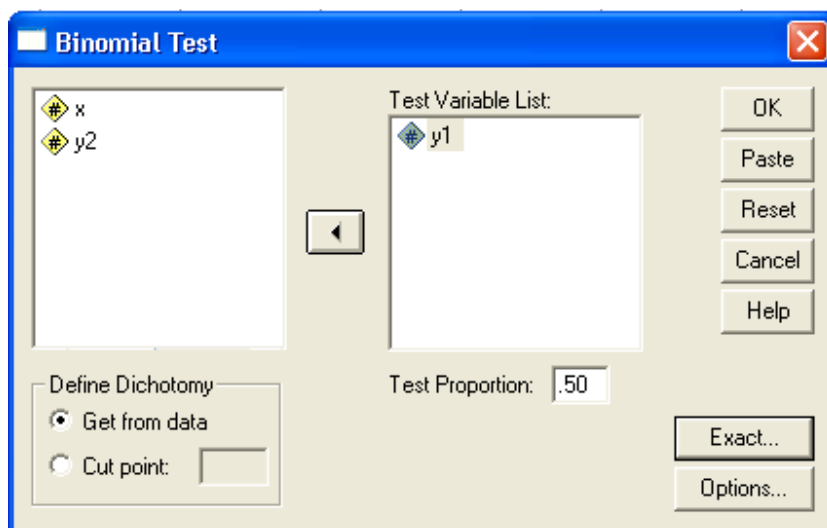
μικρότερη του 103. Οι δύο αυτές μεταβλητές παίρνουν ταυτόχρονα την τιμή 0 ή την τιμή 1 οποτεδήποτε η x έχει τιμή $\neq 103$. Για την δημιουργία μιας ψευδομεταβλητής εργαζόμαστε ως εξής: Επιλέγουμε **Transform**, **Compute** στο βασικό μενού και στο πλαίσιο διαλόγου που εμφανίζεται προσδιορίζουμε τις τιμές και το πεδίο ορισμού της ψευδομεταβλητής.



Συγκεκριμένα, στο πεδίο **Target variable**, δηλώνουμε την μεταβλητή, στην οποία θα καταχωρισθεί η ψευδομεταβλητή (π.χ. **y1**), ενώ στο πεδίο **Numeric Expression**, δηλώνουμε την λογική συνθήκη που πρέπει να ισχύει ώστε η ψευδομεταβλητή να πάρει την τιμή 1. Ανάλογη διαδικασία ακολουθούμε για την δημιουργία της ψευδομεταβλητής **y2**. Στην συνέχεια, διεξάγουμε τους δύο μονόπλευρους διωνυμικούς ελέγχους.

Υπενθυμίζεται ότι, για να θεωρηθεί η τιμή 1 ως επιτυχία, πρέπει στην πρώτη γραμμή των μεταβλητών **y1** και **y2** να έχουμε μονάδα. Αυτό το πετυχαίνουμε αλλάζοντας την θέση κάποιων παρατηρήσεων. Πρέπει,

επίσης, να προσέξουμε να δηλώσουμε επίπεδο σημαντικότητας ίσο με 0.025, αφού συνδυασμός των δύο ελέγχων θα δώσει τον έλεγχο που θέλουμε με $\alpha=0.05$. Για το πρώτο ζεύγος υποθέσεων (με $H_1 : P(X \leq 103) < 0.5$), επιλέγουμε από το βασικό μενού **Analyze**, **Nonparametric Tests**, **Binomial** και οδηγούμεθα στο εξής πλαίσιο διαλόγου:



Στο πεδίο **Test Variable List**, δηλώνουμε την ψευδομεταβλητή που αντιστοιχεί στο συγκεκριμένο ζεύγος υποθέσεων. Στο πεδίο **Define Dichotomy**, επιλέγουμε **Get from data**. Στο πεδίο **Test Proportion**, επιλέγουμε 0.5. Επίσης, ζητάμε από το πρόγραμμα να χρησιμοποιήσει την ακριβή μορφή της κατανομής της ελεγχουσυνάρτησης (διωνυμική) κατά τα γνωστά. Πιέζοντας **OK**, οδηγούμεθα στα εξής αποτελέσματα:

Binomial Test

		Category	N	Observed Prop.	Test Prop.	Exact Sig. (2-tailed)
Y1	Επιτυχία	1	6	.30	.50	.115
	Αποτυχία	0	14	.70		
	Total		20	1.00		

Το SPSS διεξάγει μόνο αμφίπλευρο έλεγχο. Έτσι, το κρίσιμο επίπεδο του πρώτου ελέγχου θα προκύψει με διαίρεση της τιμής που το SPSS δίνει (Exact Sig. (2-tailed)) δια 2. Για το δεύτερο ζεύγος υποθέσεων (με $H_1: P(X < 103) > 0.5$), τα αποτελέσματα είναι:

Binomial Test

		Category	N	Observed Prop.	Test Prop.	Exact Sig. (2-tailed)
Y2	Επιτυχία	1	4	.20	.50	.012
	Αποτυχία	0	16	.80		
	Total		20	1.00		

Εδώ, το πρόγραμμα δίνει αμφίπλευρο κρίσιμο επίπεδο. Αυτό έχει υπολογισθεί με βάση το ότι το ήμισυ της τιμής του αναφέρεται στην αριστερή ουρά της κατανομής της ελεγχουσυνάρτησης, επειδή το παρατηρούμενο ποσοστό επιτυχιών είναι μικρότερο από αυτό που δηλώνει η μηδενική υπόθεση. (Αυτό γίνεται αυτόματα από το SPSS). Άρα για τον μονόπλευρο έλεγχο του δεύτερου ζεύγους υποθέσεων, το κρίσιμο επίπεδο είναι $1 - 0.012/2 + \Pr(T=4) \approx 1$ (συνυπολογίζοντας την πιθανότητα η T να είναι ίση με την παρατηρούμενη τιμή της). Έτσι το κρίσιμο επίπεδο του συνολικού ελέγχου είναι $\hat{\alpha} = 2 \min \{\hat{\alpha}_1, \hat{\alpha}_2\} = 2 \cdot 0.058 = 0.116$. Βλέπουμε συνεπώς ότι η μηδενική υπόθεση πως η διάμεσος της κατανομής των βαθμών είναι ίση με 103 δεν απορρίπτεται σε $\alpha=0.05$.

Λύση με το SAS: Η χρησιμοποιημένη έκδοση του SAS δεν παρέχει διωνυμικό έλεγχο και η ασυμπτωτική λύση δεν προσφέρεται γιατί δεν αναμένεται να δώσει ικανοποιητικά αποτελέσματα (λόγω του μικρού αριθμού δεδομένων).

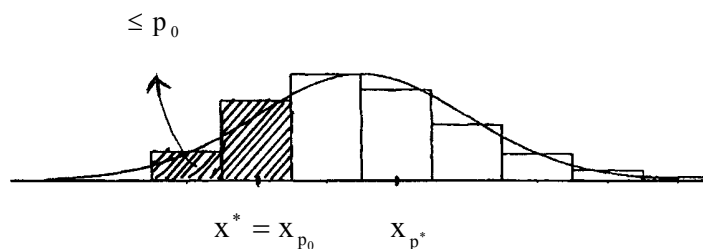
B. (Μονόπλευρος Έλεγχος)

Για τον έλεγχο των υποθέσεων

$$H_0 : x_{p^*} \geq x^*$$

$$H_1 : x_{p^*} < x^*$$

της περίπτωσης αυτής, θεωρούμε μία τιμή p_0 , τέτοια ώστε $x^* = x_{p_0}$.



Σχήμα 2.5.1

Τότε, από τον ορισμό του p_0 -ποσοστιαίου σημείου μιας κατανομής, x_{p_0} , ισχύει ότι

$$x_{p^*} \geq x^* \equiv x_{p_0} \Leftrightarrow p^* \geq p_0 \geq P(X < x_{p_0}) \equiv P(X < x^*) \Leftrightarrow P(X < x^*) \leq p^* .$$

Άρα, οι προς έλεγχο υποθέσεις είναι ισοδύναμες με τις υποθέσεις

$$H_0 : P(X < x^*) \leq p^*$$

$$H_1 : P(X < x^*) > p^* .$$

Η προφανής επιλογή στατιστικής ελεγχοσυνάρτησης για τον έλεγχο των παραπάνω υποθέσεων είναι ο αριθμός των παρατηρήσεων του δείγματος που δεν υπερβαίνουν την τιμή x^* . Δηλαδή,

$$T_2 = \# \text{ παρατηρήσεων } X_i \text{ που είναι } < x^* .$$

Επομένως, ο κανόνας απόρριψης είναι ο εξής:

Η υπόθεση H_0 απορρίπτεται σε επίπεδο σημαντικότητας περίπου ίσο με α αν $T_2 > t_2$, όπου η τιμή t_2 ορίζεται από την σχέση $P(T_2 > t_2 \mid n, p = p^*) \cong \alpha$.

Παράδειγμα 2.5.3: Αν στο προηγούμενο παράδειγμα με τις βαθμολογίες των μαθητών είχαμε να ελέγξουμε τις υποθέσεις

$$H_0 : x_{0.5} \geq 103$$

$$H_1 : x_{0.5} < 103 ,$$

θα θεωρούσαμε ισοδύναμα τις υποθέσεις

$$H_0 : P(X < 103) \leq 0.5$$

$$H_1 : P(X < 103) > 0.5 .$$

Η κρίσιμη περιοχή του ελέγχου ορίζεται από την ανισότητα $T_2 > t_2$, όπου η τιμή t_2 ορίζεται έτσι ώστε $P(T_2 > t_2 \mid n = 20, p = 0.5) = \alpha_2 \cong \alpha = 0.05$. Από τον πίνακα της διωνυμικής κατανομής (πίνακας 1 του παραρτήματος) έχουμε ότι $t_2 = 13$

αν $\alpha_2 = 0.0577$. Δεδομένου ότι η παρατηρηθείσα τιμή $\tau_2=4$ της στατιστικής συνάρτησης T_2 δεν υπερβαίνει την κρίσιμη τιμή, η μηδενική υπόθεση δεν απορρίπτεται σε επίπεδο σημαντικότητας 0.0577.

Το κρίσιμο επίπεδο του ελέγχου είναι $\hat{\alpha} = P(T_2 \geq 4 | n = 20, p = 0.5) = 1 - P(T_2 \leq 3 | n = 20, p = 0.5) = 1 - 0.0013 = 0.9987$

Λύση με το MINITAB: Θα διεξαγάγουμε διωνυμικό έλεγχο, δηλώνοντας 20 δοκιμές, 4 «επιτυχίες» και προσδιορίζοντας ως εναλλακτική υπόθεση ότι το ποσοστό επιτυχίας είναι μεγαλύτερο του 0.5. Τα παραπάνω μπορούν να γίνουν από την επιλογή **Options** του πλαισίου διαλόγου **Stat, Basic Statistics, 1 proportion**. Τα αποτελέσματα συνοψίζονται στον πίνακα που ακολουθεί:

Test of $p = 0.5$ vs $p > 0.5$

Sample	X	N	Sample p	95,0 % CI	P-Value
1	4	20	0.200000	(0.057334; 0.436614)	0.999

Λύση με το SPSS: Θα διεξαγάγουμε διωνυμικό έλεγχο πάνω στις παρατηρήσεις της y_2 που δημιουργήσαμε στο προηγούμενο παράδειγμα με τον ίδιο γνωστό τρόπο. Στο παρόν παράδειγμα, τα αποτελέσματα του ελέγχου είναι τα εξής:

	Category	N	Observed Prop.	Test Prop.	Exact Sig. (2-tailed)
Y2	Επιτυχία	4	.20	.50	.012
	Αποτυχία	16	.80		
	Total	20	1.00		

Εδώ, το πρόγραμμα δίνει την τιμή του αμφίπλευρου κρίσιμου επιπέδου. Αυτό έχει υπολογισθεί με βάση το ότι το ήμισυ της τιμής του αναφέρεται στην αριστερή ουρά της κατανομής της ελεγχοσυνάρτησης, επειδή το παρατηρούμενο ποσοστό επιτυχιών είναι μικρότερο από αυτό που δηλώνει η μηδενική υπόθεση. (Αυτό γίνεται αυτόματα από το SPSS). Επομένως, για τον μονόπλευρο έλεγχο που διεξαγάγαμε, το κρίσιμο επίπεδο είναι $1 - 0.012/2 + \Pr(T=4) \approx 1$ (συνυπολογίζοντας την πιθανότητα η T να είναι ίση με την παρατηρούμενη τιμή της). Άρα, σε επίπεδο σημαντικότητας $\alpha=0.05$ η μηδενική υπόθεση ότι η διάμεσος της κατανομής των βαθμών είναι μεγαλύτερη ή ίση της τιμής 103 αποτελεί μια εύλογη υπόθεση.

Γ. (Μονόπλευρος Έλεγχος)

Για τον έλεγχο των υποθέσεων

$$H_0 : x_{p^*} \leq x^*$$

$$H_1 : x_{p^*} > x^*,$$

θεωρούμε πάλι μία τιμή p_0 τέτοια ώστε $x^* = x_{p_0}$ και παρατηρούμε ότι, από τον ορισμό του p_0 -ποσοστιαίου σημείου, x_{p_0} , ισχύει ότι

$$x_{p^*} \leq x^* \equiv x_{p_0} \Leftrightarrow p^* \leq p_0 \leq P(X \leq x_{p_0}) \equiv P(X \leq x^*).$$

Άρα οι προς έλεγχο υποθέσεις είναι ισοδύναμες με τις υποθέσεις

$$H_0 : P(X \leq x^*) \geq p^*$$

$$H_1 : P(X \leq x^*) < p^*.$$

Η προφανής επιλογή ελεγχοσυνάρτησης για τον έλεγχο των παραπάνω υποθέσεων είναι η $T_1 = \#$ παρατηρήσεων X_i που είναι $\leq x^*$ και η κρίσιμη περιοχή του ελέγχου ορίζεται από την ανισότητα $T_1 \leq t_1$, όπου η τιμή t_1 ορίζεται από την σχέση $P(T_1 \leq t_1 \mid n, p = p^*) = \alpha_1 \equiv \alpha$.

Παράδειγμα 2.5.4: Αν στο προηγούμενο παράδειγμα η εναλλακτική υπόθεση ήταν $x_{0.5} > 103$, οι προς έλεγχο υποθέσεις

$$\begin{aligned}H_0 &: x_{0.5} \leq 103 \\H_1 &: x_{0.5} > 103\end{aligned}$$

θα ήταν ισοδύναμες με τις υποθέσεις

$$H_0 : P(X \leq 103) \geq 0.5$$

$$H_1 : P(X \leq 103) < 0.5.$$

Ο κανόνας απόφασης έχει την μορφή:

Η υπόθεση H_0 απορρίπτεται σε επίπεδο σημαντικότητας περίπου ίσο με α , αν $T_1 \leq t_1$, όπου η τιμή t_1 ορίζεται από την σχέση

$$P(T_1 \leq t_1 \mid n = 20, p = 0.5) = \alpha_1 \cong \alpha.$$

Από τον πίνακα 1 του παραρτήματος προκύπτει ότι $t_1 = 6$, αν $\alpha_1 = 0.0577$. Η παρατηρηθείσα τιμή της στατιστικής συνάρτησης T_1 είναι $t_1 = 6$. Επομένως, η μηδενική υπόθεση H_0 απορρίπτεται σε επίπεδο σημαντικότητας 0.0577.

Το κρίσιμο επίπεδο του ελέγχου είναι $\hat{\alpha} = P(T_1 \leq 6 \mid n = 20, p = 0.5) = 0.0577$.

Λύση με το MINITAB: Θα διεξαγάγουμε διωνυμικό έλεγχο, δηλώνοντας 20 δοκιμές, 6 «επιτυχίες» και προσδιορίζοντας ως εναλλακτική υπόθεση ότι το ποσοστό επιτυχίας είναι μικρότερο του 0.5. Τα παραπάνω μπορούν να γίνουν πιέζοντας το πλήκτρο **Options** του πλαισίου διαλόγου που προκύπτει με την επιλογή **Stat, Basic Statistics, 1 proportion**. Τα αποτελέσματα συνοψίζονται στον πίνακα που ακολουθεί:

Test of $p = 0.5$ vs $p < 0.5$

Exact					
Sample	X	N	Sample p	95.0 % CI	P-Value
1	6	20	0.300000	(0.118932; 0.542789)	0.058

Λύση με το SPSS: Θα διεξαγάγουμε διωνυμικό έλεγχο πάνω στις παρατηρήσεις της y_1 που δημιουργήσαμε στο παράδειγμα 2.5.2 κατά τον ήδη γνωστό τρόπο. Στο παρόν παράδειγμα, τα αποτελέσματα του ελέγχου είναι τα εξής:

		Category	N	Observed Prop.	Test Prop.	Exact Sig. (2-tailed)
Y1	Επιτυχία	1	6	.30	.50	.115
	Αποτυχία	0	14	.70		
	Total		20	1.00		

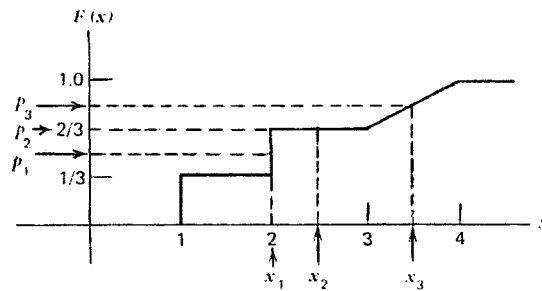
Εδώ, το πρόγραμμα δίνει την τιμή του αμφίπλευρου κρίσιμου επιπέδου. Αυτό έχει υπολογισθεί με βάση το ότι το ήμισυ της τιμής του αναφέρεται στην αριστερή ουρά της κατανομής της ελεγκοσυνάρτησης επειδή το παρατηρούμενο ποσοστό επιτυχιών είναι μικρότερο από αυτό που δηλώνει η μηδενική υπόθεση. (Αυτό γίνεται αυτόματα από το SPSS). Επομένως, για τον μονόπλευρο έλεγχο που διεξαγάγαμε, το κρίσιμο επίπεδο είναι $0.115/2=0.057$.

2.5.1 ΕΚΤΙΜΗΣΗ ΠΟΣΟΣΤΙΑΙΩΝ ΣΗΜΕΙΩΝ ΜΙΑΣ ΚΑΤΑΝΟΜΗΣ

Η μέθοδος κατασκευής διαστήματος εμπιστοσύνης για την πιθανότητα p που περιγράφεται στην προηγούμενη ενότητα μπορεί να χρησιμοποιηθεί για την κατασκευή διαστημάτων εμπιστοσύνης για συγκεκριμένες τιμές της συνάρτησης κατανομής $F_X(x)$ μιας μεταβλητής X . Αναφερόμενοι στην γραφική παράσταση μιας συνάρτησης κατανομής, όπως για παράδειγμα του σχήματος 2.5.2, μπορούμε να πούμε ότι, δοθείσης μίας τιμής x_0 , μπορούμε να κατασκευάσουμε ένα «κατακόρυφο» διάστημα εμπιστοσύνης για την άγνωστη πιθανότητα $F_X(x_0)$. Ας υποθέσουμε τώρα ότι ενδιαφερόμαστε για το «αντίστροφο» πρόβλημα, δηλαδή ότι, δοθείσης μίας τιμής της πιθανότητας $F_X(x_0)$, έστω p_0 , ενδιαφερόμαστε να κατασκευάσουμε ένα «οριζόντιο» διάστημα εμπιστοσύνης για την άγνωστη τιμή x_0 . Επειδή κάθε τιμή x της τυχαίας μεταβλητής X είναι τιμή ενός ποσοστιαίου σημείου, x_p της κατανομής της μεταβλητής X (είναι, δηλαδή, ένα p -ποσοστιαίο σημείο για κάποια τιμή του p), η τιμή x_0 μπορεί να θεωρηθεί ως η τιμή του p_0 -ποσοστιαίου σημείου της μεταβλητής X . Τότε, το πρόβλημα που μας ενδιαφέρει μπορεί ισοδύναμα να διατυπωθεί ως εξής:

Δοθείσης της τιμής p_0 της πιθανότητας $F_X(x_0)$, να κατασκευασθεί ένα «οριζόντιο» διάστημα εμπιστοσύνης για το άγνωστο p_0 -ποσοστιαίο σημείο x_{p_0} . Αυτού του είδους το διάστημα εμπιστοσύνης, κατασκευάζεται όταν επιθυμούμε να εκτιμήσουμε ένα συγκεκριμένο ποσοστιαίο σημείο, όπως είναι η διάμεσος, το άνω ή

κάτω τεταρτημόριο ή, γενικότερα, ένα οποιοδήποτε p -ποσοστιαίο σημείο, όπου p είναι μία συγκεκριμένη σταθερά, $0 \leq p \leq 1$.



Σχήμα 2.5.2

Γράφημα μιας τυχούσας συνάρτησης κατανομής

Εστω τυχαίο δείγμα X_1, X_2, \dots, X_n παρατηρήσεων πάνω στην μεταβλητή X . Η κλίμακα μέτρησης των $X_i, i = 1, 2, \dots, n$ είναι τουλάχιστον κλίμακα διάταξης. Το δείγμα που προκύπτει αν οι παρατηρήσεις διαταχθούν κατά αύξουσα σειρά μεγέθους είναι $X^{(1)} \leq X^{(2)} \leq \dots \leq X^{(n)}$. Η κατασκευή ενός διαστήματος εμπιστοσύνης για την τιμή x_p με συντελεστή εμπιστοσύνης περίπου ίσο με $1-\alpha$ ισοδυναμεί με τον προσδιορισμό δύο ακεραίων αριθμών r και s ($r \leq s$) και, ακολούθως, του διαστήματος $[X^{(r)}, X^{(s)}]$ με την ιδιότητα

$$P(X^{(r)} \leq x_p \leq X^{(s)}) = 1-\alpha, \text{ αν } X \text{ συνεχής}$$

$$P(X^{(r)} \leq x_p < X^{(s)}) \geq 1-\alpha, \text{ αν } X \text{ διακριτή.}$$

Συνεχής Περίπτωση

Υπενθυμίζεται ότι το p -ποσοστιαίο σημείο x_p μίας συνεχούς κατανομής ορίζεται ως; η τιμή x_p με την ιδιότητα $P(X < x_p) = p$.

Όπως είναι γνωστό,

$$P(X^{(r)} \leq x_p \leq X^{(s)}) = P(X^{(s)} \geq x_p) - P(X^{(r)} > x_p).$$

Επομένως, θέτοντας

$$P(X^{(s)} \geq x_p) = 1 - \alpha_2 \text{ και } P(X^{(r)} > x_p) = \alpha_1,$$

οι ακέραιες τιμές r και s που θα προσδιορισθούν, έτσι ώστε:

$$P(X^{(r)} > x_p) = \alpha_1 \cong \alpha/2$$

και

$$P(X^{(s)} \geq x_p) = 1 - \alpha_2 \cong 1 - \alpha/2,$$

οδηγούν σε διάστημα της μορφής $[X^{(r)}, X^{(s)}]$ με την ιδιότητα

$$P(X^{(r)} \leq x_p \leq X^{(s)}) = 1 - \alpha_1 - \alpha_2 \cong 1 - \alpha.$$

Το διάστημα, δηλαδή, που προσδιορίζεται με την παραπάνω διαδικασία, αποτελεί ένα διάστημα εμπιστοσύνης για το ποσοστιαίο σημείο x_p με συντελεστή εμπιστοσύνης περίπου ίσο με $100(1-\alpha)\%$.

Προσδιορισμός των τιμών r και s : Για τον προσδιορισμό της τιμής r , απαιτείται ο καθορισμός της πιθανότητας $P(X^{(r)} > x_p)$.

Παρατηρούμε ότι, για $r = 1$,

$$\begin{aligned} P(X^{(1)} > x_p) &= P(\min\{X_1, X_2, \dots, X_n\} > x_p) \\ &= P(X_1 > x_p, X_2 > x_p, \dots, X_n > x_p). \end{aligned}$$

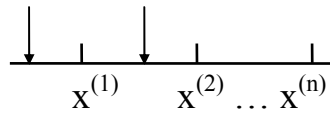
Επειδή οι παρατηρήσεις X_i , $i = 1, 2, \dots, n$ είναι ανεξάρτητες και ισόνομες τυχαίες μεταβλητές, η παραπάνω ισότητα παίρνει την μορφή

$$P(X^{(1)} > x_p) = \prod_{i=1}^n P(X_i > x_p),$$

η οποία, από τον ορισμό του p -ποσοστιαίου σημείου, γράφεται

$$P(X^{(1)} > x_p) = (1 - p)^n.$$

Για $r = 2$, το ενδεχόμενο $\{X^{(2)} > x_p\}$ ισοδυναμεί με την ένωση των ενδεχομένων $\{\text{ακριβώς } n-1 \text{ παρατηρήσεις υπερβαίνουν την τιμή } x_p\}$ και $\{\text{ακριβώς } n \text{ παρατηρήσεις υπερβαίνουν την τιμή } x_p\}$. Αυτό είναι άμεση συνέπεια του γεγονότος ότι, όπως προκύπτει από το σχήμα 2.5.3, δύο είναι οι δυνατές θέσεις της τιμής x_p σε σχέση με τις παρατηρήσεις του δείγματος: Αριστερά της τιμής $X^{(1)}$ ή μεταξύ των τιμών $X^{(1)}$ και $X^{(2)}$.



Σχήμα 2.5.3

Κατά συνέπεια,

$$\begin{aligned} P(X^{(2)} > x_p) &= P(\{\text{όλα τα } X_i > x_p\} \cup \{n-1 \text{ από τα } X_i > x_p\}) \\ &= P(\text{τουλάχιστον } n-1 \text{ παρατηρήσεις } > x_p) \\ &= P(\text{το πολύ } 1 \text{ παρατήρηση } < x_p). \end{aligned}$$

Από τον ορισμό του p -ποσοστιαίου σημείου, ισχύει ότι $P(X_i < x_p) = p$, $i = 1, 2, \dots, n$. Επομένως,

$$P(X^{(2)} > x_p) = \sum_{x=0}^1 \binom{n}{x} p^x (1-p)^{n-x}.$$

Συνεχίζοντας με τον ίδιο τρόπο, αποδεικνύεται ότι

$$P(X^{(r)} > x_p) = \sum_{x=0}^{r-1} \binom{n}{x} p^x (1-p)^{n-x}.$$

Το δεξί μέλος της παραπάνω ισότητας ταυτίζεται με την αθροιστική συνάρτηση κατανομής μίας διωνυμικής μεταβλητής, έστω X_b , με παραμέτρους n, p . Άρα,

$$P(X^{(r)} > x_p) = P(X_b \leq r-1 \mid n, p).$$

Επειδή η μεταβλητή X είναι συνεχής,

$$P(X^{(s)} \geq x_p) = P(X^{(s)} > x_p).$$

Επομένως, με όμοιο τρόπο, μπορεί να αποδειχθεί ότι

$$P(X^{(s)} \geq x_p) = P(X_b \leq s-1 | n, p).$$

Άρα, οι τιμές r και s μπορούν να προσδιορισθούν με βάση τον πίνακα της διωνυμικής κατανομής (πίνακα 1 του παραρτήματος) για τις δοθείσες τιμές των α , n και p έτσι ώστε:

$$P(X_b \leq r-1 | n, p) = \alpha_1 \cong \alpha/2,$$

$$P(X_b \leq s-1 | n, p) = 1 - \alpha_2 \cong 1 - \alpha/2.$$

Παρατήρηση: Η κατασκευή ενός διαστήματος εμπιστοσύνης της μορφής $[X^{(r)}, X^{(s)}]$ με συντελεστή εμπιστοσύνης περίπου ίσο με $1 - \alpha_1 - \alpha_2 \cong 1 - \alpha$ ισοδυναμεί με τον έλεγχο των υποθέσεων

$$H_0 : x_p = x^*$$

$$H_1 : x_p \neq x^*$$

σε επίπεδο σημαντικότητας $\alpha_1 + \alpha_2 \cong \alpha$.

Παράδειγμα 2.5.5: Ένα τυχαίο δείγμα 15 παρατηρήσεων πάνω στο βάρος ενός πληθυσμού (σε κιλά) έδωσε τα εξής αποτελέσματα:

(11) 72.6	(3) 70.0	(8) 71.3	(4) 70.5	(5) 70.9
(1) 60.8	(13) 76.0	(15) 90.1	(10) 72.5	(7) 71.2
(6) 71.1	(14) 80.6	(9) 71.9	(12) 72.8	(2) 65.7

(Οι αριθμοί που εμφανίζονται στην παρένθεση αναφέρονται στην κατ' αύξουσα σειρά διάταξη των παρατηρήσεων). Σε επίπεδο σημαντικότητας περίπου 10%, να ελεγχθούν οι εξής υποθέσεις για την διάμεσο $x_{0.5}$ του πληθυσμού:

$$H_0 : x_{0.5} = 71.2$$

$$H_1 : x_{0.5} \neq 71.2$$

Λύση: Ο έλεγχος των υποθέσεων αυτών ισοδυναμεί με την κατασκευή ενός διαστήματος εμπιστοσύνης για την διάμεσο $x_{0.5}$ του πληθυσμού με συντελεστή εμπιστοσύνης περίπου ίσο με 90%, δηλαδή, ενός τυχαίου διαστήματος της μορφής $[X^{(r)}, X^{(s)}]$ με $P(X^{(r)} \leq x_{0.5} \leq X^{(s)}) \cong 90\%$.

Προφανώς, οι τιμές r και s ορίζονται αντίστοιχα από τις σχέσεις

$$P(X_b \leq r-1 | n=15, p=0.5) = \alpha_1 \cong 0.05,$$

$$P(X_b \leq s-1 | n=15, p=0.5) = 1 - \alpha_2 \cong 0.95,$$

όπου X_b ορίζεται όπως στα προηγούμενα.

Από τον πίνακα της διωνυμικής μεταβλητής προκύπτει ότι

$$r-1 = 4 \text{ (ισοδύναμα, } r = 5) \text{ όταν } \alpha_1 = 0.0592$$

και

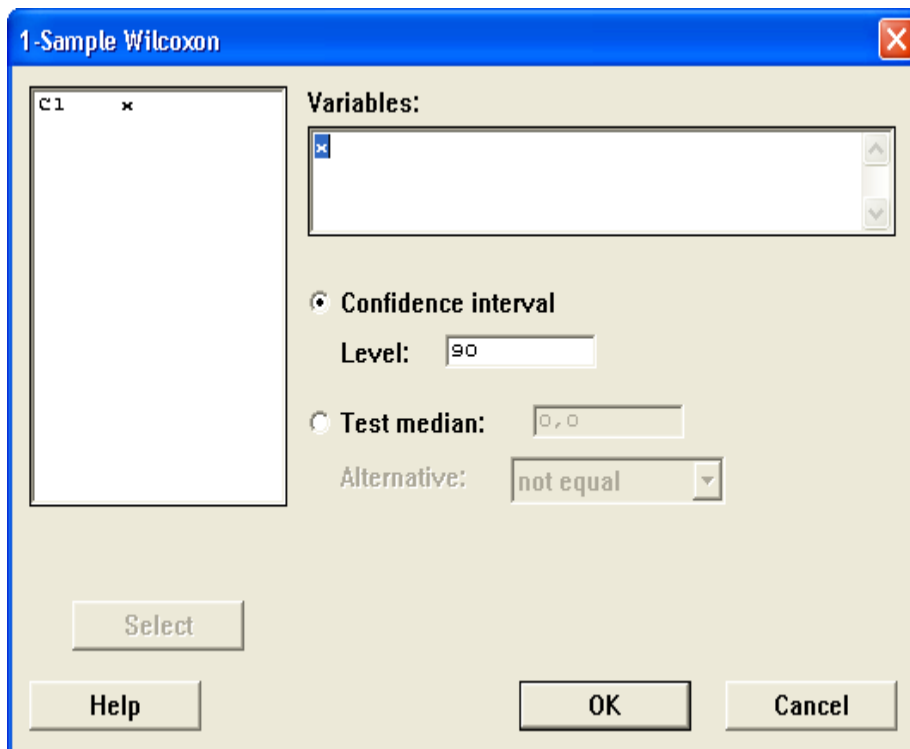
$$s-1 = 10 \text{ (ισοδύναμα, } s = 11) \text{ όταν } \alpha_2 = 0.0592.$$

Επομένως $P(X^{(5)} \leq x_{0.5} \leq X^{(11)}) = 1 - 0.0592 - 0.0592 = 0.8816 \cong 90\%$.

Κατασκευάσθηκε, δηλαδή, ένα διάστημα εμπιστοσύνης για την διάμεσο $x_{0.5}$ του πληθυσμού με άκρα την 5η και την 11η κατά αύξουσα σειρά μεγέθους παρατήρηση, συγκεκριμένα το διάστημα $[70.9, 72.6]$, με συντελεστή εμπιστοσύνης 88.16%. Το διάστημα αυτό περιέχει την τιμή 71.2. Αυτό συνεπάγεται ότι, σε επίπεδο σημαντικότητας $\alpha = \alpha_1 + \alpha_2 = 0.0592 + 0.0592 = 0.484$, δεν απορρίπτεται η μηδενική υπόθεση.

Λύση με το MINITAB: Ο έλεγχος των υποθέσεων του παραδείγματος, με την κατασκευή ενός διαστήματος εμπιστοσύνης για την διάμεσο του πληθυσμού είναι δυνατός μέσω του πακέτου αυτού

μόνο με την χρήση μιας στατιστικής συνάρτησης στην οποία βασίζεται ένας μη παραμετρικός έλεγχος, γνωστός ως έλεγχος Wilcoxon, ο οποίος ορίζεται στο επόμενο κεφάλαιο. Η λύση που παρατίθεται στην συνέχεια, στηρίζεται σ' αυτήν την στατιστική συνάρτηση και δεν αποτελεί το ακριβές ανάλογο της λύσης που δόθηκε παραπάνω. Δίνεται όμως με σκοπό να δείξει τον έλεγχο υποθέσεων με την κατασκευή ενός κατάλληλου μη παραμετρικού διαστήματος εμπιστοσύνης με το πακέτο MINITAB. Για την κατασκευή ενός διαστήματος εμπιστοσύνης για την διάμεσο $x_{0.5}$ με συντελεστή εμπιστοσύνης 90%, καταχωρίζουμε τις τιμές του δείγματος σε μία μεταβλητή (έστω x) και στην συνέχεια, επιλέγοντας κατά σειρά **Stat, Nonparametrics, 1-Sample Wilcoxon** στο βασικό μενού, οδηγούμεθα στο εξής πλαίσιο διαλόγου:



Στο πεδίο **Variables**, δηλώνουμε την μεταβλητή που περιέχει το δείγμα. Επίσης, επιλέγουμε **Confidence Interval** και στο πεδίο **Level** δηλώνουμε τον επιθυμητό βαθμό εμπιστοσύνης. Πιέζοντας **OK**, προκύπτουν τα εξής αποτελέσματα:

Sign confidence interval for median					
			Achieved	Confidence interval Position	
	N	Median	Confidence	Confidence interval	
x	15	71.30	0.8815	(70.90; 72.60)	
	5		0.9000	(70.86; 72.62)	
NLI			0.9648	(70.50; 72.80)	
4					

Το 90% διάστημα εμπιστοσύνης που προκύπτει περιέχει την τιμή 71.2 και, επομένως, οι ενδείξεις από το δείγμα δεν συνηγορούν υπέρ της απόρριψης της μηδενικής υπόθεσης.

Λύση με το SPSS: Το SPSS δεν υπολογίζει διαστήματα εμπιστοσύνης για την διάμεσο ενός πληθυσμού. Επομένως, ο έλεγχος των υποθέσεων του παραδείγματος με την κατασκευή διαστήματος εμπιστοσύνης για την διάμεσο με βάση διαδικασία που στηρίζεται στον διωνυμικό έλεγχο δεν είναι δυνατός με το πακέτο αυτό. (Μπορεί όμως να λυθεί με την κατασκευή διαστήματος εμπιστοσύνης μέσω διαδικασίας που στηρίζεται σε άλλο μη παραμετρικό έλεγχο γνωστό ως έλεγχο Wilcoxon-Mann-Whitney για την διάμεσο ενός πληθυσμού. Ο έλεγχος αυτός παρουσιάζεται στο επόμενο κεφάλαιο).

Λύση με το SAS: Το SAS δεν υπολογίζει διαστήματα εμπιστοσύνης για την διάμεσο ενός πληθυσμού.

Παράδειγμα 2.5.6: Ποιο είναι το 95% διάστημα εμπιστοσύνης για το 95ο ποσοστιαίο σημείο του πληθυσμού των οδηγών, από τον οποίο επελέγη το δείγμα του προηγούμενου παραδείγματος 2.5.1;

Λύση: Πρέπει να κατασκευασθεί ένα διάστημα της μορφής $[X^{(r)}, X^{(s)}]$ με $P(X^{(r)} \leq x_{0.95} \leq X^{(s)}) \cong 0.95 = 1 - \alpha$. Το αριστερό μέλος της ισότητας αυτής είναι ίσο με $P(X^{(s)} \geq x_{0.95}) - P(X^{(r)} > x_{0.95})$ ή (επειδή έχουμε συνεχή πληθυσμό) $P(X^{(s)} > x_{0.95}) - P(X^{(r)} > x_{0.95})$.

Θέτουμε

$$P(X^{(s)} > x_{0.95}) = 1 - \alpha_1 \cong 0.975$$

και

$$P(X^{(r)} > x_{0.95}) = \alpha_2 \cong 0.025.$$

Αλλά,

$$P(X^{(s)} > x_{0.95}) = P(X_b \leq s - 1 \mid n = 100, p = 0.95)$$

και

$$P(X^{(r)} > x_{0.95}) = P(X_b \leq r - 1 \mid n = 100, p = 0.95).$$

Επομένως, οι τιμές των s και r ορίζονται από τις σχέσεις

$$P(X_b \leq s - 1 \mid n = 100, p = 0.95) \cong 0.975$$

και

$$P(X_b \leq r - 1 \mid n = 100, p = 0.95) \cong 0.025.$$

Επειδή το μέγεθος του δείγματος είναι μεγάλο και ισχύει ότι $np \geq 5$ και $n(1-p) \geq 5$, μπορεί να χρησιμοποιηθεί η προσέγγιση της διωνυμικής κατανομής από την κανονική. Επομένως, οι τιμές των s και r προσδιορίζονται από τις σχέσεις

$$P\left(Z \leq \frac{s - 95}{\sqrt{4.75}}\right) = 0.975 \text{ και } P\left(Z \leq \frac{r - 95}{\sqrt{4.75}}\right) = 0.025.$$

Από την πρώτη σχέση, έχουμε,

$$\frac{s - 95}{\sqrt{4.75}} = z_{0.975} = 1.96 \Rightarrow s = 99.27 \cong 100.$$

Άρα, $X^{(s)} = X^{(100)} = 190$. Με όμοιο τρόπο, προκύπτει ότι $r = 90.72 \cong 91$. Επομένως, $X^{(r)} = X^{(91)} = 176.3$ και κατά συνέπεια, ένα 95% διάστημα εμπιστοσύνης για το $x_{0.95}$, το 95ο ποσοστιαίο σημείο της κατανομής του ύψους, είναι το διάστημα (176.3, 190).

Λύση με το MINITAB, το SPSS ή το SAS: Το παράδειγμα αυτό απαιτεί αναλυτική λύση και υπολογισμούς για την μετατροπή των ποσοστιαίων σημείων της $N(0,1)$ κατανομής σε αντίστοιχες τιμές των δεικτών s και r , για τους οποίους δεν προσφέρονται τα πακέτα MINITAB, SPSS και SAS.

Συχνά, απαιτείται η κατασκευή μονόπλευρων διαστημάτων εμπιστοσύνης, δηλαδή, διαστημάτων της μορφής $(-\infty, X^{(s)})$ με συντελεστή εμπιστοσύνης $1-\alpha_1 \cong 1-\alpha$, ή της μορφής $[X^{(r)}, +\infty)$ με συντελεστή εμπιστοσύνης $1-\alpha_2 \cong 1-\alpha$. Τα διαστήματα αυτά μπορούν να προκύψουν μέσω του προσδιορισμού της ακέραιης τιμής r , έτσι ώστε $P(X^{(r)} \leq x_p) = 1-\alpha_1 \cong 1-\alpha$ ή της ακέραιης τιμής s , έτσι ώστε $P(X^{(s)} \geq x_p) = 1-\alpha_2 \cong 1-\alpha$, αντίστοιχα.

Από την θεωρία που αναπτύχθηκε, η τιμή r προσδιορίζεται έτσι ώστε

$$P(X_b \leq r-1 | n, p) = \alpha_1 \cong \alpha,$$

ενώ η τιμή s προσδιορίζεται έτσι ώστε

$$P(X_b \leq s-1 | n, p) = 1-\alpha_2 \cong 1-\alpha.$$

Όπως στην περίπτωση των αμφίπλευρων διαστημάτων εμπιστοσύνης, και οι δύο μορφές μονόπλευρων διαστημάτων εμπιστοσύνης ισοδυναμούν με ελέγχους υποθέσεων για το p -ποσοστιαίο σημείο ενός πληθυσμού. Συγκεκριμένα, η κατασκευή του μονόπλευρου διαστήματος εμπιστοσύνης της μορφής $(-\infty, X^{(s)})$ με

συντελεστή εμπιστοσύνης $1-\alpha_1 \cong 1-\alpha$ ισοδυναμεί με τον έλεγχο των υποθέσεων

$$H_0 : x_p \leq x^*$$

$$H_1 : x_p > x^*$$

σε επίπεδο σημαντικότητας $\alpha_1 \cong \alpha$.

Αντίστοιχα, η κατασκευή διαστήματος εμπιστοσύνης της μορφής $[X^{(t)}, +\infty)$ με συντελεστή εμπιστοσύνης $1-\alpha_2 \cong 1-\alpha$ ισοδυναμεί με τον έλεγχο των υποθέσεων

$$H_0 : x_p \geq x^*$$

$$H_1 : x_p < x^*$$

σε επίπεδο σημαντικότητας $\alpha_2 \cong \alpha$.

Παράδειγμα 2.5.7: Ας υποθέσουμε ότι στο προηγούμενο παράδειγμα ενδιαφερόμαστε να ελέγξουμε την υπόθεση ότι η διάμεσος του βάρους του πληθυσμού δεν υπερβαίνει την τιμή 71.2 σε επίπεδο σημαντικότητας περίπου ίσο με 10%. Είναι προφανές ότι οι προς έλεγχο υποθέσεις έχουν την μορφή

$$H_0 : x_{0.5} \leq 71.2$$

$$H_1 : x_{0.5} > 71.2 .$$

Ο έλεγχος των υποθέσεων αυτών ισοδυναμεί με την κατασκευή ενός διαστήματος εμπιστοσύνης της μορφής $(-\infty, X^{(s)}]$ τέτοιου ώστε $P(-\infty < x_p \leq X^{(s)}) \cong 90\%$ ή, ισοδύναμα, $P(X^{(s)} \geq x_p) = 1-\alpha_1 \cong 90\%$.

Από την σχετική θεωρία που αναπτύχθηκε, η τιμή s προσδιορίζεται έτσι ώστε

$$P(X_b \leq s-1 | n = 15, p = 0.5) = 1-\alpha_1 \cong 90\% .$$

Από τον πίνακα της διωνυμικής κατανομής, προκύπτει ότι

$$s - 1 = 9 \text{ (ισοδύναμα, } s = 10) \text{ όταν } 1 - \alpha_1 = 0.8481$$

Επομένως, το πάνω άκρο του διαστήματος που αντιστοιχεί σε συντελεστή εμπιστοσύνης 84.81% είναι η 10η κατά αύξουσα σειρά μεγέθους παρατήρηση του δείγματος, δηλαδή η τιμή 72.5. Όπως παρατηρείται, η τιμή 71.2 περιλαμβάνεται στο διάστημα $(-\infty, 72.5]$ που αποτελεί ένα 84.81% διάστημα εμπιστοσύνης για την διάμεσο του πληθυσμού. Επομένως, σε επίπεδο σημαντικότητας $1 - \alpha_1 = 1 - 0.8481 = 0.1509$ η μηδενική υπόθεση δεν απορρίπτεται.

Με ανάλογο τρόπο μπορούμε να προχωρήσουμε για τον έλεγχο των εξής υποθέσεων για την διάμεσο του πληθυσμού

$$H_0 : x_{0.5} \geq 71.2$$

$$H_1 : x_{0.5} < 71.2,$$

σε επίπεδο σημαντικότητας περίπου ίσο με 10%. Όπως και προηγουμένως, για τον έλεγχο αυτών των υποθέσεων αρκεί να κατασκευασθεί ένα διάστημα $[X^{(r)}, +\infty)$ με την ιδιότητα $P(X^{(r)} \leq x_{p^*} < +\infty) \cong 0.90$ ή, ισοδύναμα, με την ιδιότητα $1 - P(X^{(r)} \geq x_{p^*}) \cong 0.90$. Η τελευταία σχέση γράφεται με την μορφή $P(X^{(r)} \geq x_{0.5}) \cong 10\%$. Επομένως, η τιμή r μπορεί να προσδιορισθεί από τον πίνακα της διωνυμικής κατανομής, έτσι ώστε:

$$P(X_b \leq r - 1 \mid n = 15, p = 0.5) = \alpha_2 \cong 10\%.$$

Πράγματι, από τον πίνακα της διωνυμικής κατανομής προκύπτει ότι $r - 1 = 5$ (ισοδύναμα, $r = 6$) όταν $\alpha_2 = 0.1509$, δηλαδή, το κάτω άκρο του διαστήματος που αντιστοιχεί σε συντελεστή εμπιστοσύνης $1 - \alpha_2 = 1 - 0.1509 = 0.8491$ είναι η 6η κατά αύξουσα σειρά μεγέθους παρατήρηση του δείγματος και έχει την τιμή 71.1. Παρατηρούμε ότι η

τιμή 71.2 ανήκει στο κατασκευασθέν διάστημα $[71.1, +\infty)$ που αποτελεί ένα διάστημα εμπιστοσύνης για την διάμεσο του πληθυσμού με συντελεστή εμπιστοσύνης $1-\alpha_2 = 1-0.1509=0.8491$. Επομένως, σε επίπεδο σημαντικότητας $\alpha_2 = 0.1509$, η μηδενική υπόθεση δεν απορρίπτεται.

Λύση με το MINITAB: Το MINITAB δεν υπολογίζει μονόπλευρα διαστήματα εμπιστοσύνης για την διάμεσο ενός πληθυσμού με βάση διαδικασία που στηρίζεται στον διωνυμικό έλεγχο. Το παράδειγμα μπορεί, εναλλακτικά, να λυθεί με την κατασκευή μονόπλευρου διαστήματος εμπιστοσύνης με βάση την τεχνική διεξαγωγής του ελέγχου Wilcoxon του επομένου κεφαλαίου.

Λύση με το SPSS: Το SPSS δεν υπολογίζει μονόπλευρα διαστήματα εμπιστοσύνης για την διάμεσο ενός πληθυσμού. Όπως και στην περίπτωση του MINITAB, ο έλεγχος μέσω υπολογισμού ορίου εμπιστοσύνης μπορεί να γίνει με το πακέτο SPSS με βάση τον έλεγχο Wilcoxon του επομένου κεφαλαίου.

Λύση με το SAS: Το SAS δεν υπολογίζει διαστήματα εμπιστοσύνης για την διάμεσο ενός πληθυσμού.

Διακριτή Περίπτωση

Υπενθυμίζεται ότι ως p -ποσοστιαίο σημείο μίας διακριτής κατανομής ορίζεται η τιμή x_p με την ιδιότητα $P(X > x_p) \leq 1-p$ και $P(X < x_p) \leq p$.

Όπως ήδη αναφέρθηκε στα προηγούμενα, στόχος είναι η κατασκευή ενός διαστήματος της μορφής $[X^{(r)}, X^{(s)}]$ με την ιδιότητα $P(X^{(r)} \leq x_p < X^{(s)}) \geq 1 - \alpha$.

Προσδιορισμός των τιμών r και s : Όπως αποδείχθηκε στην συνεχή περίπτωση, η τιμή r προσδιορίζεται από την σχέση $P(X^{(r)} > x_p) = P(X_b \leq r - 1 | n, p)$.

Από τον ορισμό, όμως, του p -ποσοστιαίου σημείου, ισχύει ότι $P(X < x_p) \leq p$. Επομένως, οι τιμές της X τείνουν να υπερβαίνουν την τιμή x_p με πιθανότητα μικρότερη από αυτήν της συνεχούς περίπτωσης. Επίσης, πάλι από τον ορισμό της τιμής x_p , ισχύει ότι $P(X > x_p) \geq 1 - p$. Άρα, οι τιμές της X τείνουν να είναι τουλάχιστον ίσες με την τιμή x_p με πιθανότητα μεγαλύτερη από αυτήν της συνεχούς περίπτωσης. Συνεπώς, ακολουθώντας τα ίδια βήματα, όπως και στην συνεχή περίπτωση, έχουμε ότι

$$P(X^{(r)} > x_p) \leq P(X_b \leq r - 1 | n, p) \leq \alpha_1.$$

Με όμοιο τρόπο

$$P(X^{(s)} \geq x_p) \leq P(X_b \leq r - 1 | n, p) \geq 1 - \alpha_2.$$

Επομένως, οι τιμές r και s μπορούν να προσδιορισθούν από τον πίνακα της διωνυμικής κατανομής για τις δοθείσες τιμές των n και p έτσι ώστε

$$P(X_b \leq r - 1 | n, p) \leq \alpha_1 \cong \alpha/2$$

και

$$P(X_b \leq s - 1 | n, p) \geq 1 - \alpha_2 \cong 1 - \alpha/2.$$

Είναι προφανές ότι το διάστημα $[X^{(r)}, X^{(s)}]$ που προκύπτει αντιστοιχεί σε διάστημα εμπιστοσύνης για την διάμεσο ενός διακριτού πληθυσμού με συντελεστή εμπιστοσύνης

$$P(X^{(r)} \leq x_p \leq X^{(s)}) = P(X^{(s)} \geq x_p) - P(X^{(r)} > x_p) = 1 - \alpha_1 - \alpha_2 \geq 1 - \alpha.$$

Παράδειγμα 2.5.8: Ας υποθέσουμε ότι ενδιαφερόμαστε να κατασκευάσουμε ένα διάστημα εμπιστοσύνης για την διάμεσο του πληθυσμού των 20 τελειοφοίτων μαθητών γυμνασίου του παραδείγματος 2.5.2 με συντελεστή εμπιστοσύνης περίπου 90%.

Σύμφωνα με τα παραπάνω, το ζητούμενο διάστημα θα είναι της μορφής $[X^{(r)}, X^{(s)}]$ με την ιδιότητα $P(X^{(r)} \leq x_p \leq X^{(s)}) \geq 0.90$, όπου r και s ορίζονται έτσι ώστε:

$$P(X_b \leq r-1 | n=20, p=0.5) \leq \alpha_1 \cong 0.05$$

και

$$P(X_b \leq s-1 | n=20, p=0.5) \geq 1 - \alpha_2 \cong 0.95.$$

Από τον πίνακα της διωνυμικής κατανομής προκύπτει ότι

$$r-1 = 6 \text{ (ισοδύναμα, } r = 7) \text{ όταν } \alpha_1 = 0.0577$$

και

$$s-1 = 13 \text{ (ισοδύναμα, } s = 14) \text{ όταν } 1 - \alpha_2 = 0.9423.$$

Επομένως, η 7η και η 14η κατά αύξουσα σειρά μεγέθους παρατήρηση του δείγματος αποτελούν τα άκρα του διαστήματος, που αποτελεί ένα διάστημα εμπιστοσύνης για την διάμεσο του πληθυσμού, συγκεκριμένα του $[117, 137]$, με συντελεστή εμπιστοσύνης τουλάχιστον ίσο με $1 - \alpha_2 - \alpha_1 = 0.88461$.

Λύση με το MINITAB: Και στο παράδειγμα αυτό ισχύουν παρατηρήσεις ανάλογες με αυτές που κάναμε στο παράδειγμα 2.5.5

όσο αφορά το ακριβές ανάλογο της λύσης που δίνεται παραπάνω με την χρήση του MINITAB. Όπως και στην συνεχή περίπτωση, η λύση που παρατίθεται στα επόμενα δίνεται με στόχο να δείξει την χρήση του πακέτου στην κατασκευή μη παραμετρικού διαστήματος εμπιστοσύνης για την διάμεσο διακριτού πληθυσμού. Εργαζόμενοι όπως στο παράδειγμα 2.5.5, τα αποτελέσματα που μας δίνει το MINITAB είναι:

	N	Median	Estimated Confidence	Achieved Confidence	Interval
x	20	125.0	90.3	(114.5;	135.5)

Βλέπουμε ότι ο ακριβής βαθμός εμπιστοσύνης είναι 90.3% και το αντίστοιχο διάστημα είναι (114.5 , 135.5).

Λύση με το SPSS: Το SPSS δεν υπολογίζει διαστήματα εμπιστοσύνης για την διάμεσο ενός πληθυσμού.

Λύση με το SAS: Το SAS δεν υπολογίζει διαστήματα εμπιστοσύνης για την διάμεσο ενός πληθυσμού.

Περίπτωση Μεγάλων Δειγμάτων ($n \geq 20$): Στην περίπτωση μεγάλων δειγμάτων είναι δυνατή η χρησιμοποίηση της προσέγγισης της διωνυμικής κατανομής από την κανονική κατανομή για τον προσδιορισμό των τιμών r και s , ως αποτέλεσμα του γεγονότος ότι

$$P(X^{(r)} > x_p) \leq P(X_b \leq r-1 | n, p) \equiv P(X_b < r | n, p) \\ \cong P(X_{N(\mu, \sigma^2)} < r | \mu = np, \sigma^2 = np(1-p)) = P\left(Z < \frac{r - np}{\sqrt{np(1-p)}}\right).$$

Τότε, για τιμή αυτής της πιθανότητας ίση με $\alpha/2$, δηλαδή,

$$P\left(Z < \frac{r - np}{\sqrt{np(1-p)}}\right) = \frac{\alpha}{2},$$

έπεται ότι

$$\frac{r - np}{\sqrt{np(1-p)}} = z_{\alpha/2} = -z_{1-\alpha/2}$$

ή, ισοδύναμα,

$$r = -z_{1-\alpha/2} \sqrt{np(1-p)} + np.$$

Με όμοιο τρόπο, προκύπτει ότι

$$s = z_{1-\alpha/2} \sqrt{np(1-p)} + np.$$

Οι τιμές των r και s οι οποίες προκύπτουν από την παραπάνω σχέση στρογγυλοποιούνται στην αμέσως μεγαλύτερη ακέραιη τιμή. Στην συνέχεια, το κατά προσέγγιση διάστημα εμπιστοσύνης είναι της μορφής $[X^{(r)}, X^{(s)}]$ με συντελεστή εμπιστοσύνης $1-\alpha$ (αν η κατανομή είναι συνεχής) ή $\geq 1-\alpha$ (αν η κατανομή είναι διακριτή).

Με ανάλογο τρόπο, αντιμετωπίζονται και οι περιπτώσεις μονόπλευρων διαστημάτων εμπιστοσύνης.

Παράδειγμα 2.5.9: Στο προηγούμενο παράδειγμα, $n=20$, $1-\alpha \cong 0.90$.

Επομένως

$$z_{1-\alpha/2} = z_{0.95} = 1.645$$

και

$$r = 20(0.5) - 1.645\sqrt{20(0.5)(0.5)} = 6.32 \cong 7$$

και

$$s = 20(0.5) + 1.645\sqrt{20(0.5)(0.5)} = 13.68 \cong 14.$$

Κατά συνέπεια, το κατά προσέγγιση διάστημα εμπιστοσύνης για την διάμεσο του πληθυσμού με συντελεστή εμπιστοσύνης περίπου ίσο με 0.90 είναι το διάστημα $[117, 134]$ του οποίου άκρα αποτελούν η 7η και η 14η κατά αύξουσα σειρά μεγέθους παρατήρηση του δείγματος.

Λύση με το MINITAB: Το MINITAB δεν δίνει προσεγγιστικό διάστημα εμπιστοσύνης για την διάμεσο. Δίνει μόνο το ακριβές διάστημα εμπιστοσύνης που το είδαμε στο προηγούμενο παράδειγμα.

Λύση με το SPSS ή το SAS: Τα πακέτα αυτά δεν δίνουν την δυνατότητα κατασκευής προσεγγιστικού διαστήματος εμπιστοσύνης για την διάμεσο.

2.6 ΟΡΙΑ ΑΝΟΧΗΣ

Το διάστημα εμπιστοσύνης παρέχει μία εκτίμηση μιας άγνωστης παραμέτρου με την μορφή διαστήματος και ένα συγκεκριμένο βαθμό εμπιστοσύνης ότι το διάστημα αυτό, με τον τρόπο που κατασκευάστηκε, θα περιέχει την πραγματική τιμή της άγνωστης παραμέτρου. Τα όρια ανοχής διαφέρουν από τα διαστήματα εμπιστοσύνης κατά το ότι τα όρια ανοχής παρέχουν ένα διάστημα μέσα στο οποίο περιλαμβάνεται ένα ποσοστό των τιμών του πληθυσμού τουλάχιστον ίσο με $100q\%$ με πιθανότητα τουλάχιστον ίση με $1-\alpha$. Μία τυπική εφαρμογή των ορίων ανοχής αποτελεί η περίπτωση όπου, προκειμένου να επιλέξουμε ένα τυχαίο δείγμα n παρατηρήσεων X_1, X_2, \dots, X_n πάνω σε μία τυχαία μεταβλητή X , θέλουμε να ξέρουμε πόσο μεγάλο θα πρέπει να είναι ώστε να μπορούμε να είμαστε τουλάχιστον 95% βέβαιοι ότι τουλάχιστον το 90% του πληθυσμού θα περιλαμβάνεται μεταξύ της μικρότερης και της μεγαλύτερης παρατήρησης του δείγματος. Γενικεύοντας, θα μπορούσαμε να θεωρήσουμε περιπτώσεις όπου επιθυμούμε να προσδιορίσουμε το μέγεθος n ενός τυχαίου δείγματος ώστε, με πιθανότητα τουλάχιστον ίση με $1-\alpha$, το τυχαίο διάστημα $[X^{(t)}, X^{(s)}]$ να περιέχει ένα ποσοστό του πληθυσμού τουλάχιστον ίσο με $100q\%$.

Συνήθως, το s εκφράζεται μέσω του n με την μορφή $n+1-m$, όπου m γνωστή σταθερή τιμή, οπότε ενδιαφερόμαστε το διάστημα $[X^{(r)}, X^{(n-m+1)}]$ να περιέχει τουλάχιστον $100q$ % των τιμών του πληθυσμού με πιθανότητα τουλάχιστον ίση με $100(1-\alpha)$ %.

Το $X^{(r)}$ ονομάζεται *κάτω όριο ανοχής* ενώ το $X^{(n-m+1)}$ ονομάζεται *πάνω όριο ανοχής*.

Οι τιμές των r και m ορίζονται έτσι ώστε $0 \leq r \leq n+1$, $0 \leq m \leq n+1$. Αν $m = 0$, ως πάνω όριο ανοχής ορίζεται το $+\infty$, δηλαδή, $X^{(n-m+1)} = X^{(n+1)} = +\infty$. Αν $m = n+1$ ή $r = 0$, ως κάτω όριο ανοχής ορίζεται το $-\infty$, δηλαδή, $X^{(0)} = -\infty$. Ο συμβατικός αυτός ορισμός επιτρέπει τον προσδιορισμό μονόπλευρων ορίων ανοχής. Αυτά είναι της μορφής «τουλάχιστον $100q$ % των τιμών του πληθυσμού είναι τουλάχιστον ίσες με την τιμή $X^{(r)}$, με πιθανότητα τουλάχιστον ίση με $100(1-\alpha)$ %» ή της μορφής «τουλάχιστον $100q$ % των τιμών του πληθυσμού είναι το πολύ ίσες με την τιμή $X^{(n-m)}$, με πιθανότητα τουλάχιστον ίση με $100(1-\alpha)$ %».

Τα μονόπλευρα όρια ανοχής ταυτίζονται με τα μονόπλευρα διαστήματα εμπιστοσύνης για τα ποσοστιαία σημεία του πληθυσμού, όπως φαίνεται στα επόμενα.

Η μέθοδος προσδιορισμού των ορίων ανοχής, η οποία περιγράφεται στην συνέχεια, βασίζεται σε δείγματα που έχουν ληφθεί από άπειρους πληθυσμούς με ή χωρίς επανάθεση, ή σε δείγματα από πεπερασμένους πληθυσμούς που έχουν ληφθεί με επανάθεση, ώστε να υπάρχει ανεξαρτησία μεταξύ των X_1, X_2, \dots, X_n . Στην περίπτωση δειγμάτων που έχουν ληφθεί με επανάθεση από πεπερασμένο πληθυσμό και των οποίων το μέγεθος είναι μικρό σε σύγκριση με αυτό του πληθυσμού, τα όρια ανοχής που προσδιορίζονται με την μέθοδο

αυτή δεν είναι ακριβή, αλλά ισχύουν κατά προσέγγιση. Η κλίμακα μέτρησης είναι τουλάχιστον κλίμακα διάταξης.

Προσδιορισμός του Πάνω Ορίου Ανοχής: Ο προσδιορισμός του άνω ορίου ανοχής συνίσταται στον προσδιορισμό του μεγέθους n του δείγματος, το οποίο απαιτείται ώστε το διάστημα $(-\infty, X^{(n-m+1)}]$ να περιέχει τουλάχιστον $100q\%$ των τιμών του πληθυσμού με πιθανότητα $\cong 100(1-\alpha)\%$, δηλαδή,

$$P(\text{τουλάχιστον } 100q\% \text{ των τιμών της } X \text{ είναι } \leq X^{(n-m+1)}) \geq 1 - \alpha .$$

Εδώ η τυχαία μεταβλητή X συμβολίζει την μεταβλητή, η οποία περιγράφει τον πληθυσμό.

Είναι, όμως, γνωστό από τον ορισμό του ποσοστιαίου σημείου x_q μίας κατανομής, ότι τουλάχιστον $100q\%$ των τιμών του πληθυσμού είναι $\leq x_q$. Δηλαδή, $P(X \leq x_q) \geq q$.

Επομένως, το ενδεχόμενο {τουλάχιστον $100q\%$ των τιμών της X είναι $\leq X^{(n-m+1)}$ } είναι ισοδύναμο με το ενδεχόμενο $\{x_q < X^{(n-m+1)}\}$.

Κατά συνέπεια, το μέγεθος n του δείγματος μπορεί να προσδιορισθεί έτσι ώστε

$$P(x_q \leq X^{(n-m+1)}) \geq 1 - \alpha .$$

Όπως, όμως, αποδείχθηκε στην προηγούμενη ενότητα, για το q -ποσοστιαίο σημείο μιας κατανομής, ισχύει ότι

$$P(X^{(n-m+1)} \geq x_q) \leq P(X_b \leq n - m \mid n, q),$$

όπου X_b παριστάνει μία διωνυμική μεταβλητή με παραμέτρους n και q . Επομένως, η τιμή του n μπορεί να προσδιορισθεί έτσι ώστε

$$P(X_b \leq n - m \mid n, q) \geq P(X^{(n-m+1)} \geq x_q) \geq 1 - \alpha ,$$

δηλαδή, έτσι ώστε

$$P(X_b \leq n - m \mid n, q) \geq 1 - \alpha$$

ή, ισοδύναμα, έτσι ώστε

$$P(X_b \leq m - 1 \mid n, 1 - q) \leq \alpha .$$

Η ισοδυναμία των δύο τελευταίων συνθηκών είναι αποτέλεσμα της ισοδυναμίας του ενδεχομένου «*το πολύ n-m επιτυχίες*» με το ενδεχόμενο «*τουλάχιστον n αποτυχίες*» το οποίο είναι συμπληρωματικό του ενδεχομένου «*το πολύ n-1 αποτυχίες*». Πράγματι, από την ισότητα

$$P(X_b \leq n - m \mid n, p) = \sum_{r=0}^{n-m} P(X_b = r \mid n, p) = 1 - \sum_{r=n-m+1}^n P(X_b = r \mid n, p),$$

θεωρώντας την αλλαγή δεικτών $n - r \rightarrow j$, προκύπτει ότι

$$P(X \leq n - m) = 1 - \sum_{j=0}^{m-1} \binom{n}{n-j} p^{n-j} (1-p)^j = 1 - P(X_b \leq m - 1 \mid n, 1 - p) .$$

Επομένως, η τιμή του n προσδιορίζεται έτσι ώστε

$$1 - P(X_b \leq m - 1 \mid n, 1 - p) \geq 1 - \alpha$$

ή, ισοδύναμα, έτσι ώστε

$$P(X_b \leq m - 1 \mid n, 1 - p) \leq \alpha .$$

Προσδιορισμός Κάτω Ορίου Ανοχής: Ενδιαφερόμαστε να προσδιορίσουμε το μέγεθος του δείγματος, το οποίο απαιτείται ώστε

$$P(\text{τουλάχιστον } 100q\% \text{ των τιμών της } X \text{ είναι } \geq X^{(t)}) \geq 1 - \alpha .$$

Από τον ορισμό του ποσοστιαίου σημείου x_q , ισχύει ότι $P(X < x_q) \leq q$. Επομένως, $P(X \geq x_q) \geq 1 - q$ ή, ισοδύναμα, $P(X \geq x_{1-q}) \geq q$. Άρα, για τον προσδιορισμό του κάτω ορίου ανοχής, αρκεί να προσδιορισθεί η τιμή r , έτσι ώστε

$P(X^{(r)} \leq x_{1-q}) \geq 1 - \alpha$ ή, ισοδύναμα, $P(X^{(r)} > x_{1-q}) \leq \alpha$. Επειδή, όμως, όπως αποδείχθηκε στην προηγούμενη ενότητα,

$$P(X^{(r)} > x_{1-q}) \leq P(X_b \leq r-1 \mid n, 1-q),$$

η τιμή r μπορεί να προσδιορισθεί έτσι ώστε

$$P(X_b \leq r-1 \mid n, 1-q) \leq \alpha.$$

Παρατήρηση: Είναι προφανές, ότι για τον προσδιορισμό των ορίων ανοχής, μπορεί να χρησιμοποιηθεί ο πίνακας της διωνυμικής κατανομής. Στην περίπτωση που $r + m = 1$, δηλαδή είτε το $r = 0$ είτε το $m = 0$, όπως στην περίπτωση μονόπλευρου ορίου ανοχής, το μέγεθος n του δείγματος μπορεί να προσδιορισθεί απ' ευθείας από τον πίνακα 5 του παραρτήματος για τις κατάλληλες τιμές των α και q . Αν $r + m = 2$, η τιμή του n μπορεί να προσδιορισθεί απ' ευθείας από τον πίνακα 6 του παραρτήματος για τις κατάλληλες τιμές των α και q . Στην περίπτωση που κανένας από τους πίνακες αυτούς δεν προσφέρεται, μπορεί να χρησιμοποιηθεί η προσέγγιση

$$n \cong \frac{1}{4} \chi_{2(r+m), 1-\alpha}^2 \frac{1+q}{1-q} + \frac{1}{2}(r+m-1),$$

όπου $\chi_{2(r+m), 1-\alpha}^2$ συμβολίζει το $(1-\alpha)$ -ποσοστιαίο σημείο της κατανομής χ^2 με $2(r+m)$ βαθμούς ελευθερίας. Η τιμή που προκύπτει από την παραπάνω σχέση, στρογγυλεύεται στον πλησιέστερο μεγαλύτερο ακέραιο.

Παράδειγμα 2.6.1: Ποιο πρέπει να είναι το μέγεθος n ενός τυχαίου δείγματος ώστε με πιθανότητα τουλάχιστον ίση με 90% η παρατήρηση $X^{(n)}$ να είναι μεγαλύτερη ή ίση από ένα ποσοστό τιμών του πληθυσμού τουλάχιστον ίσο με 85%;

Λύση: Πρέπει να κατασκευασθεί ένα διάστημα της μορφής $(-\infty, X^{(n)})$ τέτοιο ώστε

$$P(\text{το } (-\infty, X^{(n)}) \text{ να περιέχει τουλάχιστον } 85\% \text{ των τιμών της } X) \geq 0.90.$$

Προφανώς, $1-\alpha = 0.90$, $q = 0.85$, $m = 1$. Επομένως,

$$\alpha = 0.10 \geq P(X_b \leq 0 | n, p = 1-q = 0.15).$$

Από τον πίνακα της διωνυμικής κατανομής, προκύπτει ότι, για $n = 15$, $P(X_b \leq 0 | n = 15, p = 0.15) = 0.0874 \leq 0.10$. Επομένως, η προσδιοριζόμενη τιμή του n είναι $n = 15$.

Η τιμή του n μπορεί, επίσης, να προσδιορισθεί από τον πίνακα 5 του παραρτήματος για $1-\alpha = 0.90$ και $q = 0.85$. Η προκύπτουσα τιμή είναι $n = 15$.

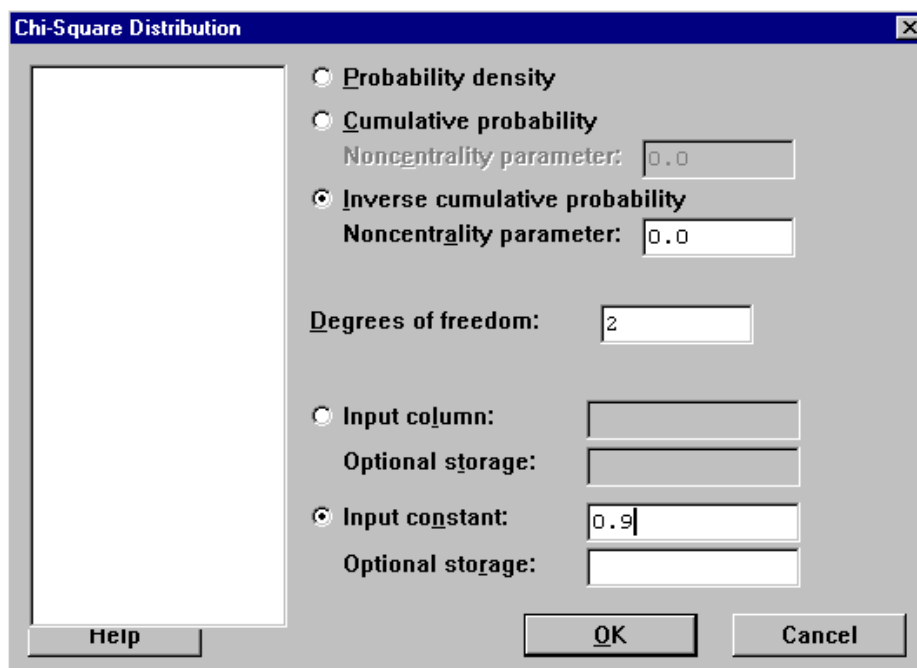
Η τιμή του n μπορεί, εναλλακτικά, να προσδιορισθεί με την χρήση του προσεγγιστικού τύπου ως εξής:

$$n \cong \frac{1}{4} \chi_{2, 0.90}^2 \frac{1+0.85}{1-0.85} + \frac{1}{2}(0+1-1) = 14.199 \cong 15,$$

όπου, από τον πίνακα της κατανομής χ^2 , $\chi_{2,0.90}^2 = 4.605$.

Λύση με το MINITAB: Η λύση του παραδείγματος απαιτεί υπολογισμούς που είναι ευκολότερο να γίνουν με απλό υπολογιστή χειρός. Το MINITAB μπορεί να βοηθήσει, όμως, στον υπολογισμό του ποσοστιαίου σημείου $\chi_{2,0.9}^2$ της κατανομής χ^2 με 2 βαθμούς ελευθερίας.

Επιλέγοντας **Calc, Probability Distributions, Chi-Square**, οδηγούμεθα στο εξής πλαίσιο διαλόγου:

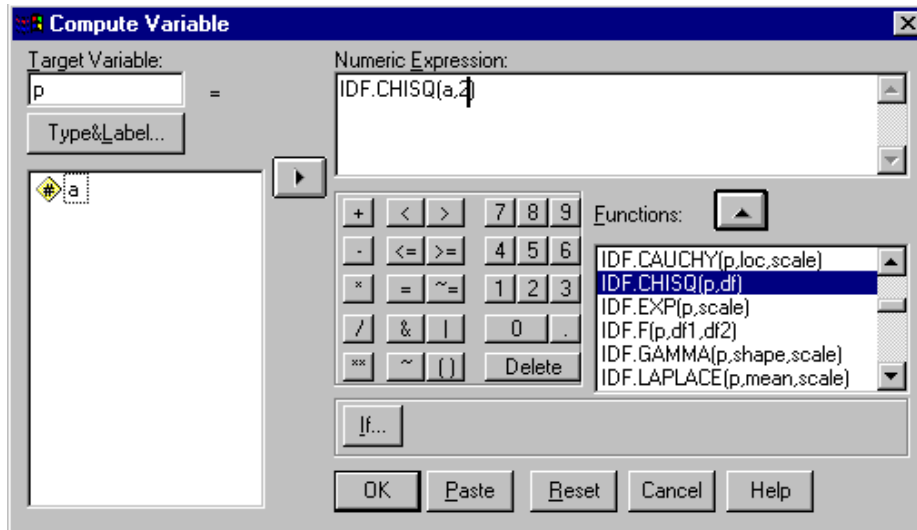


Στο πλαίσιο αυτό, δηλώνουμε ότι επιθυμούμε υπολογισμούς για την αντίστροφη αθροιστική συνάρτηση κατανομής (επιλογή **Inverse cumulative probability**), τους βαθμούς ελευθερίας της κατανομής (πεδίο **Degrees of freedom**) και την πιθανότητα $1-\alpha$ για την τιμή του α , που μας ενδιαφέρει (πεδίο **Input constant**). Το ζητούμενο 0.90 ποσοστιαίο σημείο περιέχεται στον παρακάτω πίνακα αποτελεσμάτων:

Chi-Square with 2 DF	
P(X <= x)	x
0.9000	4.6052.

Λύση με το SPSS: Όπως αναφέρθηκε παραπάνω, η λύση του παραδείγματος απαιτεί υπολογισμούς που μπορούν να γίνουν με απλό υπολογιστή χειρός. Όπως και το MINITAB, το SPSS μπορεί να χρησιμοποιηθεί για τον υπολογισμό του ποσοστιαίου σημείου $\chi^2_{2,0.90}$ της κατανομής χ^2 με 2 βαθμούς ελευθερίας.

Σε μία κενή στήλη (έστω **a**), καταχωρίζουμε την τιμή 0.9. Στην συνέχεια, επιλέγοντας **Transform, Compute**, προκύπτει το εξής πλαίσιο διαλόγου:



Στο πλαίσιο αυτό, ζητάμε να δημιουργηθεί μία νέα μεταβλητή της οποίας το όνομα (έστω **p**) δηλώνεται στο πεδίο **Target Variable**. Ο τρόπος υπολογισμού της δηλώνεται στο πεδίο **Numeric expression**. Εδώ δηλώνουμε **IDF.CHISQ(a,2)**, δηλαδή ζητάμε υπολογισμό της αντίστροφης αθροιστικής συνάρτησης κατανομής της κατανομής χ^2 με 2 βαθμούς ελευθερίας για πιθανότητα ίση με την τιμή που είναι καταχωρισμένη στην μεταβλητή **a**. Έτσι, προκύπτει ότι το ζητούμενο ποσοστιαίο σημείο είναι 4.61.

Λύση με το SAS: Το πακέτο αυτό δεν παρέχει την δυνατότητα επίλυσης προβλημάτων αυτής της μορφής με την χρήση του προσεγγιστικού τύπου.

Παράδειγμα 2.6.2: Ποιο πρέπει να είναι το μέγεθος n ενός τυχαίου δείγματος, ώστε με πιθανότητα τουλάχιστον ίση με 90%, ένα ποσοστό

τιμών του πληθυσμού τουλάχιστον ίσο με 95% να περιέχεται μεταξύ της μικρότερης και μεγαλύτερης παρατήρησης του δείγματος;

Λύση: Ζητείται να προσδιορισθεί η τιμή του n , για την οποία ισχύει

$$P[X^{(1)} \leq \text{τουλάχιστον } 0.95 \text{ του πληθυσμού} \leq X^{(n)}] \geq 0.90.$$

Εδώ, προφανώς, $r = 1$, $m = 1$, $1 - \alpha = 0.90$, $q = 0.95$. Επομένως, η τιμή του n , μπορεί να προσδιορισθεί από την σχέση

$$P(X_b \leq 1 | n, p = 1 - q = 0.05) \leq 0.1$$

με χρήση του πίνακα της διωνυμικής κατανομής. Επισκόπηση, όμως, του πίνακα αυτού δείχνει ότι η τιμή του n βρίσκεται πέρα από τις τιμές στις οποίες αναφέρεται ο πίνακας ($n > 20$). Πράγματι, κάνοντας χρήση του πίνακα 6 του παραρτήματος, για τις τιμές των $1 - \alpha$ και q , όπως προσδιορίστηκαν παραπάνω, έχουμε ότι $n = 77$. Εναλλακτικά, χρησιμοποιώντας τον προσεγγιστικό τύπο έχουμε

$$n \cong \frac{1}{4} \chi_{4,0.90}^2 \frac{1.95}{0.05} + \frac{1}{2} (1 + 1 - 1) = 76.34 \cong 77$$

όπου, από τον πίνακα της κατανομής χ^2 , $\chi_{4,0.90}^2 = 7.779$.

Παρατήρηση: Μπορεί ν' αποδειχθεί ότι, τόσο για τα αμφίπλευρα όσο και για τα μονόπλευρα όρια ανοχής, η τιμή n του μεγέθους του δείγματος προσδιορίζεται από την σχέση

$$P(X_b \leq r + m - 1 | n, 1 - q) \leq \alpha.$$

Η τιμή, δηλαδή, του n εξαρτάται μόνο από το άθροισμα $r + m$ και δεν επηρεάζεται από το εάν επιθυμούμε να επιλέξουμε ως διάστημα όλες τις τιμές δεξιά από την τιμή της $X^{(r+m)}$ ή όλες τις τιμές αριστερά από την τιμή της $X^{(n+1-r-m)}$ ή όλες τις τιμές μεταξύ των τιμών των $X^{(r)}$ και $X^{(n+1-m)}$ ή, ακόμη, όλες τις τιμές μεταξύ των τιμών των $X^{(i)}$ και $X^{(j)}$ για οποιαδήποτε ζεύγη (i, j) , $i < j$ με $j - i = n + 1 - m - r$.

Λύση με το MINITAB: Η λύση του παραδείγματος απαιτεί υπολογισμούς που είναι ευκολότερο να γίνουν με απλό υπολογιστή χειρός. Το MINITAB μπορεί να βοηθήσει, όμως, στον υπολογισμό του ποσοστιαίου σημείου $\chi^2_{4,0.9}$ της κατανομής χ^2 με 4 βαθμούς ελευθερίας. Όπως και στο προηγούμενο παράδειγμα, ο υπολογισμός αυτός γίνεται με την επιλογή **Calc, Probability Distributions, Chi-Square**. Το ζητούμενο ποσοστιαίο σημείο περιέχεται στον εξής πίνακα αποτελεσμάτων:

```
Chi-Square with 4 DF
P( X <= x )      x
0.9000           7.7794
```

Λύση με το SPSS: Όπως αναφέρθηκε παραπάνω, η λύση του παραδείγματος απαιτεί υπολογισμούς που μπορούν να γίνουν με απλό υπολογιστή χειρός. Όπως και το MINITAB, το SPSS μπορεί να χρησιμοποιηθεί για τον υπολογισμό του ποσοστιαίου σημείου $\chi^2_{4,0.9}$ της κατανομής χ^2 με 4 βαθμούς ελευθερίας.

Σε μία κενή στήλη (έστω **a**), καταχωρίζουμε την τιμή 0.9. Στην συνέχεια, επιλέγουμε **Transform, Compute** και ζητάμε να δημιουργηθεί μία νέα μεταβλητή (έστω **p**) δηλώνοντας **IDF.CHISQ(a,4)** στο πεδίο **Numeric Expression**. Έτσι, προκύπτει ότι το ζητούμενο ποσοστιαίο σημείο είναι 7.78.

Λύση με το SAS: Τα πακέτο αυτό δεν παρέχει την δυνατότητα επίλυσης προβλημάτων αυτής της μορφής με την χρήση του προσεγγιστικού τύπου.

Παράδειγμα 2.6.3: Ποιο πρέπει να είναι το n , ώστε με πιθανότητα $\geq 95\%$, ένα ποσοστό τιμών του πληθυσμού τουλάχιστον ίσο με 90% να είναι $\geq X^{(1)}$;

Λύση: Επιθυμούμε το διάστημα $[X^{(1)}, +\infty)$, δηλαδή το διάστημα $[X^{(1)}, X^{(n+1)})$ να περιέχει το 90% τουλάχιστον των τιμών του πληθυσμού με πιθανότητα $\geq 95\%$.

Σύμφωνα με την τελευταία παρατήρηση, για οποιοδήποτε τύπο ορίου ανοχής, μπορούμε να προσδιορίσουμε την τιμή του n έτσι ώστε $P(X_b \leq r + m - 1 | n, 1 - q) \leq \alpha$. Τότε, κάνοντας χρήση του προσεγγιστικού τύπου, έχουμε

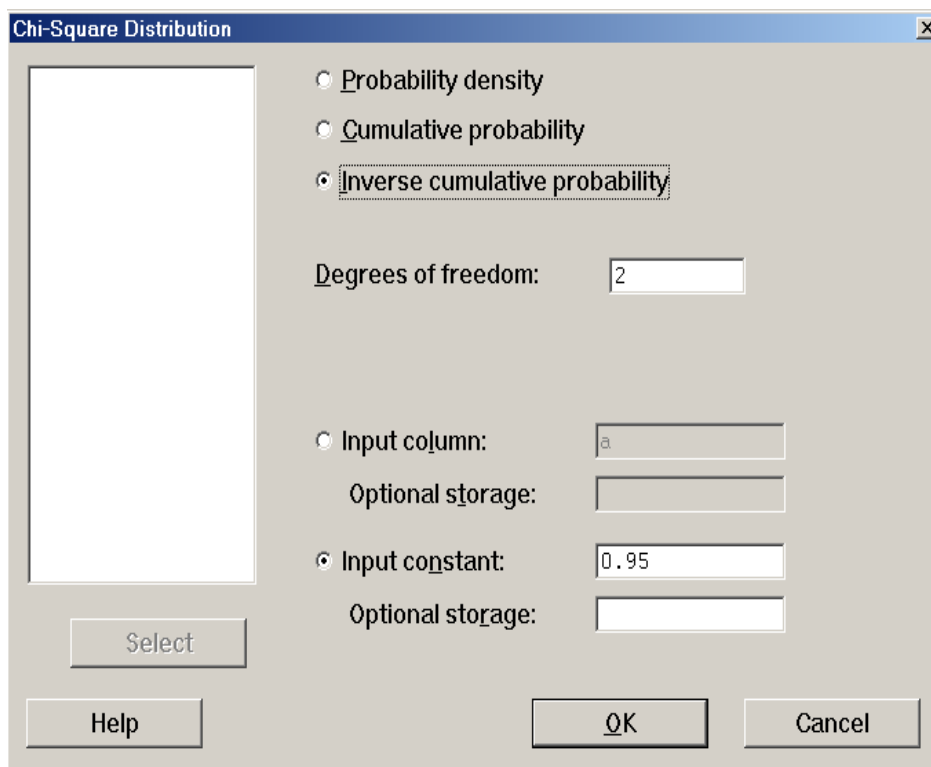
$$n = \frac{1}{4} \chi_{2(r+m), 1-\alpha}^2 \frac{1+q}{1-q} + \frac{1}{2}(r+m-1).$$

Εδώ $r=1$, $m=0$. Επομένως,

$$n = \frac{1}{4} \chi_{2, 0.95}^2 \frac{1+0.9}{1-0.9} + \frac{1}{2}(1+0-1) = 28.46 \cong 29.$$

Λύση με το MINITAB: Το παράδειγμα απαιτεί υπολογισμούς που μπορούν να γίνουν με απλό υπολογιστή χειρός. Το μόνο σημείο στο οποίο μπορεί να βοηθήσει το MINITAB είναι στον υπολογισμό του ποσοστιαίου σημείου $\chi_{2, 0.95}^2$ της κατανομής χ^2 με 2 βαθμούς ελευθερίας.

Επιλέγουμε **Calc, Probability Distributions, Chi-Square** και στο πλαίσιο διαλόγου που προκύπτει



δηλώνουμε ότι επιθυμούμε υπολογισμούς για την αντίστροφη αθροιστική συνάρτηση κατανομής (επιλογή **Inverse cumulative probability**), τους βαθμούς ελευθερίας της κατανομής (πεδίο **Degrees of freedom**) και την πιθανότητα για την οποία θέλουμε τον υπολογισμό (πεδίο **Input constant**). Το ζητούμενο ποσοστιαίο σημείο περιέχεται στον παρακάτω πίνακα αποτελεσμάτων.

Chi-Square with 2 DF

$P(X \leq x)$	x
0.9500	5.9915.

Λύση με το SPSS: Όπως τονίσθηκε παραπάνω, το παράδειγμα απαιτεί υπολογισμούς που μπορούν να γίνουν με απλό υπολογιστή χειρός. Το

σημείο στο οποίο μπορεί να βοηθήσει το SPSS είναι και πάλι στον υπολογισμό του ποσοστιαίου σημείου $\chi^2_{2,0.95}$ της κατανομής χ^2 με 2 βαθμούς ελευθερίας.

Σε μία κενή στήλη (έστω **a**) καταχωρίζουμε την τιμή 0.95. Κατόπιν, επιλέγουμε **Transform, Compute** και ζητάμε να δημιουργηθεί μία νέα μεταβλητή (έστω **p**) με **expression IDF.CHISQ(a,2)**. Έτσι, προκύπτει ότι το ζητούμενο ποσοστιαίο σημείο είναι 5.99.

Λύση με το SAS: Το SAS δεν παρέχει την δυνατότητα επίλυσης τέτοιων προβλημάτων.

ΑΣΚΗΣΕΙΣ

1. Οι μηνιαίες καθαρές εισπράξεις σε χιλιάδες δραχμές είκοσι καταστημάτων μιας περιοχής δίνονται στον πίνακα που ακολουθεί. Είναι βάσιμο να υποθέσει κάποιος ότι το ποσοστό των καταστημάτων με καθαρές μηνιαίες εισπράξεις τουλάχιστον 150 χιλ. δραχ. είναι τουλάχιστον ίσο με 75% ($\alpha=5\%$);

(Οικ. Παν/μιο Αθηνών – Εξετ. Φεβρ. 2000)

Μηνιαίες καθαρές εισπράξεις (σε χιλιάδες δραχμές)			
103	158	137	131

128	144	93	119
117	161	122	98
81	119	154	134
165	86	103	142

2. Από 15 λαμπτήρες που ελέγχθηκαν, σε ένα εργοστάσιο παραγωγής λαμπτήρων με απλή τυχαία δειγματοληψία, 5 κρίθηκαν ελαττωματικοί. Με βάση αυτό το δεδομένο, μπορεί να υποστηριχθεί η υπόθεση ότι 10% των λαμπτήρων του εργοστασίου αυτού είναι ελαττωματικοί;

(Οικ. Παν/μιο Αθηνών – Εξετ. Σεπτ. 1999)

3. Ένας φοιτητής εξετάζεται σε ένα μάθημα με βάση ένα τεστ με ερωτήσεις πολλαπλής επιλογής. Κάθε ερώτηση προσφέρει τέσσερις επιλογές, από τις οποίες μόνο μία είναι σωστή.

i) Διατυπώστε το ζεύγος υποθέσεων (μηδενική, εναλλακτική), μέσω του ελέγχου των οποίων θεωρείτε ότι θα είσαστε σε θέση να συμπεράνετε αν είναι εύλογος ο ισχυρισμός ότι ο φοιτητής δεν γνωρίζει τίποτα για το αντικείμενο που εξετάζεται και απαντά στην τύχη.

ii) Ο φοιτητής υποβάλλεται σε τεστ 15 ερωτήσεων. Για να είναι εύλογο να συμπεράνουμε (σ' επίπεδο σημαντικότητας $\alpha \approx 0.05$) ότι δεν απαντάει στην τύχη, πόσες σωστές απαντήσεις πρέπει να δώσει;

- | | |
|-------------------|------------------|
| α) Τουλάχιστον 6 | γ) Τουλάχιστον 7 |
| β) Τουλάχιστον 10 | δ) Τουλάχιστον 8 |

iii) Ας υποθέσουμε ότι ο συγκεκριμένος φοιτητής έδωσε 8 σωστές απαντήσεις σε 10 ερωτήσεις. Ποιο είναι το κρίσιμο επίπεδο του

ελέγχου της υπόθεσης ότι απαντά στην τύχη έναντι της εναλλακτικής ότι γνωρίζει το αντικείμενο;

- | | |
|-----------|-----------|
| α) 0.0001 | γ) 0.0002 |
| β) 0.0004 | δ) 0.0008 |

(Οικ. Παν/μιο Αθηνών – Εξετ. Δεκ. 2000)

4. Ένας παίκτης του ΠΑΜΕ ΣΤΟΙΧΗΜΑ πρέπει να προβλέψει το αποτέλεσμα κάποιων αγώνων ποδοσφαίρου. Σε κάθε αγώνα, έχει τρεις επιλογές, δηλαδή νίκη του γηπεδούχου, ισοπαλία ή νίκη του φιλοξενούμενου.

i) Διατυπώστε το ζεύγος υποθέσεων (μηδενική, εναλλακτική), μέσω του ελέγχου των οποίων θεωρείτε ότι θα είσαστε σε θέση να συμπεράνετε αν είναι εύλογος ο ισχυρισμός ότι ο παίκτης προβλέπει στην τύχη χωρίς να γνωρίζει τίποτε για τις ομάδες.

ii) Ο παίκτης παίζει ένα δελτίο όπου σημειώνει την πρόβλεψή του για το αποτέλεσμα 12 αγώνων. Για να είναι εύλογο να συμπεράνουμε (σ' επίπεδο σημαντικότητας $\alpha \approx 0.1$) ότι δεν προβλέπει στην τύχη, πόσες σωστές προβλέψεις πρέπει να κάνει;

- | | |
|------------------|------------------|
| α) Τουλάχιστον 6 | γ) Τουλάχιστον 8 |
| β) Τουλάχιστον 7 | δ) Τουλάχιστον 5 |

iii) Ας υποθέσουμε ότι ο συγκεκριμένος παίκτης έκανε 11 σωστές προβλέψεις σε 14 αγώνες. Ποιο είναι το κρίσιμο επίπεδο του ελέγχου της υπόθεσης ότι προβλέπει στην τύχη έναντι της εναλλακτικής ότι χρησιμοποιεί τις γνώσεις του;

- | | |
|-----------|-----------|
| α) 0.0001 | γ) 0.0011 |
| β) 0.0002 | δ) 0.0022 |

(Οικ. Παν/μιο Αθηνών – Εξετ. Δεκ. 2000)

5. Πιστεύεται ότι η χορήγηση κάποιας ορμόνης σε ποντίκια μειώνει τις πιθανότητες γέννησης αρσενικών απογόνων.

i) Διατυπώστε το ζεύγος υποθέσεων (μηδενική, εναλλακτική), μέσω του ελέγχου των οποίων θεωρείτε ότι θα είσαστε σε θέση να συμπεράνετε αν είναι εύλογη η παραπάνω επιστημονική «θεωρία».

ii) Ένα θηλυκό ποντίκι υποβάλλεται σε θεραπεία με αυτή την ορμόνη. Αργότερα, γεννάει 13 μικρά. Πόσα από αυτά πρέπει να είναι αρσενικά ώστε να είναι εύλογο να συμπεράνουμε (σ' επίπεδο σημαντικότητας $\alpha \approx 0.05$) ότι, όντως, η ορμόνη μειώνει τις πιθανότητες γέννησης αρσενικών;

- | | |
|-------------------------------|-------------------|
| α) Το πολύ 2 ή τουλάχιστον 11 | γ) Τουλάχιστον 11 |
| β) Το πολύ 2 | δ) Το πολύ 3 |

iii) Ας υποθέσουμε ότι το θηλυκό ποντίκι γεννά 12 μικρά και μόνο 2 είναι αρσενικά. Ποιο είναι το κρίσιμο επίπεδο του ελέγχου ότι η ορμόνη μειώνει τις πιθανότητες γέννησης αρσενικών;

- | | |
|----------|-----------|
| α) 0.193 | γ) 0.0033 |
| β) 0.386 | δ) 0.0066 |

(Οικ. Παν/μιο Αθηνών – Εξετ. Δεκ. 2000)

6. Στο πλαίσιο μιας μελέτης για να διερευνηθεί αν η χορήγηση συγκεκριμένης ορμόνης σε ποντίκια αυξάνει την νοημοσύνη τους, τα ποντίκια «διδάσκονται» πώς να βγαίνουν από έναν λαβύρινθο και, στην συνέχεια, εξετάζεται αν μπορούν να βρουν το σωστό δρόμο. Ο λαβύρινθος έχει 5 εναλλακτικές διαδρομές, από τις οποίες μόνο μία είναι σωστή.

i) Διατυπώστε το ζεύγος υποθέσεων (μηδενική, εναλλακτική), μέσω του ελέγχου των οποίων θεωρείτε ότι θα είσαστε σε θέση να συμπεράνετε αν είναι εύλογος ο ισχυρισμός ότι η ορμόνη αυξάνει τη νοημοσύνη των ποντικών.

ii) 15 ποντίκια υποβάλλονται σε θεραπεία με την ορμόνη αυτή, «διδάσκονται» πώς να βρίσκουν το δρόμο τους στο λαβύρινθο και μετά αφήνονται ένα – ένα μέσα σ' αυτόν. Για να είναι εύλογο να συμπεράνουμε (σ' επίπεδο σημαντικότητας $\alpha \approx 0.05$) ότι η ορμόνη αυξάνει την νοημοσύνη τους, πόσα ποντίκια πρέπει να βρουν το σωστό δρόμο;

- α) Τουλάχιστον 5 γ) Το πολύ 1
β) Τουλάχιστον 4 δ) Τουλάχιστον 6

iii) Ας υποθέσουμε ότι από 14 ποντίκια 4 βρήκαν το σωστό δρόμο μέσα στο λαβύρινθο. Ποιο είναι το κρίσιμο επίπεδο του ελέγχου ότι ο ορμόνη αυξάνει τη νοημοσύνη τους;

- α) 0.2298 γ) 0.3018
β) 0.1509 δ) 0.1149

(Οικ. Παν/μιο Αθηνών – Εξετ. Δεκ. 2000)

7. Οι τιμές που συνοψίζονται στον παρακάτω πίνακα είναι μέσες θερμοκρασίες σε μια πόλη για 12 συνεχόμενες ημέρες.

Ημέρα	Θερμοκρασία (°C)
1	20.2
2	20.4
3	20.1
4	20.3

9. Από τα 14828 νοικοκυριά μιας πόλης, επελέγη ένα τυχαίο δείγμα 290 νοικοκυριών. Κάθε οικογένεια ερωτήθηκε αν ενοικίαζε ή όχι το σπίτι στο οποίο έμενε και αν είχε αυτόνομη θέρμανση ή όχι. Τα αποτελέσματα συνοψίζονται στον πίνακα που ακολουθεί

	Ιδιοκτήτες	Ενοικιαστές
Αυτόνομη θέρμανση	6	34
Κεντρική θέρμανση	141	109

Μπορεί να υποστηριχθεί ο ισχυρισμός ότι τουλάχιστον 15% των ενοικιαζομένων κατοικιών έχουν αυτόνομη θέρμανση;

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

10. Μία μηχανή συσκευασίας λειτουργεί ικανοποιητικά όταν κατά την διαδικασία συσκευασίας του προϊόντος, το ποσοστό των ελαττωματικών συσκευασιών είναι το πολύ 1%.

i) Διατυπώστε το ζεύγος υποθέσεων (μηδενική, εναλλακτική), μέσω του ελέγχου των οποίων θεωρείτε ότι θα είσαστε σε θέση να συμπεράνετε αν είναι εύλογος ο ισχυρισμός ότι η μηχανή λειτουργεί ικανοποιητικά.

ii) Ένας στατιστικός ελέγχει την λειτουργία της μηχανής εξετάζοντας ανά τακτά χρονικά διαστήματα 1000 διαδοχικές συσκευασίες, καθώς αυτές βγαίνουν από τη μηχανή. Πόσες ελαττωματικές συσκευασίες πρέπει να παρατηρηθούν σε 1000 τεμάχια για να είναι σε θέση να συμπεράνει ότι η μηχανή δεν λειτουργεί ικανοποιητικά (σε επίπεδο σημαντικότητας α);

α) Τουλάχιστον $10 + z_{\alpha/2} \sqrt{9.9}$

γ) Τουλάχιστον $10 + z_{1-\alpha/2} \sqrt{9.9}$

β) Τουλάχιστον $10 + z_\alpha \sqrt{9.9}$

δ) Τουλάχιστον $10 + z_{1-\alpha} \sqrt{9.9}$

Όπου z_p το p -ποσοστιαίο σημείο της τυποποιημένης κανονικής κατανομής.

(Οικ. Παν/μιο Αθηνών – Εξέτ. Δεκ. 2000)

11. Οι ετήσιοι αριθμοί των θανατηφόρων τροχαίων ατυχημάτων που κατεγράφησαν παγκοσμίως για τα έτη 1976-1985 είναι (Gelman et al. (1995))

24 25 31 31 22 21 26 20 16 22.

Να ελεγχθεί αν μπορεί να υποστηριχθεί ο ισχυρισμός ότι οι αριθμοί των θανατηφόρων τροχαίων ατυχημάτων τείνουν να μειώνονται.

(Οικ. Παν/μιο Αθηνών – Εξέτ. Δεκ. 2000)

12. Από 16 αυτοκίνητα που πέρασαν από έλεγχο μηχανολογικής κατάστασης σε κάποια περιοχή, 6 κρίθηκαν ανασφαλή. Να ελεγχθεί η υπόθεση ότι το ποσοστό των αυτοκινήτων της περιοχής αυτής που είναι ανασφαλή δεν υπερβαίνει το 10%. Να εκτιμηθεί το ποσοστό των αναφορών αυτοκινήτων της περιοχής με ένα 90% διάστημα εμπιστοσύνης.

13. Οι κάτοικοι μιας περιοχής ανέφεραν στο δημοτικό συμβούλιο της περιοχής ότι τουλάχιστον 60% των κατοίκων της περιοχής είναι υπέρ της δημιουργίας ενός ψυχαγωγικού κέντρου. Το δημοτικό συμβούλιο επέλεξε ένα τυχαίο δείγμα 100 κατοίκων. Από αυτούς 48 υποστήριξαν την δημιουργία του κέντρου. Είναι βάσιμος ο ισχυρισμός;

14. Ποιο πρέπει να είναι το μέγεθος ενός τυχαίου δείγματος αν επιδιώκεται να περιέχει τον μισό πληθυσμό τουλάχιστον με πιθανότητα 95%.

15. 7 από 22 πελάτες ενός παντοπωλείου δήλωσαν ότι προτιμούν ένα είδος τυριού, 12 ένα άλλο και 3 δεν είχαν ιδιαίτερη προτίμηση στο ένα ή στο άλλο. Υπάρχει στατιστικά σημαντική διαφορά στις προτιμήσεις των πελατών;

16. Ένα ζευγάρι που παίζει μπόουλινγκ (bowling) κατέγραψε τις επιδόσεις του (σκορ) σε 10 τυχαία επιλεγμένους γύρους παιχνιδιού:

άνδρας (X_i):	147	158	131	142	183	151	196	129	155	158
γυναίκα (Y_i):	122	128	125	123	115	120	108	143	124	123

Μπορεί να υποστηριχθεί ο ισχυρισμός ότι υπάρχει συσχέτιση μεταξύ των επιδόσεων του ζευγαριού;

17. Τα στοιχεία του πίνακα που ακολουθεί παριστάνουν την ποσοστιαία μεταβολή στις τιμές των 15 μετοχών που παρουσιάζουν την μεγαλύτερη κίνηση (σε όγκο συναλλαγών) κατά το κλείσιμο του Χρηματιστηρίου της Νέας Υόρκης κατά την διάρκεια της πρώτης εβδομάδας του 1991.

Εταιρεία	Ποσοστιαία μεταβολή στην τιμή
Mfrs Hanover	-2.9
Philip Morris	-2.9
Limited Inc.	6.3
General Electric	-4.8

Schering – Plough Cp.	-3.7
Bristol Myers Sq.	-4.7
Citicorp	4.9
Pepsico Inc.	-3.4
Wal – Mart Stores	-2.1
AT & T	0.0
GTE Cp.	-0.9
Coca Cola Co.	-3.0
General Motors Cp.	-4.3
Navistar Intl.	29.4
GaGulf	2.5

Πηγή: New York Times, 7 Ιανουαρίου 1991 σελ. D9

Με βάση τα παραπάνω στοιχεία, θα μπορούσε να υποστηριχθεί ο ισχυρισμός ότι οι τιμές των μετοχών παρουσίασαν πτώση κατά την διάρκεια της εβδομάδας αυτής; ($\alpha=0.05$)

18. Μία επιστημονική ομάδα καταγράφει κάθε μήνα το μέσο μηνιαίο ρυθμό ροής του νερού ενός μικρού ποταμού (σε κυβ. πόδια ανά δευτερόλεπτο). Συγκεντρώνονται τα δεδομένα των τελευταίων 24 μηνών για να ελεγχθεί το ακόλουθο ζεύγος υποθέσεων:

H_0 : Ο ρυθμός ροής δεν μειώνεται

H_1 : Ο ρυθμός ροής μειώνεται

Τα δεδομένα, με χρονολογική σειρά από την παλαιότερη προς την πιο πρόσφατη μέτρηση είναι:

14.6	12.2	104	220	110	86.0	92.8	74.4	75.4
51.7	29.3	16.0	14.2	10.5	123	190	138	98.1
88.1	80.0	75.6	48.8	27.1	15.7			

Να ελεγχθούν οι παραπάνω υποθέσεις σε επίπεδο σημαντικότητας $\alpha=0.1$.

19. 20 ανεξάρτητες παρατηρήσεις πάνω σε μία τυχαία μεταβλητή X με άγνωστη συνάρτηση κατανομής $F(x)$ έδειξαν τις εξής τιμές:

142	134	98	119	131
103	154	122	93	137
86	119	161	144	158
165	81	117	128	103

Να κατασκευασθεί ένα 95% διάστημα εμπιστοσύνης για την $F(100)$.

(Παν/μιο Κρήτης, Ηράκλειο – Εξετ. Δεκ. 1984)

20. Οι χρόνοι ζωής (σε ώρες) 16 λυχνιών ραδιοφώνου που διαλέχθηκαν τυχαία από την παραγωγή ενός εργοστασίου είναι οι εξής:

$X^{(1)} = 46.9$	$X^{(5)} = 56.8$	$X^{(9)} = 63.3$	$X^{(13)} = 97.1$
$X^{(2)} = 47.2$	$X^{(6)} = 59.2$	$X^{(10)} = 63.4$	$X^{(14)} = 67.7$
$X^{(3)} = 49.1$	$X^{(7)} = 59.9$	$X^{(11)} = 63.7$	$X^{(15)} = 73.3$
$X^{(4)} = 56.5$	$X^{(8)} = 63.2$	$X^{(12)} = 64.1$	$X^{(16)} = 78.5$

Να κατασκευασθεί ένα διάστημα εμπιστοσύνης για το 0.75 ποσοστιαίο σημείο της κατανομής του χρόνου ζωής των λυχνιών που κατασκευάζει το εργοστάσιο με συντελεστή εμπιστοσύνης $\approx 40\%$.

(Παν/μιο Κρήτης, Ηράκλειο – Εξετ. Δεκ. 1984)

21. Το μήκος ενός βιομηχανικού προϊόντος διαφέρει από κομμάτι σε κομμάτι. Για να προσδιορισθεί το κατάλληλο μήκος που πρέπει να έχει το κουτί που θα χρησιμοποιηθεί για την αποστολή του προϊόντος αυτού, θέλουμε να προσδιορίσουμε όρια μέσα στα οποία να βρίσκεται το 80% ή περισσότερο των παραγμένων μονάδων του προϊόντος με πιθανότητα τουλάχιστον ίση με 90%. Ποιο είναι το μέγεθος n του τ.δ.

που θα απαιτηθεί ώστε οι παρατηρήσεις $X^{(1)}$ και $X^{(n)}$ ν' αποτελέσουν τα όρια αυτά;

(Παν/μιο Κρήτης, Ηράκλειο – Εξ. Δεκ. 1984)

22. Με βάση τις παρακάτω ανεξάρτητες παρατηρήσεις πάνω στην τ.μ. X , να εκτιμηθεί η $P(X \leq 98)$ με ένα 95% διάστημα εμπιστοσύνης.

142	103	86	165	119	93	144	128	134	154
119	81	131	137	158	103	98	122	161	117

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

23. Αποτελούν οι παρακάτω ετήσιες παρατηρήσεις πάνω στον μέσο αριθμό λ των θανατηφόρων τροχαίων ατυχημάτων ανά 100000 κατοίκους μιας περιοχής ένδειξη ανοδικής τάσης των τιμών του λ ;

17.3	17.9	18.4	18.1	18.3	19.6	18.6	19.2	17.7	20.0
19.0	18.8	19.3	20.2	19.9					

(Παν/μιο Κρήτης, Ηράκλειο - Εξ. Σεπτ. 1985)

24. Από 102 μωρά που γεννήθηκαν σε μία μαιευτική κλινική, 60 γεννήθηκαν νύχτα. Είναι βάσιμη η υπόθεση ότι περισσότερα μωρά γεννιούνται την νύχτα;

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

25. Δίνεται το διατεταγμένο τυχαίο δείγμα 81, 90, 92, 93, 98, 101, 111, 112, 114, 117, 126. Ποιο είναι το κρίσιμο επίπεδο για τον έλεγχο της υπόθεσης ότι το τ.δ. προέρχεται από μία κατανομή με 0.25-ποσοστιαίο σημείο μικρότερο του 105;

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

26. Με βάση τις παρακάτω παρατηρήσεις πάνω στο τυχαίο διάνυσμα (X, Y) να ελεγχθεί η υπόθεση $P(Y < X) = P(Y > X)$ εναντίον αμφίπλευρης εναλλακτικής.

(46.8, 49.5)	(48.8, 51.1)	(52.4, 49.1)	(49.5, 47.8)
(46.1, 47.7)	(48.1, 45.6)	(45.7, 46.7)	(44.8, 51.4)
(45.4, 49.2)	(47.5, 51.1)	(47.0, 49.4)	(48.1, 47.8)

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

27. Να κατασκευασθεί ένα διάστημα εμπιστοσύνης για την διάμεσο του πληθυσμού από τον οποίο προέρχεται το τυχαίο δείγμα

93 98 111 81 117 92 126 112 90 114 101

με συντελεστή εμπιστοσύνης περίπου ίσο με 95%.

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

28. Σε ένα εργοστάσιο παραγωγής στρατιωτικού υλικού, δοκιμάζεται ένας νέος τύπος τορπίλης. 20 τορπίλες του νέου τύπου εκτοξεύονται υπό τις ίδιες συνθήκες και από την ίδια απόσταση εναντίον ενός στόχου επιφανείας. Μόνο οι 15 από αυτές πετυχαίνουν το στόχο. Είναι εύλογο να ισχυρισθεί κανείς ότι η πιθανότητα προσβολής του στόχου από τον νέο τύπο τορπίλης είναι μεγαλύτερη από 80%;

Δίνεται ο πίνακας τιμών της αθροιστικής συνάρτησης κατανομής της Διωνυμικής κατανομής με $n=20$ και $p=0.8$.

29. Ένας κατασκευαστής υπολογίζει το μέσο κόστος κατασκευής του προϊόντος του (σε χιλ. δρχ. ανά τεμάχιο) για τους τελευταίους 44 μήνες και βρίσκει ότι είναι

13.65	13.41	13.53	13.23	13.58	13.43	13.73	13.4	13.7
13.58	13.8	13.4	13.63	13.69	13.92	13.68	13.72	13.42
13.66	13.98	13.81	13.60	13.32	13.45	13.27	13.26	13.28

13.29 13.1 13.09 13.36 13.4 13.35 13.53 13.66 13.1
 13.28 13.33 13.02 13.09 13.12 13.16 12.96 12.95.

Είναι εύλογο να θεωρηθεί (σε $\alpha \approx 0.05$) ότι το μέσο κόστος παρουσιάζει τάση;

30. Σε ένα πείραμα ψυχολογίας για το κατά πόσο ισχύει η θεωρία της ύπαρξης υποβολής, δείχνονται 20 ευθύγραμμο τμήματα σε δύο άτομα από τα οποία ζητείται να «εκτιμήσουν» με το μάτι το μήκος τους. Για κάθε ευθύγραμμο τμήμα, απαντάει πρώτα το άτομο Α και, στην συνέχεια, το άτομο Β, αφού ακούσει την εκτίμηση του Α. Στο άτομο Α, όμως, έχουν δοθεί οδηγίες από τους ερευνητές να υπερεκτιμήσει το μήκος των πρώτων 10 γραμμών και να υποεκτιμήσει το μήκος των τελευταίων 10. Τα σφάλματα των εκτιμήσεων (εκτίμηση – πραγματικό μήκος) των δύο ατόμων συνοψίζονται στον πίνακα που ακολουθεί:

Ευθύγραμμο Τμήμα	1	2	3	4	5	6	7	8	9	10
A	0.3	1.1	0.9	0.6	1	1.3	0.8	1.6	1.2	0.8
B	-0.1	0.6	1	0.7	0.2	0.9	-0.1	0.2	0	0.5
Ευθύγραμμο Τμήμα	11	12	13	14	15	16	17	18	19	20
A	-1.3	-1.1	-1.3	-0.7	-1.4	-1.1	-0.8	-0.5	-1.2	-1
B	-0.6	-1.2	-1	-0.7	-1	-0.1	-0.5	0	-0.4	-0.3

Σε επίπεδο σημαντικότητας $\alpha \approx 0.05$, μπορεί να θεωρηθεί ότι τα αποτελέσματα παρέχουν ενδείξεις υπέρ της θεωρίας για την ύπαρξη υποβολής;

31. Ο παρακάτω πίνακας αναφέρεται στις αντιδράσεις ενός τυχαίου δείγματος 10 ασθενών σε δύο φάρμακα X και Y.

Ασθενής	1	2	3	4	5	6	7	8	9	10
X	7	-1.6	-0.2	-1.2	-0.1	3.4	3.7	0.8	0	2
Y	1.9	0.8	1.1	0.1	-0.1	4.4	5.5	1.6	4.6	3.4

Να ελεγχθεί αν υπάρχει θετική γραμμική συσχέτιση μεταξύ των αντιδράσεων ενός ασθενούς στα δύο φάρμακα.

(Παν/μιο Κρήτης, Ηράκλειο – Εξετ. Προόδου Δεκ. 1984)

32. Ένα τυχαίο δείγμα 6 οδηγών που συνελήφθησαν να οδηγούν μεθυσμένοι υποβλήθηκε σε δύο διαφορετικά τεστ για τον καθορισμό του επιπέδου του οινοπνεύματος στο αίμα τους με τα εξής αποτελέσματα.

Οδηγός	1	2	3	4	5	6
Τεστ 1	0.170	0.190	0.200	0.183	0.187	0.178
Τεστ 2	0.197	0.178	0.150	0.176	0.205	0.153

Θα μπορούσε να ισχυρισθεί κανείς ότι τα δύο τεστ δεν δίνουν συνεπή αποτελέσματα και επομένως ένας οδηγός που θα καταδικαζόταν με βάση το ένα τεστ θα μπορούσε να αθωωθεί με βάση το άλλο;

33. Σύμφωνα με τους νόμους κληρονομικότητας του Mendel, η διασταύρωση δύο ορισμένων ειδών φυτών μπορεί να οδηγήσει σε φυτά «νάνους» με πιθανότητα 25%. Αν υποθέσουμε ότι, σε ένα σχετικό πείραμα, η διασταύρωση οδήγησε σε «απόγονους» από τους οποίους 6 ήταν «νάνου» και 11 κανονικοί, να εκτιμήσετε την πιθανότητα αυτή με βάση ένα 95% διάστημα εμπιστοσύνης. Να χρησιμοποιήσετε το διάστημα αυτό για να κρίνετε αν η υπόθεση του νόμου του Mendel είναι εύλογη για την περίπτωση αυτού του είδους των φυτών.

34. Η ποσότητα αναψυκτικού που δίνει σε κύπελλα μια αυτόματη μηχανή δεν είναι σταθερή. Ο κατασκευαστής θέλει να είναι σε θέση να εγγυηθεί ένα κατώτερο όριο για την πωλούμενη ποσότητα πριν διαθέσει το μηχάνημα στην αγορά. Πόσο μεγάλο δείγμα κυπέλλων χρειάζεται για να είναι τουλάχιστον 95% σίγουρος ότι το 95% τουλάχιστον των κυπέλλων περιέχει ποσότητα μεγαλύτερη από την 4η κατά σειρά μεγέθους παρατήρηση του δείγματος;

35. 100 άτομα ερωτήθηκαν αν στηρίζουν ή όχι την εξωτερική πολιτική μιας κυβέρνησης, 2 μήνες και 12 μήνες μετά την ορκωμοσία της, με τα εξής αποτελέσματα.

	Μετά από 2 μήνες	
Μετά από 12 μήνες	Στήριζαν	Δεν στήριζαν
Στηρίζουν	20	40
Δεν στηρίζουν	10	30

Μπορεί να θεωρηθεί ότι σημειώθηκε στατιστικά σημαντική αλλαγή στο ποσοστό των ατόμων που στηρίζουν την εξωτερική πολιτική της συγκεκριμένης κυβέρνησης;

36. Από κάθε ένα από δύο πληθυσμούς διαλέχθηκε ένα τυχαίο δείγμα 135 ατόμων για να σφυγμομετρηθεί η αντίδρασή τους σε ένα νομοσχέδιο. Αν ο αριθμός των ατόμων που δεν υποστηρίζουν το νομοσχέδιο είναι 43 για το πρώτο τυχαίο δείγμα και 37 για το δεύτερο, να ελεγχθεί αν υπάρχει στατιστικά σημαντική διαφορά στα ποσοστά

των υποστηρικτών ή μη του νομοσχεδίου στους δύο πληθυσμούς.
($\alpha=5\%$)

(Παν/μιο Κρήτης, Ηράκλειο – Εξετ. Προόδου Δεκ. 1984)

37. Given the following random sample of all scores

81 90 92 93 98 101 111 112 114 117 126

- a) What is the critical level ($\hat{\alpha}$) for a test of the hypothesis that these scores come from a population with a 0.25-quantile of less than 105?
- b) Construct an approximate 95% confidence interval for the median of the population from which these scores are a random sample. Give the exact level of confidence for your interval.

(University of Iowa, Iowa City – Εξετ. Φεβρ. 1983)

38. When bowling it is often possible to score well on the first game and then bowl poorly on the second game or vice versa. The following pairs of numbers represent the scores of the first and second game bowled by the same person on six consecutive Tuesday evenings. Is there sufficient evidence to indicate that the two paired samples of scores are negatively correlated? ($\alpha=10\%$)

Game 1	170	190	200	183	187	178
Game 2	197	178	150	176	205	153

State your exact level of significance.

(University of Iowa, Iowa City – Εξετ. Φεβρ. 1983)

39. The amount of cola dispensed by a vending machine varies. The manufacturer wants to guarantee a lower limit to the amount dispensed,

before shipping the machine. How large a sample is needed so that the manufacturer can be at least 95% confident that 99% or more of the dispensed amounts are above $X^{(4)}$?

(University of Missouri, Columbia – Εξέτ. Μαΐου 1981)