

6. Έλεγχοι Υποθέσεων

Σε αρκετές εφαρμογές παρουσιάζεται η ανάγκη λήψης αποφάσεων σχετικών με την κατανομή ενός πληθυσμού. Πιο συγκεκριμένα, σε πολλές περιπτώσεις πρέπει, βάσει ενός τ.δ. $X_1, X_2, \dots, X_n$ από έναν πληθυσμό με κατανομή $F(x; \theta)$, να αποφασίσουμε αν ευσταθεί ή όχι μία πληροφορία (υπόθεση) σχετική με την κατανομή F ή τις παραμέτρους θ .

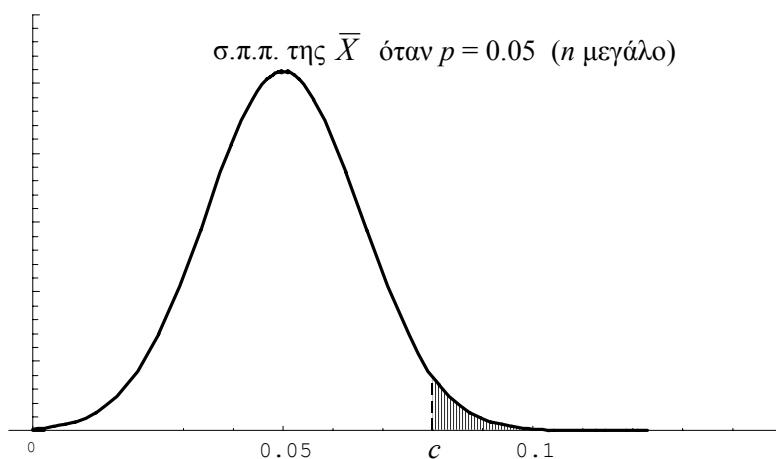
Παράδειγμα 6.1. Έστω ότι ένα εργοστάσιο παράγει κάποιες ηλεκτρονικές συσκευές (π.χ. μικροεπεξεργαστές). Κατά τακτά χρονικά διαστήματα (π.χ. ανά ώρα) γίνεται έλεγχος της ποιότητας και των n συσκευών της ωριαίας παραγωγής. Κατά τον έλεγχο αυτό απομακρύνονται όλες οι τυχόν ελαττωματικές συσκευές. Η παραγωγική διαδικασία θεωρείται ότι βρίσκεται μέσα στις προδιαγραφές της αν η πιθανότητα παραγωγής ελαττωματικής συσκευής είναι ίση (ή το πολύ) 5%. Σε περίπτωση που αυξηθεί αυτή η πιθανότητα θεωρείται ότι υπάρχει κάποιο πρόβλημα, σταματά η παραγωγή, και αναζητείται ο λόγος της ανωμαλίας.

Στο παράδειγμα αυτό θα πρέπει να κατασκευάσουμε έναν έλεγχο με βάση τον οποίο θα κρίνουμε αν για την τελευταία ώρα ισχύει για την (άγνωστη) πιθανότητα p παραγωγής ελαττωματικής μονάδας η υπόθεση ότι

$p = 5\%$ (ή $\leq 5\%$), οπότε συνεχίζεται η παραγωγή,

ή $p > 5\%$, οπότε διακόπτεται η παραγωγή.

Είναι λογικό ο έλεγχος αυτός να βασίζεται στον αριθμό των ελαττωματικών μονάδων που βρέθηκαν ανάμεσα στις n της ωριαίας παραγωγής. Αν θέσουμε $X_i=1$ ή 0 ανάλογα με το αν η i -μονάδα βρέθηκε ελαττωματική ή όχι, $i=1,2,\dots,n$ τότε το δειγματικό ποσοστό $\bar{X}$ αποτελεί κατά τα γνωστά μία εκτιμήτρια της πιθανότητας p παραγωγής ελαττωματικής μονάδας. Σύμφωνα με τα όσα γνωρίζουμε, αν το p είναι ίσο του 5% τότε αναμένουμε το δειγματικό ποσοστό $\bar{X}$ να παίρνει τιμές «κοντά» και «γύρω» από το 5%. Συνεπώς, δεδομένου ότι $p = 5\%$, είναι «απίθανο» να βρεθεί ένα $\bar{X}$ «σημαντικά» μεγαλύτερο του 5% (π.χ. να βρεθεί $\bar{X} > 10\%$ ή γενικότερα $\bar{X} > c$).



Στην περίπτωση λοιπόν που συμβεί κάτι τέτοιο ($\bar{X} > c$) είναι λογικό να θεωρήσουμε ότι το $\bar{X}$ βρέθηκε τόσο μεγάλο διότι στην πραγματικότητα δεν ισχύει ότι $p = 5\%$ αλλά $p > 5\%$. Άρα σε αυτή την περίπτωση θα πρέπει να διακόψουμε την παραγωγή (απορρίπτουμε ότι $p = 5\%$). Επομένως, κατά κάποιον τρόπο κατασκευάσαμε έναν έλεγχο της υπόθεσης $p = 5\%$ έναντι της $p > 5\%$ σύμφωνα με τον οποίον:

- αν $\bar{X} > c$ απορρίπτουμε ότι $p = 5\%$
- αν $\bar{X} \leq c$ δεχόμαστε ότι $p = 5\%$.

Το ερώτημα που τίθεται τώρα είναι: ποια θα πρέπει να είναι αυτή η τιμή c για από την οποία, δεδομένου ότι $p = 5\%$, θεωρείται «απίθανο» να συμβεί $\bar{X} > c$; Είναι φανερό ότι όσο μεγαλύτερο είναι το c , τόσο πιο απίθανο γίνεται το ενδεχόμενο $\bar{X} > c$. Από την άλλη μεριά όμως, αν πάρουμε c υπερβολικά μεγάλο (π.χ. $c = 20\%$) τότε εμφανίζεται ο κίνδυνος να ισχύει π.χ. ότι $p = 10\% > 5\%$ και εμείς να βρούμε, πράγμα σχεδόν βέβαιο, ότι $\bar{X} \leq 20\%$ και επομένως να δεχτούμε την αρχική υπόθεση ότι $p = 5\%$! Συνεπώς θα πρέπει να βρεθεί το βέλτιστο c κάτω από κάποιες συγκεκριμένες προϋποθέσεις. Ο καθορισμός αυτών των προϋποθέσεων καθώς και η εύρεση κατάλληλου ελέγχου αποτελεί αντικείμενο της θεωρίας των στατιστικών ελέγχων υποθέσεων που θα περιγράψουμε στη συνέχεια.

Το γενικό πρόβλημα - ορισμοί

Ας δούμε το παραπάνω πρόβλημα στη γενικότερή του μορφή. Έστω $X_1, X_2, \dots, X_n$ ένα τ.δ. από έναν πληθυσμό με κατανομή $F(x; \theta)$ (στο παράδειγμα είχαμε $X_1, X_2, \dots, X_n \sim B(1, p)$). Επιθυμούμε να ελέγξουμε την υπόθεση $\theta \in \Theta_0$ έναντι της $\theta \in \Theta_1$ όπου Θ_0, Θ_1 είναι υποσύνολα του παραμετρικού χώρου Θ (σύνολο επιτρεπτών τιμών της παραμέτρου θ) ενώ φυσικά $\Theta_0 \cap \Theta_1 = \emptyset$ (τα Θ_0, Θ_1 είναι ξένα). Η βασική υπόθεση $\theta \in \Theta_0$ θα καλείται *μηδενική* υπόθεση και θα συμβολίζεται με H_0 ενώ η ενάντια $\theta \in \Theta_1$ θα καλείται *εναλλακτική* υπόθεση και θα συμβολίζεται με H_1 . Συνοπτικά θα έχουμε:

H_0 : $\theta \in \Theta_0$, μηδενική (ή βασική) υπόθεση,

H_1 : $\theta \in \Theta_1$, εναλλακτική υπόθεση.

Στο παραπάνω παράδειγμα έχουμε τη μηδενική υπόθεση H_0 : $p \in \{0.05\}$ με εναλλακτική την H_1 : $p \in (0.05, 1]$. Μία υπόθεση H_i : $\theta \in \Theta_i$ θα καλείται *απλή* ή *σύνθετη* ανάλογα με το αν σύνολο Θ_i είναι μονοσύνολο ($\Theta_i = \{\theta_0\}$) ή όχι. Για παράδειγμα η υπόθεση H_0 : $p=0.05$ είναι απλή ενώ η H_1 : $p > 0.05$ σύνθετη. Επίσης μία υπόθεση θα καλείται *μονόπλευρη* αν είναι της μορφής $H_i: \theta > \theta_0$ ή $H_i: \theta < \theta_0$ ενώ θα καλείται *αμφίπλευρη* αν είναι της μορφής $H_i: \theta < \theta_0$ ή $\theta > \theta_1$ (δηλ. $\theta \in (-\infty, \theta_0) \cup (\theta_1, \infty)$)).

Στη συνέχεια, βάσει του τ.δ. $X_1, X_2, \dots, X_n$, κατασκευάζουμε μία διαδικασία ελέγχου της παραπάνω υπόθεσης. Συγκεκριμένα, χωρίζουμε το δειγματοληπτικό χώρο Ω (το σύνολο των δυνατών τιμών του δείγματος) σε δύο ξένα υποσύνολα A και K ($A \cup K = \Omega$) έτσι ώστε,

- αν $(X_1, X_2, \dots, X_n) \in K$, απορρίπτουμε την H_0 : $\theta \in \Theta_0$

- αν $(X_1, X_2, \dots, X_n) \in A$, δεχόμαστε την H_0 : $\theta \in \Theta_0$

(στο παραπάνω παράδειγμα, $K = \{(x_1, x_2, \dots, x_n) \in \{0, 1\}^n : \frac{1}{n} \sum_{i=1}^n x_i > c\}$, $A = \{(x_1, x_2, \dots, x_n) \in \{0, 1\}^n : \frac{1}{n} \sum_{i=1}^n x_i \leq c\}$, ή με απλούστερη γραφή, $K: \bar{X} > c$, $A: \bar{X} \leq c$). Η περιοχή K καλείται *κρίσιμη περιοχή* ή *περιοχή απόρριψης* της H_0 ενώ η περιοχή A καλείται *περιοχή αποδοχής* της μηδενικής υπόθεσης H_0 .

Παρατηρούμε ότι, ανάλογα με την απόφαση που θα πάρουμε, είναι πιθανό να κάνουμε τα εξής λάθη:

- *σφάλμα τύπου I*: απόρριψη της H_0 ενώ ισχύει η H_0
- *σφάλμα τύπου II*: αποδοχή της H_0 ενώ ισχύει η H_1 .

Η πιθανότητα να κάνουμε σφάλμα τύπου *I* και *II* είναι αντίστοιχα

$$P(\text{σφάλμα τύπου I}) = P((X_1, X_2, \dots, X_n) \in K / H_0)$$

$$P(\text{σφάλμα τύπου II}) = P((X_1, X_2, \dots, X_n) \in A / H_1).$$

Η ποσότητα

$$\pi(\theta) = 1 - P(\text{σφάλμα τύπου II})$$

θα καλείται *ισχύς του ελέγχου* και όπως θα δούμε εξαρτάται γενικά από την παράμετρο θ . Είναι φανερό ότι η ισχύς του ελέγχου ισούται με την πιθανότητα ορθής απόρριψης της H_0 .

Αυτό που απομένει τώρα να κάνουμε είναι να καθορίσουμε την περιοχή K (και από αυτήν και την περιοχή A αφού $A = \Omega - K$). Δηλαδή θα πρέπει να καθορίσουμε πότε, βάσει ενός τ.δ., θα απορρίπτουμε την H_0 και πότε όχι. Παρατηρούμε ότι οι $P(I)$ και $P(II)$ εξαρτώνται από την εκλογή του K . Άρα, θα ήταν λογικό να αναζητήσουμε περιοχή K απόρριψης της H_0 η οποία να ελαχιστοποιεί τις πιθανότητες των σφαλμάτων I και II . Δυστυχώς όμως διαπιστώνεται ότι στις περισσότερες περιπτώσεις μία εκλογή του K που μειώνει την $P(I)$, παράλληλα οδηγεί στην αύξηση της $P(II)$.

Επειδή συνήθως μας ενδιαφέρει περισσότερο **να μην απορρίψουμε λανθασμένα**, ζητάμε η κρίσιμη περιοχή K να ληφθεί έτσι ώστε η $P(I)$ να είναι μικρότερη από ένα προκαθορισμένο α (*επίπεδο σημαντικότητας*) και παράλληλα η $P(II)$ να είναι η ελάχιστη δυνατή (ή ισοδύναμα η ισχύς του ελέγχου να είναι η μεγαλύτερη δυνατή). Δηλαδή θα διαλέγουμε K έτσι ώστε:

$$P(\text{σφάλμα τύπου I}) = P((X_1, X_2, \dots, X_n) \in K / H_0) \leq \alpha,$$

$$P(\text{σφάλμα τύπου II}) = P((X_1, X_2, \dots, X_n) \notin K / H_1) : \text{ελάχιστο δυνατό.}$$

Συνήθως επιλέγουμε $\alpha=0.05$ ή $\alpha=0.01$. Αλλά ας το δούμε όλα τα παραπάνω στην πράξη χρησιμοποιώντας το αρχικό παράδειγμα.

Παράδειγμα (συνέχεια παραδείγματος 6.1) Στη περίπτωση που εξετάσαμε παραπάνω στο Παράδειγμα 6.1. δίνονταν ένα τ.δ. $X_1, X_2, \dots, X_n$ από την $B(1, p)$ και το ζητούμενο ήταν να πραγματοποιηθεί ο έλεγχος των υποθέσεων

$$H_0: p = 0.05$$

$$H_1: p = p_1 > 0.05,$$

ώστε να ληφθεί μία απόφαση σχετικά με το αν πρέπει να συνεχίσουμε ή να διακόψουμε την παραγωγή. Η περιοχή απόρριψης K της H_0 και η περιοχή αποδοχής A της H_0 είναι

$$K: (X_1, \dots, X_n): \bar{X} > c$$

$$A: (X_1, \dots, X_n): \bar{X} \leq c$$

Οι πιθανότητες των σφαλμάτων τύπου I και II καθώς και η ισχύς του ελέγχου θα είναι αντίστοιχα

$$P(I) = P(\text{απόρριψη της } H_0 / H_0) = P(\bar{X} > c / p=0.05)$$

$$P(II) = P(\text{αποδοχή της } H_0 / H_1) = P(\bar{X} \leq c / p = p_1 > 0.05)$$

$$\pi(p_1) = 1 - P(\bar{X} \leq c / p = p_1 > 0.05) = P(\bar{X} > c / p = p_1 > 0.05).$$

Για ευκολία ας υποθέσουμε ότι αναζητούμε την κρίσιμη περιοχή K του ελέγχου της υποθέσεως $H_0: p=0.05$ έναντι της $H_1: p=0.1$ (δηλ. έστω $p_1 = 0.1$). Επειδή $X_i \sim B(1, p)$ θα ισχύει (χρησιμοποιώντας και το Κ.Ο.Θ.) ότι, για μεγάλα n ,

$$P(\bar{X} > c) = P\left(\frac{\bar{X}_i - p}{\sqrt{p(1-p)/n}} > \frac{c - p}{\sqrt{p(1-p)/n}}\right) = 1 - \Phi\left(\frac{c - p}{\sqrt{p(1-p)/n}}\right)$$

Συνεπώς,

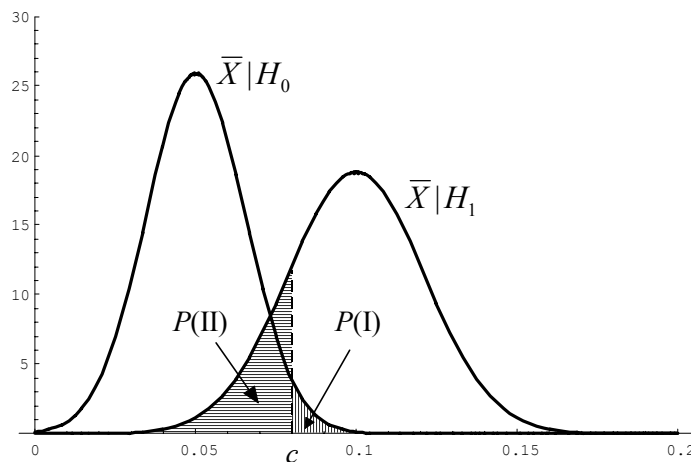
$$P(I) = P(\bar{X} > c / p = 0.05) = 1 - \Phi\left(\frac{c - 0.05}{\sqrt{0.05(1-0.05)/n}}\right)$$

$$P(II) = P(\bar{X} \leq c/p = 0.1) = \Phi\left(\frac{c - 0.1}{\sqrt{0.1(1-0.1)/n}}\right).$$

Παρατηρούμε ότι για να μειωθεί η $P(I)$ θα πρέπει να αυξήσουμε το c , δηλαδή να πάρουμε μικρότερη περιοχή απόρριψης K : $\bar{X} \in (c, \infty)$ (η συνάρτηση Φ ως συνάρτηση κατανομής είναι αύξουσα). Κάτι τέτοιο όμως δυστυχώς αυξάνει την $P(II)$. Αν υποθέσουμε π.χ. ότι $n = 200$ θα έχουμε ότι για διάφορες τιμές του c οι $P(I)$ και $P(II)$ θα είναι

c	$P(I)$	$P(II)$
0.04	0.742	0.002
0.05	0.5	0.009
0.06	0.258	0.030
0.07	0.097	0.078
0.08	0.026	0.173
0.09	0.005	0.319
0.1	0.0006	0.5
0.11	0.00005	0.681

Το δειγματικό ποσοστό $\bar{X}$ θα ακολουθεί προσεγγιστικά $N(p, p(1-p)/n)$ οπότε υπό την H_0 θα ισχύει $\bar{X} \sim N(0.05, 0.015411^2)$ ενώ υπό την H_1 θα ισχύει $\bar{X} \sim N(0.10, 0.0212132^2)$. Αν δηλαδή η πιθανότητα παραγωγής ελαττωματικής συσκευής είναι $p = 0.05$ τότε το δειγματικό ποσοστό $\bar{X}$ που θα παίρνουμε ανά ώρα (από τις $n=200$ μονάδες που παρήχθησαν) θα κατανέμεται σύμφωνα με την $N(0.05, 0.015411^2)$, ενώ αν η πιθανότητα παραγωγής ελαττωματικής συσκευής είναι $p = 0.1$ τότε το δειγματικό ποσοστό $\bar{X}$ θα κατανέμεται σύμφωνα με την $N(0.1, 0.0212132^2)$. Στο επόμενο σχήμα βλέπουμε τις κατανομές της στατιστικής συνάρτησης $\bar{X}$ κάτω από τις δύο υποθέσεις H_0 και H_1 . Αν $K: \bar{X} > c$ τότε η $P(I)$ είναι ίση με το εμβαδόν της γραμμοσκιασμένης με κάθετες γραμμές περιοχής, ενώ η $P(II)$ είναι ίση με το εμβαδόν της γραμμοσκιασμένης με οριζόντιες γραμμές περιοχής.



Επαληθεύεται εύκολα και από το σχήμα ότι αν πάρουμε μεγαλύτερο c τότε ναι μεν μειώνεται η πιθανότητα να απορρίψουμε την H_0 λανθασμένα αλλά παράλληλα αυξάνεται η πιθανότητα να δεχτούμε την H_0 λανθασμένα.

Όπως έχουμε αναφέρει και παραπάνω, μας ενδιαφέρει περισσότερο να μην απορρίψουμε κατά λάθος και συνεπώς ζητάμε η κρίσιμη περιοχή K να ληφθεί έτσι ώστε:

$$P(\text{σφάλμα τύπου I}) = P((X_1, X_2, \dots, X_n) \in K / H_0) \leq \alpha,$$

$$P(\text{σφάλμα τύπου II}) = P((X_1, X_2, \dots, X_n) \notin K / H_1) : \text{ελάχιστο δυνατό},$$

όπου α είναι το προκαθορισμένο επίπεδο σημαντικότητας. Επομένως στη συγκεκριμένη περίπτωση, αναζητούμε K έτσι ώστε,

$$P(I) = P(\bar{X} > c/p = 0.05) = 1 - \Phi\left(\frac{c - 0.05}{\sqrt{0.05(1-0.05)/n}}\right) \leq \alpha$$

$$P(II) = P(\bar{X} \leq c/p = 0.1) = \Phi\left(\frac{c - 0.1}{\sqrt{0.1(1-0.1)/n}}\right) : \text{ελάχιστο δυνατό.}$$

Λύνοντας την πρώτη ανίσωση ως προς c (για να καθορίσουμε την K : $\bar{X} > c$) θα έχουμε ότι

$$\Phi\left(\frac{c - 0.05}{\sqrt{0.05(1-0.05)/n}}\right) \geq 1 - \alpha \Rightarrow \frac{c - 0.05}{\sqrt{0.05(1-0.05)/n}} \geq \Phi^{-1}(1 - \alpha) = Z_\alpha \Rightarrow c \geq 0.05 + \sqrt{\frac{0.05(1-0.05)}{n}} Z_\alpha,$$

και επειδή η $P(II)$ μικραίνει όσο το c μικραίνει θα πάρουμε το μικρότερο επιτρεπτό c από την παραπάνω ανίσωση (ώστε να έχουμε $P(II)$: ελάχιστο δυνατό), δηλαδή

$$c = 0.05 + \sqrt{\frac{0.05(1-0.05)}{n}} Z_\alpha.$$

Αν π.χ. $n = 200$ και θέσουμε επίπεδο σημαντικότητας $\alpha = 1\%$ τότε βρίσκουμε ότι

$$c = 0.05 + \sqrt{\frac{0.05(1-0.05)}{200}} 2.33 = 0.0859.$$

Άρα τελικά ο έλεγχος ποιότητας θα γίνεται ως εξής: κάθε μία ώρα υπολογίζουμε το ποσοστό $\bar{X}$ των ελαττωματικών μονάδων ανάμεσα στις $n = 200$ παραγόμενες και αν

- $\bar{X} > 8.59\%$ τότε απορρίπτουμε ότι $H_0: p = 0.05$ και διακόπτουμε την παραγωγή,
- $\bar{X} \leq 8.59\%$ τότε δεχόμαστε ότι $H_0: p = 0.05$ και η παραγωγή συνεχίζεται κανονικά.

Αναμένεται σύμφωνα με τα παραπάνω ότι στο 1% ($=\alpha$) των ελέγχων θα έχουμε διακόψει άσκοπα την παραγωγή ($P(\text{λανθασμένης απόρριψης της } H_0) = P(I) = \alpha$). Επίσης η πιθανότητα σφάλματος τύπου II για K : $\bar{X} > 8.59\%$ θα είναι

$$P(II) = \Phi\left(\frac{0.0859 - 0.1}{\sqrt{0.1(1-0.1)/200}}\right) = \Phi(-0.66) = 1 - \Phi(0.66) = 1 - 0.7454 = 25.46\%,$$

οπότε αναμένεται ότι στο 25.46% των ελέγχων θα έχουμε συνεχίσει την παραγωγή ενώ έχει αυξηθεί η πιθανότητα παραγωγής ελαττωματικών ($P(\text{λανθασμένης αποδοχής της } H_0) = P(II)$). Σε αυτό το παράδειγμα διαπιστώνουμε ότι η διοίκηση του εργοστασίου ενδιαφέρεται πολύ περισσότερο να αποφύγει μία άσκοπη διακοπή της παραγωγής με κίνδυνο να ανακαλύψει καθυστερημένα κάποιο πρόβλημα στην παραγωγή, παρά να ανακαλύπτει αμέσως οποιοδήποτε πρόβλημα με κίνδυνο όμως να σταματά πολύ συχνά άσκοπα την παραγωγή.

Αν τέλος π.χ. είχαμε θέσει $\alpha = 0.05$ τότε βρίσκουμε όμοια ότι

$$c = 0.05 + \sqrt{\frac{0.05(1-0.05)}{200}} 1.645 = 0.0753$$

και άρα θα απορρίπτουμε όταν $\bar{X} > 7.53\%$. Σε αυτή την περίπτωση, $P(I) = \alpha = 5\%$ και

$$P(II) = \Phi\left(\frac{0.0753 - 0.1}{\sqrt{0.1(1-0.1)/200}}\right) = \Phi(-1.162) = 1 - \Phi(1.162) = 1 - 0.8770 = 12.3\%.$$

Συνοψίζοντας λοιπόν, θα έχουμε τα εξής. Έστω $X_1, X_2, \dots, X_n$ ένα τ.δ. από έναν πληθυσμό με κατανομή $F(x; \theta)$. Επιθυμούμε να ελέγξουμε την υπόθεση

$$H_0: \theta \in \Theta_0$$

$$H_1: \theta \in \Theta_1.$$

Χωρίζοντας το δειγματοληπτικό χώρο Ω σε δύο ξένα υποσύνολα K και $A = \Omega - K$, αποφασίζουμε ότι,

- αν $(X_1, X_2, \dots, X_n) \in K$ (π.χ. αν $T(X_1, X_2, \dots, X_n) > c$), απορρίπτουμε την $H_0: \theta \in \Theta_0$
- αν $(X_1, X_2, \dots, X_n) \in A$ ($\notin K$), δεχόμαστε την $H_0: \theta \in \Theta_0$,

όπου το K λαμβάνεται έτσι ώστε

$$P(\text{σφάλμα τύπου } I) = P((X_1, X_2, \dots, X_n) \in K / H_0) \leq \alpha,$$

$$P(\text{σφάλμα τύπου } II) = P((X_1, X_2, \dots, X_n) \notin K / H_1) : \text{ελάχιστο δυνατό.}$$

Στον παρακάτω πίνακα δίνονται τα δυνατά αποτελέσματα του ελέγχου μιας υποθέσεως:

Απόφαση Φυσική κατάσταση	Αποδοχή της H_0	Απόρριψη της H_0
H_0 ορθή	ορθή απόφαση, $P(H_0 H_0)=1-\alpha$	Σφάλμα τύπου I , $P(H_1 H_0)=\alpha$
H_0 όχι ορθή	Σφάλμα τύπου II , $P(H_0 H_1)=\beta$	ορθή απόφαση, $P(H_1 H_1) = 1-\beta = \pi(\theta)$ ισχύς

Τέλος, αξίζει να παρατηρήσουμε ότι, με βάση την παραπάνω διαδικασία, αυτό που μας ενδιαφέρει περισσότερο είναι να διατηρηθεί μικρή η $P(I)$ και για αυτό απαιτούμε $P(I) \leq \alpha$ (μικρή πιθανότητα εσφαλμένης απόρριψης της H_0). Η $P(II)$ μπορεί να είναι και αυτή μικρή, μπορεί όμως να είναι και αρκετά μεγάλη, ανάλογα την περίπτωση. Για το λόγο αυτό αν $(X_1, X_2, \dots, X_n) \in K$ τότε λέμε ότι «**απορρίπτουμε την H_0** » με πιθανότητα να κάνουμε λάθος μικρότερη του α , ενώ αν $(X_1, X_2, \dots, X_n) \in A$, τότε συνήθως λέμε ότι «**δεν έχουμε αρκετά στοιχεία ώστε να απορρίψουμε την H_0** ». Αποφεύγουμε δηλαδή να πούμε ότι «δεχόμαστε την H_0 » διότι η πιθανότητα να έχουμε κάνει λάθος ($P(II)$) μπορεί να είναι και μεγάλη.

Η p -τιμή ή p -value ενός ελέγχου με βάση ένα τ.δ.

Μία άλλη ποσότητα που συνήθως χρησιμοποιείται στην πράξη είναι το p -value (ή p -τιμή ή κρίσιμη τιμή). Για παράδειγμα, οι έλεγχοι στα στατιστικά πακέτα επεξεργασίας δεδομένων που βασίζονται σε H/Y γίνονται σχεδόν πάντα με τη βοήθεια του p -value. Η τιμή του p -value *εξαρτάται κάθε φορά από το τυχαίο δείγμα που παίρνουμε*.

Έστω λοιπόν τ.δ. $X_1, X_2, \dots, X_n$ από μία κατανομή $F(x; \theta)$ και έστω ότι ενδιαφερόμαστε για τον έλεγχο της υπόθεσης $H_0: \theta \in \Theta_0$ έναντι της $H_1: \theta \in \Theta_1$. Η περιοχή απόρριψης του ελέγχου θα είναι γενικά της μορφής

$$K: T(X_1, \dots, X_n) > c$$

όπου $T(X_1, \dots, X_n)$ είναι μία στατιστική συνάρτηση του δείγματος $X_1, \dots, X_n$. Αν κατά την πραγματοποίηση του πειράματος ελήφθη το δείγμα με τιμές $x_1, x_2, \dots, x_n$ τότε το p -value του συγκεκριμένου δείγματος ορίζεται ως η πιθανότητα

$$p - \text{value} = P(T(X_1, X_2, \dots, X_n) > T(x_1, x_2, \dots, x_n) / H_0).$$

Π.χ., αν στο παραπάνω παράδειγμα την τελευταία ώρα ελέγχθησαν όλες οι συσκευές και βρέθηκε ποσοστό ελαττωματικών $\bar{X} = \bar{x} = 7\%$ τότε, επειδή στη συγκεκριμένη περίπτωση $K: \bar{X} > c$, θα είναι

$$p - value = P(\bar{X} > \bar{x} / H_0) = P(\bar{X} > 0.07 / p = 0.05) = P\left(\frac{\bar{X} - 0.05}{\sqrt{0.05 \cdot 0.95/n}} > \frac{0.07 - 0.05}{\sqrt{0.05 \cdot 0.95/n}}\right) \\ \approx 1 - \Phi\left(\frac{0.07 - 0.05}{\sqrt{0.05 \cdot 0.95/200}}\right) = 1 - \Phi(1.29) = 1 - 0.9015 = 0.0985.$$

Αρα το p -value του συγκεκριμένου ελέγχου για δείγμα που έδωσε $\bar{X} = 7\%$ είναι 9.85%. Η σπουδαιότητα της τιμής του p -value γίνεται φανερή από την εξής παρατήρηση: αν επιλέξουμε επίπεδο σημαντικότητας $\alpha > p$ -value τότε θα ζητάμε κρίσιμη περιοχή $K: \bar{X} > c$ τέτοια ώστε

$$P(\bar{X} \in (c, 1] / H_0) = \alpha > p - value = P(\bar{X} \in (0.07, 1] / H_0)$$

και συνεπώς για το c που θα επιλέγαμε θα ίσχυε ότι $(c, 1] \supseteq (0.07, 1]$, δηλαδή $0.07 > c$. Αρα, αν επιλέγαμε $\alpha > p$ -value τότε, με βάση το συγκεκριμένο δείγμα που έδωσε ποσοστό ελαττωματικών 0.07, θα βρίσκαμε ότι $\bar{X} = 0.07 > c$ και άρα θα απορρίπταμε την H_0 . Από την άλλη μεριά, αν επιλέγαμε $\alpha < p$ -value τότε όμοια βρίσκουμε ότι θα πρέπει $0.07 < c$ και άρα, με βάση το παραπάνω τ.δ. θα ήταν $\bar{X} = 0.07 < c$ και θα δεχόμασταν την H_0 . Η παρατήρηση αυτή είναι εύκολο να δειχθεί ότι ισχύει για οποιοδήποτε έλεγχο και επομένως γενικότερα αν, με βάση ένα τ.δ.,

- το p -value είναι μεγαλύτερο του α τότε δεχόμαστε την H_0
- το p -value είναι μικρότερο του α τότε απορρίπτουμε την H_0

	0	α	1
p -value:	← απόρριψη της H_0 αποδοχή της H_0 →		

Π.χ. στο παραπάνω παράδειγμα ($\bar{X} = 7\%$) βρήκαμε p -value ίσο με 9.85%. Επομένως, αν επιλέξουμε να κάνουμε έλεγχο σε επίπεδο σημαντικότητας $\alpha = 5\%$, τότε επειδή p -value = 9.85% $>$ 5% = α θα δεχτούμε την H_0 : $p = 0.05$. Αντίθετα, αν επιλέξουμε $\alpha = 10\%$ τότε επειδή 9.85% $<$ 10% = α θα απορρίψουμε την H_0 : $p = 0.05$.

(Αν π.χ. την επόμενη ώρα παίρναμε ποσοστό ελαττωματικών $\bar{X} = 9\%$ τότε το αντίστοιχο p -value θα ήταν

$$p - value = P(\bar{X} > 0.09 / p = 0.05) \approx 1 - \Phi\left(\frac{0.09 - 0.05}{\sqrt{0.05 \cdot 0.95/200}}\right) = 1 - \Phi(2.59) = 1 - 0.9952 = 0.004$$

και αν $\alpha = 0.05\%$ θα απορρίπταμε ($\alpha = 5\% > 0.4\%$) την H_0 : $p = 0.5$. Παρατηρούμε ότι σε αυτή την περίπτωση θα απορρίπταμε την H_0 ακόμη και αν επιλέγαμε $\alpha = 1\%$)

Είναι ενδιαφέρον να παρατηρήσουμε ότι το p -value ενός τυχαίου δείγματος μπορεί να θεωρηθεί ίσο με την πιθανότητα να εμφανιστεί ένα τόσο ή ακόμη και πιο «ακραίο» δείγμα από αυτό που εμφανίστηκε, δεδομένου ότι ισχύει η H_0 . Πράγματι, στο παραπάνω παράδειγμα το δειγματικό ποσοστό βρέθηκε ίσο με 0.07 και η πιθανότητα να εμφανιστεί ένα ακόμη και πιο «ακραίο» δείγμα είναι η πιθανότητα να βρεθεί $\bar{X} > 0.07$ δεδομένου ότι ισχύει η H_0 δηλαδή

$$p - value = P(\bar{X} > 0.07 / p = 0.05).$$

Επομένως, αν το p -value ενός τ.δ. $X_1, X_2, \dots, X_n$, βρεθεί πολύ κοντά στο 0, τότε μπορούμε να πούμε ότι η πιθανότητα να εμφανιστεί αυτό το δείγμα (υπό την H_0) είναι μικρή. Με άλλα λόγια είναι «απίθανο», δεδομένης της H_0 , να εμφανιστεί αυτό το τόσο «ακραίο» δείγμα, και όπως είναι φυσικό φτάνουμε στο συμπέρασμα ότι μάλλον δεν πρέπει να ισχύει η H_0 . Αρα όταν το p -value είναι «μικρό», απορρίπτουμε την H_0 . Το πόσο «μικρό» πρέπει να είναι εξαρτάται όπως είδαμε παραπάνω από το επίπεδο σημαντικότητας α που επιθυμούμε να έχουμε. Η λέξη ακραίο μπαίνει σε εισαγωγικά διότι αν το p -value βρεθεί σχετικά μεγάλο τότε το δείγμα δεν είναι καθόλου ακραίο, αντίθετα μπορεί να πει κανείς ότι ήταν αναμενόμενο υπό την H_0 (και άρα δεν απορρίπτουμε την H_0).

Συνήθως στα στατιστικά πακέτα επεξεργασίας δεδομένων με H/Y (π.χ. SPSS, Statgraphics, Statistica, S-Plus κ.α.) με την εισαγωγή των τιμών του δείγματος $X_1, X_2, \dots, X_n$ και την εκλογή του ελέγχου (π.χ. $H_0: p=0.05$ έναντι της $H_1: p=0.10$) το πακέτο εμφανίζει αμέσως το p -value του δείγματος οπότε άμεσα εμείς μπορούμε (με βάση το προαποφασισμένο α) να κρίνουμε αν θα απορρίψουμε ή θα δεχτούμε την H_0 .

Λήμμα Neyman - Pearson, κριτήριο γενικευμένου λόγου πιθανοφανειών

Παρατηρούμε ότι αν και στο Παράδειγμα 6.1 που εξετάσαμε παραπάνω καταφέραμε να βρούμε κατάλληλο K για τον έλεγχο $H_0: p=0.05$ έναντι $H_1: p=0.1$, σε επίπεδο σημαντικότητας α ($K: \bar{X} > c = 0.05 + \sqrt{0.05 \cdot 0.95/n} Z_\alpha$), δεν έχουμε ακόμη περιγράψει μία γενική μέθοδο εύρεσης της κρίσιμης περιοχής K . Έστω δηλαδή ότι γενικά έχουμε ένα τ.δ. $X_1, X_2, \dots, X_n \sim F(x; \theta)$ και επιθυμούμε να ελέγξουμε την υπόθεση $H_0: \theta \in \Theta_0$ έναντι της $H_1: \theta \in \Theta_1$. Ποια θα είναι η κρίσιμη περιοχή K του ελέγχου ώστε

$$P(\text{σφάλμα τύπου I}) = P((X_1, X_2, \dots, X_n) \in K / H_0) \leq \alpha, \text{ και}$$

$$P(\text{σφάλμα τύπου II}) = P((X_1, X_2, \dots, X_n) \notin K / H_1) : \text{ελάχιστο δυνατό;}$$

Μία απάντηση σε αυτό το ερώτημα, όταν H_0, H_1 είναι απλές υποθέσεις, έρχεται να δώσει το γνωστό Λήμμα Neyman - Pearson. Αν λοιπόν $X_1, X_2, \dots, X_n$ ένα τ.δ. από κατανομή με σ.κ. $F(x; \theta)$ και σ.π. ή σ.π.π. $f(x; \theta)$, τότε ισχύει το επόμενο αποτέλεσμα.

Θεώρημα 6.1 (Λήμμα Neyman - Pearson) Έστω ότι θέλουμε να ελέγξουμε την υπόθεση

$$H_0: \theta = \theta_0 \text{ έναντι της } H_1: \theta = \theta_1$$

σε επίπεδο σημαντικότητας α . Ο έλεγχος που ορίζεται από την κρίσιμη περιοχή

$$K: \lambda(\mathbf{X}) = \frac{L(\theta_0)}{L(\theta_1)} = \frac{\prod_{i=1}^n f(X_i; \theta_0)}{\prod_{i=1}^n f(X_i; \theta_1)} < c,$$

όπου c σταθερά τέτοια ώστε $P(I) = \alpha$, παρουσιάζει τη μεγαλύτερη ισχύ ανάμεσα σε όλους τους ελέγχους με επίπεδο σημαντικότητας α .

Το πηλίκο που εμφανίζεται στο παραπάνω θεώρημα

$$\lambda(\mathbf{X}) = \frac{\prod_{i=1}^n f(X_i; \theta_0)}{\prod_{i=1}^n f(X_i; \theta_1)}$$

καλείται *λόγος πιθανοφανειών* του δείγματος.

Επομένως, σύμφωνα με το λήμμα $N - P$ για να βρούμε την κρίσιμη περιοχή K που δίνει τον καλύτερο δυνατό έλεγχο (εξασφαλίζει ότι $P(I) \leq \alpha$, ενώ έχει τη μικρότερη $P(II)$ ανάμεσα στους όλους τους ελέγχους με επίπεδο σημαντικότητας α) αρκεί να ορίσουμε την περιοχή

$$K: \lambda(\mathbf{X}) = \frac{\prod_{i=1}^n f(X_i; \theta_0)}{\prod_{i=1}^n f(X_i; \theta_1)} < c \text{ ώστε } P(I) = P(\mathbf{X} \in K / H_0) = \alpha.$$

Παρατηρούμε ότι το παραπάνω θεώρημα εφαρμόζεται μόνο στην περίπτωση που έχουμε έλεγχο απλής έναντι απλούς υποθέσεως. Στη γενικότερη περίπτωση μπορεί να χρησιμοποιηθεί το λεγόμενο κριτήριο του γενικευμένου λόγου πιθανοφανειών που θα παρουσιάσουμε στη συνέχεια. Από το κριτήριο αυτό δεν προκύπτει απαραίτητα «βέλτιστος» έλεγχος (δηλαδή με τη μεγαλύτερη ισχύ από όλους τους ελέγχους σε επίπεδο σημαντικότητας α). Κάτι τέτοιο εξάλλου δεν θα ήταν δυνατό αφού αποδεικνύεται ότι υπάρχουν περιπτώσεις (π.χ. έλεγχοι αμφίπλευρων υποθέσεων) όπου δεν μπορεί να υπάρξει τέτοιος έλεγχος. Παρόλα αυτά το κριτήριο του γενικευμένου λόγου

πιθανοφανειών είναι ένα εύχρηστο εργαλείο που στις περισσότερες περιπτώσεις δίνει αρκετά ικανοποιητικά αποτελέσματα.

Συγκεκριμένα, αν $X_1, X_2, \dots, X_n$ είναι ένα τ.δ. από κατανομή με σ.κ. $F(x; \theta)$ και σ.π. ή σ.π.π. $f(x; \theta)$, και θέλουμε να ελέγξουμε την υπόθεση $H_0: \theta \in \Theta_0$ έναντι της $H_1: \theta \in \Theta_1$ τότε επιλέγουμε ως κρίσιμη περιοχή του ελέγχου την

$$K: \lambda^*(\mathbf{X}) = \frac{\sup_{\theta \in \Theta_0} L(\theta)}{\sup_{\theta \in \Theta_0 \cup \Theta_1} L(\theta)} = \frac{\sup_{\theta \in \Theta_0} \prod_{i=1}^n f(X_i; \theta)}{\sup_{\theta \in \Theta_0 \cup \Theta_1} \prod_{i=1}^n f(X_i; \theta)} < c$$

όπου c σταθερά τέτοια ώστε $P(I)=a$. Όπως γίνεται φανερό και από την ονομασία του, το κριτήριο γενικευμένου λόγου πιθανοφανειών μπορεί να θεωρηθεί ότι αποτελεί γενίκευση του λόγου πιθανοφανειών που χρησιμοποιείται στο παραπάνω Λήμμα N-P. Είναι προφανές ότι το συγκεκριμένο κριτήριο θα χρησιμοποιείται μόνο σε περιπτώσεις όπου δεν μπορούμε να αξιοποιήσουμε το Λήμμα N-P.

Ας δούμε όμως πως μπορούμε να εφαρμόσουμε τους παραπάνω κανόνες σε ορισμένες ενδιαφέρουσες περιπτώσεις.

α. Έλεγχος υποθέσεων για το μέσο κανονικής κατανομής όταν το σ^2 είναι γνωστό.

Έστω $X_1, X_2, \dots, X_n$ ένα τ.δ. από την $N(\mu, \sigma^2)$ με σ^2 γνωστό. Επιθυμούμε αρχικά να ελέγξουμε την υπόθεση

$$H_0: \mu = \mu_0 \text{ έναντι της } H_1: \mu = \mu_1 \text{ όταν } \mu_1 > \mu_0.$$

Διαπιστώνουμε ότι πρόκειται για τον έλεγχο απλής υποθέσεως έναντι απλής οπότε μπορούμε να εφαρμόσουμε το Λήμμα N - P. Θα έχουμε ότι $\theta_0 = \mu_0$ και $\theta_1 = \mu_1$ και άρα

$$\begin{aligned} K: \lambda(\mathbf{x}) &= \frac{\prod_{i=1}^n f(x_i; \mu_0)}{\prod_{i=1}^n f(x_i; \mu_1)} = \frac{\prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_i - \mu_0)^2}{2\sigma^2}\right)}{\prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_i - \mu_1)^2}{2\sigma^2}\right)} = \frac{\exp\left(-\sum_{i=1}^n \frac{(x_i - \mu_0)^2}{2\sigma^2}\right)}{\exp\left(-\sum_{i=1}^n \frac{(x_i - \mu_1)^2}{2\sigma^2}\right)} \\ &= e^{-\sum_{i=1}^n \frac{(x_i - \mu_0)^2}{2\sigma^2} + \sum_{i=1}^n \frac{(x_i - \mu_1)^2}{2\sigma^2}} = e^{-\frac{1}{2\sigma^2} \left(\sum_{i=1}^n (x_i - \mu_0)^2 - \sum_{i=1}^n (x_i - \mu_1)^2 \right)}. \end{aligned}$$

Παρατηρούμε στη συνέχεια ότι

$$\begin{aligned} \lambda(\mathbf{x}) < c &\Leftrightarrow \sum_{i=1}^n (x_i - \mu_0)^2 - \sum_{i=1}^n (x_i - \mu_1)^2 > -2\sigma^2 \ln c \equiv c' \\ &\Leftrightarrow \left(\sum_{i=1}^n x_i^2 + n\mu_0^2 - 2\mu_0 \sum_{i=1}^n x_i \right) - \left(\sum_{i=1}^n x_i^2 + n\mu_1^2 - 2\mu_1 \sum_{i=1}^n x_i \right) > c' \\ &\Leftrightarrow 2(\mu_1 - \mu_0) \sum_{i=1}^n x_i > c' - n\mu_0^2 + n\mu_1^2 \equiv c'' \end{aligned}$$

και συνεπώς, επειδή $\mu_1 - \mu_0 > 0$, συμπεραίνουμε ότι η ανισότητα $\lambda(\mathbf{X}) < c$ είναι ισοδύναμη με την

$$\sum_{i=1}^n X_i > \frac{c''}{2(\mu_1 - \mu_0)} \equiv c_0.$$

Άρα η κρίσιμη περιοχή του ελέγχου (περιοχή απόρριψης) θα έχει τη μορφή

$$K: \sum_{i=1}^n X_i > c_0$$

για κάποιο c_0 το οποίο θα καθορίζεται από την $P(I) = P(X \in K / H_0) = \alpha$. Επομένως, το c_0 θα πρέπει να είναι τέτοιο ώστε

$$P(I) = P\left(\sum_{i=1}^n X_i > c_0 / H_0\right) = P\left(\frac{\sum_{i=1}^n X_i - n\mu_0}{\sqrt{n\sigma^2}} > \frac{c_0 - n\mu_0}{\sqrt{n\sigma^2}} / H_0\right) = \alpha$$

και επειδή, κάτω από την H_0 , η τ.μ.

$$Z = \frac{\sum_{i=1}^n X_i - n\mu_0}{\sqrt{n\sigma^2}} \sim N(0,1)$$

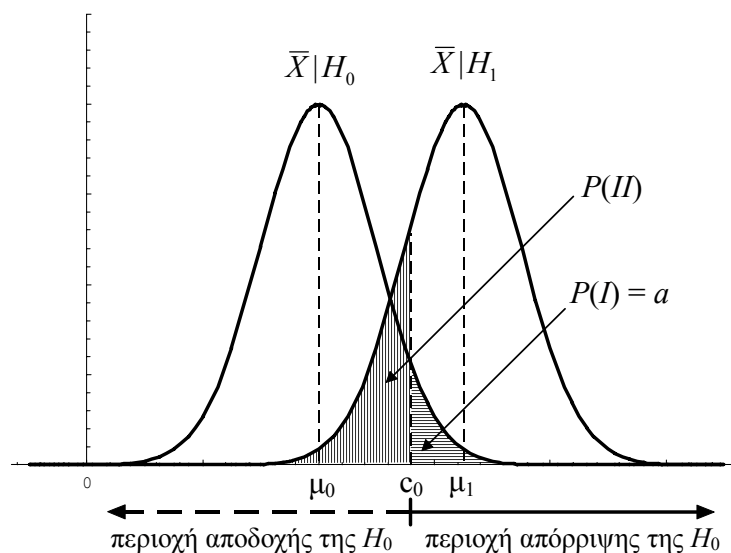
($\sum_{i=1}^n X_i \sim N(n\mu_0, n\sigma^2)$) συμπεραίνουμε τελικά ότι

$$\frac{c_0 - n\mu_0}{\sqrt{n\sigma^2}} = Z_\alpha \Rightarrow c_0 = n\mu_0 + \sqrt{n\sigma^2} Z_\alpha.$$

Άρα, σύμφωνα με το Λήμμα N-P, ο καλύτερος έλεγχος σε επίπεδο σημαντικότητας α για την υπόθεση $H_0: \mu = \mu_0$ έναντι της $H_1: \mu = \mu_1, \mu_1 > \mu_0$, θα έχει περιοχή απόρριψης:

$$\bar{X} > \mu_0 + \sqrt{\frac{\sigma^2}{n}} Z_\alpha \text{ ή ισοδύναμα } K: \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} > Z_\alpha.$$

Σχηματικά, η περιοχή απόρριψης θα είναι:



Η πιθανότητα σφάλματος τύπου I αυτού του ελέγχου θα είναι φυσικά ίση με α ενώ η πιθανότητα σφάλματος τύπου II θα είναι ίση με

$$P(\bar{X} \leq \mu_0 + \sqrt{\frac{\sigma^2}{n}} Z_\alpha / H_1) = P(\bar{X} \leq \mu_0 + \sqrt{\frac{\sigma^2}{n}} Z_\alpha / \mu = \mu_1) = P\left(\frac{\bar{X} - \mu_1}{\sqrt{\sigma^2/n}} \leq \frac{\mu_0 - \mu_1}{\sqrt{\sigma^2/n}} + Z_\alpha / \mu = \mu_1\right)$$

και επειδή κάτω από την H_1 , ισχύει ότι

$$Z' = \frac{\bar{X} - \mu_1}{\sqrt{\sigma^2/n}} \sim N(0,1),$$

προκύπτει ότι

$$P(II) = \Phi\left(\frac{\mu_0 - \mu_1}{\sqrt{\sigma^2/n}} + Z_\alpha\right).$$

Επίσης, η ισχύς του ελέγχου θα είναι

$$\pi(\mu_1) = 1 - P(II) = 1 - \Phi\left(\frac{\mu_0 - \mu_1}{\sqrt{\sigma^2/n}} + Z_a\right) = \Phi\left(\frac{\mu_1 - \mu_0}{\sqrt{\sigma^2/n}} - Z_a\right).$$

Παρατηρούμε ότι η κρίσιμη περιοχή K δεν εξαρτάται από την τιμή μ_1 και επομένως η ίδια κρίσιμη περιοχή μπορεί να χρησιμοποιηθεί και για τον έλεγχο της υπόθεσης $H_0: \mu = \mu_0$ έναντι της $H_1: \mu > \mu_0$ (απλή έναντι μονόπλευρης) οδηγώντας και πάλι σε βέλτιστο έλεγχο. Η τιμή της μ_1 ωστόσο επηρεάζει την $P(II)$ καθώς και την ισχύ του ελέγχου. Όταν η μ_1 αυξάνεται τότε η $P(II)$ μειώνεται και η ισχύς $\pi(\mu_1)$ του ελέγχου αυξάνεται. Αυτό μπορεί εύκολα να γίνει αντιληπτό και από το παραπάνω σχήμα. Αποδεικνύεται επίσης ότι η ίδια κρίσιμη περιοχή K οδηγεί σε βέλτιστο έλεγχο της υπόθεσης $H_0: \mu \leq \mu_0$ έναντι της $H_1: \mu > \mu_0$ (μονόπλευρη έναντι μονόπλευρης).

Αντίστοιχα με τα ζεύγη των υποθέσεων

$$H_0: \mu = \mu_0, H_1: \mu = \mu_1, \mu_1 > \mu_0 \quad \text{ή} \quad H_0: \mu = \mu_0, H_1: \mu > \mu_0, \quad \text{ή} \quad H_0: \mu \leq \mu_0, H_1: \mu > \mu_0$$

μπορούμε να αντιμετωπίσουμε και τα ζεύγη

$$H_0: \mu = \mu_0, H_1: \mu = \mu_1, \mu_1 < \mu_0 \quad \text{ή} \quad H_0: \mu = \mu_0, H_1: \mu < \mu_0 \quad \text{ή} \quad H_0: \mu \geq \mu_0, H_1: \mu < \mu_0.$$

Σε αυτές τις περιπτώσεις, αν εργαστούμε όμοια με παραπάνω καταλήγουμε στην κρίσιμη περιοχή

$$K: \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} < -Z_a$$

η οποία ορίζει έναν βέλτιστο έλεγχο (με τη μεγαλύτερη δυνατή ισχύ) σε επίπεδο σημαντικότητας α .

Τέλος, θα εξετάσουμε την περίπτωση $H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$, (απλή έναντι αμφίπλευρης υπόθεσης). Σε αυτή την περίπτωση δεν μπορούμε να χρησιμοποιήσουμε το λήμμα N-P διότι δεν έχουμε απλή έναντι απλής υπόθεσης. Θα χρησιμοποιήσουμε το κριτήριο του γενικευμένου λόγου πιθανοφανειών. Σύμφωνα με αυτό, η κρίσιμη περιοχή θα έχει τη μορφή

$$K: \lambda^*(\mathbf{x}) = \frac{\sup_{\theta \in \Theta_0} \prod_{i=1}^n f(X_i; \theta)}{\sup_{\theta \in \Theta_0 \cup \Theta_1} \prod_{i=1}^n f(X_i; \theta)} = \frac{\prod_{i=1}^n f(X_i; \mu_0)}{\sup_{\mu \in \mathbb{R}} \prod_{i=1}^n f(X_i; \mu)} < c$$

και επειδή η συνάρτηση πιθανοφάνειας $L(\mu) = \prod_{i=1}^n f(X_i; \mu)$ έχει μέγιστο για μ ίσο με την εκτιμήτρια μέγιστης πιθανοφάνειας θα έχουμε ότι

$$K: \frac{\prod_{i=1}^n f(X_i; \mu_0)}{\prod_{i=1}^n f(X_i; \hat{\mu})} = \frac{\prod_{i=1}^n f(X_i; \mu_0)}{\prod_{i=1}^n f(X_i; \bar{X})} < c$$

Τελικά μετά από σχετικές πράξεις προκύπτει ότι

$$K: |\bar{X} - \mu_0| > c_0$$

Επειδή θα πρέπει $P(I) = \alpha$ τελικά βρίσκουμε ότι η περιοχή απόρριψης της $H_0: \mu = \mu_0$ έναντι της $H_1: \mu \neq \mu_0$, θα είναι η

$$K: |\bar{X} - \mu_0| > \frac{\sigma}{\sqrt{n}} Z_{\alpha/2} \quad \text{ή ισοδύναμα} \quad K: \frac{|\bar{X} - \mu_0|}{\sigma/\sqrt{n}} > Z_{\alpha/2}.$$

Αξίζει εδώ να παρατηρήσουμε ότι ο παραπάνω αμφίπλευρος έλεγχος σε επίπεδο σημαντικότητας α είναι ισοδύναμος με τον έλεγχο μέσω ενός διαστήματος εμπιστοσύνης συντελεστού $1 - \alpha$. Πράγματι, στο προηγούμενο κεφάλαιο είδαμε ότι ένα δ.ε. για το μ συντελεστού $1 - \alpha$ είναι το

$$[\bar{X} - \frac{\sigma}{\sqrt{n}} Z_{\alpha/2}, \bar{X} + \frac{\sigma}{\sqrt{n}} Z_{\alpha/2}]$$

και απορρίπταμε ότι $\mu = \mu_0$ με «συντελεστή εμπιστοσύνης» $1-\alpha$ αν το μ_0 δεν περιέχονταν στο δ.ε. Δηλαδή αν

$$\mu_0 > \bar{X} + \frac{\sigma}{\sqrt{n}} Z_{\alpha/2} \quad \text{ή} \quad \mu_0 < \bar{X} - \frac{\sigma}{\sqrt{n}} Z_{\alpha/2} \quad \text{ή ισοδύναμα αν} \quad \frac{|\bar{X} - \mu_0|}{\sigma/\sqrt{n}} > Z_{\alpha/2}$$

που, σύμφωνα με τα παραπάνω, συμπίπτει με την περιοχή απόρριψης K του ελέγχου της υπόθεσης $H_0: \mu = \mu_0$ έναντι της $H_1: \mu \neq \mu_0$ σε επίπεδο σημαντικότητας α . Η παρατήρηση αυτή ισχύει γενικότερα και έτσι μπορούμε να πούμε ότι ο έλεγχος της υπόθεσης $\theta = \theta_0$ μέσω ενός δ.ε. για το θ συντελεστού $1-\alpha$ (αν $\theta_0 \notin \delta.ε.$ απορρίπτουμε ότι $\theta = \theta_0$) είναι ισοδύναμος με τον αμφίπλευρο έλεγχο $H_0: \theta = \theta_0, H_1: \theta \neq \theta_0$ σε επίπεδο σημαντικότητας α .

Άσκηση 6.1. Γνωρίζουμε ότι το μήνα Ιανουάριο η μέση τιμή πώλησης ενός προϊόντος σε μία περιοχή είναι 1000δρχ. Η μέση τιμή πώλησης του ίδιου προϊόντος σε ένα τ.δ. $n=10$ καταστημάτων το μήνα Φεβρουάριο βρέθηκε ίση με 1030 δρχ. Μπορούμε σε επίπεδο σημαντικότητας 5% να πούμε ότι η μέση τιμή πώλησης αυξήθηκε κατά το μήνα Φεβρουάριο; (γνωρίζουμε ότι η τιμή του προϊόντος ακολουθεί $N(\mu, \sigma=50\deltaρχ)$). Να βρεθεί το p -value του δείγματος και η συνάρτηση ισχύος. Αν τελικά ισχυε ότι $\mu=1040$ με τι πιθανότητα θα παίρναμε σωστή απόφαση;

Λύση. Στη συγκεκριμένη περίπτωση έχουμε στην ουσία ένα τ.δ. $X_1, X_2, \dots, X_n, n=10$ από την $N(\mu, \sigma^2=50^2)$ και ζητείται ο έλεγχος της υπόθεσης

$$H_0: \mu = \mu_0 (=1000) \text{ έναντι της } H_1: \mu > \mu_0 (=1000)$$

σε επίπεδο σημαντικότητας α . Σύμφωνα με τα παραπάνω, γνωρίζουμε ότι η κρίσιμη περιοχή (περιοχή απόρριψης της H_0) ενός τέτοιου ελέγχου είναι η

$$K: \bar{X} > \mu_0 + \sqrt{\frac{\sigma^2}{n}} Z_{\alpha}.$$

Αντικαθιστώντας θα έχουμε ότι,

$$\mu_0 + \sqrt{\frac{\sigma^2}{n}} Z_{\alpha} = 1000 + \sqrt{\frac{50^2}{10}} 1.645 \approx 1026$$

και επειδή $\bar{X} = 1030 > 1026$ απορρίπτουμε την $H_0: \mu = \mu_0 = 1000\deltaρχ$ έναντι της $H_1: \mu > 1000\deltaρχ$. Το p -value του ελέγχου θα είναι

$$\begin{aligned} P(\bar{X} > 1030 / H_0) &= P\left(\frac{\bar{X} - \mu_0}{\sqrt{\sigma^2/n}} > \frac{1030 - \mu_0}{\sqrt{\sigma^2/n}} / H_0\right) = 1 - \Phi\left(\frac{1030 - \mu_0}{\sqrt{\sigma^2/n}}\right) \\ &= 1 - \Phi\left(\frac{30}{50/\sqrt{10}}\right) = 1 - \Phi(1.89) = 0.029 = p\text{-value}. \end{aligned}$$

Επειδή το p -value = 2.9% είναι μικρότερο του δοθέντος επιπέδου σημαντικότητας 5%, θα πρέπει να απορρίψουμε την H_0 (όπως έχουμε ήδη κάνει παραπάνω). Παρατηρούμε ότι ακόμη και αν πάρουμε $\alpha = 3\%$ εξακολουθούμε να απορρίπτουμε την H_0 . Όπως έχουμε αναφέρει παραπάνω, μπορούμε να δούμε το p -value ως την πιθανότητα να εμφανιστεί ένα τόσο ακραίο δείγμα ενώ ισχύει η H_0 . Επομένως μόνο στο 2.9% των περιπτώσεων που ισχύει η H_0 θα μπορούσαμε να πάρουμε ένα τόσο μεγάλο (σε σχέση με το $\mu = 1000$) δειγματικό μέσο. Άρα, φυσιολογικά απορρίψαμε ότι ισχύει η H_0 (δηλαδή ότι $\mu = 1000$) με πιθανότητα να κάνουμε λάθος μικρότερο του 3%.

Επίσης, η ισχύς του ελέγχου θα είναι

$$\begin{aligned}\pi(\mu) &= 1 - P(II) = 1 - P(\bar{X} \leq \mu_0 + \sqrt{\frac{\sigma^2}{n}} Z_a / H_1) = 1 - P\left(\frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}} \leq \frac{\mu_0 - \mu}{\sqrt{\sigma^2/n}} + Z_a / H_1\right) \\ &= 1 - \Phi\left(\frac{\mu_0 - \mu}{\sqrt{\sigma^2/n}} + Z_a\right) = \Phi\left(\frac{\mu - \mu_0}{\sqrt{\sigma^2/n}} - Z_a\right) = \Phi\left(\frac{\mu - 1000}{50/\sqrt{10}} - 1.645\right).\end{aligned}$$

Τέλος, αν ισχυρε η H_1 με $\mu=1040$ τότε

$$\pi(1040) = \Phi\left(\frac{1040 - 1000}{50/\sqrt{10}} - 1.645\right) = \Phi\left(\frac{40}{50/\sqrt{10}} - 1.645\right) = \Phi(0.88) = 0.8106,$$

η οποία είναι στην ουσία η πιθανότητα ορθής απόρριψης της H_0 (πιθανότητα σωστής απόφασης) αν ισχυρε ότι $\mu=1040$.

Άσκηση 6.2. Έστω τ.δ. $X_1, \dots, X_n \sim N(\mu, \sigma^2)$ με σ^2 γνωστό. Ποιο θα πρέπει να είναι το μέγεθος του δείγματος n ώστε για τον έλεγχο της υπόθεσης

$$H_0: \mu = \mu_0, \quad H_1: \mu = \mu_1, \quad \mu_1 < \mu_0$$

οι πιθανότητες σφάλματος τύπου I και II να είναι αντίστοιχα α και β . ($\alpha=0.05$, $\beta=0.1$, $\mu_0=600$, $\mu_1=560$, $\sigma^2=6400$). Ποια θα είναι τότε η κρίσιμη περιοχή;

Λύση. Για τον έλεγχο της συγκεκριμένης υπόθεσης θα έχουμε από τα παραπάνω την κρίσιμη περιοχή (περιοχή απόρριψης)

$$K: \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} < -Z_a$$

για την οποία ως γνωστό ισχύει ότι $P(I)=\alpha$. Θα πρέπει και $P(II)=\beta$. Δηλαδή

$$\begin{aligned}P(II) &= P(\text{αποδοχή της } H_0 / \text{ισχύει η } H_1) = P\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \geq -Z_a / \mu = \mu_1\right) \\ &= P\left(\bar{X} \geq \mu_0 - \frac{\sigma}{\sqrt{n}} Z_a / \mu = \mu_1\right) = P\left(\frac{\bar{X} - \mu_1}{\sigma/\sqrt{n}} \geq \frac{\mu_0 - \mu_1}{\sigma/\sqrt{n}} - Z_a / \mu = \mu_1\right) = \beta\end{aligned}$$

Επομένως $\left(\frac{\bar{X} - \mu_1}{\sigma/\sqrt{n}} \sim N(0,1) / H_1\right)$,

$$\frac{\mu_0 - \mu_1}{\sigma/\sqrt{n}} - Z_a = Z_\beta \quad \text{και ισοδύναμα} \quad n = \sigma^2 \frac{(Z_a + Z_\beta)^2}{(\mu_0 - \mu_1)^2}.$$

Για $\alpha=0.05$, $\beta=0.1$, $\mu_0=600$, $\mu_1=560$, $\sigma^2=6400$ βρίσκουμε ότι

$$n = 6400 \frac{(1.65 + 1.28)^2}{40^2} \approx 35.$$

Η κρίσιμη περιοχή θα είναι η

$$K: \bar{X} < \mu_0 - \frac{\sigma}{\sqrt{n}} Z_a = 600 - \frac{80}{\sqrt{35}} 1.65 = 577.68.$$

Άσκηση 6.3. Από κανονικό πληθυσμό $N(\mu, \sigma^2=900)$ λαμβάνεται τυχαίο δείγμα μεγέθους $n=9$, προκειμένου να ελεγχθεί η υπόθεση $H_0: \mu=100$ έναντι της $H_1: \mu \neq 100$ σε επίπεδο σημαντικότητας

$\alpha=0.05$. i) Να προσδιοριστεί η περιοχή απόρριψης. ii) Να υπολογιστεί η πιθανότητα σφάλματος τύπου II αν ισχύει ότι $\mu=105$.

Λύση. Η περιοχή απόρριψης σύμφωνα με τα παραπάνω θα είναι

$$K: \frac{|\bar{X} - \mu_0|}{\sigma/\sqrt{n}} > Z_{\alpha/2}$$

δηλαδή

$$K: \bar{X} > \mu_0 + \frac{\sigma}{\sqrt{n}} Z_{\alpha/2} \quad \text{ή} \quad \bar{X} < \mu_0 - \frac{\sigma}{\sqrt{n}} Z_{\alpha/2}$$

και άρα η περιοχή απόρριψης θα είναι η ($\mu_0=100$, $Z_{0.025}=1.96$, $n=9$, $\sigma=30$)

$$K: \bar{X} > 119.6 \quad \text{ή} \quad \bar{X} < 80.4$$

Η πιθανότητα σφάλματος τύπου II θα είναι

$$\begin{aligned} P(II) &= P(\text{αποδοχή της } H_0 / \text{ισχύει η } H_1) = P(80.4 \leq \bar{X} \leq 119.6 / \mu=105) \\ &= P\left(\frac{80.4-105}{\sigma/\sqrt{n}} \leq \frac{\bar{X}-105}{\sigma/\sqrt{n}} \leq \frac{119.6-105}{\sigma/\sqrt{n}} / \mu=105\right) = \Phi\left(\frac{119.6-105}{30/\sqrt{9}}\right) - \Phi\left(\frac{80.4-105}{30/\sqrt{9}}\right) \\ &= \Phi(1.46) - \Phi(-2.46) = \Phi(1.46) - 1 + \Phi(2.46) = 0.9279 + 0.9931 - 1 = 0.921. \end{aligned}$$

β. Έλεγχος υποθέσεων για το μέσο κανονικής κατανομής όταν το σ^2 είναι άγνωστο.

Σε αυτή την περίπτωση δεν μπορούμε να χρησιμοποιήσουμε το λήμμα N-P για τον έλεγχο π.χ. της υπόθεσης $H_0: \mu=\mu_0$, $H_1: \mu=\mu_1$, $\mu_1 > \mu_0$. Αυτό συμβαίνει διότι το σ^2 είναι άγνωστο και επομένως σε αυτή την περίπτωση έχουμε στην ουσία να ελέγξουμε την υπόθεση $H_0: (\mu, \sigma^2) \in \{(\mu_0, x), x \in \mathbb{R}^+\}$ έναντι της $H_1: (\mu, \sigma^2) \in \{(\mu_1, x), x \in \mathbb{R}^+\}$ όπου προφανώς δεν πρόκειται για περίπτωση απλής έναντι απλούς υποθέσεως. Χρησιμοποιώντας λοιπόν το κριτήριο του γενικευμένου λόγου πιθανοφανειών προκύπτουν τα παρακάτω:

- για τον έλεγχο σε επίπεδο σημαντικότητας α των υποθέσεων:

$$\begin{array}{lll} H_0: \mu=\mu_0, & \text{ή} & H_0: \mu=\mu_0, & \text{ή} & H_0: \mu \leq \mu_0, \\ H_1: \mu=\mu_1, \mu_1 > \mu_0 & & H_1: \mu > \mu_0, & & H_1: \mu > \mu_0 \end{array}$$

λαμβάνεται ως κρίσιμη περιοχή η

$$K: \frac{\bar{X} - \mu_0}{S/\sqrt{n}} > t_{n-1}(\alpha).$$

- για τον έλεγχο σε επίπεδο σημαντικότητας α των υποθέσεων:

$$\begin{array}{lll} H_0: \mu=\mu_0, & \text{ή} & H_0: \mu=\mu_0, & \text{ή} & H_0: \mu > \mu_0, \\ H_1: \mu=\mu_1, \mu_1 < \mu_0 & & H_1: \mu < \mu_0, & & H_1: \mu \leq \mu_0 \end{array}$$

λαμβάνεται ως κρίσιμη περιοχή η

$$K: \frac{\bar{X} - \mu_0}{S/\sqrt{n}} < -t_{n-1}(\alpha).$$

- για τον έλεγχο σε επίπεδο σημαντικότητας α της υπόθεσης:

$$H_0: \mu=\mu_0,$$

$$H_1: \mu \neq \mu_1,$$

λαμβάνεται ως κρίσιμη περιοχή η

$$K: \frac{|\bar{X} - \mu_0|}{S/\sqrt{n}} > t_{n-1}(a/2).$$

Άσκηση 6.4. Ο σύλλογος καταναλωτών, για να εξακριβώσει αν τα κουτιά των 100gr καφέ περιέχουν πραγματικά μέσο βάρος 100gr, πήρε 9 κουτιά τυχαία από διάφορα καταστήματα και βρήκε $\bar{X}=96\text{gr}$ και $S=3.8\text{gr}$.

i) Είναι τα δεδομένα αρκετά για να υποστηρίξουμε ότι τα κουτιά των 100gr είναι ελλιποβαρή; ($\alpha=0.05$)

ii) Βρείτε το p -value του ελέγχου που χρησιμοποιήσατε.

iii) Αν τελικά ίσχυε ότι το πραγματικό μέσο βάρος ήταν 95gr, με τι πιθανότητα θα παίρναμε σωστή απόφαση (υποθ. ότι $S \approx \sigma$);

Λύση. Στη συγκεκριμένη περίπτωση έχουμε στην ουσία ένα τ.δ. $X_1, X_2, \dots, X_n$, $n=9$ από την $N(\mu, \sigma^2)$ και ζητείται ο έλεγχος της υπόθεσης

$$H_0: \mu = \mu_0 = 100 \text{ έναντι της } H_1: \mu < \mu_0 = 100$$

σε επίπεδο σημαντικότητας $\alpha=0.05$. Σύμφωνα με τα παραπάνω, γνωρίζουμε ότι η κρίσιμη περιοχή (περιοχή απόρριψης της H_0) ενός τέτοιου ελέγχου είναι η

$$K: \frac{\bar{X} - \mu_0}{S/\sqrt{n}} < -t_{n-1}(\alpha).$$

Αντικαθιστώντας θα έχουμε ότι,

$$\frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{96 - 100}{3.8/\sqrt{9}} = -3.15 < -1.860 = -t_8(0.05)$$

και συνεπώς απορρίπτουμε την $H_0: \mu = \mu_0 = 100\text{gr}$ έναντι της $H_1: \mu < 100\text{gr}$. Επομένως, η διαφορά των $100 - 96 = 4\text{gr}$ θεωρείται «στατιστικά σημαντική» και τα κουτιά κρίνονται ελλιποβαρή με πιθανότητα να κάνουμε λάθος $\alpha=5\%$. Το p -value θα είναι

$$p\text{-value} = P\left(\frac{\bar{X} - \mu_0}{S/\sqrt{n}} < -3.15 / H_0\right) = F_{t_{n-1}}(-3.15) = 1 - F_{t_8}(3.15)$$

Η $F_{t_8}(3.15)$ είναι η συνάρτηση κατανομής της t_n με $n=8$ β.ε. στη θέση 3.15. Από πίνακες της t_8 βρίσκουμε ότι $F_{t_8}(2.896)=0.99$ και συνεπώς $p\text{-value}=1 - F_{t_8}(3.15) < 1 - F_{t_8}(2.896)=0.01$. Άρα το p -value είναι πολύ μικρό, μικρότερο του $\alpha=0.05\%$ (και άρα σωστά απορρίψαμε). Παρατηρούμε ότι αφού $p\text{-value} < 0.01$ θα απορρίπταμε την H_0 ακόμη και αν είχαμε θέσει $\alpha=0.01$.

Τέλος, αν ίσχυε η H_1 με $\mu = \mu_1 = 95$ τότε θα παίρναμε σωστή απόφαση αν απορρίπταμε την H_0 . Η πιθανότητα σωστής απόφασης είναι

$$P(\text{απόρριψη της } H_0 / \text{ισχύει η } H_1) = 1 - P(II) = \pi(\mu_1).$$

Άρα τελικά η πιθανότητα ορθής απόφασης όταν ισχύει η H_1 είναι η ισχύς:

$$\pi(\mu_1) = 1 - P(II) = P\left(\frac{\bar{X} - \mu_0}{S/\sqrt{n}} < -t_{n-1}(\alpha) / \mu = \mu_1\right) \approx P\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} < -t_{n-1}(\alpha) / \mu = \mu_1\right)$$

$$\begin{aligned}
&= P\left(\frac{\bar{X}-\mu_1}{\sigma/\sqrt{n}} < \frac{\mu_0-\mu_1}{\sigma/\sqrt{n}} - t_{n-1}(a)/\mu = \mu_1\right) = \Phi\left(\frac{\mu_0-\mu_1}{\sigma/\sqrt{n}} - t_{n-1}(a)\right) \approx \Phi\left(\frac{100-95}{3.8/\sqrt{9}} - 1.86\right) \\
&= \Phi(2.08) = 0.9812.
\end{aligned}$$

Σημειώνεται ότι αν στη συγκεκριμένη περίπτωση θέλαμε να υπολογίσουμε την τιμή της ισχύος ακριβώς, θα έπρεπε να χρησιμοποιήσουμε την λεγόμενη «μη-κεντρική» κατανομή t_{n-1} (συμβ. και με t_{n-1}^δ) η οποία ορίζεται ως η κατανομή ενός πηλίκου δύο ανεξάρτητων τ.μ., μία που ακολουθεί $N(\delta, 1)$ δια μία που ακολουθεί $\sqrt{\chi_{n-1}^2/(n-1)}$ κατανομή (δ : παράμετρος μη-κεντρικότητας). Αν $\delta = 0$ προκύπτει η συνήθης t_{n-1} κατανομή. Συγκεκριμένα θα βρίσκαμε ότι $\pi(\mu_1) = F_{t_{n-1}^\delta}(-t_{n-1}(a))$ όπου $\delta = \frac{\mu_1 - \mu_0}{\sigma/\sqrt{n}}$.

γ. Έλεγχος υποθέσεων για τη διασπορά κανονικής κατανομής όταν το μ είναι γνωστό.

Έστω $X_1, X_2, \dots, X_n$ ένα τ.δ. από την $N(\mu, \sigma^2)$ με μ γνωστό. Επιθυμούμε να ελέγξουμε την υπόθεση

$$H_0: \sigma = \sigma_0 \text{ έναντι της } H_1: \sigma = \sigma_1 \text{ όταν } \sigma_1 > \sigma_0.$$

Πρόκειται για τον έλεγχο απλής υποθέσεως έναντι απλής με $\theta_0 = \sigma_0$ και $\theta_1 = \sigma_1$ και άρα η κρίσιμη σύμφωνα με το Λήμμα N - P θα είναι της μορφής

$$K: \lambda(\mathbf{x}) = \frac{\prod_{i=1}^n f(x_i; \sigma_0)}{\prod_{i=1}^n f(x_i; \sigma_1)} < c$$

μετά από σχετικά απλές αλγεβρικές πράξεις καταλήγουμε στην ακόλουθη ισοδύναμη ανισότητα

$$K: \sum_{i=1}^n (X_i - \mu)^2 > c_0.$$

Το c_0 καθορίζεται από την

$$P(I) = P\left(\sum_{i=1}^n (X_i - \mu)^2 > c_0 / H_0\right) = P\left(\sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma_0}\right)^2 > \frac{c_0}{\sigma_0^2} / H_0\right) = a$$

και επειδή $\sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma_0}\right)^2 \sim \chi_n^2 / H_0$ συμπεραίνουμε τελικά ότι

$$\frac{c_0}{\sigma_0^2} = \chi_n^2(a) \Rightarrow c_0 = \sigma_0^2 \chi_n^2(a).$$

Άρα, σύμφωνα με το Λήμμα N-P, ο καλύτερος έλεγχος σε επίπεδο σημαντικότητας α για την υπόθεση $H_0: \sigma = \sigma_0$ έναντι της $H_1: \sigma = \sigma_1, \sigma_1 > \sigma_0$ (μ γνωστό) θα έχει περιοχή απόρριψης:

$$K: \sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma_0}\right)^2 > \chi_n^2(a)$$

Η ίδια κρίσιμη περιοχή μπορεί να χρησιμοποιηθεί και για τον έλεγχο της υπόθεσης $H_0: \sigma = \sigma_0$ έναντι της $H_1: \sigma > \sigma_0$ (απλή έναντι μονόπλευρης) καθώς και της $H_0: \sigma \leq \sigma_0$ έναντι της $H_1: \sigma > \sigma_0$ (μονόπλευρη έναντι μονόπλευρης).

Αντίστοιχα με τις παραπάνω περιπτώσεις, μπορούμε να αντιμετωπίσουμε και την περίπτωση που $H_0: \sigma = \sigma_0, H_1: \sigma = \sigma_1, \sigma_1 < \sigma_0$ ή $H_0: \sigma = \sigma_0, H_1: \sigma < \sigma_0$ ή $H_0: \sigma \geq \sigma_0, H_1: \sigma < \sigma_0$. Σε αυτή την περίπτωση, αν εργαστούμε όπως παραπάνω, καταλήγουμε στην κρίσιμη περιοχή

$$K: \sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma_0} \right)^2 < \chi_n^2(1-a)$$

η οποία ορίζει έναν βέλτιστο έλεγχο (με τη μεγαλύτερη δυνατή ισχύ) σε επίπεδο σημαντικότητας α .

Τέλος, θα εξετάσουμε την περίπτωση $H_0: \sigma = \sigma_0$, $H_1: \sigma \neq \sigma_0$, (απλή έναντι αμφίπλευρης υπόθεσης). Σε αυτή την περίπτωση χρησιμοποιώντας το κριτήριο του γενικευμένου λόγου πιθανοφαινοίων καταλήγουμε σε μία κρίσιμη περιοχή της μορφής

$$K: \sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma_0} \right)^2 > c_0 \text{ ή } \sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma_0} \right)^2 < c_1$$

και επειδή θα πρέπει $P(I)=a$ τελικά βρίσκουμε ότι η περιοχή απόρριψης της $H_0: \sigma = \sigma_0$ έναντι της $H_1: \sigma \neq \sigma_0$, θα είναι η

$$K: \sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma_0} \right)^2 < \chi_n^2(1 - \frac{a}{2}) \text{ ή } \sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma_0} \right)^2 > \chi_n^2(\frac{a}{2})$$

Όπως έχουμε ήδη επισημάνει και παραπάνω, ο συγκεκριμένος αμφίπλευρος έλεγχος σε επίπεδο σημαντικότητας α είναι ισοδύναμος με τον έλεγχο $\sigma = \sigma_0$ μέσω ενός διαστήματος εμπιστοσύνης για το σ^2 συντελεστού $1-\alpha$ (αν $\sigma_0^2 \notin \delta.ε.$ απορρίπτουμε ότι $\sigma = \sigma_0$).

δ. Έλεγχος υποθέσεων για τη διασπορά κανονικής κατανομής όταν το μ είναι άγνωστο.

Όμοια με το (β) παραπάνω χρησιμοποιούμε και εδώ το κριτήριο του γενικευμένου λόγου πιθανοφαινοίων. Όπως ήταν αναμενόμενο, τελικά προκύπτουν ανάλογες με το (γ) κρίσιμες περιοχές. Συγκεκριμένα, όπου στο (γ) είχαμε $\sum_{i=1}^n (X_i - \mu)^2$ τώρα θα έχουμε $\sum_{i=1}^n (X_i - \bar{X})^2 = (n-1)S^2$, και όπου είχαμε $\chi_n^2(a)$ τώρα θα έχουμε $\chi_{n-1}^2(a)$. Άρα,

- για τον έλεγχο σε επίπεδο σημαντικότητας α των υποθέσεων:

$$\begin{array}{lll} H_0: \sigma = \sigma_0, & \text{ή} & H_0: \sigma = \sigma_0, & \text{ή} & H_0: \sigma \leq \sigma_0, \\ H_1: \sigma = \sigma_1, \sigma_1 > \sigma_0 & & H_1: \sigma > \sigma_0, & & H_1: \sigma > \sigma_0 \end{array}$$

λαμβάνεται ως κρίσιμη περιοχή η

$$K: \frac{(n-1)S^2}{\sigma_0^2} > \chi_{n-1}^2(a)$$

- για τον έλεγχο σε επίπεδο σημαντικότητας α των υποθέσεων:

$$\begin{array}{lll} H_0: \sigma = \sigma_0, & \text{ή} & H_0: \sigma = \sigma_0, & \text{ή} & H_0: \sigma > \sigma_0, \\ H_1: \sigma = \sigma_1, \sigma_1 < \sigma_0 & & H_1: \sigma < \sigma_0, & & H_1: \sigma \leq \sigma_0 \end{array}$$

λαμβάνεται ως κρίσιμη περιοχή η

$$K: \frac{(n-1)S^2}{\sigma_0^2} < \chi_{n-1}^2(1-a).$$

- για τον έλεγχο σε επίπεδο σημαντικότητας α της υπόθεσης:

$$\begin{array}{l} H_0: \sigma = \sigma_0, \\ H_1: \sigma \neq \sigma_0, \end{array}$$

λαμβάνεται ως κρίσιμη περιοχή η

$$K: \frac{(n-1)S^2}{\sigma_0^2} < \chi_{n-1}^2(1-\frac{a}{2}) \quad \text{ή} \quad \frac{(n-1)S^2}{\sigma_0^2} > \chi_{n-1}^2(\frac{a}{2})$$

Άσκηση 6.5. Έστω ότι ο χρόνος ζωής ενός τύπου μπαταριών ακολουθεί κανονική κατανομή $N(\mu, \sigma^2)$. Ο κατασκευαστής ισχυρίζεται ότι $\sigma^2=100$. Μπορούμε να απορρίψουμε τον ισχυρισμό αυτό έναντι του $\sigma^2 > 100$ σε επίπεδο σημαντικότητας 5% αν έχουμε την πληροφορία ότι η δειγματική διασπορά S^2 των χρόνων ζωής ενός τυχαίου δείγματος 20 μπαταριών βρέθηκε ίση με 130. Να βρεθεί το p -value του ελέγχου.

Λύση. Πρόκειται για τον έλεγχο της υπόθεσης

$$H_0: \sigma^2 = \sigma_0^2 \quad \text{έναντι της} \quad H_1: \sigma^2 > \sigma_0^2$$

για τη διασπορά ενός κανονικού πληθυσμού με άγνωστη. Από την παράγραφο (δ) παραπάνω η περιοχή απόρριψης αυτού του ελέγχου θα είναι η

$$K: \frac{(n-1)S^2}{\sigma_0^2} > \chi_{n-1}^2(a)$$

Αντικαθιστώντας βρίσκουμε ότι

$$\frac{(n-1)S^2}{\sigma_0^2} = \frac{19 \cdot 130}{100} = 24.7 \leq 30.14 = \chi_{19}^2(0.05)$$

και επομένως δεν μπορούμε να απορρίψουμε την υπόθεση $H_0: \sigma^2 = \sigma_0^2 = 100$ σε επίπεδο σημαντικότητας 0.05.

Το p -value του ελέγχου θα είναι η πιθανότητα

$$p\text{-value} = P\left(\frac{(n-1)S^2}{\sigma_0^2} > 24.7 / H_0\right) = 1 - F_{\chi_{19}^2}(24.7) \approx 0.17$$

Παρατηρούμε ότι το p -value είναι αρκετά μεγάλο, μεγαλύτερο του $\alpha=0.05$ (και άρα σωστά δεχτήκαμε την H_0).

ε. Έλεγχος υποθέσεων για ποσοστό.

Έστω $X_1, X_2, \dots, X_n$ τ.δ. από την $B(1, p)$, δηλαδή $P(X_i=1)=p$, $P(X_i=0)=1-p$ ή συνοπτικά $f(x; p) = P(X=x) = p^x(1-p)^{1-x}$, $x=0,1$. Επιθυμούμε αρχικά να κατασκευάσουμε έναν έλεγχο για την υπόθεση

$$H_0: p = p_0 \quad \text{έναντι της} \quad H_1: p = p_1, \quad p_1 > p_0.$$

Πρόκειται για έλεγχο απλής έναντι απλούς και συνεπώς από το λήμμα N-P η κρίσιμη περιοχή θα έχει μορφή

$$\begin{aligned} K: \lambda(\mathbf{x}) &= \frac{\prod_{i=1}^n f(x_i; p_0)}{\prod_{i=1}^n f(x_i; p_1)} = \frac{\prod_{i=1}^n p_0^{x_i} (1-p_0)^{1-x_i}}{\prod_{i=1}^n p_1^{x_i} (1-p_1)^{1-x_i}} = \frac{\prod_{i=1}^n \left(\frac{p_0}{1-p_0}\right)^{x_i} (1-p_0)}{\prod_{i=1}^n \left(\frac{p_1}{1-p_1}\right)^{x_i} (1-p_1)} \\ &= \frac{\left(\frac{p_0}{1-p_0}\right)^{\sum_{i=1}^n x_i} p_0^{1-n}}{\left(\frac{p_1}{1-p_1}\right)^{\sum_{i=1}^n x_i} (1-p_1)^n} = \left(\frac{p_0(1-p_1)}{p_1(1-p_0)}\right)^{\sum_{i=1}^n x_i} \left(\frac{1-p_0}{1-p_1}\right)^n < c \end{aligned}$$

ή ισοδύναμα,

$$\sum_{i=1}^n x_i \ln \frac{p_0(1-p_1)}{p_1(1-p_0)} < \ln c - \ln \left(\frac{1-p_0}{1-p_1} \right)^n$$

Επειδή $p_1 > p_0$ θα είναι $\frac{p_0(1-p_1)}{p_1(1-p_0)} < 1 \Leftrightarrow \ln \frac{p_0(1-p_1)}{p_1(1-p_0)} < 0$ και άρα η κρίσιμη περιοχή θα έχει την μορφή

$$\sum_{i=1}^n X_i > \frac{\ln c - \ln \left(\frac{1-p_0}{1-p_1} \right)^n}{\ln \frac{p_0(1-p_1)}{p_1(1-p_0)}} \equiv c_0$$

Σύμφωνα με τα παραπάνω, το c_0 λαμβάνεται τέτοιο ώστε να ισχύει $P(I)=a$. Συνεπώς,

$$P(I) = P\left(\sum_{i=1}^n X_i > c_0 / H_0\right) = P\left(\frac{\sum_{i=1}^n X_i - np_0}{\sqrt{np_0(1-p_0)}} > \frac{c_0 - np_0}{\sqrt{np_0(1-p_0)}} / H_0\right) = a$$

και επειδή, κάτω από την H_0 , και για μεγάλο n (Κ.Ο.Θ., $n > 30$) η τ.μ.

$$Z = \frac{\sum_{i=1}^n X_i - np_0}{\sqrt{np_0(1-p_0)}} \sim N(0,1)$$

συμπεραίνουμε τελικά ότι

$$\frac{c_0 - np_0}{\sqrt{np_0(1-p_0)}} = Z_a \Rightarrow c_0 = np_0 + \sqrt{np_0(1-p_0)} Z_a.$$

Άρα, σύμφωνα με το Λήμμα N-P, ο καλύτερος έλεγχος σε επίπεδο σημαντικότητας α για την υποθέση $H_0: p=p_0$ έναντι της $H_1: p=p_1, p_1 > p_0$, θα έχει περιοχή απόρριψης ($n > 30$):

$$\sum_{i=1}^n X_i > np_0 + \sqrt{np_0(1-p_0)} Z_a \quad \text{ή ισοδύναμα} \quad K: \frac{\bar{X} - p_0}{\sqrt{p_0(1-p_0)/n}} > Z_a.$$

Η πιθανότητα σφάλματος τύπου I αυτού του ελέγχου θα είναι φυσικά ίση με α ενώ η πιθανότητα σφάλματος τύπου II θα είναι ίση με

$$\begin{aligned} P(II) &= P\left(\frac{\bar{X} - p_0}{\sqrt{p_0(1-p_0)/n}} \leq Z_a / H_1\right) = P\left(\bar{X} \leq p_0 + \sqrt{\frac{p_0(1-p_0)}{n}} Z_a \mid p=p_1\right) \\ &= P\left(\frac{\bar{X} - p_1}{\sqrt{p_1(1-p_1)/n}} \leq \frac{p_0 - p_1}{\sqrt{p_1(1-p_1)/n}} + \frac{\sqrt{p_0(1-p_0)}}{\sqrt{p_1(1-p_1)}} Z_a \mid p=p_1\right) \\ &\approx \Phi\left(\frac{p_0 - p_1}{\sqrt{p_1(1-p_1)/n}} + \frac{\sqrt{p_0(1-p_0)}}{\sqrt{p_1(1-p_1)}} Z_a\right). \end{aligned}$$

Τέλος, η ισχύς του ελέγχου θα είναι $\pi(p_1) = 1 - P(II)$.

Παρατηρούμε και εδώ ότι η κρίσιμη περιοχή K δεν εξαρτάται από την τιμή p_1 και επομένως η ίδια κρίσιμη περιοχή μπορεί να χρησιμοποιηθεί και για τον έλεγχο της υπόθεσης $H_0: p = p_0$ έναντι της $H_1: p > p_0$ (απλή έναντι μονόπλευρης). Αποδεικνύεται επίσης ότι η ίδια κρίσιμη περιοχή K οδηγεί σε βέλτιστο έλεγχο της υπόθεσης $H_0: p \leq p_0$ έναντι της $H_1: p > p_0$ (μονόπλευρη έναντι μονόπλευρης).

Αντίστοιχα με τα ζεύγη των υποθέσεων

$$H_0: p=p_0, H_1: p=p_1, p_1 > p_0 \quad \text{ή} \quad H_0: p=p_0, H_1: p > p_0, \quad \text{ή} \quad H_0: p \leq p_0, H_1: p > p_0$$

μπορούμε να αντιμετωπίσουμε και τα ζεύγη

$$H_0: p=p_0, H_1: p=p_1, p_1 < p_0 \quad \text{ή} \quad H_0: p=p_0, H_1: p < p_0 \quad \text{ή} \quad H_0: p \geq p_0, H_1: p < p_0.$$

Σε αυτές τις περιπτώσεις, αν εργαστούμε όμοια με παραπάνω καταλήγουμε στην κρίσιμη περιοχή ($n > 30$)

$$K: \frac{\bar{X} - p_0}{\sqrt{p_0(1-p_0)/n}} < -Z_\alpha.$$

η οποία ορίζει έναν βέλτιστο έλεγχο (με τη μεγαλύτερη δυνατή ισχύ) σε επίπεδο σημαντικότητας α .

Τέλος, θα εξετάσουμε την περίπτωση $H_0: p=p_0, H_1: p \neq p_0$, (απλή έναντι αμφίπλευρης υπόθεσης). Όπως και παραπάνω, εδώ θα χρησιμοποιήσουμε το κριτήριο του γενικευμένου λόγου πιθανοφανειών. Σύμφωνα με αυτό, η κρίσιμη περιοχή θα έχει τελικά τη μορφή ($n > 30$)

$$K: \frac{|\bar{X} - p_0|}{\sqrt{p_0(1-p_0)/n}} > Z_{\alpha/2}.$$

Αξίζει εδώ να παρατηρήσουμε ότι ο παραπάνω αμφίπλευρος έλεγχος σε επίπεδο σημαντικότητας α δεν είναι ακριβώς ισοδύναμος με τον έλεγχο μέσω του διαστήματος εμπιστοσύνης για το p συντελεστού $1-\alpha$. Αυτό συμβαίνει διότι στην κατασκευή του δ.ε. για το p είχαμε για λόγους ευκολίας αντικαταστήσει το $p(1-p)$ με το $\bar{X}(1-\bar{X})$. Για το λόγο αυτό ο έλεγχος της υπόθεσης $H_0: p=p_0$, έναντι της $H_1: p \neq p_0$, είναι προτιμότερο να γίνεται χρησιμοποιώντας την παραπάνω κρίσιμη περιοχή και όχι μέσω δ.ε.

Άσκηση 6.6. Ένα κόμμα ισχυρίζεται ότι στις εκλογές που έγιναν σήμερα θα λάβει ποσοστό p μεγαλύτερο του 10%. Έχουμε αρκετά στοιχεία για να απορρίψουμε τον παραπάνω ισχυρισμό, αν σε ένα τ.δ. $n=200$ ψηφοφόρων (οι οποίοι ρωτήθηκαν μετά την έξοδό τους από τυχαία επιλεγμένα εκλογικά τμήματα) μόνο το 6% απάντησαν ότι ψήφισαν το κόμμα. ($\alpha=0.01$). Αν βρεθεί ότι δεν έχουμε αρκετά στοιχεία, πόσο θα έπρεπε να είναι το ελάχιστο μέγεθος του δείγματος ώστε, με ίδιο δειγματικό ποσοστό, να απορρίπταμε τον παραπάνω ισχυρισμό;

Λύση. Στη συγκεκριμένη περίπτωση ζητείται ο έλεγχος της υπόθεσης

$$H_0: p=p_0 \text{ έναντι της } H_1: p < p_0$$

όπου $p_0=10\%$. Γνωρίζουμε από τα παραπάνω ότι η κρίσιμη περιοχή θα είναι

$$K: \frac{\bar{X} - p_0}{\sqrt{p_0(1-p_0)/n}} < -Z_\alpha.$$

Αντικαθιστώντας βρίσκουμε ότι

$$\frac{\bar{X} - p_0}{\sqrt{p_0(1-p_0)/n}} = \frac{0.06-0.1}{\sqrt{0.1(1-0.1)/200}} = -1.88 > -Z_{0.01} = -2.33.$$

Επομένως δεν έχουμε αρκετά στοιχεία για να απορρίψουμε τον παραπάνω ισχυρισμό.

Παρατηρούμε ότι για να απορρίπταμε την $H_0: p=0.1$ λαμβάνοντας ένα μεγαλύτερο δείγμα n' θα έπρεπε να ισχύει ότι

$$\frac{\bar{X}' - p_0}{\sqrt{p_0(1-p_0)/n'}} < -Z_\alpha \text{ δηλαδή } n' > \frac{p_0(1-p_0)}{(\bar{X}' - p_0)^2} Z_\alpha^2$$

όπου $\bar{X}'$ είναι το αντίστοιχο δειγματικό ποσοστό που θα βρίσκαμε. Βρίσκοντας, σύμφωνα με την εκφώνηση, και πάλι $\bar{X}'=0.06$ θα έπρεπε

$$n' > \frac{p_0(1-p_0)}{(\bar{X}' - p_0)^2} Z_a^2 = \frac{0.1(1-0.1)}{(0.06-0.1)^2} 2.33^2 = 305.37$$

Συνεπώς για να απορρίψουμε την $H_0: p=10\%$ έναντι της $H_1: p<10\%$ σε επίπεδο σημαντικότητας 1% θα έπρεπε να είχαμε βρει το δειγματικό ποσοστό 6% από ένα δείγμα τουλάχιστον 306 ψηφοφόρων.

Άσκηση 6.7. Έστω ότι έχουμε δύο όμοιες εξωτερικά κάλπες Α και Β. Η κάλπη Α περιέχει 6 άσπρες και 4 μαύρες σφαίρες ενώ η κάλπη Β περιέχει 4 άσπρες και 6 μαύρες σφαίρες. Εκλέγουμε διαδοχικά $n=100$ σφαίρες με επανάθεση από τη μία κάλπη χωρίς να γνωρίζουμε αν είναι η κάλπη Α ή η Β. i) Αν 52 από τις 100 σφαίρες βρέθηκαν μαύρες, μπορούμε να απορρίψουμε σε επίπεδο σημαντικότητας $\alpha=0.01$ ότι οι σφαίρες που πήραμε ήταν από την κάλπη Α; Να βρεθεί η ισχύς του ελέγχου. ii) Πόσο θα έπρεπε να είναι το n ώστε οι πιθανότητες των σφαλμάτων τύπου Ι και ΙΙ να είναι ακριβώς α και β αντίστοιχα.

Λύση. i) Έστω $X_i=1$ αν ή i -εκλογή σφαίρας είναι μαύρη και $X_i=0$ αν ή i -εκλογή σφαίρας είναι λευκή, $i=1,2,\dots,n$. Δηλαδή οι τ.μ. $X_1, X_2, \dots, X_n$ αποτελούν ένα τ.δ. από την $B(1, p)$. Στην ουσία εδώ ζητείται ο έλεγχος της υπόθεσης

$$H_0: p=p_A=4/10 \text{ έναντι της } H_1: p=p_B=6/10$$

Γνωρίζουμε από τα παραπάνω ότι η κρίσιμη περιοχή αυτού του ελέγχου θα είναι η

$$K: \frac{\bar{X} - p_A}{\sqrt{p_A(1-p_A)/n}} > Z_a.$$

Αντικαθιστώντας βρίσκουμε ότι

$$\frac{\bar{X} - p_A}{\sqrt{p_A(1-p_A)/n}} = \frac{0.52-0.4}{\sqrt{0.4(1-0.4)/100}} = 2.45 > Z_{0.01} = 2.33$$

και συνεπώς απορρίπτουμε την $H_0: p = p_A = 0.4$ με πιθανότητα να κάνουμε λάθος ίση με $\alpha = 1\%$. Η ισχύς του ελέγχου είναι ίση με

$$\begin{aligned} \pi(p) &= 1 - P(II) = P(\text{απόρριψη της } H_0 / \text{ισχύει ή } H_1) = P\left(\frac{\bar{X} - p_A}{\sqrt{p_A(1-p_A)/n}} > Z_a / p = 0.6\right) \\ &= P\left(\bar{X} > p_A + \sqrt{\frac{p_A(1-p_A)}{n}} Z_a / p = p_B\right) \\ &= P\left(\frac{\bar{X} - p_B}{\sqrt{p_B(1-p_B)/n}} > \frac{p_A - p_B}{\sqrt{p_B(1-p_B)/n}} + \frac{\sqrt{p_A(1-p_A)}}{\sqrt{p_B(1-p_B)}} Z_a / p = p_B\right) \\ &= 1 - \Phi\left(\frac{p_A - p_B}{\sqrt{p_B(1-p_B)/n}} + \frac{\sqrt{p_A(1-p_A)}}{\sqrt{p_B(1-p_B)}} Z_a\right) = 1 - \Phi\left(\frac{0.4-0.6}{\sqrt{0.6(1-0.6)/100}} + 2.33\right) \\ &= 1 - \Phi(-1.75) = \Phi(1.75) = 0.96. \end{aligned}$$

Παρατηρούμε λοιπόν ότι αν $n=100$, η $P(II) = 1 - 0.96 = 4\%$.

ii) Έχουμε τον έλεγχο της υπόθεσης

$$H_0: p=p_A=4/10 \text{ έναντι της } H_1: p=p_B=6/10$$

Γνωρίζουμε ότι αν επιλέξουμε ως κρίσιμη περιοχή την

$$K: \frac{\bar{X} - p_A}{\sqrt{p_A(1-p_A)/n}} > Z_\alpha$$

τότε η $P(I)$ θα είναι ίση με α . Για να ισχύει και $P(II)=\beta$ θα πρέπει (βλ. (i)) ότι

$$P(II)=\Phi\left(\frac{p_A - p_B}{\sqrt{p_B(1-p_B)/n}} + \frac{\sqrt{p_A(1-p_A)}}{\sqrt{p_B(1-p_B)}} Z_\alpha\right) = \beta$$

δηλαδή

$$\frac{p_A - p_B}{\sqrt{p_B(1-p_B)/n}} + \frac{\sqrt{p_A(1-p_A)}}{\sqrt{p_B(1-p_B)}} Z_\alpha = -Z_\beta$$

Λύνοντας ως προς n βρίσκουμε ότι

$$n = \left(\frac{\sqrt{p_B(1-p_B)} Z_\beta + \sqrt{p_A(1-p_A)} Z_\alpha}{p_B - p_A} \right)^2$$

Αν π.χ. $\alpha=\beta=1\%$ τότε θα πρέπει να πάρουμε δείγμα ($p_A=0.4, p_B=0.6$) μεγέθους $n \approx 130$.

Έλεγχοι υποθέσεων για τις παραμέτρους δύο ανεξάρτητων κανονικών πληθυσμών

Έστω $X_1, X_2, \dots, X_{n_1}$ και $Y_1, Y_2, \dots, Y_{n_2}$ δύο ανεξάρτητα μεταξύ τους τ.δ. από τις κατανομές $N(\mu_1, \sigma_1^2)$ και $N(\mu_2, \sigma_2^2)$ αντίστοιχα. Οι κρίσιμες περιοχές ελέγχων υποθέσεων για τις παραμέτρους δύο ανεξάρτητων κανονικών πληθυσμών προκύπτουν με μεθόδους παρόμοιες με αυτές που χρησιμοποιήθηκαν παραπάνω (είτε με τη χρήση τη λήμματος $N-P$, είτε με τη χρήση του γενικευμένου λόγου πιθανοφανειών). Συνοπτικά προκύπτουν οι παρακάτω έλεγχοι σε επίπεδο σημαντικότητας α .

α. Έλεγχος για τη σύγκριση των μ_1, μ_2 όταν τα σ_1^2, σ_2^2 είναι γνωστά

Θεωρούμε τη στατιστική συνάρτηση

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

η οποία όταν $\mu_1 = \mu_2$ ακολουθεί $N(0,1)$ κατανομή. Είναι εύκολο να δειχθεί ότι στις παρακάτω υποθέσεις αντιστοιχούν οι κρίσιμες περιοχές:

Υποθέσεις	Κρίσιμη περιοχή
$H_0: \mu_1 = \mu_2$ έναντι της $H_1: \mu_1 > \mu_2$	$T > Z_\alpha$
$H_0: \mu_1 = \mu_2$ έναντι της $H_1: \mu_1 < \mu_2$	$T < -Z_\alpha$
$H_0: \mu_1 \leq \mu_2$ έναντι της $H_1: \mu_1 > \mu_2$	$T > Z_\alpha$
$H_0: \mu_1 \geq \mu_2$ έναντι της $H_1: \mu_1 < \mu_2$	$T < -Z_\alpha$
$H_0: \mu_1 = \mu_2$ έναντι της $H_1: \mu_1 \neq \mu_2$	$ T > Z_{\alpha/2}$

Στην περίπτωση που θέλουμε να ελέγξουμε την υπόθεση $H_0: \mu_1 - \mu_2 = \gamma$, τότε χρησιμοποιούμε τη στατιστική συνάρτηση

$$T = \frac{\bar{X} - \bar{Y} - \gamma}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}.$$

β. Έλεγχος για τη σύγκριση των μ_1, μ_2 όταν τα σ_1^2, σ_2^2 είναι άγνωστα και ίσα

Θεωρούμε τη στατιστική συνάρτηση

$$T = \frac{\sqrt{\frac{n_1 n_2}{n_1 + n_2}} (\bar{X} - \bar{Y})}{\sqrt{\frac{\sum_{i=1}^{n_1} (X_i - \bar{X})^2 + \sum_{j=1}^{n_2} (Y_j - \bar{Y})^2}{n_1 + n_2 - 2}}} = \frac{\sqrt{\frac{n_1 n_2}{n_1 + n_2}} (\bar{X} - \bar{Y})}{\sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}}.$$

η οποία όταν $\mu_1 = \mu_2$ ακολουθεί κατανομή $t_{n_1+n_2-2}$. Είναι εύκολο να δειχθεί ότι στις παρακάτω υποθέσεις αντιστοιχούν οι κρίσιμες περιοχές:

Υποθέσεις	Κρίσιμη περιοχή
$H_0 : \mu_1 = \mu_2$ έναντι της $H_1 : \mu_1 > \mu_2$	$T > t_{n_1+n_2-2}(a)$
$H_0 : \mu_1 = \mu_2$ έναντι της $H_1 : \mu_1 < \mu_2$	$T < -t_{n_1+n_2-2}(a)$
$H_0 : \mu_1 \leq \mu_2$ έναντι της $H_1 : \mu_1 > \mu_2$	$T > t_{n_1+n_2-2}(a)$
$H_0 : \mu_1 \geq \mu_2$ έναντι της $H_1 : \mu_1 < \mu_2$	$T < -t_{n_1+n_2-2}(a)$
$H_0 : \mu_1 = \mu_2$ έναντι της $H_1 : \mu_1 \neq \mu_2$	$ T > t_{n_1+n_2-2}(a)$

Στην περίπτωση που θέλουμε να ελέγξουμε την υπόθεση $H_0 : \mu_1 - \mu_2 = \gamma$, τότε χρησιμοποιούμε τη στατιστική συνάρτηση

$$T = \frac{\sqrt{\frac{n_1 n_2}{n_1 + n_2}} (\bar{X} - \bar{Y} - \gamma)}{\sqrt{\frac{\sum_{i=1}^{n_1} (X_i - \bar{X})^2 + \sum_{j=1}^{n_2} (Y_j - \bar{Y})^2}{n_1 + n_2 - 2}}} = \frac{\sqrt{\frac{n_1 n_2}{n_1 + n_2}} (\bar{X} - \bar{Y} - \gamma)}{\sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}}.$$

γ. Έλεγχος για τη σύγκριση των σ_1^2, σ_2^2 όταν τα μ_1, μ_2 είναι γνωστά

Θεωρούμε τη στατιστική συνάρτηση

$$T = \frac{n_2 \sum_{i=1}^{n_1} (X_i - \mu_1)^2}{n_1 \sum_{j=1}^{n_2} (Y_j - \mu_2)^2}$$

η οποία όταν $\sigma_1 = \sigma_2$ ακολουθεί κατανομή F_{n_1, n_2} . Είναι εύκολο να δειχθεί ότι στις παρακάτω υποθέσεις αντιστοιχούν οι κρίσιμες περιοχές:

Υποθέσεις	Κρίσιμη περιοχή
$H_0 : \sigma_1^2 = \sigma_2^2$ έναντι της $H_1 : \sigma_1^2 > \sigma_2^2$	$T > F_{n_1, n_2}(a)$

$$\begin{array}{ll}
H_0: \sigma_1^2 = \sigma_2^2 \text{ \acute{e}\nu\alpha\nu\tau\iota \tau\eta\varsigma } H_1: \sigma_1^2 < \sigma_2^2 & T < F_{n_1, n_2} (1-a) \\
H_0: \sigma_1^2 \leq \sigma_2^2 \text{ \acute{e}\nu\alpha\nu\tau\iota \tau\eta\varsigma } H_1: \sigma_1^2 > \sigma_2^2 & T > F_{n_1, n_2} (a) \\
H_0: \sigma_1^2 \geq \sigma_2^2 \text{ \acute{e}\nu\alpha\nu\tau\iota \tau\eta\varsigma } H_1: \sigma_1^2 < \sigma_2^2 & T < F_{n_1, n_2} (1-a) \\
H_0: \sigma_1^2 = \sigma_2^2 \text{ \acute{e}\nu\alpha\nu\tau\iota \tau\eta\varsigma } H_1: \sigma_1^2 \neq \sigma_2^2 & T < F_{n_1, n_2} (1-a/2) \text{ \acute{\eta} } T > F_{n_1, n_2} (a/2)
\end{array}$$

Όταν θέλουμε να ελέγξουμε την υπόθεση $H_0: \sigma_1^2 / \sigma_2^2 = \delta^2$, χρησιμοποιούμε τη στατιστική συνάρτηση

$$T = \frac{1}{\delta^2} \frac{n_2}{n_1} \frac{\sum_{i=1}^{n_1} (X_i - \mu_1)^2}{\sum_{j=1}^{n_2} (Y_j - \mu_2)^2}.$$

δ. Έλεγχος για τη σύγκριση των σ_1^2, σ_2^2 όταν τα μ_1, μ_2 είναι άγνωστα

Θεωρούμε τη στατιστική συνάρτηση

$$T = \frac{n_2 - 1}{n_1 - 1} \frac{\sum_{i=1}^{n_1} (X_i - \bar{X})^2}{\sum_{j=1}^{n_2} (Y_j - \bar{Y})^2} = \frac{S_1^2}{S_2^2}.$$

η οποία όταν $\sigma_1 = \sigma_2$ ακολουθεί κατανομή F_{n_1-1, n_2-1} . Είναι εύκολο να δειχθεί ότι στις παρακάτω υποθέσεις αντιστοιχούν οι κρίσιμες περιοχές:

Υποθέσεις	Κρίσιμη περιοχή
$H_0: \sigma_1^2 = \sigma_2^2 \text{ \acute{e}\nu\alpha\nu\tau\iota \tau\eta\varsigma } H_1: \sigma_1^2 > \sigma_2^2$	$T > F_{n_1-1, n_2-1} (a)$
$H_0: \sigma_1^2 = \sigma_2^2 \text{ \acute{e}\nu\alpha\nu\tau\iota \tau\eta\varsigma } H_1: \sigma_1^2 < \sigma_2^2$	$T < F_{n_1-1, n_2-1} (1-a)$
$H_0: \sigma_1^2 \leq \sigma_2^2 \text{ \acute{e}\nu\alpha\nu\tau\iota \tau\eta\varsigma } H_1: \sigma_1^2 > \sigma_2^2$	$T > F_{n_1-1, n_2-1} (a)$
$H_0: \sigma_1^2 \geq \sigma_2^2 \text{ \acute{e}\nu\alpha\nu\tau\iota \tau\eta\varsigma } H_1: \sigma_1^2 < \sigma_2^2$	$T < F_{n_1-1, n_2-1} (1-a)$
$H_0: \sigma_1^2 = \sigma_2^2 \text{ \acute{e}\nu\alpha\nu\tau\iota \tau\eta\varsigma } H_1: \sigma_1^2 \neq \sigma_2^2$	$T < F_{n_1-1, n_2-1} (1-a/2) \text{ \acute{\eta} } T > F_{n_1-1, n_2-1} (a/2)$

Όταν θέλουμε να ελέγξουμε την υπόθεση $H_0: \sigma_1^2 / \sigma_2^2 = \delta^2$, χρησιμοποιούμε τη στατιστική συνάρτηση

$$T = \frac{1}{\delta^2} \frac{n_2 - 1}{n_1 - 1} \frac{\sum_{i=1}^{n_1} (X_i - \bar{X})^2}{\sum_{j=1}^{n_2} (Y_j - \bar{Y})^2} = \frac{1}{\delta^2} \frac{S_1^2}{S_2^2}.$$

ε. Έλεγχος για τη σύγκριση δύο ποσοστών

Τέλος, έστω $X_1, X_2, \dots, X_{n_1}$ και $Y_1, Y_2, \dots, Y_{n_2}$ δύο ανεξάρτητα μεταξύ τους τ.δ. από τις κατανομές $B(1, p_1)$ και $B(1, p_2)$ αντίστοιχα. Επιθυμούμε να συγκρίνουμε τα ποσοστά p_1, p_2 . Ακολουθώντας μεθόδους παρόμοιες με αυτές που χρησιμοποιήθηκαν παραπάνω προκύπτουν οι ακόλουθοι έλεγχοι σε επίπεδο σημαντικότητας α . Θεωρώντας τη στατιστική συνάρτηση ($n_1, n_2 > 30$):

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{P(1-P)(1/n_1 + 1/n_2)}}, \quad P = \frac{n_1 \bar{X} + n_2 \bar{Y}}{n_1 + n_2}$$

η οποία όταν $p_1 = p_2$ προσεγγιστικά ακολουθεί κατανομή $N(0,1)$. Είναι εύκολο να δειχθεί ότι στις παρακάτω υποθέσεις αντιστοιχούν οι κρίσιμες περιοχές:

Υποθέσεις	Κρίσιμη περιοχή
$H_0 : p_1 = p_2$ έναντι της $H_1 : p_1 > p_2$ ή $H_0 : p_1 \leq p_2$ έναντι της $H_1 : p_1 > p_2$	$K: T > Z_\alpha$
$H_0 : p_1 = p_2$ έναντι της $H_1 : p_1 < p_2$ ή $H_0 : p_1 \geq p_2$ έναντι της $H_1 : p_1 < p_2$	$K: T < -Z_\alpha$
$H_0 : p_1 = p_2$ έναντι της $H_1 : p_1 \neq p_2$	$K: T > Z_{\alpha/2}$