ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ & ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ

ΣΗΜΕΙΩΣΕΙΣ ΠΑΡΑΔΟΣΕΩΝ ΤΟΥ ΜΑΘΗΜΑΤΟΣ

«ΣΤΑΤΙΣΤΙΚΗ ΙΙ»

Μ. Κούτρας – Μ. Μπούτσικας

ΙΑΝΟΥΑΡΙΟΣ 2011

Εισαγωγή

Κύρια επιδίωξη της *Στατιστικής* ΙΙ αποτελεί ο **έλεγχος μιας υπόθεσης** σχετικά με τις παραμέτρους της κατανομής κάποιου χαρακτηριστικού ενός πληθυσμού. Πιο συγκεκριμένα, συνήθως:

• Υποθέτουμε ότι το χαρακτηριστικό X ενός πληθυσμού που μας ενδιαφέρει ακολουθεί κάποια *γνωστή* κατανομή $F(x;\boldsymbol{\theta})$ με *άγνωστες* όμως παραμέτρους $\boldsymbol{\theta} = (\theta_1,\theta_2,...,\theta_k)$ (π.χ. κανονική κατανομή $N(\mu,\sigma^2)$)

• Με βάση ένα τυχαίο δείγμα $X_1, X_2, \ldots, X_\nu$ του χαρακτηριστικού αυτού από τον πληθυσμό, επιθυμούμε να ελέγξουμε αν ισχύει κάποια υπόθεση σχετικά με τις παραμέτρους $\boldsymbol{\theta}$ (π.χ. να ελέγξουμε αν $\mu = 100$ ή $\mu > 100$).

Πριν περάσουμε στην μελέτη του παραπάνω προβλήματος, γίνεται μια σύντομη επισκόπηση κάποιων βασικών εννοιών της Στατιστικής που θα θεωρούνται γνωστές στη συνέχεια.

ΣΥΝΤΟΜΗ ΕΠΙΣΚΟΠΗΣΗ ΒΑΣΙΚΩΝ ΕΝΝΟΙΩΝ ΤΗΣ ΕΚΤΙΜΗΤΙΚΗΣ

1. Βασικοί ορισμοί

Ορισμός 1. *Τυχαίο δείγμα* μεγέθους ν από την κατανομή $F(x;\boldsymbol{\theta})$ (ή τη σ.π. ή σ.π.π. $f(x;\boldsymbol{\theta})$) καλείται μια συλλογή ανεξάρτητων και ισόνομων (αν.ισ. ή iid) τυχαίων μεταβλητών $X_1, X_2, ..., X_\nu \sim F(x;\boldsymbol{\theta})$ (ή $\sim f(x;\boldsymbol{\theta})$).

Ορισμός 2. *Δειγματοληπτικός χώρος* καλείται το σύνολο των δυνατών τιμών του δείγματος (π.χ. αν $X_i \in \mathbf{R}$, τότε ο δειγματοληπτικός χώρος είναι ο $\mathbf{R}^\nu$). Επίσης, *παραμετρικός χώρος* Θ καλείται το σύνολο των επιτρεπτών τιμών των παραμέτρων $\boldsymbol{\theta}$ (π.χ. αν $\boldsymbol{\theta} = (\mu,\sigma^2)$ τότε ο παραμετρικός χώρος είναι ο $\mathbf{R}\times(0,\infty)$).

Ορισμός 3. Έστω τυχαίο δείγμα $X_1,X_2,...,X_\nu \sim F(x;\boldsymbol{\theta})$. *Στατιστική* (*ή δειγματική*) *συνάρτηση* (σ.σ.) καλείται κάθε συνάρτηση

$$T = T(\mathbf{X}) = T(X_1,X_2, ..., X_\nu)$$

των $X_1,X_2,...,X_\nu$, που δεν εξαρτάται από άγνωστες παραμέτρους.

Οι πλέον συνήθεις στατιστικές συναρτήσεις που εμφανίζονται σε προβλήματα της Στατιστικής είναι οι επόμενες:

$$\overline{X} = \frac{1}{\nu}\sum_{i=1}^{\nu} X_i \quad (\textit{δειγματικός μέσος})$$

$$m_r = \frac{1}{\nu}\sum_{i=1}^{\nu} X_i^r \quad (\textit{δειγματικές ροπές τάξεως } r \geq 1)$$

$$S^2 = \frac{1}{\nu-1}\sum_{i=1}^{\nu}(X_i - \overline{X})^2 = \frac{1}{\nu-1}(\sum_{i=1}^{\nu} X_i^2 - \nu\,\overline{X}^2) \quad (\textit{δειγματική διασπορά})$$

$$R = \max\{X_1,...,X_\nu\} - \min\{X_1,...,X_\nu\} = X_{(\nu)} - X_{(1)} \quad (\textit{δειγματικό εύρος})$$

Ορισμός 4. *Εκτιμήτρια συνάρτηση* της παραμέτρου θ (ή των παραμέτρων $\boldsymbol{\theta}$ ή μιας παραμετρικής συνάρτησης g($\boldsymbol{\theta}$)) θα καλείται μία σ.σ. $T = T(\mathbf{X})$ η οποία χρησιμοποιείται για την εκτίμηση του θ (ή των $\boldsymbol{\theta}$ ή του g($\boldsymbol{\theta}$)). ΟΙ τιμές που παίρνει μια στατιστική συνάρτηση θα πρέπει να βρίσκονται εντός του παραμετρικού χώρου Θ στον οποίο κινείται η παράμετρος θ (ή εντός του $g(\Theta)$ αν αυτό που μας ενδιαφέρει είναι η παραμετρικής συνάρτησης g($\boldsymbol{\theta}$))

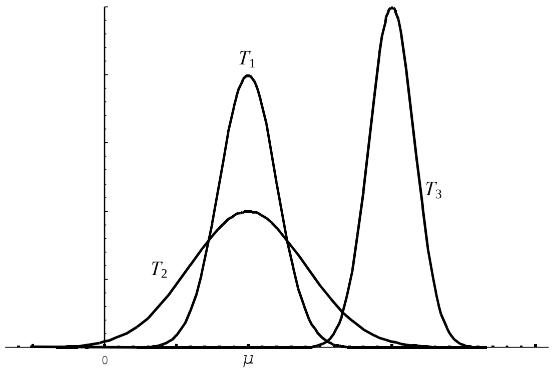
2. Ιδιότητες εκτιμητριών

Προφανώς, μία εκτιμήτρια συνάρτηση $T(\mathbf{X})$ είναι και αυτή μία τυχαία μεταβλητή (κάθε φορά που παίρνουμε ένα άλλο τυχαίο δείγμα **X** η T δίνει διαφορετική τιμή). Είναι επίσης φανερό ότι από ένα τυχαίο δείγμα **X** μπορούμε να κατασκευάσουμε πολλές εκτιμήτριες για μια παράμετρο θ (π.χ. οι $\overline{X}$, X_1, $(X_{(\nu)} - X_{(1)})/2$, θα μπορούσαν να προταθούν ως εκτιμήτριες του μέσου μ ενός κανονικού πληθυσμού). Το ερώτημα είναι ποιές θεωρούνται «καλές» ή ποιά είναι η «καλύτερη» από όλες. Διαισθητικά, αναμένουμε ότι μια «καλή εκτιμήτρια» T του θ θα λαμβάνει τιμές

α. «γύρω» από το θ και

β. «κοντά» στο θ με «μεγάλη» πιθανότητα.

Για το (α) απαιτούμε η τυχαία μεταβλητή T να έχει **μέση τιμή** θ ή «σχεδόν» θ και για το (β) απαιτούμε η τυχαία μεταβλητή T να έχει **«μικρή» διασπορά**. Π.χ. αν έχουμε τρεις εκτιμήτριες T_1, T_2, T_3 για τη μέση τιμή μ ενός πληθυσμού με συναρτήσεις πυκνότητας πιθανότητας (σ.π.π.) της μορφής:



τότε θα προτιμήσουμε την T_1 γιατί παίρνει τιμές κοντά στο μ με μεγαλύτερη πιθανότητα από ό,τι οι T_2, T_3. Η T_2 παίρνει τιμές γύρω από το μ αλλά αυτές μπορεί να διαφέρουν αρκετά από το μ (έχει μεγαλύτερη διασπορά από την T_1), ενώ η T_3 παίρνει τιμές που είναι μακριά από το μ (αν και έχει μικρή διασπορά, έχει μέση τιμή μακριά από το μ).

3. Αμερόληπτες εκτιμήτριες

Ορισμός 5. Μία εκτιμήτρια συνάρτηση T της $g(\boldsymbol{\theta})$ καλείται *αμερόληπτη εκτιμήτρια* (α.ε.) εάν

$$E(T) = E(T(\mathbf{X})) = g(\boldsymbol{\theta}) \text{ για κάθε } \boldsymbol{\theta}.$$

Η T θα θεωρείται *ασυμπτωτικά αμερόληπτη* εάν $\lim_{\nu\to\infty} E(T(\mathbf{X})) = g(\boldsymbol{\theta})$ για κάθε $\boldsymbol{\theta}$. Επίσης, η διαφορά

$$bias(T) = E(T) - g(\boldsymbol{\theta})$$

καλείται *μεροληψία* της εκτιμήτριας T. Η μεροληψία μιας α.ε. είναι 0.

Είναι φανερό ότι μας συμφέρει να χρησιμοποιήσουμε αμερόληπτες ή σχεδόν αμερόληπτες εκτιμήτριες γιατί διαφορετικά έχουμε υποεκτίμηση ή υπερεκτίμηση του $g(\boldsymbol{\theta})$.

Αν έχουμε ένα τυχαίο δείγμα $X_1, X_2, ..., X_\nu \sim F(x;\boldsymbol{\theta})$ με *μέση τιμή* μ (=μ($\boldsymbol{\theta}$)) και *διασπορά* σ^2 (=$\sigma^2(\boldsymbol{\theta})$) τότε αποδεικνύονται οι ακόλουθες προτάσεις:

Πρόταση 1. *Ο δειγματικός μέσος* $\overline{X}$ *είναι α.ε. της μέσης τιμής μ με διασπορά* $V(\overline{X}) = \sigma^2/\nu$.

Πρόταση 2. *Η δειγματική διασπορά* S^2 *είναι α.ε. της διασποράς* σ^2.

Για τη σύγκριση μεταξύ εκτιμητριών δίνονται οι ακόλουθοι ορισμοί.

Ορισμός 6. Αν T_1, T_2 είναι δύο α.ε. της g($\boldsymbol{\theta}$), η T_1 θα καλείται *αποτελεσματικότερη* της T_2 εάν ισχύει ότι $V(T_1) < V(T_2)$.

Ορισμός 7. Αν μία α.ε. $T = T(\mathbf{X})$ έχει (για κάθε $\boldsymbol{\theta}$) τη μικρότερη διασπορά μεταξύ όλων των α.ε. του $g(\boldsymbol{\theta})$ (που μπορούν να κατασκευαστούν από το τυχαίο δείγμα $\mathbf{X}$) τότε θα καλείται *άριστη* ή *αμερόληπτη ομοιόμορφα ελαχίστης διασποράς* (Α.Ο.Ε.Δ.) *εκτιμήτρια* του $g(\boldsymbol{\theta})$.

Προφανώς, αν T_1 είναι Α.Ο.Ε.Δ. και T_2 είναι μια άλλη α.ε. του $g(\boldsymbol{\theta})$, τότε $V(T_1) \leq V(T_2)$. Συνήθως, προκειμένου να εκτιμήσουμε το $g(\boldsymbol{\theta})$, ως βέλτιστη εκτιμήτρια θεωρείται η Α.Ο.Ε.Δ. εκτιμήτρια (αν υπάρχει).

Ορισμός 8. Αν T είναι μια εκτιμήτρια του $g(\boldsymbol{\theta})$, η ποσότητα

$$mse(T) = E((T - g(\boldsymbol{\theta}))^2) = V(T) + bias(T)^2$$

καλείται *μέσο τετραγωνικό σφάλμα* της T από την $g(\boldsymbol{\theta})$.

Ορισμός 9. Αν T_1, T_2 είναι δύο εκτιμήτριες της $g(\boldsymbol{\theta})$, η T_1 θα καλείται *αποτελεσματικότερη* της T_2 εάν ισχύει ότι $mse(T_1) < mse(T_2)$.

Από ένα σύνολο εκτιμητριών του $g(\boldsymbol{\theta})$, καλύτερη θεωρείται αυτή που έχει το μικρότερο μέσο τετραγωνικό σφάλμα.

4. Συνέπεια

Μια ακόμη ιδιότητα που διαισθητικά αναμένουμε να έχει μια «καλή» εκτιμήτρια είναι αυτή της συνέπειας. Μια συνεπής εκτιμήτρια βελτιώνεται με την αύξηση του μεγέθους του δείγματος, και για πολύ μεγάλο δείγμα γίνεται πρακτικά ίση με την υπό εκτίμηση ποσότητα. Συγκεκριμένα δίνεται ο ακόλουθος ορισμός.

Ορισμός 10. *Μία εκτιμήτρια* $T_\nu = T(X_1,X_2,\ldots,X_\nu)$ *μιας παραμετρικής συνάρτησης* $g(\boldsymbol{\theta})$ *καλείται* ***συνεπής*** *αν ισχύει ότι* $\lim_{\nu\to\infty} P(|T_\nu - g(\boldsymbol{\theta})| < \varepsilon) = 1$ *για κάθε ε > 0, δηλαδή έχουμε* $T_\nu \to g(\boldsymbol{\theta})$ *καθώς* $\nu \to \infty$ (*σύγκλιση κατά πιθανότητα*).

Ένα απλό κριτήριο για τη συνέπεια μιας εκτιμήτριας δίνεται στην ακόλουθη πρόταση.

Πρόταση 3. *Μία εκτιμήτρια* $T_\nu = T(X_1,X_2,\ldots,X_\nu)$ *μιας παραμετρικής συνάρτησης* $g(\boldsymbol{\theta})$ *είναι* ***συνεπής*** *αν ισχύουν οι παρακάτω συνθήκες*

$$\text{i) } \lim_{\nu\to\infty} E(T_\nu) = g(\boldsymbol{\theta}) \ \ (\text{δηλαδή } E(T_\nu) \underset{\nu\to\infty}{\to} g(\boldsymbol{\theta})) \qquad \text{ii) } \lim_{\nu\to\infty} V(T_\nu) = 0.$$

Για παράδειγμα, ο δειγματικός μέσος $\overline{X}_\nu$ ενός τυχαίου δείγματος $X_1,X_2,\ldots,X_\nu$, είναι συνεπής εκτιμήτρια του μέσου μ των X_i.

5. Η μέθοδος μέγιστης πιθανοφάνειας.

Η πιο γνωστή μέθοδος εύρεσης μιας εκτιμήτριας για τις παραμέτρους $\boldsymbol{\theta}$ μιας κατανομής F είναι η μέθοδος μέγιστης πιθανοφάνειας. Η μέθοδος αυτή είναι αρκετά ισχυρή διότι, με μία σχετικά απλή διαδικασία, οδηγεί σε εκτιμήτριες με πολύ καλές ιδιότητες.

Θεωρούμε ένα τυχαίο δείγμα $X_1,X_2,...,X_\nu$ από μία κατανομή με συνάρτηση σ.π. ή σ.π.π. $f(x;\boldsymbol{\theta})$ που εξαρτάται από τις παραμέτρους $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)$, και επιθυμούμε να εκτιμήσουμε το $\boldsymbol{\theta}$.

Ορισμός 11. *Συνάρτηση πιθανοφάνειας ή πιθανοφάνεια* (*Likelihood*) *του δείγματος* $X_1, X_2, ..., X_\nu$ *καλείται η από κοινού σ.π.π ή σ.π. των* $X_1, X_2, ..., X_\nu$ *θεωρούμενη ως συνάρτηση του* $\boldsymbol{\theta}$, *δηλαδή η*

$$L_{\mathbf{x}}(\boldsymbol{\theta}) = f(x_1, x_2, ..., x_\nu; \boldsymbol{\theta}) = \prod_{i=1}^{\nu} f(x_i; \boldsymbol{\theta}).$$

Ορισμός 12. *Μία στατιστική συνάρτηση* $\hat{\boldsymbol{\theta}}$ *καλείται εκτιμήτρια μέγιστης πιθανοφάνειας* (ΕΜΠ ή MLE) *των παραμέτρων* $\boldsymbol{\theta}$ *αν μεγιστοποιεί την πιθανοφάνεια* $L_{\mathbf{X}}(\boldsymbol{\theta})$, *δηλαδή, αν ισχύει ότι*

$$L_{\mathbf{X}}(\hat{\boldsymbol{\theta}}) = \sup_{\boldsymbol{\theta}\in\Theta} L_{\mathbf{X}}(\boldsymbol{\theta}).$$

Συνήθως, αντί να αναζητούμε το σημείο μεγίστου της $L(\boldsymbol{\theta})$, είναι πιο εύκολο να αναζητούμε το σημείο μεγίστου της $l(\boldsymbol{\theta}) = \ln L(\boldsymbol{\theta})$ (έχουν το ίδιο σημείο μεγίστου διότι η συνάρτηση ln είναι αύξουσα).

Αν $k = 1$ (δηλ. $\boldsymbol{\theta} = \theta$) και η συνάρτηση $l(\theta)$ παραγωγίζεται σε ολόκληρο τον παραμετρικό χώρο Θ (και Θ ανοικτό διάστημα) μπορούμε να βρούμε το σημείο μεγίστου μέσα από τη λύση της εξίσωσης $l'(\theta) = 0$, ελέγχοντας παράλληλα ότι η δεύτερη παράγωγος $l''(\theta) < 0$. Αξίζει να αναφέρουμε ότι, υπό συγκεκριμένες συνθήκες ομαλότητας της f, αποδεικνύεται γενικά ότι ασυμπτωτικά (για $\nu\to\infty$), ισχύει $\hat{\theta} \sim N(\theta, I(\theta)^{-1})$, όπου $I(\theta) = E(-l''(\theta))$ (πληροφορία κατά Fisher). Δηλαδή, η ΕΜΠ του θ ακολουθεί ασυμπτωτικά κανονική κατανομή μέση τιμή θ (δηλ. είναι ασυμπτωτικά αμερόληπτη) και διασπορά ίση με τον αντίστροφο της πληροφορίας (δηλ. είναι ασυμπτωτικά Α.Ο.Ε.Δ. εκτιμήτρια αφού η διασπορά της συγκλίνει στο φράγμα Cramer-Rao).

Αντίστοιχα αποτελέσματα ισχύουν και για $k \geq 2$, δηλαδή όταν $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)$. Στην περίπτωση αυτή η ΕΜΠ του $\boldsymbol{\theta}$ συνήθως λαμβάνεται από την λύση του συστήματος k εξισώσεων

$$\frac{\partial l}{\partial \theta_1}(\boldsymbol{\theta}) = 0, \; ..., \; \frac{\partial l}{\partial \theta_k}(\boldsymbol{\theta}) = 0.$$

(δεδομένου ότι η $l(\boldsymbol{\theta})$ είναι παραγωγίσιμη στο Θ, και Θ ανοικτό σύνολο του $\mathbf{R}^k$). Και πάλι, υπό συγκεκριμένες συνθήκες ομαλότητας της f, η ΕΜΠ ακολουθεί ασυμπτωτικά πολυδιάστατη κανονική κατανομή με διάνυσμα μέσων τιμών $\boldsymbol{\theta}$ (δηλ. είναι ασυμπτωτικά αμερόληπτη) και πίνακα διασποράς τον αντίστροφο του $k\times k$ πίνακα πληροφορίας (δηλ. είναι ασυμπτωτικά Α.Ο.Ε.Δ. εκτιμήτρια).

1. ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ ΕΛΕΓΧΩΝ ΥΠΟΘΕΣΕΩΝ

1.1. Εισαγωγή

Σε αρκετές εφαρμογές παρουσιάζεται η ανάγκη λήψης αποφάσεων σχετικά με την κατανομή F του πληθυσμού. Θα πρέπει, βάσει του τυχαίου δείγματος $X_1, X_2, ..., X_\nu \sim F(x;\boldsymbol{\theta})$, να αποφασίσουμε αν ευσταθεί ή όχι μία υπόθεση σχετικά με την κατανομή F ή τις παραμέτρους $\boldsymbol{\theta}$. Για παράδειγμα,

- θέλουμε να ελέγξουμε αν ο μέσος μ του πληθυσμού είναι ίσος με 100 ή μεγαλύτερος.
- θέλουμε να ελέγξουμε αν μια παράμετρος θ είναι ίση με 1 ή όχι.
- θέλουμε να ελέγξουμε αν ένα ποσοστό p είναι μικρότερο ή μεγαλύτερο του 50%

κ.ο.κ. Πριν προχωρήσουμε στον ορισμό των βασικών εννοιών των ελέγχων υποθέσεων, ας δούμε τις έννοιες αυτές μέσα από ένα απλό παράδειγμα.

Παράδειγμα 1. Ένα εργοστάσιο παράγει κάποιες συσκευές και κάθε ώρα γίνεται έλεγχος των ν συσκευών της ωριαίας παραγωγής. Η παραγωγική διαδικασία θεωρείται ότι βρίσκεται μέσα στις προδιαγραφές της αν η πιθανότητα p *παραγωγής ελαττωματικής συσκευής* είναι ≤ 5%. Σε περίπτωση που το p αυξηθεί (p > 5%), θεωρείται ότι υπάρχει κάποιο πρόβλημα, *σταματά η παραγωγή*, και αναζητούνται τα αίτια. Επομένως θα πρέπει να *κατασκευάσουμε έναν έλεγχο* ώστε να κρίνουμε αν ισχύει

p = 5% (ή ≤ 5%), οπότε *συνεχίζεται η παραγωγή*, ή

p > 5%, όποτε *διακόπτεται η παραγωγή*.

Συνήθως καλούμε τις παραπάνω ***βασική*** ή ***μηδενική*** υπόθεση (H_0) και ***εναλλακτική*** υπόθεση (H_1) αντίστοιχα. Θέτουμε $X_i = 1$ αν η i-μονάδα βρέθηκε ελαττωματική και 0 διαφορετικά, $i = 1, 2, ..., \nu$. Για να αποφασίσουμε αν ισχύει η μία ή ή άλλη υπόθεση μπορούμε να βασιστούμε στο δειγματικό ποσοστό $\overline{X}$. Συγκεκριμένα, παρατηρούμε ότι:

- Αν ισχύει ότι H_0: p = 5% τότε είναι πολύ πιθανό, αν υπολογίσουμε τη δειγματική μέση τιμή $\overline{X}$, να βρεθεί κοντά στο 5% (δηλαδή $\overline{X} \approx 5\%$) ή να είναι και λίγο μικρότερο από το 5%.
- Αν ισχύει ότι H_1: p > 5% τότε είναι πολύ πιθανό, αν υπολογίσουμε τη δειγματική μέση τιμή $\overline{X}$, να λάβει τιμή πολύ μεγαλύτερη από το $\overline{X} \approx 5\%$ (δηλαδή $\overline{X}$ >> 5%).

Δηλαδή μια μεγάλη σχετικά τιμή του $\overline{X}$ οδηγεί στο συμπέρασμα ότι μάλλον δεν ισχύει η H_0. Επομένως για να αποφασίσουμε αν p = 5% ή p > 5% φαίνεται λογικός ο κανόνας:

- Αν $\overline{X} \le c$ («μικρή» τιμή του $\overline{X}$) τότε μάλλον ισχύει ότι p = 5% (*δεχόμαστε την* H_0)

- Αν $\overline{X} > c$ («μεγάλη» τιμή του $\overline{X}$) τότε μάλλον ισχύει ότι p > 5% (*απορρίπτουμε την* H_0)

για κάποιο κατώφλι c το οποίο απομένει να καθοριστεί. Παρατηρούμε ότι αν πάρουμε μεγάλο c (π.χ. 10% ή 20%) υπάρχει ο κίνδυνος να δεχθούμε ότι p = 5% ενώ στην πραγματικότητα ισχύει ότι p > 5% (π.χ. p = 7%), ενώ αν πάρουμε μικρό c (π.χ. 5%) τότε υπάρχει ο κίνδυνος να απορρίψουμε ότι p = 5% (διότι π.χ. $\overline{X}$ = 6%) ενώ στην πραγματικότητα ισχύει p = 5%. Γενικά, ανάλογα με την απόφαση που θα πάρουμε (αποδοχή η απόρριψη της H_0) ενδέχεται να κάνουμε ένα από τα εξής σφάλματα:

- *Λανθασμένη απόρριψη* της $H_0 : p = 5\%$ (όταν $\overline{X} > c$ ενώ $p = 5\%$) : ***Σφάλμα τύπου Ι***

- *Λανθασμένη αποδοχή της* $H_0 : p = 5\%$ (όταν $\overline{X} \le c$ ενώ $p > 5\%$): ***Σφάλμα τύπου ΙΙ***

Οι πιθανότητες πραγματοποίησης των παραπάνω σφαλμάτων είναι

$$P[I] = P(\textit{σφάλμα τύπου } \mathrm{I}) = P(\overline{X} > c \mid \textit{ισχύει η } \mathrm{H}_0) = P(\overline{X} > c \mid p = 5\%)$$

$$P[II] = P(\textit{σφάλμα τύπου } \mathrm{II}) = P(\overline{X} \le c \mid \textit{ισχύει η } \mathrm{H}_1) = P(\overline{X} \le c \mid p > 5\%)$$

Π.χ. αν v = 100 τότε μπορούμε να υπολογίσουμε τις P[I], P[II] για διάφορες τιμές των ν, p (διότι η τυχαία μεταβλητή $v\overline{X} = X_1 + \ldots + X_v$ ακολουθεί τη διωνυμική κατανομή b(v, p))

	p = 5% (H_0)	p = 7.5% (H_1)	p = 10% (H_1)
c = 5%	$P[I]$ = 0.384	$P[II]$ = 0.230	$P[II]$ = 0.057
c = 6%	$P[I]$ = 0.234	$P[II]$ = 0.370	$P[II]$ = 0.117
c = 7%	$P[I]$ = 0.128	$P[II]$ = 0.521	$P[II]$ = 0.206
c = 8%	$P[I]$ = 0.063	$P[II]$ = 0.664	$P[II]$ = 0.320
c = 9%	$P[I]$ = 0.028	$P[II]$ = 0.783	$P[II]$ = 0.451
c = 10%	$P[I]$ = 0.011	$P[II]$ = 0.870	$P[II]$ = 0.583

Για παράδειγμα, αν θέσουμε c = 8% και ισχύει ότι p = 5% (H_0) τότε παίρνουμε λάθος απόφαση (λανθασμένη απόρριψη της H_0 – σφάλμα τύπου Ι) με πιθ. $P[I]$ = 0.063. Αν ισχύει ότι p = 10% (H_1) τότε παίρνουμε λάθος απόφαση (λανθασμένη αποδοχή της H_0 – σφάλμα τύπου ΙΙ) με $P[II]$ = 0.320.

Από τον παραπάνω πίνακα παρατηρούμε ότι όσο αυξάνεται το c, τόσο μειώνεται η $P[I]$ αλλά παράλληλα αυξάνεται η $P[II]$. Επομένως δεν υπάρχει c που να ελαχιστοποιεί τις πιθανότητες $P[I]$, $P[II]$ ταυτόχρονα. Για τον λόγο αυτό συνήθως δίνεται μεγαλύτερη βαρύτητα στην αποφυγή του σφάλματος τύπου Ι, δηλαδή θέλουμε σπάνια να απορρίπτουμε την H_0 λανθασμένα. Πιο συγκεκριμένα απαιτούμε η $P[I]$ να είναι μικρή, συνήθως $P[I]$ = 1% ή 5%. Δηλαδή προκαθορίζουμε την πιθανότητα $P[I]$ και βρίσκουμε το c που δίνει τη συγκεκριμένη $P[I]$.

Αν π.χ. θέλουμε $P[I] \approx 6\%$ (και v = 100) τότε θα πρέπει να πάρουμε c = 8%. Σε αυτή την περίπτωση

- Αν $\overline{X} \leq 0.08$ τότε ***δεχόμαστε την* H$_0$**: $p = 5\%$
- Αν $\overline{X} > 0.08$ τότε ***απορρίπτουμε την* H$_0$**: $p = 5\%$ (έναντι της H_1: $p > 5\%$)

Ο παραπάνω κανόνας οδηγεί σε λανθασμένη απόρριψη της H_0 (δηλαδή διακοπή της παραγωγικής διαδικασίας χωρίς λόγο – false alarm) στο 6% περίπου των ελέγχων.

1.2. Το γενικό πρόβλημα: στατιστικές υποθέσεις, σφάλματα, ισχύς.

Ας δούμε το παραπάνω πρόβλημα στη γενικότερή του μορφή. Έστω τυχαίο δείγμα $X_1,X_2,...,X_\nu \sim F(x;\boldsymbol{\theta})$ και έστω ότι επιθυμούμε να ελέγξουμε την υπόθεση

H_0: $\boldsymbol{\theta}\in\Theta_0$, ***μηδενική*** (ή βασική) υπόθεση, έναντι της

H_1: $\boldsymbol{\theta}\in\Theta_1$, ***εναλλακτική*** υπόθεση.

όπου Θ_0, Θ_1 είναι υποσύνολα του παραμετρικού χώρου Θ (σύνολο επιτρεπτών τιμών των άγνωστων παραμέτρων $\boldsymbol{\theta}$) ενώ φυσικά $\Theta_0\cap\Theta_1=\varnothing$ (τα Θ_0, Θ_1 είναι ξένα). Στο Παράδειγμα 1 είχαμε $X_1,X_2,...,X_\nu \sim$ b(1, p), δηλαδή κατανομή *Bernoulli* με παράμετρο p) και H_0: $p\in\{0.05\}$ με H_1: $p\in(0.05, 1]$.

Μία υπόθεση H_i: $\boldsymbol{\theta}\in\Theta_i$ καλείται ***απλή*** αν είναι της μορφής $\Theta_i = \{\boldsymbol{\theta}_i\}$. Διαφορετικά καλείται ***σύνθετη.*** Για παράδειγμα, η H_0: $p = 0.05$ είναι απλή ενώ η H_1: $p > 0.05$ είναι σύνθετη. Επίσης, μία υπόθεση καλείται ***μονόπλευρη*** αν είναι της μορφής H_i: $\theta > \theta_0$ ή της μορφής H_i: $\theta < \theta_0$, και ***αμφίπλευρη*** αν είναι της μορφής H_i: $\theta < \theta_0$ ή $\theta > \theta_1$,δηλ. $\theta\in(-\infty, \theta_0) \cup (\theta_1, \infty)$ (η πιο συνηθισμένη περίπτωση είναι η H_i: $\theta < \theta_0$ ή $\theta > \theta_0$,δηλ. $\theta\in(-\infty, \theta_0) \cup (\theta_0, \infty)$ οπότε θα γράφουμε H_i: $\theta \neq \theta_0$).

Χρησιμοποιώντας το τυχαίο δείγμα $X_1,X_2,...,X_\nu$, 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 $\mathbf{X} = (X_1,X_2,...,X_\nu)\in K$, ***απορρίπτουμε*** την H_0: $\theta\in\Theta_0$
- αν $\mathbf{X} = (X_1,X_2,...,X_\nu)\in A$, ***δεχόμαστε*** την H_0: $\theta\in\Theta_0$

Η περιοχή K καλείται ***κρίσιμη περιοχή*** ή ***περιοχή απόρριψης*** της H_0 ενώ η περιοχή A καλείται ***περιοχή αποδοχής*** της μηδενικής υπόθεσης H_0. Η περιοχή απόρριψης καθορίζεται από την τιμή κατάλληλης *στατιστικής συνάρτησης* $T = T(\mathbf{X})$, και συνήθως είναι της μορφής:

$$K: T > c \qquad (\text{και άρα } A: T \leq c).$$

(στο παραπάνω παράδειγμα, K: $\overline{X} > c$, A: $\overline{X} \le c$).

Ανάλογα τώρα με την απόφαση που θα πάρουμε, είναι πιθανό να προκύψει ένα από τα εξής δύο σφάλματα:

- *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 σφάλματος τύπου *I* και *II* αντίστοιχα είναι

$$P[I] = P(\textit{σφάλμα τύπου } I) = P(\mathbf{X}\in K \mid H_0) = P(T > c \mid H_0)$$

$$P[II] = P(\textit{σφάλμα τύπου } II) = P(\mathbf{X}\in A \mid H_1) = P(T \le c \mid H_1).$$

Επίσης, η ποσότητα

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

καλείται ***ισχύς του ελέγχου*** και όπως θα δούμε γενικά εξαρτάται από τις άγνωστες παραμέτρους. Παρατηρούμε ότι η ισχύς του ελέγχου ισούται με την *πιθανότητα ορθής απόρριψης* της H_0:

$$\pi(\boldsymbol{\theta}) = 1 - P(\textit{σφάλμα τύπου } II) = 1 - P(\mathbf{X}\in A \mid \mathrm{H}_1) = P(\mathbf{X}\in K \mid \mathrm{H}_1).$$

Στον επόμενο πίνακα δίνονται τα δυνατά αποτελέσματα ενός ελέγχου.

Απόφαση / ***Φυσική κατάσταση (τι ισχύει πραγματικά)***	**Αποδοχή της $\mathbf{H_0}$**	**Απόρριψη της $\mathbf{H_0}$**
$\mathbf{H_0}$ ορθή	*Ορθή απόφαση*	*Σφάλμα τύπου I.*
$\mathbf{H_0}$ όχι ορθή	*Σφάλμα τύπου II*	*Ορθή απόφαση.*

Απομένει να καθορίσουμε την κρίσιμη περιοχή K (και από αυτήν την A αφού $A = \Omega - K$). Αν π.χ. είναι της μορφής $T > c$, θα πρέπει να βρούμε κατάλληλη σ.σ. T και να προσδιορίσουμε και το c. Παρατηρούμε ότι οι $P[I]$, $P[II]$ εξαρτώνται από την επιλογή της περιοχής K, και άρα θα ήταν λογικό να αναζητήσουμε την K που τις ελαχιστοποιεί. Όπως όμως διαπιστώσαμε και στο Παράδειγμα 1, μία επιλογή της κρίσιμης περιοχής που μειώνει την $P[I]$, παράλληλα αυξάνει την $P[II]$. Συνήθως μας ενδιαφέρει περισσότερο ***να μην απορρίψουμε λανθασμένα την*** $\mathbf{H_0}$, και συνεπώς θα ζητάμε η κρίσιμη περιοχή K να είναι έτσι ώστε να ισχύει

$$P(\textit{σφάλμα τύπου } I) = P(T > c \mid H_0) \le a,$$

και η πιθανότητα

$$P(\textit{σφάλμα τύπου II}) = P(T \le c \mid H_1)$$

να είναι η ελάχιστη δυνατή.

Συνήθως επιλέγουμε $a = 0.05$ ή $a = 0.01$. Αν έχουμε προσδιορίσει τη σ.σ. T, και έχουμε προεπιλέξει το a, το οποίο καλείται *επίπεδο σημαντικότητας του ελέγχου*, μπορούμε να προσδιορίσουμε και το c. Συγκεκριμένα, αν συμβολίσουμε με $F_{T|H_0}$ τη συνάρτηση κατανομής της σ.σ. T όταν ισχύει η H_0, και η περιοχή απόρριψης της H_0 είναι της μορφής $T > c$, τότε θα πρέπει

$$P[\mathrm{I}] \le a \Leftrightarrow P(T > c \mid H_0) \le a \Leftrightarrow 1 - F_{T|H_0}(c) \le a \Leftrightarrow F_{T|H_0}(c) \ge 1 - a \Leftrightarrow c \ge F^{-1}_{T|H_0}(1-a)$$

όπου με $F^{-1}_{T|H_0}(1-a)$ συμβολίσαμε το *άνω a-σημείο της κατανομής της T όταν ισχύει η H_0*.

Συνήθως το c λαμβάνεται *ίσο* με την παραπάνω ποσότητα για να εξασφαλίσουμε ελάχιστη $P[\mathrm{II}]$. Δηλαδή τελικά απορρίπτουμε την H_0 όταν

$$T > c = F^{-1}_{T|H_0}(1-a)$$

Παλαιότερα υπήρχαν πίνακες με τιμές των άνω a-σημείων για συγκεκριμένες κατανομές και a = 0.1%, 1%, 5%. Σήμερα, στην πράξη ο έλεγχος γίνεται (ισοδύναμα) μέσω Η/Υ χρησιμοποιώντας το *p-value* (ή *significance value*) το οποίο θα εξετάσουμε σε επόμενη παράγραφο.

Στον επόμενο πίνακα δίνονται τα δυνατά αποτελέσματα ενός ελέγχου μαζί με τους συμβολισμούς που έχουμε εισάγει για τα σφάλματα και τις πιθανότητες σωστής απόφασης.

Απόφαση / ***Φυσική κατάσταση (τι ισχύει πραγματικά)***	**Αποδοχή της H_0**	**Απόρριψη της H_0**
H_0 ορθή	*Ορθή απόφαση*, $P(\textit{Αποδ.}\ H_0 \mid H_0) \ge 1-a$	*Σφάλμα τύπου I*, $P(\textit{Απόρ.}\ H_0 \mid H_0) \le a$
H_0 όχι ορθή	*Σφάλμα τύπου II*, $P(\textit{Αποδ.}\ H_0 \mid H_1) = \beta$	*Ορθή απόφαση*, $P(\textit{Απόρ.}\ H_0 \mid H_1) = 1 - \beta = \pi(\theta)$

Υπογραμμίζεται σε αυτό το σημείο ότι οι συμβολισμοί $P(\,\cdot \mid H_0)$, $P(\,\cdot \mid H_1)$ που χρησιμοποιούμε **δεν υποδηλώνουν δεσμευμένη πιθανότητα**, διότι στην κλασική στατιστική οι υποθέσεις H_0, H_1 δεν αποτελούν ενδεχόμενα. Οι συμβολισμοί αυτοί χρησιμοποιούνται για να υποδηλωθεί η κατανομή από την οποία προέρχεται το δείγμα, όταν ισχύει η H_0 και η H_1 αντίστοιχα. Θα ήταν ίσως σωστότερο να γράφουμε $P_{H_0}(\cdot)$ και $P_{H_1}(\cdot)$, αλλά στην συνέχεια θα εξακολουθήσουμε να χρησιμοποιούμε για λόγους απλότητας τους συμβολισμούς $P(\,\cdot \mid H_0)$, $P(\,\cdot \mid H_1)$ καθώς επίσης και τα σύμβολα $E(\,\cdot \mid H_0)$,

$E(\cdot \mid H_1)$ αντί $E_{H_0}(\cdot)$ και $E_{H_1}(\cdot)$ για να δηλώσουμε τη μέση τιμή των αντίστοιχων τυχαίων μεταβλητών όταν ισχύει η H_0 και η H_1 αντίστοιχα.

Αξίζει εδώ να αναφέρουμε και την επόμενη ερμηνεία της πιθανότητας σφάλματος τύπου I. Έστω ότι χρησιμοποιούμε έναν έλεγχο με επίπεδο σημαντικότητας a και ότι η μηδενική υπόθεση είναι αληθής. Τότε, από όλα τα δείγματα μεγέθους ν που μπορούμε να συγκεντρώσουμε από τον πληθυσμό, ένα ποσοστό ίσο με a (το πολύ) θα οδηγήσουν σε τιμή της στατιστικής συνάρτησης ελέγχου $T = T(\mathbf{x})$ τέτοια που να αναγκασθούμε να απορρίψουμε τη μηδενική υπόθεση (εν προκειμένω εσφαλμένα).

Σε έναν έλεγχο υποθέσεων, αυτό που μας ενδιαφέρει περισσότερο είναι να διατηρηθεί μικρή η $P[I]$, και για αυτό απαιτούμε $P[I] \le a$. Κάτω από αυτό τον περιορισμό, η $P[II]$ μπορεί να είναι μικρή, θα μπορούσε όμως και να πάρει μεγάλες τιμές (κάτι το οποίο είναι φυσικά ανεπιθύμητο). Για το λόγο αυτό

- αν $\mathbf{X} \in K$ λέμε ότι «*απορρίπτουμε την* H_0» με πιθ. λάθους $\le a$,

- αν $\mathbf{X} \in A$, 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 λάθους, $P[II]$, μπορεί να είναι μεγάλη.

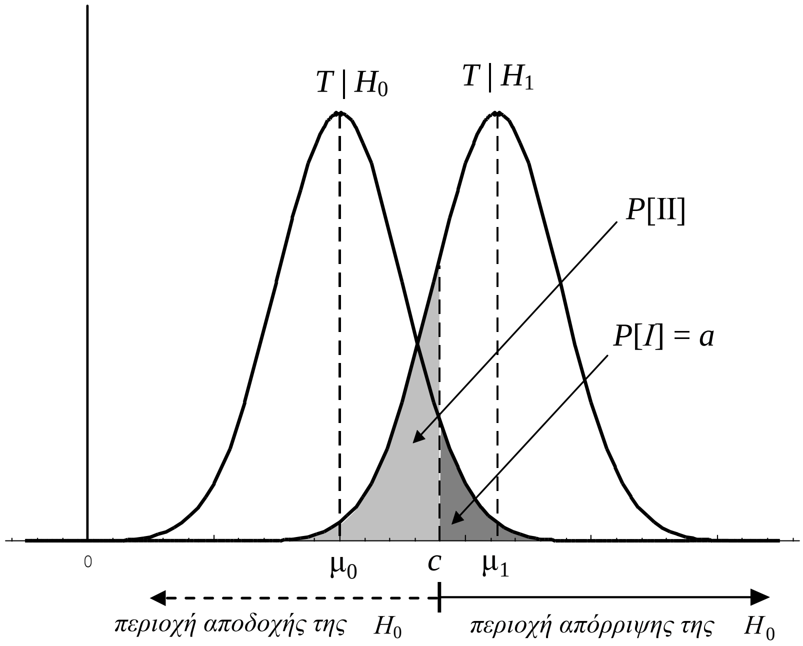
Παράδειγμα 2. Έστω ότι έχουμε έναν πληθυσμό N(μ,1) (δηλαδή το χαρακτηριστικό του πληθυσμού που μας ενδιαφέρει ακολουθεί κανονική κατανομή με μέση τιμή μ και διασπορά 1) και θέλουμε να ελέγξουμε την υπόθεση

$$H_0: \mu = 10 \quad \textit{έναντι της} \quad H_1: \mu = 11.$$

Επιλέγουμε τυχαίο δείγμα $X_1, X_2, \ldots, X_\nu$ από τον πληθυσμό αυτό και αποφασίζουμε να χρησιμοποιήσουμε την σ.σ. $T = T(\mathbf{x}) = \sqrt{\nu}(\overline{X} - 10)$. Παρατηρούμε ότι,

- όταν ισχύει η H_0: μ = 10, *τότε* $T \sim$ N(0,1)
- όταν ισχύει η H_1: μ = 11, *τότε η T λαμβάνει «μεγάλες» τιμές* (*ακολουθεί* την κατανομή N($\sqrt{\nu}$,1))

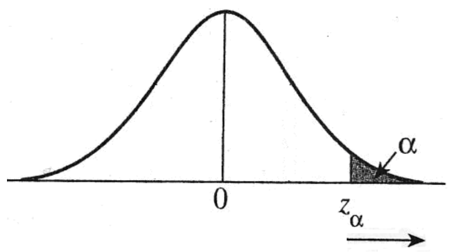
(τα παραπάνω ισχύουν διότι όταν $X_i \sim \mathrm{N}(\mu,\sigma^2)$ τότε $\overline{X} \sim \mathrm{N}(\mu, \sigma^2/\nu)$). Επομένως, είναι λογικό να απορρίπτουμε την H_0 όταν $T > c$. Στο ακόλουθο σχήμα απεικονίζεται η μορφή της σ.π.π. της T υπό την H_0 και υπό την H_1, καθώς και η περιοχή απόρριψης.



• Για να έχουμε $P[I] = a$, θα πρέπει

$$P(T > c \mid H_0) = a \Leftrightarrow c = \Phi^{-1}(1-a) \equiv z_a$$

όπου με Φ συμβολίζεται η σ.κ. της N(0,1) και με z_a συμβολίζουμε το άνω a-σημείο της τυποποιημένης κανονικής κατανομής (βλέπε ακόλουθο σχήμα)



Τελικά ο έλεγχος θα είναι της μορφής:

- Αν $T = \sqrt{\nu}(\overline{X} - 10) > z_a$ τότε *απορρίπτουμε* την H_0
- Αν $T = \sqrt{\nu}(\overline{X} - 10) \leq z_a$ τότε *δεν απορρίπτουμε* την H_0.

• Η $P[II]$ στη συγκεκριμένη περίπτωση θα είναι

$$\begin{aligned} P[II] &= P(T \leq c \mid H_1) = P(T \leq z_a \mid T \sim N(\sqrt{\nu}, 1)) \\ &= P(T - \sqrt{\nu} \leq z_a - \sqrt{\nu} \mid T \sim N(\sqrt{\nu}, 1)) = \Phi(z_a - \sqrt{\nu}) \end{aligned}$$

η οποία μειώνεται όσο αυξάνεται το μέγεθος του δείγματος. Η ισχύς θα είναι ίση με

$$\pi = 1 - P[II].$$

Για $P[I] = a$ = 5% προκύπτει $c = \Phi^{-1}(1-a) \equiv z_a \approx 1.645$ και παίρνουμε τις επόμενες τιμές για την πιθανότητα σφάλματος τύπου II.

ν	P[II]
5	0.277237
10	0.0645983
15	0.0129408
20	0.00234832
25	0.000396825

Έστω τώρα ότι πήραμε συγκεκριμένο δείγμα μεγέθους ν = 9 και καταγράφηκαν οι τιμές

10.45, 11.54, 11.81, 11.03, 8.88, 10.48, 9.77, 10.49, 11.5.

Οι συγκεκριμένες τιμές του δείγματος *μετά* την πραγματοποίηση του τυχαίου πειράματος (δειγματοληψία) δεν θεωρούνται τυχαίες μεταβλητές (είναι γνωστές) και συνήθως συμβολίζονται με x_1, x_2, …, x_ν (*πριν* την πραγματοποίηση του τυχαίου πειράματος συμβολίζονται με X_1, X_2, …, X_ν, διότι τότε θεωρούνται τυχαίες μεταβλητές). Όμοια, ο δειγματικός μέσος και η τιμή της σ.σ. $T = T(\mathbf{x})$ από τα πραγματικά δεδομένα x_1, x_2, …x_ν, συνήθως συμβολίζονται με $\bar{x}$ και $t = T(x_1,\ldots,x_\nu) = T(\mathbf{x})$ αντίστοιχα. Από το παραπάνω δείγμα υπολογίζουμε τώρα ότι $\bar{x} \approx 10.6611$, και

$$t = \sqrt{\nu}(\bar{x} - 10) \approx 1.9833$$

το οποίο είναι μεγαλύτερο του $c = z_a = 1.645$ οπότε το δείγμα προέρχεται από την περιοχή K. Επομένως απορρίπτουμε την $H_0 : \mu = 10$ έναντι της $H_1 : \mu = 11$ σε επίπεδο σημαντικότητας a = 5% ($P[I]$ = 5%, $P[II] = \Phi(z_a - \sqrt{\nu}) \approx$ 8.77%).

Παρατηρούμε ότι, από το συγκεκριμένο δείγμα, η σ.σ. $T = T(\mathbf{x})$ έλαβε την τιμή $t = T(\mathbf{x})$ = 1.9833. Ένα ενδιαφέρον και γόνιμο ερώτημα εδώ είναι το εξής: πόσο πιθανή είναι μια τέτοια τιμή της σ.σ. (και ακόμη πιο «ακραία» από αυτή) όταν ισχύει η H_0; Η πιθανότητα αυτή είναι

$$P(T(\mathbf{X}) > T(\mathbf{x}) \mid H_0) = P(T \geq t \mid H_0) = P(T > 1.9833 \mid H_0) = 1 - \Phi(1.9833) \approx 0.02366.$$

(το «≥» μπορεί να γίνει «>» διότι η T είναι συνεχής τυχαία μεταβλητή). Μπορεί να θεωρηθεί ότι η παραπάνω τιμή *εκφράζει την πιθανότητα να εμφανιστεί το δείγμα που πήραμε, ενώ ισχύει η* H_0. Μπορούμε επομένως να πούμε ότι, αν ισχύει η H_0 παίρνουμε ένα τέτοιο δείγμα μόλις στο 2.3% περίπου των περιπτώσεων. Η παραπάνω ποσότητα καλείται *p-value* ή significance value του δείγματος. Θα εξετάσουμε γενικότερα την έννοια αυτή στην ακόλουθη παράγραφο.

1.3 *p-value*.

Αν η περιοχή απόρριψης της H_0 είναι της μορφής $T > c$ τότε ως *p-value* (ή significance value) των τιμών ενός συγκεκριμένου δείγματος $x_1, x_2, \ldots, x_\nu$, ορίζεται η τιμή

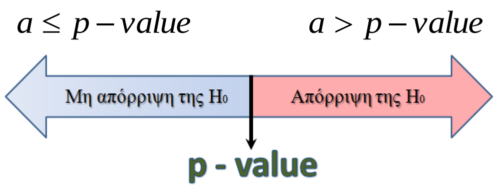
$$p-value = P(T(\mathbf{X}) > T(\mathbf{x}) \mid H_0) = P(T \geq t \mid H_0) = 1 - F_{T|H_0}(t),$$

όπου $t = T(x_1,\ldots,x_\nu) = T(\mathbf{x})$ είναι η τιμή της σ.σ. $T = T(\mathbf{x})$ με βάση το δείγμα $x_1, x_2, \ldots, x_\nu$. Το $p-value$ μπορεί να θεωρηθεί ότι εκφράζει την πιθανότητα *να πάρουμε την τιμή t και ακόμη πιο «ακραία» από αυτήν, ενώ ισχύει η* H_0. Συνήθως η T είναι συνεχής τυχαία μεταβλητή οπότε μπορεί να θεωρηθεί ότι έχουμε «>» μέσα στην παραπάνω πιθανότητα. Διαισθητικά, αν το *p-value* είναι «κοντά» στο 0 τότε συμπεραίνουμε ότι είναι «απίθανο», δεδομένης της H_0, να εμφανιστεί το συγκεκριμένο δείγμα $x_1, x_2, \ldots, x_\nu$, και όπως είναι φυσικό φτάνουμε στο συμπέρασμα ότι μάλλον δεν πρέπει να ισχύει η H_0. Για «μικρό» λοιπόν *p-value* είναι λογικό να απορρίπτουμε την H_0. Τι σχέση όμως έχει αυτή η απόρριψη με την απόρριψη που γίνεται με βάση την κρίσιμη περιοχή που είδαμε παραπάνω (δηλ. όταν $t > c$); Σχετικά παρατηρούμε ότι

$$t > c \Leftrightarrow P(T \geq t \mid H_0) < P(T > c \mid H_0) \Leftrightarrow p-value < a.$$

Επομένως αντί να εξετάζουμε αν $t > c$, ισοδύναμα μπορούμε να εξετάζουμε αν ισχύει η ανισότητα *p-value*<a. Ειδικότερα μπορούμε να αποφασίζουμε με βάση τον εξής κανόνα (βλέπε και διπλανό σχήμα),

- αν το *p-value* < a : *απορρίπτουμε* την H_0, ενώ
- αν *p-value* ≥ a : *δεν απορρίπτουμε* την H_0.



Κανόνας απόφασης με βάση την τιμή του *p-value* ενός ελέγχου

Αν το *p-value* είναι πάρα πολύ μικρό (π.χ. 0.00001) τότε απορρίπτουμε την H_0 χωρίς επιφυλάξεις ενώ αν το *p-value* είναι σχετικά μικρό (π.χ. «κοντά» στο 0.045 με a = 0.05) τότε ναι μεν απορρίπτουμε την H_0 αλλά με κάποια επιφύλαξη (σε αυτή την περίπτωση στην πράξη, για να είμαστε πιο σίγουροι, χρειάζεται περισσότερη πληροφορία, π.χ. μεγαλύτερο δείγμα).

Όπως είναι φανερό από το παραπάνω σχήμα, το *p-value* μπορεί να ορισθεί και ως η ελάχιστη τιμή του επιπέδου σημαντικότητας για την οποία απορρίπτεται η H_0.

Το *p-value* είναι ένα μέτρο το οποίο εκφράζει *πόσο ισχυρές είναι οι ενδείξεις που προκύπτουν από το δείγμα, εναντίον της* H_0. Έτσι, υπολογίζοντας το *p-value* ενός δείγματος, για συγκεκριμένο έλεγχο υποθέσεων, γνωρίζουμε πόσο πιθανή ήταν η εμφάνιση του δείγματος που πήραμε αν η μηδενική υπόθεση H_0 **ήταν** αληθής. Επομένως, **όσο πιο μικρό** είναι το *p-value* τόσο ισχυρότερες ενδείξεις **εναντίον** της H_0 προκύπτουν από το συγκεκριμένο τυχαίο δείγμα ή αλλιώς **τόσο πιο σημαντική** είναι **η τιμή** που δίνει το δείγμα στη στατιστική συνάρτηση ελέγχου.

Στα στατιστικά πακέτα, μετά την εισαγωγή των τιμών $x_1, x_2, \ldots, x_\nu$, του δείγματος και την επιλογή του επιθυμητού ελέγχου, εμφανίζεται η τιμή του *p-value* που αντιστοιχεί στο $x_1, x_2, \ldots, x_\nu$. Σύμφωνα με τα παραπάνω, αν η τιμή αυτή είναι μικρή (μικρότερη του προαποφασισμένου επίπεδου σημαντικότητας $a = 0.01$ ή 0.05) τότε απορρίπτουμε την H_0 σε ε.σ. a. Το πλεονέκτημα από την χρήση του *p-value* είναι ότι δεν απορρίπτουμε ή δεχόμαστε απλώς την H_0, αλλά μπορούμε να δούμε και πόσο πιθανή ήταν η εμφάνιση του δείγματος $x_1, x_2, \ldots, x_\nu$, που πήραμε (υπό την H_0) ενώ επίσης μπορούμε να την συγκρίνουμε άμεσα με όποιο a και αν επιλέξουμε. Ο λόγος για τον οποίο το *p-value* συνήθως προϋποθέτει τη χρήση Η/Υ είναι διότι χωρίς τον Η/Υ δεν είναι πάντοτε εύκολο να υπολογιστεί ή να πινακοποιηθεί η $P(T \geq t \mid H_0)$ για κάθε τιμή του t.

Στη διεθνή βιβλιογραφία, καθώς και σε ορισμένα στατιστικά πακέτα, χρησιμοποιείται για το *p-value* και ο όρος, παρατηρούμενο επίπεδο σημαντικότητας (observed significance level). Στα πλαίσια των παρόντων σημειώσεων δεν θα χρησιμοποιηθεί ο όρος αυτός, αλλά καλό θα είναι ο αναγνώστης να τον έχει υπόψη του.

Παράδειγμα 3. Επιθυμούμε να ελέγξουμε αν ο μέσος μ ενός κανονικού πληθυσμού (με γνωστή διασπορά σ^2) είναι ίσος με μ_0 ή είναι μεγαλύτερος του μ_0, δηλαδή έχουμε τις υποθέσεις

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

Λαμβάνουμε ένα τυχαίο δείγμα $X_1, X_2, \ldots, X_\nu$ από τον πληθυσμό αυτό και χρησιμοποιούμε τη στατιστική συνάρτηση

$$T = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{\nu}},$$

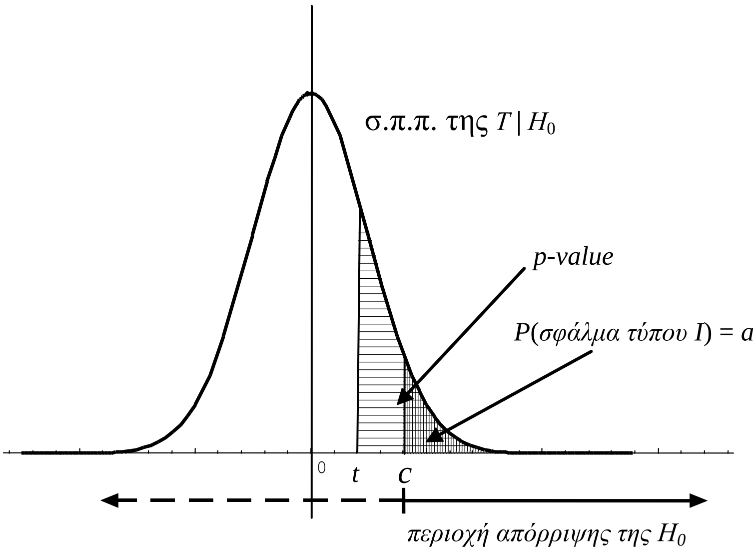
η οποία, όταν ισχύει η H_0, ακολουθεί την τυποποιημένη κανονική κατανομή N(0,1) (διότι $\overline{X} \sim N(\mu, \sigma^2/\nu)$) ενώ όταν ισχύει η H_1: $\mu > \mu_0$, λαμβάνει «μεγάλες» τιμές. Συνεπώς απορρίπτουμε την H_0 όταν

$$t = \frac{\overline{x} - \mu_0}{\sigma / \sqrt{\nu}} > c = z_a.$$

Σημειώνεται ότι λάβαμε $c = z_a$, για να εξασφαλίσουμε $P[I] = a$. Το *p-value* των τιμών ενός τυχαίου δείγματος $x_1, x_2, \ldots, x_\nu$, εδώ θα είναι

$$p - value = P(T > t \mid H_0) = 1 - \Phi(t) = 1 - \Phi\left(\frac{\overline{x} - \mu_0}{\sigma / \sqrt{\nu}}\right).$$

Στο ακόλουθο σχήμα φαίνεται το *p-value* και το επίπεδο σημαντικότητας a στο συγκεκριμένο παράδειγμα. Από το σχήμα αυτό είναι προφανές ότι αν $t < c$ τότε για τα αντίστοιχα εμβαδά θα είναι *p-value* > a, και αντίστροφα.



Επίσης είναι εύκολο να δούμε εδώ ότι

$$P[II] = P(T \le z_a \mid H_1 : \mu > \mu_0) = P(\frac{\overline{X} - \mu_0}{\sigma / \sqrt{\nu}} \le z_a \mid H_1) = P(\frac{\overline{X} - \mu}{\sigma / \sqrt{\nu}} \le z_a - \frac{\mu - \mu_0}{\sigma / \sqrt{\nu}} \mid H_1) = \Phi(z_a - \frac{\mu - \mu_0}{\sigma / \sqrt{\nu}}),$$

και

$$\pi(\mu) = 1 - \Phi(z_a - \frac{\mu - \mu_0}{\sigma / \sqrt{\nu}}) = \Phi(\frac{\mu - \mu_0}{\sigma / \sqrt{\nu}} - z_a),$$

για $\mu > \mu_0$ (υπενθυμίζεται ότι $\Phi(-x) = 1 - \Phi(x)$).

Έστω τώρα ότι από ένα δείγμα μεγέθους $\nu = 25$ πήραμε $\bar{x} = 101$ (με $\sigma = 5$) και θέλουμε να ελέγξουμε αν ισχύει η H_0: $\mu = 100$, ή H_1: $\mu > 100$ (δηλ. $\mu_0 = 100$), σε ε.σ. $a = 5\%$. Σύμφωνα με τα παραπάνω, η περιοχή απόρριψης της H_0 καθορίζεται από την ανισότητα

$$K:\ t = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{\nu}} > z_a,$$

Επειδή όμως $t = 1$, $z_{0.05} = 1.645$, η παραπάνω *δεν ισχύει* και επομένως δεν μπορούμε να απορρίψουμε την H_0. Επίσης,

$$p - value = P(T > 1 \mid H_0) = 1 - \Phi(1) \approx 0.158.$$

Μπορούμε (λιγότερα αυστηρά) να πούμε ότι στο 15.8% των περιπτώσεων που ισχύει η H_0 λαμβάνουμε ένα τέτοιο $\bar{x}$, δηλαδή το δείγμα αυτό μπορεί να θεωρηθεί σχετικά «φυσιολογικό» υπό την H_0. Επειδή λοιπόν το *p-value* είναι μεγαλύτερο του $a = 5\%$, δεν απορρίπτεται την H_0. Σε αυτό το σημείο αξίζει και πάλι να παρατηρήσουμε ότι συνήθως, όταν είμαστε στην περιοχή Α, αποφεύγουμε να πούμε ότι «δεχόμαστε την H_0» και προτιμούμε όπως παραπάνω να λέμε ότι «δεν απορρίπτουμε» ή «δεν έχουμε αρκετά στοιχεία να απορρίψουμε» την H_0 (π.χ. χρειαζόμαστε μεγαλύτερο δείγμα). Αυτό συμβαίνει διότι αν πούμε ότι «δεχόμαστε την H_0», υπάρχει το ενδεχόμενο να έχουμε πάρει λάθος απόφαση με πιθανότητα P(*αποδ.* H_0 | ισχ. H_1) = $P[II]$, η οποία, αντίθετα με την $P[I]$, μπορεί να είναι αρκετά μεγάλη. Π.χ. αν στο παράδειγμα ισχύει η H_1 με μ = 101 τότε

$$P[II] = \Phi(z_a - \frac{\mu - \mu_0}{\sigma / \sqrt{\nu}}) = \Phi(1.645 - \frac{101-100}{5/\sqrt{25}}) = \Phi(0.645) \approx 74\%.$$

Κλείνοντας την παράγραφο αυτή αξίζει να αναφερθεί ότι, θέτοντας μικρότερο επίπεδο σημαντικότητας σε ένα έλεγχο, απαιτούμε για την απόρριψη της H_0 πιο «σημαντικές ενδείξεις» από το παρατηρηθέν δείγμα μας ότι η συμπεριφορά των τιμών που συγκεντρώσαμε αποκλίνει από τη συμπεριφορά που θα αναμέναμε όταν η H_0 είναι αληθής. Έτσι, μπορεί, σε κάποιο επίπεδο σημαντικότητας a, π.χ. a = 0.05, να απορρίπτουμε την H_0 και σε κάποιο μικρότερο, π.χ. a = 0.01, να μην την απορρίπτουμε γιατί απαιτούμε σημαντικότερες αποδείξεις. Όσο πιο μικρό είναι το επίπεδο σημαντικότητας στο οποίο μπορούμε να απορρίψουμε την H_0, δηλαδή το *p-value* του δείγματος, τόσο πιο σημαντική είναι η τιμή της στατιστικής συνάρτησης ελέγχου που παρατηρείται στο δείγμα, με την έννοια ότι δίνει πιο ισχυρές αποδείξεις εναντίον της H_0. Άρα, όσο **πιο μικρό** είναι το επίπεδο σημαντικότητας στο οποίο μπορούμε να απορρίψουμε την H_0, τόσο **πιο σημαντικό**, στατιστικά, είναι το αποτέλεσμα του ελέγχου. Τέλος, είναι προφανές, ότι

- αν η H_0 απορρίπτεται σε κάποιο επίπεδο σημαντικότητας a, τότε επίσης απορρίπτεται σε οποιοδήποτε μεγαλύτερο, ενώ
- αν δεν απορρίπτεται σε κάποιο επίπεδο σημαντικότητας a, τότε επίσης δεν απορρίπτεται σε οποιοδήποτε μικρότερο.

1.4. Τυχαιοποιημένοι έλεγχοι

Σε ορισμένες περιπτώσεις που δεν μπορούμε να επιτύχουμε ακριβώς το επίπεδο σημαντικότητας που θέλουμε (π.χ. όταν η σ.σ. που χρησιμοποιούμε για τον έλεγχο είναι διακριτή), μπορούμε εναλλακτικά να πραγματοποιήσουμε τυχαιοποιημένο έλεγχο. Για απλότητα, η μέθοδος αυτή θα παρουσιαστεί μέσα από ένα παράδειγμα.

Παράδειγμα 4. (*συνέχεια Παραδείγματος* 1 - *τυχαιοποιημένος έλεγχος*). Σε ένα εργοστάσιο γίνεται έλεγχος των ν = 100 συσκευών της ωριαίας παραγωγής. Η διαδικασία βρίσκεται εντός προδιαγραφών αν η πιθ. *p παραγωγής ελαττωματικής συσκευής* ≤ 5%. Για να αποφασίσουμε αν ισχύει ότι H_0: p = 5% ή H_1: p > 5%, εξετάζουμε αν ισχύει $\overline{X} > c$ (όπου $X_i = 1$ αν η i-μονάδα είναι ελαττωματική και 0 διαφορετικά) και αν ισχύει, *απορρίπτουμε την* H_0. Θέλοντας να βρούμε το c τέτοιο ώστε $P[I] = a = 5\%$, παρατηρούμε ότι

$$P[I] = P(\overline{X} > c \mid H_0\text{: } p = 5\%) = P(\sum X_i > [\nu c] \mid H_0\text{: } p = 5\%) = \sum_{i=[\nu c]+1}^{\nu} \binom{\nu}{i} p^i (1-p)^{\nu-i},$$

και κατασκευάζουμε τον πίνακα

$[\nu c]$	**p = 5% (H_0)**
7	$P[I] = 0.128$
8	$P[I] = 0.063$
9	$P[I] = 0.028$
10	$P[I] = 0.011$

από τον οποίο γίνεται αντιληπτό ότι δεν μπορούμε να επιτύχουμε ακριβώς $P[I] = a = 0.05$. Αν ως περιοχή απόρριψης οριστεί η $\sum X_i > 9$ τότε έχουμε $P[I] = 0.028$, ενώ αν οριστεί η $\sum X_i > 8$ τότε $P[I] = 0.063$. Συνήθως, σε αυτή την περίπτωση θέτουμε ως περιοχή απόρριψης αυτή που οδηγεί σε $P[I]$ πιο κοντά στο $a = 0.05$ (και $P[I] < a$). Οι έλεγχοι αυτού του τύπου (με $P[I] < a$) καλούνται *συντηρητικοί* διότι απορρίπτουν την H_0 σπανιότερα (δηλ. με πιθανότητα μικρότερη από την ονομαστική a).

Εάν όμως για κάποιο λόγο επιθυμούμε να κάνουμε τον έλεγχο της H_0 με $P[I]$ ίση ακριβώς με a τότε μπορούμε να κάνουμε το εξής:

- Αν $\sum X_i > 9$ απορρίπτουμε την H_0
- Αν $\sum X_i = 9$ τότε απορρίπτουμε την H_0 με πιθανότητα γ.

Παρατηρώντας ότι τώρα,

$$P[I] = P(\sum X_i > 9 \mid H_0) + \gamma P(\sum X_i = 9 \mid H_0) = 0.028 + \gamma\, 0.035,$$

μπορούμε να βρούμε το γ που θα οδηγεί σε $P[I] = a$ από την εξίσωση $0.028 + \gamma\, 0.035 = 0.05$. Προκύπτει ότι $\gamma = 0.6285$.

Ο παραπάνω έλεγχος κατά τον οποίο απορρίπτουμε την H_0 με κάποια πιθανότητα καλείται *τυχαιοποιημένος έλεγχος*. Στην πράξη σπανίως κατασκευάζεται τυχαιοποιημένος έλεγχος. Όπως αναφέραμε και παραπάνω συνήθως αρκούμαστε σε έναν συντηρητικό έλεγχο.

1.5. Ελεγχοσυνάρτηση

Για την ευκολότερη διατύπωση κάποιων εννοιών και αποδείξεων ορίζουμε την ελεγχοσυνάρτηση ενός ελέγχου.

Ορισμός. 1. *Ελεγχοσυνάρτηση* ή *κρίνουσα συνάρτηση* για τον έλεγχο της H_0 έναντι της H_1 καλείται η στατιστική συνάρτηση (συνάρτηση του τυχαίου δείγματος $\mathbf{X} = (X_1, X_2, \ldots, X_v)$):

$$\phi(\mathbf{X}) = \begin{cases} 1, & \mathbf{X} \in K \\ 0, & \mathbf{X} \in A \end{cases}$$

ή γενικότερα (περιλαμβάνοντας και την περίπτωση τυχαιοποιημένων ελέγχων)

$$\phi(\mathbf{X}) = \begin{cases} 1, & \mathbf{X} \in K \\ 0, & \mathbf{X} \in A \\ \gamma, & \mathbf{X} \in M \end{cases}$$

Στην περίπτωση τυχαιοποιημένου έλεγχου ο δ.χ. χωρίζεται σε τρία ξένα υποσύνολα K, A, $M = \Omega - K - A$. Η ελεγχοσυνάρτηση μπορεί να θεωρηθεί ότι εκφράζει την πιθανότητα με την οποία θα απορρίψουμε την H_0, δεδομένου ότι λάβαμε το δείγμα $\mathbf{X}$. Η πιθανότητα απόρριψης της H_0 εκφράζεται μέσω της ελεγχοσυνάρτησης ως εξής:

$$P(\textit{απόρριψης της } H_0) = 1P(\phi(\mathbf{X}) = 1) + 0P(\phi(\mathbf{X}) = 0) + \gamma\, P(\phi(\mathbf{X}) = \gamma) = E(\phi(\mathbf{X})) ,$$

δηλαδή είναι ίση με την αναμενόμενη τιμή της ελεγχοσυνάρτησης. Επομένως, γενικά,

$$P[I] = E(\phi(\mathbf{X}) \mid H_0) \quad \text{και} \quad \pi(\theta) = E(\phi(\mathbf{X}) \mid H_1) .$$

1.6. Ασκήσεις

1. Έστω $X_1 = X$ ένα τυχαίο δείγμα μεγέθους $v = 1$ από μια κατανομή με συνάρτηση πυκνότητας

$$f(x;\theta) = (1+\theta)x^{\theta}, \quad 0 < x < 1.$$

Για τον έλεγχο της υπόθεσης $H_0 : \theta = 1$ έναντι της $H_1 : \theta = 2$ χρησιμοποιούμε την κρίσιμη περιοχή $K : X > c$ όπου c κατάλληλη σταθερά.

α. Να υπολογισθούν οι πιθανότητες σφάλματος τύπου *I* και τύπου *II* στην περίπτωση που επιλέγουμε $c = 1/2$.

β. Να βρεθεί η τιμή της σταθεράς c έτσι ώστε η πιθανότητα εμφάνισης σφάλματος τύπου *I* να είναι ίση με 0.19. Στη συνέχεια να υπολογισθεί η πιθανότητα σφάλματος τύπου *II.*

γ. Αν $x = 0.8$ να κάνετε τον έλεγχο με $a = 0.19$ μέσω του p-value.

2. Έστω $X_1, X_2, ..., X_\nu$ ένα τυχαίο δείγμα από την κανονική κατανομή $N(0,\sigma^2)$. Για τον έλεγχο της υπόθεση $H_0 : \sigma^2 = 1$ έναντι της $H_1 : \sigma^2 = 4$ χρησιμοποιούμε την κρίσιμη περιοχή

$$K : \sum_{i=1}^{\nu} X_i^2 > c .$$

α. Αφού βρεθεί η σταθερά c ώστε ο έλεγχος να έχει πιθανότητα σφάλματος τύπου *I* ίση με α=5%, να υπολογισθεί η πιθανότητα σφάλματος τύπου *II*.

β. Αν $x_1 = 2.3$, $x_2 = -4.1$, $x_3 = 1.8$, να κάνετε τον έλεγχο σε ε.σ. $a = 5\%$.

3. Σε μια κάλπη με 10 σφαίρες είναι γνωστό ότι υπάρχουν είτε 3 Λευκές και 7 Μαύρες σφαίρες είτε 7 Λευκές και 3 Μαύρες. Για τον έλεγχο της υπόθεσης

H_0: Η κάλπη έχει 3 Λευκές και 7 Μαύρες σφαίρες

H_1: Η κάλπη έχει 7 Λευκές και 3 Μαύρες σφαίρες

εξάγονται 3 σφαίρες και η H_0 απορρίπτεται όταν και οι 3 σφαίρες είναι Λευκές. Να υπολογισθούν οι πιθανότητες σφάλματος τύπου *I* και τύπου *II* στην περίπτωση που η εξαγωγή των 3 σφαιρών γίνεται

α. με επανάθεση, β. χωρίς επανάθεση.

4. Για τον έλεγχο της υπόθεσης $H_0 : \theta = 1$ έναντι της $H_1 : \theta = 2$ όπου θ η παράμετρος της κατανομής με συνάρτηση πυκνότητας

$$f(x;\theta) = \theta x^{\theta-1}, \quad 0 < x < 1.$$

χρησιμοποιείται τυχαίο δείγμα δύο παρατηρήσεων X_1, X_2 και η H_0 απορρίπτεται όταν ισχύει $X_1 X_2 > 3/4$.

α. Να βρεθεί η κατανομή της τυχαίας μεταβλητής $S = -\ln X_1 - \ln X_2$.

β. Να υπολογισθούν οι πιθανότητες σφάλματος τύπου *I* και τύπου *II*.

5. Έστω $X_1, X_2, ..., X_\nu$ ένα τυχαίο δείγμα από την κατανομή $N(\mu,1)$. Για τον έλεγχο της υπόθεσης H_0:μ=μ_0 έναντι της H_1: μ=μ_0+1 χρησιμοποιείται η επόμενη κρίσιμη περιοχή

$$K: \ \overline{X} > c,$$

όπου c κατάλληλη σταθερά. Να δειχθεί ότι το ελάχιστο μέγεθος ν του δείγματος που απαιτείται ώστε οι πιθανότητες σφαλμάτων τύπου *I* και *II* του ελέγχου να είναι το πολύ ίσες με α και β αντίστοιχα, δίνεται από τον τύπο

$$\nu_0 = [(z_a - z_{1-\beta})^2] + 1$$

όπου το σύμβολο [*x*] παριστάνει το ακέραιο μέρος του *x*.

6. Έστω p η πιθανότητα εμφάνισης της ένδειξης "γράμματα" κατά τη ρίψη ενός νομίσματος. Για τον έλεγχο της υπόθεσης $H_0 : p = 1/2$ κατά της $H_0 : p \neq 1/2$ ρίχνουμε το νόμισμα 10 φορές και έστω X ο αριθμός των φορών που εμφανίζεται "γράμματα". Ποια είναι η πιθανότητα σφάλματος τύπου I και ποια η ισχύς π(3/4) του ελέγχου με κρίσιμη περιοχή

$$K : X < 2 \quad ή \quad X > 8;$$

7. Για να αποφασίσουμε κατά πόσο ένα νόμισμα είναι αμερόληπτο ($H_0 : p = 1/2$) ή ευνοεί την εμφάνιση της ένδειξης "γράμματα" ($H_1 : p > 1/2$), ρίχνουμε το νόμισμα 5 φορές και απορρίπτουμε την H_0 αν είχαμε 4 ή 5 εμφανίσεις γραμμάτων. Να υπολογισθεί η πιθανότητα σφάλματος τύπου I και η ισχύς του ελέγχου για p=3/4.

8. Έστω $X_1, X_2, ..., X_\nu$ ένα τυχαίο δείγμα από την κατανομή $N(\mu, 4)$. Για τον έλεγχο της υπόθεσης H_0:μ=μ_0 έναντι της H_1: μ>μ_0 (όπου μ_0 είναι ένας δοσμένος πραγματικός αριθμός) χρησιμοποιείται η επόμενη κρίσιμη περιοχή

$$K : \sum_{i=1}^{\nu} X_i > c$$

όπου c κατάλληλη σταθερά. Να βρεθεί το c και το μέγεθος του δείγματος ν ώστε για τη συνάρτηση ισχύος $\pi(\mu)$ του ελέγχου να ισχύει $\pi(\mu_0) = \alpha = 0.05$ και $\pi(\mu_0 + 5) = 0.90$.

2. ΚΑΤΑΣΚΕΥΗ ΕΛΕΓΧΩΝ ΥΠΟΘΕΣΕΩΝ

2.1. Έλεγχοι απλών στατιστικών υποθέσεων. Το λήμμα Neyman – Pearson.

Ένα ερώτημα που δεν έχει απαντηθεί μέχρι τώρα είναι πως προσδιορίζουμε την περιοχή απόρριψης της H_0. Σε προηγούμενα παραδείγματα δίνονταν η μορφή της περιοχής K , π.χ. $T > c$ για συγκεκριμένη σ.σ. $T = T(\mathbf{X})$, και απέμενε να προσδιοριστεί το c έτσι ώστε $P[I] = a$. Πως όμως προσδιορίζουμε την σ.σ. T και πότε η επιλογή της είναι βέλτιστη; Για έλεγχο απλής H_0 έναντι απλής H_1 αποδεικνύεται ότι μπορεί να βρεθεί βέλτιστη επιλογή της σ.σ. T, που οδηγεί στη μεγαλύτερη δυνατή ισχύ. Αν $X_1, X_2, ..., X_\nu$ ένα τυχαίο δείγμα ~ $F(x;\theta)$ με σ.π. ή σ.π.π. $f(x;\theta)$, ισχύει το ακόλουθο αποτέλεσμα.

Θεώρημα 1 (*Λήμμα Neyman - Pearson*) *Έστω ότι θέλουμε να ελέγξουμε την υπόθεση*

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

σε επίπεδο σημαντικότητας a. Αν συμβολίσουμε με $\lambda(\mathbf{x})$ *το λόγο*

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

τότε ο έλεγχος που ορίζεται από την κρίσιμη περιοχή

$$K: \ \lambda(\mathbf{X}) < c ,$$

όπου $c > 0$ σταθερά τέτοια ώστε $P[I] = a$, παρουσιάζει τη μεγαλύτερη ισχύ ανάμεσα σε όλους τους ελέγχους με επίπεδο σημαντικότητας a. (ΙΕ - Ισχυρότατος έλεγχος)

Το πηλίκο $\lambda(\mathbf{x})$ καλείται *λόγος πιθανοφανειών* του δείγματος. Ισοδύναμα με την διατύπωση του παραπάνω θεωρήματος μπορούμε να πούμε ότι η κρίνουσα συνάρτηση

$$\phi(\mathbf{X}) = \begin{cases} 1, & \lambda(\mathbf{X}) < c \\ 0, & \lambda(\mathbf{X}) \geq c \end{cases}$$

με $c > 0$ τέτοιο ώστε να ισχύει

$$P[I] = E(\phi(\mathbf{X}) | H_0) = a ,$$

οδηγεί σε έλεγχο με τη μεγαλύτερη ισχύ ανάμεσα σε όλους τους ελέγχους με ε.σ. a. Διαισθητικά, όταν ο *λόγος πιθανοφανειών* $\lambda(\mathbf{X})$ λάβει μικρή τιμή (μικρότερη ενός c) τότε ο παρονομαστής (πιθανοφάνεια του δείγματος υπό την H_1) πρέπει να είναι σχετικά μεγαλύτερος από τον αριθμητή (πιθανοφάνεια του δείγματος υπό την H_0), από όπου συμπεραίνουμε ότι, σε αυτή την περίπτωση μάλλον θα πρέπει να ισχύει η H_1 (είμαστε στην περιοχή K). Ακολουθεί αυστηρή απόδειξη.

Απόδειξη. Έστω ένας άλλος έλεγχος ο οποίος έχει κρίνουσα συνάρτηση την φ^* και $P[I] = a$. Αν συμβολίσουμε με π, π^* την ισχύ του ελέγχου με βάση την φ και την φ^* αντίστοιχα, και $K = \{\mathbf{x}: \lambda(\mathbf{x}) < c\}$, $A = \{\mathbf{x}: \lambda(\mathbf{x}) \geq c\}$, τότε,

$$\pi - \pi^* = E(\phi(\mathbf{X}) \mid H_1) - E(\phi^*(\mathbf{X}) \mid H_1) = E(\phi(\mathbf{X}) - \phi^*(\mathbf{X}) \mid H_1) = \int_{\mathbf{x}\in R^\nu} (\phi(\mathbf{x}) - \phi^*(\mathbf{x}))\, f(\mathbf{x};\theta_1)\, d\mathbf{x}$$

$$= \int_{\mathbf{x}\in K} (\phi(\mathbf{x}) - \phi^*(\mathbf{x}))\, f(\mathbf{x};\theta_1)\, d\mathbf{x} + \int_{\mathbf{x}\in A} (\phi(\mathbf{x}) - \phi^*(\mathbf{x}))\, f(\mathbf{x};\theta_1)\, d\mathbf{x}$$

(με $\int_{\mathbf{x}\in R^\nu} g(\mathbf{x}) d\mathbf{x}$ συμβολίσουμε το ν-πλό ολοκλήρωμα ως προς $x_1, x_2, \ldots, x_\nu$). Όταν

$$\mathbf{x} \in K \;\; \textit{έχουμε ότι} \;\; \phi(\mathbf{x}) - \phi^*(\mathbf{x}) = 1 - \phi^*(\mathbf{x}) \geq 0, \;\; \textit{και} \;\; f(\mathbf{x};\theta_1) > \tfrac{1}{c} f(\mathbf{x};\theta_0)$$

$$\mathbf{x} \in A \;\; \textit{έχουμε ότι} \;\; \phi(\mathbf{x}) - \phi^*(\mathbf{x}) = 0 - \phi^*(\mathbf{x}) \leq 0, \;\; \textit{και} \;\; f(\mathbf{x};\theta_1) < \tfrac{1}{c} f(\mathbf{x};\theta_0)$$

Επομένως

$$\pi - \pi^* \geq \tfrac{1}{c} \int_{\mathbf{x}\in K} (\phi(\mathbf{x}) - \phi^*(\mathbf{x}))\, f(\mathbf{x};\theta_0)\, d\mathbf{x} + \tfrac{1}{c} \int_{\mathbf{x}\in A} (\phi(\mathbf{x}) - \phi^*(\mathbf{x}))\, f(\mathbf{x};\theta_0)\, d\mathbf{x}$$

$$= \tfrac{1}{c} \int_{\mathbf{x}\in R^\nu} (\phi(\mathbf{x}) - \phi^*(\mathbf{x}))\, f(\mathbf{x};\theta_0)\, d\mathbf{x} = \tfrac{1}{c} E(\phi(\mathbf{X}) - \phi^*(\mathbf{X}) \mid H_0) = \tfrac{1}{c}(a - a) = 0,$$

δηλαδή, $\pi \geq \pi^*$. □

Αν επιθυμούμε να κατασκευάσουμε έναν τυχαιοποιημένο έλεγχο (π.χ. όταν δεν μπορούμε να επιτύχουμε ακριβώς το a που θέλουμε όπως σε παραπάνω παράδειγμα), τότε το λήμμα Neyman – Pearson διατυπώνεται με τον αντίστοιχο τρόπο. Συγκεκριμένα, σε αυτή την περίπτωση, η κρίνουσα συνάρτηση

$$\phi(\mathbf{X}) = \begin{cases} 1, & \lambda(\mathbf{X}) < c \\ 0, & \lambda(\mathbf{X}) > c \\ \gamma, & \lambda(\mathbf{X}) = c \end{cases}$$

με $c > 0$ και $\gamma \in (0,1)$: $P[I] = E(\phi(\mathbf{X}) \mid H_0) = a$, οδηγεί σε έναν έλεγχο με τη μεγαλύτερη ισχύ ανάμεσα σε όλους τους ελέγχους με ε.σ. a. Η απόδειξη είναι ανάλογη αυτής του μη τυχαιοποιημένου ελέγχου.

Ας θεωρήσουμε στη συνέχεια ένα τυχαίο δείγμα $X_1, X_2, \ldots, X_\nu \sim N(\mu, \sigma^2)$, όπου η παράμετρος σ είναι γνωστή και ας προσπαθήσουμε να προσδιορίσουμε μέσω του λήμματος Neyman – Pearson τον ισχυρότατο έλεγχο (ΙΕ) σε επίπεδο σημαντικότητας (ε.σ.) a για τον έλεγχο της

$$H_0: \mu = \mu_0 \quad \text{έναντι της} \quad H_0: \mu = \mu_1$$

για δεδομένα μ_0, μ_1, με $\mu_1 > \mu_0$. Για το λόγο πιθανοφανειών $\lambda(\mathbf{x})$ έχουμε

$$\lambda(\mathbf{x})=\frac{\prod_{i=1}^{\nu} f(x_i;\mu_0)}{\prod_{i=1}^{\nu} f(x_i;\mu_1)}=\frac{\prod_{i=1}^{\nu}\frac{1}{\sqrt{2\pi\sigma^2}}\exp(-\frac{(x_i-\mu_0)^2}{2\sigma^2})}{\prod_{i=1}^{\nu}\frac{1}{\sqrt{2\pi\sigma^2}}\exp(-\frac{(x_i-\mu_1)^2}{2\sigma^2})}=\frac{\exp(-\sum_{i=1}^{\nu}\frac{(x_i-\mu_0)^2}{2\sigma^2})}{\exp(-\sum_{i=1}^{\nu}\frac{(x_i-\mu_1)^2}{2\sigma^2})}$$

οπότε

$$\ln\lambda(\mathbf{x})=-\sum_{i=1}^{\nu}\frac{(x_i-\mu_0)^2}{2\sigma^2}+\sum_{i=1}^{\nu}\frac{(x_i-\mu_1)^2}{2\sigma^2}=-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{\nu}(x_i-\mu_0)^2-\sum_{i=1}^{\nu}(x_i-\mu_1)^2\right).$$

Παρατηρούμε ότι

$$\lambda(\mathbf{x})<c\Leftrightarrow\ln\lambda(\mathbf{x})<\ln c'\Leftrightarrow\sum_{i=1}^{\nu}(x_i-\mu_0)^2-\sum_{i=1}^{\nu}(x_i-\mu_1)^2>-2\sigma^2\ln c\equiv c''$$

$$\Leftrightarrow\left(\sum_{i=1}^{\nu}x_i^2+\nu\mu_0^2-2\mu_0\sum_{i=1}^{\nu}x_i\right)-\left(\sum_{i=1}^{\nu}x_i^2+\nu\mu_1^2-2\mu_1\sum_{i=1}^{\nu}x_i\right)>c''$$

$$\Leftrightarrow 2(\mu_1-\mu_0)\sum_{i=1}^{\nu}x_i>c'-\nu\mu_0^2+\nu\mu_1^2\equiv c'''$$

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

$$\sum_{i=1}^{\nu}x_i>\frac{c'''}{2(\mu_1-\mu_0)}\equiv c_a.$$

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

$$K:\ \sum_{i=1}^{\nu}X_i>c_a$$

για κάποιο c_a το οποίο θα καθορίζεται από την $P(I) = P(\mathbf{X}\in K \mid H_0) = \alpha$. Επομένως, το c_a θα πρέπει α είναι τέτοιο ώστε

$$P(I)=P(\sum_{i=1}^{\nu}X_i>c_a\mid H_0)=P(\frac{\sum_{i=1}^{\nu}X_i-\nu\mu_0}{\sqrt{\nu\sigma^2}}>\frac{c_a-\nu\mu_0}{\sqrt{\nu\sigma^2}}/H_0)=a$$

και επειδή, κάτω από την H_0, ισχύει ότι $\sum_{i=1}^{\nu}X_i\sim N(\nu\mu_0,\nu\sigma^2)$, συμπεραίνουμε τελικά ότι

$$\frac{c_a-\nu\mu_0}{\sqrt{\nu\sigma^2}}=z_a\ \Rightarrow\ c_a=\nu\mu_0+\sqrt{\nu\sigma^2}\,z_a.$$

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:

$$K:\ T>z_a.$$

όπου $T=T(\mathbf{X})$ είναι η στατιστική συνάρτηση

$$T=T(\mathbf{X})=\frac{\overline{X}-\mu_0}{\sigma/\sqrt{\nu}}.$$

Η πιθανότητα σφάλματος τύπου ΙΙ και η ισχύς $\pi(\mu_1)$ του παραπάνω ελέγχου είναι ίσες με

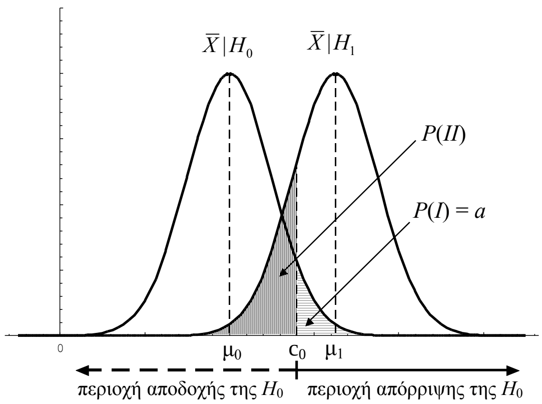
$$P(II) = (\frac{\overline{X}-\mu_0}{\sqrt{\sigma^2/\nu}} \le z_a | \mu=\mu_1) = P(\frac{\overline{X}-\mu_1}{\sqrt{\sigma^2/\nu}} \le \frac{\mu_0-\mu_1}{\sqrt{\sigma^2/\nu}} + z_a | \mu=\mu_1) = \Phi(\frac{\mu_0-\mu_1}{\sqrt{\sigma^2/\nu}} + z_a).$$

και

$$\pi(\mu_1) = 1 - P(II) = 1 - \Phi(\frac{\mu_0-\mu_1}{\sqrt{\sigma^2/\nu}} + z_a) = \Phi(\frac{\mu_1-\mu_0}{\sqrt{\sigma^2/\nu}} - z_a).$$

αντίστοιχα.

Στο επόμενο σχήμα, φαίνεται διαγραμματικά η περιοχή απόρριψης της υπόθεσης H_0: $\mu = \mu_0$ έναντι της H_1: $\mu = \mu_1$, με $\mu_1 > \mu_0$ καθώς και οι πιθανότητες των σφαλμάτων τύπου Ι και τύπου ΙΙ



Αν παράλληλα με την πιθανότητα σφάλματος τύπου Ι θέλαμε να περιορίσουμε και την πιθανότητα σφάλματος τύπου ΙΙ, θα μπορούσαμε να το επιτύχουμε, με την προϋπόθεση ότι είχαμε τη δυνατότητα να συλλέξουμε πολλές παρατηρήσεις (δηλαδή να αυξήσουμε το μέγεθος του δείγματος). Πράγματι, αν θέλαμε το σφάλμα τύπου ΙΙ να μην ξεπεράσει την τιμή δ ($0 < \delta < 1$, π.χ. δ=1%) θα έπρεπε να ισχύει

$$P[II] \le \delta \Leftrightarrow \Phi(\frac{\mu_0-\mu_1}{\sqrt{\sigma^2/\nu}} + z_a) \le \delta$$

από όπου παίρνουμε διαδοχικά

$$\Phi(\frac{\mu_0-\mu_1}{\sqrt{\sigma^2/\nu}} + z_a) \le \Phi(-z_\delta) \Leftrightarrow \frac{\mu_0-\mu_1}{\sigma/\sqrt{\nu}} + z_a \le -z_\delta \Leftrightarrow \nu \ge \frac{\sigma^2 (z_a + z_\delta)^2}{(\mu_1-\mu_0)^2}.$$

Στον έλεγχο που κατασκευάσαμε για την υπόθεση H_0: $\mu=\mu_0$ έναντι της H_1: $\mu=\mu_1$, με $\mu_1>\mu_0$, παρατηρούμε ότι η κρίσιμη περιοχή

$$K: \frac{\overline{X}-\mu_0}{\sigma/\sqrt{\nu}} > z_a.$$

δεν εξαρτάται από την τιμή μ_1. 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.

Προφανώς η τιμή της μ_1 επηρεάζει την $P[II]$ καθώς και την ισχύ του ελέγχου. Πιο συγκεκριμένα, όταν η τιμή του μ_1 αυξάνεται τότε η

$$P(II) = \Phi\left(\frac{\mu_0 - \mu_1}{\sqrt{\sigma^2 / \nu}} + z_a\right)$$

μειώνεται, ενώ η ισχύς

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

αυξάνεται. Δεδομένου ότι, οποιοδήποτε και αν είναι το μ_0, προκύπτει κρίσιμη περιοχή της ίδιας μορφής (K: $\sum X_i > c_a$ ή ισοδύναμα K: $\overline{X} > c$), η ίδια περιοχή K μπορεί να χρησιμοποιηθεί και για τον έλεγχο της υπόθεσης

$$H_0: \mu \leq \mu_0 \text{ έναντι της } H_1: \mu > \mu_0$$

(μονόπλευρη έναντι μονόπλευρης) οδηγώντας και πάλι σε βέλτιστο έλεγχο. Στην περίπτωση αυτή το c θα καθορίζεται μέσω της ισότητας

$$\sup P[I] = \sup_{\mu \leq \mu_0} P(\overline{X} > c \mid H_0) = a$$

και αφού

$$\sup_{\mu \leq \mu_0} P(\overline{X} > c \mid H_0) = \sup_{\mu \leq \mu_0} \left(1 - \Phi\left(\tfrac{c-\mu}{\sigma / \sqrt{\nu}}\right)\right) = 1 - \Phi\left(\tfrac{c-\mu_0}{\sigma / \sqrt{\sqrt{\nu}}}\right)$$

θα πρέπει να έχουμε

$$1 - \Phi\left(\tfrac{c-\mu_0}{\sigma / \sqrt{n}}\right) = a \Leftrightarrow \frac{c - \mu_0}{\sigma / \sqrt{\nu}} = z_a \Leftrightarrow c = \mu_0 + \frac{\sigma}{\sqrt{\nu}} z_a$$

από όπου προκύπτει και πάλι η ίδια κρίσιμη περιοχή

$$K: \frac{\overline{X} - \mu_0}{\sigma / \sqrt{\nu}} > z_a .$$

Επομένως ο ισχυρότατος έλεγχος (ΙΕ) για την υπόθεση H_0: $\mu = \mu_0$ έναντι της H_1: $\mu > \mu_0$ καθώς και ο ομοιόμορφα ισχυρότατος έλεγχος (ΟΙΕ) για την υπόθεση H_0: $\mu \leq \mu_0$ έναντι της H_1: $\mu > \mu_0$ έχει ως κρίσιμη περιοχή την ίδια με τον έλεγχο της απλής H_0: $\mu = \mu_0$ έναντι της απλής H_1: $\mu = \mu_1$ με $\mu_1 > \mu_0$.

Εργαζόμενοι με παρόμοιο τρόπο μπορούμε να κατασκευάσουμε την κρίσιμη περιοχή για τον έλεγχο της H_0: $\mu = \mu_0$ έναντι της H_1: $\mu = \mu_1$ με $\mu_1 < \mu_0$. Πιο συγκεκριμένα, θα έχουμε και πάλι

$$\lambda(\mathbf{x}) < c \Leftrightarrow 2(\mu_1 - \mu_0) \sum_{i=1}^{\nu} x_i > c'''$$

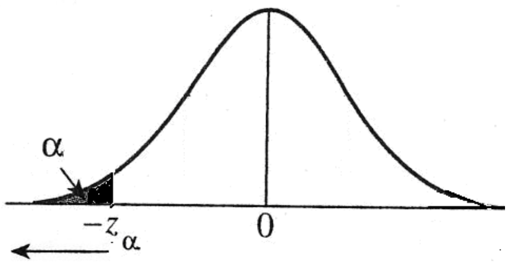
και επειδή τώρα ισχύει $\mu_1 - \mu_0 < 0$, η ανισότητα $\lambda(\mathbf{x}) < c$ είναι ισοδύναμη με την

$$\sum_{i=1}^{\nu} x_i < \frac{c'''}{2(\mu_1 - \mu_0)} \equiv c_a \, .$$

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

$$K\colon \ \sum_{i=1}^{\nu} X_i < c_a$$

και η τιμή της σταθεράς c_a βρίσκεται ίση με $c_a = \nu\mu_0 - \sqrt{\nu\sigma^2}\, z_a$ (βλέπε και ακόλουθο σχήμα).



Επομένως ο καλύτερος έλεγχος σε επίπεδο σημαντικότητας α για την υπόθεση H_0: $\mu=\mu_0$ έναντι της H_1: $\mu = \mu_1$, με $\mu_1 < \mu_0$, θα έχει περιοχή απόρριψης:

$$K\colon \ T < -z_a \, .$$

όπου $T = T(\mathbf{X})$ είναι η ίδια στατιστική συνάρτηση που χρησιμοποιήσαμε και για τους προηγούμενους ελέγχους. Ομοίως διαπιστώνεται ότι ή ίδια κρίσιμη περιοχή παρέχει ισχυρότατο έλεγχο (ΙΕ) για την υπόθεση H_0: $\mu = \mu_0$ έναντι της H_1: $\mu < \mu_0$ καθώς και ομοιόμορφα ισχυρότατο έλεγχο (ΟΙΕ) για την υπόθεση H_0: $\mu \geq \mu_0$ έναντι της H_1: $\mu < \mu_0$.

Συνοψίζοντας δίνουμε τον ακόλουθο πίνακα ο οποίος περιλαμβάνει όλες τις περιπτώσεις μονόπλευρων ελέγχων για το μέσο μ ενός κανονικού πληθυσμού, όταν το σ^2 είναι γνωστό.

Έλεγχος για το μ όταν το σ^2 είναι γνωστό (με ε.σ. a)				
$H_0{:}\mu=\mu_0$, $H_1{:}\mu=\mu_1$, $\mu_1>\mu_0$	$H_0{:}\mu=\mu_0$, $H_1{:}\mu>\mu_0$	$H_0{:}\mu\leq\mu_0$, $H_1{:}\mu>\mu_0$	$K\colon\ T > z_a$	Στατιστική συνάρτηση
$H_0{:}\mu=\mu_0$, $H_1{:}\mu=\mu_1$, $\mu_1<\mu_0$	$H_0{:}\mu=\mu_0$, $H_1{:}\mu<\mu_0$	$H_0{:}\mu\geq\mu_0$, $H_1{:}\mu<\mu_0$	$K\colon\ T < -z_a$	$T = \dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{\nu}}$

Στην περίπτωση που η διασπορά, σ^2, του πληθυσμού είναι γνωστή και το μέγεθος ν του δείγματος είναι μεγάλο (θεωρητικά $\nu \to +\infty$, στην πράξη $\nu > 30$), οι παραπάνω περιοχές απόρριψης ισχύουν για οποιονδήποτε πληθυσμό, όχι κατ' ανάγκη κανονικό αφού τότε μπορεί να εφαρμοσθεί το Κ.Ο.Θ. Όμως, στην περίπτωση αυτή, οι αντίστοιχοι έλεγχοι είναι **κατά προσέγγιση** επιπέδου σημαντικότητας a, γιατί η κατανομή της στατιστικής συνάρτησης ελέγχου, $T = (\overline{X} - \mu_0)/(\sigma/\sqrt{\nu})$, **δεν είναι,** στην περίπτωση αυτή, κανονική αλλά προσεγγίζεται από την κανονική. Φυσικά, όσο μεγαλύτερο είναι το δείγμα, τόσο καλύτερη είναι η προσέγγιση.

Ας θεωρήσουμε τώρα ένα τυχαίο δείγμα $X_1, X_2, \ldots, X_\nu \sim N(\mu,\sigma^2)$, όπου η παράμετρος μ είναι γνωστή και ας προσπαθήσουμε να προσδιορίσου μέσω του λήμματος Neyman – Pearson τον ισχυρότατο έλεγχο (ΙΕ) σε επίπεδο σημαντικότητας a για τον έλεγχο της

$$H_0: \sigma = \sigma_0 \text{ έναντι της } H_0: \sigma = \sigma_1$$

για δεδομένα σ_0, σ_1 με $\sigma_1 > \sigma_0$. Για το λόγο πιθανοφανειών $\lambda(\mathbf{x})$ έχουμε

$$\lambda(\mathbf{x})=\frac{\prod_{i=1}^{\nu} f(x_i;\sigma_0)}{\prod_{i=1}^{\nu} f(x_i;\sigma_1)}=\frac{\prod_{i=1}^{\nu}\frac{1}{\sqrt{2\pi\sigma_0^2}}\exp(-\frac{(x_i-\mu)^2}{2\sigma_0^2})}{\prod_{i=1}^{\nu}\frac{1}{\sqrt{2\pi\sigma_1^2}}\exp(-\frac{(x_i-\mu)^2}{2\sigma_1^2})}=\left(\frac{\sigma_1^2}{\sigma_0^2}\right)^{\nu/2}\frac{\exp(-\sum_{i=1}^{\nu}\frac{(x_i-\mu)^2}{2\sigma_0^2})}{\exp(-\sum_{i=1}^{\nu}\frac{(x_i-\mu)^2}{2\sigma_1^2})}$$

οπότε

$$\ln\lambda(\mathbf{x})=\frac{\nu}{2}\ln\left(\frac{\sigma_1^2}{\sigma_0^2}\right)-\left(\frac{1}{2\sigma_0^2}-\frac{1}{2\sigma_1^2}\right)\sum_{i=1}^{\nu}(x_i-\mu)^2 .$$

Παρατηρούμε ότι

$$\lambda(\mathbf{x})<c \Leftrightarrow \ln\lambda(\mathbf{x})<\ln c' \Leftrightarrow \frac{\nu}{2}\ln\left(\frac{\sigma_1^2}{\sigma_0^2}\right)-\left(\frac{1}{2\sigma_0^2}-\frac{1}{2\sigma_1^2}\right)\sum_{i=1}^{\nu}(x_i-\mu)^2<\ln c'\equiv c''$$

$$\Leftrightarrow -\left(\frac{1}{2\sigma_0^2}-\frac{1}{2\sigma_1^2}\right)\sum_{i=1}^{\nu}(x_i-\mu)^2<c''-\frac{\nu}{2}\ln\left(\frac{\sigma_1^2}{\sigma_0^2}\right)=c''' \Leftrightarrow \frac{\sigma_0^2-\sigma_1^2}{2\sigma_0^2\sigma_1^2}\sum_{i=1}^{\nu}(x_i-\mu)^2<c'''$$

και επειδή $\sigma_0^2-\sigma_1^2<0$, η ανισότητα $\lambda(\mathbf{x}) < c$ είναι ισοδύναμη με την

$$\sum_{i=1}^{\nu}(x_i-\mu)^2>c_a .$$

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

$$K:\ \sum_{i=1}^{\nu}(X_i-\mu)^2>c_a$$

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

$$P[I]=P\left(\sum_{i=1}^{\nu}(X_i-\mu)^2>c_a\,|\,H_0\right)=a$$

και επειδή, κάτω από την H_0, ισχύει ότι $\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma_0}\right)^2\sim\chi_\nu^2$, θα έχουμε

$$P[I]=\alpha \Leftrightarrow P\left(\sum_{i=1}^{\nu}(X_i-\mu)^2>c_a\,|\,H_0\right)=\alpha \Leftrightarrow P\left(\sum_{i=1}^{\nu}\left(\frac{X_i-\mu}{\sigma_0}\right)^2>\frac{c_a}{\sigma_0^2}\,|\,H_0\right)=a \Leftrightarrow \frac{c_a}{\sigma_0^2}=\chi_\nu^2(a)$$

οπότε

$$\frac{c_a}{\sigma_0^2} = \chi_\nu^2(a) \Leftrightarrow c_a = \sigma_0^2 \chi_\nu^2(a).$$

όπου με $\chi_\nu^2(a)$ συμβολίζεται το άνω a-σημείο της χι τετράγωνο κατανομής με ν β.ε. Άρα, σύμφωνα με το Λήμμα N-P, ο καλύτερος έλεγχος σε επίπεδο σημαντικότητας a για την υπόθεση H_0: σ = σ_0 έναντι της H_0: σ = σ_1 με $\sigma_1 > \sigma_0$.θα έχει περιοχή απόρριψης

$$K\colon \sum_{i=1}^{\nu} (X_i - \mu)^2 > \sigma_0^2 \chi_\nu^2(a)$$

ή ισοδύναμα K: $T > \chi_\nu^2(a)$ όπου $T = T(\mathbf{X})$ είναι η στατιστική συνάρτηση

$$T = T(\mathbf{X}) = \sum_{i=1}^{\nu} \left(\frac{X_i - \mu}{\sigma_0} \right)^2.$$

Η πιθανότητα σφάλματος τύπου II και η ισχύς $\pi(\sigma_1^2)$ του παραπάνω ελέγχου είναι ίσες με

$$P[II] = P(\sum_{i=1}^{\nu} (X_i - \mu)^2 \le \sigma_0^2\, \chi_\nu^2(a) \,|\, H_1) = P(\sum_{i=1}^{\nu} (\frac{X_i - \mu}{\sigma_1})^2 \le \frac{c_a}{\sigma_1^2} \,|\, H_1) = F_{\chi_\nu^2}(\frac{c_a}{\sigma_1^2}) = F_{\chi_\nu^2}\left(\frac{\sigma_0^2}{\sigma_1^2}\, \chi_\nu^2(a) \right),$$

$$\pi(\sigma_1^2) = 1 - F_{\chi_\nu^2}\left(\frac{\sigma_0^2}{\sigma_1^2}\, \chi_\nu^2(a) \right)$$

αντίστοιχα, όπου με $F_{\chi_\nu^2}(\cdot)$ έχουμε συμβολίσει την αθροιστική συνάρτηση κατανομής της χι τετράγωνο κατανομής με ν β.ε.

Στον έλεγχο που κατασκευάσαμε για την υπόθεση H_0: σ = σ_0 έναντι της H_0: σ = σ_1 με $\sigma_1 > \sigma_0$, παρατηρούμε ότι η κρίσιμη περιοχή

$$K\colon \sum_{i=1}^{\nu} (X_i - \mu)^2 > \sigma_0^2 \chi_\nu^2(a)$$

δεν εξαρτάται από την τιμή σ_1. Επομένως η ίδια κρίσιμη περιοχή μπορεί να χρησιμοποιηθεί και για τον έλεγχο της υπόθεσης

H_0: σ = σ_0 έναντι της H_0: σ > σ_0

(απλή έναντι μονόπλευρης) οδηγώντας και πάλι σε βέλτιστο έλεγχο. Λαμβάνοντας υπόψη ότι η ισχύς

$$\pi(\sigma_1^2) = 1 - F_{\chi_\nu^2}\left(\frac{\sigma_0^2}{\sigma_1^2}\, \chi_\nu^2(a) \right)$$

είναι αύξουσα συνάρτηση του σ_1,προκύπτει ότι η ίδια περιοχή K μπορεί να χρησιμοποιηθεί και για τον έλεγχο της υπόθεσης

H_0: $\sigma^2 \le \sigma_0^2$ έναντι της H_1: $\sigma^2 > \sigma_0^2$

(μονόπλευρη έναντι μονόπλευρης) οδηγώντας και πάλι σε βέλτιστο έλεγχο με

$$\sup P[I] = \sup_{\sigma \le \sigma_0} P(\sum_{i=1}^{\nu} (X_i - \mu)^2 > \sigma_0^2 \chi_\nu^2(a) \,|\, H_0) = a.$$

Επομένως ο ισχυρότατος έλεγχος (ΙΕ) για την υπόθεση H_0: σ = $σ_1$ με $σ_1 > σ_0$ καθώς και ο ομοιόμορφα ισχυρότατος έλεγχος (ΟΙΕ) για την υπόθεση H_0: $σ^2 \le σ_0^2$ έναντι της H_1: $σ^2 > σ_0^2$ έχει ως κρίσιμη περιοχή την ίδια με τον έλεγχο της απλής H_0: σ = $σ_0$ έναντι της απλής H_0: σ = $σ_1$ με $σ_1 > σ_0$.

Εργαζόμενοι με παρόμοιο τρόπο μπορούμε να κατασκευάσουμε την κρίσιμη περιοχή για τον έλεγχο της H_0: σ = $σ_0$ έναντι της H_0: σ = $σ_1$ με $σ_1 < σ_0$. Πιο συγκεκριμένα, θα έχουμε και πάλι

$$\lambda(\mathbf{x}) < c \Leftrightarrow \Leftrightarrow \frac{\sigma_0^2 - \sigma_1^2}{2\sigma_0^2\sigma_1^2}\sum_{i=1}^{\nu}(x_i - \mu)^2 < c'''$$

και επειδή τώρα ισχύει $\sigma_0^2 - \sigma_1^2 > 0$, η ανισότητα λ(**x**) < c είναι ισοδύναμη με την

$$\sum_{i=1}^{\nu}(x_i - \mu)^2 < c_a .$$

Άρα η κρίσιμη περιοχή του ελέγχου θα έχει τη μορφή

$$K\colon \sum_{i=1}^{\nu}(X_i - \mu)^2 < c_a$$

και η τιμή της σταθεράς c_a βρίσκεται ίση με $c_a = \sigma_0^2\chi_\nu^2(1-a)$. Επομένως ο καλύτερος έλεγχος σε επίπεδο σημαντικότητας $α$ για την υπόθεση H_0: σ = $σ_0$ έναντι της H_0: σ = $σ_1$ με $σ_1 < σ_0$.θα έχει περιοχή απόρριψης

$$K\colon T < \chi_\nu^2(1-a) .$$

όπου $T = T(\mathbf{X})$ είναι η ίδια στατιστική συνάρτηση που χρησιμοποιήσαμε και για τους προηγούμενους ελέγχους. Ομοίως διαπιστώνεται ότι ή ίδια κρίσιμη περιοχή παρέχει ισχυρότατο έλεγχο (ΙΕ) για την υπόθεση H_0: σ = $σ_0$ έναντι της H_0: σ < $σ_0$.καθώς και σε ομοιόμορφα ισχυρότατο έλεγχο (ΟΙΕ) για την υπόθεση H_0: σ ≥ $σ_0$ έναντι της H_0: σ < $σ_0$.

Συνοψίζοντας δίνουμε τον ακόλουθο πίνακα ο οποίος περιλαμβάνει όλες τις περιπτώσεις μονόπλευρων ελέγχων για το μέσο μ ενός κανονικού πληθυσμού, όταν το $σ^2$ είναι γνωστό.

Έλεγχοι για το *σ* όταν το *μ* είναι γνωστό				
H_0: $\sigma=\sigma_0$, H_1: $\sigma=\sigma_1$, $\sigma_1>\sigma_0$	H_0: $\sigma=\sigma_0$, H_1: $\sigma>\sigma_0$	H_0: $\sigma\le\sigma_0$, H_1: $\sigma>\sigma_0$	$K\colon T > \chi_\nu^2(a)$	Στατιστική συνάρτηση $T = \sum_{i=1}^{\nu}\left(\frac{X_i - \mu}{\sigma_0}\right)^2$
H_0: $\sigma=\sigma_0$, H_1: $\sigma=\sigma_1$, $\sigma_1<\sigma_0$	H_0: $\sigma=\sigma_0$, H_1: $\sigma<\sigma_0$	H_0: $\sigma\ge\sigma_0$, H_1: $\sigma<\sigma_0$	$K\colon T < \chi_\nu^2(1-a)$	

2.2. Κατασκευή Ομοιόμορφα Ισχυρότατων Ελέγχων (ΟΙΕ) μέσω της Εκθετικής Οικογένειας Κατανομών (ΕΟΚ)

Σε αυτή την παράγραφο προτείνεται μια μέθοδος κατασκευής βέλτιστων ελέγχων για μονόπλευρες υποθέσεις μέσω της (μονοπαραμετρικής) εκθετικής οικογένειας κατανομών.

Ορισμός 2. *Εκθετική οικογένεια κατανομών* (ΕΟΚ) *καλείται η οικογένεια των κατανομών με σ.π.π. ή σ.π. της μορφής*

$$f(x;\theta)=\beta(\theta)e^{\eta(\theta)T(x)}h(x)\,,\ x\in\mathrm{R} \qquad (\text{ή ισοδύναμα, } \ln f(x;\theta)=-B(\theta)+\eta(\theta)T(x)+H(x))$$

όπου β, η, T, h είναι πραγματικές συναρτήσεις με h > 0 *και* β > 0 (*και το στήριγμα* {*x*: *f* (*x*;*θ*) > 0} *της f δεν εξαρτάται από την παράμετρο θ*).

Αποδεικνύεται ότι αν μια τυχαία μεταβλητή ανήκει στην εκθετική οικογένεια κατανομών με με σ.π.π. ή σ.π. της μορφής $f(x;\theta)=\exp(-B(\theta)+\eta(\theta)T(x)+H(x))$, τότε η μέση τιμή και η διακύμανση της τυχαίας μεταβλητής $T(X)$ δίνονται από τους τύπους

$$E[T(X)]=\frac{B'(\theta)}{\eta'(\theta)}, \qquad V[T(X)]=\left(\frac{B'(\theta)}{\eta'(\theta)}\right)'\frac{1}{\eta'(\theta)}.$$

Αν ένα τυχαίο δείγμα $X_1, X_2, \ldots, X_\nu \sim f(x;\theta)\in$ΕΟΚ τότε για τον έλεγχο της H_0: $\theta=\theta_0$ έναντι της H_1: $\theta=\theta_1>\theta_0$ σε ε.σ. *a*, ο ισχυρότατος έλεγχος (ΙΕ) έχει περιοχή απόρριψης της μορφής $K:\lambda(\mathbf{X})<c$, όπου $c>0$: $P[I]=a$ (από το λήμμα Neyman-Pearson). Επειδή η $f(x;\theta)$ ανήκει στην εκθετική οικογένεια κατανομών θα είναι

$$K:\ \lambda(\mathbf{x})=\frac{\prod_{i=1}^{\nu} f(x_i;\theta_0)}{\prod_{i=1}^{\nu} f(x_i;\theta_1)}=\frac{\prod_{i=1}^{\nu}\beta(\theta_0)e^{\eta(\theta_0)T(x_i)}h(x_i)}{\prod_{i=1}^{\nu}\beta(\theta_1)e^{\eta(\theta_1)T(x_i)}h(x_i)}=\left(\frac{\beta(\theta_0)}{\beta(\theta_1)}\right)^{\nu}e^{(\eta(\theta_0)-\eta(\theta_1))\sum_{i=1}^{\nu}T(x_i)}<c$$

και επομένως η ανισότητα $\lambda(\mathbf{x})<c$, οδηγεί στη συνθήκη

$$(\eta(\theta_0)-\eta(\theta_1))\sum_{i=1}^{\nu}T(x_i)<c'.$$

Παρατηρούμε ότι αν η συνάρτηση $\eta(\theta)$ είναι αύξουσα (↑) τότε $\eta(\theta_0)-\eta(\theta_1)<0$ και επομένως η κρίσιμη περιοχή (περιοχή απόρριψης της H_0) θα έχει τη μορφή

$$K:\ \sum_{i=1}^{\nu}T(X_i)>c.$$

Αντίθετα, αν η $\eta(\theta)$ είναι φθίνουσα (↓) τότε

$$K: \ \sum_{i=1}^{\nu} T(X_i) < c,$$

όπου $c > 0$: $P[I] = a$. Άρα αν $f(x;\theta) \in EOK$ τότε προκύπτει άμεσα η περιοχή απόρριψης που οδηγεί σε ΙΕ για την H_0: $\theta = \theta_0$ έναντι της H_1: $\theta = \theta_1 > \theta_0$, αρκεί να προσδιορίσουμε την μονοτονία της η. Όμοια προκύπτει ΙΕ για την H_0: $\theta = \theta_0$ έναντι της H_1: $\theta = \theta_1 < \theta_0$ (η ανισότητα στην περιοχή απόρριψης θα είναι αντίθετη). Γενικότερα, μπορούμε να αποδείξουμε ότι οι παραπάνω περιοχές απόρριψης οδηγούν σε ομοιόμορφα ισχυρότατους ελέγχους για μονόπλευρες υποθέσεις της μορφής H_0: $\theta \leq \theta_0$ με H_1: $\theta > \theta_0$ και H_0: $\theta \geq \theta_0$ με H_1: $\theta < \theta_0$ (βλ. θεώρημα που ακολουθεί). Ένας έλεγχος καλείται *ομοιόμορφα ισχυρότατος* (ΟΙΕ) αν, για κάθε τιμή του θ υπό την H_1, *έχει τη μεγαλύτερη ισχύ ανάμεσα σε όλους τους ελέγχους με* ε.σ. a.

Θεώρημα 2 (*ΟΙΕ στην ΕΟΚ*). *Έστω* $f(x;\theta) \in EOK$ με $\eta(\theta) \uparrow$, *και ότι θέλουμε να ελέγξουμε την υπόθεση*

$$H_0: \theta \leq \theta_0 \ \text{ έναντι της } H_1: \theta > \theta_0,$$

σε ε.σ. a (δηλ. $P[I] = \max_{H_0} P_\theta[I] = a$). *Ο έλεγχος με κρίσιμη περιοχή* (*απόρριψης της* H_0) *της μορφής*

$$K: \ \sum_{i=1}^{\nu} T(X_i) > c,$$

όπου $c > 0$ *τέτοιο ώστε* $P_{\theta_0}[I] = a$, είναι *Ομοιόμορφα Ισχυρότατος Έλεγχος* (*ΟΙΕ*). *Μάλιστα, ισχύει ότι* $\max_{H_0} P_\theta[I] = P_{\theta_0}[I]$.

Ισοδύναμα μπορούμε να πούμε ότι η κρίνουσα συνάρτηση

$$\phi(\mathbf{X}) = \begin{cases} 1, & \sum_{i=1}^{\nu} T(X_i) > c \\ 0, & \sum_{i=1}^{\nu} T(X_i) \leq c \end{cases}$$

με $c > 0$ *τέτοιο ώστε* $P_{\theta_0}[I] = E(\phi(\mathbf{X}) \mid \theta_0) = a$, οδηγεί σε ΟΙΕ.

Απόδειξη. Αφού $f(x;\theta) \in EOK$ με $\eta(\theta) \uparrow$, για $\theta < \theta_0$ θα ισχύει ότι $\eta(\theta) < \eta(\theta_0)$, και για $\mathbf{x} \in K$,

$$\frac{f(\mathbf{x};\theta_0)}{f(\mathbf{x};\theta)} = \frac{\prod_{i=1}^{\nu} \beta(\theta_0) e^{\eta(\theta_0)T(x_i)} h(x_i)}{\prod_{i=1}^{\nu} \beta(\theta) e^{\eta(\theta)T(x_i)} h(x_i)} = \frac{\beta(\theta_0)^\nu}{\beta(\theta)^\nu} e^{(\eta(\theta_0)-\eta(\theta))\sum_{i=1}^{\nu} T(x_i)} > \frac{\beta(\theta_0)^\nu}{\beta(\theta)^\nu} e^{(\eta(\theta_0)-\eta(\theta))c} \equiv c' > 0$$

και όμοια $f(\mathbf{x};\theta_0)/f(\mathbf{x};\theta) < c'$ για $\mathbf{x} \in A$. Αν συμβολίσουμε $b = P_\theta[I] = E(\phi(\mathbf{X}) \mid \theta)$ τότε

$$P_{\theta_0}[I]-P_\theta[I]=E(\phi(\mathbf{X})\,|\,\theta_0)-b\ =\int_{\mathbf{x}}\phi(\mathbf{x})\,f(\mathbf{x};\theta_0)\,d\mathbf{x}-b\int_{\mathbf{x}} f(\mathbf{x};\theta_0)\,d\mathbf{x}$$

$$=\int_{\mathbf{x}\in K}(\phi(\mathbf{x})-b)\,f(\mathbf{x};\theta_0)\,d\mathbf{x}+\int_{\mathbf{x}\in A}(\phi(\mathbf{x})-b)\,f(\mathbf{x};\theta_0)\,d\mathbf{x}$$

$$\geq c'\int_{\mathbf{x}\in K}(\phi(\mathbf{x})-b)\,f(\mathbf{x};\theta)\,d\mathbf{x}+c'\int_{\mathbf{x}\in A}(\phi(\mathbf{x})-b)\,f(\mathbf{x};\theta)\,d\mathbf{x}$$

$$=c'(E(\phi(\mathbf{X})\,|\,\theta)-b)=c'(b-b)=0\,,$$

για $\theta\leq\theta_0$. Επομένως, $\max_{\theta\leq\theta_0}P_\theta[I]=P_{\theta_0}[I]=a$. Έστω τώρα άλλος έλεγχος με κρίνουσα συνάρτηση φ^* και $\max_{H_0}P^*_\theta[I]=a$. Συμβολίζουμε με $\pi(\theta)$, $\pi^*(\theta)$, για $\theta>\theta_0$ την ισχύ των φ, φ^* αντίστοιχα. Όμοια με παραπάνω, για $\theta>\theta_0$, βρίσκουμε ότι

$$\frac{f(\mathbf{x};\theta_0)}{f(\mathbf{x};\theta)}<c',\ \text{για } \mathbf{x}\in K,\quad \text{και}\quad \frac{f(\mathbf{x};\theta_0)}{f(\mathbf{x};\theta)}>c',\ \text{για } \mathbf{x}\in A.$$

Επομένως, για $\theta>\theta_0$,

$$\pi(\theta)-\pi^*(\theta)=E(\phi(\mathbf{X})\,|\,\theta)-E(\phi^*(\mathbf{X})\,|\,\theta)=E(\phi(\mathbf{X})-\phi^*(\mathbf{X})\,|\,\theta)$$

$$=\int_{\mathbf{x}\in K}(\phi(\mathbf{x})-\phi^*(\mathbf{x}))\,f(\mathbf{x};\theta)\,d\mathbf{x}+\int_{\mathbf{x}\in A}(\phi(\mathbf{x})-\phi^*(\mathbf{x}))\,f(\mathbf{x};\theta)\,d\mathbf{x}$$

$$\geq\tfrac{1}{c'}\int_{\mathbf{x}\in K}(\phi(\mathbf{x})-\phi^*(\mathbf{x}))\,f(\mathbf{x};\theta_0)\,d\mathbf{x}+\tfrac{1}{c'}\int_{\mathbf{x}\in A}(\phi(\mathbf{x})-\phi^*(\mathbf{x}))\,f(\mathbf{x};\theta_0)\,d\mathbf{x}$$

$$=\tfrac{1}{c'}E(\phi(\mathbf{X})-\phi^*(\mathbf{X})\,|\,\theta_0)=\tfrac{1}{c'}(P_{\theta_0}[I]-P^*_{\theta_0}[I])=\tfrac{1}{c'}(a-P^*_{\theta_0}[I])\geq 0\,,$$

δηλαδή, $\pi(\theta)\geq\pi^*(\theta)$ για κάθε $\theta>\theta_0$ (H_1). (*Για διακριτές κατανομές ισχύει παρόμοια απόδειξη με αθροίσματα*). □

Όμοια καλύπτονται όλες οι περιπτώσεις μονόπλευρων ελέγχων. Συγκεκριμένα, αν $f(x;\theta)\in$: ΕΟΚ (δηλαδή, $f(x;\theta)=\beta(\theta)e^{\eta(\theta)T(x)}h(x)$), ο ακόλουθος πίνακας περιλαμβάνει τις περιοχές απόρριψης της H_0 που οδηγούν σε ΟΙΕ, σε όλες τις δυνατές περιπτώσεις μονόπλευρων ελέγχων.

	$\eta(\theta)\uparrow$,	$\eta(\theta)\downarrow$,	$\eta(\theta)\uparrow$,	$\eta(\theta)\downarrow$,
H	H_0: $\theta\leq(=)\,\theta_0$ H_1: $\theta>\theta_0$	H_0: $\theta\leq(=)\,\theta_0$ H_1: $\theta>\theta_0$	H_0: $\theta\geq(=)\,\theta_0$ H_1: $\theta<\theta_0$	H_0: $\theta\geq(=)\,\theta_0$ H_1: $\theta<\theta_0$
ΟΙΕ	$K:\ \sum_{i=1}^{\nu}T(X_i)>c$	$K:\ \sum_{i=1}^{\nu}T(X_i)<c$	$K:\ \sum_{i=1}^{\nu}T(X_i)<c$	$K:\ \sum_{i=1}^{\nu}T(X_i)>c$

Η σταθερά c προσδιορίζεται έτσι ώστε να ισχύει $P_{\theta_0}[I]=a$.

Παράδειγμα 5: Έστω τυχαίο δείγμα X_1, X_2, …X_ν από τη Διωνυμική κατανομή $b(n,p)$ με n γνωστό. Ισχύει ότι

$$\ln f(x;p) = \ln\binom{n}{x} + x\ln p + (n-x)\ln(1-p) =$$
$$= n\ln(1-p) + \left(\ln\frac{p}{1-p}\right)x + \ln\binom{n}{x} = B(p) + \eta(p)T(x) + H(x)$$

όπου

$$B(p) = n\ln(1-p), \quad \eta(p) = \ln\frac{p}{1-p}, \quad T(x) = x, \quad H(x) = \ln\binom{n}{x}.$$

(ισοδύναμα θα μπορούσαμε να γράψουμε

$$f(x) = \binom{n}{x}p^x(1-p)^{n-x} = (1-p)^n e^{\ln\frac{p}{1-p}\cdot x}\binom{n}{x} = \beta(p)e^{\eta(p)T(x)}h(x)$$

οπότε οδηγούμαστε στους ίδιους τύπους για τα $B(p), \eta(p), H(x)$). Επειδή

$$\eta(p) = \ln\frac{p}{1-p}, \quad \eta'(p) = \frac{1}{p} + \frac{1}{1-p} > 0$$

η συνάρτηση $\eta(\theta)$ είναι αύξουσα και για τον έλεγχο (π.χ.) H_0: $p \le p_0$ έναντι της H_1: $p > p_0$, η κρίσιμη περιοχή

$$K: \sum_{i=1}^{\nu} T(X_i) = \sum_{i=1}^{\nu} X_i > c$$

θα ορίζει ΟΙΕ.

Παράδειγμα 6: Έστω τυχαίο δείγμα X_1, X_2, …X_ν ~ $N(\mu,\sigma^2)$ με σ^2 γνωστό. Ισχύει ότι

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}} = \left(\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{\mu^2}{2\sigma^2}}\right)e^{\frac{\mu}{\sigma^2}x}\left(e^{-\frac{x^2}{2\sigma^2}}\right) = \beta(\mu)e^{\eta(\mu)T(x)}h(x)$$

και επειδή η συνάρτηση $\eta(\mu) = \mu/\sigma^2$, είναι αύξουσα, για τον έλεγχο H_0: $\mu \le \mu_0$ έναντι της H_1: $\mu > \mu_0$, η κρίσιμη περιοχή

$$K: \sum_{i=1}^{\nu} T(X_i) = \sum_{i=1}^{\nu} X_i > c$$

ορίζει ΟΙΕ. Αν θέλουμε $P[\mathrm{I}] = a$ τότε θα πρέπει,

$$P[I] = P(\sum_{i=1}^{\nu} X_i > c \mid H_0) = a \Leftrightarrow P(\frac{\overline{X}-\mu_0}{\sigma/\sqrt{\nu}} > \frac{c/\nu-\mu_0}{\sigma/\sqrt{\nu}} \mid H_0) = a \Leftrightarrow \frac{c/\nu-\mu_0}{\sigma/\sqrt{\nu}} = z_a \Leftrightarrow c = \nu\mu_0 + \sigma\sqrt{\nu}z_a,$$

όπου με z_a συμβολίζεται το άνω a-σημείο της N(0,1). Τελικά ο ΟΙΕ σε ε.σ. a, της υπόθεσης H_0: $\mu \leq \mu_0$ έναντι της H_1: $\mu > \mu_0$, θα έχει κρίσιμη περιοχή,

$$K: \ \frac{\overline{X} - \mu_0}{\sigma / \sqrt{\nu}} > z_a .$$

Με τον ίδιο τρόπο, βρίσκουμε ότι ο ΟΙΕ σε ε.σ. a, της υπόθεσης H_0: $\mu \geq \mu_0$ έναντι της H_1: $\mu < \mu_0$, θα έχει κρίσιμη περιοχή,

$$K: \ \frac{\overline{X} - \mu_0}{\sigma / \sqrt{\nu}} < -z_a .$$

Προφανώς τα αποτελέσματα αυτά συμφωνούν με εκείνα που βρέθηκαν στην Παράγραφο 2.3.

Παράδειγμα 7: Έστω τυχαίο δείγμα $X_1, X_2, \ldots X_\nu$ από την κατανομή Γάμμα

$$f(x) = \frac{\lambda^a}{\Gamma(a)} x^{a-1} e^{-\lambda x},$$

όπου το a θεωρείται γνωστό και το λ άγνωστη παράμετρος. Ισχύει ότι

$$\ln f(x;\theta) = \ln \lambda^a - \lambda x + (\ln x^{a-1} - \ln \Gamma(a)) = -B(\lambda) + \eta(\lambda) T(x) + H(x)$$

όπου $B(\lambda) = -\ln \lambda^a, \eta(\lambda) = -\lambda, T(x) = x, H(x) = \ln x^{a-1} - \ln \Gamma(a)$. Επειδή $\eta(\lambda) = -\lambda$ $\downarrow$, π.χ. για τον έλεγχο

$$H_0: \lambda \geq \lambda_0 \quad \text{έναντι} \quad H_1: \lambda < \lambda_0,$$

η κρίσιμη περιοχή που οδηγεί σε ΟΙΕ είναι της μορφής

$$K: \ \sum_{i=1}^{\nu} T(X_i) = \sum_{i=1}^{\nu} X_i > c .$$

Ας θεωρήσουμε στη συνέχεια ένα τυχαίο δείγμα $X_1, X_2, \ldots, X_\nu$ από την εκθετική κατανομή με συνάρτηση πυκνότητας .πιθανότητας

$$f(x;\theta) = \frac{1}{\theta} e^{-\frac{x}{\theta}}, \ x > 0, \theta > 0.$$

και ας εξετάσουμε το πρόβλημα της κατασκευής ομοιόμορφα ισχυρότατου έλεγχου σε ε.σ. a της

$$\mathrm{H}_0 : \theta \leq \theta_0 \text{ έναντι της } \mathrm{H}_1 : \theta > \theta_0.$$

Δεδομένου ότι η εκθετική κατανομή αποτελεί ειδική περίπτωση της κατανομής Γάμμα με $a = 1$ και λ=1/θ , θα μπορούσαμε να γράψουμε την υπόθεση που μας ενδιαφέρει στην ισοδύναμη μορφή

$$\mathrm{H}_0 : \lambda = 1/\theta \geq 1/\theta_0 = \lambda_0 \ \text{ έναντι της } \ \mathrm{H}_1 : \lambda = 1/\theta < 1/\theta_0 = \lambda_0$$

και να συμπεράνουμε ότι η κρίσιμη περιοχή του ΟΙΕ είναι της μορφής $K: \sum_{i=1}^{\nu} X_i > c$. Στην περίπτωση που διαθέτουμε μεγάλο δείγμα (ν >30) ο προσδιορισμός της σταθεράς c ώστε να επιτευχθεί συγκεκριμένο ε.σ. a, μπορεί να γίνει εύκολα με χρήση του Κεντρικού Οριακού Θεωρήματος (ΚΟΘ). Από το ΚΟΘ γνωρίζουμε ότι ισχύει (κατά προσέγγιση)

$$\sum_{i=1}^{\nu} X_i \sim N(\nu\theta, \nu\theta^2)$$

και επομένως,

$$P[I] = P\left(\sum_{i=1}^{\nu} X_i > c_a \mid H_0\right) = P\left(\frac{\sum_{i=1}^{\nu} X_i - \nu\theta_0}{\sqrt{\nu\theta_0^2}} > \frac{c_a - \nu\theta_0}{\sqrt{\nu\theta_0^2}} \mid H_0\right) = 1 - \Phi\left(\frac{c_a - \nu\theta_0}{\sqrt{\nu\theta_0^2}}\right).$$

Η συνθήκη $P[I] = a$ δίνει διαδοχικά

$$P[I] = a \Leftrightarrow 1 - \Phi\left(\frac{c_a - \nu\theta_0}{\sqrt{\nu\theta_0^2}}\right) = a \Leftrightarrow \Phi\left(\frac{c_a - \nu\theta_0}{\sqrt{\nu\theta_0^2}}\right) = 1 - a = \Phi(z_a) \Leftrightarrow \frac{c_a - \nu\theta_0}{\sqrt{n\theta_0^2}} = z_a$$

οπότε $c_a = \nu\theta_0 + \theta_0\sqrt{\nu} z_a$ και η κρίσιμη περιοχή του ΟΙΕ είναι η

$$K: \sum_{i=1}^{\nu} X_i > \nu\theta_0 + \theta_0\sqrt{\nu} z_a \text{ ή ισοδύναμα } K: \bar{X} > \theta_0 + \frac{\theta_0}{\sqrt{\nu}} z_a$$

Τέλος, η πιθανότητα σφάλματος τύπου ΙΙ και η ισχύς θα δίνεται (προσεγγιστικά) από τους τύπους

$$P[II] = P\left(\sum_{i=1}^{\nu} X_i \le c_a \mid H_1\right) = P\left(\frac{\sum_{i=1}^{\nu} X_i - \nu\theta_1}{\sqrt{\nu\theta_1^2}} \le \frac{c_a - \nu\theta_1}{\sqrt{\nu\theta_1^2}} \mid H_1\right) = \Phi\left(\frac{c_a - \nu\theta_1}{\sqrt{\nu\theta_1^2}}\right) = \Phi\left(\frac{\nu(\theta_0 - \theta_1)}{\theta_1\sqrt{\nu}} + \frac{\theta_0}{\theta_1} z_a\right)$$

$$\pi(\theta) = 1 - P(II) = \Phi\left(\frac{\nu(\theta_1 - \theta_0)}{\theta_1\sqrt{\nu}} - \frac{\theta_0}{\theta_1} z_a\right)$$

2.3. Ασκήσεις

1. Έστω τυχαίο δείγμα $X_1, X_2, \ldots, X_\nu \sim N(\mu,\sigma^2)$, όπου ν = 16, σ = 6 και μ είναι άγνωστη παράμετρος.

α. Να βρεθεί η περιοχή απόρριψης K για τον έλεγχο της υπόθεσης

H_0: μ = μ_0 έναντι της H_0: μ = μ_1

για μ_0 = 50, μ_1 = 55 σε ε.σ. a = 0.05.

β. Να βρείτε την πιθανότητα σφάλματος τύπου ΙΙ και την ισχύ $\pi(\mu_1)$ του παραπάνω ελέγχου. Επίσης να βρεθεί πόσο θα έπρεπε να ήταν το μέγεθος του δείγματος ν ώστε η ισχύς να είναι τουλάχιστον $1 - \beta = 0.95$ όταν $\mu_0 = 0$, $\mu_1 = 1$, $a = 0.01$, σ = 1.

2. Επιθυμούμε να εξετάσουμε αν ο μέσος μ ενός κανονικού πληθυσμού $N(\mu,\sigma^2)$ είναι μ = 100 (H_0) ή μ > 100 (H_1). Λαμβάνουμε τυχαίο δείγμα μεγέθους ν = 10, στο οποίο βρίσκουμε μέσο $\bar{x} = 102$.

α. Αν σ = 5 να κάνετε τον συγκεκριμένο έλεγχο σε ε.σ. α = 5%.

β. Να υπολογίσετε το *p-value* του παραπάνω δείγματος και μέσω αυτού να επανελέγξετε την H_0 έναντι της H_1. Τι μπορεί να θεωρηθεί ότι εκφράζει το *p-value*;

γ. Αν ο ίδιος δειγματικός μέσος $\bar{x} = 102$ προέκυπτε από δείγμα $\nu = 50$, τι απόφαση θα παίρναμε; Να βρείτε το αντίστοιχο *p-value*.

δ. Πόσο πρέπει να είναι το ν ώστε $\pi(101) = 1 - \beta \approx 0.95$;

3. Η μέση γεωργική παραγωγή ανά στρέμμα ενός συγκεκριμένου προϊόντος σε μια αγροτική περιοχή είναι 35 τόνοι με τυπική απόκλιση 3 τόνους. Μετά την χρησιμοποίηση ενός νέου λιπάσματος, η μέση παραγωγή σε 100 τυχαία επιλεγμένα στρέμματα ανήλθε στους 35.6 τόνους. Να ελέγξετε σε ε.σ. 5% αν η συγκεκριμένη αύξηση είναι στατιστικά σημαντική, δηλαδή αν το συγκεκριμένο λίπασμα ευνοεί την αύξηση της μέσης παραγωγής (θεωρείστε ότι η διασπορά της παραγωγής έχει παραμείνει η ίδια). Να βρείτε το αντίστοιχο *p-value* του δείγματος.

4. Με βάση ένα τυχαίο δείγμα $X_1, X_2, \ldots, X_\nu \sim N(\mu, 8^2)$ θέλουμε να ελέγξουμε σε ε.σ. α = 5% την H_0: μ = 80 έναντι της H_1: μ = 76. Να δειχθεί ότι η κρίσιμη περιοχή του ΙΕ είναι της μορφής K: $\bar{x} < c_a$. Να βρεθεί το ελάχιστο ν και το c_a έτσι ώστε $P[\text{I}] \le 0.05$, $P[\text{II}] \le 0.10$.

5. Έστω ένα τυχαίο δείγμα $X_1, X_2, \ldots, X_\nu \sim N(0, \sigma^2)$.

α. Αν ν = 10, να δοθεί η κρίσιμη περιοχή του ελέγχου σε ε.σ. a=5% της υπόθεσης H_0: σ = 1 έναντι της H_0: σ = 2. Στη συνέχεια να βρεθεί η πιθανότητα σφάλματος τύπου ΙΙ και η ισχύς του παραπάνω ελέγχου.

β. Να βρεθεί το ελάχιστο μέγεθος του δείγματος ν ώστε η ισχύς να ελέγχου σε ε.σ. a=5% της υπόθεσης H_0: σ = 1 έναντι της H_0: σ = 2 να είναι τουλάχιστον 0.95.

6. Επιθυμούμε να ελέγξουμε αν η τιμή πώλησης ενός αγαθού έχει διασπορά ίση με 9 ή μεγαλύτερη από 9. Λαμβάνοντας τυχαίο δείγμα μεγέθους $\nu = 30$ βρίσκουμε σε αυτό (δειγματική) διασπορά 12 και (δειγματικό) μέσο 101. Να ελέγξετε την παραπάνω υπόθεση σε ε.σ. 1% (μέσω κατάλληλης κρίσιμης περιοχής K και μέσω του αντίστοιχου *p-value*) θεωρώντας ότι η τιμή του αγαθού ακολουθεί κανονική κατανομή με μέσο 100.

7. Έστω τυχαίο δείγμα $X_1, X_2, \ldots, X_\nu$ όπου οι τυχαίες μεταβλητές X_i ακολουθούν την κατανομή Poisson με παράμετρο λ.

α. Να δείξετε ότι η περιοχή απόρριψης του ισχυρότατου ελέγχου της $H_0 : \lambda = \lambda_0$ έναντι της $H_1 : \lambda = \lambda_1$ με $\lambda_1 > \lambda_0$ είναι της μορφής

$$\sum_{i=1}^{\nu} X_i > c_a .$$

β. Αν $\lambda_0 = 0.3$, $\lambda_1 = 0.4$, $\alpha = 0.05$, $\nu = 20$, να βρεθεί ο μικρότερος ακέραιος c_a που οδηγεί σε P[Ι] $\leq a$ (συντηρητικός έλεγχος με P[Ι] όσο το δυνατό πιο κοντά στο a). Να βρείτε την αντίστοιχη ισχύ.

γ. Αν $\lambda_0 = 0.3$, $\lambda_1 = 0.4$, $\alpha = 0.05$, $\nu = 20$, να κατασκευάσετε έναν τυχαιοποιημένο έλεγχο (c_a,γ) ώστε P[Ι] = α, και να βρείτε την ισχύ του.

δ. Σε ένα σημείο της εθνικής οδού ο μέσος αριθμός ατυχημάτων που συμβαίνουν την περίοδο Μάιο - Σεπτέμβριο είναι $\lambda_0 = 0.3$ ανά εβδομάδα. Μετά από κάποια έργα που έγιναν στο σημείο αυτό (τα οποία ολοκληρώθηκαν αρχές Μαΐου) καταγράφηκαν τα ακόλουθα ατυχήματα τις επόμενες 20 εβδομάδες: 1, 1, 1, 2, 0, 0, 1, 0, 0, 0, 0, 1, 2, 0, 0, 0, 1, 0, 1, 2. Αυξήθηκε η επικινδυνότητα του συγκεκριμένου σημείου; Να δώσετε το *p-value* του παραπάνω δείγματος.

8. Έστω τυχαίο δείγμα μεγέθους n από την κατανομή με σ.π.π.

$$f(x;\theta) = (\theta+1)x^{\theta},\ \ x \in [0,1],\ \theta > -1.$$

Να βρεθεί ΟΙΕ σε ε.σ. a για την υπόθεση H_0: $\theta = \theta_0$ έναντι της H_1: $\theta > \theta_0$ και να δοθεί η αντίστοιχη συνάρτηση ισχύος $\pi(\theta)$.

9. Έστω $X_1, X_2, \ldots, X_\nu$ τυχαίο δείγμα από την εκθετική κατανομή με σ.π.π.

$$f(x;\theta) = \frac{1}{\theta} e^{-\frac{x}{\theta}},\ \ x > 0,\ \theta > 0.$$

α. Να βρεθεί ομοιόμορφα ισχυρότατος έλεγχος σε ε.σ. a της

$$H_0 : \theta \leq \theta_0 \text{ έναντι της } H_1 : \theta > \theta_0.$$

β. Να βρεθεί η συνάρτηση ισχύος $\pi(\theta)$.

γ. Ο μέσος χρόνος διάρκειας μιας τηλεφωνικής κλήσης σε ένα δίκτυο είναι 2 λεπτά. Μετά από μια μείωση της χρονοχρέωσης των κλήσεων στο δίκτυο αυτό, καταγράφηκε η διάρκεια 300 κλήσεων η οποία ήταν συνολικά 660 λεπτά (δηλ. δειγματική μέση διάρκεια κλήσης 2.2 λεπτά). Υποθέτοντας ότι η διάρκεια μιας κλήσης ακολουθεί εκθετική κατανομή, να ελέγξετε σε ε.σ. α = 0.01 αν η αύξηση αυτή είναι στατιστικά σημαντική (δηλαδή αν επήλθε πραγματική αύξηση της μέσης διάρκειας μιας κλήσης μετά την μείωση της χρέωσης). Να βρεθεί το αντίστοιχο *p-value* του δείγματος.

10. Έστω $X_1, X_2, \ldots, X_\nu$ τυχαίο δείγμα από την κατανομή Weibull με σ.π.π.

$$f(x;\theta) = \beta \frac{1}{\theta} x^{\beta-1} e^{-x^{\beta}/\theta},\ \ x > 0,\ \theta > 0 \text{, και } \beta > 0 \text{ γνωστό.}$$

α. Να βρεθεί ομοιόμορφα ισχυρότατος έλεγχος σε ε.σ. a της

$$H_0 : \theta \leq \theta_0 \quad \text{έναντι της} \quad H_1 : \theta > \theta_0.$$

β. Να βρεθεί η συνάρτηση ισχύος $\pi(\theta)$.

11. Έστω ένα τυχαίο δείγμα: 0, 0, 0, 1, 2, 0, 1, 0, 2, 1, το οποίο προέρχεται από τη γεωμετρική κατανομή με σ.π.

$$f(x;\theta) = p(1-p)^x \quad x = 0,1,2,\ldots \quad p \in (0,1)$$

Να ελέγξετε σε ε.σ. a = 5% αν ισχύει η H_0: $p \leq p_0 = 1/3$ έναντι της H_1: $p > p_0 = 1/3$ μέσω του *p-value* του δείγματος.

3. ΓΕΝΙΚΕΥΜΕΝΟΣ ΛΟΓΟΣ ΠΙΘΑΝΟΦΑΝΕΙΩΝ - ΕΛΕΓΧΟΙ ΥΠΟΘΕΣΕΩΝ ΓΙΑ ΤΙΣ ΠΑΡΑΜΕΤΡΟΥΣ ΕΝΟΣ ΚΑΝΟΝΙΚΟΥ ΠΛΗΘΥΣΜΟΥ

3.1. Γενικευμένος λόγος πιθανοφανειών

Στην Ενότητα 2 είδαμε ότι μέσω του λήμματος Neyman-Pearson μπορούμε να κατασκευάζουμε ισχυρότατους ελέγχους απλών υποθέσεων. Διαπιστώσαμε επίσης ότι σε συνδυασμό με το λήμμα αυτό, μπορεί να αναπτυχθεί κατάλληλη μεθοδολογία για τη δημιουργία ΟΙΕ στην ΕΟΚ για τον έλεγχο μονόπλευρων υποθέσεων της μορφής H_0: $\theta \leq \theta_0$, H_1: $\theta > \theta_0$, και H_0: $\theta \geq \theta_0$, H_1: $\theta < \theta_0$.

Ωστόσο, στη γενικότερη περίπτωση όπου π.χ. μπορεί να έχουμε αμφίπλευρες υποθέσεις της μορφής H_0: $\theta = \theta_0$, H_1: $\theta \neq \theta_0$ ή μπορεί στους ελέγχους να εμπλέκονται άγνωστες παράμετροι (π.χ. στον έλεγχο για το μ κανονικής με το σ^2 να είναι άγνωστο) δεν μπορούμε να εργαστούμε μέσω του λήμματος Neyman-Pearson για τη δημιουργία ΟΙΕ. Σε αυτή την περίπτωση θα δούμε ότι μπορεί να χρησιμοποιηθεί μια επέκταση του λήμματος Neyman-Pearson που, αν και δεν οδηγεί πάντα σε βέλτιστο έλεγχο, δίνει αρκετά ικανοποιητικά αποτελέσματα στις περισσότερες περιπτώσεις.

Έστω τυχαίο δείγμα $X_1, X_2, ..., X_\nu$ από κατανομή με σ.π. ή σ.π.π. $f(x;\boldsymbol{\theta})$, $\boldsymbol{\theta} = (\theta_1,\ldots,\theta_k)$ και ότι θέλουμε να ελέγξουμε την

$$H_0\text{: } \boldsymbol{\theta}\in\Theta_0 \;\; \textit{έναντι της} \;\; H_1\text{: } \boldsymbol{\theta}\in\Theta_1.$$

Τότε, γενικεύοντας το λήμμα Neyman-Pearson, επιλέγουμε ως περιοχή απόρριψης της H_0 αυτήν που ορίζεται από την ανισότητα

$$\lambda(\mathbf{x}) \;=\; \frac{\sup\limits_{\boldsymbol{\theta}\in\Theta_0} L(\boldsymbol{\theta})}{\sup\limits_{\boldsymbol{\theta}\in\Theta_1} L(\boldsymbol{\theta})} < c,$$

όπου $L(\boldsymbol{\theta}) = f(\mathbf{x};\boldsymbol{\theta}) = \prod_{i=1}^{\nu} f(x_i;\boldsymbol{\theta})$, είναι η πιθανοφάνεια του δείγματος. Στην πράξη, για την απλοποίηση της διαδικασίας, τις περισσότερες φορές στον παρονομαστή λαμβάνεται το $\sup L(\boldsymbol{\theta})$ για όλα τα $\boldsymbol{\theta}\in\Theta_0\cup\Theta_1$. Επομένως, σύμφωνα κριτήριο του ***γενικευμένου λόγου πιθανοφανειών*** (ΓΛΠ) θα απορρίπτεται η H_0: $\boldsymbol{\theta}\in\Theta_0$ έναντι της H_1: $\boldsymbol{\theta}\in\Theta_1$ σε ε.σ. a όταν ισχύει

$$\lambda(\mathbf{x}) \;=\; \frac{\sup\limits_{\boldsymbol{\theta}\in\Theta_0} L(\boldsymbol{\theta})}{\sup\limits_{\boldsymbol{\theta}\in\Theta_0\cup\Theta_1} L(\boldsymbol{\theta})} = \frac{\sup\limits_{\boldsymbol{\theta}\in\Theta_0} \prod\limits_{i=1}^{\nu} f(X_i;\boldsymbol{\theta})}{\sup\limits_{\boldsymbol{\theta}\in\Theta_0\cup\Theta_1} \prod\limits_{i=1}^{\nu} f(X_i;\boldsymbol{\theta})} < c$$

όπου το c καθορίζεται έτσι ώστε $\sup_{H_0} P_{\boldsymbol{\theta}}[I] = a$. Έτσι, η κρίσιμη περιοχή του ελέγχου θα είναι η

$$K: \quad \lambda(\mathbf{X}) < c.$$

Από το κριτήριο γενικευμένου λόγου πιθανοφανειών δεν προκύπτει απαραίτητα βέλτιστος έλεγχος (δηλαδή με τη μεγαλύτερη ισχύ από όλους τους ελέγχους σε επίπεδο σημαντικότητας a, για κάθε $\boldsymbol{\theta}\in \Theta_1$). 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. Στις επόμενες παραγράφους θα εφαρμόσουμε το κριτήριο του γενικευμένου λόγου πιθανοφανειών για να κατασκευάσουμε ελέγχους για τις παραμέτρους κανονικών πληθυσμών.

3.2. Έλεγχος υποθέσεων για το μέσο ενός κανονικού πληθυσμού όταν η διασπορά είναι γνωστή.

Έστω $X_1, X_2, \ldots, X_\nu$ τυχαίο δείγμα από την κανονική κατανομή $N(\mu,\sigma^2)$ με σ^2 γνωστό. Θα αναζητήσουμε την κρίσιμη περιοχή του ελέγχου της H_0: $\mu = \mu_0$ έναντι της H_1: $\mu \neq \mu_0$, μέσω του γενικευμένου λόγου πιθανοφανειών (ΓΛΠ), σε ε.σ. a. Σε αυτή την περίπτωση ουσιαστικά αναζητούμε έλεγχο των υποθέσεων

$$H_0 : \mu\in\Theta_0 = \{\mu_0\} \;\; \textit{έναντι της} \;\; H_1 : \mu\in \Theta_1 = (-\infty, \mu_0) \cup (\mu_0, \infty),$$

και επομένως η περιοχή απόρριψης σύμφωνα με το ΓΛΠ θα είναι της μορφής

$$K: \ \lambda(\mathbf{X}) \ = \ \frac{\sup\limits_{\mu\in\Theta_0} L(\mu)}{\sup\limits_{\mu\in\Theta_0\cup\Theta_1} L(\mu)} = \frac{L(\mu_0)}{\sup\limits_{\mu\in R} L(\mu)} < c,$$

όπου $L(\mu)$ είναι η πιθανοφάνεια του τυχαίου δείγματος μεγέθους ν,

$$L(\theta) = \prod_{i=1}^{\nu} f(x_i,\theta) = \prod_{i=1}^{\nu} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x_i-\mu)^2}{\sigma^2}} = (2\pi\sigma^2)^{-\nu/2} e^{-\frac{1}{2\sigma^2}\sum\limits_{i=1}^{\nu}(x_i-\mu)^2} .$$

Στον παρονομαστή του $\lambda(\mathbf{X})$ θα πρέπει να βρούμε το μ που μεγιστοποιεί την πιθανοφάνεια $L(\mu)$ (ή ισοδύναμα την $\ln L(\mu)$) σε όλο τον παραμετρικό χώρο ($\mu\in$R). Είναι όμως γνωστό ότι η πιθανοφάνεια μεγιστοποιείται στο μ που είναι ίσο με την εκτιμήτρια μέγιστης πιθανοφάνειας (ΕΜΠ) του μ (εξ ορισμού της ΕΜΠ). Δηλαδή η $L(\mu)$ μεγιστοποιείται στο $L(\hat{\mu})$, όπου $\hat{\mu} = \overline{X}$. Πράγματι,

$$\frac{d}{d\mu}\ln L(\mu) = 0 \;\Leftrightarrow\; \frac{1}{\sigma^2}\sum_{i=1}^{\nu}(x_i - \mu) = 0 \;\Rightarrow\; \hat{\mu} = \bar{x} .$$

(και η δεύτερη παράγωγος στο $\bar{x}$ είναι αρνητική). Επομένως,

$$K:\ \lambda(\mathbf{X})=\frac{L(\mu_0)}{\sup\limits_{\mu\in R}L(\mu)}=\frac{L(\mu_0)}{L(\bar{X})}=\frac{e^{-\frac{1}{2\sigma^2}\sum\limits_{i=1}^{\nu}(X_i-\mu_0)^2}}{e^{-\frac{1}{2\sigma^2}\sum\limits_{i=1}^{\nu}(X_i-\bar{X})^2}}<c\ \Leftrightarrow\ \sum_{i=1}^{\nu}(X_i-\bar{X})^2-\sum_{i=1}^{\nu}(X_i-\mu_0)^2<c'$$

ή, ισοδύναμα,

$$-\nu\,\bar{X}^2-\nu\mu_0^2+2\nu\mu_0\bar{X}<c'\Leftrightarrow-\nu(\bar{X}-\mu_0)^2<c'\Leftrightarrow\left|\bar{X}-\mu_0\right|>c_a.$$

Το c_a καθορίζεται έτσι ώστε

$$P[I]=a\Leftrightarrow P(\left|\bar{X}-\mu_0\right|>c_a\mid H_0)=a\Leftrightarrow 1-P(-\frac{c_a}{\sigma/\sqrt{\nu}}\le\frac{\bar{X}-\mu_0}{\sigma/\sqrt{\nu}}\le\frac{c_a}{\sigma/\sqrt{\nu}}\mid H_0)=a$$

δηλαδή,

$$\Phi(\frac{c_a}{\sigma/\sqrt{\nu}})-\Phi(-\frac{c_a}{\sigma/\sqrt{\nu}})=1-a\Leftrightarrow 2\Phi(\frac{c_a}{\sigma/\sqrt{\nu}})-1=1-a\Leftrightarrow\frac{c_a}{\sigma/\sqrt{\nu}}=\Phi^{-1}(1-\frac{a}{2})=z_{a/2}$$

όπου με z_a συμβολίζεται το άνω a-σημείο της N(0,1).

Τελικά, η περιοχή απόρριψης της H_0 θα είναι

$$K:|T|>z_{a/2},\quad T=\frac{\bar{X}-\mu_0}{\sigma/\sqrt{\nu}}\quad.$$

Σε προηγούμενη παράγραφο είχαμε δώσει σε πίνακα τις κρίσιμες περιοχές για τις αντίστοιχες μονόπλευρες υποθέσεις. Συμπληρώνοντας και τον αμφίπλευρο έλεγχο παίρνουμε τον ακόλουθο πίνακα ο οποίος περιλαμβάνει όλες τις περιπτώσεις ελέγχων για τον μέσο μ κανονικού πληθυσμού, όταν το σ^2 είναι γνωστό.

<table>
<tr><th colspan="5">Έλεγχος για το μ όταν το σ² είναι γνωστό (με ε.σ. a)</th></tr>
<tr><td>$H_0:\mu=\mu_0$,
$H_1:\mu=\mu_1,\ \mu_1>\mu_0$</td><td>$H_0:\mu=\mu_0$,
$H_1:\mu>\mu_0$</td><td>$H_0:\mu\le\mu_0$,
$H_1:\mu>\mu_0$</td><td>$K:\ T>z_a$</td><td rowspan="3">Στατιστική συνάρτηση
$T=\frac{\bar{X}-\mu_0}{\sigma/\sqrt{\nu}}$</td></tr>
<tr><td>$H_0:\mu=\mu_0$,
$H_1:\mu=\mu_1,\ \mu_1<\mu_0$</td><td>$H_0:\mu=\mu_0$,
$H_1:\mu<\mu_0$</td><td>$H_0:\mu\ge\mu_0$,
$H_1:\mu<\mu_0$</td><td>$K:\ T<-z_a$</td></tr>
<tr><td colspan="3">$H_0:\ \mu=\mu_0$,
$H_1:\ \mu\ne\mu_0$,</td><td>$K:\ |T|>z_{a/2}$</td></tr>
</table>

3.3. Έλεγχος υποθέσεων για το μέσο ενός κανονικού πληθυσμού όταν η διασπορά είναι άγνωστη.

Η υπόθεση που κάναμε στην προηγούμενη ενότητα ότι η διασπορά, σ^2, του κανονικού πληθυσμού που μελετάμε είναι γνωστή, δεν είναι μια ρεαλιστική υπόθεση. Στην πράξη, η διασπορά του πληθυσμού είναι συνήθως άγνωστη. Στην παρούσα ενότητα, θα δούμε πώς εργαζόμαστε όταν η διασπορά του πληθυσμού είναι άγνωστη.

Έστω X_1, X_2, …, X_ν τυχαίο δείγμα από την κανονική κατανομή $N(\mu,\sigma^2)$ με σ^2 άγνωστο. Θα αναζητήσουμε και πάλι την κρίσιμη περιοχή του ελέγχου της H_0: $\mu = \mu_0$ έναντι της H_1: $\mu \neq \mu_0$, μέσω του γενικευμένου λόγου πιθανοφανειών (ΓΛΠ), σε ε.σ. a. Σε αυτή την περίπτωση ουσιαστικά αναζητούμε έλεγχο των υποθέσεων

$$H_0 : \boldsymbol{\theta} = (\mu,\sigma^2)\in\Theta_0 = \{\mu_0\}\times R_+ \text{ } \textit{έναντι της} \text{ } H_1 : \boldsymbol{\theta} = (\mu,\sigma^2)\in \Theta_1 = (-\infty,\mu_0)\cup(\mu_0,\infty)\times R_+$$

και η περιοχή απόρριψης θα προκύπτει μέσω της ανισότητας $\lambda(\mathbf{x}) < c$, όπου $\lambda(\mathbf{x})$ ο ΓΛΠ

$$\lambda(\mathbf{x}) = \frac{\sup_{\boldsymbol{\theta}\in\Theta_0} L(\boldsymbol{\theta})}{\sup_{\boldsymbol{\theta}\in\Theta_0\cup\Theta_1} L(\boldsymbol{\theta})} = \frac{\sup_{\sigma^2} L(\mu_0,\sigma^2)}{\sup_{\mu,\sigma^2} L(\mu,\sigma^2)}.$$

Η πιθανοφάνεια $L(\mu,\sigma^2)$ του τυχαίου δείγματος μεγέθους ν, είναι ίση με

$$L(\mu,\sigma^2) = \prod_{i=1}^{\nu} f(x_i,\mu,\sigma^2) = \prod_{i=1}^{\nu} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x_i-\mu)^2}{\sigma^2}} = (2\pi\sigma^2)^{-\nu/2} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{\nu}(x_i-\mu)^2}.$$

Η $L(\mu_0,\sigma^2)$ μεγιστοποιείται ως προς $\theta = \sigma^2$ όταν

$$\frac{d}{d\theta}\ln L = \frac{d}{d\theta}\left(-\frac{\nu}{2}\ln(2\pi\theta) - \frac{1}{2\theta}\sum_{i=1}^{\nu}(x_i-\mu_0)^2\right) = -\frac{\nu}{2\theta} + \frac{1}{2\theta^2}\sum_{i=1}^{\nu}(x_i-\mu_0)^2 = 0$$

Επομένως, μεγιστοποιείται στο

$$\hat{\theta} = \hat{\sigma}_0^2 = \tfrac{1}{\nu}\sum_{i=1}^{\nu}(x_i-\mu_0)^2.$$

Όμοια (εξισώνοντας τις μερικές παραγώγους ως προς μ, σ^2 με το 0), βρίσκουμε ότι η $L(\mu,\sigma^2)$ μεγιστοποιείται ως προς μ, σ^2 όταν

$$\hat{\mu} = \bar{x},\ \hat{\sigma}^2 = \tfrac{1}{\nu}\sum_{i=1}^{\nu}(x_i-\bar{x})^2.$$

Επομένως, η ανισότητα $\lambda(\mathbf{x}) < c$ οδηγεί , οδηγεί στη συνθήκη

$$\frac{\sup\limits_{\sigma^2} L(\mu_0,\sigma^2)}{\sup\limits_{\mu,\sigma^2} L(\mu,\sigma^2)} = \frac{\sup\limits_{\sigma^2} L(\mu_0,\hat{\sigma}_0^2)}{\sup\limits_{\mu,\sigma^2} L(\bar{x},\hat{\sigma}^2)} = \left(\frac{\sum_{i=1}^{\nu}(x_i-\mu_0)^2}{\sum_{i=1}^{\nu}(x_i-\bar{x})^2} \right)^{-\nu/2} < c$$

από την οποία παίρνουμε διαδοχικά

$$\frac{\sum_{i=1}^{\nu}(x_i-\mu_0)^2}{\sum_{i=1}^{\nu}(x_i-\bar{x})^2} > c' \Leftrightarrow \frac{\sum_{i=1}^{\nu}(x_i-\bar{x})^2+\nu(\bar{x}-\mu_0)^2}{\sum_{i=1}^{\nu}(x_i-\bar{x})^2} > c' \Leftrightarrow \frac{(\bar{x}-\mu_0)^2}{\frac{1}{\nu}\sum_{i=1}^{\nu}(x_i-\bar{x})^2} > c'-1 \Leftrightarrow \frac{|\bar{x}-\mu_0|}{s/\sqrt{\nu}} > c_a$$

όπου θέσαμε $s^2 = \frac{1}{\nu-1}\sum_{i=1}^{\nu}(x_i-\bar{x})^2$. Τελικά η περιοχή απόρριψης της H_0 θα έχει την μορφή,

$$T = T(\mathbf{X}) = \frac{|\bar{X}-\mu_0|}{S/\sqrt{\nu}} > c_a, \quad \text{όπου} \quad S^2 = \frac{1}{\nu-1}\sum_{i=1}^{\nu}(X_i-\bar{X}).$$

Παρατηρούμε τώρα ότι η τυχαία μεταβλητή γράφεται ως πηλίκο

$$T = \frac{\bar{X}-\mu_0}{S/\sqrt{\nu}} = \frac{\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{\nu}}}{\sqrt{\dfrac{1}{\nu-1}\sum_{i=1}^{\nu}\left(\dfrac{X_i-\bar{X}}{\sigma}\right)^2}} = \frac{Z}{\sqrt{Y/(\nu-1)}}$$

όπου για τις τυχαίες μεταβλητές,

$$Z = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{\nu}}, \qquad Y = \sum_{i=1}^{\nu}\left(\frac{X_i-\bar{X}}{\sigma}\right)^2$$

αποδεικνύεται ότι είναι ανεξάρτητες και επίσης είναι γνωστό ότι $Z \sim N(0,1)$ (υπό την H_0) και $Y \sim \chi^2_{\nu-1}$ (χι-τετράγωνο με $\nu-1$ βαθμούς ελευθερίας). Επομένως, υπό την H_0, η τυχαία μεταβλητή T ακολουθεί την κατανομή του *student* (από τον ορισμό της κατανομής αυτής) με $\nu-1$ βαθμούς ελευθερίας (συμβ. με $t_{\nu-1}$). Χρησιμοποιώντας το γεγονός αυτό μπορούμε να βρούμε το c_a έτσι ώστε $P[I]=a$. Πιο συγκεκριμένα θα έχουμε,

$$P[I]=a \Leftrightarrow P(|T|>c_a \mid H_0)=a \Leftrightarrow P(-c_a \le T \le c_a \mid H_0) = 1-a \Leftrightarrow c_a = F^{-1}_{T|H_0}(1-a/2) = t_{\nu-1}(a/2)$$

όπου με $t_{\nu-1}(a)$ συμβολίσουμε το άνω a-σημείο της κατανομής $t_{\nu-1}$. Τελικά, θα απορρίπτεται η H_0 σε ε.σ. a όταν,

$$K:\ |T| > t_{\nu-1}(a/2)$$

όπου

$$T = \frac{\bar{X}-\mu_0}{S/\sqrt{\nu}}, \quad S^2 = \frac{1}{\nu-1}\sum_{i=1}^{\nu}(X_i-\bar{X})^2.$$

Για τις μονόπλευρες υποθέσεις μπορούμε και πάλι να χρησιμοποιήσουμε την ίδια στατιστική συνάρτηση. Οι περιοχές απόρριψης της H_0 θα είναι της μορφής $T > c$ για H_1: $\mu>\mu_0$ και $T < c$ για H_1:

$\mu<\mu_0$. Για να έχουμε $P[I]$ = a μπορούμε εύκολα να διαπιστώσουμε ότι το c θα πρέπει να είναι ίσο με $t_{\nu-1}(a)$ και $-t_{\nu-1}(a)$ αντίστοιχα. Συγκεκριμένα, θα έχουμε τον επόμενο πίνακα ο οποίος περιλαμβάνει όλες τις περιπτώσεις ελέγχων για το μέσο μ κανονικού πληθυσμού, όταν το σ^2 είναι άγνωστο.

<table>
<tr><td colspan="5">Έλεγχοι για το μ όταν το σ^2 είναι άγνωστο (με ε.σ. a)</td></tr>
<tr><td>H_0:μ=μ_0,
H_1: μ=μ_1, μ_1>μ_0</td><td>H_0:μ=μ_0,
H_1:μ>μ_0,</td><td>H_0:$\mu\leq\mu_0$,
H_1:μ>μ_0</td><td>K: $T > t_{\nu-1}(a)$</td><td rowspan="3">Στατιστική συνάρτηση
$T = \dfrac{\overline{X}-\mu_0}{S/\sqrt{\nu}}$</td></tr>
<tr><td>H_0:μ=μ_0,
H_1:μ=μ_1, μ_1<μ_0</td><td>H_0:μ=μ_0,
H_1:μ<μ_0</td><td>H_0:$\mu\geq\mu_0$,
H_1:μ<μ_0</td><td>K: $T < -t_{\nu-1}(a)$</td></tr>
<tr><td colspan="3">H_0: μ=μ_0,
H_1: $\mu\neq\mu_0$,</td><td>K: $|T| > t_{\nu-1}(a/2)$</td></tr>
</table>

Αξίζει να σημειωθεί στην περίπτωση που διαθέτουμε τυχαίο δείγμα $X_1,\ X_2,...X_\nu$ από ένα πληθυσμό που ακολουθεί κανονική κατανομή με άγνωστη διασπορά σ^2, δεν θα μπορούσαμε να χρησιμοποιήσουμε ως στατιστική συνάρτηση ελέγχου την $\sqrt{\nu}(\overline{X}-\mu_0)/\sigma$ (με την οποία πραγματοποιήσαμε στην Ενότητα 3.1 τους αντίστοιχους ελέγχους όταν διασπορά σ^2 ήταν γνωστή) αφού τώρα πλέον δεν μπορούμε να υπολογίσουμε την τιμή της με χρήση του δείγματος. Εκείνο λοιπόν το οποίο κάνουμε τώρα είναι ότι χρησιμοποιούμε το δείγμα για να εκτιμήσουμε την άγνωστη διασπορά σ^2, μέσω της δειγματικής διασπορά S^2 οποία είναι αμερόληπτη εκτιμήτριά της, και ως στατιστική συνάρτηση ελέγχου χρησιμοποιήσαμε πλέον την $\sqrt{\nu}(\overline{X}-\mu_0)/s$ η οποία, υπό την H_0, ακολουθεί την κατανομή του *student* με ν–1 βαθμούς ελευθερίας (αντί της N(0, 1) που ακολουθεί η $\sqrt{\nu}(\overline{X}-\mu_0)/\sigma$).

3.4. Έλεγχος υποθέσεων για τη διασπορά ενός κανονικού πληθυσμού όταν η μέση τιμή είναι γνωστή.

Έστω τυχαίο δείγμα $X_1, X_2, ..., X_\nu \sim N(\mu,\sigma^2)$ με μ γνωστό και επιθυμούμε να ελέγξουμε την υπόθεση H_0: $\sigma^2 = \sigma_0^2$, H_1: $\sigma^2 \neq \sigma_0^2$. Θα αναζητήσουμε την περιοχή απόρριψης της H_0 (σε ε.σ. α) χρησιμοποιώντας το κριτήριο του γενικευμένου λόγου πιθανοφανειών. Οι υποθέσεις θα είναι της μορφής

$$H_0: \sigma^2 \in \Theta_0 = \{\sigma_0^2\},$$

$$H_1: \sigma^2 \in \Theta_1 = R_+ \backslash \{\sigma_0^2\}$$

και από το ΓΛΠ η περιοχή απόρριψης θα καθορίζεται μέσω της ανισότητας

$$\lambda(\mathbf{x}) \ = \ \frac{\sup\limits_{\sigma^2\in\Theta_0} L(\sigma^2)}{\sup\limits_{\sigma^2\in\Theta_0\cup\Theta_1} L(\sigma^2)} = \frac{L(\sigma_0^2)}{\sup\limits_{\sigma^2\in R_+} L(\sigma^2)} < c,$$

όπου $L(\theta)$, $\theta = \sigma^2$, είναι η πιθανοφάνεια του τυχαίου δείγματος μεγέθους ν. Έχουμε

$$L(\theta) = (2\pi\,\theta)^{-\nu/2} e^{-\frac{1}{2\theta}\sum\limits_{i=1}^{\nu}(x_i-\mu)^2}$$

και για να βρούμε το θ που μεγιστοποιεί την πιθανοφάνεια $L(\theta)$ σε όλο τον παραμετρικό χώρο ($\theta > 0$) θα έχουμε

$$\frac{d}{d\theta}\ln L(\theta) = 0 \Leftrightarrow \frac{d}{d\theta}\left(-\frac{\nu}{2}\ln(2\pi\,\theta) - \frac{1}{2\theta}\sum_{i=1}^{\nu}(x_i-\mu)^2\right) = 0 \Leftrightarrow \hat{\theta} = \frac{1}{\nu}\sum_{i=1}^{\nu}(x_i-\mu)^2.$$

(με δεύτερη παράγωγο αρνητική). Επομένως,

$$\lambda(\mathbf{x}) < c \Leftrightarrow \frac{L(\sigma_0^2)}{\sup\limits_{\theta\in R_+} L(\theta)} < c \Leftrightarrow \frac{L(\sigma_0^2)}{L(\hat{\theta})} < c \Leftrightarrow \frac{e^{-\frac{1}{2}\sum\limits_{i=1}^{\nu}(\frac{x_i-\mu}{\sigma_0})^2}}{(\frac{1}{\nu}\sum\limits_{i=1}^{\nu}(\frac{x_i-\mu}{\sigma_0})^2)^{-\nu/2} e^{-\frac{\nu}{2}}} < c$$

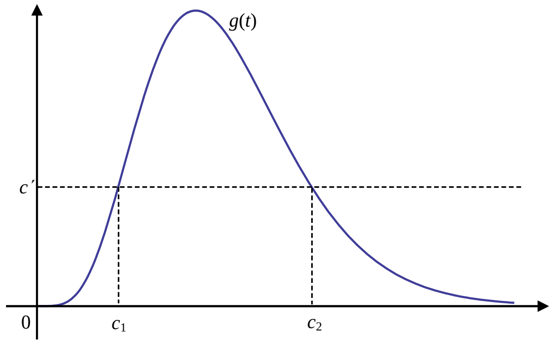
από όπου παίρνουμε

$$\lambda(\mathbf{x}) < c \Leftrightarrow T(\mathbf{x})^{\frac{\nu}{2}} e^{-\frac{1}{2}T(\mathbf{x})} < c' \Leftrightarrow g(T(\mathbf{x})) < c'$$

όπου

$$g(t) = t^{\nu/2} e^{-\frac{1}{2}t}, \qquad T(\mathbf{x}) = \sum_{i=1}^{\nu}(\frac{x_i-\mu}{\sigma_0})^2 \ .$$

Από το επόμενο σχήμα βλέπουμε ότι για να ισχύει η ανισότητα $g(t) < c'$ θα πρέπει να είναι $t < c_1$ ή $t > c_2$, για κάποια c_1, c_2.



Επομένως, και η κρίσιμη περιοχή θα πρέπει να είναι της μορφής

$$K\colon\ T > c_2 \ \text{ ή } \ T < c_1 .$$

Παρατηρούμε ότι υπό την H_0, η σ.σ. $T \sim \chi^2_\nu$. Άρα για να έχουμε $P[I] = a$, η περιοχή απόρριψης της H_0: $\sigma^2 = \sigma_0^2$ έναντι της H_1: $\sigma^2 \neq \sigma_0^2$ (με μ γνωστό), θα είναι η

$$K\colon\ T < \chi^2_\nu(1-\frac{a}{2}) \ \text{ ή } \ T > \chi^2_\nu(\frac{a}{2}) \quad \text{όπου} \quad T = \sum_{i=1}^{\nu}(\frac{X_i-\mu}{\sigma_0})^2 .$$

Για τους αντίστοιχους μονόπλευρους ελέγχους μπορούμε να βρούμε ΟΙΕ μέσω της ΕΟΚ (Θεώρημα 2). Πράγματι, διαπιστώνουμε ότι

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}} = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\left(-\frac{1}{2\sigma^2}\right)(x-\mu)^2} \cdot 1 = \beta(\sigma^2) e^{\eta(\sigma^2)T(x)} h(x),$$

και επειδή η συνάρτηση $\eta(\sigma^2) = -1/(2\sigma^2)$, είναι αύξουσα, για τον έλεγχο H_0: $\sigma^2 \leq \sigma_0^2$ έναντι της H_1: $\sigma^2 > \sigma_0^2$, η κρίσιμη περιοχή

$$K:\ \sum_{i=1}^{\nu} T(X_i) = \sum_{i=1}^{\nu}(X_i-\mu)^2 > c ,$$

ορίζει ΟΙΕ. Αν θέλουμε $P[I] = a$ τότε θα πρέπει,

$$P[I] = P(\sum_{i=1}^{\nu}(X_i-\mu)^2 > c_a \mid H_0) = P(\sum_{i=1}^{\nu}(\frac{X_i-\mu}{\sigma_0})^2 > \frac{c_a}{\sigma_0^2} \mid H_0) = a \Leftrightarrow \frac{c_a}{\sigma_0^2} = \chi^2_\nu(a) ,$$

όπου με $\chi^2_\nu(a)$ συμβολίζεται το άνω a-σημείο της χι τετράγωνο κατανομής με ν β.ε. Τελικά ο ΟΙΕ σε ε.σ. a, της υπόθεσης H_0: $\sigma^2 \leq \sigma_0^2$ έναντι της H_1: $\sigma^2 > \sigma_0^2$, θα έχει κρίσιμη περιοχή,

$$K:\ \frac{1}{\sigma_0^2}\sum_{i=1}^{\nu}(X_i-\mu)^2 > \chi^2_\nu(a) .$$

Με τον ίδιο τρόπο, βρίσκουμε ότι ο ΟΙΕ σε ε.σ. a, της υπόθεσης H_0: $\sigma^2 \geq \sigma_0^2$ έναντι της H_1: $\sigma^2 < \sigma_0^2$, θα έχει κρίσιμη περιοχή,

$$K: \ \frac{1}{\sigma_0^2}\sum_{i=1}^{\nu}(X_i-\mu)^2 < \chi_\nu^2(1-a).$$

ή ισοδύναμα

$$K: |T| > z_{a/2}, \quad T = \frac{\overline{X}-\mu_0}{\sigma/\sqrt{\nu}} \quad .$$

Σε προηγούμενη παράγραφο είχαμε δώσει σε πίνακα τις κρίσιμες περιοχές για τις αντίστοιχες μονόπλευρες υποθέσεις. Συμπληρώνοντας και τον αμφίπλευρο έλεγχο παίρνουμε τον ακόλουθο πίνακα ο οποίος περιλαμβάνει όλες τις περιπτώσεις ελέγχων για τη διασπορά σ^2 ενός κανονικού πληθυσμού, όταν το μ είναι γνωστό.

<table>
<tr><th colspan="5">Έλεγχοι για το σ όταν το μ είναι γνωστό</th></tr>
<tr><td>H_0: $\sigma=\sigma_0$,
H_1: $\sigma=\sigma_1$, $\sigma_1>\sigma_0$</td><td>H_0: $\sigma=\sigma_0$,
H_1: $\sigma>\sigma_0$</td><td>H_0: $\sigma\le\sigma_0$,
H_1: $\sigma>\sigma_0$</td><td>K: $T > \chi_\nu^2(a)$</td><td rowspan="3">Στατιστική συνάρτηση
$T=\sum_{i=1}^{\nu}\left(\frac{X_i-\mu}{\sigma_0}\right)^2$</td></tr>
<tr><td>H_0: $\sigma=\sigma_0$,
H_1: $\sigma=\sigma_1$, $\sigma_1<\sigma_0$</td><td>H_0: $\sigma=\sigma_0$,
H_1: $\sigma<\sigma_0$</td><td>H_0: $\sigma\ge\sigma_0$,
H_1: $\sigma<\sigma_0$</td><td>K: $T < \chi_\nu^2(1-a)$</td></tr>
<tr><td colspan="3">H_0: $\sigma=\sigma_0$,
H_1: $\sigma\ne\sigma_0$</td><td>K: $T < \chi_\nu^2(1-\frac{a}{2})$ ή $T > \chi_\nu^2(\frac{a}{2})$</td></tr>
</table>

3.5. Έλεγχος υποθέσεων για τη διασπορά ενός κανονικού πληθυσμού όταν η μέση τιμή είναι άγνωστη.

Έστω τυχαίο δείγμα X_1, X_2, ..., X_ν ~ N(μ,σ^2) με μ άγνωστό και επιθυμούμε να ελέγξουμε την υπόθεση H_0: $\sigma^2 = \sigma_0^2$, H_1: $\sigma^2 \ne \sigma_0^2$. Θα αναζητήσουμε και πάλι την περιοχή απόρριψης της H_0 (σε ε.σ. α) χρησιμοποιώντας το κριτήριο του ΓΛΠ. Οι υποθέσεις τώρα θα είναι

$$H_0: \boldsymbol{\theta} = (\mu,\sigma^2)\in\Theta_0 = R\times\{\sigma_0^2\}$$

$$H_1: \boldsymbol{\theta} = (\mu,\sigma^2)\in\Theta_1 = R\times(R_+\backslash\{\sigma_0^2\})$$

Η περιοχή απόρριψης θα καθορίζεται μέσω της ανισότητας

$$\lambda(\mathbf{x}) = \frac{\sup_{\boldsymbol{\theta}\in\Theta_0} L(\boldsymbol{\theta})}{\sup_{\boldsymbol{\theta}\in\Theta_0\cup\Theta_1} L(\boldsymbol{\theta})} = \frac{\sup_{\mu} L(\mu,\sigma_0^2)}{\sup_{\mu,\sigma^2} L(\mu,\sigma^2)} = \frac{L(\overline{x},\sigma_0^2)}{L(\overline{x},\frac{1}{\nu}\sum(x_i-\overline{x})^2)} < c$$

όπου

$$L(\mu,\sigma^2)=(2\pi\sigma^2)^{-\nu/2}e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{\nu}(x_i-\mu)^2}.$$

Έχουμε,

$$\lambda(\mathbf{x})<c\Leftrightarrow\frac{(2\pi\sigma_0^2)^{-\nu/2}e^{-\frac{1}{2\sigma_0^2}\sum_{i=1}^{\nu}(x_i-\bar{x})^2}}{(2\pi s^2)^{-\nu/2}e^{-\frac{1}{2s^2}\sum_{i=1}^{\nu}(x_i-\bar{x})^2}}<c\Leftrightarrow\left(\frac{(\nu-1)s^2}{\sigma_0^2}\right)^{\nu/2}e^{-\frac{1}{2}\frac{(\nu-1)s^2}{\sigma_0^2}}<c'\Leftrightarrow T^{\nu/2}e^{-\frac{1}{2}T}<c'\Leftrightarrow g(T)<c'$$

όπου

$$g(t)=t^{\nu/2}e^{-\frac{1}{2}t},\qquad T=\frac{(\nu-1)s^2}{\sigma_0^2}=\sum_{i=1}^{\nu}\left(\frac{x_i-\bar{x}}{\sigma_0}\right)^2.$$

Όπως και στην προηγούμενη παράγραφο, από την ανισότητα $g(t)<c'$ προκύπτει ότι $t<c_1$ ή $t>c_2$, και επομένως, η κρίσιμη περιοχή θα πρέπει να έχει την μορφή,

$$K:\ T>c_2\ \text{ ή }\ T<c_1.$$

Εδώ, η σ.σ. T υπό την H_0 ακολουθεί κατανομή $\chi^2_{\nu-1}$. Άρα για να ισχύει $P[\mathrm{I}]=a$, η περιοχή απόρριψης της H_0: $\sigma^2=\sigma_0^2$ έναντι της H_1: $\sigma^2\neq\sigma_0^2$ (μ άγνωστο), θα είναι η

$$K:\ T<\chi^2_{\nu-1}\left(1-\frac{a}{2}\right)\ \text{ ή }\ T>\chi^2_{\nu-1}\left(\frac{a}{2}\right)\quad\text{όπου}\quad T=\sum_{i=1}^{\nu}\left(\frac{x_i-\bar{x}}{\sigma_0}\right)^2=\frac{\nu-1}{\sigma_0^2}s^2.$$

Για τις μονόπλευρες υποθέσεις μπορούμε και πάλι να χρησιμοποιήσουμε την ίδια στατιστική συνάρτηση. Οι περιοχές απόρριψης της H_0 θα είναι της μορφής $T>c$ για H_1: $\sigma^2>\sigma_0^2$ και $T<c$ για H_1: $\sigma^2<\sigma_0^2$. Για να εξασφαλίσουμε ότι $P[I]=a$, διαπιστώνουμε ότι το c θα πρέπει να είναι ίσο με $\chi^2_{\nu-1}(a)$ και $\chi^2_{\nu-1}(1-a)$ αντίστοιχα. Συγκεκριμένα, θα έχουμε τον παρακάτω πίνακα ο οποίος περιλαμβάνει όλες τις περιπτώσεις ελέγχων για το σ^2 κανονικού πληθυσμού, όταν το μ είναι άγνωστο.

<table>
<tr><th colspan="5">Έλεγχοι για το σ όταν το μ είναι άγνωστο</th></tr>
<tr><td>H_0: $\sigma=\sigma_0$,
H_1: $\sigma=\sigma_1$, $\sigma_1>\sigma_0$</td><td>H_0: $\sigma=\sigma_0$,
H_1: $\sigma>\sigma_0$</td><td>H_0: $\sigma\leq\sigma_0$,
H_1: $\sigma>\sigma_0$</td><td>K: $T>\chi^2_{\nu-1}(a)$</td><td rowspan="3">Στατιστική συνάρτηση
$T=\frac{(\nu-1)S^2}{\sigma_0^2}=$
$=\sum_{i=1}^{\nu}\left(\frac{X_i-\bar{X}}{\sigma_0}\right)^2$</td></tr>
<tr><td>H_0: $\sigma=\sigma_0$,
H_1: $\sigma=\sigma_1$, $\sigma_1<\sigma_0$</td><td>H_0: $\sigma=\sigma_0$,
H_1: $\sigma<\sigma_0$</td><td>H_0: $\sigma\geq\sigma_0$,
H_1: $\sigma<\sigma_0$</td><td>K: $T<\chi^2_{\nu-1}(1-a)$</td></tr>
<tr><td colspan="3">H_0: $\sigma=\sigma_0$,
H_1: $\sigma\neq\sigma_0$</td><td>K: $T<\chi^2_{\nu-1}(1-\frac{a}{2})$ ή $T>\chi^2_{\nu-1}(\frac{a}{2})$</td></tr>
</table>

3.6. Ασκήσεις

1. Έστω $X_1,X_2, \ldots, X_\nu$ τυχαίο δείγμα από την κανονική κατανομή $N(\mu,\sigma^2)$ με σ^2 γνωστό.

α. Χρησιμοποιώντας την κρίσιμη περιοχή του ελέγχου της H_0: $\mu = \mu_0$ έναντι της H_1: $\mu \neq \mu_0$ να προσδιορίσετε το σύνολο των τιμών του μ_0 για τις οποίες η υπόθεση H_0: $\mu = \mu_0$ δεν απορρίπτεται (σε ε.σ. a, έναντι της H_1). Να συγκρίνετε το σύνολο αυτό με το γνωστό διάστημα εμπιστοσύνης για το μ, συντελεστού $1 - a$. Πως μπορούμε να χρησιμοποιήσουμε το δ.ε. για να ελέγξουμε τις παραπάνω υποθέσεις;

β. Να εκφράσετε το *p-value* ενός δείγματος $x_1,x_2,\ldots,x_\nu$ για τον έλεγχο της H_0: $\mu = \mu_0$ έναντι της H_1: $\mu \neq \mu_0$.

γ. Να βρείτε την συνάρτηση ισχύος του ελέγχου του ελέγχου της H_0: $\mu = \mu_0$ έναντι της H_1: $\mu \neq \mu_0$ αν $\nu = 100$, $\mu_0 = 10$, $\sigma = 10$, $\alpha = 5\%$. Αν ισχύει $\mu = 12$, με τι πιθανότητα θα πάρουμε σωστή απόφαση;

2. Ο δειγματικός μέσος με βάση ένα δείγμα μεγέθους $\nu = 100$ από την κανονική κατανομή $N(\mu,\sigma^2)$ βρέθηκε ίσος με 11, ενώ γνωρίζουμε ότι $\sigma = 4$. Να ελέγξετε αν ισχύει η H_0: $\mu = 10$ έναντι της H_1: $\mu \neq 10$ σε ε.σ. $a = 5\%$:

α. μέσω της κρίσιμης περιοχής *K*,

β. μέσω κατάλληλου δ.ε. για το μ, και

γ. μέσω του *p-value*

Δίνεται ότι $z_{0.025} \approx 1.96$.

3. Από τυχαίο δείγμα 150 εμπορικών καταστημάτων της πόλεως Α προέκυψε ότι το μέσο ετήσιο ενοίκιο τους είναι 13632 ευρώ, ενώ από απογραφή που έγινε στην πόλη Β προέκυψε ότι το μέσο ετήσιο ενοίκιο όλων των εμπορικών καταστημάτων της είναι 13452 ευρώ με τυπική απόκλιση 1864 ευρώ.

α. Να εξεταστεί αν μπορούμε να δεχθούμε, σε επίπεδο σημαντικότητας 0.10 ότι τα ενοίκια των εμπορικών καταστημάτων της πόλεως Α δε διαφέρουν από εκείνα της πόλεως Β, όταν είναι γνωστό ότι οι δύο κατανομές έχουν την ίδια διακύμανση.

β. Υπό τις ίδιες παραδοχές και δεδομένα, να εξεταστεί αν τα ενοίκια των εμπορικών καταστημάτων της Α είναι υψηλότερα από εκείνα της Β, σε ε.σ. 0.10.

Δίνεται ότι $z_{0.05} \approx 1.645$, $z_{0.1} \approx 1.281$.

4. Να κάνετε και πάλι τα Ερωτήματα (α), (β), (γ) της Άσκησης 1, αυτή τη φορά θεωρώντας ότι το σ δεν είναι γνωστό.

5. Ο δειγματικός μέσος και η δειγματική διασπορά από δείγμα μεγέθους ν =20 που προέρχεται από την κανονική κατανομή βρέθηκαν ίσα με 12 και 22.8 αντίστοιχα. Να ελέγξετε αν ισχύει η H$_0$: μ = 10 έναντι της H$_1$: μ ≠ 10 σε ε.σ. a = 5%

α. μέσω της κρίσιμης περιοχής K,

β. μέσω κατάλληλου δ.ε. για το μ, και

γ. μέσω του *p-value*.

Δίνεται ότι $t_{19}(0.025) \approx 2.09302$.

6. Έστω ότι ο αριθμός μηνιαίων πωλήσεων ενός προϊόντος από τους αντιπροσώπους μιας εταιρείας αυτοκινήτων ακολουθεί την N(40, 100). Το προσωπικό των αντιπροσώπων της εταιρείας παρακολουθεί κάποια σεμινάρια και έστω $X_1, X_2, \ldots, X_\nu$ οι πωλήσεις των αντιπροσώπων τον επόμενο μήνα. Πως θα ελέγχατε σε ε.σ. a αν μεταβλήθηκαν οι μέσες πωλήσεις ή όχι. Ποια θα ήταν η απάντησή σας στο παραπάνω ερώτημα αν για ν = 5 αντιπροσώπους είχαμε x_i = 35, 45, 38, 40, 43. (a = 5%). Να δώσετε το αντίστοιχο *p-value*.

7. Κατά το έτος 2008 η μέση μηνιαία καταναλωτική δαπάνη των οικογενειών μιας πόλης ήταν μ$_0$ =14.4 εκατοντάδες ευρώ. Κατά το επόμενο έτος συγκεντρώθηκαν στοιχεία μέσης μηνιαίας καταναλωτικής δαπάνης (X, σε εκατοντάδες ευρώ) από τυχαίο δείγμα 10 οικογενειών της πόλης αυτής, τα οποία έδωσαν τα ακόλουθα αθροίσματα:

$$\sum_{i=1}^{\nu} x_i = 153.2 \quad \text{και} \quad \sum_{i=1}^{\nu} x_i^2 = 2448.6 .$$

Θεωρώντας ότι μηνιαία καταναλωτική δαπάνη των οικογενειών μια πόλης ακολουθεί N(μ,σ^2), μπορούμε να δεχθούμε, σε επίπεδο σημαντικότητας 0.05, ότι κατά το έτος 2009 η μέση μηνιαία καταναλωτική δαπάνη μ όλων των οικογενειών της πόλης έμεινε, στην πραγματικότητα, αμετάβλητη σε σχέση με το προηγούμενο έτος; Να θεωρήσετε ως εναλλακτική υπόθεση

α. την μ ≠ μ$_0$, β. την μ < μ$_0$, γ. την μ > μ$_0$.

8. Έστω τυχαίο δείγμα $X_1, X_2, \ldots, X_\nu$ από την N(μ,σ^2) με σ^2 άγνωστο και έστω ότι θέλουμε να ελέγξουμε την H$_0$: μ = μ$_0$. Αν συμβολίσουμε με $pvalue_{>}$, $pvalue_{<}$, και $pvalue_{\neq}$, τα *p-values* που αντιστοιχούν στον έλεγχο της H$_0$: μ = μ$_0$ με εναλλακτικές τις H$_1$: $\mu > \mu_0$, H$_1$: $\mu < \mu_0$, H$_1$: μ ≠ μ$_0$ αντίστοιχα, να δείξετε ότι

$$pvalue_{>} = \begin{cases} pvalue_{\neq}/2, & t \geq 0 \\ 1 - pvalue_{\neq}/2, & t < 0 \end{cases} \quad \text{και} \quad pvalue_{<} = \begin{cases} pvalue_{\neq}/2, & t < 0 \\ 1 - pvalue_{\neq}/2, & t \geq 0 \end{cases}$$

όπου $t = \sqrt{\nu}(\bar{x} - \mu_0)/s$.

9. Θεωρώντας ότι η τιμή πώλησης ενός προϊόντος σε μια περιοχή ακολουθεί $N(\mu,\sigma^2)$, και προκειμένου να ελεγχθεί η H_0: $\mu = 1000$ ευρώ έναντι της H_1 : $\mu \neq 1000$ ευρώ, καταγράφηκε η τιμή πώλησης του προϊόντος αυτού σε ένα τυχαίο δείγμα ν καταστημάτων. Από αυτό το τυχαίο δείγμα βρέθηκε μέση τιμή πώλησης 1030 ευρώ με *p-value* = 8%. Αν αντί της εναλλακτικής H_1:$\mu \neq 1000$ θέλαμε να ελέγξουμε την H_1 : $\mu > 1000$ ποιο θα ήταν το αντίστοιχο *p-value*; Ποιο θα ήταν το αντίστοιχο *p-value* όταν H_1 : $\mu < 1000$;

10. Σε ν = 20 καπνιστές που αποφάσισαν να διακόψουν το κάπνισμα μετρήθηκε το σωματικό τους βάρος λίγο πριν και τρείς μήνες μετά τη διακοπή του καπνίσματος, και βρέθηκε ότι $\bar{x} = 1.8$ και $s_x = 0.8$, όπου x_i είναι η διαφορά του σωματικού βάρους (σε κιλά) του i ατόμου *μετά – πρίν* τη διακοπή του καπνίσματος. Να ελέγξετε αν, με βάση το δείγμα αυτό, μπορούμε να πούμε ότι η διακοπή του καπνίσματος συνδέεται με μια μεταβολή του σωματικού βάρους. Να κάνετε τον αμφίπλευρο και τους δύο μονόπλευρους ελέγχους μέσω των *p-value*s.

11. Έστω τυχαίο δείγμα $X_1, X_2, ..., X_\nu \sim N(\mu,\sigma^2)$ με μ γνωστό.

α. Χρησιμοποιώντας το αποτέλεσμα που βρέθηκε για τον έλεγχο της υπόθεσης H_0: $\sigma = \sigma_0$ έναντι της H_1: $\sigma \neq \sigma_0$ (σε ε.σ. α) με το κριτήριο του γενικευμένου λόγου πιθανοφανειών να προσδιορίσετε το σύνολο των τιμών του σ_0 για τις οποίες δεν απορρίπτεται η H_0 και να το συγκρίνετε με το γνωστό διάστημα εμπιστοσύνης για το σ, συντελεστού $1 - a$ (όταν μ γνωστό).

12. Να επαναλάβετε την Άσκηση 11, αυτή την φορά θεωρώντας ότι το μ είναι άγνωστο.

13. 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 της H_1: $\sigma^2 > 100$ 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.

14. Η δοκιμή ενός συστήματος παρακολουθήσεως της παραγωγής είχε αποτυχία 4 φορές τον πρώτο μήνα, 3 φορές το δεύτερο μήνα, 7 φορές τον τρίτο μήνα και 8 φορές τον τέταρτο μήνα. Μπορούμε να δεχθούμε, σε επίπεδο σημαντικότητας 5%, ότι, κατά μέσο όρο, οι αποτυχίες του συστήματος είναι, γενικά, λιγότερες από 10 το μήνα; Θεωρήστε ότι η κατανομή των μηνιαίων αποτυχιών του συστήματος μπορεί να προσεγγισθεί από την κανονική κατανομή.

15. Έστω $X_1, X_2, \ldots, X_\nu$ ένα τυχαίο δείγμα από την Εκθετική κατανομή με μέση τιμή θ. Να βρεθεί έλεγχος για την H_0: $\theta = \theta_0$ έναντι της H_1: $\theta \neq \theta_0$, σε ε.σ. a, μέσω του γενικευμένου λόγου πιθανοφανειών.

4. ΕΛΕΓΧΟΙ ΥΠΟΘΕΣΕΩΝ ΓΙΑ ΤΙΣ ΠΑΡΑΜΕΤΡΟΥΣ ΔΥΟ ΚΑΝΟΝΙΚΩΝ ΠΛΗΘΥΣΜΩΝ

4.1. Έλεγχος υποθέσεων για τη διαφορά $\mu_1 - \mu_2$ των μέσων δυο ανεξάρτητων πληθυσμών

Σε αυτή την παράγραφο, θέλουμε να ελέγξουμε αν ο μέσος μ_1 ενός κανονικού πληθυσμού είναι ίσος με τον μέσο μ_2 ενός άλλου (ανεξάρτητου) κανονικού πληθυσμού. Έχουμε δηλαδή την υπόθεση

$$H_0: \mu_1 - \mu_2 = 0 \quad \textit{έναντι της} \quad H_1: \mu_1 - \mu_2 \neq 0.$$

Θεωρούμε ότι έχουμε δυο *ανεξάρτητα* τυχαία δείγματα

$$X_1, X_2, \ldots, X_{\nu_1} \sim \mathrm{N}(\mu_1, \sigma_1^2), \quad Y_1, Y_2, \ldots, Y_{\nu_2} \sim \mathrm{N}(\mu_2, \sigma_2^2)$$

από τους δύο αυτούς πληθυσμούς και διακρίνουμε τις ακόλουθες περιπτώσεις:

(α). *Τα* σ_1^2, σ_2^2 ***είναι γνωστά.*** Σε αυτή την περίπτωση, έχουμε

$$\mathrm{H}_0 : \boldsymbol{\theta} = (\mu_1, \mu_2) \in \Theta_0 = \{(\mu, \mu), \mu \in \mathrm{R}\} \ \textit{έναντι της} \ \mathrm{H}_1 : \boldsymbol{\theta} = (\mu_1, \mu_2) \in \mathrm{R}^2 - \Theta_0$$

και η περιοχή απόρριψης της H_0,σύμφωνα με το ΓΛΠ, θα καθορίζεται μέσω της ανισότητας

$$\lambda(\mathbf{x}, \mathbf{y}) = \frac{\sup\limits_{\boldsymbol{\theta} \in \Theta_0} L(\boldsymbol{\theta})}{\sup\limits_{\boldsymbol{\theta} \in \Theta_0 \cup \Theta_1} L(\boldsymbol{\theta})} = \frac{\sup_\mu L(\mu, \mu)}{\sup\limits_{\mu_1, \mu_2} L(\mu_1, \mu_2)} < c,$$

όπου

$$L(\mu_1, \mu_2) = \prod_{i=1}^{\nu_1} f_X(x_i, \mu_1) \prod_{i=1}^{\nu_2} f_Y(y_i, \mu_2) = (2\pi\sigma_1^2)^{-\frac{\nu_1}{2}} e^{-\frac{1}{2\sigma_1^2}\sum_{i=1}^{\nu_1}(x_i - \mu_1)^2} (2\pi\sigma_2^2)^{-\frac{\nu_2}{2}} e^{-\frac{1}{2\sigma_2^2}\sum_{i=1}^{\nu_1}(y_i - \mu_2)^2}$$

είναι η πιθανοφάνεια των δύο ανεξάρτητων δειγμάτων. Βρίσκουμε ότι

$$\frac{\partial}{\partial \mu} \ln L(\mu, \mu) = 0 \Leftrightarrow \frac{1}{\sigma_1^2} \sum_{i=1}^{\nu_1} (x_i - \mu) + \frac{1}{\sigma_2^2} \sum_{i=1}^{\nu_1} (y_i - \mu) = 0 \Rightarrow \hat{\mu}_0 = \frac{\frac{\nu_1}{\sigma_1^2}\bar{x} + \frac{\nu_2}{\sigma_2^2}\bar{y}}{\frac{\nu_1}{\sigma_1^2} + \frac{\nu_2}{\sigma_2^2}},$$

και επίσης βρίσκουμε ότι sup $L(\mu_1, \mu_2) = L(\bar{x}, \bar{y})$. Έπειτα από κάποιες πράξεις και απλοποιήσεις προκύπτει

$$\lambda(\mathbf{x}, \mathbf{y}) < \mathrm{c} \Leftrightarrow \frac{\sup_\mu L(\mu, \mu)}{\sup\limits_{\mu_1, \mu_2} L(\mu_1, \mu_2)} < c \Leftrightarrow \frac{L(\hat{\mu}_0, \hat{\mu}_0)}{L(\bar{X}, \bar{Y})} < c \Leftrightarrow \ldots \Leftrightarrow |\bar{x} - \bar{y}| > c'$$

οπότε η κρίσιμη περιοχή θα πρέπει να έχει την μορφή

$$|\bar{X} - \bar{Y}| > c'.$$

Για να ισχύει $P[I] = a$, θα πρέπει $P(|\bar{X} - \bar{Y}| > c' \mid H_0) = a$. Παρατηρούμε ότι η σ.σ.

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{\dfrac{\sigma_1^2}{\nu_1} + \dfrac{\sigma_2^2}{\nu_2}}}$$

ακολουθεί, υπό την H_0, την N(0,1), και επομένως η περιοχή απόρριψης της H_0 σε ε.σ. a θα είναι

$$K: \ |T| > z_{\alpha/2}$$

Όμοια βρίσκουμε τις κρίσιμες περιοχές για τον έλεγχο των μονόπλευρων υποθέσεων. Συγκεντρωτικά λαμβάνουμε τον επόμενο πίνακα:

Υποθέσεις	**Κρίσιμη περιοχή**
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 : $\mu_1 \neq \mu_2$	$\lvert T \rvert > z_{\alpha/2}$

Αν γενικότερα, H_0: $\mu_1 = \mu_2 + \delta$, τότε αντί της παραπάνω σ.σ. χρησιμοποιούμε την

$$T = \frac{\bar{X} - \bar{Y} - \delta}{\sqrt{\frac{\sigma_1^2}{\nu_1} + \frac{\sigma_2^2}{\nu_2}}},$$

η οποία υπό την H_0 ακολουθεί N(0,1) (οι περιοχές απόρριψης μένουν ίδιες).

(β). ***Τα*** σ_1^2, σ_2^2 ***είναι άγνωστα και ίσα (ομοσκεδαστικότητα).*** Εδώ θα έχουμε

$$H_0 : \boldsymbol{\theta} = (\mu_1,\mu_2,\sigma^2) \in \Theta_0 = \{(\mu,\mu),\ \mu \in R\} \times R_+ \ \textit{έναντι της}\ H_1 : \boldsymbol{\theta} = (\mu_1,\mu_2,\sigma^2) \in R^2 \times (R_+ - \Theta_0)$$

και επομένως η περιοχή απόρριψης της H_0 σύμφωνα με το ΓΛΠ, θα καθορίζεται μέσω της ανισότητας

$$\lambda(\mathbf{x},\mathbf{y}) = \frac{\sup\limits_{\boldsymbol{\theta} \in \Theta_0} L(\boldsymbol{\theta})}{\sup\limits_{\boldsymbol{\theta} \in \Theta_0 \cup \Theta_1} L(\boldsymbol{\theta})} = \frac{\sup_{\mu,\sigma^2} L(\mu,\mu,\sigma^2)}{\sup\limits_{\mu_1,\mu_2,\sigma^2} L(\mu_1,\mu_2,\sigma^2)} < c.$$

Έπειτα από αρκετές πράξεις και απλοποιήσεις, προκύπτει τελικά ότι η περιοχή απόρριψης της H_0 σε ε.σ. a είναι η

$$K: \ |T| > t_{\nu_1+\nu_2-2}(a/2)$$

όπου

$$T = T(\mathbf{X},\mathbf{Y}) = \frac{\bar{X} - \bar{Y}}{\sqrt{\dfrac{1}{\nu_1} + \dfrac{1}{\nu_2}} \sqrt{\dfrac{(\nu_1 - 1)S_1^2 + (\nu_2 - 1)S_2^2}{\nu_1 + \nu_2 - 2}}}$$

Σημειώνουμε ότι η T, υπό την H_0: $\mu_1 = \mu_2$ (με $\sigma_1 = \sigma_2$), ακολουθεί την κατανομή του student με $\nu_1+\nu_2-2$ βαθμούς ελευθερίας ($t_{\nu_1+\nu_2-2}$). Επίσης είναι φανερό ότι, όταν ισχύει η H_1, η T θα λαμβάνει «μεγάλες» (αρνητικές ή θετικές) τιμές.

Όμοια βρίσκουμε τις κρίσιμες περιοχές για τον έλεγχο των μονόπλευρων υποθέσεων. Συγκεντρωτικά λαμβάνουμε τον επόμενο πίνακα:

Υποθέσεις	**Κρίσιμη περιοχή**
$H_0 : \mu_1 \leq(=) \mu_2$ *έναντι της* $H_1 : \mu_1 > \mu_2$	$T > t_{\nu_1+\nu_2-2}(a)$
$H_0 : \mu_1 \geq(=) \mu_2$ *έναντι της* $H_1 : \mu_1 < \mu_2$	$T < -\, t_{\nu_1+\nu_2-2}(a)$
$H_0 : \mu_1 = \mu_2$ *έναντι της* $H_1 : \mu_1 \neq \mu_2$	$\lvert T \rvert > t_{\nu_1+\nu_2-2}(a/2)$

Στην περίπτωση που θέλουμε να ελέγξουμε την υπόθεση $H_0 : \mu_1 - \mu_2 = \gamma$, τότε χρησιμοποιούμε την ίδια στατιστική συνάρτηση T, με $\overline{X} - \overline{Y} - \gamma$ στον αριθμητή δηλαδή την

$$T = \frac{(\overline{X} - \overline{Y}) - \gamma}{\sqrt{\frac{1}{\nu_1} + \frac{1}{\nu_2}}\sqrt{\frac{(\nu_1 - 1)S_1^2 + (\nu_2 - 1)S_2^2}{\nu_1 + \nu_2 - 2}}}$$

(οι περιοχές απόρριψης μένουν ίδιες).

(γ) *Τα σ_1^2, σ_2^2 είναι άγνωστα και δεν είναι (ή δεν γνωρίζουμε αν είναι) ίσα.* Χωρίς την υπόθεση της ομοσκεδαστικότητας, χρησιμοποιούμε τη στατιστική συνάρτηση

$$T'(\mathbf{X}, \mathbf{Y}) = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{S_1^2}{\nu_1} + \frac{S_2^2}{\nu_2}}},$$

η οποία υπό την H_0: $\mu_1 = \mu_2$ αποδεικνύεται ότι *προσεγγιστικά* ακολουθεί κατανομή t_n ενώ όταν ισχύει η H_1 λαμβάνει «μεγάλες» (θετικές ή αρνητικές) τιμές. Οι βαθμοί ελευθερίας n εξαρτώνται από τα σ_1, σ_2 και επομένως είναι και αυτοί άγνωστοι. Αποδεικνύεται όμως ότι μπορούν να εκτιμηθούν από την στατιστική συνάρτηση

$$\hat{n} = \frac{(\frac{S_1^2}{\nu_1} + \frac{S_2^2}{\nu_2})^2}{\frac{S_1^4}{\nu_1^2(\nu_1 - 1)} + \frac{S_2^4}{\nu_2^2(\nu_2 - 1)}}.$$

Και εδώ, μπορούμε να χρησιμοποιήσουμε τη στατιστική συνάρτηση $|T'(\mathbf{X},\mathbf{Y})|$ η οποία, υπό την H_1, λαμβάνει «μεγάλες» θετικές τιμές. Απορρίπτουμε την H_0 (σε ε.σ. *προσεγγιστικά* a) όταν

$$|T'(\mathbf{x},\mathbf{y})| > c = t_{\hat{n}}(a/2)$$

Εάν θέλουμε να κάνουμε μονόπλευρο έλεγχο με εναλλακτική H_1: $\mu_1 > \mu_2$ ή H_1: $\mu_1 < \mu_2$ εργαζόμαστε όπως και στις προηγούμενες περιπτώσεις.

4.2. Έλεγχος υποθέσεων για το λόγο των διασπορών σ_1^2/σ_2^2 δυο ανεξάρτητων πληθυσμών

Θέλουμε να ελέγξουμε αν η διασπορά σ_1^2 ενός κανονικού πληθυσμού είναι ίση με τη διασπορά σ_2^2 ενός άλλου κανονικού πληθυσμού. Έχουμε δηλαδή την υπόθεση

$$H_0: \sigma_1^2/\sigma_2^2 = 1 \quad \textit{έναντι της} \quad H_1: \sigma_1^2/\sigma_2^2 \neq 0.$$

Με βάση δυο *ανεξάρτητα* τυχαία δείγματα

$$X_1, X_2, \ldots, X_{\nu_1} \sim \mathrm{N}(\mu_1, \sigma_1^2), \quad Y_1, Y_2, \ldots, Y_{\nu_2} \sim \mathrm{N}(\mu_2, \sigma_2^2)$$

από τους δύο αυτούς πληθυσμούς διακρίνουμε τις περιπτώσεις:

(α). Τα μ_1, μ_2 **είναι γνωστά**

Σε αυτή την περίπτωση έχουμε τις υποθέσεις

$$\mathrm{H}_0 : \boldsymbol{\theta} = (\sigma_1^2, \sigma_2^2) \in \Theta_0 = \{(\sigma^2, \sigma^2),\ \sigma^2 > 0\}$$

$$\mathrm{H}_1 : \boldsymbol{\theta} = (\sigma_1^2, \sigma_2^2) \in \mathrm{R}^2 - \Theta_0$$

και επομένως η περιοχή απόρριψης της H_0 σύμφωνα με το κριτήριο του γενικευμένου λόγου πιθανοφανειών θα καθορίζεται μέσω της ανισότητας

$$\lambda(\mathbf{x},\mathbf{y}) = \frac{\sup_{\boldsymbol{\theta}\in\Theta_0} L(\boldsymbol{\theta})}{\sup_{\boldsymbol{\theta}\in\Theta_0\cup\Theta_1} L(\boldsymbol{\theta})} = \frac{\sup_{\sigma^2} L(\sigma^2, \sigma^2)}{\sup_{\sigma_1^2,\sigma_2^2} L(\sigma_1^2, \sigma_2^2)} < c,$$

όπου

$$L(\sigma_1^2, \sigma_2^2) = (2\pi\sigma_1^2)^{-\frac{\nu_1}{2}} e^{-\frac{1}{2\sigma_1^2}\sum_{i=1}^{\nu_1}(x_i-\mu_1)^2} (2\pi\sigma_2^2)^{-\frac{\nu_2}{2}} e^{-\frac{1}{2\sigma_2^2}\sum_{i=1}^{\nu_1}(y_i-\mu_2)^2}.$$

Έπειτα από αρκετές πράξεις και απλοποιήσεις, προκύπτει τελικά ότι η περιοχή απόρριψης της H_0 σε ε.σ. a είναι η

$$T < c_1 \text{ ή } T > c_2$$

όπου

$$T = T(\mathbf{X},\mathbf{Y}) = \frac{\frac{1}{\nu_1}\sum_{i=1}^{\nu_1}(X_i - \mu_1)^2}{\frac{1}{\nu_2}\sum_{j=1}^{\nu_2}(Y_j - \mu_2)^2}.$$

Είναι φανερό ότι ο λόγος

$$\frac{\frac{1}{\nu_1}\sum_{i=1}^{\nu_1}\left(\frac{X_i - \mu_1}{\sigma_1}\right)^2}{\frac{1}{\nu_2}\sum_{j=1}^{\nu_2}\left(\frac{Y_j - \mu_2}{\sigma_2}\right)^2} = \frac{Y_1/\nu_1}{Y_2/\nu_2}$$

ακολουθεί την κατανομή του Snedecor (ή κατανομή F) με ν_1,ν_2 β.ε. αφού οι τυχαίες μεταβλητές Y_1, Y_2 είναι ανεξάρτητες και ακολουθούν κατανομές χι-τετράγωνο με ν_1 και ν_2 β.ε. αντίστοιχα (ως αθροίσματα τετραγώνων ανεξάρτητων τυποποιημένων κανονικών τυχαίων μεταβλητών). Δοθέντος ότι, όταν ισχύει η H_0: $\sigma_1^2/\sigma_2^2 = 1$, έχουμε

$$\frac{\frac{1}{\nu_1}\sum_{i=1}^{\nu_1}\left(\frac{X_i - \mu_1}{\sigma_1}\right)^2}{\frac{1}{\nu_2}\sum_{j=1}^{\nu_2}\left(\frac{Y_j - \mu_2}{\sigma_2}\right)^2} = \frac{\sigma_2^2}{\sigma_1^2}\,\frac{\frac{1}{\nu_1}\sum_{i=1}^{\nu_1}(X_i - \mu_1)^2}{\frac{1}{\nu_2}\sum_{j=1}^{\nu_2}(Y_j - \mu_2)^2} = \frac{\frac{1}{\nu_1}\sum_{i=1}^{\nu_1}(X_i - \mu_1)^2}{\frac{1}{\nu_2}\sum_{j=1}^{\nu_2}(Y_j - \mu_2)^2} = T(\mathbf{X},\mathbf{Y}) = T$$

συμπεραίνουμε ότι, υπό την H_0, η στατιστική συνάρτηση T ακολουθεί την κατανομή του Snedecor (F ή F-Ratio) με ν_1,ν_2 β.ε. και αν συμβολίσουμε με $F_{\nu_1,\nu_2}(a)$ το άνω a - σημείο της κατανομής αυτής, τότε, για να έχουμε $P[\text{I}] = a$ η περιοχή απόρριψης της H_0 πρέπει να έχει τη μορφή

$$K: \ T < F_{\nu_1,\nu_2}(1-a/2) \text{ ή } T > F_{\nu_1,\nu_2}(a/2).$$

Οι κρίσιμες περιοχές για τον έλεγχο των μονόπλευρων υποθέσεων προκύπτουν με τον ίδιο τρόπο. Συγκεντρωτικά λαμβάνουμε τον επόμενο πίνακα.

Υποθέσεις	**Κρίσιμη περιοχή**
$H_0:\ \sigma_1^2 = \sigma_2^2$ έναντι της $H_1:\ \sigma_1^2 \neq \sigma_2^2$	$T < F_{\nu_1,\nu_2}(1-a/2)$ ή $T > F_{\nu_1,\nu_2}(a/2)$
$H_0: \sigma_1^2 = (\leq)\,\sigma_2^2$ έναντι της $H_1:\ \sigma_1^2 > \sigma_2^2$	$T > F_{\nu_1,\nu_2}(a)$
$H_0: \sigma_1^2 = (\geq)\,\sigma_2^2$ έναντι της $H_1:\ \sigma_1^2 < \sigma_2^2$	$T < F_{\nu_1,\nu_2}(1-a)$

Τέλος, αν θέλουμε να ελέγξουμε την υπόθεση $H_0: \sigma_1^2/\sigma_2^2 = \delta^2$, μπορούμε να χρησιμοποιήσουμε τις ίδιες κρίσιμες περιοχές που αναφέρονται στον προηγούμενο πίνακα, με τη στατιστική συνάρτηση $T = T(\mathbf{x})$ όμως να δίνεται από τον τύπο

$$T = \frac{1}{\delta^2}\frac{\frac{1}{\nu_1}\sum_{i=1}^{\nu_1}(X_i - \mu_1)^2}{\frac{1}{\nu_2}\sum_{j=1}^{\nu_2}(Y_j - \mu_2)^2}.$$

(β). Τα μ_1, μ_2 είναι άγνωστα

Με τον ίδιο τρόπο όπως παραπάνω, από το κριτήριο του γενικευμένου λόγου πιθανοφανειών, προκύπτει ότι πρέπει να χρησιμοποιήσουμε την στατιστική συνάρτηση

$$T = T(\mathbf{X},\mathbf{Y}) = \frac{\frac{1}{\nu_1-1}\sum_{i=1}^{\nu_1}(X_i - \overline{X})^2}{\frac{1}{\nu_2-1}\sum_{j=1}^{\nu_2}(Y_j - \overline{Y})^2} = \frac{S_1^2}{S_2^2},$$

όπου με S_1^2, S_2^2 συμβολίσαμε τις δειγματικές διασπορές

$$S_1^2 = \frac{1}{\nu_1 - 1}\sum_{i=1}^{\nu_1}(X_i - \overline{X})^2,\; S_2^2 = \frac{1}{\nu_2 - 1}\sum_{j=1}^{\nu_2}(Y_j - \overline{Y})^2.$$

Ο λόγος

$$\frac{\frac{1}{\nu_1 - 1}\sum_{i=1}^{\nu_1}\left(\frac{X_i - \overline{X}}{\sigma_1}\right)^2}{\frac{1}{\nu_2 - 1}\sum_{j=1}^{\nu_2}\left(\frac{Y_j - \overline{Y}}{\sigma_2}\right)^2} = \frac{Y_1/(\nu_1 - 1)}{Y_2/(\nu_2 - 1)} = \frac{\sigma_2^2}{\sigma_1^2}\frac{S_1^2}{S_2^2} = \frac{\sigma_2^2}{\sigma_1^2}T(\mathbf{X},\mathbf{Y})$$

ακολουθεί την κατανομή του Snedecor (ή κατανομή F) με ν_1,ν_2 β.ε. αφού οι τυχαίες μεταβλητές Y_1, Y_2 είναι ανεξάρτητες και ακολουθούν κατανομές χι-τετράγωνο με $\nu_1 - 1$ και $\nu_2 - 1$ β.ε. αντίστοιχα. Δοθέντος ότι, όταν ισχύει η H_0: $\sigma_1^2/\sigma_2^2 = 1$, έχουμε

$$\frac{\frac{1}{\nu_1 - 1}\sum_{i=1}^{\nu_1}\left(\frac{X_i - \overline{X}}{\sigma_1}\right)^2}{\frac{1}{\nu_2 - 1}\sum_{j=1}^{\nu_2}\left(\frac{Y_j - \overline{Y}}{\sigma_2}\right)^2} = \frac{\sigma_2^2}{\sigma_1^2}T(\mathbf{X},\mathbf{Y}) = T(\mathbf{X},\mathbf{Y})$$

συμπεραίνουμε ότι, υπό την H_0, η στατιστική συνάρτηση T ακολουθεί την κατανομή του Snedecor με $\nu_1 - 1$, $\nu_2 - 1$ β.ε.

Μπορούμε τώρα εύκολα να διαπιστώσουμε ότι σε αυτή την περίπτωση αντιστοιχούν οι επόμενες κρίσιμες περιοχές (ε.σ. a) ανάλογα με τις υποθέσεις που θέλουμε να ελέγξουμε κάθε φορά.

Υποθέσεις	Κρίσιμη περιοχή
$H_0: \sigma_1^2 = \sigma_2^2$ έναντι της $H_1: \sigma_1^2 \neq \sigma_2^2$	$T < F_{\nu_1-1,\nu_2-1}(1-a/2)$ ή $T > F_{\nu_1-1,\nu_2-1}(a/2)$
$H_0: \sigma_1^2 =(\leq) \sigma_2^2$ έναντι της $H_1: \sigma_1^2 > \sigma_2^2$	$T > F_{\nu_1-1,\nu_2-1}(a)$
$H_0: \sigma_1^2 =(\geq) \sigma_2^2$ έναντι της $H_1: \sigma_1^2 < \sigma_2^2$	$T < F_{\nu_1-1,\nu_2-1}(1-a)$

Αν θέλουμε να ελέγξουμε την $H_0: \sigma_1^2/\sigma_2^2 = \delta^2$, μπορούμε να χρησιμοποιήσουμε τις ίδιες κρίσιμες περιοχές που αναφέρονται στον προηγούμενο πίνακα, με τη στατιστική συνάρτηση $T = T(\mathbf{x})$ όμως να δίνεται από τον τύπο

$$T = \frac{1}{\delta^2}\frac{S_1^2}{S_2^2}.$$

(οι περιοχές απόρριψης μένουν ίδιες).

4.3. Έλεγχος υποθέσεων για τη διαφορά των μέσων για ζευγαρωτές παρατηρήσεις

Θεωρούμε ν ανεξάρτητα ζεύγη τυχαων μεταβλητών (X_1,Y_1), (X_2,Y_2), …, (X_ν,Y_ν) και υποθέτουμε ότι κάθε ζεύγος (X_i,Y_i) ακολουθεί διδιάστατη κανονική κατανομή με

$$E(X_i) = \mu_i,\ E(Y_i) = \mu_i + \delta,\ \ V(X_i) = \sigma_1^2,\ V(Y_i) = \sigma_2^2,\ \rho(X_i,Y_i) = \rho \neq 0,\ i = 1, 2, \ldots, \nu.$$

Επιθυμούμε να ελέγξουμε αν η διαφορά $\delta = E(Y_i) - E(X_i)$ μεταξύ των μέσων των X_i και των Y_i είναι μηδενική (θεωρώντας άγνωστες τις παραμέτρους σ_1, σ_2, ρ). Η συγκεκριμένη περίπτωση είναι διαφορετική από αυτήν της προηγούμενης παραγράφου (ακόμη και αν $\mu_i = \mu$) διότι εδώ τα τυχαία δε-ίγματα $X_1,X_2,\ldots,X_\nu$ και $Y_1,Y_2,\ldots,Y_\nu$ δεν είναι ανεξάρτητα (διότι $\rho(X_i,Y_i) = \rho \neq 0$). Η περίπτωση αυτή εμφανίζεται π.χ. όταν μελετάται η ***επίδραση κάποιου παράγοντα*** (π.χ. θεραπείας) σε ν άτομα. Συνή-θως τα X_i, Y_i εκφράζουν την τιμή μιας μεταβλητής που αφορά το i-άτομο «**πριν**» (X_i) και «**μετά**» (Y_i) την επίδραση του παράγοντα. Ένα άλλο παράδειγμα όπου εφαρμόζεται είναι στις λεγόμενες case-control έρευνες στις οποίες εφαρμόζεται μια θεραπεία σε ν ασθενείς (cases). Για να ελεγχθεί η επίδ-ραση της θεραπείας επιλέγονται και άλλα ν άτομα (άτομα ελέγχου - controls) τα οποία δεν έχουν α-κολουθήσει την συγκεκριμένη θεραπεία ή λαμβάνουν placebo (π.χ. λαμβάνουν ένα ανενεργό σκεύ-ασμα ίδιο στη μορφή με αυτό που έλαβαν οι ν ασθενείς ώστε να εξαλειφθεί η επίδραση ψυχολογικών παραγόντων). Η επιλογή γίνεται έτσι ώστε ο i-ασθενής να έχει τα ίδια χαρακτηριστικά (π.χ. ηλικία, φύλο, περιβάλλον, κατάσταση υγείας κ.ο.κ.) με το i-άτομο ελέγχου, $i = 1,2,\ldots,\nu$. Σε αυτή την περίπτωση η X_i εκφράζει την τιμή της μεταβλητής που μας ενδιαφέρει στον i-ασθενή (μετά την θεραπεία) ενώ η Y_i εκφράζει την τιμή της ίδιας μεταβλητής στο i-άτομο ελέγχου.

Κάτω από τις παραπάνω υποθέσεις, οι τυχαίες μεταβλητές $Z_i = Y_i - X_i$, $i = 1, 2, \ldots, \nu$, είναι ανεξάρτητες και

$$Z_i = Y_i - X_i \sim N(\delta, \sigma^2)$$

με $\sigma^2 = \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2$. Επομένως, ισοδύναμα μπορούμε να ελέγξουμε αν η μέση τιμή δ των Z_i είναι ίση με $\mu_0 = 0$ (H_0: $\delta = 0$, H_1: $\delta \neq 0$). Αυτό γίνεται εύκολα χρησιμοποιώντας την μεθοδολογία που αναπτύχθηκε παραπάνω (έλεγχος υποθέσεων για το μέσο μ ενός πληθυσμού). Συγκεκριμένα, χρησιμοποιούμε την στατιστική συνάρτηση

$$T = \frac{\bar{Z}}{S_Z / \sqrt{\nu}} \sim_{H_0} t_{n-1}, \quad \bar{Z} = \frac{1}{\nu}\sum_{i=1}^{\nu} Z_i, \; S_Z^2 = \frac{1}{\nu - 1}\sum_{i=1}^{\nu}(Z_i - \bar{Z})^2 .$$

Απορρίπτουμε την H_0 σε ε.σ. a, όταν

$$|T| > c = t_{\nu-1}(a/2),$$

Εάν θέλαμε να κάνουμε μονόπλευρο έλεγχο με εναλλακτική H_1: $\mu_1 > \mu_2$ ή H_1: $\mu_1 < \mu_2$ εργαζόμαστε όπως και παραπάνω.

Κλείνοντας, αξίζει να σημειωθεί ότι, σε όλες τις παραπάνω περιπτώσεις που αφορούν υποθέσεις για τους μέσους (σε έναν ή δυο πληθυσμούς), οι έλεγχοι μπορεί να γίνουν (προσεγγιστικά) ακόμη και όταν οι πληθυσμοί δεν είναι κανονικοί, αλλά τα δείγματα είναι αρκετά μεγάλα (συνήθως > 30 ή καλύτερα >100). Αυτό ισχύει διότι, για μεγάλο δείγμα, οι δειγματικοί μέσοι ακολουθούν (προσεγγιστικά, από το Κεντρικό Οριακό Θεώρημα) κανονική κατανομή, ανεξάρτητα από την αρχική κατανομή του τιμών του δείγματος.

4.4. Ασκήσεις

1. Οι χρόνοι συναρμολόγησης ενός προϊόντος από δύο συγκεκριμένους εργάτες, ακολουθούν κανονική κατανομή με μέση τιμή μ_1 και μ_2 αντίστοιχα. Αν

234, 99, 234, 174, 188, 107, 173, 172 και 105, 194, 77, 33, 159, 150, 167, 127, 169, 166

είναι δειγματοληπτικά κάποιοι χρόνοι (σε min) συναρμολόγησης των δύο αυτών εργατών αντίστοιχα, μπορούμε σε επίπεδο σημαντικότητας α=0.05 να πούμε ότι α) οι δύο εργάτες έχουν διαφορετική απόδοση; β) ο πρώτος εργάτης έχει χειρότερη απόδοση από το δεύτερο; γ) ο πρώτος εργάτης συναρμολογεί το προϊόν 10 λεπτά αργότερα από το δεύτερο; Θεωρείστε ότι οι τυπικές αποκλίσεις σ_1, σ_2

των χρόνων συναρμολόγησης του προϊόντος από τους εργάτες είναι γνωστές και ίσες με σ_1=50, σ_2=50.

2. Από τυχαίο δείγμα 50 οικογενειών της πόλεως Α προέκυψε ότι το μέσο ετήσιο εισόδημά τους ήταν 15562 ευρώ. Επίσης από τυχαίο δείγμα 100 οικογενειών της πόλεως Β προέκυψε ότι το μέσο ετήσιο εισόδημά τους ήταν 15126 ευρώ. Μπορούμε να δεχθούμε, σε επίπεδο σημαντικότητας 0.01, ότι οι οικογένειες των δύο πόλεων από τις οποίες προέρχονται τα δείγματα έχουν το ίδιο μέσο εισόδημα, αν είναι γνωστό ότι στους δύο αυτούς πληθυσμούς η τυπική απόκλιση είναι ίση με 830;

3. Μια εταιρεία κατασκευής σκελετών σκαφών βρίσκει ότι τα δένδρα τύπου Α που προμηθεύεται από ένα δάσος αποδίδουν κατά μέσο όρο 64 kgr ξύλα περισσότερο από τα δένδρα τύπου Β. Η διακύμανση της απόδοσης ξύλου βρέθηκε ίση με 115 kgr^2. Ένα άλλο δάσος περιέχει και τους δύο τύπους ξύλων. Παίρνοντας δείγμα 100 δένδρων για κάθε τύπο (από το δεύτερο δάσος) οι αποδόσεις σε ξύλο βρέθηκαν 1390 kgr και 1332 kgr αντίστοιχα ενώ οι διακυμάνσεις έμεναν αμετάβλητες. Μπορούμε να δεχθούμε σε ε.σ. 1% ότι η διαφορά των μέσων αποδόσεων ξύλου παραμένει στο ίδιο επίπεδο δηλ. 64 kgr;

4. Έστω μ_1, μ_2 οι μέσες τιμές πώλησης ενός προϊόντος στις περιοχές Α, Β αντίστοιχα. Επιλέγουμε τυχαία 4 τιμές από την περιοχή Α και 4 τιμές από την περιοχή Β. Αν οι τιμές αυτές είναι

1.12, 1.06, 1.21, 1.1 (περιοχή Α) και 1.07, 0.93, 0.97, 0.99 (περιοχή Β),

μπορούμε, σε επίπεδο σημαντικότητας 5%, να πούμε ότι η μέση τιμή πώλησης στην περιοχή Α είναι υψηλότερη από την αντίστοιχη στην περιοχή Β; (H_0: $\mu_1 = \mu_2$, H_1: $\mu_1 > \mu_2$) (υποθέτουμε ότι οι τιμές κατανέμονται κανονικά και με ίση διασπορά στις δύο περιοχές).

5. Από τυχαίο δείγμα 12 εμπορικών καταστημάτων ειδών ενδυμασίας προέκυψε ότι η μέση ημερήσια δαπάνη τους για διαφήμιση είναι 51 ευρώ με δειγματική τυπική απόκλιση 3.6 ευρώ. Επίσης από τυχαίο δείγμα 12 εμπορικών καταστημάτων ειδών υποδήσεως προέκυψε ότι η μέση ημερήσια δαπάνη τους για διαφήμιση είναι 48 ευρώ με δειγματική τυπική απόκλιση 4 ευρώ. Μπορούμε να δεχθούμε, σε ε.σ. 1%, ότι οι δύο κατηγορίες καταστημάτων δαπανούν, κατά μέσο όρο το ίδιο ποσόν για διαφήμιση; Υποθέτουμε ότι οι δαπάνες για κάθε κατάστημα κατανέμονται κανονικά και με ίση διασπορά.

6. α. Έστω μ_1, μ_2 οι μέσες τιμές πώλησης ενός προϊόντος στις περιοχές Α, Β αντίστοιχα. Η μέση τιμή και η διασπορά ενός τυχαίου δείγματος 10 τιμών πώλησης από την περιοχή Α βρέθηκε 100.9 και 8.76667 αντίστοιχα. Επίσης, η μέση τιμή και η διασπορά ενός τυχαίου δείγματος 20 τιμών πώλησης από την περιοχή Β βρέθηκε 104.45 και 12.9974 αντίστοιχα. Αν υποθέσουμε ότι οι τιμές κατανέμονται κανονικά και με ίση (αλλά άγνωστη) διασπορά και στις δύο περιοχές, να ελέγξετε σε επίπεδο σημαντικότητας 5%

i. Αν η μέση τιμή πώλησης στην A είναι διαφορετική από την αντίστοιχη στην B.

ii. Αν η μέση τιμή πώλησης στην A είναι χαμηλότερη από την αντίστοιχη στην B.

β. Παραπάνω υποθέσαμε ότι οι διασπορές των τιμών στις περιοχές αυτές είναι ίσες. Να ελέγξετε σε επίπεδο σημαντικότητας 5% αν όντως οι δύο αυτές διασπορές μπορούν να θεωρηθούν ίσες.

7. Έστω ότι οι χρόνοι ζωής δύο τύπων εξαρτημάτων A, B ακολουθούν κανονική κατανομή $N(\mu_1,\sigma_1^2)$ και $N(\mu_2, \sigma_2^2)$ αντίστοιχα. Αν σε ένα τυχαίο δείγμα από εξαρτήματα τύπου A βρέθηκε ότι $S_1 = 100$ ($\nu_1 = 50$) και σε ένα τυχαίο δείγμα από εξαρτήματα τύπου B βρέθηκε ότι $S_2 = 120$ ($\nu_2 = 50$), μπορούμε σε ε.σ. 5% να ισχυριστούμε ότι οι χρόνοι ζωής των εξαρτημάτων A παρουσιάζουν μικρότερη μεταβλητότητα από τους χρόνους του B; Δίνεται ότι $F_{49,49}(0.05) = 1.605$

8. Από δύο δείγματα άγαμων και έγγαμων εργατριών συγκεντρώθηκαν τα ακόλουθα στοιχεία που αφορούν ώρες απουσίας από την εργασίας τους λόγω ασθένειας σε ένα έτος:

Άγαμοι: 88, 68, 77, 82, 63, 80, 78, 71, 68, Έγγαμοι : 73, 77, 67, 70, 74, 64, 71, 71, 72

Να ελεγχθεί σε ε.σ. 5% η υπόθεση ότι η διακύμανση των ωρών απουσίας των άγαμων

α. είναι ίση με τη διακύμανση των ωρών απουσίας των έγγαμων εργατριών, έναντι της εναλλακτικής υποθέσεως ότι η διασπορά των ωρών απουσίας των άγαμων είναι μεγαλύτερη από εκείνη των έγγαμων.

β. είναι τετραπλάσια της διακύμανσης των ωρών απουσίας των έγγαμων εργατριών.

Δίνεται ότι $F_{8,8}(0.05) = 3.44, F_{8,8}(0.025) = 4.43$.

9. Από δύο ανεξάρτητους πληθυσμούς παίρνουμε τα τυχαία δείγματα 7,6,4,8,5,3,5 και 6,5,3,7,8,3,6,3,4 αντίστοιχα. Να ελεγχθεί σε ε.σ. 5% η υπόθεση ότι οι πληθυσμοί έχουν ίσες διακυμάνσεις έναντι της εναλλακτικής ότι η διακύμανση του πρώτου πληθυσμού είναι μεγαλύτερη της διακύμανσης του δεύτερου. Δίνεται ότι $F_{8,6}(0.95) = 0.28$.

10. Για τον έλεγχο της αποτελεσματικότητας ενός σκευάσματος που καταπολεμά την παχυσαρκία, χορηγήθηκε συγκεκριμένη ποσότητά του σε 20 κατάλληλα πειραματόζωα. Σε καθένα από αυτά καταγράφηκε το βάρος του αμέσως πριν και μια εβδομάδα μετά την χορήγηση του σκευάσματος. Καταγράφηκαν τα παρακάτω σωματικά βάρη (σε kgr):

	1	2	3	4	5	6	7	8	9	10
Πριν	81.3	81.7	89.6	85.1	83.3	81.5	80.6	81.9	69.5	80.2
Μετά	81.8	80.1	84.2	82.1	75.8	81.6	76.1	79.9	70.9	81.2

	11	12	13	14	15	16	17	18	19	20
Πριν	77.3	80.8	71.5	84.7	75.4	81.9	79.8	82	94.3	77
Μετά	81	83.3	70	75.5	72.4	79.9	79	80.7	91.9	76.3

Υπάρχει θετική μέση επίδραση του συγκεκριμένου σκευάσματος (σε ε.σ. 5%) στο βάρος των πειραματόζωων;

11. (Παράδειγμα του Student) Η επίδραση δύο φαρμάκων Α και Β στο χρόνο ύπνου δοκιμάστηκε σε 10 ασθενείς που έπασχαν από αϋπνία. Και έστω x_i, y_i η παράταση του ύπνου του ασθενούς (σε ώρες) που οφείλεται στην επίδραση του φαρμάκου Α, Β αντίστοιχα.

x_i	1.9	0.8	1.1	0.1	–0.1	4.4	5.5	1.6	4.6	3.4
y_i	0.7	–1.6	–0.2	–1.2	–0.1	3.4	3.7	0.8	0.0	2.0
$z_i = x_i - y_i$	1.2	2.4	1.3	1.3	0.0	1.0	1.8	0.8	4.6	1.4

Είναι λογικό να δεχτούμε ότι

α. η επίδραση του φαρμάκου Α είναι διαφορετική από την επίδραση του φαρμάκου Β σε ε.σ 1%,

β. η παράταση του ύπνου που προκαλεί το φάρμακο Α είναι κατά 1 ώρα μεγαλύτερη της παράτασης που προκαλεί το φάρμακο Β;

5. ΠΡΟΣΕΓΓΙΣΤΙΚΟΙ ΕΛΕΓΧΟΙ ΥΠΟΘΕΣΕΩΝ ΓΙΑ ΠΟΣΟΣΤΑ

5.1. Έλεγχος υποθέσεων για ένα ποσοστό.

Έστω τυχαίο δείγμα $X_1, X_2, ..., X_\nu$ ~ Bernoulli(p), δηλαδή $P(X_i = 1) = p$, $P(X_i = 0) = 1 - p$, και συνάρτηση πιθανότητας

$$f(x;p) = P(X = x) = p^x(1-p)^{1-x},\ x = 0, 1.$$

Επιθυμούμε να κατασκευάσουμε έναν έλεγχο για την υπόθεση

$$H_0\colon p = p_0 \ \textit{έναντι της}\ H_1\colon p = p_1,\ \ p_1 > p_0.$$

Σύμφωνα με το λήμμα Neyman-Pearson, η κρίσιμη περιοχή θα καθορίζεται από την ανισότητα $\lambda(\mathbf{x}) < c$ όπου

$$\lambda(\mathbf{x}) = \frac{\prod_{i=1}^{\nu} f(x_i;p_0)}{\prod_{i=1}^{\nu} f(x_i;p_1)} = \frac{\prod_{i=1}^{\nu} p_0^{x_i}(1-p_0)^{1-x_i}}{\prod_{i=1}^{\nu} p_1^{x_i}(1-p_1)^{1-x_i}} = \left(\frac{p_0(1-p_1)}{p_1(1-p_0)}\right)^{\sum_{i=1}^{\nu} x_i}\left(\frac{1-p_0}{1-p_1}\right)^{\nu}.$$

Έχουμε

$$\lambda(\mathbf{x}) < c \Leftrightarrow \ln\lambda(\mathbf{x}) < \ln c = c' \Leftrightarrow \sum_{i=1}^{\nu} x_i \ln\frac{p_0(1-p_1)}{p_1(1-p_0)} < \ln c - \ln\left(\frac{1-p_0}{1-p_1}\right)^{\nu} = c''.$$

Επειδή $p_1 > p_0$ θα ισχύει ότι

$$\tfrac{p_0(1-p_1)}{p_1(1-p_0)} < 1 \ \Leftrightarrow\ \ln\tfrac{p_0(1-p_1)}{p_1(1-p_0)} < 0$$

οπότε

$$\sum_{i=1}^{\nu} x_i \ln\frac{p_0(1-p_1)}{p_1(1-p_0)} < c'' \Leftrightarrow \sum_{i=1}^{\nu} x_i > \frac{c''}{\ln\frac{p_0(1-p_1)}{p_1(1-p_0)}} \equiv c_a$$

και τελικά η κρίσιμη περιοχή θα έχει την μορφή

$$\sum_{i=1}^{\nu} X_i > c_a.$$

Το c_a λαμβάνεται τέτοιο ώστε να ισχύει $P[I] = a$. Συνεπώς,

$$P[I] = P(\sum_{i=1}^{\nu} X_i > c_a \mid H_0) = a \Leftrightarrow P(Y > c_a \mid Y \sim b(\nu, p_0)) = a \ \Leftrightarrow \sum_{i=c_a+1}^{\nu}\binom{\nu}{i}p_0^i(1-p_0)^{\nu-1} = a.$$

Η παραπάνω συνθήκη δεν λύνεται εύκολα ως προς c_a. Στην πράξη, για μεγάλα ν, χρησιμοποιείται προσέγγιση της διωνυμικής από την κανονική κατανομή. Πιο συγκεκριμένα, υπό την H_0 και για ν>30 (ή καλύτερα για ν>100), το Κ.Ο.Θ. εξασφαλίζει ότι η τυχαία μεταβλητή

$$T = \frac{\sum_{i=1}^{\nu} X_i - \nu p_0}{\sqrt{\nu p_0(1-p_0)}}$$

ακολουθεί (προσεγγιστικά) την $N(0,1)$. Επομένως

$$P[I] = a \Leftrightarrow P(\sum_{i=1}^{\nu} X_i > c_a \mid H_0) = P(\frac{\sum_{i=1}^{\nu} X_i - \nu p_0}{\sqrt{\nu p_0(1-p_0)}} > \frac{c_a - \nu p_0}{\sqrt{\nu p_0(1-p_0)}} \mid H_0) = a$$

από όπου προκύπτει

$$\frac{c_a - \nu\, p_0}{\sqrt{\nu\, p_0(1-p_0)}} = z_a \quad \Rightarrow \quad c_a = \nu\, p_0 + \sqrt{\nu\, p_0(1-p_0)}\,z_a\,.$$

Άρα, για $\nu > 30$, θα απορρίπτεται η H_0 (σε ε.σ. ***περίπου a***) όταν

$$K:\quad T = \frac{\overline{X} - p_0}{\sqrt{p_0(1-p_0)/\nu}} > z_a\,.$$

Η ίδια κρίσιμη περιοχή μπορεί να χρησιμοποιηθεί και για τον έλεγχο της υπόθεσης

H_0: p =(≤) p_0 έναντι της H_1: $p > p_0$,

οδηγώντας σε βέλτιστο έλεγχο (από Θεώρημα 2 κατασκευής ΟΙΕ στην ΕΟΚ). Συγκεκριμένα, εδώ, οι τυχαίες μεταβλητές X_i ακολουθούν την κατανομή Bernoulli=$b(1,p)$, οπότε

$$f(x) = p^x(1-p)^{1-x} = (1-p)e^{(\ln\frac{p}{1-p})x} = \beta(p)e^{\eta(p)T(x)}h(x)$$

με $\eta(p) = \ln\dfrac{p}{1-p}$. Αφού ισχύει

$$,\quad \eta'(p) = \frac{1}{p} + \frac{1}{1-p} > 0\,,$$

η συνάρτηση $\eta(p)$ θα είναι αύξουσα. Επομένως, για την H_0: $p \le p_0$ έναντι της H_1: $p > p_0$, η κρίσιμη περιοχή

$$K:\ \sum_{i=1}^{\nu} T(X_i) = \sum_{i=1}^{\nu} X_i > c_a$$

θα ορίζει ΟΙΕ.

Για να κατασκευάσουμε ελέγχους για τις υποθέσεις

H_0: $p \ge (=)\ p_0$, H_1: $p < p_0$

εργαζόμαστε όμοια με παραπάνω και καταλήγουμε στην κρίσιμη περιοχή ($\nu > 30$)

$$K:\ T = \frac{\overline{X} - p_0}{\sqrt{p_0(1-p_0)/\nu}} < -z_a$$

η οποία ορίζει ΟΙΕ σε επίπεδο σημαντικότητας (περίπου) a. Τέλος, για την περίπτωση αμφίπλευρου ελέγχου της μορφής H_0: $p = p_0$, H_1: $p \ne p_0$ χρησιμοποιούμε και πάλι την ίδια σ.σ.,

$$T = \frac{\overline{X} - p_0}{\sqrt{p_0(1-p_0)/\nu}}$$

η οποία ακολουθεί (προσεγγιστικά) την $N(0,1)$ και απορρίπτουμε την H_0 σε ε.σ. (περίπου) a όταν

$$|T| > z_{a/2}\,.$$

5.2. Έλεγχος υποθέσεων για τη σύγκριση δύο ποσοστών

Έστω δύο ανεξάρτητα τυχαία δείγματα $X_1, X_2, ..., X_{\nu_1}$ και $Y_1, Y_2, ..., Y_{\nu_2}$ από την κατανομή *Bernoulli* με παραμέτρους p_1 και p_2 αντίστοιχα. Επιθυμούμε να κατασκευάσουμε έναν έλεγχο για την υπόθεση

$$H_0\colon p_1 = p_2 \ \textit{έναντι της}\ H_1\colon p_1 \neq p_2.$$

Παρατηρούμε ότι προσεγγιστικά (για ν_1, ν_2 >30), ισχύει

$$T = \overline{X} - \overline{Y} \sim N(p_1 - p_2, \frac{p_1(1-p_1)}{\nu_1} + \frac{p_2(1-p_2)}{\nu_2})$$

οπότε, υπό την H_0: $p_1 = p_2 = p$, θα έχουμε

$$T = \overline{X} - \overline{Y} \sim N(0, p(1-p)(\frac{1}{\nu_1} + \frac{1}{\nu_2})).$$

Επομένως (όταν $p_1 = p_2 = p$), η τυχαία μεταβλητή,

$$\frac{\overline{X} - \overline{Y}}{\sqrt{p(1-p)(1/\nu_1 + 1/\nu_2)}},$$

ακολουθεί προσεγγιστικά N(0,1) κατανομή. Επειδή όμως το p που εμφανίζεται στον παραπάνω τύπο της T δεν είναι γνωστό, το αντικαθιστούμε από την εκτίμησή του

$$P = \frac{\nu_1 \overline{X} + \nu_2 \overline{Y}}{\nu_1 + \nu_2}$$

(πρόκειται για τον μέσο του δείγματος των $\nu_1+\nu_2$ παρατηρήσεων) και έτσι τελικά χρησιμοποιούμε τη στατιστική συνάρτηση

$$T = \frac{\overline{X} - \overline{Y}}{\sqrt{P(1-P)(1/\nu_1 + 1/\nu_2)}},$$

η οποία (επειδή $P \approx p$ για μεγάλο ν), ακολουθεί και αυτή προσεγγιστικά την κατανομή N(0,1). Για τον έλεγχο της υπόθεσης

$$H_0\colon p_1 = p_2 \ \textit{έναντι της}\ H_1\colon p_1 \neq p_2.$$

σε ε.σ. a (περίπου) μπορεί να χρησιμοποιηθεί η κρίσιμη περιοχή

$$K\colon |T| > z_{a/2}.$$

Στον επόμενο πίνακα δίνονται οι κρίσιμες περιοχές (σε ε.σ. περίπου a) τόσο για την αμφίπλευρη υπόθεση που εξετάσαμε όσο και για τις αντίστοιχες μονόπλευρες.

Υποθέσεις	Κρίσιμη περιοχή
$H_0 : p_1 = (\leq)\, p_2$ έναντι της $H_1 : p_1 > p_2$	K: $T > z_a$
$H_0 : p_1 = (\geq)\, p_2$ έναντι της $H_1 : p_1 < p_2$	K: $T < -z_a$
$H_0 : p_1 = p_2$ έναντι της $H_1 : p_1 \neq p_2$	K: $\lvert T\rvert > z_{a/2}$

5.3. Ασκήσεις

1. Ένας υποψήφιος δήμαρχος ισχυρίζεται ότι στις εκλογές που έγιναν σήμερα θα λάβει ποσοστό μεγαλύτερο του 50%. Έχουμε αρκετά στοιχεία για να απορρίψουμε τον παραπάνω ισχυρισμό, αν σε ένα τυχαίο δείγμα $\nu = 300$ ψηφοφόρων (που ελήφθη με exit poll, δηλαδή οι ψηφοφόροι ρωτήθηκαν αμέσως μετά την έξοδό τους από τυχαία επιλεγμένα εκλογικά τμήματα) μόνο το 44% δήλωσαν ότι ψήφισαν τον συγκεκριμένο υποψήφιο ($a = 0.01$). Αν βρεθεί ότι δεν έχουμε αρκετά στοιχεία, πόσο θα έπρεπε να είναι το ελάχιστο μέγεθος του δείγματος ώστε, με το ίδιο δειγματικό ποσοστό, να απορρίπταμε τον παραπάνω ισχυρισμό;

2. Σε τυχαίο δείγμα 400 τηλεθεατών οι 100 δήλωσαν ότι παρακολουθούν μια ορισμένη τηλεοπτική σειρά. Μπορούμε να δεχθούμε σε ε.σ 5% ότι εκείνοι που παρακολουθούν τη σειρά από το σύνολο των τηλεθεατών είναι λιγότεροι του 30%;

3. Έστω τυχαίο δείγμα $X_1, X_2, \ldots, X_\nu$ με $P(X_i = 1) = p$, $P(X_i = 0) = 1 - p$. Ποιο θα πρέπει να είναι το μέγεθος του δείγματος ν ώστε για τον έλεγχο της υπόθεσης H_0: $p = p_0$, H_1: $p = p_1$ ($p_1 > p_0$), οι πιθανότητες σφάλματος τύπου I και II να είναι αντίστοιχα α και β (χρησιμοποιήστε προσέγγιση μέσω κανονικής κατανομής).

4. Επιλέγοντας τυχαία 400 προϊόντα από μία μηχανή που τα κατασκευάζει, τα 16 βρέθηκαν ελαττωματικά, ενώ επιλέγοντας τυχαία 300 προϊόντα από μία άλλη μηχανή, τα 24 βρέθηκαν ελαττωματικά. Να ελέγξετε σε επίπεδο σημαντικότητας 1%

α αν υπάρχει διαφορά στην παραγωγή ελαττωματικών μεταξύ των δύο μηχανών και

β. αν η δεύτερη μηχανή είναι «χειρότερη» από την πρώτη.

5. Ένας δημοσιογράφος σε μία τηλεοπτική συζήτηση ισχυρίζεται ότι συμβαίνει κάτι ύποπτο στην εταιρία μέτρησης τηλεθέασης GBA. Η εταιρία έχει διαθέσει 1150 μηχανάκια σε ισάριθμους τηλεθεατές, 150 από τα οποία έχουν δοθεί σε νέους συνεργάτες (τηλεθεατές που έχουν ξεκινήσει τη συνεργασία τους με την GBA το τελευταίο εξάμηνο). Ο δημοσιογράφος διαπιστώνει ότι, σύμφωνα με την GBA, το κανάλι VEGA παρουσιάζει τηλεθέαση 30% στο σύνολο (στους 1000+150 τηλεθεατές) ενώ έχει μόνο 22% στους νέους (στους 150) τηλεθεατές. Ο δημοσιογράφος το θεωρεί αυτό πολύ ύποπτο και μάλιστα κάποιος πολιτικός που βρίσκεται στην συζήτηση δηλώνει ότι «οι δύο αυτές τιμές δεν συμφωνούν καθόλου» (πρόκειται για πραγματικό περιστατικό). Είναι πράγματι αυτό «πολύ ύποπτο»; Θα μπορούσε δηλαδή η διαφορά αυτή να είναι τυχαία; Συγκεκριμένα, αν p_1, p_2 είναι τα ποσοστά της τηλεθέασης του VEGA στους πληθυσμούς από τους οποίους προέρχονται οι δύο ομάδες τηλεθεατών («νέοι» και «παλαιοί» συνεργάτες), ελέγξτε σε ε.σ. $\alpha = 0.01$ την υπόθεση H_0: $p_1 = p_2$ έναντι της H_1: $p_1 > p_2$ (*αν απορρίψουμε την* H_0 *μπορούμε πράγματι να μιλήσουμε για κάτι «ύποπτο», διότι αναμένουμε να ισχύει* $p_1 = p_2$ *αφού και οι δύο ομάδες συνεργατών προέρχονται από τον ίδιο πληθυσμό*).

6. Τον μήνα Ιανουάριο επιλέχθηκαν τυχαία 400 ψηφοφόροι από μία περιοχή και ανάμεσά τους βρέθηκαν 100 οι οποίοι προτίθενται να ψηφίσουν το κόμμα Α. Τον μήνα Φεβρουάριο, επιλέχθηκαν τυχαία 500 ψηφοφόροι από την ίδια περιοχή και ανάμεσά τους βρέθηκαν τώρα 150 οι οποίοι προτίθενται να ψηφίσουν το κόμμα Α. Να ελέγξετε σε ε.σ. 5% αν το ποσοστό των ψηφοφόρων του κόμματος Α στην περιοχή αυτή αυξήθηκε κατά τον μήνα Φεβρουάριο. ($H_0 : p_1 = p_2$ έναντι της $H_1 : p_1 < p_2$)

7. Για να εκτιμηθεί η δύναμη ενός κόμματος στις αστικές και αγροτικές περιοχές λήφθηκε δείγμα 100 ατόμων από κάθε περιοχή. Υπέρ του κόμματος τάχθηκαν 55 άτομα της αστικής περιοχής και 45 της αγροτικής. Μπορούμε σε ε.σ 5% να δεχθούμε ότι η εκλογική δύναμη του κόμματος είναι η ίδια στις δύο περιοχές;