Μάθημα:

ΕΚΤΙΜΗΤΙΚΗ ΚΑΙ ΕΛΕΓΧΟΙ ΥΠΟΘΕΣΕΩΝ

Διδάσκων: Θ. ΔΗΜΗΤΡΑΚΟΣ

Τμήμα Στατιστικής, Ο. Π. Α.

ΕΛΕΓΧΟΙ ΥΠΟΘΕΣΕΩΝ

ΕΙΣΑΓΩΓΗ: (βλέπε βιβλίο, σελ. 359-373).

Οι έλεγχοι υποθέσεων αποτελούν ένα από τα βασικότερα πεδία της Στατιστικής. Ας ξεκινήσουμε με ένα παράδειγμα. Υποθέτουμε ότι υπάρχει ένα φάρμακο με το οποίο, αν γίνει θεραπεία, αναμένεται να εμφανιστούν τα πρώτα αποτελέσματα βελτίωσης, για τους ασθενείς, μέσα σε περίπου δέκα ημέρες. Έστω ότι ανακαλύπτεται ένα καινούργιο φάρμακο για την ίδια ασθένεια και η κατασκευαστική εταιρεία ισχυρίζεται ότι φέρνει αποτελέσματα σε συντομότερο χρονικό διάστημα. Υποθέτουμε ότι η τυχαία μεταβλητή X εκφράζει το χρόνο (σε ημέρες) μέσα στον οποίον, εμφανίζονται τα πρώτα αποτελέσματα βελτίωσης των ασθενών στους οποίους χορηγούνται τα φάρμακα. Έστω ότι $X \sim N(\mu, 4)$. Αν ο πληθυσμός των ασθενών στους οποίους χορηγούνται τα φάρμακα είναι μεγάλος, η μελέτη ολόκληρου του πληθυσμού είναι οικονομικά ασύμφορη ή και πρακτικά αδύνατη. Για το λόγο αυτό, μελετούμε ένα μέρος του πληθυσμού, λαμβάνοντας ένα τυχαίο δείγμα μεγέθους n, $X_1, \dots, X_n$, όπου με X_i, $i=1,\ldots,n$, συμβολίζουμε την τυχαία μεταβλητή (τ.μ.) που αναπαριστά το χρόνο (σε ημέρες) μέσα στον οποίον, εμφανίζονται τα πρώτα αποτελέσματα βελτίωσης του ασθενούς i, στον οποίον χορηγούνται τα φάρμακα.

Με βάση το τυχαίο δείγμα (τ.δ.) $X_1,\ldots,X_n$ μεγέθους n από την κανονική κατανομή $N(\mu,4)$, επιθυμούμε να ελέγξουμε τον ισχυρισμό της κατασκευαστικής εταιρείας. Εάν ο μέσος χρόνος μέσα στον οποίον, εμφανίζονται τα πρώτα αποτελέσματα βελτίωσης των ασθενών του πληθυσμού, από τη χρήση του πρώτου φαρμάκου, είναι $\mu=10$ ημέρες, επιθυμούμε να ελέγξουμε εάν πράγματι, σύμφωνα με τον ισχυρισμό της κατασκευαστικής εταιρείας, για το δεύτερο φάρμακο, ο αντίστοιχος μέσος χρόνος είναι μικρότερος, δηλαδή εάν $\mu<10$. Η υπόθεση $\mu=10$ συμβολίζεται με H_0 και καλείται <u>μηδενική υπόθεση</u> και η υπόθεση $\mu<10$ η οποία ελέγχεται ως προς την H_0 λέγεται <u>εναλλακτική υπόθεση</u> και συνήθως συμβολίζεται με H_1.

Για να λάβουμε μία απόφαση ποια από τις δύο υποθέσεις μπορεί να γίνει αποδεκτή και ποια να απορριφθεί, θεωρούμε ένα τυχαίο δείγμα $X_1,\ldots,X_n$ από την κανονική κατανομή $N(\mu,4)$ και βασιζόμαστε στις παρατηρούμενες τιμές $x_1,x_2,\ldots,x_n$ των τ.μ. $X_1,\ldots,X_n$ δηλαδή στα <u>δεδομένα</u>.

Γενικά, με τον όρο στατιστική υπόθεση θεωρούμε οποιαδήποτε υπόθεση που αναφέρεται στην συμπεριφορά τυχαίων μεταβλητών για τις οποίες μπορούμε, να

έχουμε

παρατηρήσεις (δεδομένα).

Η απόφασή μας να απορρίψουμε ή να μην απορρίψουμε τη μηδενική υπόθεση H_0 βασίζεται σε πληροφορίες από παρατηρηθείσες τιμές μίας τυχαίας μεταβλητής. Στη διαδικασία λήψης της απόφασης χρησιμοποιείται μία στατιστική συνάρτηση που καλείται στατιστική συνάρτηση ελέγχου ή ελεγχοσυνάρτηση (test statistic).

Η ελεγχοσυνάρτηση είναι μία τ.μ. που ακολουθεί μία συγκεκριμένη κατανομή ανάλογα με τη στατιστική υπόθεση που επιθυμούμε να ελέγξουμε. Λαμβάνει μία συγκεκριμένη τιμή ανάλογα με τις τιμές του δείγματος που χρησιμοποιούνται για τον έλεγχο. Η τιμή αυτή καθορίζει, για τα συγκεκριμένα δεδομένα, δηλαδή για το συγκεκριμένο δείγμα, την απόρριψη ή την αποδοχή της H_0.

Το σύνολο των τιμών που μπορεί να πάρει η ελεγχοσυνάρτηση για διαφορετικά δείγματα μπορεί να διαχωρισθεί σε δύο περιοχές:

α. Την περιοχή απόρριψης (rejection region) της H_0.

β. Την περιοχή αποδοχής (acceptance region) της H_0.

Εάν η τιμή της στατιστικής συνάρτησης, για το συγκεκριμένο δείγμα, δηλαδή για τα συγκεκριμένα δεδομένα που χρησιμοποιήσαμε για τον έλεγχο, βρίσκεται στην περιοχή απόρριψης, τότε η H_0 απορρίπτεται. Εάν η τιμή της στατιστικής συνάρτησης βρεθεί στην περιοχή αποδοχής, τότε η H_0 <u>δεν</u> απορρίπτεται.

Εκείνη η τιμή της παραμέτρου που διαχωρίζει την περιοχή αποδοχής από την περιοχή απόρριψης καλείται κρίσιμο σημείο (critical point) και συμβολίζεται με C.

Σε κάθε έλεγχο διατρέχουμε δύο ειδών κινδύνους:

1. Κίνδυνος ή σφάλμα (λάθος) τύπου I που συνίσταται στην απόρριψη της H_0 **ενώ** αυτή είναι σωστή (αληθινή).
2. Κίνδυνος ή σφάλμα (λάθος) τύπου II που συνίσταται στην απόρριψη της H_1 ενώ αυτή είναι σωστή (αληθινή).

Οι αποφάσεις που μπορούν να ληφθούν σε ένα στατιστικό έλεγχο υποθέσεων και τα λάθη που μπορούν να γίνουν εμφανίζονται παραστατικά **στον** ακόλουθο πίνακα:

	Επιλέγω H_0	Επιλέγω H_1
H_0 αληθινή	**Σωστή** απόφαση	**Λάθος** τύπου I
H_1 αληθινή	Λάθος τύπου II	Σωστή απόφαση

Η απόφασή μας, σε κάθε έλεγχο υποθέσεων στηρίζεται σε ένα μόνο δείγμα παρατηρήσεων. Επομένως, υπάρχει πιθανότητα να κάνουμε, με την απόφασή μας, λάθος τύπου I. Επιπλέον, υπάρχει κάποια άλλη πιθανότητα να κάνουμε, με την απόφασή μας, λάθος τύπου II.

Ορίζουμε τις πιθανότητες:

$$\alpha = P(\text{Λάθος τύπου I}) = P\left(\begin{array}{l}\text{απορρίπτω την } H_0 \text{ όταν}\\ \text{είναι αληθινή}\end{array}\right)$$

$$= P(\cancel{H_0} \mid H_0)$$

$$\beta = P(\text{Λάθος τύπου II}) = P\left(\begin{array}{l}\text{δέχομαι την } H_0 \text{ όταν}\\ \text{είναι αληθινή η } H_1\end{array}\right)$$

$$= P(H_0 \mid H_1)$$

Η τιμή της πιθανότητας α καλείται <u>επίπεδο</u> <u>στατιστικής</u> σημαντικότητας (level of significance).

Η συμπληρωματική πιθανότητα της β δηλαδή η πιθανότητα να απορριφθεί η H_0 όταν πράγματι η H_0 δεν είναι αληθινή καλείται ισχύς (power) του στατιστικού ελέγχου.

Είναι: $1-\beta = P(\cancel{H_0} / H_1)$.

Μία στατιστική υπόθεση καλείται <u>απλή (simple)</u> αν η H_0 αναφέρεται σε <u>μία</u> <u>μόνο</u> <u>συγκεκριμένη</u> <u>τιμή</u> της παραμέτρου. Για παράδειγμα, αν η H_0 αναφέρεται στη μέση τιμή ενός πληθυσμού, η απλή στατιστική υπόθεση θα είναι της μορφής $H_0: \mu = \mu_0$, όπου μ_0 είναι η συγκεκριμένη τιμή της παραμέτρου μ.

Η στατιστική υπόθεση θα καλείται <u>σύνθετη (composite)</u> αν αναφέρεται σε περισσότερες από μία ενδεχόμενες τιμές της υπό έλεγχο παραμέτρου. Μία σύνθετη υπόθεση για τη μέση τιμή ενός πληθυσμού θα είναι της μορφής:

$H_0: \mu \leq \mu_0$ ή $H_0: \mu \geq \mu_0$, όπου μ_0 είναι μία συγκεκριμένη τιμή της παραμέτρου μ.

Ο έλεγχος μίας στατιστικής υπόθεσης θα καλείται μονόπλευρος (one sided) αν η εναλλακτική υπόθεση H_1 είναι της μορφής $H_1: \mu < \mu_0$ (ή $H_1: \mu > \mu_0$) και αντίστοιχα, θα καλείται αμφίπλευρος αν η εναλλακτική υπόθεση H_1 είναι της μορφής $H_1: \mu \neq \mu_0$.

Για τη διεξαγωγή ενός στατιστικού ελέγχου υποθέσεων θα πρέπει να ορίσουμε κάποιον κανόνα βάσει του οποίου θα λάβουμε μία <u>απόφαση</u>, όταν οι παρατηρηθείσες τιμές $x_1, \dots, x_n$ (του δείγματος) των τ.μ. $X_1, \dots, X_n$ έχουν καθοριστεί.

Έστω ότι μας ενδιαφέρει να ελέγξουμε την

$H_0: \mu = 0$ έναντι της εναλλακτικής $H_1: \mu > 0$

για ένα πληθυσμό. Μία τ.μ. X αναπαριστά ένα υπό μελέτη χαρακτηριστικό του πληθυσμού. Υποθέτουμε ότι ο πληθυσμός κατανέμεται κανονικά, δηλ. θεωρούμε ότι: $X \sim N(\mu, 36)$.

Έστω ότι αποφασίζουμε να απορρίψουμε την H_0 αν η δειγματική μέση τιμή $\bar{X}$ είναι μεγαλύτερη από ένα κρίσιμο σημείο C. Για να υπολογίσουμε το <u>κρίσιμο σημείο</u> C θα πρέπει να ορίσουμε το <u>επίπεδο</u> <u>στατιστικής σημαντικότητας</u> α του ελέγχου.

Υποθέτουμε ότι $\alpha = 5\% = 0.05$

Όμως $\alpha = P(\text{Λάθος τύπου I})$

$$= P(\text{απόρριψης της } H_0 \mid H_0 \text{ ισχύει})$$

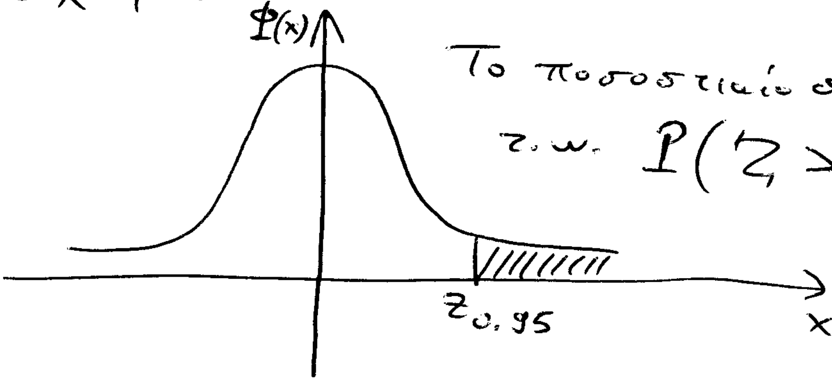
$$= P(\bar{X} > C \mid H_0) = 0.05$$

Στην περίπτωσή μας, $0.05 = P(\bar{X} > C \mid \mu = 0)$

$$\Rightarrow 0.05 = P\left(\frac{\bar{X} - 0}{\sqrt{\frac{36}{n}}} > \frac{C - 0}{\sqrt{\frac{36}{n}}}\right)$$

$$\Rightarrow 0.05 = P\left(Z > \frac{C}{\sqrt{\frac{36}{n}}}\right), \text{ όπου } Z = \frac{\bar{X}}{\sqrt{\frac{36}{n}}}$$

'Αρα, θα πρέπει: $z_{0.95} = 1.645 = \dfrac{C}{\sqrt{\frac{36}{n}}}$, όπως φαίνεται στο σχήμα.



Το ποσοστιαίο σημείο $z_{0.95}$ είναι τ.ω. $P(Z > z_{0.95}) = 0.05$

$\Phi(x)$: συνάρτηση πυκνότητας της $N(0,1)$.

Επομένως $\dfrac{C}{\sqrt{\frac{36}{n}}} = 1.645 \Rightarrow C = 1.645 \cdot \sqrt{\dfrac{36}{n}}$

Σελ. 374-383 (Βιβλίο)

Γενικά, έστω ότι έχουμε ένα κανονικό πληθυσμό $N(\mu, \sigma^2)$, όπου σ^2 γνωστό. Θέλουμε να ελέγξουμε την στατιστική υπόθεση: $H_0: \mu = \mu_0$ έναντι της $H_1: \mu > \mu_0$. Για να καθορίσουμε το κρίσιμο σημείο C, χρειάζεται να καθορίσουμε είτε το α είτε το β. Συνήθως, καθορίζουμε το α.

'Οπως προηγουμένως, αν C είναι το κρίσιμο σημείο και α το επίπεδο στατιστικής σημαντικότητας, διαδοχικά θα έχουμε:

$$\alpha = P(\cancel{H_0} \mid H_0) = P(\bar{X} > C \mid \mu = \mu_0)$$

$$= P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} > \frac{C - \mu}{\sigma/\sqrt{n}} \,\Big|\, \mu = \mu_0 \right)$$

$$= P\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} > \frac{C - \mu_0}{\sigma/\sqrt{n}} \right)$$

Θα πρέπει: $\dfrac{C - \mu_0}{\sigma/\sqrt{n}} = z_{1-\alpha}$

Ισοδύναμα,

$$C' = \mu_0 + z_{1-\alpha}\frac{\sigma}{\sqrt{n}}$$

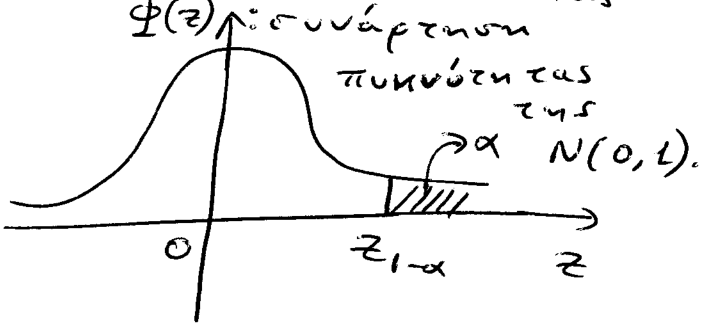
Συνεπώς, για κάθε στατιστικό έλεγχο της υπόθεσης $H_0: \mu = \mu_0$ έναντι της $H_1: \mu > \mu_0$, σε επίπεδο στατιστικής σημαντικότητας α, θα απορρίπτουμε την H_0, αν $\overline{X} > C'$.

Ισοδύναμα,

$$\overline{X} > \mu_0 + z_{1-\alpha}\frac{\sigma}{\sqrt{n}}$$

Με βάση την παραπάνω σχέση, είναι ισοδύναμο να πούμε ότι απορρίπτουμε την H_0 σε επίπεδο στατιστικής σημαντικότητας α αν

$$\frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} > z_{1-\alpha}$$

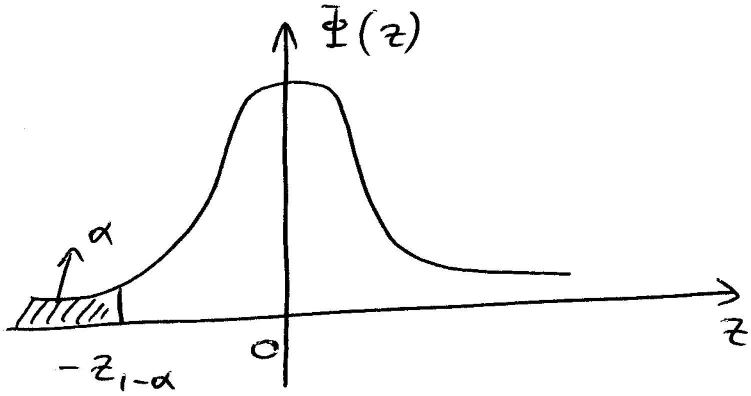


Θέτουμε $Z_0 = \dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$. Η Z_0 καλείται τυποποιημένη στατιστική συνάρτηση ελέγχου (Z_0 standard statistic).

Αν θέλουμε να ελέγξουμε τη στατιστική υπόθεση:

$H_0: \mu = \mu_0$ έναντι της $H_1: \mu < \mu_0$, θα απορρίπτουμε την H_0 σε επίπεδο στατιστικής σημαντικότητας α, αν

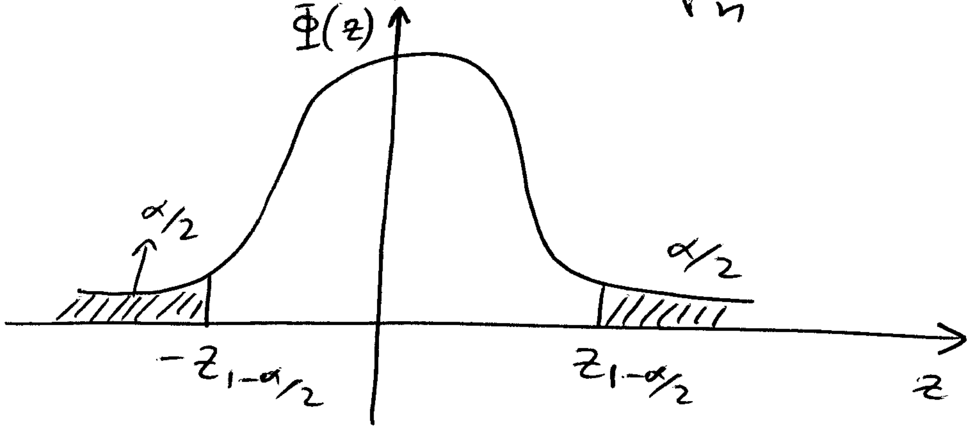
$$Z_0 = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} < -z_{1-\alpha} \text{ ή ισοδύναμα αν } \overline{X} < \mu_0 - z_{1-\alpha}\frac{\sigma}{\sqrt{n}}.$$



Αν η H_1 είναι αμφίπλευρη
δηλ. αν $H_0: \mu=\mu_0$
$H_1: \mu \neq \mu_0$

Θα απορρίπτουμε την H_0 σε επίπεδο στατιστικής σημαντικότητας α, αν $Z_0 < -z_{1-\alpha/2}$ ή $Z_0 > z_{1-\alpha/2}$

Ισοδύναμα, αν $\overline{X} < \mu_0 - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}$ ή $\overline{X} > \mu_0 + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}$



Σε γενικό πλαίσιο:
Τα στάδια ενός στατιστικού ελέγχου υποθέσεων είναι τα εξής:

1. Ορισμός της μηδενικής υπόθεσης H_0 (π.χ. $\mu=\mu_0$, $\mu_1=\mu_2$, $P=P_0$ κλπ.)
2. Ορισμός της εναλλακτικής υπόθεσης H_1 (μπορεί να είναι μονόπλευρη ή δίπλευρη).
3. Υπολογισμός της κατάλληλης στατιστικής συνάρτησης ελέγχου (π.χ. της Z_0).
4. Σύγκριση της τιμής της στατιστικής συνάρτησης ελέγχου για το συγκεκριμένο δείγμα με το αντίστοιχο ποσοστιαίο σημείο της κατανομής της στατιστικής συνάρτησης κάτω από την υπόθεση ότι η H_0 είναι αληθινή.
5. Συμπερασματολογία - Λήψη απόφασης.

Σε περίπτωση άγνωστης διασποράς σ^2 ενός κανονικού πληθυσμού, χρησιμοποιούμε την εκτιμήτρια της διασποράς

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n}$$

ή την αμερόληπτη εκτιμήτρια

$$S^{*2} = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}.$$

Η στατιστική συνάρτηση $T_0 = \dfrac{\overline{X} - \mu_0}{S/\sqrt{n-1}} = \dfrac{\overline{X} - \mu_0}{S^*/\sqrt{n}} \sim t_{n-1}$
(T_0-ελεγχοσυνάρτηση)

κάτω από την $H_0: \mu = \mu_0$.

Για παράδειγμα, αν $H_1: \mu \neq \mu_0$ είναι:

$$\alpha = P(\cancel{H_0} \mid H_0) = P\left(T_0 < \frac{c_1 - \mu_0}{S^*/\sqrt{n}} \text{ ή } T_0 > \frac{c_2 - \mu_0}{S^*/\sqrt{n}} \right) \quad (1)$$

όπου c_1, c_2 τα κρίσιμα σημεία. Η (1) προκύπτει με παρόμοιο τρόπο όπως και στην περίπτωση γνωστής διασποράς (βλέπε Βιβλίο, σελ. 395-398).

<u>Παράδειγμα 1</u> Θεωρούμε 25 τιμές που λαμβάνονται από έναν κανονικό πληθυσμό $N(\mu, 4)$. Θέλουμε να ελέγξουμε αν $\mu = 5$ έναντι της εναλλακτικής $\mu > 5$. Η δειγματική μέση τιμή είναι 5.45. Είναι $H_0: \mu = 5$ έναντι της $H_1: \mu > 5$.
Ισχύει ότι: $\overline{X} \sim N(5, 4/25)$, κάτω από την H_0.

Η Z_0 στατιστική συνάρτηση ελέγχου:

$$Z_0 = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{\overline{x} - 5}{2/\sqrt{25}} = \frac{5.45 - 5}{2/5} = 1.125.$$

Συγκρίνουμε την τιμή Z_0 με το ποσοστιαίο σημείο $Z_{1-\alpha}$ της $N(0,1)$.

Για $\alpha = 10\%$ έχουμε: $Z_{1-0.10} = Z_{0.90} = 1.28$

Για $\alpha = 5\%$ έχουμε: $Z_{1-0.05} = Z_{0.95} = 1.645$

Αφού $Z_0 = 1.125$ και στις δύο περιπτώσεις δεν έχουμε ενδείξεις κατά της H_0.

Παρατηρούμενο Επίπεδο Σημαντικότητας ή p-τιμή (observed level of significance ή p-value), σελ. 372-373.

Η p-τιμή μπορεί να χρησιμοποιηθεί αντί για το επίπεδο στατιστικής σημαντικότητας α.

Ορίζουμε ως p-τιμή την πιθανότητα η στατιστική συνάρτηση ελέγχου να πάρει μία τιμή τόσο ακραία ή περισσότερο ακραία από αυτήν που πήρε για το συγκεκριμένο δείγμα κάτω από την μηδενική υπόθεση.

Μικρές τιμές για την p-τιμή οδηγούν σε απόρριψη της H_0 ενώ μεγάλες τιμές για την p-τιμή οδηγούν σε μη-απόρριψη της H_0.

Η p-τιμή είναι το μικρότερο επίπεδο στατιστικής σημαντικότητας που μπορούμε να χρησιμοποιήσουμε για να απορρίψουμε την H_0. Για κάθε επίπεδο σημαντικότητας μικρότερο από την p-τιμή δεν έχουμε ενδείξεις για να απορρίψουμε την H_0. Για κάθε επίπεδο σημαντικότητας μεγαλύτερο από την p-τιμή έχουμε ενδείξεις για να απορρίψουμε την H_0.

Για οποιοδήποτε $\alpha \geq p$-τιμής, η τιμή της ελεγχοσυνάρτησης που λάβαμε για το συγκεκριμένο δείγμα, θα βρίσκεται στην περιοχή απόρριψης του στατιστικού ελέγχου.

Επειδή η p-τιμή εξαρτάται από τα δεδομένα του δείγματος που χρησιμοποιούνται για τον κάθε έλεγχο, καλείται "παρατηρούμενο" επίπεδο σημαντικότητας.

Για το Παράδειγμα 1, η p-τιμή, υπολογίζεται, ως εξής:

$$p\text{-τιμή} = P(\bar{X} > 5.45 \mid \mu = 5) =$$

$$= P\left(\frac{\bar{X}-5}{2/\sqrt{25}} > \frac{5.45-5}{2/\sqrt{25}}\right) = P(Z > 1.125)$$

$$= 0.13$$

όπου $Z = \dfrac{\bar{X}-5}{2/\sqrt{25}}$. Επειδή η p-τιμή είναι μεγαλύτερη από 5%, για $\alpha = 5\%$, δεν έχουμε ενδείξεις για να απορρίψουμε την H_0.

Αν ο έλεγχος ήταν αμφίπλευρος δηλαδή αν θέλαμε να ελέγξουμε την $H_0: \mu = 5$ έναντι της $H_1: \mu \neq 5$, τότε η τιμή της στατιστικής συνάρτησης ελέγχου θα ήταν πάλι $Z_0 = 1.125$ και

για $\alpha = 10\%$, $Z_{1-\frac{0.1}{2}} = Z_{0.95} = 1.645$

$$-Z_{0.95} = -1.645$$

για $\alpha = 5\%$, $Z_{1-\frac{0.05}{2}} = Z_{0.975} = 1.96$, $-Z_{0.975} = -1.96$

Η p-τιμή θα ήταν:

p-τιμή $= P(Z_0 > 1.125) + P(Z_0 < -1.125) = 2 \cdot 0.13$

$= 0.26$

Για οποιαδήποτε τιμή του α μεγαλύτερη του 0.26, η H_0 (με βάση το συγκεκριμένο δείγμα) θα πρέπει να απορριφθεί. Επομένως, για $\alpha = 5\%$ και $\alpha = 10\%$ δεν υπάρχουν ισχυρές ενδείξεις απόρριψης της H_0. (Παραδείγματα με p-τιμή, Βιβλίο, σελ. 388-405).

Παράδειγμα 2 Τυχαίο δείγμα από 26 τράπεζες έδωσε μέση απόδοση επενδύσεων $\bar{x} = 10.5$ με τυπική απόκλιση $S^* = 0.714$. Να ελεγχθεί αν η μέση απόδοση είναι μεγαλύτερη από 10.2%.

Ελέγχουμε την $H_0: \mu = 10.2$ έναντι της $H_1: \mu > 10.2$

Εκτιμώμενο τυπικό σφάλμα της δειγματικής μέσης τιμής $\bar{x}$ είναι: $\dfrac{S^*}{\sqrt{n}} = \dfrac{0.714}{\sqrt{26}} = 0.14$.

Η στατιστική συνάρτηση ελέγχου T_0 δίνει:

$$T_0 = \frac{10.5 - 10.2}{0.14} = 2.14$$

Για $\alpha = 5\%$ το ποσοστιαίο σημείο της t_{25}-κατανομής (από πίνακες) είναι $t_{25, 0.95} = 1.708$.

Για $\alpha = 1\%$, $t_{25, 0.99} = 2.485$

Συγκρίνοντας τα παραπάνω ποσοστιαία σημεία της t_{25}-κατανομής με την τιμή της T_0, συμπεραίνουμε ότι:

η H_0 απορρίπτεται, σε επίπεδο στατιστικής σημαντικότητας $\alpha = 5\%$ αλλά δεν απορρίπτεται σε επίπεδο στατιστικής σημαντικότητας $\alpha = 1\%$.

Η p-τιμή του ελέγχου είναι:

p-τιμή $= P(T_0 > 2.14) = 0.021$ (η τιμή υπολογίζεται από πίνακες της t_{25}-κατανομής). Παρατηρούμε ότι για $\alpha = 1\%$, επιβεβαιώνεται και μέσω της p-τιμής, η μη-απόρριψη της H_0.

Παράδειγμα 3 Μία εταιρεία γνωρίζει ότι τα φωτοτυπικά μηχανήματα είναι σε επισκευή κατά μέσο όρο για 12 ημέρες το χρόνο. Υπάρχει η άποψη ότι ο αριθμός των ημερών έχει αυξηθεί. Για να ελεγχθεί αυτός ο ισχυρισμός λάβαμε τυχαίο δείγμα από 9 μηχανήματα τα οποία έδωσαν τους παρακάτω χρόνους (σε ημέρες) όταν εστάλησαν για επισκευή τον προηγούμενο χρόνο: 16, 10, 21, 22, 8, 17, 19, 14, 19. Υπάρχουν ενδείξεις ότι υπάρχει αύξηση στους χρόνους επισκευής;

Ελέγχουμε $H_0: \mu = 12$ έναντι της $H_1: \mu > 12$

Είναι $\bar{X} = 16.22$, $S^* = 4.79$

Το εκτιμώμενο τυπικό σφάλμα της δειγματικής μέσης τιμής $\bar{X}$ είναι: $\dfrac{S^*}{\sqrt{n}} = \dfrac{4.79}{\sqrt{9}} = 1.597$

Η τιμή της στατιστικής συνάρτησης ελέγχου T_0 είναι:

$$T_0 = \frac{16.22 - 12}{1.597} = 2.64$$

Για $\alpha=5\%$, $t_{8,0.95}=1.860$

Για $\alpha=1\%$, $t_{8,0.99}=2.8965$

Επομένως, απορρίπτουμε την H_0 σε επίπεδο στατιστικής σημαντικότητας $\alpha=5\%$ αλλά <u>δεν</u> απορρίπτουμε την H_0 σε επίπεδο στατιστικής σημαντικότητας $\alpha=1\%$.

Η p-τιμή του ελέγχου είναι:

$$p\text{-τιμή} = P(T_0 > 2.64) = 0.015 \text{ (από πίνακες)}.$$

Για κάθε επίπεδο στατιστικής σημαντικότητας α μεγαλύτερο από 1.5%, υπάρχουν ενδείξεις απόρριψης της H_0, δηλαδή υπάρχουν ενδείξεις αύξησης του χρόνου επισκευής των φωτοτυπικών μηχανημάτων.

<u>Γενικά Σχόλια για τους ελέγχους υποθέσεων</u>

Ένας βέλτιστος έλεγχος υποθέσεων είναι εκείνος που ελαχιστοποιεί τα μεγέθη των σφαλμάτων τύπου I και II. Η ταυτόχρονη ελαχιστοποίηση των σφαλμάτων τύπου I και II είναι συνήθως αδύνατη διότι το σφάλμα τύπου I σχετίζεται με τιμές της παραμέτρου που αφορούν την H_0 ενώ το σφάλμα τύπου II σχετίζεται με τιμές της παραμέτρου που αφορούν την H_1. Όταν το μέγεθος του σφάλματος τύπου I ελαττώνεται δεν μπορούμε να πούμε με σιγουριά αν ελαττώνεται ή αυξάνεται το μέγεθος σφάλματος τύπου II. Το μόνο που μπορούμε να πούμε είναι ότι όταν μας δοθεί μία τιμή για το επίπεδο στατιστικής σημαντικότητας α μπορούμε να βρούμε μία ελεγχοσυνάρτηση που να μεγιστοποιεί την ισχύ.

Με άλλα λόγια, για δοσμένο επίπεδο στατιστικής σημαντικότητας α (συνήθη α είναι $\alpha = 0.005, 0.01, 0.05, 0.1$) βρίσκουμε εκείνη την ελεγχοσυνάρτηση που μεγιστοποιεί την ισχύ. Συνήθως, ως H_0 ορίζουμε εκείνη την υπόθεση, η λανθασμένη απόρριψη της οποίας, προκαλεί μεγαλύτερους κινδύνους από ότι η λανθασμένη αποδοχή της, διότι την πιθανότητα λανθασμένης απόρριψης την περιορίζουμε <u>εμείς</u> σ' όποιο επίπεδο θέλουμε. Ένας έλεγχος που για συγκεκριμένη τιμή του α ελαχιστοποιεί το β ή ισοδύναμα μεγιστοποιεί την ισχύ καλείται ισχυρότατος έλεγχος (βλέπε σελ. 367-371, βιβλίου). Για δεδομένο μέγεθος δείγματος n, όσο αυξάνεται το α, ελαττώνεται το β. Αν αντίστοιχα, ελαττώνουμε το α, το β θα αυξάνεται. Για συγκεκριμένη τιμή του α, το β ενδεχομένως ελαττώνεται (η ισχύς του ελέγχου αυξάνεται) με αύξηση του μεγέθους n του δείγματος. Βέβαια, αύξηση του μεγέθους n του δείγματος, δεν είναι πάντοτε εφικτή, για οικονομικούς λόγους.

(σελ. 406-408).

<u>Έλεγχοι υποθέσεων για αναλογίες, για μεγάλα μεγέθη δείγματος</u>

Ελέγχουμε την $H_0: P = P_0$ έναντι της $H_1: P > P_0$ ή της $H_1: P < P_0$ ή της $H_1: P \neq P_0$, όπου P είναι το ποσοστό των ατόμων (ή των στοιχείων) ενός πληθυσμού που έχουν κάποιο συγκεκριμένο χαρακτηριστικό και P_0 είναι μία συγκεκριμένη τιμή του ποσοστού. Έστω δείγμα μεγέθους n από αυτόν τον πληθυσμό και $\hat{P} = \dfrac{X}{n}$, η εκτιμήτρια του ποσοστού που λαμβάνεται μέσω του δείγματος και X είναι ο αριθμός των ατόμων στο δείγμα οι οποίοι έχουν το υπό μελέτη χαρακτηριστικό.

Η X είναι τ.μ. για την οποία (για μεγάλα n), ισχύει ότι: $X \sim N(nP, nP(1-P))$, κατά προσέγγιση. Η τ.μ. Z είναι:

$$Z = \frac{X - nP}{\sqrt{nP(1-P)}} = \frac{\frac{X}{n} - P}{\sqrt{\frac{P(1-P)}{n}}}$$

(τυποποιημένη μορφή της X)

Για μεγάλα n ισχύει ότι $Z \sim N(0,1)$, κατά προσέγγιση. Για $P = P_0$ δηλαδή κάτω από την H_0 ισχύει ότι:

$$Z_0 = \frac{\frac{X}{n} - P_0}{\sqrt{\frac{P_0(1-P_0)}{n}}} \underset{\text{προσεγγιστικά}}{\overset{n \text{ μεγάλο}}{\sim}} N(0,1),$$

όπου $\sqrt{\frac{P_0(1-P_0)}{n}}$ είναι το τυπικό σφάλμα της εκτιμήτριας $\hat{P} = \frac{X}{n}$.

Για την απόρριψη ή μη της H_0 συγκρίνουμε την τιμή της Z_0 με κατάλληλα ποσοστιαία σημεία της $N(0,1)$.

<u>Παράδειγμα 4</u> Η αναλογία των πελατών που ενδιαφέρονται για ένα νέο τραπεζικό σύστημα εκτιμάται ότι είναι $68/150 = 0.453$. Άλλες έρευνες ισχυρίζονται ότι αυτό το ποσοστό για ένα πληθυσμό είναι 40%. Ελέγξτε τον ισχυρισμό των ερευνών.

Ελέγχουμε την $H_0: P = 0.4$ έναντι της $H_1: P \neq 0.4$, όπου P είναι το ποσοστό των ατόμων του πληθυσμού που ενδιαφέρεται για το νέο τραπεζικό σύστημα.

Η στατιστική συνάρτηση ελέγχου δίνει:

$$Z_0 = \frac{0.453 - 0.4}{\sqrt{\frac{0.4 \cdot 0.6}{150}}} = 1.325$$

Η H_0 απορρίπτεται σε επίπεδο στατιστικής σημαντικότητας $\alpha = 0.05 = 5\%$ αν $Z_0 > Z_{1-\alpha/2}$ ή $Z_0 < -Z_{1-\alpha/2}$

Όμως $Z_{1-\frac{0.05}{2}} = Z_{1-0.025} = Z_{0.975} = 1.96$

Δεν έχουμε ενδείξεις για να απορρίψουμε την H_0 σε $\alpha = 5\%$.

Η p-τιμή είναι:

$$\text{p-τιμή} = P(Z_0 > 1.325) + P(Z_0 < -1.325)$$
$$= 0.0934 + 0.0934 = 0.1868 \text{ (από πίνακες).}$$

Η p-τιμή είναι το μικρότερο επίπεδο σημαντικότητας που μπορούμε να χρησιμοποιήσουμε για να απορρίψουμε την H_0. Για οποιοδήποτε $\alpha > 18.68\%$ υπάρχουν ενδείξεις απόρριψης της H_0.

Έλεγχοι υποθέσεων για τη σύγκριση μέσων τιμών δύο κανονικών πληθυσμών

Ανεξάρτητα δείγματα (σελ. 409-414).

Έστω $x_1, \ldots, x_n$ και $y_1, \ldots, y_m$ παρατηρήσεις από τυχαία δείγματα μεγέθους n και m αντίστοιχα, από κανονικούς πληθυσμούς $N(\mu_X, \sigma_X^2)$ και $N(\mu_Y, \sigma_Y^2)$.

Μας ενδιαφέρει να ελέγξουμε τις υποθέσεις:

$H_0: \mu_X = \mu_Y$ ή $H_0: \mu_X - \mu_Y = 0$ έναντι της

$H_1: \mu_X \neq \mu_Y$ ή $H_1: \mu_X > \mu_Y$ ή $H_1: \mu_X < \mu_Y$

Για ανεξάρτητα δείγματα, διακρίνουμε τις περιπτώσεις:

1η περίπτωση: Οι διακυμάνσεις των πληθυσμών είναι γνωστές, σ_X^2 και σ_Y^2, αντίστοιχα.

Χρησιμοποιούμε τη στατιστική συνάρτηση

$$Z_1 = \frac{\overline{X} - \overline{Y} - (\mu_X - \mu_Y)}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}}$$

(η $\overline{X} - \overline{Y}$ είναι η αμερόληπτη εκτιμήτρια των μ_X, μ_Y).

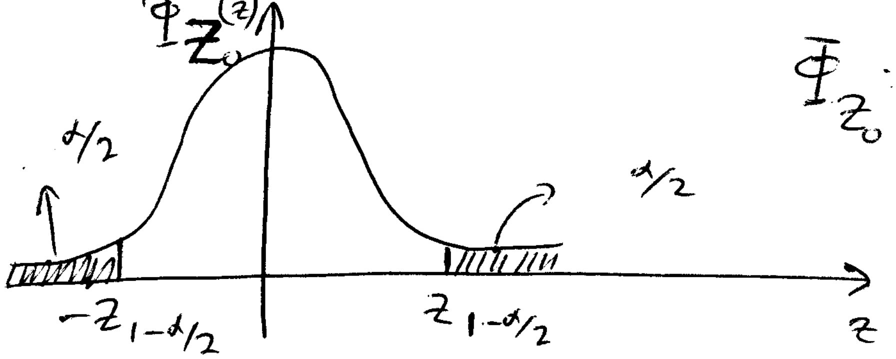
Κάτω από την Ho, η στατιστική συνάρτηση ελέγχου θα γίνει:

$$Z_0 = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}} \quad \text{και} \quad Z_0 \sim N(0,1)$$

Για τις διάφορες περιπτώσεις εναλλακτικών υποθέσεων η τιμή της Z_0 συγκρίνεται με τιμές από την $N(0,1)$ για συγκεκριμένο επίπεδο στατιστικής σημαντικότητας α.

Πιο συγκεκριμένα:

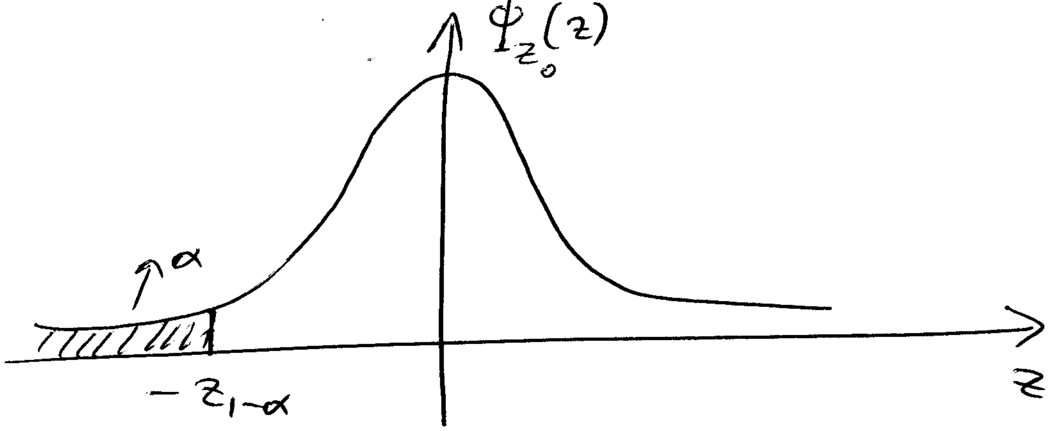
1. Αν $H_1: \mu_X \neq \mu_Y$ ή $H_1: \mu_X - \mu_Y \neq 0$, τότε απορρίπτουμε την H_0 αν $Z_0 < -Z_{1-\alpha/2}$ ή $Z_0 > Z_{1-\alpha/2}$



Φ_{Z_0}: συνάρτηση πυκνότητας της τ.μ. $Z_0 \sim N(0,1)$.

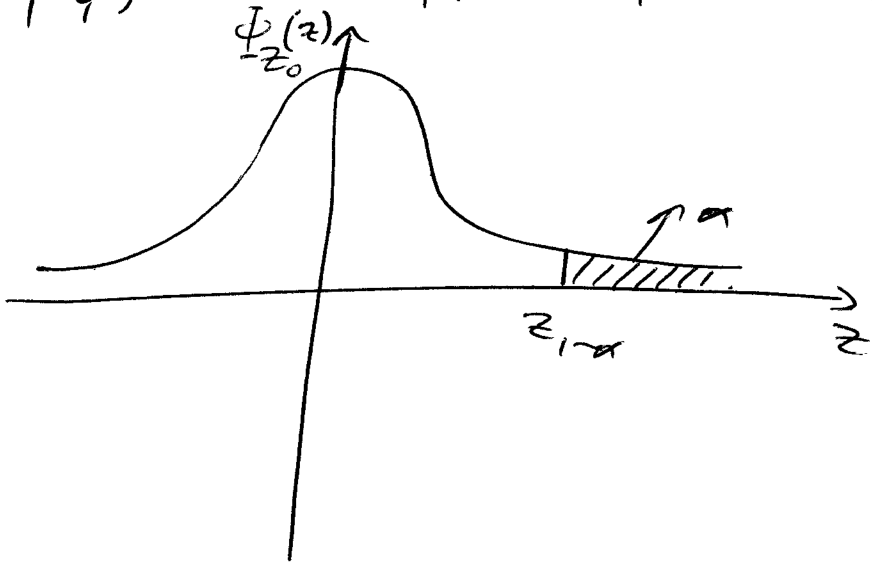
2. Αν $H_1: \mu_X < \mu_Y$, τότε απορρίπτουμε την H_0, αν

$Z_0 < -Z_{1-\alpha}$



3. Αν $H_1: \mu_X > \mu_Y$, τότε απορρίπτουμε την H_0, αν

$Z_0 > Z_{1-\alpha}$



<u>2η περίπτωση:</u> Άγνωστες διακυμάνσεις. Υποθέτουμε ότι $\sigma_X^2 = \sigma_Y^2$ και ότι οι πληθυσμοί είναι κανονικοί.

Χρησιμοποιείται η στατιστική συνάρτηση:

$$T = \frac{\overline{X} - \overline{Y} - (\mu_X - \mu_Y)}{\sqrt{\frac{S_p^2}{n} + \frac{S_p^2}{m}}}$$

Κάτω από την μηδενική υπόθεση:

$$T_0 = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{S_p^2}{n} + \frac{S_p^2}{m}}} = \frac{\overline{X} - \overline{Y}}{S_p\sqrt{\frac{1}{n} + \frac{1}{m}}}, \text{ όπου,}$$

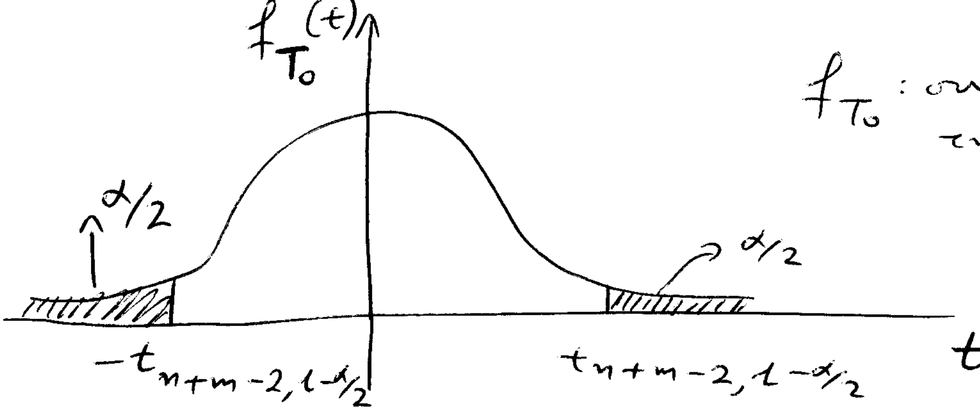
$$S_p^{*2} = \frac{(n-1)S_X^2 + (m-1)S_Y^2}{n+m-2}$$ είναι η σταθμισμένη διακύμανση (εκτιμήτρια της κοινής διακύμανσης $\sigma^2 = \sigma_X^2 = \sigma_Y^2$ του πληθυσμού).

Ισχύει ότι: $T_0 \sim t_{n+m-2}$

Για τις διάφορες μορφές της εναλλακτικής υπόθεσης, συγκρίνουμε την τιμή της T_0 με τα κατάλληλα ποσοστιαία σημεία της t_{n+m-2} κατανομής.

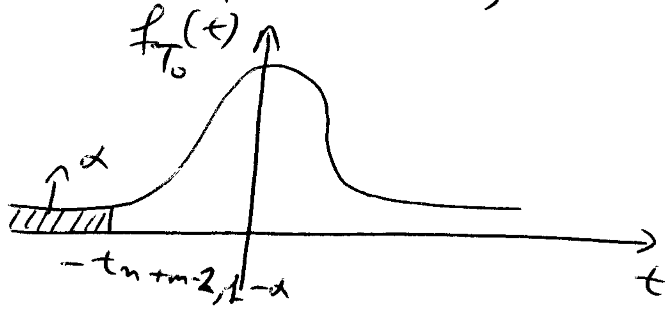
Πιο συγκεκριμένα:

1. Αν $H_1: \mu_X \neq \mu_Y$, τότε απορρίπτουμε την H_0 αν
$T_0 < -t_{n+m-2,1-\alpha/2}$ ή $T_0 > t_{n+m-2,1-\alpha/2}$

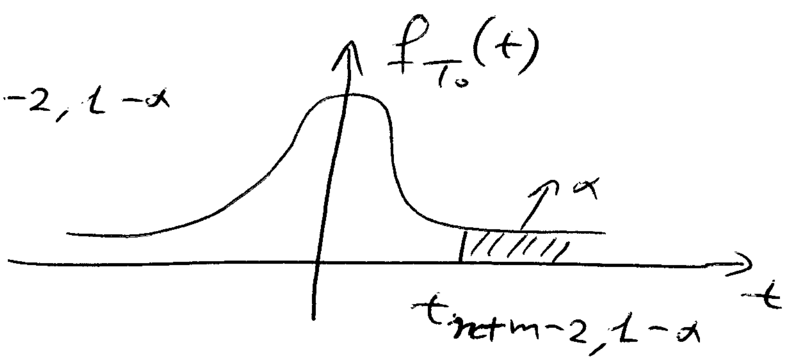


f_{T_0}: συνάρτηση πυκνότητας της τ.μ. $T_0 \sim t_{n+m-2}$.

2. Αν $H_1: \mu_X < \mu_Y$, τότε απορρίπτουμε την H_0, αν
$T_0 < -t_{n+m-2,1-\alpha}$



3. Αν $H_1: \mu_X > \mu_Y$, τότε απορρίπτουμε την H_0, αν $T_0 > t_{n+m-2,1-\alpha}$



<u>3η περίπτωση:</u> Αν τα μεγέθη των δειγμάτων n και m είναι μεγάλα, χρησιμοποιούμε είτε τις αμερόληπτες εκτιμήτριες $S_X'^{*2}$, $S_Y'^{*2}$ είτε τις μεροληπτικές εκτιμήτριες $S_X'^2$, $S_Y'^2$ των διακυμάνσεων σ_X^2, σ_Y^2, αντίστοιχα, των δύο πληθυσμών.

Οι στατιστικές συναρτήσεις ελέγχου θα είναι:

$$Z = \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sqrt{\frac{S_X'^{*2}}{n} + \frac{S_Y'^{*2}}{m}}}$$

ή $$Z = \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sqrt{\frac{S_X'^2}{(n-1)} + \frac{S_Y'^2}{(m-1)}}}$$

Κάτω από την H_0, έχουμε:

$$Z_0 = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_X'^{*2}}{n} + \frac{S_Y'^{*2}}{m}}} \quad \text{ή} \quad Z_0 = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_X'^2}{(n-1)} + \frac{S_Y'^2}{(m-1)}}}$$

Σύμφωνα με το Κ.Ο.Θ. (Κεντρικό Οριακό Θεώρημα) για μεγάλα δείγματα, $Z_0 \sim N(0,1)$.

Σ' αυτήν την περίπτωση θεωρούμε ότι οι διακυμάνσεις των πληθυσμών είναι άγνωστες και άνισες δηλ. $\sigma_X^2 \neq \sigma_Y^2$ και σ_X^2, σ_Y^2 άγνωστα. Τα μεγέθη των δειγμάτων θεωρούνται "μεγάλα" (δηλ. $n, m \geq 30$).

Συγκρίνουμε τις τιμές της Z_0 με τα κατάλληλα ποσοστιαία σημεία της $N(0,1)$ για να λάβουμε μία απόφαση αν θα απορρίψουμε ή αν δεν θα απορρίψουμε την $H_0: \mu_X = \mu_Y$ για τις τρεις εναλλακτικές μορφές της H_1,

$$H_1: \mu_X < \mu_Y, \; H_1: \mu_X > \mu_Y, \; H_1: \mu_X \neq \mu_Y.$$

Παράδειγμα 5 Τα δεδομένα του παρακάτω πίνακα δίνουν τους χρόνους ανταπόκρισης (σε ημέρες) των πελατών στα προϊόντα που προωθούν στην αγορά δύο διαφορετικές εταιρείες. Υπάρχουν ενδείξεις ότι αυτοί οι χρόνοι διαφέρουν από τη μία εταιρεία στην άλλη.

	Μέγεθος δείγματος πελατών	Μέση τιμή δείγματος	Τυπική απόκλιση δείγματος
Εταιρεία Α	12	8.5	3.6
Εταιρεία Β	10	4.8	2.1

Ελέγχουμε τις διαφορές στους μέσους χρόνους απόκρισης των πελατών στα προϊόντα που προωθούν οι δύο εταιρείες. Έστω μ_A, μ_B, αντίστοιχα οι μέσοι χρόνοι ανταπόκρισης του πληθυσμού των πελατών των εταιρειών Α, Β.

$$H_0: \mu_A = \mu_B \quad \text{έναντι της} \quad H_1: \mu_A \neq \mu_B$$

Κάτω από την H_0 χρησιμοποιούμε τη στατιστική συνάρτηση ελέγχου: $T_0 = \dfrac{\bar{X} - \bar{Y}}{S_p^* \sqrt{\frac{1}{n} + \frac{1}{m}}}$, όπου

$$S_p^{*2} = \frac{11\cdot 3.6^2 + 9\cdot 2.1^2}{12+10-2} = 9.11$$

και $\bar{x}, \bar{y}$ είναι οι μέσοι των δειγμάτων για τις εταιρείες A, B. Είναι $n=12$, $m=10$.

Το τυπικό σφάλμα είναι:

$$S_p^{*}\sqrt{\frac{1}{n}+\frac{1}{m}} = \sqrt{9.11\left(\frac{1}{12}+\frac{1}{10}\right)} = 1.29$$

Επομένως, η στατιστική συνάρτηση ελέγχου T_0 γίνεται:

$$T_0 = \frac{8.5-4.8}{1.29} = \frac{3.7}{1.29} = 2.85$$

Για $\alpha = 5\%$ από τους πίνακες της t_{20}-κατανομής έχουμε $t_{20,1-\alpha/2} = t_{20,0.975} = 2.086$.

Επειδή $T_0 > t_{20,0.975}$, σε $\alpha = 5\%$ υπάρχουν ενδείξεις για να απορρίψουμε την H_0.

Στο παράδειγμά μας, θεωρήσαμε ότι οι πληθυσμοί των πελατών των δύο εταιρειών είναι κανονικοί και ότι οι διακυμάνσεις των δύο πληθυσμών είναι ίσες.

<u>Παράδειγμα 6</u> Δεδομένα συλλέχθησαν για το επίπεδο του IQ σε παιδιά που έχουν εκτεθεί σε χαμηλά επίπεδα μόλυβδου στην Καλιφόρνια των Ηνωμένων Πολιτειών. Μας ενδιαφέρει να ελέγξουμε αν υπάρχουν διαφορές στα μέσα επίπεδα IQ σε παιδιά που εκτέθηκαν σε δύο διαφορετικά χαμηλά επίπεδα μόλυβδου χωριζόμενα σε δύο ομάδες A, B.

	Μέγεθος δείγματος	Μέση τιμή δείγματος	Τυπική απόκλιση δείγματος
Ομάδα A (ομάδα μόλυβδου υψηλότερο ποσό)	24	96.44	13.74
Ομάδα B (ομάδα ελεγχόμενης ποσότητας μόλυβδου)	26	103.29	17.87

Ελέγχουμε: $H_0: \mu_A = \mu_B$ έναντι της $H_1: \mu_A > \mu_B$

όπου μας ενδιαφέρει αν η ομάδα A των παιδιών που εκτέθηκαν σε υψηλότερα ποσά μόλυβδου έχουν ή όχι υψηλότερα μέσα επίπεδα IQ από ότι η ομάδα των παιδιών B που εκτέθηκαν σε ελεγχόμενες ποσότητες μολύβδου. Με μ_A και μ_B συμβολίζουμε τα αντίστοιχα μέσα επίπεδα IQ του πληθυσμού των παιδιών των ομάδων A και B, αντίστοιχα. Υποθέτουμε ότι οι πληθυσμοί των ομάδων A και B είναι κανονικοί και ότι οι διακυμάνσεις στους δύο πληθυσμούς είναι ίσες.

Είναι: $$S_p^{*2} = \frac{23 \cdot 13.74^2 + 25 \cdot 17.87^2}{24+26-2} = 256.78$$

Τυπικό σφάλμα $$\sqrt{256.78\left(\frac{1}{24} + \frac{1}{26}\right)} = 4.54$$

Στατιστική Συνάρτηση Ελέγχου:

$$T_0 = \frac{103.29 - 96.44}{4.54} = 1.5088$$

Συγκρίνουμε την τιμή της T_0 με το ποσοστιαίο σημείο της t_{48}-κατανομής. Είναι $t_{48,1-\alpha} = 2.403$, για $\alpha = 0.01$.

Επομένως, σε $\alpha = 1\%$, δεν υπάρχουν ενδείξεις για να απορρίψουμε την H_0.

<u>Σύγκριση αναλογιών δύο πληθυσμών - Ανεξάρτητα δείγματα (Σελ. 434 - 438)</u>

Ενδιαφερόμαστε να συγκρίνουμε αναλογίες στοιχείων (ή ατόμων) **δύο** πληθυσμών που έχουν κάποια συγκεκριμένη ιδιότητα. Για παράδειγμα, εξετάζουμε τα ποσοστά των πολιτών που στηρίζουν ένα κόμμα σε δύο διαφορετικές περιοχές της χώρας. Επιλέγουμε δύο τυχαία δείγματα μεγέθους n και m, αντίστοιχα, από τους δύο πληθυσμούς και έστω X και Y ο αριθμός των στοιχείων (ή των ατόμων) σε καθένα από τα δείγματα που έχουν την ιδιότητα που μας ενδιαφέρει. Έστω P_X, P_Y οι αντίστοιχες αναλογίες των ατόμων των δύο πληθυσμών που έχουν την υπό μελέτη ιδιότητα.

Ελέγχουμε την $H_0: P_X = P_Y$ έναντι της:

$$H_1: P_X > P_Y \quad \text{ή} \quad H_1: P_X < P_Y \quad \text{ή} \quad H_1: P_X \neq P_Y$$

Ισχύει ότι: $X \sim B(n, P_X)$ και $Y \sim B(m, P_Y)$

Για μεγάλα δείγματα, οι εκτιμήτριες $\hat{P}_X$, $\hat{P}_Y$ των

p_X και p_Y, αντίστοιχα, ακολουθούν κατά προσέγγιση, την κανονική κατανομή.

$$\hat{p}_X = \frac{X}{n} \underset{\text{προσεγγιστικά}}{\overset{(n\to\infty)}{\sim}} N\left(p_X, \frac{p_X(1-p_X)}{n}\right)$$

$$\hat{p}_Y = \frac{Y}{m} \underset{\text{προσεγγιστικά}}{\overset{(n\to\infty)}{\sim}} N\left(p_Y, \frac{p_Y(1-p_Y)}{m}\right)$$

Αφού τα δείγματα είναι ανεξάρτητα,

$$\hat{p}_X - \hat{p}_Y = \frac{X}{n} - \frac{Y}{m} \underset{\text{προσεγγιστικά}}{\sim} N\left(p_X - p_Y, \frac{p_X(1-p_X)}{n} + \frac{p_Y(1-p_Y)}{m}\right)$$

Χρησιμοποιούμε την στατιστική συνάρτηση

$$Z = \frac{\hat{p}_X - \hat{p}_Y - (p_X - p_Y)}{\sqrt{\frac{p_X(1-p_X)}{n} + \frac{p_Y(1-p_Y)}{m}}}$$

Κάτω από την H_0 ισχύει ότι $p_X = p_Y = p$. Μπορούμε να εκτιμήσουμε το ποσοστό p με τη σταθμισμένη εκτιμήτρια $\hat{p}$ που προέρχεται από τα δύο δείγματα.

$$\hat{p} = \frac{n\hat{p}_X + m\hat{p}_Y}{n+m} = \frac{X+Y}{n+m}$$

Κάτω από την H_0 η στατιστική συνάρτηση ελέγχου Z_0 θα γίνει:

$$Z_0 = \frac{\hat{p}_X - \hat{p}_Y}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n}+\frac{1}{m}\right)}} = \frac{\frac{X}{n} - \frac{Y}{m}}{\sqrt{\left(\frac{X}{n}+\frac{Y}{m}\right)\left(1-\frac{X}{n}-\frac{Y}{m}\right)\left(\frac{1}{n}+\frac{1}{m}\right)}}$$

Για μεγάλα δείγματα, $Z_0 \sim N(0,1)$. Επομένως, η λήψη της απόφασης αν θα απορρίψουμε ή <u>δεν</u> θα απορρίψουμε την H_0 πραγματοποιείται με σύγκριση της τιμής της Z_0 με τα κατάλληλα ποσοστιαία σημεία από την κανονική κατανομή $N(0,1)$, για διάφορα επίπεδα στατιστικής σημαντικότητας α.

Πιο συγκεκριμένα:

1. Αν $H_1: p_X \neq p_Y$ τότε απορρίπτουμε την H_0, σε επίπεδο α, αν $Z_0 < -z_{1-\alpha/2}$ ή $Z_0 > z_{1-\alpha/2}$.
2. Αν $H_1: p_X < p_Y$ τότε απορρίπτουμε την H_0 σε επίπεδο α αν $Z_0 < -z_{1-\alpha}$.
3. Αν $H_1: p_X > p_Y$ τότε απορρίπτουμε την H_0 σε επίπεδο α αν $Z_0 > z_{1-\alpha}$.

<u>Παράδειγμα 7</u> Θέλουμε να ελέγξουμε την επίδραση που είχαν στους ερωτώμενους οι ερωτήσεις ενός τύπου ερωτηματολογίου για τον ρατσισμό. Σε 500 ερωτηματολόγια που δεν υπήρχαν ερωτήσεις για τον ρατσισμό υπήρξαν 240 απαντήσεις. Σε 450 ερωτηματολόγια που υπήρχαν ερωτήσεις για τον ρατσισμό είχαμε 200 απαντήσεις.

Υπάρχουν ενδείξεις ότι οι ερωτήσεις για τον ρατσισμό οδήγησαν σε μικρότερη συμμετοχή στην συμπλήρωση του ερωτηματολογίου;

Έστω P_1, P_2 τα ποσοστά του πληθυσμού που απάντησαν στους δύο τύπους ερωτηματολογίων (χωρίς ερωτήσεις για τον ρατσισμό, με ερωτήσεις για τον ρατσισμό).

Ελέγχουμε την $H_0: P_1 = P_2$ έναντι της $H_1: P_1 > P_2$.

Η διαφορά $P_1 - P_2$ εκτιμάται από τη διαφορά:

$$\frac{240}{500} - \frac{200}{450}.$$

Το τυπικό σφάλμα εκτίμησης είναι:

$$\sqrt{\hat{P}(1-\hat{P})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}, \text{ όπου}$$

$n_1 = 500$, $n_2 = 450$ είναι τα δύο μεγέθη δειγμάτων από τους δύο τύπους ερωτηματολογίων.

$$\hat{P} = \frac{240+200}{500+450} = 0.463$$

Η στατιστική συνάρτηση ελέγχου Z_0 κάτω από την H_0 δίνει:

$$Z_0 = \frac{\frac{240}{500} - \frac{200}{450}}{\sqrt{0.463 \cdot 0.537\left(\frac{1}{500} + \frac{1}{450}\right)}} = 1.097$$

Σε επίπεδο στατιστικής σημαντικότητας $\alpha = 5\% = 0.05$, απορρίπτουμε την H_0 αν $Z_0 > Z_{1-\alpha}$. Όμως

$$Z_{1-\alpha} = Z_{1-0.05} = Z_{0.95} = 1.65$$

Άρα, η H_0 δεν απορρίπτεται σε $\alpha = 5\%$.

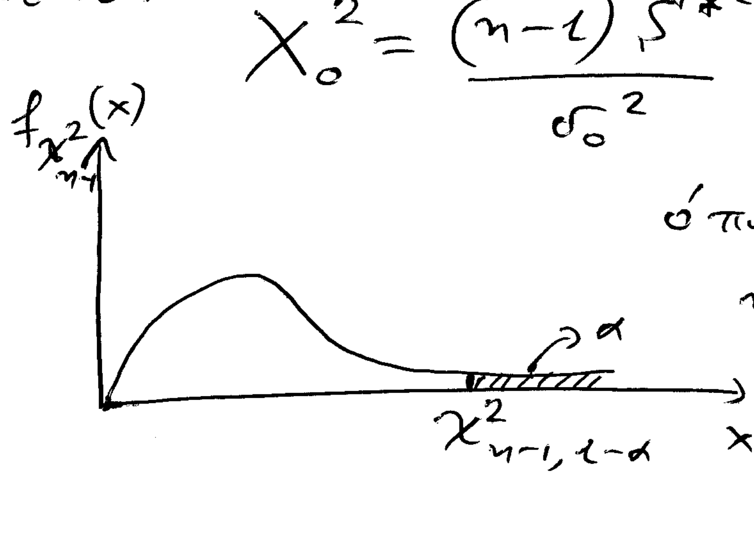
Η p-τιμή είναι: $P(Z_0 > 1.097) = 0.1379$ (από πίνακες)

Επομένως, η H_0 απορρίπτεται σε οποιοδήποτε επίπεδο στατιστικής σημαντικότητας μεγαλύτερο ή ίσο του 13.79%.

<u>Έλεγχοι υποθέσεων για διακυμάνσεις (σελ. 439-448)</u>

Αν ο υπό μελέτη πληθυσμός είναι κανονικός τότε η στατιστική συνάρτηση:
$$X^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{\sigma^2} = \frac{(n-1)S^{*2}}{\sigma^2} = \frac{nS^2}{\sigma^2} \sim \chi^2_{n-1}$$

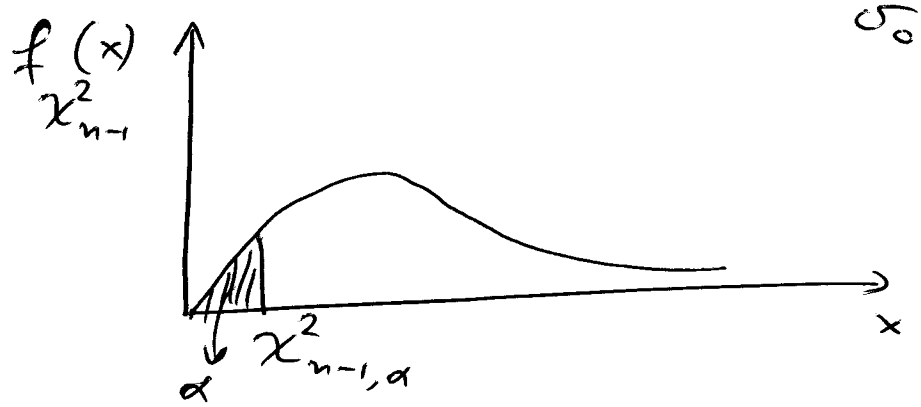
Επομένως για τον έλεγχο της $H_0: \sigma^2 = \sigma_0^2$ έναντι της $H_1: \sigma^2 > \sigma_0^2$ θα απορρίπτουμε την H_0 σε επίπεδο στατιστικής σημαντικότητας α όταν η στατιστική συνάρτηση ελέγχου κάτω από την H_0 είναι τέτοια ώστε:
$$X_0^2 = \frac{(n-1)S^{*2}}{\sigma_0^2} > \chi^2_{n-1,1-\alpha}$$



όπου $f_{\chi^2_{n-1}}(x)$ είναι η συνάρτηση πυκνότητας της τ.μ. X_0^2 που ακολουθεί την χ^2_{n-1}.

Για τον έλεγχο:

$H_0: \sigma^2 = \sigma_0^2$ έναντι της $H_1: \sigma^2 < \sigma_0^2$, θα απορρίπτουμε την H_0 όταν $X_0^2 = \dfrac{(n-1)\, S'^{*2}}{\sigma_0^2} < \chi^2_{n-1,\alpha}$

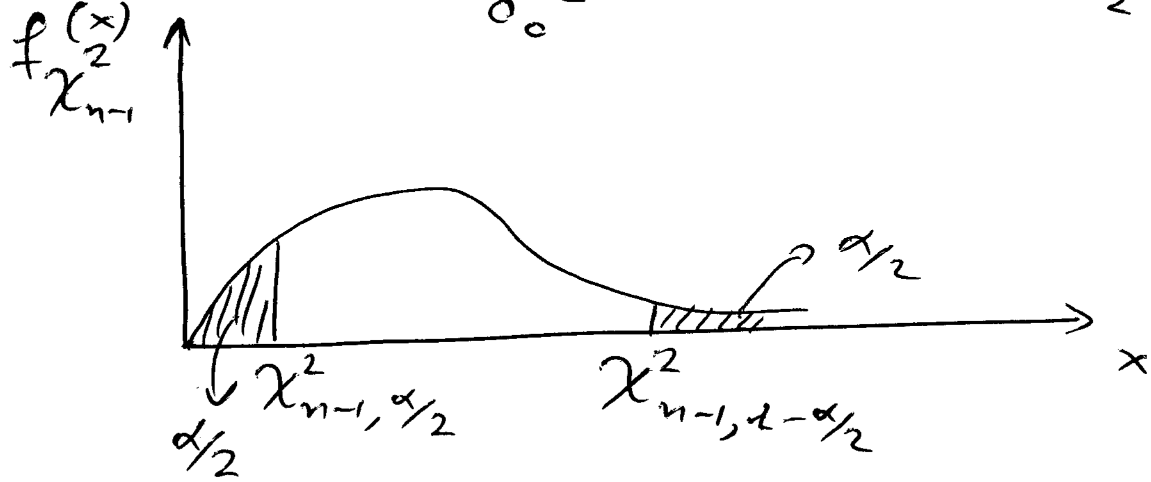


Για τον έλεγχο:

$H_0: \sigma^2 = \sigma_0^2$ έναντι της $H_1: \sigma^2 \neq \sigma_0^2$, θα απορρίπτουμε την H_0 όταν

$$X_0^2 = \frac{(n-1)\, S'^{*2}}{\sigma_0^2} > \chi^2_{n-1,1-\alpha/2}$$

ή $$X_0^2 = \frac{(n-1)\, S'^{*2}}{\sigma_0^2} < \chi^2_{n-1,\alpha/2}$$



<u>Παράδειγμα 8</u> Η τυπική απόκλιση του βάρους του περιεχομένου στα corn-flakes ορισμένης μάρκας ισούται με $\sigma = 22.7$ gr και η παραγωγός επιχείρηση, θεωρώντας ότι η διασπορά του βάρους είναι πολύ μεγάλη, κάνει νέα

ρύθμιση του μηχανήματος συσκευασίας. Αν σε τυχαίο δείγμα $n=25$ κουτιών, τα οποία συσκευάστηκαν μετά την ρύθμιση, υπολογίσαμε $s^*=21.3$ gr (δειγματική τυπική απόκλιση, αμερόληπτη εκτίμηση), να ελεγχθεί η υπόθεση ότι η διακύμανση του πληθυσμού δεν μειώθηκε σε επίπεδο στατιστικής σημαντικότητας $\alpha=0.01$. Η κατανομή του βάρους μπορεί να υποτεθεί κανονική.

<u>Λύση</u>

Είναι $\sigma^2=515.29$

Ελέγχουμε την $H_0: \sigma^2=515.29$ έναντι της:

$H_1: \sigma^2<515.29$

Για $\alpha=0.01$ και για $\sigma_0^2=515.29$ χρησιμοποιούμε τη στατιστική συνάρτηση ελέγχου:

$$X_0^2=\frac{(n-1)\,s'^{*2}}{\sigma_0^2}=\frac{(24)(21.3)^2}{515.29}=21.13$$

Από τους πίνακες της χ^2_{24}-κατανομής είναι: $\chi^2_{24,0.99}=10.86$

Δεν υπάρχουν ενδείξεις για να απορρίψουμε την H_0 σε $\alpha=0.01$.

<u>Σύγκριση διακυμάνσεων δύο κανονικών πληθυσμών</u>

Η μηδενική υπόθεση $H_0: \sigma_X^2=\sigma_Y^2$ μπορεί να γραφεί στη μορφή: $H_0: \dfrac{\sigma_X^2}{\sigma_Y^2}=1$

Η στατιστική συνάρτηση:

$$\frac{\dfrac{(n-1)S_X^{*2}}{\sigma_X^2}\Big/(n-1)}{\dfrac{(m-1)S_Y^{*2}}{\sigma_Y^2}\Big/(m-1)}\sim F_{n-1,m-1}$$

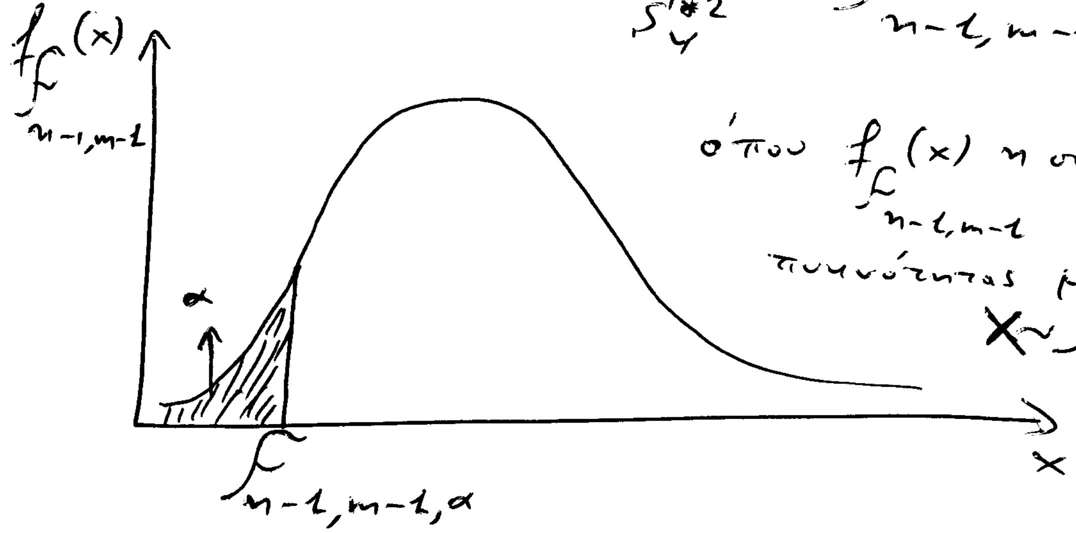
Κάτω από τη μηδενική υπόθεση $H_0: \sigma_X^2 = \sigma_Y^2$, η παραπάνω στατιστική συνάρτηση γράφεται:

$$F_0 = \frac{S_X'^{*2}}{S_Y'^{*2}} \sim F_{n-1,m-1}$$

Για τις διάφορες εναλλακτικές υποθέσεις έχουμε:

1. Για τον έλεγχο της $H_0: \frac{\sigma_X^2}{\sigma_Y^2} = 1$ έναντι της $H_1: \frac{\sigma_X^2}{\sigma_Y^2} < 1$

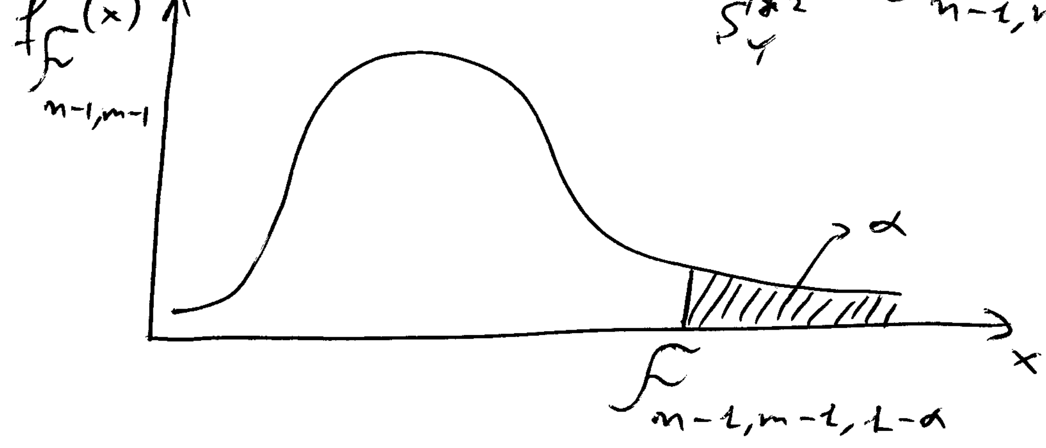
Θα απορρίπτουμε την H_0 σε επίπεδο στατιστικής σημαντικότητας α αν $F_0 = \frac{S_X'^{*2}}{S_Y'^{*2}} < F_{n-1,m-1,\alpha}$.



όπου $f_{F_{n-1,m-1}}(x)$ η συνάρτηση πυκνότητας μιας τ.μ. $X \sim F_{n-1,m-1}$

2. Για τον έλεγχο της $H_0: \frac{\sigma_X^2}{\sigma_Y^2} = 1$ έναντι της $H_1: \frac{\sigma_X^2}{\sigma_Y^2} > 1$

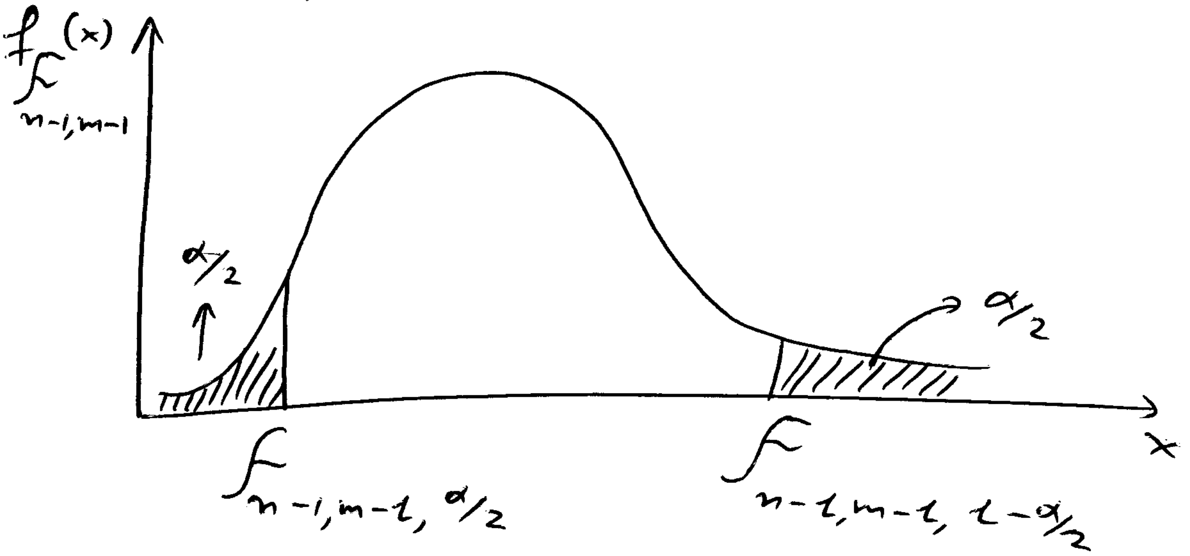
Θα απορρίπτουμε την H_0 σε επίπεδο στατιστικής σημαντικότητας α αν $F_0 = \frac{S_X'^{*2}}{S_Y'^{*2}} > F_{n-1,m-1,1-\alpha}$



3. Για τον έλεγχο της $H_0: \frac{\sigma_X^2}{\sigma_Y^2} = 1$ έναντι της $H_1: \frac{\sigma_X^2}{\sigma_Y^2} \neq 1$

Θα απορρίπτουμε την H_0 σε επίπεδο στατιστικής σημαντικότητας α, αν $F_0 = \frac{S_X'^{*2}}{S_Y'^{*2}} > F_{n-1,m-1,1-\alpha/2}$

ή $F_0 = \frac{S_X'^{*2}}{S_Y'^{*2}} < F_{n-1,m-1,\alpha/2}$



Τα δύο ανεξάρτητα δείγματα μεγέθους n, m, αντίστοιχα, προέρχονται από δύο ανεξάρτητους κανονικούς πληθυσμούς.

<u>Παράδειγμα 9</u> Μηχάνημα αυτόματης συσκευασίας, όταν εμφανίζει κάποια προβλήματα συσκευάζει το γάλα με μεγάλη διασπορά στο περιεχόμενο. Σε τυχαίο δείγμα $n_1 = 22$ κουτιών του 0.5 lt υπολογίσαμε τυπική απόκλιση $S_1^* = 12$ ml, ενώ σε τυχαίο δείγμα $n_2 = 11$ κουτιών του 1 lt υπολογίσαμε τυπική απόκλιση $S_2^* = 33$ ml. Να ελεγχθεί η υπόθεση ότι η διασπορά στις δύο συσκευασίες δεν διαφέρει σημαντικά σε επίπεδο στατιστικής σημαντικότητας $\alpha = 10\%$.

<u>Λύση</u>

Υποθέτουμε ότι η κατανομή του περιεχομένου συσκευασίας μπορεί να υποτεθεί κανονική.

Ελέγχουμε: $H_0: \frac{\sigma_1^2}{\sigma_2^2} = 1$ έναντι της $H_1: \frac{\sigma_1^2}{\sigma_2^2} \neq 1$

Για $\alpha = 0.10$, για $\nu_1 = 22-1 = 21$, $\nu_2 = 11-1 = 10$

από πίνακες: $F_{21,10,0.05} = 2.76$ και

$$F_{21,10,0.95} = \frac{1}{F_{10,21,0.05}} = \frac{1}{2.32} = 0.43$$

Όμως

$$F_0 = \frac{s_1'^{*2}}{s_2'^{*2}} = \frac{(12)^2}{(33)^2} = 0.13 < 0.43$$

επομένως, απορρίπτουμε την μηδενική υπόθεση της ίσης διασποράς στις δύο συσκευασίες. Η μεγάλη συσκευασία φαίνεται να έχει μεγαλύτερη διασπορά.