ΚΕΦΑΛΑΙΟ 18

ΕΛΕΓΧΟΙ ΥΠΟΘΕΣΕΩΝ ΓΙΑ ΜΕΣΕΣ ΤΙΜΕΣ ΚΑΙ ΑΝΑΛΟΓΙΕΣ ΚΑΝΟΝΙΚΩΝ ΠΛΗΘΥΣΜΩΝ

Στο κεφάλαιο αυτό θα μας απασχολήσουν έλεγχοι στατιστικών υποθέσεων που αναφέρονται στις μέσες τιμές και αναλογίες πληθυσμών που περιγράφονται, είτε με ακρίβεια, είτε κατά προσέγγιση, από την κανονική κατανομή.

ΕΛΕΓΧΟΙ ΥΠΟΘΕΣΕΩΝ ΓΙΑ ΤΗΝ ΜΕΣΗ ΤΙΜΗ

Α. Περίπτωση Γνωστών Διακυμάνσεων

Καθορισμός του Κρίσιμου Σημείου

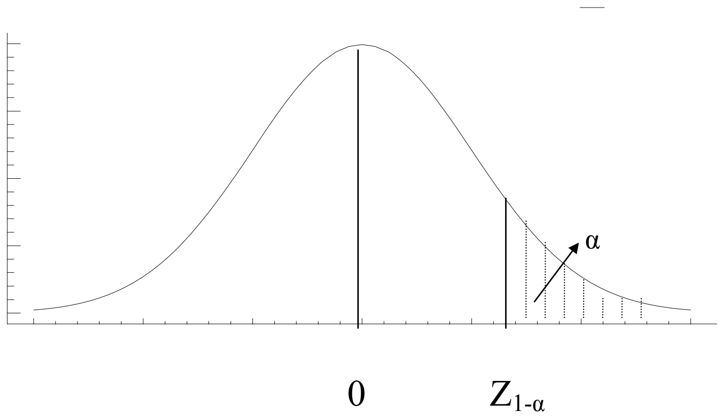
Όπως είπαμε και προηγουμένως, το συνηθέστερο πρόβλημα στον έλεγχο υποθέσεων είναι ο προσδιορισμός της κρίσιμης περιοχής. Για τον σκοπό αυτό, χρειάζεται να καθορίσουμε ή το α ή το β. Στις περισσότερες περιπτώσεις, εκείνο το οποίο καθορίζουμε είναι το α. Τότε μπορούμε να βρούμε το κρίσιμο σημείο (δηλαδή να καθορίσουμε την κρίσιμη περιοχή).

Παράδειγμα: Έστω ότι έχουμε ένα κανονικό πληθυσμό $N(\mu, \sigma^2)$, όπου το σ^2 είναι γνωστό. Έστω ότι θέλουμε να ελέγξουμε την στατιστική υπόθεση

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu > \mu_0$$

Από όσα είπαμε προηγουμένως, θα έχουμε, αν με c συμβολίσουμε το κρίσιμο σημείο,

$$\begin{aligned}
\alpha &= P(\cancel{H_0} \mid H_0) \\
&= P(\overline{X} > c \mid \mu = \mu_0) \\
&= P\left[\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} > \frac{c - \mu}{\sigma/\sqrt{n}} \mid \mu = \mu_0 \right] \\
&= P\left[Z > \frac{c - \mu_0}{\sigma/\sqrt{n}} \right]
\end{aligned}$$



Προφανώς,

$$\frac{c-\mu_0}{\sigma/\sqrt{n}} = Z_{1\text{-}a}$$

ή ισοδύναμα,

$$c = \mu_0 + Z_{1-a}\frac{\sigma}{\sqrt{n}}$$

Επομένως, θα απορρίπτουμε την H_0 αν,

$$\overline{X} > \mu_0 + Z_{1-a}\frac{\sigma}{\sqrt{n}}$$

σε επίπεδο σημαντικότητας α.

Σημείωση: Συνήθως χρησιμοποιούμε, ισοδύναμα, την τυποποιημένη ελεγχοσυνάρτηση και απορρίπτουμε την H_0 σε επίπεδο σημαντικότητας α αν,

$$Z_0 = \frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} > Z_{1\text{-}\alpha}$$

Σημείωση: Η τυποποιημένη συνάρτηση ελέγχου, ή αλλιώς Z_0-στατιστική συνάρτηση (Z_0 statistic)

$$Z_0 = \frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$$

που χρησιμοποιούμε στους ελέγχους της μέσης τιμής κανονικών πληθυσμών με γνωστή διακύμανση είναι το κλασσικό παράδειγμα τυποποιημένης ελεγχοσυνάρτησης. (Ο δείκτης 0 χρησιμοποιείται για να δώσει έμφαση στο γεγονός ότι η τυποποίηση γίνεται σε σχέση με την συγκεκριμένη τιμή της παραμέτρου που μελετάται με την μηδενική υπόθεση). Σε κάθε περίπτωση ελέγχου υποθέσεων για μια παράμετρο ενός πληθυσμού, χρησιμοποιούμε ως τυποποιημένη συνάρτηση ελέγχου την τυποποιημένη τιμή της εκτιμήτριας της παραμέτρου που μελετάμε, κάτω από την μηδενική υπόθεση. (Εδώ, παράμετρος κάτω από την H_0 είναι το μ_0, εκτιμήτριά του είναι το $\overline{X}$ και τυπική απόκλιση της κατανομής του $\overline{X}$ είναι το $\sigma/\sqrt{n}$).

Οι έλεγχοι που χρησιμοποιούν την Z_0-στατιστική συνάρτηση ονομάζονται *Z-έλεγχοι (Z-tests).*

> *Η τιμή της Z_0-στατιστικής συνάρτησης δηλώνει τον αριθμό των τυπικών αποκλίσεων που μια παρατηρηθείσα τιμή της συνάρτησης ελέγχου απέχει από την μέση της τιμή, όπου η μέση αυτή τιμή υπολογίζεται κάτω από την μηδενική υπόθεση.*

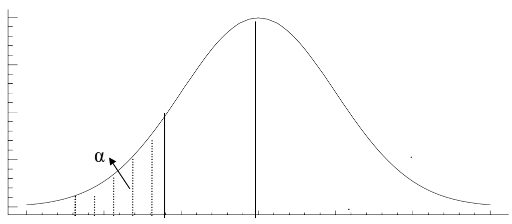
Με την ίδια λογική, καταλήγουμε στο συμπέρασμα ότι αν έχουμε να ελέγξουμε την στατιστική υπόθεση

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu < \mu_0$$

θα απορρίπτουμε την H_0 σε επίπεδο σημαντικότητας α αν

$$Z_0 = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} < -Z_{1-\alpha}$$

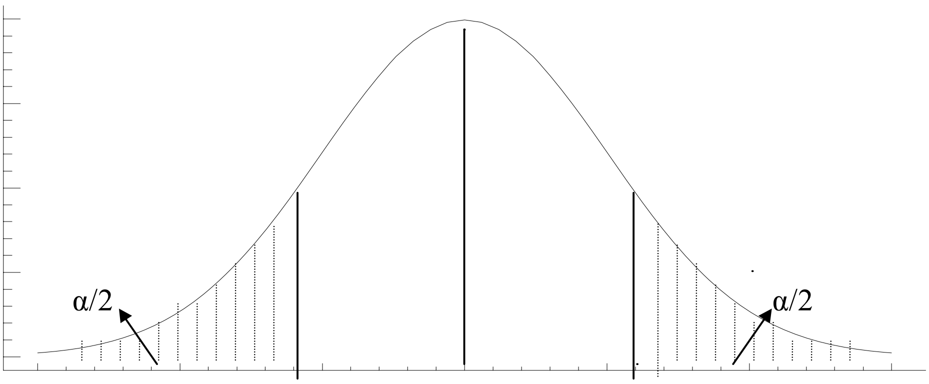
[ισοδύναμα, αν $\overline{X} < \mu_0 - Z_{1-\alpha}\dfrac{\sigma}{\sqrt{n}}$].



$Z_\alpha = -Z_{1-\alpha}$ 0

Τέλος, αν η εναλλακτική στατιστική υπόθεση είναι αμφίπλευρη, αν δηλαδή

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu \neq \mu_0$$



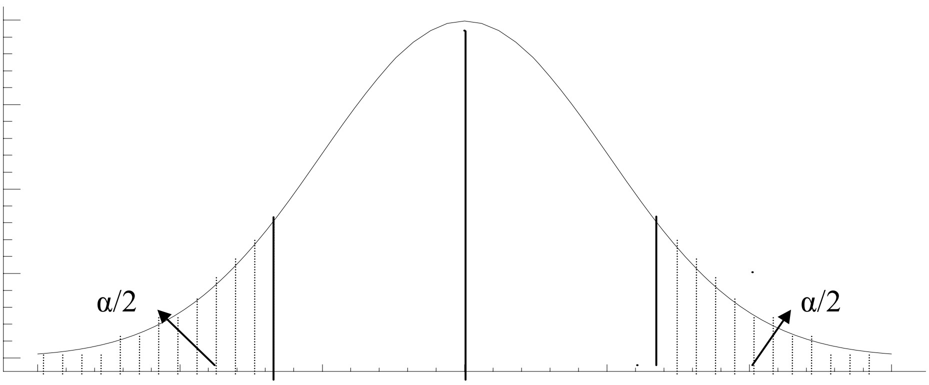
$-Z_{1-\alpha/2}$ 0 $Z_{1-\alpha/2}$

θα απορρίπτουμε την H_0 αν,

$$Z < -Z_{1-\alpha/2} \text{ ή } Z > Z_{1-\alpha/2}$$

[ή ισοδύναμα

$$\overline{X} < \mu_0 - Z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \qquad \text{ή} \qquad \overline{X} > \mu_0 + Z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \;]$$

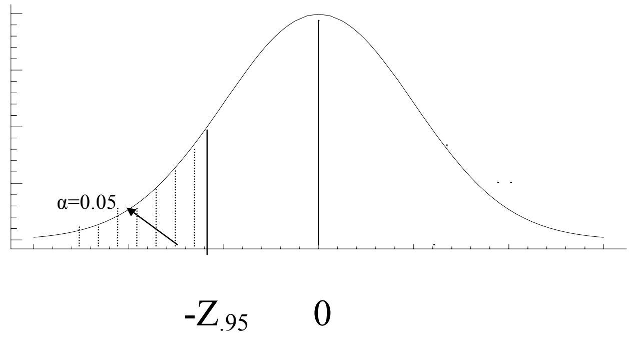


c_1 μ_0 c_2

Παράδειγμα: Ένας κατασκευαστής ηλεκτρικών λαμπτήρων ισχυρίζεται ότι η διάρκεια ζωής (σε ώρες) κάποιου συγκεκριμένου τύπου ηλεκτρικών λαμπτήρων που κατασκευάζει έχει μέση τιμή (μέση ζωή) 700 ώρες. Από προηγούμενη εμπειρία, είναι γνωστό ότι ο χρόνος ζωής των ηλεκτρικών λαμπτήρων ακολουθεί την κανονική κατανομή. Ας υποθέσουμε ότι η διασπορά του χρόνου ζωής των λαμπτήρων της κατασκευής αυτής είναι $\sigma^2 = (46.14)^2$. Μια εταιρεία προστασίας καταναλωτών, προκειμένου να ελέγξει τον ισχυρισμό αυτό του κατασκευαστή, επιλέγει ένα τυχαίο δείγμα 10 λαμπτήρων του συγκεκριμένου είδους και τους ελέγχει. Ο έλεγχος δείχνει ότι ο μέσος χρόνος ζωής των λαμπτήρων αυτών είναι $\overline{X} = 689.8$ ώρες.
Με βάση τα αποτελέσματα του συγκεκριμένου αυτού δείγματος, τί θα μπορούσε να πει κανείς για τον ισχυρισμό του κατασκευαστή σε επίπεδο σημαντικότητας $\alpha = 0.05$;

Λύση: Είναι προφανές ότι ο μέσος χρόνος ζωής των λαμπτήρων, ο οποίος προέκυψε από το δείγμα, είναι μικρότερος από αυτόν που ο κατασκευαστής ισχυρίζεται. Το στατιστικό όμως ερώτημα είναι αν ο μέσος αυτός χρόνος είναι σημαντικά μικρότερος (στην στατιστική ορολογία) από τον ισχυρισμό του κατασκευαστή ή αν η διαφορά που προέκυψε είναι αποτέλεσμα της τυχαίας δειγματοληψίας και δεν είναι αρκετή ώστε να οδηγήσει τον σύνδεσμο καταναλωτών στο να απορρίψει τον ισχυρισμό του κατασκευαστή με στατιστικά ισχυρά κριτήρια.
Η στατιστική υπόθεση που θα πρέπει να ελεγχθεί είναι,

$$H_0 : \mu = 700$$
$$H_1 : \mu < 700$$



Η τυποποιημένη τιμή της ελεγχοσυνάρτησης για το πρόβλημά μας, κάτω από την υπόθεση H_0 είναι,

$$Z_0 = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$$

$$= \frac{680.8 - 700}{46.14/\sqrt{10}}$$

$$= -0.699$$

Δοθέντος ότι $Z > -Z_{.95}$ το συγκεκριμένο δείγμα δεν δίνει αρκετές ενδείξεις που να οδηγούν στην απόρριψη της H_0.

Παράδειγμα: Μια βιομηχανία παρασκευής και συσκευασίας καφέ χρησιμοποιεί αεροστεγείς συσκευασίες που περιέχουν 368gr καφέ. Όπως είναι φυσικό, δεν είναι δυνατό να επιτυγχάνεται πάντοτε συσκευασία που να περιέχει ακριβώς το περιεχόμενο αυτό.

1. Ο υπεύθυνος της συσκευασίας προκειμένου να ελέγξει το κατά πόσο η επιδίωξη αυτή επιτυγχάνεται, επιλέγει ένα τυχαίο δείγμα 25 πακέτων που έχουν συσκευασθεί με τον τρόπο αυτό. Μετρώντας το περιεχόμενο στις συσκευασίες αυτές διαπιστώνει ότι η μέση ποσότητα καφέ που περιέχεται στις συσκευασίες αυτές είναι 364.1gr. ($\overline{x}$=364.1gr). Με βάση το στοιχείο αυτό σε τί συμπέρασμα μπορεί να καταλήξει ο προϊστάμενος της εταιρείας όσον αφορά την επιδίωξή του;
2. Μια εταιρεία προστασίας καταναλωτών ενδιαφέρεται να ελέγξει κατά πόσον ο ισχυρισμός αυτός της συγκεκριμένης εταιρείας (για περιεχόμενο στις συσκευασίες 368gr καφέ) ισχύει. Με βάση το παραπάνω δείγμα σε τι συμπέρασμα μπορεί να καταλήξει η εταιρεία προστασίας των καταναλωτών;

(Από προηγούμενη εμπειρία, είναι γνωστό ότι η ποσότητα καφέ η οποία περιέχεται στις συσκευασίες αυτές ακολουθεί την κανονική κατανομή με τυπική απόκλιση σ = 15 gr).
Ως επίπεδο σημαντικότητας να χρησιμοποιηθεί το α = 0.05.

Λύση:

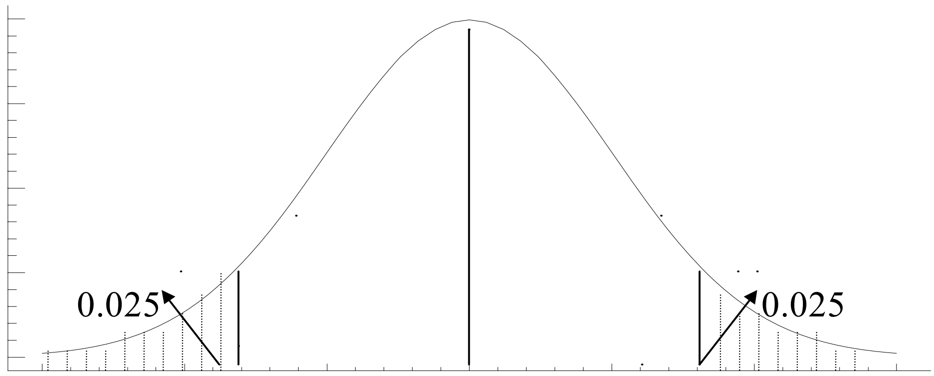
1. Όσον αφορά το πρώτο ερώτημα ο υπεύθυνος ελέγχου ποιότητας της εταιρείας θα πρέπει να ελέγξει την υπόθεση,

$$H_0 : \mu = 368$$

$$H_1 : \mu \neq 368$$

Η τιμή της στατιστικής συνάρτησης ελέγχου για το συγκεκριμένο δείγμα είναι,

$$Z_0 = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{364.1 - 368}{15/\sqrt{25}} = -1.30$$



$Z_{.025} = -Z_{0.975} = -1.96.$ $\quad$ $Z_{0.975} = 1.96$

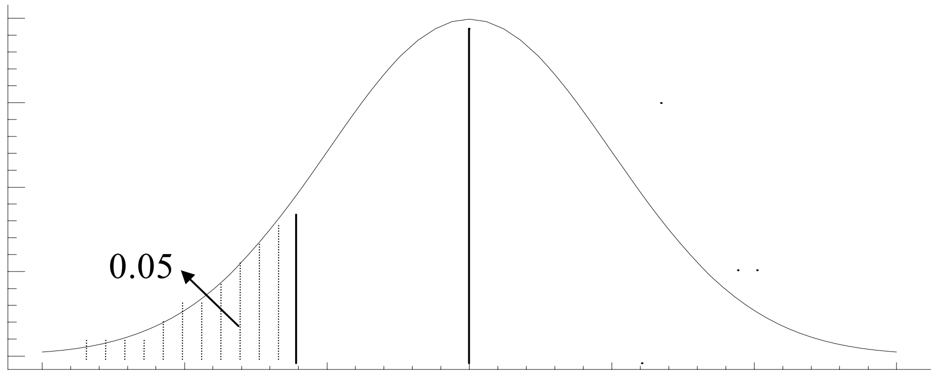
Όπως προκύπτει από τους πίνακες, η κρίσιμη περιοχή για το συγκεκριμένο πρόβλημα είναι αυτή που αναφέρεται σε τιμές της στατιστικής συνάρτησης ελέγχου

$$|Z| > Z_{.975} = 1.96$$

Δοθέντος ότι η κρίσιμη περιοχή για το πρόβλημα αυτό περιλαμβάνει τις τιμές της στατιστικής συνάρτησης ελέγχου $Z < -Z_{.975}$ και τις τιμές $Z > Z_{.975}$, καταλήγουμε στο συμπέρασμα ότι για το συγκεκριμένο δείγμα η παρατηρηθείσα τιμή της στατιστικής συνάρτησης ελέγχου δεν βρίσκεται στην κρίσιμη περιοχή. Επομένως, δεν υπάρχουν ισχυρές στατιστικές ενδείξεις που να οδηγούν στο συμπέρασμα ότι η μηδενική υπόθεση θα πρέπει να απορριφθεί στο δοθέν επίπεδο σημαντικότητας. Δεν υπάρχουν δηλαδή ενδείξεις ότι η μέση ποσότητα που περιέχεται στις συσκευασίες αυτές διαφέρει στατιστικά σημαντικά από τα 368gr.

2. Όσον αφορά τις επιδιώξεις της εταιρείας προστασίας του καταναλωτή, ο κατάλληλος έλεγχος υπόθεσης είναι:

$$H_0 : \mu = 368$$
$$H_1 : \mu < 368$$



$$Z_{.05} = - Z_{.95} = - 1.645$$

Και στην περίπτωση αυτή, η τιμή της στατιστικής συνάρτησης ελέγχου για το συγκεκριμένο δείγμα είναι $Z > - Z_{.95}$ δηλαδή, η τιμή αυτή της στατιστικής συνάρτησης ελέγχου δεν βρίσκεται στην κρίσιμη περιοχή. Επομένως, και η εταιρεία προστασίας καταναλωτή δεν έχει ισχυρές ενδείξεις που να την οδηγούν στο συμπέρασμα ότι τα συμφέροντα των καταναλωτών που προτιμούν το συγκεκριμένο προϊόν θίγονται (ότι δηλαδή η ποσότητα του καφέ που περιέχεται στην συγκεκριμένη συσκευασία υπολείπεται στατιστικά σημαντικά των 368gr).

Καθορισμός του Επιπέδου Σημαντικότητας

Ελέγχθη ήδη ότι στα προβλήματα ελέγχου υποθέσεων, συνήθως, ο ερευνητής καθορίζει το επίπεδο σημαντικότητας ώστε να ελέγχει το μέγεθος του ενδεχομένου λάθους και με βάση αυτό προσδιορίζει τον κανόνα απόρριψης της μηδενικής υπόθεσης. Σε ορισμένες όμως περιπτώσεις ξεκινάμε από τον καθορισμό του κανόνα απόρριψης και προσδιορίζουμε το επίπεδο σημαντικότητας.

Παράδειγμα: (υπολογισμού του επιπέδου σημαντικότητας α). Ας υποθέσουμε ότι θέλουμε να ελέγξουμε τη στατιστική υπόθεση,

$$H_0: \mu = 10$$

έναντι της

$$H_1: \mu > 10$$

για έναν πληθυσμό που ακολουθεί την κανονική κατανομή με μέση τιμή μ και διασπορά 4 $[X \sim N(\mu,4)]$. Ας υποθέσουμε ότι ο κανόνας απόφασης είναι τέτοιος που αν $\overline{X}<12.5$, δεν απορρίπτουμε την H_0, ενώ αν $\overline{X}\geq 12.5$, απορρίπτουμε την H_0. (Δηλαδή η τιμή 12.5 είναι το κρίσιμο σημείο). Έστω ότι παίρνουμε ένα δείγμα μεγέθους $n = 4$.
Για τον κανόνα απόφασης, όπως ορίσθηκε παραπάνω, έχουμε

$$\alpha = P(\not{H_0}|H_0) = P(\overline{X}\geq 12.5 \mid \mu = 10)$$

$$= P\left[\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \geq \frac{12.5-\mu}{\sigma/\sqrt{n}} \mid \mu = 10\right]$$

$$= P\left[Z \geq \frac{12.5-10}{2/\sqrt{4}}\right] = P(Z \geq 2.5)$$

$$= 1 - \Phi(2.5) = 1 - 0.9938 = 0.0062$$

Στην περίπτωση αυτή επίσης,

$$\beta = P(H_0 \mid H_1)$$
$$= P(\overline{X} < 12.5 \mid \mu > 10)$$

Η συνάρτηση 1-β για τις διάφορες τιμές του μ τις μεγαλύτερες του 10 θα μας δίνει την ισχύ του συγκεκριμένου ελέγχου.

Σημείωση: Όπως έχουμε ήδη αναφέρει, η συνήθης πρακτική δεν είναι να υπολογίζεται το α αλλά να καθορίζεται και στην συνέχεια να υπολογίζεται το κρίσιμο σημείο. Ο λόγος που δώσαμε το παράδειγμα αυτό είναι για να αντιληφθούμε την μεθοδολογία που ακολουθείται στα προβλήματα αυτά και η οποία είναι η ίδια για όλα τα προβλήματα.

Καθορισμός του Μεγέθους του Δείγματος με Βάση τα α και β

Σε μια διαδικασία λήψης απόφασης όπως ο έλεγχος υποθέσεων και ιδιαίτερα στην περίπτωση που αναφερόμαστε σε μονόπλευρους ελέγχους, είναι δυνατόν να καθορίσουμε το μέγεθος του δείγματος που απαιτείται ώστε να επιτευχθεί ένα ελεγχόμενο επιθυμητό επίπεδο για το α και το β (και επομένως για την ισχύ του ελέγχου).
Η διαδικασία σχεδιασμού που απαιτείται για αυτό υποθέτει τα εξής:

1. Το μέγεθος του δείγματος που καθορίζεται τελικά είναι σχετικά μεγάλο.
2. Ο πληθυσμός από τον οποίο θα προέλθει το δείγμα είτε είναι άπειρος, είτε, στην περίπτωση που είναι πεπερασμένος, είναι μεγάλος σε σχέση με το μέγεθος του δείγματος στο οποίο θα καταλήξουμε.

Η διαδικασία σχεδιασμού απαιτεί, φυσικά, τον καθορισμό των α και β. Επιπλέον, δοθέντος ότι το μέγεθος του δείγματος είναι υπό προσδιορισμό, δεν έχουμε ακόμα διαθέσιμη την δειγματική τυπική απόκλιση που χρειάζεται στους υπολογισμούς. Με δεδομένο επίσης ότι η τυπική απόκλιση σ του πληθυσμού είναι, εν γένει, άγνωστη θα πρέπει να χρησιμοποιήσουμε κάποια τιμή σχεδιασμού για το σ, ώστε να προσδιορίσουμε το μέγεθος του δείγματος.

Το απαιτούμενο μέγεθος του δείγματος, ώστε να ελέγχονται αμφότερα τα α και β για δοθείσα τιμή σ της τυπικής απόκλισης του πληθυσμού είναι

$$n = \frac{\sigma^2\left(|Z_1| + |Z_0|\right)}{|\mu_1 - \mu_0|}$$

όπου:

- σ είναι η τυπική απόκλιση του πληθυσμού (ή η τιμή σχεδιασμού της τυπικής αυτής απόκλισης)
- μ_0 και μ_1 είναι οι τιμές του μ όταν ελέγχονται, αντίστοιχα, τα α και β
- Z_0 και Z_1 είναι οι z-τιμές που αντιστοιχούν στα καθορισμένα α και β, αντίστοιχα και ορίζονται ως ακολούθως για κάθε είδος ελέγχου:

α. *Δεξιά μονόπλευρος έλεγχος* (*one-sided upper tail test*)

$$(H_0 : \mu \le \mu_0 \, , \, H_1 : \mu > \mu_0)$$

$$Z_0 = Z_{1-\alpha} \quad , \qquad Z_1 = Z_\beta$$

β. *Αριστερά μονόπλευρος έλεγχος* (*one-sided lower tail test*)

$$(H_0 : \mu \ge \mu_0 \, , \, H_1 : \mu < \mu_0)$$

$$Z_0 = Z_{\alpha} \quad , \qquad Z_1 = Z_{1-\beta}$$

γ. *Αμφίπλευρος έλεγχος (two-sided test)*

$$(H_0 : \mu = \mu_0 \, , \, H_1 : \mu \ne \mu_0)$$

$$Z_0 = Z_{1-\alpha/2} \quad , \qquad Z_1 = Z_\beta$$

Πριν προχωρήσουμε στην απόδειξη του τύπου αυτού θα δούμε πώς ο τύπος προκύπτει με ένα παράδειγμα.

Παράδειγμα: Ένας κατασκευαστής μηχανημάτων υπερήχων χρησιμοποιεί σε μια από τις μηχανές ένα εξάρτημα που θα πρέπει να έχει τη δυνατότητα να αντέχει έντονη πίεση από κραδασμούς. Στο παρελθόν, εχρησιμοποιείτο ένα μεταλλικό μηχάνημα για την δουλειά αυτή. Μακροχρόνια εμπειρία από την χρήση του μεταλλικού αυτού εξαρτήματος έχει δείξει ότι ο μέσος χρόνος ζωής του είναι 1100 ώρες. Το ερευνητικό τμήμα της εταιρείας κατασκεύασε πρόσφατα πειραματικά ένα εξάρτημα από μέταλλο και πλαστικό. Ο κατασκευαστής ενδιαφέρεται να μάθει κατά πόσο η μέση ζωή μ του νέου αυτού εξαρτήματος ξεπερνά τη μέση ζωή των 1100 ωρών του μεταλλικού εξαρτήματος. Κάτω από τα δεδομένα αυτά, θα πρέπει να ελεγχθεί η υπόθεση

$$H_0 : \mu \leq 1100$$
$$H_1 : \mu > 1100$$

Λύση: Έστω ότι ο κατασκευαστής θέλει να ελέγξει το επίπεδο σημαντικότητας στην τιμή $\alpha=0.01$ όταν $\mu = \mu_0 = 1100$. Το αντίστοιχο ποσοστιαίο σημείο z_0 στο επίπεδο α είναι

$$z_0 = z_{1-\alpha} = z_{0.99} = 2.326$$

Έστω ότι ο κατασκευαστής έχει αποφασίσει ότι όταν ο μέσος χρόνος ζωής του νέου εξαρτήματος είναι $\mu=\mu_1=1250$, ότι δηλαδή είναι σημαντικά μακρύτερος από τον μέσο χρόνο ζωής του πλήρως μεταλλικού εξαρτήματος, θα πρέπει να υπάρχει πιθανότητα μόνο 0.10 αποδοχής της H_0 με βάση τον έλεγχο αυτό. Αυτό είναι ισοδύναμο με το να απαιτείται ότι

$$P\,(H_0 \mid \mu=\mu_1 = 1250) = 0.10$$

Έστω ότι ο κατασκευαστής πιστεύει από προηγούμενη εμπειρία ότι η τιμή $\sigma=250$ ώρες είναι μια λογική τιμή σχεδιασμού για την τυπική απόκλιση του πληθυσμού.

Η απαίτηση ότι $\beta=0.10$ καθορίζει ότι το αριστερό άκρο της περιοχής αποδοχής θα είναι ίσο με 0.10 και επομένως το κρίσιμο σημείο στον άξονα των $\overline{X}$ αντιστοιχεί στην τιμή $z_0 = z_{0.10} = -1.282$ στην κλίμακα του z. Την τιμή αυτή της τυποποιημένης κανονικής που αντιστοιχεί στον έλεγχο του β συμβολίζουμε με z_1.
Επομένως

$$z_1 = z_\beta = z_{0.10} = -\,1.282$$

Προκειμένου να καθορίσουμε το μέγεθος του δείγματος, παρατηρούμε ότι η τυπική απόκλιση της δειγματικής κατανομής του $\overline{X}$ (τυπικό σφάλμα) είναι

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}} = \frac{250}{\sqrt{n}}$$

Στην συνέχεια, παρατηρούμε ότι το μέγεθος του δείγματος για το οποίο ενδιαφερόμαστε θα πρέπει να είναι τέτοιο ώστε το διάστημα πάνω στην κλίμακα του $\overline{X}$ από το σημείο μ_0 =1100 ως το μ_1 θα πρέπει να είναι ίσο με

$$|\mu_0 - \mu_1| = 2.326\ \sigma_{\bar{x}} + 1.282\ \sigma_{\bar{x}} =$$
$$= 3.608\ \sigma_{\bar{x}}$$

Επομένως,

$$150 = 3.608\ \frac{\sigma}{\sqrt{n}}$$

Λύνοντας την εξίσωση αυτή ως προς n, βρίσκουμε ότι n=36.

Από την στιγμή που το μέγεθος του δείγματος έχει καθορισθεί και τα δειγματικά αποτελέσματα συγκεντρώνονται με βάση δείγμα αυτού του μεγέθους, μπορεί να εφαρμοσθεί η διαδικασία του ελέγχου με τον τρόπο που είναι ήδη γνωστός χρησιμοποιώντας την δειγματική τυπική απόκλιση σε αντικατάσταση του σ.

Σημείωση: Στο ίδιο συμπέρασμα θα καταλήγαμε προφανώς αν χρησιμοποιούσαμε απευθείας τον τύπο για τον προσδιορισμό του δείγματος.

Πράγματι, για την περίπτωσή μας, έχουμε

$$z_0 = z_{1-\alpha} = z_{0.99} = 2.326$$

Δηλαδή

$$|z_0| = 2.326$$

$$z_1 = z_\beta = z_{0.10} = -1.282$$

Δηλαδή

$$|z_1| = 1.282$$

Επίσης

$$|\mu_1 - \mu_0| = |\ 1250 - 1100\ | = 150$$

Τέλος

$$\sigma = 250$$

και επομένως, αντικαθιστώντας στον τύπο, έχουμε

$$n = \frac{(250)^2(1.282 + 2.326)^2}{(150)^2} = 36$$

Σημείωση: Ο τύπος για το μέγεθος του δείγματος εξαρτάται, προφανώς, από τις προκαθορισμένες τιμές των α και β. Όσο μικρότερες είναι οι επιθυμητές τιμές του α και β, τόσο μεγαλύτερες είναι οι τιμές $|z_0|$ και $|z_1|$, αντίστοιχα, και συνεπώς τόσο μεγαλύτερη είναι η ζητούμενη τιμή του n. Για τον ίδιο λόγο, όσο πλησιέστερα είναι το μ_1 στο μ_0, τόσο μικρότερη είναι η απόσταση $|\mu_1 - \mu_0|$ και τόσο μεγαλύτερο είναι το απαιτούμενο μέγεθος του δείγματος. Τέλος, όσο μεγαλύτερη είναι η τυπική απόκλιση του πληθυσμού (είτε η πραγματική, αν αυτή είναι γνωστή, είτε η τιμή σχεδιασμού της, αν δεν είναι γνωστή), τόσο μεγαλύτερο είναι το ζητούμενο μέγεθος του δείγματος n ώστε να επιτευχθεί ο έλεγχος των α και β.

Απόδειξη του τύπου που αναφέρεται στις καμπύλες ισχύος: Η απόσταση του μ_0 από το μ_1 είναι $|\mu_1 - \mu_0|$. Το μήκος του διαστήματος από το μ_0 έως το κρίσιμο σημείο α είναι

$$|z_0|\sigma_{\bar{x}}$$

Αντίστοιχα, το μήκος του διαστήματος από το α έως το μ_1 είναι

$$|z_1|\sigma_{\bar{x}}$$

Είναι προφανές ότι ισχύει η ισότητα:

$$|\mu_1 - \mu_0| = |z_1|\sigma_{\bar{x}} + |z_0|\sigma_{\bar{x}} =$$
$$= (|z_1| + |z_0|)\sigma_{\bar{x}}$$

Δοθέντος ότι

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

θα έχουμε

$$|\mu_1 - \mu_0| = (|z_1| + |z_0|)\left(\frac{\sigma}{\sqrt{n}}\right)$$

Η λύση της εξίσωσης αυτής ως προς n δίνει τον τύπο που έπρεπε να αποδείξουμε.

Παράδειγμα: Στο πρόβλημα των συσκευασιών καφέ, έστω ότι θέλουμε να ελέγξουμε την υπόθεση

$$H_0 : \mu \geq 368$$

$$H_1 : \mu < 368$$

Έστω επίσης ότι θέλουμε να έχουμε 80% ισχύ, δηλαδή 0.8 πιθανότητα απόρριψης της μηδενικής υπόθεσης (των 368gr), όταν ο μέσος του πληθυσμού είναι στην πραγματικότητα $\mu=\mu_1=360$gr. Εστω επίσης ότι είμαστε διατεθειμένοι να δεχθούμε ως τιμή του α το 0.05. Ποιό είναι το μέγεθος του δείγματος που εξασφαλίζει τις τιμές αυτές των α και β με δεδομένο ότι σ=15gr;

Για το συγκεκριμένο πρόβλημα έχουμε ότι

$$|z_0| = |z_{0.05}| = 1.645$$

$$|z_1| = |z_{0.80}| = 0.84$$

Επομένως,

$$n = \frac{(15)^2(1.645+0.84)^2}{(368-360)^2}$$

Άρα,

$$n = 22$$

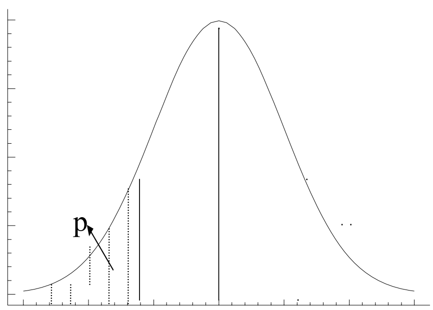
Δηλαδή, απαιτείται ένα δείγμα 22 συσκευασιών εάν μας ενδιαφέρει να έχουμε 0.05 κίνδυνο να κάνουμε λάθος τύπου Ι και 80% πιθανότητα να απορρίψουμε την μηδενική υπόθεση των 368gr και να συμπεράνουμε ότι η μέση τιμή του πληθυσμού έχει στην πραγματικότητα μετακινηθεί στην τιμή των 360gr.

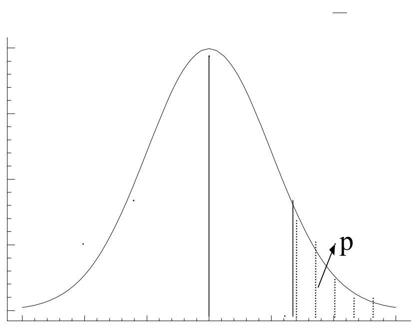
Έλεγχοι για την Μέση Τιμή με την Χρήση της p-Τιμής (Γνωστή Διακύμανση)

Για τις περιπτώσεις ελέγχου υποθέσεων που έχουμε μέχρι τώρα εξετάσει και που αναφέρονται σε κανονικό πληθυσμό με γνωστή διασπορά, οι p-τιμές εμφανίζονται στα σχήματα που ακολουθούν:

$H_0 : \mu = \mu_0$

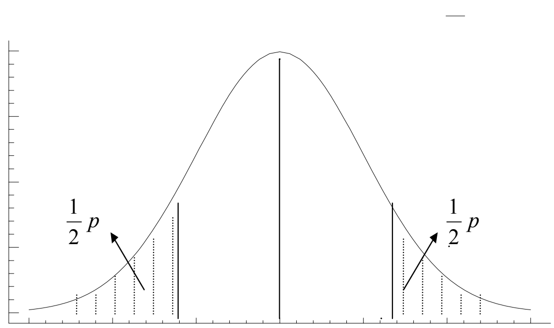
$H_1 : \mu < \mu_0$ $\qquad\qquad$ $H_1 : \mu > \mu_0$





z_0 $\qquad\qquad$ z_0

$H_1 : \mu \neq \mu_0$



$-z_0$ $\qquad\qquad$ z_0

Η τιμή z_0 που εμφανίζεται στα παραπάνω σχήματα είναι η τιμή της τυποποιημένης στατιστικής συνάρτησης ελέγχου Z_0 κάτω από την μηδενική υπόθεση. Δηλαδή, η τιμή

$$z_0 = \frac{\overline{x} - \mu_0}{\sigma / \sqrt{n}}$$

είναι η τιμή της Z_0-στατιστικής συνάρτησης

$$Z_0 = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$$

Έτσι έχουμε:

1. Για την περίπτωση $H_1 : \mu < \mu_0$

   p-τιμή = $P(Z_0 \leq z_0)$

2. Για την περίπτωση $H_1 : \mu > \mu_0$

   p-τιμή = $P(Z_0 \geq z_0)$

3. Για την περίπτωση $H_1 : \mu \neq \mu_0$

   p-τιμή = $P(|Z_0| \geq z_0)$

Παράδειγμα: Ας επανέλθουμε στο παράδειγμα της εταιρείας παραγωγής και τυποποίησης προϊόντος που ισχυρίζεται ότι κάθε πακέτο του συγκεκριμένου προϊόντος της που κυκλοφορεί στην αγορά έχει βάρος 368 gr. (H_0: μ=368).

Όπως είχαμε δει, η τιμή της Z-στατιστικής συνάρτησης ελέγχου για το συγκεκριμένο δείγμα που είχε επιλεγεί ήταν,

$$z_0 = -1.30$$

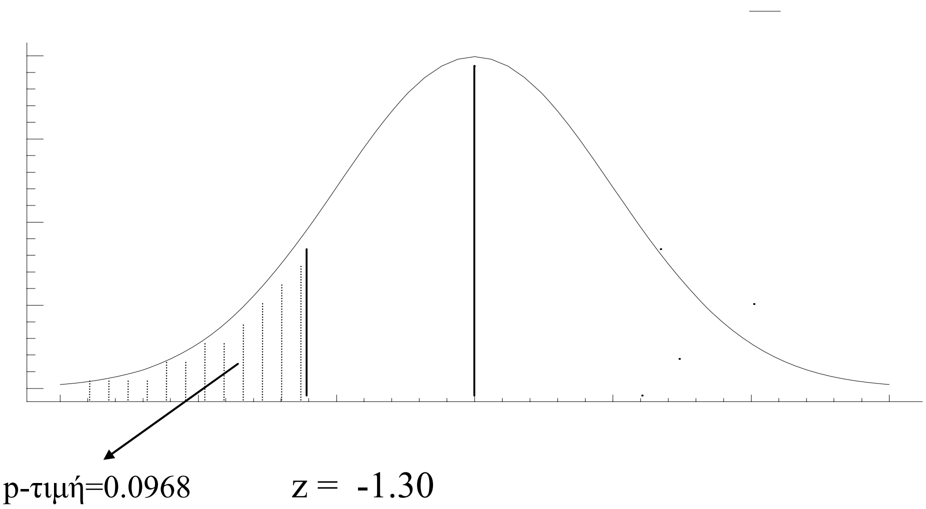
Για την περίπτωση που το πρόβλημα εξεταζόταν από τη σκοπιά της εταιρείας προστασίας καταναλωτών, (δηλαδή για την περίπτωση που H_1: μ<368), η p-τιμή (δηλαδή το παρατηρούμενο επίπεδο σημαντικότητας) θα είναι,

$$P(Z \leq -1.30) = 0.0968$$

(όπως βρίσκουμε από τους σχετικούς πίνακες της κανονικής κατανομής).

Δηλαδή, για την περίπτωση αυτή, η μηδενική υπόθεση (H_0: μ=368) θα πρέπει να απορριφθεί για οποιαδήποτε τιμή του επιπέδου σημαντικότητας α μεγαλύτερη από το 0.0968.

Αυτό γιατί, όπως φαίνεται από το σχήμα που ακολουθεί,



για οποιοδήποτε α ≥ 0.0968, η τιμή της στατιστικής συνάρτησης ελέγχου που πήραμε για το συγκεκριμένο δείγμα θα βρίσκεται στην περιοχή απόρριψης του στατιστικού ελέγχου.

Είναι προφανές ότι το επίπεδο σημαντικότητας α=0.05 που είχαμε επιλέξει είναι μικρότερο από την p-τιμή ($\alpha < p$) και γι' αυτό εκεί δεν είχαμε απορρίψει την μηδενική υπόθεση.

Σημείωση: Αν η τιμή της Z_0-στατιστικής συνάρτησης στο πρόβλημα ήταν $z_0 = -2$, αυτό θα αποτελούσε ισχυρότερη ένδειξη για το ότι οι παρατηρήσεις (το δείγμα) δεν συνάδουν με την μηδενική υπόθεση. (2 τυπικές αποκλίσεις μακριά από το μ_0 (368gr) είναι περισσότερο έντονη ένδειξη από 1.30 αποκλίσεις από το μ_0). <u>Το εμβαδόν αριστερά από το -1.30 εκφράζει τα δείγματα που δίνουν περισσότερο ακραίες z_0-τιμές από αυτήν που δίνει το συγκεκριμένο δείγμα</u> που παρατηρήσαμε (και έδωσε $z_0 = -1.30$) και τα οποία περιέχουν ισχυρότερες ενδείξεις κατά της μηδενικής υπόθεσης.

Η διαπίστωση αυτή εξηγεί και τον ορισμό της p-τιμής που δόθηκε στο προηγούμενο κεφάλαιο ως της πιθανότητας να παρατηρήσουμε μια στατιστική συνάρτηση ελέγχου τόσο ακραία, ή περισσότερο ακραία, από αυτήν που παρατηρήθηκε. Το “ακραία” αναφέρεται σε σχέση με την τιμή της παραμέτρου κάτω από την μηδενική υπόθεση. Αυτό γιατί η πιθανότητα αυτή υπολογίζεται κάτω από την παραδοχή ότι ισχύει η μηδενική υπόθεση. Όσο μικρότερη είναι η πιθανότητα αυτή τόσο ισχυρότερες είναι οι ενδείξεις εναντίον της μηδενικής υπόθεσης.

Σημείωση: Δοθέντος ότι η Z_0-στατιστική συνάρτηση ελέγχου εξαρτάται από τα δεδομένα, το ίδιο ισχύει για την p-τιμή. Αυτό εξηγεί και τον εναλλακτικό όρο "παρατηρούμενο επίπεδο σημαντικότητας" για την p-τιμή.

Παρατήρηση: Τα προαναφερθέντα εξηγούν ίσως καλύτερα και την λογική της Z_0-στατιστικής συνάρτησης. Η επιχειρηματολογία αυτή στηρίζεται στην λογική της αντίφασης και χρησιμοποιείται για να δείξει ότι η αποδοχή της μηδενικής υπόθεσης θα οδηγήσει σε ένα παράλογο συμπέρασμα και επομένως θα πρέπει να απορριφθεί. Από άποψη μεθοδολογίας, αυτό σημαίνει ότι παρατηρούμε τα δεδομένα, υπολογίζουμε την στατιστική συνάρτηση ελέγχου και το παρατηρούμενο επίπεδο σημαντικότητας (την p-τιμή).

Ας θεωρήσουμε, για παράδειγμα, ένα πείραμα που οδηγεί σε μια p-τιμή ίση με 0.001 (1 στα 1000). Για να εξηγήσουμε τον αριθμό αυτό, ξεκινάμε υποθέτοντας ότι η μηδενική υπόθεση είναι σωστή. Στην συνέχεια, ας φαντασθούμε πολλούς άλλους ερευνητές που επαναλαμβάνουν το πείραμα αυτό. Αυτό που λέει το 1 στα 1000 (0.001) είναι ότι η στατιστική συνάρτηση ελέγχου είναι "πολύ μακριά" από αυτό που η μηδενική υπόθεση ισχυρίζεται. Μόνο 1 στα 1000 πειράματα θα έδινε στατιστική συνάρτηση ελέγχου τόσο ακραία, ή περισσότερο ακραία, από αυτήν που υπολογίσαμε με βάση τις παρατηρήσεις μας (το δείγμα μας). Δηλαδή, μόνο 1 στα 1000 πειράματα θα έδινε τιμή της στατιστικής συνάρτησης ελέγχου τόσο μακριά, ή περισσότερο μακριά, από την τιμή της παραμέτρου κάτω από την μηδενική υπόθεση, από αυτήν που εμείς υπολογίσαμε με βάση το δείγμα μας. Η μηδενική υπόθεση επομένως, στην περίπτωση αυτή, οδηγεί σε παραλογισμούς και θα πρέπει να απορριφθεί.

Γενικά, όσο μικρότερη είναι η p-τιμή τόσο ισχυρότερες είναι οι ενδείξεις εναντίον της μηδενικής υπόθεσης και επομένως συνηγορούν στην απόρριψή της. Η διατύπωση "απόρριψη της μηδενικής" δίνει έμφαση στο ότι στον έλεγχο σημαντικότητας η επιχειρηματολογία στηρίζεται στην αντίφαση.

Σημείωση: Για την αποφυγή παρανοήσεων, θα πρέπει να τονισθεί ότι η p-τιμή <u>δεν είναι</u> η πιθανότητα ότι η μηδενική υπόθεση είναι σωστή. Εκφράζει την πιθανότητα να οδηγηθούμε σε "μεγάλη" τιμή της στατιστικής συνάρτησης ελέγχου κάτω από την παραδοχή ότι η μηδενική υπόθεση είναι σωστή. Στην κλασσική στατιστική θεωρία δεν υπάρχει τρόπος να προσδιορίσουμε την πιθανότητα να είναι σωστή η μηδενική υπόθεση. Η μηδενική υπόθεση θα είναι ή πάντα σωστή ή πάντα λάθος. Αυτό που παρέχει η p-τιμή είναι η πιθανότητα να βρούμε ενδείξεις αντίθετες με την μηδενική υπόθεση τόσο ισχυρές, ή και ακόμα περισσότερο ισχυρές από αυτές που έχουμε διαθέσιμες, αν η μηδενική υπόθεση ίσχυε.

Αν αντιμετωπίζαμε το πρόβλημα από την σκοπιά της εταιρείας παραγωγής του προϊόντος (H_1: $\mu \neq 368$), η p-τιμή θα ήταν,

$$
\begin{aligned}
\text{p-τιμή} &= P(Z \leq -1.30) + P(Z \geq 1.30) \\
&= 2\,(0.0968) \\
&= 0.1936
\end{aligned}
$$

Επομένως, για οποιαδήποτε τιμή του επιπέδου σημαντικότητας α μεγαλύτερη του 0.1936, η μηδενική υπόθεση (με βάση το συγκεκριμένο δείγμα) θα πρέπει να απορριφθεί.
Για το επίπεδο σημαντικότητας α=0.05 που είχαμε επιλέξει, παρατηρούμε και πάλι ότι είναι σημαντικά μικρότερο από την p-τιμή και επομένως, για το επίπεδο αυτό σημαντικότητας και με βάση τις πληροφορίες του συγκεκριμένου δείγματος, δεν έχουμε ισχυρές ενδείξεις για να απορρίψουμε την μηδενική υπόθεση.

Παρατήρηση: Όσα αναφέρθηκαν μέχρι τώρα για τον ορισμό του επιπέδου σημαντικότητας α και της p-τιμής μπορούν να εφαρμοσθούν αν η μηδενική υπόθεση είναι απλή (H_0: $\mu = \mu_0$). Αυτό προκύπτει από τους αντίστοιχους ορισμούς των δύο εννοιών.

Αν η μηδενική υπόθεση είναι σύνθετη (H_0: $\mu \geq \mu_0$ ή H_1: $\mu \leq \mu_0$), τότε η τιμή του α και η p-τιμή ορίζονται ως οι μέγιστες πιθανότητες των ενδεχομένων για τα οποία αυτές ορίσθηκαν στην περίπτωση απλής μηδενικής υπόθεσης.

Έτσι, συγκεκριμένα αν

$$H_0 : \mu \geq \mu_0$$
$$H_1 : \mu < \mu_0$$

ορίζουμε ως

$$\begin{aligned} \alpha &= \max P(\cancel{H_0} \mid H_0) \\ &= \max P(\cancel{H_0} \mid \mu \geq \mu_0) \\ &= P(\cancel{H_0} \mid \mu = \mu_0) \end{aligned}$$

αντίστοιχα,

$$\begin{aligned} \text{p-τιμή} &= \max P\left(\overline{X} \leq \overline{x} \mid H_0\right) \\ &= \max P\left(\overline{X} \leq \overline{x} \mid \mu \geq \mu_0\right) \\ &= P\left(\overline{X} \leq \overline{x} \mid \mu = \mu_0\right) \\ &= P(Z_0 \leq z_0) \end{aligned}$$

Οι υπολογισμοί γίνονται όπως προηγουμένως.

Μετά την αναλυτική παρουσίαση και τα γενικά σχόλια για τους ελέγχους υποθέσεων που αναφέρονται σε κανονικούς πληθυσμούς με γνωστή διακύμανση, θα προχωρήσουμε στην ανάπτυξη των στατιστικών ελέγχων υποθέσεων που αναφέρονται σε πληθυσμούς με άγνωστη διακύμανση, σε ελέγχους που αναφέρονται σε αναλογίες, σε διασπορές, όπως επίσης και σε συγκρίσεις μέσων τιμών, αναλογιών και διασπορών.

Β. Περίπτωση Αγνώστων Διακυμάνσεων

Στα περισσότερα πρακτικά προβλήματα, η διακύμανση σ^2 (ή αντίστοιχα η τυπική απόκλιση σ) του πληθυσμού είναι άγνωστη. Η εκτιμήτρια της διακύμανσης είναι η

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

Όπως έχουμε δει, το S^2 δεν είναι αμερόληπτη εκτιμήτρια του σ^2. Αμερόληπτη εκτιμήτρια του σ^2 είναι η

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

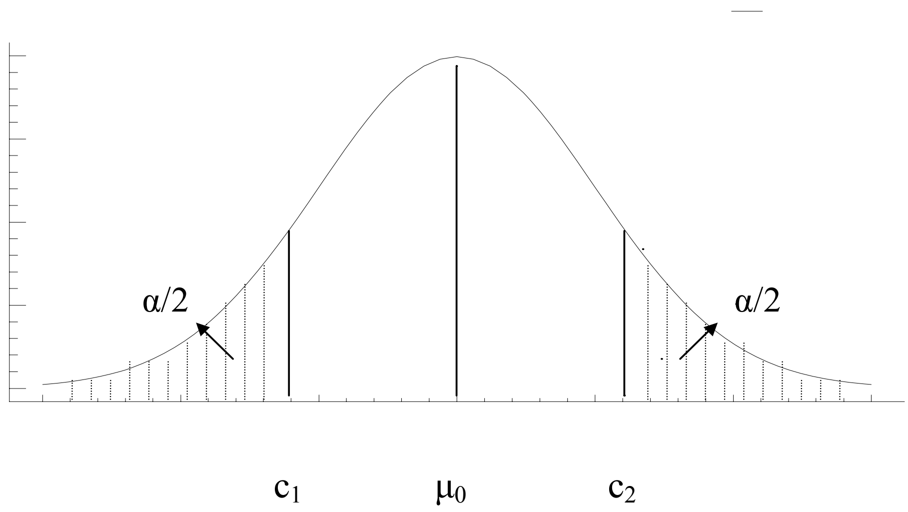
όπου και στις δύο περιπτώσεις n είναι το μέγεθος του δείγματος.
Όπως γνωρίζουμε στη περίπτωση αυτή, με την προϋπόθεση ότι ο πληθυσμός είναι κανονικός $X \sim N(\mu, \sigma^2)$, η στατιστική συνάρτηση

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n-1}} = \frac{\overline{X} - \mu}{S^*/\sqrt{n}} \sim t_{n-1}$$

Η συνάρτηση αυτή θα χρησιμοποιηθεί και για έλεγχο υποθέσεων που αναφέρονται σε κανονικούς πληθυσμούς με άγνωστη διακύμανση.
Έστω ότι θέλουμε να ελέγξουμε την υπόθεση

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu \neq \mu_0$$

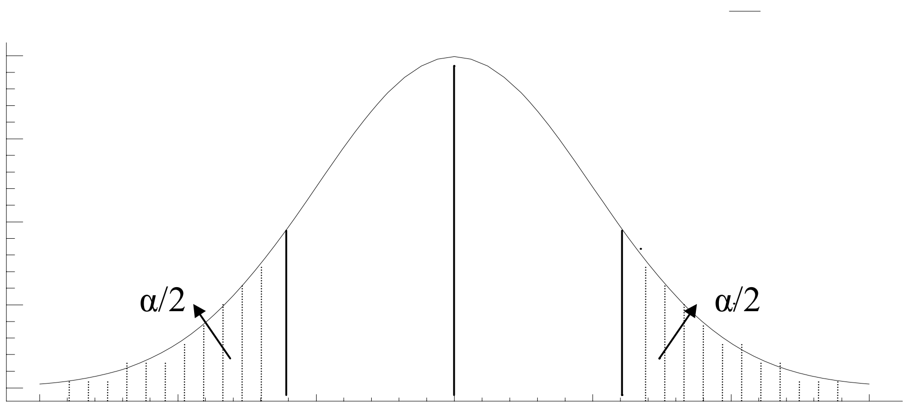
για την μέση τιμή ενός κανονικού πληθυσμού του οποίου η διακύμανση είναι άγνωστη σε επίπεδο σημαντικότητας α. Σύμφωνα με τα όσα έχουμε ήδη εκθέσει, η κατάλληλη στατιστική συνάρτηση ελέγχου είναι η $\overline{X}$. Η τυποποιημένη τιμή της ελεγχοσυνάρτησης αυτής κάτω από την H_0 συμβολίζεται με T_0 και ονομάζεται *T_0 συνάρτηση ελέγχου* (ή, *T_0 ελεγχοσυνάρτηση*). Στην συνέχεια, χρειάζεται να καθορίσουμε τις κρίσιμες τιμές c_1 και c_2 που θα χρησιμοποιηθούν για τον έλεγχο. (Εκείνα τα σημεία που τιμές της στατιστικής συνάρτησης μεγαλύτερές τους θα οδηγούν στην απόρριψη της μηδενικής υπόθεσης. Τα σημεία δηλαδή που χωρίζουν την περιοχή απόρριψης από την περιοχή αποδοχής).



Ακολουθώντας τη λογική που έχουμε ήδη χρησιμοποιήσει, έχουμε

$$
\begin{aligned}
\alpha &= P\left(\cancel{H_0} \mid H_0\right) = P(\overline{X} < c_1 \;\; ή \;\; \overline{X} > c_2 \mid \mu = \mu_0) \\
&= P\left(\frac{\overline{X}-\mu}{S^*/\sqrt{n}} < \frac{c_1-\mu}{S^*/\sqrt{n}} \; ή \; \frac{\overline{X}-\mu}{S^*/\sqrt{n}} > \frac{c_2-\mu}{S^*/\sqrt{n}} \,\middle|\, -\mu = \mu_0 \right) \\
&= P\left(\frac{\overline{X}-\mu_0}{S^*/\sqrt{n}} < \frac{c_1-\mu_0}{S^*/\sqrt{n}} \; ή \; \frac{\overline{X}-\mu_0}{S^*/\sqrt{n}} > \frac{c_2-\mu_0}{S^*/\sqrt{n}} \right) \\
&= P\left(T_0 < \frac{c_1-\mu_0}{S^*/\sqrt{n}} \; ή \; T_0 > \frac{c_2-\mu_0}{S^*/\sqrt{n}} \right)
\end{aligned}
$$

Δοθέντος ότι η περιοχή απόρριψης έχει συνολικό εμβαδόν α θα έχουμε, όπως προκύπτει από το σχήμα που ακολουθεί,



$-t_{n-1,\,1-\alpha/2}$ $\qquad\qquad$ $t_{n-1,\,1-\alpha/2}$

$= t_{n-1,\,\alpha/2}$

$$\frac{c_1-\mu_0}{S^*/\sqrt{n}} = -t_{n-1,\,1-\alpha/2}\,, \qquad \frac{c_2-\mu_0}{S^*/\sqrt{n}} = t_{n-1,\,1-\alpha/2}$$

Λύνοντας ως προς c_1 και c_2, θα έχουμε,

$$c_1 = \mu_0 - t_{n-1,\,1-\frac{\alpha}{2}}\frac{S^*}{\sqrt{n}}$$

$$c_2 = \mu_0 + t_{n-1,\,1-\frac{\alpha}{2}}\frac{S^*}{\sqrt{n}}$$

Επομένως, θα απορρίπτουμε την μηδενική υπόθεση σε επίπεδο σημαντικότητας α αν,

$$\overline{X} < \mu_0 - t_{n-1,\,1-\frac{\alpha}{2}}\frac{S^*}{\sqrt{n}} \quad \text{ή} \quad \overline{X} > \mu_0 + t_{n-1,\,1-\frac{\alpha}{2}}\frac{S^*}{\sqrt{n}}$$

Ισοδύναμα, χρησιμοποιώντας την τυποποιημένη τιμή της στατιστικής συνάρτησης ελέγχου, θα απορρίπτουμε την μηδενική υπόθεση αν,

$$T_0 < -t_{n-1,\,1-\alpha/2} \quad \text{ή} \quad T_0 > t_{n-1,\,1-\alpha/2}$$

Σημείωση: Η διαδικασία αυτή ισχύει με την προϋπόθεση ότι ο πληθυσμός είναι κανονικός. Μπορούμε να την εφαρμόσουμε και στην περίπτωση που ο πληθυσμός δεν είναι κανονικός με την προϋπόθεση ότι το μέγεθος του δείγματος είναι αρκετά μεγάλο ($n \geq 30$). Αυτό γιατί σύμφωνα με το κεντρικό οριακό θεώρημα, και στην περίπτωση αυτή η στατιστική συνάρτηση Τ ακολουθεί την t_{n-1}.

Παρατήρηση: Χρειάζεται προσοχή στην χρησιμοποίηση της στατιστικής συνάρτησης Τ. Αν για την εκτίμηση του σ^2 έχει χρησιμοποιηθεί ο τύπος της αμερόληπτης εκτιμήτριας (S^{*2}), τότε προχωράμε όπως παραπάνω. Αν όμως έχει χρησιμοποιηθεί η S^2, στον τύπο της στατιστικής συνάρτησης ελέγχου, όπου χρησιμοποιείται το S, η διαίρεση θα γίνεται με το $\sqrt{n-1}$ (και όχι με το $\sqrt{n}$).

Παράδειγμα: Στο παράδειγμα της εταιρείας παραγωγής τυποποιημένων προϊόντων, ας υποθέσουμε ότι ένας νέος διευθυντής του τμήματος ελέγχου ποιότητας αρνείται να θεωρήσει ως δεδομένο ότι η τυπική απόκλιση του βάρους των προϊόντων που περιέχονται στη συγκεκριμένη συσκευασία είναι 15gr. Αποφασίζει τότε να χρησιμοποιήσει την τυπική απόκλιση του δείγματος την οποία υπολογίζει ότι είναι s=17.3gr προκειμένου να ελέγξει την υπόθεση

$$H_0 : \mu = 368 \text{ gr}$$
$$H_1 : \mu \neq 368 \text{ gr}$$

σε επίπεδο σημαντικότητας α = 0.05. Υπολογίζει την τιμή της στατιστικής συνάρτησης ελέγχου, η οποία για n = 25 είναι,

$$t_0 = \frac{364.1 - 368}{17.3 / \sqrt{25}} = -1.127$$

Δοθέντος ότι, όπως προκύπτει από τους πίνακες,

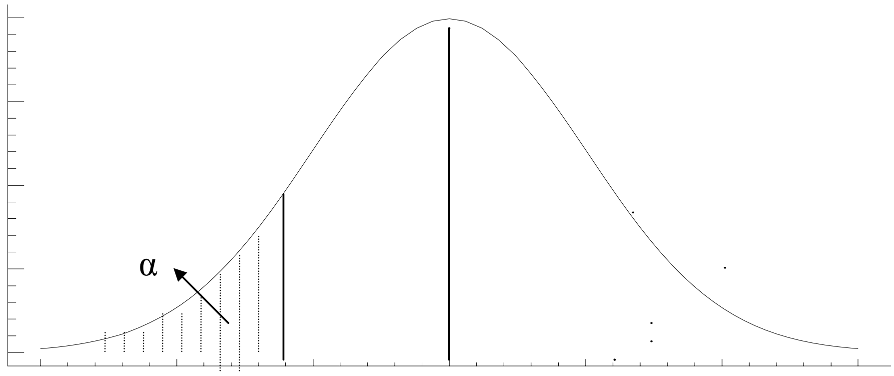
$$t_{24,.975} = 2.0639$$

βλέπουμε ότι η τιμή της στατιστικής συνάρτησης ελέγχου για το συγκεκριμένο δείγμα δεν βρίσκεται στην περιοχή απόρριψης. Δεν έχουμε επομένως ισχυρές ενδείξεις για να απορρίψουμε την μηδενική υπόθεση σε επίπεδο σημαντικότητας α=0.05.

Χρησιμοποιώντας την ίδια λογική κατασκευάζουμε ελέγχους υποθέσεων για την περίπτωση κανονικού πληθυσμού με άγνωστη διασπορά όταν η εναλλακτική υπόθεση είναι μονόπλευρη. Έτσι, στην περίπτωση της υπόθεσης,

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu < \mu_0$$

απορρίπτουμε την μηδενική υπόθεση σε επίπεδο σημαντικότητας α αν



$-t_{n-1,\,1-\alpha}$

$$\overline{X} < \mu_0 - t_{n-1,\,1-\alpha}\,\frac{S^*}{\sqrt{n}}$$

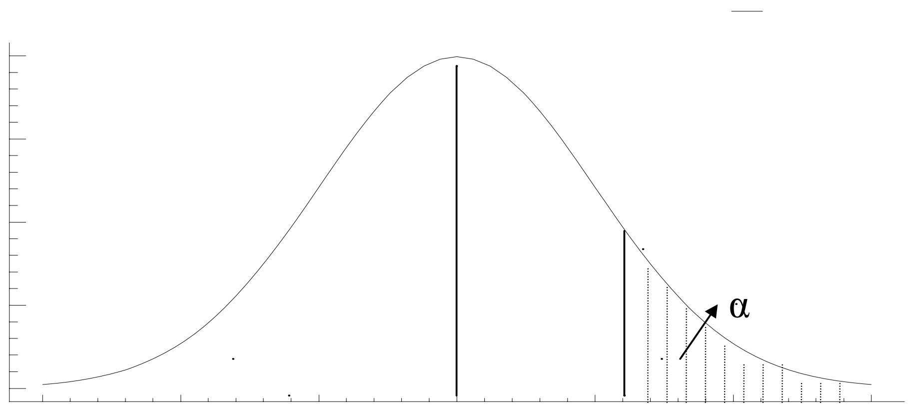
ή, ισοδύναμα, αν

$$T_0 < -t_{n-1,\,1-\alpha}$$

Επίσης, προκειμένου να ελέγξουμε την υπόθεση

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu > \mu_0$$

σε επίπεδο σημαντικότητας α, θα απορρίπτουμε την μηδενική υπόθεση αν



$$\overline{X} > \mu_0 + t_{n-1,\,1-\alpha}\,\frac{S^*}{\sqrt{n}}$$

ή, ισοδύναμα, αν

$$T_0 > t_{n-1,\,1-\alpha}$$

Παράδειγμα: Ο ιδιοκτήτης ενός ορυχείου ενδιαφέρεται να αξιολογήσει μια νέα μέθοδο παραγωγής συνθετικών διαμαντιών. Η μελέτη του κόστους που συνεπάγεται η διαδικασία κατασκευής, έχει καταλήξει στο συμπέρασμα ότι για να είναι επικερδής η νέα αυτή μέθοδος, θα πρέπει το μέσο βάρος των συνθετικών διαμαντιών να είναι περισσότερο από 0.5 καράτια. Προκειμένου να αξιολογηθεί η διαδικασία κατασκευής, επιλέγεται δείγμα από 6 συνθετικά διαμάντια που έχουν κατασκευασθεί με τη νέα μέθοδο κατασκευής. Το βάρος τους βρίσκεται ότι είναι: 0.46, 0.61, 0.52, 0.48, 0.57 και 0.54 καράτια αντίστοιχα.
Να καθορισθεί σε επίπεδο σημαντικότητας α=0.05 με βάση τις πληροφορίες από το δείγμα αυτό αν η νέα μέθοδος είναι επικερδής.

Λύση: Θεωρώντας ότι το βάρος των συνθετικών διαμαντιών ακολουθεί την κανονική κατανομή, θα πρέπει να ελέγξουμε την υπόθεση:

$$H_0 : \mu = 0.05$$
$$H_1 : \mu > 0.05$$

Από τα στοιχεία του δείγματος βρίσκουμε ότι:

$$\overline{x} = 0.53 \text{ και } s^* = 0.0559$$

Επομένως, η τιμή της στατιστικής συνάρτησης ελέγχου κάτω από τη μηδενική υπόθεση είναι:

$$t_0 = \frac{\overline{x} - \mu_0}{s^*/\sqrt{n}} = \frac{0.53 - 0.5}{0.0559/\sqrt{6}} = 1.31$$

Απο τους πίνακες της κατανομής t έχουμε ότι,

$$t_{5,.95} = 2.015$$

Με βάση τα στοιχεία αυτά, παρατηρούμε ότι η τιμή της στατιστικής συνάρτησης ελέγχου δεν βρίσκεται στη κρίσιμη περιοχή και επομένως δεν έχουμε ισχυρές ενδείξεις που να οδηγούν στην απόρριψη της μηδενικής υπόθεσης σε επίπεδο σημαντικότητας 0.05. Επομένως, με βάση τα στοιχεία του δείγματος δεν μπορούμε να ισχυρισθούμε ότι η νέα μέθοδος είναι κερδοφόρα.

Λύση με το πακέτο SPSS

Ο έλεγχος υποθέσεων για την μέση τιμή κανονικού πληθυσμού με άγνωστη διακύμανση με το στατιστικό πακέτο SPSS γίνεται ως εξής:

- Εισάγουμε τα δεδομένα σε κάποια μεταβλητή π.χ. VAR00001
- Από την επιλογή Statistics επιλέγουμε Compare means.
- Επιλέγουμε One sample T test
- Στο παράθυρο που εμφανίζεται επιλέγουμε τη μεταβλητή μας (VAR00001) και την τοποθετούμε στο πεδίο test variable
- Στο πεδίο Test value γράφουμε την τιμή για την οποία επιθυμούμε να ελέγξουμε αν ταυτίζεται με τον μέσο του πληθυσμού (δηλ. για το συγκεκριμένο παράδειγμα την τιμή 0.5) και πατάμε OK.

Τα αποτελέσματα που δίνει το SPSS είναι τα ακόλουθα

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
VAR00001	6	.5300	5.586E-02	2.280E-02

One-Sample Test

	Test Value = 0.5					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
VAR00001	1.316	5	.245	3.000E-02	-2.86E-02	8.862E-02

Στον πρώτο πίνακα αποτελεσμάτων, φαίνεται το όνομα της μεταβλητής (VAR00001), ο αριθμός των παρατηρήσεων (N), ο δειγματικός μέσος (Mean), η τυπική απόκλιση (Std. Deviation) και το τυπικό σφάλμα του μέσου (Std. Error Mean).

Στον δεύτερο πίνακα αποτελεσμάτων, δίνεται το όνομα της μεταβλητής, η τιμή της στατιστικής συνάρτησης ελέγχου (t), οι βαθμοί ελευθερίας (df), η p-τιμή για τον αμφίπλευρο έλεγχο (Sig. (2-tailed)), η τιμή της διαφοράς του δειγματικού μέσου από την τιμή που ελέγχουμε, δηλ. $\overline{X}-\mu_0$ (Mean Difference), καθώς και ένα 95% διάστημα εμπιστοσύνης για την τιμή του μέσου (95% Confidence Interval of the Difference).

Σημείωση: Το πακέτο SPSS δεν δίνει την p-τιμή για μονόπλευρους ελέγχους, καθώς, όπως προαναφέραμε, η p-τιμή που δίνεται στον πίνακα με τα αποτελέσματα είναι η p-τιμή του αμφίπλευρου ελέγχου. Επειδή στο παράδειγμα μας ο έλεγχος είναι μονόπλευρος, η p-τιμή είναι εκείνη που δίνεται στον πίνακα διαιρεμένη με το δύο δηλ. 0.245/2=0.1225.

<u>Λύση με το πακέτο Minitab</u>

Ο έλεγχος υποθέσεων για την μέση τιμή κανονικού πληθυσμού με άγνωστη διακύμανση με το στατιστικό πακέτο MINITAB γίνεται ως εξής:

- Στο παράθυρο Data, εισάγουμε τα δεδομένα σε κάποια μεταβλητή π.χ. C1.
- Από την επιλογή Stat, επιλέγουμε Basic Statistics.
- Επιλέγουμε 1-Sample t…

- Στο παράθυρο που ανοίγει, επιλέγουμε την μεταβλητή στην οποία έχουμε περάσει τα δεδομένα και την τοποθετούμε στο πεδίο Variables.
- Επιλέγουμε Test mean και στο πεδίο σχετικό πεδίο γράφουμε την τιμή για την οποία επιθυμούμε να ελέγξουμε αν ταυτίζεται με την μέση τιμή του πληθυσμού (δηλ. για το συγκεκριμένο παράδειγμα την τιμή 0.5)
- Στο πεδίο Alternative, δηλώνουμε την μορφή της εναλλακτικής υπόθεσης. Όταν αυτή είναι αμφίπλευρη (δηλ. $\mu \neq \mu_0$), επιλέγουμε not equal. Αντίθετα, όταν έχει την μορφή $\mu > \mu_0$, επιλέγουμε greater than, ενώ τέλος όταν έχει την μορφή $\mu < \mu_0$, επιλέγουμε less than. Για το παράδειγμα μας, επιλέγουμε greater than και στην συνέχεια πατάμε OK.

Το Minitab δίνει τα παρακάτω αποτελέσματα:

TEST OF MU = 0.5000 VS MU G.T. 0.5000

	N	MEAN	STDEV	SE MEAN	T	P VALUE
C1	6	0.5300	0.0559	0.0228	1.32	0.12

Στην πρώτη γραμμή του παραπάνω output, περιγράφεται ο ζητούμενος έλεγχος. Επίσης στον πίνακα δίνεται η μεταβλητή με τα δεδομένα (C1), ο αριθμός των παρατηρήσεων (N), ο δειγματικός μέσος (Mean), η δειγματική τυπική απόκλιση (Stdev), το τυπικό σφάλμα του μέσου (SE Mean), η τιμή της στατιστικής συνάρτησης ελέγχου (t) και η p-τιμή (p-value).

Έλεγχοι για την Μέση Τιμή με την Χρήση της p-Τιμής (Άγνωστη Διακύμανση)

Και στην περίπτωση που η διακύμανση του κανονικού πληθυσμού είναι άγνωστη, ισχύει ο ορισμός της p-τιμής (παρατηρούμενου επίπεδου σημαντικότητας) τον ορισμό της οποίας δώσαμε στην προηγούμενη ενότητα.

Έτσι, στο τελευταίο παράδειγμα που αναφέρεται στην παραγωγή των συνθετικών διαμαντιών, είχαμε υπολογίσει ότι, για την μηδενική και την εναλλακτική υπόθεση, έτσι όπως ορίσθηκαν, η τιμή της στατιστικής συνάρτησης ελέγχου ήταν $t_0 = 1.31$. Η p-τιμή γι' αυτό τον έλεγχο υποθέσεων, με βάση της πληροφορίες του συγκεκριμένου δείγματος, θα είναι:

$$p = P(t_5 > 1.31 \mid H_0) = P(t_5 > 1.31 \mid \mu = 0.5)$$

Το παράδειγμα αυτό εξηγεί τον λόγο για τον οποίο ο υπολογισμός της p-τιμής σε ένα έλεγχο υποθέσεων που χρησιμοποιείται μια t στατιστική συνάρτηση ελέγχου δεν είναι εύκολο να καθορισθεί με ακρίβεια. Αυτό γιατί, όπως προκύπτει από την μελέτη των πινάκων της κατανομής t, οι πίνακες αυτοί δεν δίνουν την ακριβή τιμή της πιθανότητας για όλες τις διαφορετικές δυνατές τιμές του t. Από τους πίνακες, παρατηρούμε ότι η τιμή αυτή είναι μεταξύ του 0.25 (που αντιστοιχεί στο σημείο $t_5 = 0.7267$) και 0.1 (που αντιστοιχεί στο σημείο $t_5 = 1.4759$).

Η δυσκολία στον ακριβή καθορισμό της p-τιμής ήταν αυτή που εμπόδιζε τους ερευνητές να την χρησιμοποιούν αντί της συνήθους διαδικασίας του εκ των προτέρων καθορισμού του επιπέδου σημαντικότητας α. Η χρήση όμως διαφόρων στατιστικών πακέτων μεγάλων αλλά και προσωπικών υπολογιστών έχει ξεπεράσει πλήρως αυτό το πρόβλημα. Έτσι, όλα τα στατιστικά πακέτα δίνουν, σε κάθε περίπτωση, και την p-τιμή (το παρατηρούμενο επίπεδο σημαντικότητας) που αντιστοιχεί σε κάθε έλεγχο υπόθεσης.

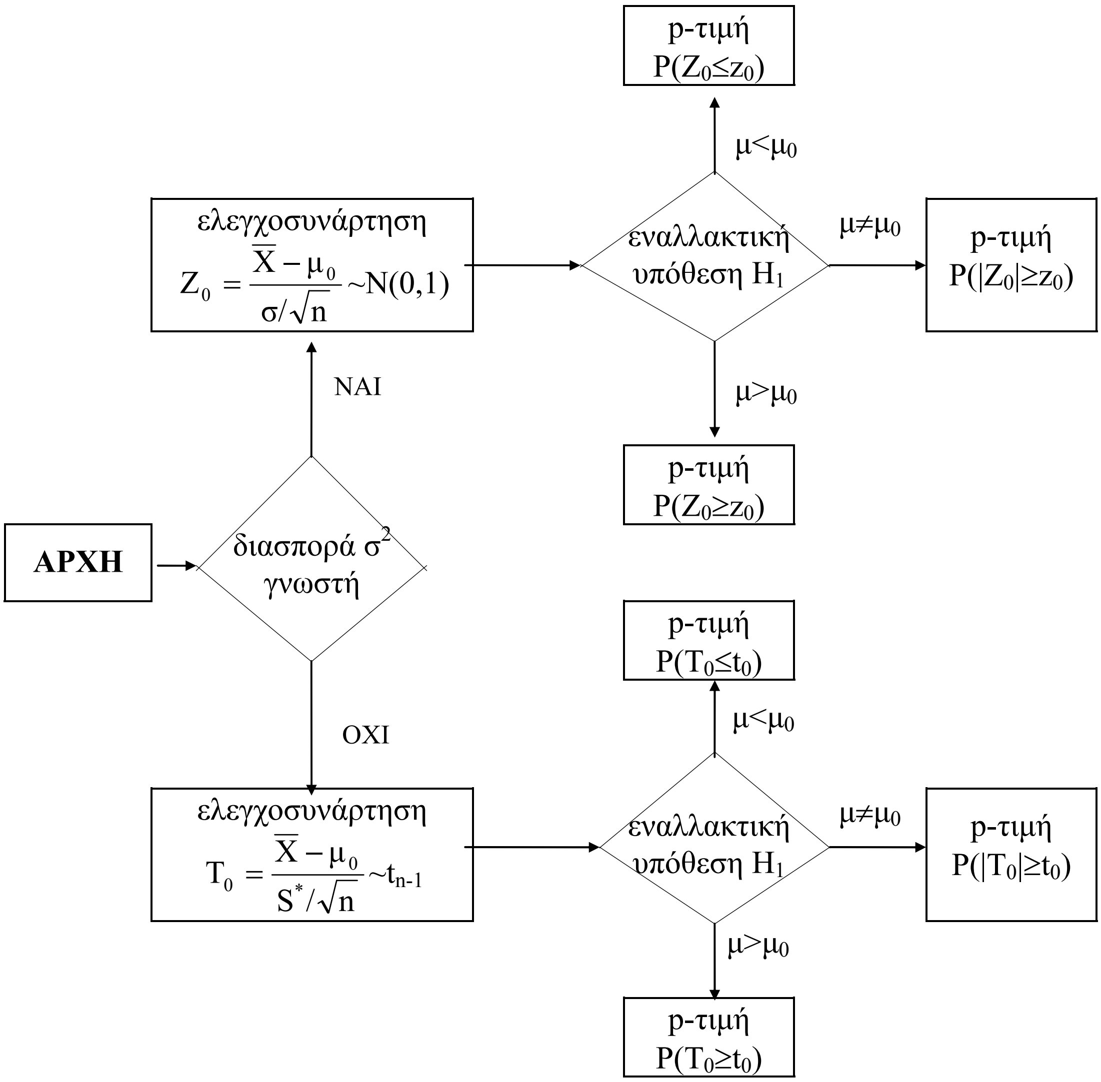
Για παράδειγμα, στο πρόβλημά μας το στατιστικό πακέτο Statgraphics δίνει τιμή p = 0.2454. Επομένως, μόνο για επίπεδο σημαντικότητας μεγαλύτερο από 0.2454 θα απορρίπταμε την μηδενική υπόθεση του προβλήματος με βάση τα στοιχεία του συγκεκριμένου δείγματος.

Σημείωση: Για την περίπτωση που η μηδενική υπόθεση είναι σύνθετη ($H_0 : \mu \geq \mu_0$ ή $H_0 : \mu \leq \mu_0$), το επίπεδο σημαντικότητας α και η p-τιμή ορίζονται με τον ίδιο τρόπο όπως στην προηγούμενη ενότητα.

Σχηματική Παρουσίαση της Διαδικασίας Ελέγχου Υποθέσεων για την μέση Τιμή Κανονικού Πληθυσμού

Η διαδικασία που ακολουθείται για τον έλεγχο υποθέσεων για την μέση τιμή κανονικού πληθυσμού και οι ενέργειες που πρέπει να γίνουν, ανάλογα με το εάν η διακύμανση είναι γνωστή ή όχι μπορούν να παρουσιασθούν με το διάγραμμα που ακολουθεί.

Έλεγχος Υπόθεσης H_0: $\mu=\mu_0$
(Με την χρήση της p-τιμής)



ΕΛΕΓΧΟΙ ΥΠΟΘΕΣΕΩΝ ΓΙΑ ΑΝΑΛΟΓΙΕΣ

Για ελέγχους υποθέσεων που αναφέρονται σε αναλογίες πληθυσμών, είναι δυνατόν να χρησιμοποιήσουμε τις ιδιότητες της κανονικής κατανομής δεδομένου ότι, όπως έχουμε δει στα διαστήματα εμπιστοσύνης, λόγω της προσέγγισης της διωνυμικής κατανομής από την κανονική κατανομή, ισχύει ότι

$$X \underset{appr.}{\sim} N(np, npq)$$

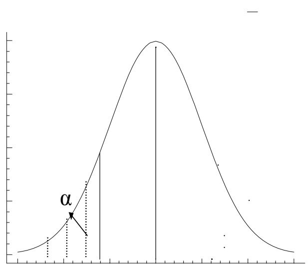
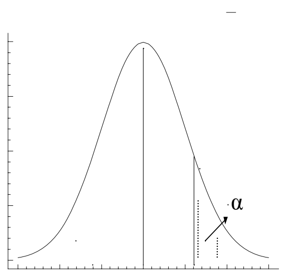
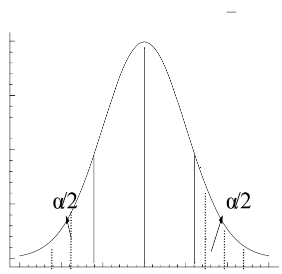
όπου X ο αριθμός των επιτυχιών σε μία ακολουθία δοκιμών Bernoulli με πιθανότητα επιτυχίας p και επομένως,

$$Z = \frac{X - np}{\sqrt{npq}} = \frac{\frac{X}{n} - p}{\sqrt{\frac{pq}{n}}} \underset{appr.}{\sim} N(0, 1)$$

Δηλαδή, η εκτιμήτρια $\hat{p} = \frac{X}{n}$ του ποσοστού επιτυχιών στον πληθυσμό που προσδιορίζεται από το ποσοστό επιτυχιών στο δείγμα μπορεί να προσεγγισθεί από την τυπική κανονική κατανομή.

Επομένως, όπως και στους ελέγχους υποθέσεων της κανονικής κατανομής, μπορούμε κατά προσέγγιση να χρησιμοποιήσουμε στους ελέγχους υποθέσεων για αναλογίες τους κανόνες που συνοψίζονται στον πίνακα που ακολουθεί.

Έλεγχοι Υποθέσεων για Αναλογίες

$H_0 : p = p_0$		
$H_1 : p > p_0$	**$H_1 : p < p_0$**	**$H_1 : p \neq p_0$**
Απορρίπτω την H_0 αν $Z_0 = \dfrac{\frac{X}{n} - p_0}{\sqrt{\frac{p_0 q_0}{n}}} > Z_{1-\alpha}$ ή ισοδύναμα, αν $\frac{X}{n} > p_0 + Z_{1-\alpha}\sqrt{\frac{p_0 q_0/n}{n}}$	Απορρίπτω την H_0 αν $Z_0 = \dfrac{\frac{X}{n} - p_0}{\sqrt{\frac{p_0 q_0}{n}}} < Z_{1-\alpha}$ ή ισοδύναμα, αν $\frac{X}{n} < p_0 + Z_{1-\alpha}\sqrt{\frac{p_0 q_0/n}{n}}$	Απορρίπτω την H_0 αν $Z_0 > Z_{1-\alpha/2}$ ή αν $Z_0 < Z_{1-\alpha/2}$ ή ισοδύναμα, αν $\frac{X}{n} < p_0 + Z_{1-\alpha/2}\sqrt{\frac{p_0 q_0/n}{n}}$ ή αν $\frac{X}{n} > p_0 + Z_{1-\alpha/2}\sqrt{\frac{p_0 q_0/n}{n}}$
 $-Z_{1-\alpha}$	 $Z_{1-\alpha}$	 $-Z_{1-\alpha/2}$ $Z_{1-\alpha/2}$

Σημείωση: Η p-τιμή (το παρατηρούμενο επίπεδο σημαντικότητας) υπολογίζεται στην περίπτωση ελέγχου υποθέσεων αναλογιών με τον ίδιο τρόπο που υπολογίζεται στον έλεγχο υποθέσεων για την μέση τιμή πληθυσμού που ακολουθεί την κανονική κατανομή με γνωστή διακύμανση.

Παράδειγμα: Ο προϊστάμενος του ελέγχου ποιότητας μιας βιομηχανίας ενδιαφέρεται να διερευνήσει μια νέα μέθοδο συσκευασίας. Συγκεκριμένα, τον ενδιαφέρει να εξετάσει αν το ποσοστό των ελαττωματικών συσκευασιών με την καινούργια μέθοδο μειώνεται κάτω από 0.10 που ήταν το ποσοστό των ελαττωματικών συσκευασιών στο παρελθόν. Για να ελέγξει την υπόθεση αυτή, επιλέγει ένα τυχαίο δείγμα 200 πακέτων του προϊόντος που έχουν συσκευασθεί με την ίδια μέθοδο και παρατηρεί ότι 11 από αυτά είναι ελαττωματικά. Τί συμπέρασμα μπορεί να βγάλει ο υπεύθυνος του ελέγχου ποιότητας με βάση το δείγμα αυτό;

Λύση: Η υπόθεση που θα πρέπει να ελεγχθεί είναι,

$$H_0 : p = 0.10$$
$$H_1 : p < 0.10$$

Η τιμή της στατιστικής συνάρτησης ελέγχου, κάτω από την μηδενική υπόθεση είναι,

$$z_0 = \frac{0.055 - 0.10}{\sqrt{\frac{(.10)(.90)}{200}}} = \frac{-0.045}{\sqrt{0.00045}} = -2.12$$

Η προσέγγιση ισχύει γιατί το δείγμα είναι μεγάλο και, ταυτόχρονα $np \geq 5$ και $n(1\text{-}p) \geq 5$.

α) Αν ακολουθήσουμε τον κλασσικό τρόπο, τότε, σε επίπεδο σημαντικότητας α=0.05, απορρίπτουμε την μηδενική υπόθεση δοθέντος ότι,

$$z = -2.12 < -z_{.950} = -1.645$$

Σε επίπεδο σημαντικότητας δηλαδή 0.05 ο υπεύθυνος καταλήγει στο συμπέρασμα ότι υπάρχουν αρκετές ενδείξεις που να δείχνουν ότι το ποσοστό των ελαττωματικών με το νέο σύστημα είναι μικρότερο από 0.10.