B. 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

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

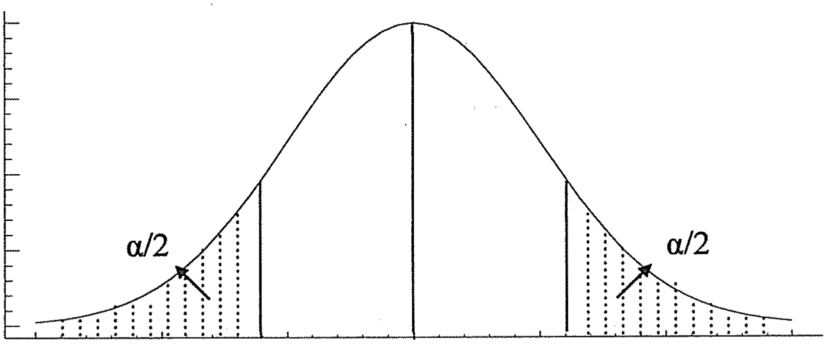
όπου και στις δύο περιπτώσεις n είναι το μέγεθος του δείγματος.
Όπως γνωρίζουμε στη περίπτωση αυτή, με την προϋπόθεση ότι ο πληθυσμός είναι κανονικός $X \sim N(\mu, \sigma^2)$, η στατιστική συνάρτηση

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

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

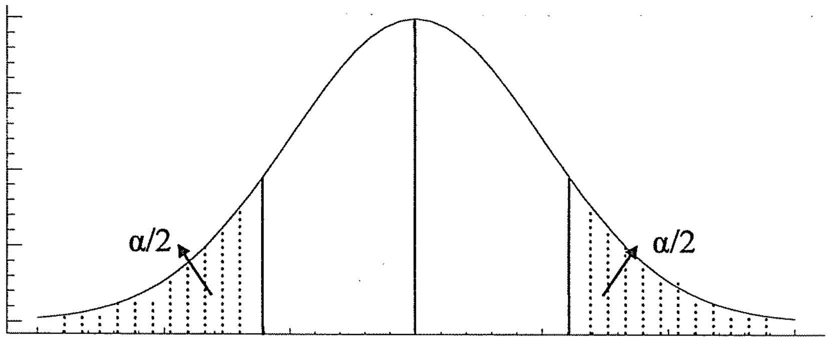
$$\alpha = P(\bar{H}_0 \mid H_0) = P(\overline{X} < c_1 \text{ ή } \overline{X} > c_2 \mid \mu = \mu_0)$$

$$= P\left(\frac{\overline{X}-\mu}{S^*/\sqrt{n}} < \frac{c_1-\mu}{S^*/\sqrt{n}} \text{ ή } \frac{\overline{X}-\mu}{S^*/\sqrt{n}} > \frac{c_2-\mu}{S^*/\sqrt{n}} \middle| -\mu = \mu_0 \right)$$

$$= P\left(\frac{\overline{X}-\mu_0}{S^*/\sqrt{n}} < \frac{c_1-\mu_0}{S^*/\sqrt{n}} \text{ ή } \frac{\overline{X}-\mu_0}{S^*/\sqrt{n}} > \frac{c_2-\mu_0}{S^*/\sqrt{n}} \right)$$

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

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



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

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

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,

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

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

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

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 T ακολουθεί την t_{n-1}.

Παρατήρηση: Χρειάζεται προσοχή στην χρησιμοποίηση της στατιστικής συνάρτησης T. 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 $\alpha = 0.05$. 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,

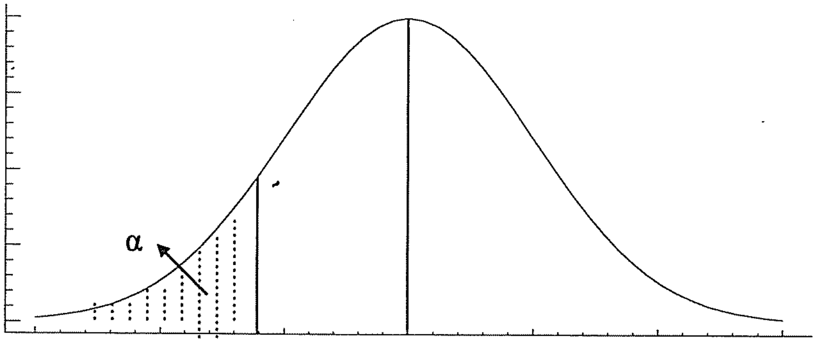
$$t_{24,.975} = 2.0639$$

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



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

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

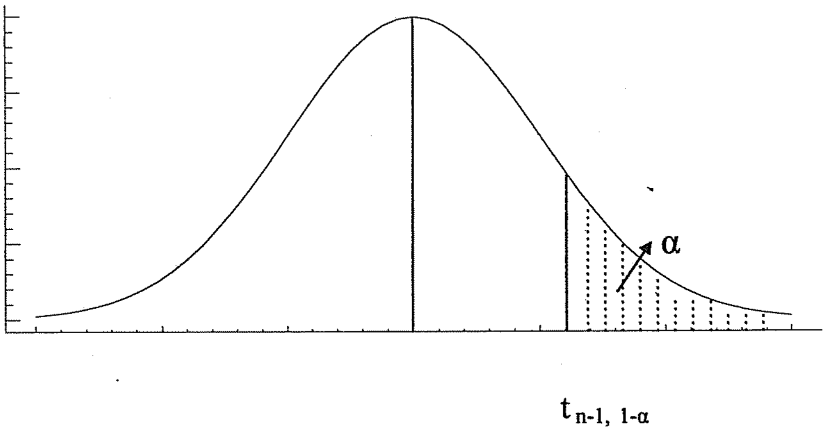
ή, ισοδύναμα, αν

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

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.