5.3 Έλεγχος της διαφοράς δύο μέσων. Ανεξάρτητα δείγματα

5.3.1 Γνωστές διακυμάνσεις

Έστω δύο τυχαίες μεταβλητές X_1 και X_2 και θέλουμε να ελέγξουμε την υπόθεση της ισότητας των μέσων τους με βάση την παρατηρούμενη διαφορά των δειγματικών μέσων δύο τυχαίων και ανεξάρτητων δειγμάτων μεγέθους n_1 και n_2, αντίστοιχα. Αν $X_1 \sim N(\mu_1, \sigma_1^2)$ και $X_2 \sim N(\mu_2, \sigma_2^2)$ τότε η τυχαία μεταβλητή

$$Z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

ακολουθεί την $N(0,1)$.

Αν οι X_1 και X_2 δεν ακολουθούν την κανονική κατανομή τότε, για $n_1, n_2 > 30$ η κατανομή της Z είναι κατά προσέγγιση η κανονική. Επομένως ο έλεγχος της ισότητας των μ_1 και μ_2 θα γίνει σταδιακά ως εξής:

α. Υποθέσεις που γίνονται δεκτές:

(i) Οι δύο κατανομές πληθυσμού είναι κανονικές. Η ισχύς της υπόθεσης αυτής μπορεί να αγνοηθεί αν $n_1, n_2 > 30$,

(ii) τα δείγματα είναι τυχαία και ανεξάρτητα, και

(iii) οι διακυμάνσεις πληθυσμού γνωστές.

β. Η μηδενική και η εναλλακτική υπόθεση:

$$\begin{array}{lll} H_0: & & \mu_1 = \mu_2 \\ H_\varepsilon: & \text{(i)} & \mu_1 \neq \mu_2 \\ & \text{(ii)} & \mu_1 > \mu_2 \\ & \text{(iii)} & \mu_1 < \mu_2 \end{array}$$

Στην πραγματικότητα ο έλεγχος της υπόθεσης (iii) είναι ίδιος με αυτόν της (ii) αρκεί να αλλάξουμε τα μ_1 και μ_2, γι' αυτό και στο εξής δεν θα αναφέρεται.

γ. Το κριτήριο αποφάσεως: Υπολογίζουμε το στατιστικό:

$$z_\pi = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

δ. Η απόφαση: Η μηδενική υπόθεση θα απορριφθεί σε επίπεδο σημαντικότητας α, για τον δίπλευρο και τον μονόπλευρο έλεγχο, αντίστοιχα, αν

(i) $|z_\pi| > z_{1-\alpha/2}$

(ii) $z_\pi > z_{1-\alpha}$

Παράδειγμα 5.5

Πολυεθνική εταιρεία θέλει να ελέγξει αν η απόδοση των εργατών της στο εργοστάσιο Α είναι ίδια με την απόδοση των εργατών στο εργοστάσιο Β. Σε τυχαίο δείγμα $\eta_1=30$ εργατών του εργοστασίου Α υπολόγισε μέσο χρόνο εκτέλεσης ορισμένου έργου 54min ενώ σε τυχαίο δείγμα $\eta_2=20$ εργατών του εργοστασίου Β ο μέσος χρόνος εκτέλεσης του ίδιου έργου ήταν 57min. Να ελεγχθεί η υπόθεση της ισότητας του μέσου χρόνου για το σύνολο των εργατών στα δύο εργοστάσια σε επίπεδο σημαντικότητας α=0.05 αν είναι γνωστό ότι ο χρόνος εκτέλεσης του έργου αυτού ακολουθεί την κανονική κατανομή και οι διακυμάνσεις είναι αντίστοιχα ίσες με $\sigma_1^2=36$ και $\sigma_2^2=64$.

α. Υποθέσεις που γίνονται δεκτές: Η κατανομή του χρόνου εκτέλεσης του έργου είναι κανονική και στους δύο πληθυσμούς, δηλαδή $X_1 \sim N(\mu_1, 36)$, $X_2 \sim N(\mu_2, 64)$.

β. Η μηδενική και η εναλλακτική υπόθεση:

$$H_0: \mu_1 = \mu_2$$
$$H_\varepsilon: \mu_1 \neq \mu_2$$

γ. Το κριτήριο αποφάσεως: Έχουμε $z_{1-\alpha/2}=1.96$. Επομένως αν

$$|z_\pi| = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{36}{30}+\frac{64}{20}}} > 1.96$$

θα απορρίψουμε την H_0 υπέρ της H_ε.

δ. Η απόφαση: Επειδή

$$|z_\pi| = \frac{|54-57|}{\sqrt{\frac{36}{30}+\frac{64}{20}}} = 1.43 < 1.96$$

δεν μπορούμε να απορρίψουμε την H_0·