5.7 Έλεγχος της διαφοράς δύο αναλογίων

Έστω ότι έχουμε δύο πληθυσμούς με αναλογία p_1 και p_2 αντίστοιχα. Αν $\hat{p}_1$ είναι η αναλογία σε τυχαίο δείγμα n_1 στοιχείων από τον πρώτο πληθυσμό και $\hat{p}_2$ η αναλογία σε τυχαίο και ανεξάρτητο δείγμα n_2 στοιχείων από τον δεύτερο πληθυσμό, όταν τα δείγματα είναι μεγάλα, η τυχαία μεταβλητή

$$Z=\frac{(\hat{p}_{X_1}-\hat{p}_{X_2})-(p_1-p_2)}{\sqrt{\frac{p_1(1-p_1)}{n_1}+\frac{p_2(1-p_2)}{n_2}}}$$

ακολουθεί την N(0, 1).

Όταν $p_1=p_2=p$ τότε η παραπάνω σχέση γράφεται ισοδύναμα ως εξής:

$$Z=\frac{\hat{p}_{X_1}-\hat{p}_{X_2}}{\sqrt{p(1-p)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

Επομένως, ο έλεγχος της ισότητας δύο αναλογιών πληθυσμού με βάση τις παρατηρήσεις δύο δειγμάτων, ένα από κάθε πληθυσμό, θα γίνει σταδιακά ως εξής:

α. Υποθέσεις που γίνονται δεκτές: Τα δείγματα είναι τυχαία και ανεξάρτητα.

β. Η μηδενική και η εναλλακτική υπόθεση:

$$
\begin{aligned}
&H_0: && p_1=p_2=p\\
&H_\varepsilon: && \text{(i)}\ p_1\neq p_2\\
& && \text{(ii)}\ p_1>p_2
\end{aligned}
$$

γ. Το κριτήριο αποφάσεως: Έστω $\hat{p}_1$ και $\hat{p}_2$ η τιμή της δειγματικής αναλογίας στο πρώτο και το δεύτερο δείγμα αντίστοιχα. Συνδυάζοντας τις πληροφορίες των δύο δειγμάτων εκτιμούμε την κοινή αναλογία ρ με την:

$$\hat{p}=\frac{n_1\hat{p}_{x_1}+n_2\hat{p}_{x_2}}{n_1+n_2}$$

και υπολογίζουμε τη στατιστικό

$$z_\pi=\frac{\hat{p}_{x_1}-\hat{p}_{x_2}}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}.$$

δ. Η απόφαση: Η μηδενική υπόθεση απορρίπτεται για τα δύο-είδη ελέγχου, αντίστοιχα, αν:

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

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

Παράδειγμα 5.10

Ένας επιχειρηματίας θέλει να δει αν το διαφημιστικό μήνυμα για το προϊόν του προκαλεί το ενδιαφέρον εξίσου στους άνδρες και τις γυναίκες. Σε δύο τυχαία και ανεξάρτητα δείγματα από άνδρες και γυναίκες, δέκτες του μηνύματος, πήρε τα εξής αποτελέσματα. Σε δείγμα n_1=200 ανδρών, ενδιαφέρθηκαν για το μήνυμα οι 50, ενώ σε δείγμα n_2=200 γυναικών ενδιαφέρθηκαν για το μήνυμα οι 70. Να ελεγχθεί η υπόθεση της ισότητας των δύο αναλογιών πληθυσμού σε επίπεδο σημαντικότητας α = 0.05.

α. Υποθέσεις που γίνονται δεκτές: Τα δύο δείγματα είναι τυχαία και ανεξάρτητα.

β. Η μηδενική και η εναλλακτική υπόθεση:

$$H_0:\ p_1 = p_2 = p$$
$$H_\varepsilon:\ p_1 \neq p_2$$

γ. Το κριτήριο αποφάσεως: Για α = 0.05 και α/2=0.025 έχουμε $z_{1-\alpha/2}$=1.96. Επομένως αν

$$|z_\pi| = \frac{|\hat{p}_{x_1} - \hat{p}_{x_2}|}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{200}+\frac{1}{200}\right)}} > 1.96$$

θα απορρίψουμε την H_0.

δ. Η απόφαση: Εκτιμούμε την κοινή αναλογία p

$$\hat{p} = \frac{50+70}{400} = 0.3$$

$$|z_\pi| = \frac{|0.25-0.35|}{\sqrt{(0.3)(0.7)\left(\frac{2}{200}\right)}} = 2.18 > 1.96$$

απορρίπτουμε την H_0 υπέρ της H_ε.

Επομένως σε επίπεδο σημαντικότητας α = 0.05 δεν γίνεται δεκτή η υπόθεση της ισότητας των δύο αναλογιών. Το ελάχιστο επίπεδο σημαντικότητας στο οποίο απορρίπτεται η H_0 δηλαδή η p-τιμή του ελέγχου ισούται με

$$\begin{aligned} 2P(Z > 2.18) &= 2[1 - F_Z(2.18)] \\ &= 2(1 - 0.98537) \\ &= 0.02926 \end{aligned}$$