Δύο ανεξάρτητα δείγματα, διαφορά μέσων τιμών
(πολλές παρατηρήσεις)

Για τους 35 φοιτητές που εξετάστηκαν στο μάθημα της Βιοστατιστικής τον Ιούνιο του 2012 ο δειγματικός μέσος και η δειγματική διασπορά των βαθμών τους ήταν $\overline{X_I}$ =6.04 και S_I^2=4.93 αντιστοίχως. Για τους 55 φοιτητές που εξετάσθηκαν τον Σεπτέμβριο του ιδίου έτους τα αντίστοιχα νούμερα ήταν $\overline{X_\Sigma}$ =6.45 και S_Σ^2=3.79. Να ελέγξετε εάν, σε επίπεδο σημαντικότητας α=10%, δεχόμαστε την υπόθεση ότι η μέση επίδοση των φοιτητών τον Σεπτέμβριο ήταν ίδια με τη μέση επίδοσή τους τον Ιούνιο ή όχι.

Λύση

Ο έλεγχος $H_o : \mu_I = \mu_\Sigma$ έναντι της $H_1 : \mu_I \neq \mu_\Sigma$ είναι αμφίπλευρος (και σύνθετος) και ισοδύναμα γίνεται $H_o : \mu_I - \mu_\Sigma = 0$ έναντι της $H_1 : \mu_I - \mu_\Sigma \neq 0$.

(Τρόπος 1^{ος}) Θα κατασκευάσω δ.ε., ίσων ουρών, για τη διαφορά $\mu_I - \mu_\Sigma$, σ.ε. $1-\alpha = 0.90$. Απορρίπτουμε την H_o όταν το διάστημα δεν περιέχει το μηδέν.

Επειδή $n_I = 35$ $n_\Sigma = 55$ (αρκετά μεγάλα μεγέθη δειγμάτων) το διάστημα θα δίδεται από τον τύπο:

$$\left(\overline{X_I} - \overline{X_\Sigma} - z_{\frac{\alpha}{2}}\sqrt{\frac{S_I^2}{n_I} + \frac{S_\Sigma^2}{n_\Sigma}}, \overline{X_I} - \overline{X_\Sigma} + z_{\frac{\alpha}{2}}\sqrt{\frac{S_I^2}{n_I} + \frac{S_\Sigma^2}{n_\Sigma}}\right) \tag{2}$$

όπου $z_{\frac{\alpha}{2}} \overset{\text{εδώ}}{=} z_{0.05} = 1.645$, εφόσον, $1-\alpha = 0.90 \Rightarrow \frac{\alpha}{2} = 0.05$ και $z_{0.05}$ είναι τέτοιο ώστε:
$P(Z > z_{0.05}) = 0.05 \Rightarrow \Phi(z_{0.05}) = P(Z \leq z_{0.05}) = 1 - P(Z > z_{0.05}) = 1 - 0.05 \Rightarrow \Phi(z_{0.05}) = 0.95$, όμως $\Phi(1.645) = 0.95 \Rightarrow z_{0.05} = 1.645$, για $Z \sim N(0,1)$.

Συνεπώς το δ.ε. (2) γίνεται:

$$\left(6.04 - 6.45 - 1.645\sqrt{0.1408 + 0.06890},\ \ 6.04 - 6.45 + 1.645\sqrt{0.1408 + 0.06890}\right) =$$

$$(-0.410 - 1.645 \cdot 0.4580,\ \ -0.410 + 1.645 \cdot 0.4580) = (-0.410 - 0.7533,\ \ -0.410 - 0.7533) =$$

$$\underline{\underline{(-1.163347,\ \ 0.343347)}} \qquad (3)$$

Επειδή το δ.ε. (3) περιέχει το 0, η H_0 γίνεται δεκτή σε ε.σ. 10%.

(Τρόπος 2^{ος}) Απορρίπτουμε την H_o όταν ισχύει $T = \left| \dfrac{\overline{X_I} - \overline{X_\Sigma} - 0}{\sqrt{\frac{S_I^2}{n_I} + \frac{S_\Sigma^2}{n_\Sigma}}} \right| > z_{\frac{\alpha}{2}}$.

Εδώ $T = \left|\frac{-0.410}{0.4580}\right| = |-0.8951| = \underline{\underline{0.8951}} < 1.645$. (το p-value είναι 0.3706)

η H_0 γίνεται δεκτή σε ε.σ. 10%

Δύο ανεξάρτητα δείγματα, διαφορά ποσοστών
(πολλές παρατηρήσεις)

Κατά τη διάρκεια μιας επιδημίας της κοινής γρίπης ένας ερευνητής, για να μελετήσει τη σχέση ανάμεσα στην ομάδα αίματος και στην προσβολή από τον ιό της γρίπης, επέλεξε τυχαία 226 άτομα και πήρε τα παρακάτω αποτελέσματα

Ομάδα Αίματος	A	B	AB	0
Προσβολή	54	14	7	43
Μη Προσβολή	56	15	16	21

Εάν συμβολίσουμε με p_A το ποσοστό των ατόμων με ομάδα αίματος A που προσβλήθηκαν από τον ιό της γρίπης και με p_0 το αντίστοιχο ποσοστό για τα άτομα με ομάδα αίματος 0, να ελέγξετε εάν σε επίπεδο σημαντικότητας α=5% δεχόμαστε την υπόθεση $H_o : p_A = p_0$ έναντι της $H_1 : p_A \neq p_0$.

Λύση

Εδώ $\widehat{p_A} = \frac{54}{54+56} = 0.4909$, $\widehat{p_0} = \frac{43}{43+21} = 0.6719$. Ο έλεγχος $H_o : p_A = p_0$ έναντι της $H_1 : p_A \neq p_0$ είναι αμφίπλευρος (και σύνθετος) και ισοδύναμα γίνεται $H_o : p_A - p_0 = 0$ έναντι της $H_1 : p_A - p_0 \neq 0$.

(Τρόπος $1^{ος}$)

Για αυτό, θα κατασκευάσω δ.ε., ίσων ουρών, για τη διαφορά $p_A - p_0$, σ.ε. $1 - \alpha = 0.95$.

Δύο ανεξάρτητα δείγματα, διαφορά ποσοστών (συνέχεια…)

Επειδή $n_A = 110$ $n_0 = 64$ (αρκετά μεγάλα μεγέθη δειγμάτων) το διάστημα θα δίδεται από τον τύπο:

$$\left(\widehat{p_A} - \widehat{p_0} - z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n_A} + \frac{\widehat{p}(1-\widehat{p})}{n_0}}\,, \widehat{p_A} - \widehat{p_0} + z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n_A} + \frac{\widehat{p}(1-\widehat{p})}{n_0}}\right) \quad (5)$$

όπου p είναι (υπό την υπόθεση H_0) η κοινή πιθανότητα $p = p_A = p_0$ με $\widehat{p} = \frac{54+43}{110+64} = 0.5575$,
και $z_{\frac{\alpha}{2}} \overset{\varepsilon\delta\acute{\omega}}{=} z_{0.025} = 1.96$.

(για $Z \sim N(0,1)$, $P(Z > z_{0.025}) = 0.025 \Rightarrow \Phi(z_{0.025}) = P(Z \leq z_{0.025}) = 1 - 0.025 \Rightarrow \Phi(z_{0.025}) = 0.975$, όμως $\Phi(1.96) = 0.975$) Συνεπώς το δ.ε. (5) γίνεται:

$$\left(0.4909 - 0.6719 - 1.96\sqrt{0.0022 + 0.0039},\;\; 0.4909 - 0.6719 + 1.96\sqrt{0.0022 + 0.0039}\right) =$$

$$(-0.1809 - 1.96 \cdot 0.0781,\;\; -0.1809 - 1.96 \cdot 0.0781) = (-0.1809 - 0.153,\;\; -0.1809 - 0.153) =$$

$$(-0.3340,\;\; -0.0279) \quad (6)$$

Επειδή το δ.ε. (6) δεν περιέχει το 0, η H_0 απορρίπτεται σε ε.σ. 5%.

Σημείωση

Εάν είχε χρησιμοποιηθεί (εσφαλμένα) ο τύπος: $\widehat{p_A} - \widehat{p_0} \pm z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{p_A}(1-\widehat{p_A})}{n_A} + \frac{\widehat{p_0}(1-\widehat{p_0})}{n_0}}$, θα παίρναμε το διάστημα: $0.4909 - 0.6719 \pm 1.96\sqrt{0.0023 + 0.0034} = -0.1809 \pm 1.96 \cdot 0.0756 = -0.1809 \pm 0.1482 =$ $(-0.3292\,,\; -0.0328)$, και θα καταλήξουμε στο ίδιο συμπέρασμα.

Δύο ανεξάρτητα δείγματα, διαφορά ποσοστών (συνέχεια…)

(Τρόπος $2^{ος}$) Απορρίπτουμε την H_o όταν ισχύει $T = \left| \dfrac{\widehat{p_A} - \widehat{p_0} - 0}{\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n_A} + \frac{\widehat{p}(1-\widehat{p})}{n_0}}} \right| > z_{\frac{\alpha}{2}}$.

όπου p είναι (υπό την υπόθεση H_0) η κοινή πιθανότητα $p = p_A = p_0$ με $\widehat{p} = \frac{54+43}{110+64} = 0.5575$

Εδώ $T = \left| \frac{-0.1809}{0.0756} \right| = |-2.3928| = \underline{\underline{2.3928}} > 1.96$ (το p-value είναι 0.0167)

η H_0 απορρίπτεται σε ε.σ. 5%