

1) Έλεγχος υπόθεσης για διαφορά δύο αναλογιών του πληθυσμού (P1 - P2)

Δύο φάρμακα: Τύπου A και Τύπου B

Ποιο από τα δύο είναι πιο αποτελεσματικό; Υπάρχει η πεποίθηση πως το φάρμακο τύπου B έχει μεγαλύτερη επιτυχία.

Παίρνουμε δύο δείγματα:

$n_A = 1.250$ άτομα (δείγμα A)

$n_B = 1.565$ άτομα (δείγμα B)

Επιτυχία θεραπείας (δειγματικές αναλογίες):

$r_A = 1.156$ άτομα $\rightarrow r_A / n_A = 1.156 / 1.250 = 0,9248$ (92,48%)

$r_B = 1.495$ άτομα $\rightarrow r_B / n_B = 1.495 / 1.565 = 0,9552$ (95,52%)

- Βήμα 1ο: Διατύπωση υποθέσεων

$H_0: (P1 - P2) = 0$

$H_1: (P1 - P2) < 0$ (μονοκατάληκτος)

- Βήμα 2ο: Υπολογισμός στατιστικής τιμής ελέγχου (Z)

$$Z = \frac{(p_1 - p_2) - 0}{\sqrt{p^*(1-p)(\frac{1}{n_1} + \frac{1}{n_2})}} = \frac{(0,9248 - 0,9552)}{\sqrt{0,9417 * 0,058 * (0,0008 + 0,0006)}}$$
$$= -3,48$$

όπου $p^* = (r_1 + r_2) / (n_1 + n_2) = 2.651 / 2.815 = 0,9417$

- Βήμα 3ο: Σύγκριση της στατιστικής τιμής ελέγχου (Z-stat) με την κριτική τιμή (Z-crit) από τους πίνακες

$$Z_{\alpha} = -1,64$$

$$-3,48 < -1,64$$

Απορρίπτουμε την H_0 (για $\alpha=5\%$)

$$Z_{\alpha} = -2,29$$

$$-3,48 < -2,29$$

**Απορρίπτουμε την H_0
(ακόμη και για $\alpha=1\%$)**

2) Έλεγχος υπόθεσης για διαφορά δύο αναλογιών του πληθυσμού (P1 - P2)

Επιμορφωτικά Προγράμματα και Απασχόληση

Μια ομάδα ανέργων συμμετέχει σε ένα κρατικό πρόγραμμα επιμόρφωσης. Υπάρχει η πεποίθηση πως οι προοπτικές απασχόλησης βελτιώνονται μετά τη συμμετοχή στην επιμόρφωση.

Παίρνουμε δύο δείγματα:

$n_A = 850$ άτομα που επιμορφώθηκαν

$n_B = 565$ άτομα που ΔΕΝ επιμορφώθηκαν

Επίπεδα απασχόλησης σήμερα:

$r_A = 706$ άτομα $\rightarrow r_A / n_A = 706 / 850 = 0,8305$ (83,05%)

$r_B = 395$ άτομα $\rightarrow r_B / n_B = 395 / 565 = 0,6991$ (69,91%)

- Βήμα 1ο: Διατύπωση υποθέσεων

$$H_0: (P1 - P2) = 0$$

$$H_1: (P1 - P2) > 0 \quad (\text{μονοκατάληκτος})$$

- Βήμα 2ο: Υπολογισμός στατιστικής τιμής ελέγχου

$$Z = \frac{(p_1 - p_2)}{\sqrt{p^*(1-p)(\frac{1}{n_1} + \frac{1}{n_2})}} = \frac{(0,8305 - 0,6991)}{\sqrt{0,7780 * 0,2219 * (0,0011 + 0,0017)}}$$
$$= 5,94$$

$$\text{όπου } p^* = (r_1 + r_2) / (n_1 + n_2) = 1.101 / 1.415 = 0,7780$$

- Βήμα 3ο: Σύγκριση της στατιστικής τιμής ελέγχου (Z-stat) με την κριτική τιμή (Z-crit) από τους πίνακες

$$Z_\alpha = 1,64$$

$$5,94 > 1,64$$

Απορρίπτουμε την H_0 (για $\alpha=5\%$)

$$Z_\alpha = 2,29$$

$$5,94 > 2,29$$

**Απορρίπτουμε την H_0
(ακόμη και για $\alpha=1\%$)**

- 3) Μια ομάδα ανέργων συμμετέχει σε ένα κρατικό πρόγραμμα επιμόρφωσης. **Ήταν στατιστικά σημαντικός παράγοντας η συμμετοχή των ανέργων στην επιμόρφωση;**

$$H_0: (P_1 - P_2) = 0$$

$$H_1: (P_1 - P_2) \neq 0 \quad (\text{δικοτάληκτος})$$

$$Z\text{-στατ.} = 5,94$$

$$Z_{\alpha/2} = 1,96$$

$$5,94 > 1,96$$

Απορρίπτουμε την H_0 (για $\alpha=5\%$)

$$Z_{\alpha/2} = 2,57$$

$$5,94 > 2,57$$

**Απορρίπτουμε την H_0
(ακόμη και για $\alpha=1\%$)**

- 4) Μια ομάδα ανέργων συμμετέχει σε ένα κρατικό πρόγραμμα επιμόρφωσης. **Είχε ως αποτέλεσμα η συμμετοχή των ανέργων στην επιμόρφωση το άνοιγμα μιας ψαλίδας ΑΝΩ του 20%;**

$$H_0: (P_1 - P_2) = 0,2$$

$$H_1: (P_1 - P_2) > 0,2 \quad (\text{μονοκατάληκτος})$$

Διόρθωση: εδώ δεν υποθέτουμε κοινή αναλογία. Άρα, το τυπικό σφάλμα στον παρανομαστή δεν έχει αυτό το κοινό p. Στον παρανομαστή θα υπάρχει η τετραγωνική ρίζα του: $[p_1*(1-p_1)]/n_1 + [p_2*(1-p_2)]/n_2$

$$Z = \frac{(p_1 - p_2) - 0,2}{\sqrt{p^*(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0,8305 - 0,6991) - 0,2}{\sqrt{0,7780 * 0,2219 * (0,0011 + 0,0017)}} = -3,12$$

$$\text{όπου } p^* = (r_1 + r_2) / (n_1 + n_2) = 1.101 / 1.415 = 0,7780$$

$$Z_\alpha = 1,64$$

$$-3,12 < 1,64$$

ΔΕΝ μπορούμε να απορρίψουμε την H_0 (για $\alpha=5\%$)

$$Z_\alpha = 2,29$$

$$-3,12 < 2,29$$

ΔΕΝ μπορούμε να απορρίψουμε την H_0 (ακόμη και για $\alpha=1\%$)

- 5) Μια ομάδα ανέργων συμμετέχει σε ένα κρατικό πρόγραμμα επιμόρφωσης. **Είχε ως αποτέλεσμα η συμμετοχή των ανέργων στην επιμόρφωση το άνοιγμα μιας ψαλίδας ΑΝΩ του 10%;**

$$H_0: (P_1 - P_2) = 0,1$$

$$H_1: (P_1 - P_2) > 0,1 \quad (\text{μονοκατάληκτος})$$

Διόρθωση: εδώ δεν υποθέτουμε κοινή αναλογία. Άρα, το τυπικό σφάλμα στον παρανομαστή δεν έχει αυτό το κοινό p. Στον παρανομαστή θα υπάρχει η τετραγωνική ρίζα του:

$$[p_1^*(1-p_1)]/n_1 + [p_2^*(1-p_2)]/n_2$$

$$Z = \frac{(p_1 - p_2) - 0,1}{\sqrt{p^*(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0,8305 - 0,6991) - 0,1}{\sqrt{0,7780 * 0,2219 * (0,0011 + 0,0017)}}$$
$$= 1,43$$

$$\text{όπου } p^* = (r_1 + r_2) / (n_1 + n_2) = 1.101 / 1.415 = 0,7780$$

$$Z_{\alpha} = 1,64$$

$$1,43 < 1,64$$

ΔΕΝ μπορούμε να απορρίψουμε την H_0 (για $\alpha=5\%$)

$$Z_{\alpha} = 1,28$$

$$1,43 > 1,28$$

Μπορούμε να απορρίπτουμε την H_0 (για $\alpha=10\%$)

Έλεγχος υπόθεσης για διαφορά δύο μέσων του πληθυσμού: ($\mu_1 - \mu_2$)

Έμφυλη Ανισότητα στην Απασχόληση

Συμμετέχουν περισσότερο οι άντρες συγκριτικά με τις γυναίκες στην αγορά εργασίας;

- **Βήμα 1ο:** Διατύπωση υποθέσεων

Δίπλευρος

$H_0: (\mu_{\text{ανδ}} - \mu_{\text{γυν}}) = 0 \rightarrow$ ισότιμη συμμετοχή

$H_1: (\mu_{\text{ανδ}} - \mu_{\text{γυν}}) \neq 0 \rightarrow$ ΜΗ ισότιμη συμμετοχή (συμμετέχουν περισσότερο ή λιγότερο ;;;)

Μονόπλευρος

$H_0: (\mu_{\text{ανδ}} - \mu_{\text{γυν}}) = 0 \rightarrow$ ισότιμη συμμετοχή

$H_1: (\mu_{\text{ανδ}} - \mu_{\text{γυν}}) > 0 \rightarrow$ ΜΗ ισότιμη συμμετοχή (οι άνδρες συμμετέχουν περισσότερο)

Μονόπλευρος

$H_0: (\mu_{\text{ανδ}} - \mu_{\text{γυν}}) = 0 \rightarrow$ ισότιμη συμμετοχή

$H_1: (\mu_{\text{ανδ}} - \mu_{\text{γυν}}) < 0 \rightarrow$ ΜΗ ισότιμη συμμετοχή (οι γυναίκες συμμετέχουν περισσότερο)

- **Βήμα 2ο:** Υπολογισμός στατιστικής τιμής ελέγχου

Γνωστές διακυμάνσεις πληθυσμού (σ_1, σ_2)

$$Z - stat = \frac{(\bar{X}_{men} - \bar{X}_{women})}{\sqrt{\frac{\sigma_{men}^2}{n_{men}} + \frac{\sigma_{women}^2}{n_{women}}}}$$

Άγνωστες διακυμάνσεις πληθυσμού (σ), αλλά μεγάλα δείγματα (n)

$$Z - stat = \frac{(\bar{X}_{men} - \bar{X}_{women})}{\sqrt{\frac{S_{men}^2}{n_{men}} + \frac{S_{women}^2}{n_{women}}}}$$

Άγνωστες διακυμάνσεις πληθυσμού (σ), αλλά μικρά δείγματα (n)

$$t - stat = \frac{(\bar{X}_{men} - \bar{X}_{women})}{\sqrt{\frac{S_{men}^2}{n_{men}} + \frac{S_{women}^2}{n_{women}}}}$$

ΠΡΟΣΟΧΗ !!!

$$\text{Βαθμοί Ελευθερίας: } df = \frac{(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2})^2}{\frac{(\frac{S_1^2}{n_1})^2}{n_1-1} + \frac{(\frac{S_2^2}{n_2})^2}{n_2-1}}$$

.....
Άγνωστες διακυμάνσεις πληθυσμού (σ), αλλά μικρά δείγματα (n)

$$t - stat = \frac{(\bar{X}_{men} - \bar{X}_{women})}{S^* \sqrt{\frac{1}{n_{men}} + \frac{1}{n_{women}}}}$$

ΠΡΟΣΟΧΗ !!!

Βαθμοί Ελευθερίας: $df = n_1 + n_2 - 2$

$$\text{όπου: } S^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} \quad (\text{κοινή διακύμανση})$$

.....

- **Βήμα 3ο:** Σύγκριση της στατιστικής τιμής ελέγχου με τις κριτικές τιμές