

Πιθανότητες & Στατιστική

Έλεγχοι Υποθέσεων

Δημήτρης Μητσούδης

Στο παράδειγμα αυτό, η εταιρεία δημοσκοπήσεων θα πρέπει να πάρει μια απόφαση σχετικά με το αν θα εκλεγεί ο Χ ή όχι. Άρα έχουμε δύο αντικρουόμενες απόψεις:

- H_0 : Ο Χ δεν θα εκλεγεί την πρώτη Κυριακή.
- H_1 : Ο Χ θα εκλεγεί την πρώτη Κυριακή.

Στο παράδειγμά μας, όπως και σε κάθε στατιστικό έλεγχο, η H_0 ονομάζεται **μηδενική υπόθεση**, ενώ η H_1 **εναλλακτική υπόθεση**.

Εισαγωγή¹

Έστω (και πάλι) ότι επιλέγουμε ένα τυχαίο δείγμα από ένα πληθυσμό. Τώρα, όμως, αντί για να εκτιμήσουμε (μέσω του δείγματος) **άγνωστες** παραμέτρους του πληθυσμού θα χρησιμοποιήσουμε το δείγμα για να **ελέγξουμε συγκεκριμένες υποθέσεις** αναφορικά με τις παραμέτρους αυτές.

Παράδειγμα: Πριν τις εκλογές, ο Χ, υποψήφιος δήμαρχος μιας πόλης παρήγγειλε δημοσκόπηση από την οποία προέκυψε ότι σε σύνολο 1000 ερωτηθέντων, 600 άτομα θα τον ψηφίσουν. Από το ποσοστό αυτό φαίνεται λογικό πώς ο Χ θα εκλεγεί από την πρώτη Κυριακή. Είναι αυτό σωστό; Πώς θα πρέπει να το ελέγξει η εταιρεία δημοσκοπήσεων;

¹ Στο κεφάλαιο αυτό ακολουθούμε κυρίως τις σημειώσεις: Εισαγωγή στις Πιθανότητες και τη Στατιστική, Χ. Δαμιανού, Ν. Παπαδάτος, Χ. Χαραλαμπίδης, Τμ. Μαθηματικών, Παν. Αθηνών, Αθήνα 2003.

Στη διεθνή πρακτική έχουν καθιερωθεί δύο βασικοί (εμπειρικοί) κανόνες για τον καθορισμό της H_0 . (Η H_1 καθορίζεται αυτόματα ως η άρνηση της H_0 .)

- **1ος Κανόνας:** Θέτουμε ως μηδενική την υπόθεση εκείνη της οποίας η εσφαλμένη απόρριψη εγκυμονεί τους περισσότερους κινδύνους.
- **2ος Κανόνας:** (Απλούστερος.) Θέτουμε ως μηδενική την υπόθεση εκείνη την οποία επιθυμούμε να απορρίψουμε (ώστε να αποφασίσουμε υπέρ της H_1).

1. Καθορίζεται η μηδενική υπόθεση H_0 .
2. Καθορίζεται μια μικρή πιθανότητα α (συνήθως $\alpha = 0.01, 0.05, 0.1$ ή 0.001) που είναι η **μέγιστη αποδεκτή πιθανότητα εσφαλμένης απόρριψης της μηδενικής H_0** .
Το α λέγεται **επίπεδο σημαντικότητας (ε.σ.) του ελέγχου**.
3. Λαμβάνεται ένα τυχαίο δείγμα (τ.δ.) $X_1, X_2, \dots, X_n$, το οποίο υποτίθεται ότι ακολουθεί κατανομή τέτοια ώστε να δικαιολογείται η H_0 . Κατασκευάζεται κατάλληλη στατιστική συνάρτηση T , της οποίας η κατανομή υπό την H_0 είναι συνήθως γνωστή.

4. Ανάλογα με τη μορφή της H_0 , υπολογίζονται οι **σταθερές αποκοπής c_0 ή c_1, c_2** για τις οποίες

$$P(T < c_0) \geq 1 - \alpha \quad \text{ή} \quad P(T > c_0) \geq 1 - \alpha \quad \text{ή} \quad P(c_1 < T < c_2) \geq 1 - \alpha.$$

(Το ποιά από τις 3 χρησιμοποιείται καθορίζεται από τη μορφή της H_0 .)

5. Ανάλογα με τη μορφή της H_0 , απορρίπτουμε την H_0 όταν για την τιμή της T ισχύει:

$T \geq c_0$ στην α' περίπτωση

$T \leq c_0$ στην β' περίπτωση

$T \leq c_1$ ή $T \geq c_2$ στην γ' περίπτωση

Π.χ., έστω ότι $\alpha = 0.05$ και κοιτάμε την α' περίπτωση όπου $P(T < c_0) \geq 1 - 0.05$.

Τότε αν ισχύει η H_0 :

$$P(T \geq c_0) = 1 - P(T < c_0) \leq 1 - (1 - 0.05) = 0.05.$$

Με άλλα λόγια, η σχέση $T \geq c_0$ ισχύει το πολύ στο 5% των περιπτώσεων που θα πάρουμε τ.δ. (εφόσον ισχύει η H_0). Αυτή την πιθανότητα την θεωρούμε αρκετά μικρή και έτσι αποφασίζουμε (με ρίσκο το πολύ 5%) ότι δεν ισχύει η H_0 .

Άρα, αν τελικά η τιμή της T ξεπεράσει τη σταθερά αποκοπής c_0 , τότε

- είτε τα δεδομένα μας δεν προέρχονται από την H_0 (δηλ. δεν ισχύει η H_0)
- ή τα δεδομένα μας προέρχονται από την H_0 αλλά, για κακή μας τύχη, προέκυψε μια σπάνια τιμή της T ($T \geq c_0$), η οποία συμβαίνει με πιθανότητα το πολύ 5%.

Σφάλμα τύπου I και II

Προφανώς στη λήψη μιας απόφασης υπάρχουν οι ακόλουθες 4 δυνατότητες:

	Ισχύει η H_0	Ισχύει η H_1
Αποφασίζουμε υπέρ H_0	Ορθή αποδοχή της H_0	Σφάλμα τύπου II
Αποφασίζουμε υπέρ H_1	Σφάλμα τύπου I	Ορθή απόρριψη της H_0

Παρατηρήστε ότι έχουμε ορίσει ως επίπεδο σημαντικότητας του ελέγχου την πιθανότητα να έχουμε Σφάλμα Τύπου I.

Συμπέρασμα

- Όταν απορρίπτουμε την H_0 σε ε.σ. α , τότε είμαστε τουλάχιστον $100 \cdot (1 - \alpha)\%$ σίγουροι ότι δεν ισχύει η H_0 ($\Leftrightarrow$ ισχύει η H_1).
- Όταν αποδεχόμαστε την H_0 σε ε.σ. α , τότε απλά δεν κατορθώσαμε να την απορρίψουμε, δηλ. από τα δεδομένα μας δεν προκύπτουν αρκετά στοιχεία εναντίον της. Αυτό όμως δεν σημαίνει ότι ισχύει η H_0 .

Έλεγχοι υποθέσεων για τον μέσο ενός πληθυσμού

Έλεγχος για το μέσο κανονικής με γνωστή διασπορά

Έστω $X_1, X_2, \dots, X_n$ τ.δ. από κανονική κατανομή $\mathcal{N}(\mu, \sigma^2)$ με γνωστή διασπορά σ^2 . Τότε

- Για τον έλεγχο της

$$H_0: \mu \leq \mu_0 \text{ έναντι της } H_1: \mu > \mu_0$$

σε ε.σ. α , απορρίπτουμε την H_0 όταν

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \geq z_\alpha,$$

όπου z_α το α -άνω ποσοστιαίο σημείο της τυποποιημένης κανονικής κατανομής.

- Για τον έλεγχο της

$$H_0: \mu \geq \mu_0 \text{ έναντι της } H_1: \mu < \mu_0$$

σε $\epsilon.σ. \alpha$, απορρίπτουμε την H_0 όταν

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \leq -z_\alpha.$$

- Για τον (αμφίπλευρο) έλεγχο της

$$H_0: \mu = \mu_0 \text{ έναντι της } H_1: \mu \neq \mu_0$$

σε $\epsilon.σ. \alpha$, απορρίπτουμε την H_0 όταν

$$|Z| = \left| \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \right| \geq z_{\alpha/2},$$

όπου $z_{\alpha/2}$ το $\alpha/2$ -άνω ποσοστιαίο σημείο της τυποποιημένης κανονικής κατανομής.

Συνοπτικά για το μέσο κανονικής με γνωστή διασπορά σ

Αν

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

και με R συμβολίζουμε το λεγόμενο **χωρίο απόρριψης**, τότε για τους ακόλουθους ελέγχους σε $\epsilon.σ. \alpha$ ισχύουν:

$H_0: \mu \leq \mu_0$ $H_1: \mu > \mu_0$	$H_0: \mu \geq \mu_0$ $H_1: \mu < \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$
$R = \{Z > z_\alpha\}$	$R = \{Z < -z_\alpha\}$	$R = \{ Z > z_{\alpha/2}\}$

Παράδειγμα

Έστω ότι ένα σήμα με τιμή μ μεταδίδεται από τη θέση A και η τιμή που λαμβάνεται στη θέση B ακολουθεί κανονική κατανομή με μέση τιμή μ και διασπορά 4. (Δηλαδή, ο τυχαίος θόρυβος που προστίθεται στο σήμα ακολουθεί κανονική κατανομή $\mathcal{N}(0, 4)$.)

Οι άνθρωποι που λαμβάνουν το σήμα στη θέση B υποψιάζονται ότι σήμερα θα σταλεί σήμα με $\mu = 8$. Ελέγξτε την υπόθεση αυτή σε $\epsilon.σ. \alpha = 0.05$, δεδομένου ότι το ίδιο σήμα στάλθηκε ανεξάρτητα 5 φορές και η μέση τιμή που ελήφθη στη θέση B ήταν $\bar{X} = 9.5$.

Λύση

Η ερώτηση αυτή γράφεται ισοδύναμα στη μορφή του ακόλουθου ελέγχου:

$$H_0: \mu = 8 \text{ έναντι της } H_1: \mu \neq 8$$

Υπολογίζουμε το

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{9.5 - 8}{\frac{2}{\sqrt{5}}} = 1.68.$$

Εδώ $\alpha = 0.05 \Rightarrow \alpha/2 = 0.025$, και

$$z_{\alpha/2} = z_{0.025} = 1.96.$$

Επειδή $|Z| = 1.68 < 1.96 = z_{\alpha/2}$ αποδεχόμαστε την H_0 .

Έλεγχος για το μέσο κανονικής με άγνωστη διασπορά

Έστω $X_1, X_2, \dots, X_n$ τ.δ. από κανονική κατανομή $\mathcal{N}(\mu, \sigma^2)$ με άγνωστη διασπορά σ^2 . Τότε

- Για τον έλεγχο της

$$H_0: \mu \leq \mu_0 \text{ έναντι της } H_1: \mu > \mu_0$$

σε ε.σ. α , απορρίπτουμε την H_0 όταν

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \geq t_{n-1, \alpha},$$

όπου $t_{n-1, \alpha}$ το α -άνω ποσοστιαίο σημείο της t_{n-1} και

$$S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}$$
 η δειγματική τυπική απόκλιση.

- Για τον έλεγχο της

$$H_0: \mu \geq \mu_0 \text{ έναντι της } H_1: \mu < \mu_0$$

σε ε.σ. α , απορρίπτουμε την H_0 όταν

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \leq -t_{n-1, \alpha}.$$

- Για τον (αμφίπλευρο) έλεγχο της

$$H_0: \mu = \mu_0 \text{ έναντι της } H_1: \mu \neq \mu_0$$

σε ε.σ. α , απορρίπτουμε την H_0 όταν

$$|T| = \left| \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \right| \geq t_{n-1, \alpha/2},$$

όπου $t_{n-1, \alpha/2}$ το $\alpha/2$ -άνω ποσοστιαίο σημείο της t_{n-1} .

Συνοπτικά για το μέσο κανονικής με άγνωστη διασπορά

Αν

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

και με R συμβολίζουμε το λεγόμενο **χωρίο απόρριψης**, τότε για τους ακόλουθους ελέγχους σε ε.σ. α ισχύουν:

$H_0: \mu \leq \mu_0$ $H_1: \mu > \mu_0$	$H_0: \mu \geq \mu_0$ $H_1: \mu < \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$
$R = \{T > t_{n-1, \alpha}\}$	$R = \{T < -t_{n-1, \alpha}\}$	$R = \{ T > t_{n-1, \alpha/2}\}$

Παράδειγμα

Ο κατασκευαστής ενός καινούργιου τύπου ελαστικού υποστηρίζει ότι αυτό έχει μέση διάρκεια ζωής τουλάχιστον 40 000 μίλια. Για να ελέγξουμε τον ισχυρισμό του δοκιμάζουμε 12 ελαστικά. Η διάρκεια ζωής τους δίνεται στον ακόλουθο πίνακα (σε χιλιάδες μίλια) :

1	2	3	4	5	6	7	8	9	10	11	12
36.1	40.2	33.8	38.5	42	35.8	37	41	36.8	37.2	33	36

Ελέγξτε τον ισχυρισμό του κατασκευαστή σε ε.σ. $\alpha = 0.05$.

Λύση

Το πρόβλημα αυτό γράφεται στη μορφή του ακόλουθου ελέγχου:

$$H_0: \mu \geq 40000 \text{ έναντι της } H_1: \mu < 40000$$

Υπολογίζουμε το δειγματικό μέσο και τη δειγματική τυπική απόκλιση:

$$\bar{X} = 37283.3, \quad S = 2731.9.$$

Στη συνέχεια υπολογίζουμε το

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{37283.3 - 40000}{\frac{2731.9}{\sqrt{12}}} = -3.4448.$$

$$\text{Εδώ } \alpha = 0.05 \Rightarrow t_{n-1, \alpha} = t_{11, 0.05} = 1.796 \Rightarrow -t_{11, 0.05} = -1.796.$$

Επειδή $T = -3.4448 < -1.796 = -t_{11, 0.05}$ απορρίπτουμε την H_0 σε ε.σ. $\alpha = 0.05$.

Έλεγχος για το μέσο μ οποιασδήποτε κατανομής, για μεγάλο n

Έστω $X_1, X_2, \dots, X_n$ τ.δ. από οποιαδήποτε κατανομή με μέσο $\mu = E(X_i)$ και διασπορά $\sigma^2 = \text{Var}(X_i)$. Υποθέτουμε ότι το μέγεθος δείγματος n είναι μεγάλο (θεωρητικά $n \rightarrow \infty$, στην πράξη $n \geq 30$).

- Για τον έλεγχο της

$$H_0: \mu \leq \mu_0 \text{ έναντι της } H_1: \mu > \mu_0$$

σε ε.σ. α , απορρίπτουμε την H_0 όταν

$$Z = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \geq z_\alpha,$$

όπου z_α το α -άνω ποσοστιαίο σημείο της τυποποιημένης κανονικής κατανομής και S η δειγματική τυπική απόκλιση.

- Για τον έλεγχο της

$$H_0: \mu \geq \mu_0 \text{ έναντι της } H_1: \mu < \mu_0$$

σε ε.σ. α , απορρίπτουμε την H_0 όταν

$$Z = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \leq -z_\alpha.$$

- Για τον (αμφίπλευρο) έλεγχο της

$$H_0: \mu = \mu_0 \text{ έναντι της } H_1: \mu \neq \mu_0$$

σε ε.σ. α , απορρίπτουμε την H_0 όταν

$$|Z| = \left| \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \right| \geq z_{\alpha/2},$$

όπου $z_{\alpha/2}$ το $\alpha/2$ -άνω ποσοστιαίο σημείο της τυποποιημένης κανονικής κατανομής.

Έλεγχος για το άγνωστο ποσοστό p στην κατανομή Bernoulli

Έστω $X_1, X_2, \dots, X_n$ τ.δ. από κατανομή Bernoulli με άγνωστο ποσοστό p και ας υποθέσουμε ότι $n \geq 30$.

- Για τον έλεγχο της

$$H_0: p \leq p_0 \text{ έναντι της } H_1: p > p_0$$

σε ε.σ. α , απορρίπτουμε την H_0 όταν

$$Z = \frac{\bar{X} - p_0}{\sqrt{\frac{\bar{X}(1-\bar{X})}{n}}} \geq z_\alpha,$$

όπου z_α το α -άνω ποσοστιαίο σημείο της τυποποιημένης κανονικής κατανομής.

- Για τον έλεγχο της

$$H_0: p \geq p_0 \text{ έναντι της } H_1: p < p_0$$

σε ε.σ. α , απορρίπτουμε την H_0 όταν

$$Z = \frac{\bar{X} - p_0}{\sqrt{\frac{\bar{X}(1-\bar{X})}{n}}} \leq -z_\alpha.$$

- Για τον έλεγχο της

$$H_0: p = p_0 \text{ έναντι της } H_1: p \neq p_0$$

σε ε.σ. α , απορρίπτουμε την H_0 όταν

$$|Z| = \left| \frac{\bar{X} - p_0}{\sqrt{\frac{\bar{X}(1-\bar{X})}{n}}} \right| \geq z_{\alpha/2}.$$

Έστω $X_i = 1$ αν ο i -οστός ψηφοφόρος ψηφίζει τον Χ και $X_i = 0$ αν δεν τον ψηφίζει. Έστω p το ποσοστό που θα πάρει ο Χ την πρώτη Κυριακή. Άρα οι υποθέσεις H_0 και H_1 γράφονται στη μορφή:

$$H_0: p \leq 0.5 \text{ έναντι της } H_1: p > 0.5$$

Ας επιλέξουμε ε.σ. $\alpha = 0.05$. Από τα δεδομένα μας έχουμε $\bar{X} = \frac{600}{1000} = 0.6$ και $n = 1000$. Άρα

$$Z = \frac{\bar{X} - p_0}{\sqrt{\frac{\bar{X}(1-\bar{X})}{n}}} = \frac{0.6 - 0.5}{\sqrt{\frac{0.24}{1000}}} = 6.455.$$

Επειδή

$$Z = 6.455 > z_{0.05} = 1.645$$

απορρίπτουμε την H_0 σε ε.σ. 5%. Δηλαδή, είμαστε κατά 95% βέβαιοι ότι ο Χ θα εκλεγεί από την πρώτη Κυριακή.

Παράδειγμα

Πριν τις εκλογές, ο Χ, υποψήφιος δήμαρχος μιας πόλης παρήγγειλε δημοσκοπήηση από την οποία προέκυψε ότι σε σύνολο 1000 ερωτηθέντων, 600 άτομα θα τον ψηφίσουν. Από το ποσοστό αυτό φαίνεται λογικό πώς ο Χ θα εκλεγεί από την πρώτη Κυριακή. Είναι αυτό σωστό; Πώς θα πρέπει να το ελέγξει η εταιρεία δημοσκοπήσεων;

Όπως ήδη είπαμε, η εταιρεία δημοσκοπήσεων θα πρέπει να πάρει μια απόφαση σχετικά με το αν θα εκλεγεί ο Χ ή όχι. Έστω λοιπόν:

- H_0 : Ο Χ δεν θα εκλεγεί την πρώτη Κυριακή.
- H_1 : Ο Χ θα εκλεγεί την πρώτη Κυριακή.

Έλεγχοι υποθέσεων για τη διαφορά των μέσων από δύο δείγματα

Πολλές φορές θέλουμε να συγκρίνουμε δύο ανεξάρτητους πληθυσμούς με βάση δύο αντίστοιχα ανεξάρτητα δείγματα.

Π.χ., έστω $X_1, X_2, \dots, X_n$ τ.δ. από τον πληθυσμό 1 που έχει σ.κ. F_1 και $Y_1, Y_2, \dots, Y_n$ τ.δ. από τον πληθυσμό 2 που έχει σ.κ. F_2 . Στη συνέχεια θα δούμε τρόπους να ελέγχουμε τη διαφορά των αντίστοιχων μέσων.

Τέτοια προβλήματα παρουσιάζονται π.χ. όταν συγκρίνουμε δύο νέα φάρμακα, ή δυο νέες διδακτικές μεθόδους κλπ.

- Για τον έλεγχο της

$$H_0: \mu_1 \geq \mu_2 \text{ έναντι της } H_1: \mu_1 < \mu_2$$

σε **ε.σ. α**, απορρίπτουμε την H_0 όταν

$$Z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \leq -z_\alpha.$$

- Για τον (αμφίπλευρο) έλεγχο της

$$H_0: \mu_1 = \mu_2 \text{ έναντι της } H_1: \mu_1 \neq \mu_2$$

σε **ε.σ. α**, απορρίπτουμε την H_0 όταν

$$|Z| = \frac{|\bar{X} - \bar{Y}|}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \geq z_{\alpha/2},$$

όπου $z_{\alpha/2}$ το $\alpha/2$ -άνω ποσοστιαίο σημείο της τυποποιημένης κανονικής κατανομής.

Έλεγχος για τη διαφορά δύο μέσων από την κανονική με γνωστές διασπορές

Έστω τα ανεξάρτητα τ.δ. $X_1, X_2, \dots, X_{n_1}$ από **κανονική κατανομή** $\mathcal{N}(\mu_1, \sigma_1^2)$ και $Y_1, Y_2, \dots, Y_{n_2}$ τ.δ. από **κανονική κατανομή** $\mathcal{N}(\mu_2, \sigma_2^2)$ με **γνωστές διασπορές** σ_1^2, σ_2^2 . Τότε

- Για τον έλεγχο της

$$H_0: \mu_1 \leq \mu_2 \text{ έναντι της } H_1: \mu_1 > \mu_2$$

σε **ε.σ. α**, απορρίπτουμε την H_0 όταν

$$Z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \geq z_\alpha,$$

όπου z_α το α -άνω ποσοστιαίο σημείο της τυποποιημένης κανονικής κατανομής.

Έλεγχος για τη διαφορά δύο μέσων από την κανονική με άγνωστες αλλά ίσες διασπορές

Έστω τα ανεξάρτητα τ.δ. $X_1, X_2, \dots, X_{n_1}$ από **κανονική κατανομή** $\mathcal{N}(\mu_1, \sigma^2)$ και $Y_1, Y_2, \dots, Y_{n_2}$ τ.δ. από **κανονική κατανομή** $\mathcal{N}(\mu_2, \sigma^2)$ με **άγνωστη διασπορά** σ^2 .

Θέτουμε

$$S_1^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (X_i - \bar{X})^2, \quad S_2^2 = \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (Y_i - \bar{Y})^2$$

και

$$S_p^2 = \frac{1}{n_1 + n_2 - 2} \left((n_1 - 1)S_1^2 + (n_2 - 1)S_2^2 \right).$$

Τότε

- Για τον έλεγχο της

$$H_0: \mu_1 \leq \mu_2 \text{ έναντι της } H_1: \mu_1 > \mu_2$$

σε **ε.σ. α**, απορρίπτουμε την H_0 όταν

$$T = \frac{\bar{X} - \bar{Y}}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \geq t_{n_1+n_2-2, \alpha},$$

όπου $t_{n_1+n_2-2, \alpha}$ το α -άνω ποσοστιαίο σημείο της $t_{n_1+n_2-2}$.

- Για τον έλεγχο της

$$H_0: \mu_1 \geq \mu_2 \text{ έναντι της } H_1: \mu_1 < \mu_2$$

σε **ε.σ. α**, απορρίπτουμε την H_0 όταν

$$T = \frac{\bar{X} - \bar{Y}}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \leq -t_{n_1+n_2-2, \alpha}.$$

- Για τον (αμφίπλευρο) έλεγχο της

$$H_0: \mu_1 = \mu_2 \text{ έναντι της } H_1: \mu_1 \neq \mu_2$$

σε **ε.σ. α**, απορρίπτουμε την H_0 όταν

$$|T| = \frac{|\bar{X} - \bar{Y}|}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \geq t_{n_1+n_2-2, \alpha/2}.$$

Έλεγχος για τη διαφορά δύο μέσων από οποιεσδήποτε κατανομές, για $n_1, n_2 \rightarrow \infty$

Έστω τα ανεξάρτητα τ.δ. $X_1, X_2, \dots, X_{n_1}$ και $Y_1, Y_2, \dots, Y_{n_2}$ από κατανομές F_1 και F_2 , αντίστοιχα, με $E(X_i) = \mu_1$, $E(Y_j) = \mu_2$ και $\text{Var}(X_i) = \sigma_1^2$, $\text{Var}(Y_j) = \sigma_2^2$. Επιπλέον, υποθέτουμε ότι **τα μεγέθη των δύο δειγμάτων n_1, n_2 είναι μεγάλα**. (Στην πράξη $n_1, n_2 \geq 30$.)

Τότε

- Για τον έλεγχο της

$$H_0: \mu_1 \leq \mu_2 \text{ έναντι της } H_1: \mu_1 > \mu_2$$

σε **ε.σ. α**, απορρίπτουμε την H_0 όταν

$$Z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \geq z_\alpha.$$

- Για τον έλεγχο της

$$H_0: \mu_1 \geq \mu_2 \text{ έναντι της } H_1: \mu_1 < \mu_2$$

σε **ε.σ. α**, απορρίπτουμε την H_0 όταν

$$Z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \leq -z_\alpha.$$

- Για τον (αμφίπλευρο) έλεγχο της

$$H_0: \mu_1 = \mu_2 \text{ έναντι της } H_1: \mu_1 \neq \mu_2$$

σε **ε.σ. α**, απορρίπτουμε την H_0 όταν

$$|Z| = \frac{|\bar{X} - \bar{Y}|}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \geq z_{\alpha/2},$$

όπου $z_{\alpha/2}$ το $\alpha/2$ -άνω ποσοστιαίο σημείο της τυποποιημένης κανονικής κατανομής.

Παράδειγμα

Οι βαθμολογίες 100 και 200 φοιτητών που διδάχθηκαν με τις μεθόδους Α και Β, αντίστοιχα, ήταν

$$A : X_1, X_2, \dots, X_{100}$$

$$B : Y_1, Y_2, \dots, Y_{200}$$

Από τα δεδομένα υπολογίσθηκε ότι:

$$\sum_{i=1}^{100} X_i = 780, \sum_{i=1}^{200} Y_i = 1400, \sum_{i=1}^{100} X_i^2 = 10000, \sum_{i=1}^{200} Y_i^2 = 20000.$$

Μπορούμε να συμπεράνουμε ότι η μέθοδος Α οδηγεί σε καλύτερα αποτελέσματα από τη Β, σε ε.σ. $\alpha = 5\%$;

$$\begin{aligned} S_1^2 &= \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (X_i - \bar{X})^2 = \frac{1}{n_1 - 1} \left(\sum_{i=1}^{n_1} X_i^2 - n_1 \bar{X}^2 \right) \\ &= \frac{1}{99} \left(\sum_{i=1}^{100} X_i^2 - 100 \cdot (7.8)^2 \right) = \frac{1}{99} (10000 - 6084) = 39.56 \\ S_2^2 &= \frac{1}{n_2 - 1} \left(\sum_{i=1}^{n_2} Y_i^2 - n_2 \bar{Y}^2 \right) \\ &= \frac{1}{199} (20000 - 200 \cdot 7^2) = \frac{1}{199} (20000 - 9800) = 51.26 \end{aligned}$$

Λύση

Έχουμε δύο ανεξάρτητα δείγματα και αν μ_1, μ_2 είναι οι άγνωστοι πληθυσμιακοί μέσοι των δύο μεθόδων Α, Β, αντίστοιχα, τότε η έκφραση «**καλύτερα αποτελέσματα για την Α**» σημαίνει $\mu_1 > \mu_2$.

Επομένως, θα ελέγξουμε την

$$H_0: \mu_1 \leq \mu_2 \text{ έναντι της } H_1: \mu_1 > \mu_2$$

σε ε.σ. $\alpha = 0.05$.

Από τα δεδομένα μας έχουμε ότι:

$$\bar{X} = \frac{\sum_{i=1}^{100} X_i}{100} = \frac{780}{100} = 7.8, \quad \bar{Y} = \frac{\sum_{i=1}^{200} Y_i}{200} = \frac{1400}{200} = 7,$$

Θα απορρίψουμε την H_0 όταν

$$Z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \geq z_\alpha \Leftrightarrow Z = \frac{7.8 - 7}{\sqrt{\frac{39.56}{100} + \frac{51.26}{200}}} \geq z_{0.05}$$

Άρα

$$Z = \frac{0.8}{0.807} = 0.991 \not\geq 1.645 = z_{0.05}.$$

Συνεπώς δεν μπορούμε να απορρίψουμε την H_0 , δηλ. δεν έχουμε αρκετά στοιχεία ώστε να αποφασίσουμε υπέρ της μεθόδου Α σε ε.σ. $\alpha = 0.05$.

Έλεγχος για τη διαφορά των ποσοστών από ανεξάρτητα δείγματα Bernoulli

Έστω τα ανεξάρτητα τ.δ. $X_1, X_2, \dots, X_{n_1}$ και $Y_1, Y_2, \dots, Y_{n_2}$ από κατανομές Bernoulli με άγνωστα ποσοστά p_1, p_2 , αντίστοιχα. Επιπλέον, υποθέτουμε ότι **τα μεγέθη των δύο δειγμάτων n_1, n_2 είναι μεγάλα**. (Στην πράξη $n_1, n_2 \geq 30$.) Τότε

- Για τον έλεγχο της

$$H_0: p_1 \leq p_2 \text{ έναντι της } H_1: p_1 > p_2$$

σε **ε.σ. α** , απορρίπτουμε την H_0 όταν

$$Z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\bar{X}(1-\bar{X})}{n_1} + \frac{\bar{Y}(1-\bar{Y})}{n_2}}} \geq z_\alpha.$$

Δ.Α. Μητσούδης Πιθανότητες & Στατιστική - Τμ. Ναυπηγών 41 / 53

- Για τον έλεγχο της

$$H_0: p_1 \geq p_2 \text{ έναντι της } H_1: p_1 < p_2$$

σε **ε.σ. α** , απορρίπτουμε την H_0 όταν

$$Z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\bar{X}(1-\bar{X})}{n_1} + \frac{\bar{Y}(1-\bar{Y})}{n_2}}} \leq -z_\alpha.$$

- Για τον (αμφίπλευρο) έλεγχο της

$$H_0: p_1 = p_2 \text{ έναντι της } H_1: p_1 \neq p_2$$

σε **ε.σ. α** , απορρίπτουμε την H_0 όταν

$$|Z| = \frac{|\bar{X} - \bar{Y}|}{\sqrt{\frac{\bar{X}(1-\bar{X})}{n_1} + \frac{\bar{Y}(1-\bar{Y})}{n_2}}} \geq z_{\alpha/2}.$$

Δ.Α. Μητσούδης Πιθανότητες & Στατιστική - Τμ. Ναυπηγών 42 / 53

Παράδειγμα

Πριν τη διαφημιστική εκστρατεία, 300 ψηφοφόροι θα ψήφιζαν ένα κόμμα σε σύνολο 1000 ερωτηθέντων. Μετά τη διαφημιστική εκστρατεία, σε ανεξάρτητο δείγμα 2000 ψηφοφόρων, οι 700 θα ψήφιζαν το κόμμα. Μπορούμε να πούμε ότι η διαφήμιση ήταν αποδοτική;

Δ.Α. Μητσούδης Πιθανότητες & Στατιστική - Τμ. Ναυπηγών 43 / 53

Λύση

Εδώ $X_1, X_2, \dots, X_{n_1}$ και $Y_1, Y_2, \dots, Y_{n_2}$ είναι ανεξάρτητα τ.δ από κατανομές Bernoulli με άγνωστα ποσοστά ψηφοφόρων p_1, p_2 , που θα ψήφιζαν το κόμμα αυτό πριν, και μετά, τη διαφημιστική εκστρατεία, αντίστοιχα.

Από τα δεδομένα μας έχουμε ότι:

$$\bar{X} = \frac{300}{1000} = 0.3, \quad \bar{Y} = \frac{700}{2000} = 0.35,$$

και $n_1 = 1000, n_2 = 2000$ (**πολύ μεγάλα!**).

Ο έλεγχος που θα πραγματοποιήσουμε είναι:

$$H_0: p_1 \geq p_2 \text{ έναντι της } H_1: p_1 < p_2$$

και ας επιλέξουμε επίπεδο σημαντικότητας $\alpha = 0.05$.

Δ.Α. Μητσούδης Πιθανότητες & Στατιστική - Τμ. Ναυπηγών 44 / 53

Παρατήρηση

Άρα θα απορρίψουμε την H_0 αν

$$Z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\bar{X}(1-\bar{X})}{n_1} + \frac{\bar{Y}(1-\bar{Y})}{n_2}}} \leq -z_\alpha$$
$$\Leftrightarrow Z = \frac{0.3 - 0.35}{\sqrt{\frac{(0.3) \cdot (0.7)}{1000} + \frac{(0.35) \cdot (0.65)}{2000}}} \leq -z_{0.05}$$
$$Z = -\frac{0.05}{0.018} = -2.77 \leq -1.645 = -z_{0.05} \quad \text{😊}$$

Επομένως, απορρίπτουμε την H_0 και είμαστε κατά 95% βέβαιοι ότι η διαφήμιση απέδωσε.

Από το προηγούμενο παράδειγμα είναι φανερό ότι **όσο μικραίνουμε το ε.σ. α τόσο δυσκολότερο γίνεται να απορριφθεί η H_0 .**

Αυτό είναι λογικό αφού γινόμαστε ολοένα και πιο «ανελαστικοί» αποδεχόμενοι μικρότερη πιθανότητα σφάλματος τύπου Ι.

Αν κάναμε τον έλεγχο σε ε.σ. $\alpha = 0.01$, το μόνο που θα άλλαζε θα ήταν η σταθερά $-z_\alpha = -z_{0.01} = -2.33$, και επειδή

$$Z = -2.77 \leq -2.33 = -z_{0.01}, \quad \text{😊}$$

θα απορρίπταμε και πάλι την H_0 .

Αν όμως κάναμε τον έλεγχο σε ε.σ. $\alpha = 0.001$, και πάλι το μόνο που θα άλλαζε θα ήταν η σταθερά $-z_\alpha = -z_{0.001} = -3.09$, αλλά τότε

$$Z = -2.77 \not\leq -3.09 = -z_{0.001}, \quad \text{😞}$$

οπότε σε αυτή την περίπτωση δεν θα μπορούσαμε να απορρίψουμε την H_0 .

Έλεγχοι υποθέσεων με χρήση της p -value

p -value

p -value (p -τιμή) ή παρατηρούμενο επίπεδο σημαντικότητας ενός ελέγχου, ονομάζεται η πιθανότητα να παρατηρηθεί μια τιμή της στατιστικής συνάρτησης ελέγχου τουλάχιστον τόσο ακραία όσο αυτή που έδωσε το δείγμα (προς την κατεύθυνση της H_1), δεδομένου ότι η H_0 είναι αληθής.

Η p -value ουσιαστικά μετράει το ενδεχόμενο οι διαφορές που βλέπουμε να είναι απλή σύμπτωση.

Για τον έλεγχο συγκρίνουμε την p -value με το προκαθορισμένο επίπεδο σημαντικότητας (α).

- 1. Αν $p \leq \alpha$, λέμε ότι η διαφορά είναι **στατιστικά σημαντική** (statistically significant) και απορρίπτουμε την H_0 ,
- 2. διαφορετικά, συμπεραίνουμε ότι δεν έχουμε επαρκή στοιχεία ώστε να απορρίψουμε την H_0 (στατιστικά μη σημαντικό αποτέλεσμα).

Παράδειγμα

Τα τσιγάρα που υπάρχουν αυτή τη στιγμή στην αγορά έχουν μέση περιεκτικότητα σε νικοτίνη τουλάχιστον 1.6 mg ανά τσιγάρο. Μια καπνοβιομηχανία ισχυρίζεται ότι χρησιμοποιεί μια καινούργια κατεργασία στα φύλλα καπνού που έχει ως αποτέλεσμα το κάθε τσιγάρο να έχει μέση περιεκτικότητα σε νικοτίνη λιγότερο από 1.6 mg. Για να ελεγχθεί ο ισχυρισμός αυτός λαμβάνεται και αναλύεται ένα δείγμα 20 τσιγάρων. Αν είναι γνωστό ότι η περιεκτικότητα σε νικοτίνη ανά τσιγάρο είναι κανονικά κατανομημένη με μέση τιμή μ και τυπική απόκλιση 0.8 mg, τι συμπεράσματα εξαγονται σε ε.σ. 5%, αν η μέση περιεκτικότητα σε νικοτίνη των 20 τσιγάρων του δείγματος είναι 1.54;

Λύση (με χρήση της p -value)

Στο παράδειγμα αυτό θέλουμε να αποφασίσουμε υπέρ του ισχυρισμού του κατασκευαστή μόνο αν έχουμε σημαντικά στοιχεία υπέρ του. Δηλαδή θα θεωρήσουμε τον ισχυρισμό αυτό ως την εναλλακτική υπόθεση. Άρα ο έλεγχος που θα πραγματοποιήσουμε είναι:

$$H_0: \mu \geq 1.6 \text{ έναντι της } H_1: \mu < 1.6$$

Επειδή είμαστε στην περίπτωση **ελέγχου μέσου κανονικής κατανομής με γνωστή διασπορά**, η στατιστική συνάρτηση ελέγχου παίρνει την τιμή:

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} = \frac{1.54 - 1.6}{\frac{0.8}{\sqrt{20}}} = -0.335.$$

Εδώ η p -value ορίζεται ως:

$$p\text{-value} = P(Z < -0.335) = \Phi(-0.335) = 1 - \Phi(0.335) = 0.3688.$$

Επειδή

$$p\text{-value} = 0.3688 > \alpha = 0.05$$

τα δεδομένα μας δεν επιτρέπουν την απόρριψη της H_0 , επομένως δεν αποδεχόμαστε τον ισχυρισμό του κατασκευαστή.

Παρατηρήστε ότι το ίδιο θα ίσχυε και για ε.σ. $\alpha = 0.001$, 0.01 ή 0.1.

p -value για έλεγχο μέσου κανονικής με γνωστή διασπορά

Αν $X_1, \dots, X_n$ είναι ένα τ.δ. από πληθυσμό $\sim \mathcal{N}(\mu, \sigma^2)$ με σ^2 γνωστό, τότε σε ε.σ. α έχουμε:

H_0	H_1	Συν. Ελέγχου (TS)	Χωρίς απόρριψης R	p -value αν $TS = t$
$\mu \leq \mu_0$	$\mu > \mu_0$	$\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$	$R = \{TS > z_\alpha\}$	$P(Z \geq t)$
$\mu \geq \mu_0$	$\mu < \mu_0$	$\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$	$R = \{TS < -z_\alpha\}$	$P(Z \leq t)$
$\mu = \mu_0$	$\mu \neq \mu_0$	$\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$	$R = \{ TS > z_{\alpha/2}\}$	$2P(Z \geq t)$

όπου Z είναι η τυποποιημένη κανονική κατανομή.

Ανάλογα αποτελέσματα μπορούν να διατυπωθούν και για τις υπόλοιπες περιπτώσεις που μέχρι τώρα εξετάσαμε με την παραδοσιακή μέθοδο των χωρίων απόρριψης.