Στατιστική συμπερασματολογία

Ορίζεται ως ο κλάδος της στατιστικής που πραγματεύεται την εξαγωγή νόμων, κανόνων και συμπερασμάτων τα οποία ξεπερνούν το επίπεδο των παρατηρήσεων και γενικεύονται στο σύνολο του πληθυσμού.

Βεβαία, για να είναι δυνατή η γενίκευση των συμπερασμάτων που προκύπτουν από ένα δείγμα σε ολόκληρο τον πληθυσμό, θα πρέπει το δείγμα να είναι όσο το δυνατόν πιο αντιπροσωπευτικό.

ΜΑΘΗΜΑ 9

ΣΤΑΤΙΣΤΙΚΗ ΥΠΟΘΕΣΗ

Ορίζεται ως οποιαδήποτε πρόταση στατιστικής φύσεως (περιγραφικά μέτρα, κατανομές των πληθυσμών, στοχαστικές διαδικασίες κλπ) την ορθότητα της οποίας επιθυμούμε να ελέγξουμε με βάση τα δείγματα που διαθέτουμε.

Έστω ότι μελετάμε μία παράμετρο Θ για την οποία υποθέτουμε πως μπορεί να λάβει μία συγκεκριμένη τιμή $\Theta = \Theta_0$. Αυτή η υπόθεση ονομάζεται μηδενική υπόθεση και συμβολίζεται με $\boxed{H_0}$ Επομένως: $H_0 : \Theta = \Theta_0$

Εάν η επεξεργασία των δεδομένων οδηγήσει στο συμπέρασμα πως η μηδενική υπόθεση ΔΕΝ ισχύει και κατά συνέπεια Θ θα πρέπει να απορριφθεί, θα ισχύει αναγκαστικά η εναλλακτική υπόθεση (2)

H_1 (εναλλακτική υπόθεση) η οποία μπορεί να λάβει μία από τις ακόλουθες μορφές:

$H_1: \theta \neq \theta_0$ (δίπλευρος έλεγχος)

$H_1: \theta < \theta_0$ (μονόπλευρος έλεγχος)

$H_1: \theta > \theta_0$ (μονόπλευρος έλεγχος)

Ο τύπος ελέγχου που θα επιλέξουμε εξαρτάται από το πρόβλημα.

Ανάλυση

Εάν η επεξεργασία των αποτελεσμάτων οδηγεί σε τιμές της θ κοντά στην θ_0 η υπόθεση H_0 γίνεται δεκτή ενώ εάν οι τιμές της θ είναι μακριά από την H_0 η πρόταση απορρίπτεται. Η περιοχή τιμών της θ για την οποία η H_0 απορρίπτεται ονομάζεται

περιοχή απόρριψης.

Με ποιον τρόπο λαμβάνεται η απόφαση σχετικά με την απόρριψη ή όχι της H_0? Χρησιμοποιούνται οι έννοιες του επιπέδου σημαντικότητας α και του επιπέδου εμπιστοσύνης $1-\alpha$. Το επίπεδο σημαντικότητας μπορεί να θεωρηθεί ως μία διαχωριστική γραμμή ανάμεσα στις περιοχές απόρριψης ή μη. Όταν η τιμή του α μικραίνει, το ίδιο συμβαίνει

και με την περιοχή απορριψης R.
Συνηθως επιλεγουμε α = 5% που οδηγει σε επιπεδο εμπιστοσυνης 95%.

Στατιστική ελεγχου υποθεσης

- Παραμετρικοί ελεγχοι: συνήθως στηριζονται στη χρήση της κανονικής κατανομής και κατα συνεπεια προκειμενου να αποφασισουμε αν θα χρησιμοποιήσουμε ή οχι τετοιους ελεγχους εξεταζουμε εαν οι παρατηρήσεις ακολουθουν ή οχι κανονική κατανομη.
- Μή παραμετρικοί ελεγχοι: στηριζονται σε αλλες ιδιοτητες της παραμετρου

Στάδια στατιστικού ελέγχου

- οριζεται η μηδενική υποθεση H_0
- οριζεται η εναλλακτική υποθεση H_1
- οριζεται το επιπεδο σημαντικότητας α
- επιλεγεται η καταλληλη συνάρτηση ελεγχου
- οριζεται η περιοχή απορριψης R της H_0

Σφάλματα που εμφανίζονται στον έλεγχο υποθέσεων

Σφάλμα τύπου I: ορίζεται ως η πιθανότητα απόρριψης της μηδενικής υπόθεσης ενώ αυτή είναι σωστή.

Σφάλμα τύπου II: ορίζεται ως η πιθανότητα αποδοχής της μηδενικής υπόθεσης ενώ αυτή είναι εσφαλμένη

Ισχύς στατιστικού ελέγχου ονομάζεται η πιθανότητα απόρριψης της H_0 ενώ αυτή είναι πράγματι ψευδής, δηλαδή η πιθανότητα να αποφύγουμε ένα σφάλμα τύπου II.

Επίπεδο σημαντικότητας ελέγχου ονομάζεται η πιθανότητα να παρατηρηθεί για τιμή του στατιστικού ελέγχου μεγαλύτερη από αυτή που ενώσε το δείγμα των παρατηρήσεων.

Ο έλεγχος (δοκιμασία) υποθέσεων μπορεί να εφαρμοσθεί για αρκετά στατιστικά μέτρα όπως είναι ο αριθμητικός μέσος μ, η διακύμανση σ^2 και η παράμετρος p της διωνυμικής κατανομής.

Δοκιμασία υποθέσεων για τη μέση τιμή μ

- $n > 30$, σ γνωστό
- Η δειγματική κατανομή $\bar{X}$ είναι κανονική

Διαδικασία

1ο Βήμα

- Περιγραφή των παραμέτρων του πληθυσμού
- Μηδενική υπόθεση $H_0: \mu = \mu_0$
- Εναλλακτική υπόθεση $H_1: \mu > \mu_0$ ή $\mu \neq \mu_0$
  $\mu < \mu_0$

2ο Βήμα

- Έλεγχος προϋποθέσεων
- Αναγνώριση κατανομής πιθανότητας και του στατιστικού ελέγχου:

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

- Καθορισμός του επιπέδου σημαντικότητας α

3ο Βήμα

- Συλλογή πληροφοριών από το δείγμα.
- Υπολογισμός της τιμής του στατιστικού ελέγχου

4ο Βήμα

- Υπολογισμός της κρίσιμης τιμής για δεδομένο α από τους πίνακες της κανονικής κατανομής
- Σύγκριση της τιμής Z με την κρίσιμη τιμή

5ο Βήμα

- Λήψη απόφασης σχετικά με την H_0

  Απορρίπτουμε την H_0 αν $Z > Z_\alpha$ ή $Z < -Z_\alpha$

  ή $|Z| > Z_{\alpha/2}$ ανάλογα με την

  εναλλακτική υπόθεση
- Συμπεράσματα σχετικά με την H_1

ΠΑΡΑΔΕΙΓΜΑ

Μια εταιρεία θέλει να ελέγξει τον βαθμό ικανοποίησης των εργαζόμενων μέσω ερωτηματολογίων. Στο πιο πρόσφατο δείγμα των 36 ατόμων το αποτέλεσμα ήταν $\bar{x} = 69$. Εάν κατά το παρελθόν ο μέσος βαθμός ικανοποίησης όλων των εργαζόμενων στην επιχείρηση ήταν $\mu = 68$ με τυπική απόκλιση

$\sigma = 4$, να ελεγχθεί η υπόθεση ότι ο βαθμός ικανοποίησης των εργαζόμενων άλλαξε. Θεωρήστε κανονική κατανομή με $\alpha = 0.05$

ΛΥΣΗ

Βήμα 1

- Θέλουμε να ελέγξουμε εάν ο βαθμός ικανοποίησης των εργαζομένων έχει αλλάξει σε σχέση με τον παρελθόντα χρόνο.
- Μηδενική υπόθεση: $H_0 : \mu = 68$
- Εναλλακτική υπόθεση : $H_1 : \mu \neq \mu_0$

Βήμα 2

- Κανονική κατανομή, μεγάλο δείγμα, γνωστό σ
- $Z = \dfrac{(\bar{X} - \mu_0)\sqrt{n}}{\sigma}$
- $\alpha = 0.05$ (επίπεδο σημαντικότητας)

Βήμα 3

- Μέγεθος δείγματος $n = 36$, $\bar{X} = 69$
- $Z = \dfrac{(\bar{X} - \mu_0)\sqrt{n}}{\sigma} = \dfrac{(69 - 68)\sqrt{36}}{4} = 1.5$

Βήμα 4

- Για $\alpha = 0.05$ και δίπλευρο έλεγχο, είναι

$$Z_{\alpha/2} = Z_{0.025} = 1.96 \text{ (από πίνακες)}.$$

• $|Z| = 1.5 < 1.96 = Z_{0.025}$

Βήμα 5

• Δεχόμαστε τη μηδενική υπόθεση
• ο βαθμός ικανοποίησης <u>ΔΕΝ</u> έχει αλλάξει.

<u>$n > 30$ και άγνωστο σ</u>

Εφαρμόζουμε όπως και πριν αλλά χρησιμοποιούμε τον τύπο

$$Z = \frac{(\bar{x} - \mu_0)\sqrt{n}}{s}$$

<u>$n \leq 30$ και άγνωστο σ</u>

Χρησιμοποιούμε την τυπική απόκλιση του δείγματος για να εκτιμήσουμε την αντίστοιχη του πληθυσμού και στη συνέχεια τον τύπο

$$t = \frac{(\bar{x} - \mu_0)\sqrt{n}}{s}$$ με $n-1$ βαθμούς ελευθερίας.

Άρα έχουμε κατανομή t και η H_0 απορρίπτεται αν

$t > t_\alpha$ ή $t < -t_\alpha$ ή $|t| > t_{\alpha/2}$

αναλόγως με την εναλλακτική υπόθεση.

Έλεγχος υποθέσεων για τη διαφορά των μέσων δύο δειγμάτων

Ανεξάρτητα μεγάλα δείγματα ($n_1 > 30$, $n_2 > 30$)

Βήμα 1

- Περιγραφή των παραμέτρων μ_1 και μ_2 των πληθυσμών
- Καθορισμός των υποθέσεων H_0 και H_1

H_0: $\mu_1 = \mu_2$ ή $\mu_1 - \mu_2 = 0$

H_1: $\mu_1 > \mu_2$ ή $\mu_1 - \mu_2 > 0$

$\mu_1 < \mu_2$ ή $\mu_1 - \mu_2 < 0$

$\mu_1 \neq \mu_2$ ή $\mu_1 - \mu_2 \neq 0$.

Βήμα 2

- Έλεγχος προϋποθέσεων
- Αναγνώριση της κατανομής πιθανότητας και του στατιστικού TEST που θα χρησιμοποιηθεί.

Εδώ θα είναι

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

Αν δεν είναι γνωστές οι διακυμάνσεις των πληθυσμών χρησιμοποιούμε τις διακυμάνσεις των αντίστοιχων δειγμάτων όπου είναι

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

- Καθορισμός του επιπέδου σημαντικότητας α

Βήμα 3

- Συλλογή πληροφοριών από τα δείγματα
- Υπολογισμός της τιμής του στατιστικού test

Βήμα 4

- Υπολογισμός της κρίσιμης τιμής για δεδομένο α από τους πίνακες της κατανομής
- Σύγκριση της τιμής Z με την κρίσιμη τιμή

Βήμα 5

- Λήψη απόφασης σχετικά με την H_0

  Απορρίπτουμε την H_0 αν $Z > Z_\alpha$ ή $Z < -Z_\alpha$ ή $|Z| > Z_{\alpha/2}$ ανάλογα με την εναλλακτική υπόθεση

- Εξάγουμε συμπεράσματα σχετικά με την H_1

ΠΑΡΑΔΕΙΓΜΑ

Έστω δύο δείγματα μαθητών φροντιστηρίου

ΔΕΙΓΜΑ Α $n_1 = 31, \ \bar{x}_1 = 82.5, \ s_1^2 = 64$

ΔΕΙΓΜΑ Β $n_2 = 40, \ \bar{x}_2 = 78, \ s_2^2 = 100$

$\alpha = 0.05$

H_0: ΔΕΝ υπάρχει διαφορά στη μέση επίδοση του συνόλου των μαθητών των 2 φροντιστηρίων.

Μηδενική υπόθεση $H_0: \mu_1 = \mu_2$ ή $\mu_1 - \mu_2 = 0$

Εναλλακτική υπόθεση $H_1: \mu_1 \neq \mu_2$ ή $\mu_1 - \mu_2 \neq 0$.

n_1, n_2 μεγάλα δείγματα με κανονική κατανομή

$\alpha = 0.05$.

Δεν είναι γνωστά τα σ_1^2 και σ_2^2 άρα χρησιμοποιούμε τα s_1^2 και s_2^2.

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(82.5 - 78) - 0}{\sqrt{\frac{64}{31} + \frac{100}{40}}} = 2.1$$

$\alpha = 0.05 \Rightarrow Z_{0.025} = 1.96$

$|Z| = 2.1 > 1.96 = Z_{0.025}$:

$|Z| > Z_{\alpha/2} \rightarrow H_0$ απορρίπτεται.

Ανεξάρτητα μικρά δείγματα

$n_1 \leq 30$ και / ή $n_2 \leq > 30$

Χρησιμοποιείται κατανομή t με $n_1 + n_2 - 2$ βαθμούς ελευθερίας

Ισχύουν τα ίδια με τη διαφορά

Βήμα 2Β Χρησιμοποιείται η κανονική κατανομή

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad \text{όπου}$$

$$S^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

η εκτίμηση της κοινής διακύμανσης των 2 πληθυσμών

Έστω πρόγραμμα που βοηθά τους αναλυτές να ελαττώσουν το χρόνο ανάπτυξης εφαρμογών.

Δείγμα $n_1 = 12$ (παλαιά τεχνολογία) $\quad \alpha = 0.05$

$n_2 = 12$ (νέα τεχνολογία)

$\bar{X}_1 = 325$ ώρες, $S_1 = 40$ ώρες

$\bar{X}_2 = 288$ ώρες, $S_2 = 44$ ώρες

H_0: το νέο πρόγραμμα να είναι καλύτερο

Μικρά δείγματα με ίσες διακυμάνσεις

$$s^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} =$$

$$= \frac{(12-1)\,40^2 + (12-1)\,44^2}{12+12-2} = 1.768$$ Ομοίως

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(325-288)-0}{\sqrt{1.768\left(\frac{1}{12}+\frac{1}{12}\right)}} =$$

$$= 2.16$$

Από τους πίνακες της t-κατανομής:

$$t_{0.05} = 1.717$$

$$t = 2.16 > 1.717$$

Άρα Απορρίπτουμε την μηδενική H_0 και αποδεχόμαστε την εναλλακτική H_1

Εξαρτημένα δείγματα

$$\bar{d} = \frac{1}{n}\sum d_i$$

Τώρα είναι

$$t = \frac{(\bar{d} - \mu_d)\sqrt{n}}{\sigma_d}$$

όπου

$$s_d = \sqrt{\frac{\sum d_i^2 - \left[\frac{(\sum d_i)^2}{n}\right]}{n-1}}$$

Δοκιμασία υποθέσεων για την αναλογία p του πληθυσμού.

Τα δεδομένα είναι ασυνεχή και ως εκ τούτου τυπικά απαιτείται η χρήση της διωνυμικής κατανομής. Ωστόσο όταν το μέγεθος του δείγματος είναι αρκετά μεγάλο, η διωνυμική κατανομή προσεγγίζει την κανονική κατανομή. Ειδικότερα, για να αντικατασταθεί η διωνυμική με την κανονική κατανομή θα πρέπει να πληρούνται οι επόμενες συνθήκες: $n > 20$, $np > 5$, $nq > 5$

Μηδενική υπόθεση

$$H_0 : p = p_0$$

Εναλλακτική υπόθεση:

$$H_1 : p \neq p_0 \quad (\text{διπλευρος έλεγχος})$$

$$\left.\begin{array}{l} p < p_0 \\ p > p_0 \end{array}\right\} \text{μονόπλευρος έλεγχος.}$$

Ορίζουμε

$$Z = \frac{(\bar{p} - p_0)}{\sigma_{\bar{p}}} \quad \text{όπου} \quad \bar{p} = \frac{x}{n},$$

x: πλήθος επιτυχιών
n: μέγεθος δείγματος

και $\sigma_{\bar{p}} = \sqrt{\dfrac{p_0(1-p_0)}{n}}$

Επίσης ορίζουμε το <u>επίπεδο</u> σημαντικότητας α.

Όπως και πριν, συλλέγουμε πληροφορίες από το δείγμα και υπολογίζουμε την τιμή του στατιστικού ελέγχου. Στη συνέχεια υπολογίζουμε την κρίσιμη τιμή και συγκρίνουμε τις δύο τιμές.

Η H_0 απορρίπτεται εάν είναι (ανάλογα με την περίπτωση)

$$Z > Z_\alpha \quad \text{ή} \quad Z < -Z_\alpha$$

$$\text{ή} \quad |Z| > Z_{\alpha/2}$$

Διαδικασία υποθέσεων για τη διαφορά αναλογιών δύο πληθυσμών

Εάν <u>ΔΕΝ</u> υπάρχει διαφορά μεταξύ των αναλογιών P_1 και P_2 των 2 πληθυσμών ενώ οι δειγματικές κατανομές είναι κανονικές ή κατά προσέγγιση κανονικές, το στατιστικό test που μπορούμε να εφαρμόσουμε είναι το

$$Z = \frac{(\bar{p}_1 - \bar{p}_2) - (p_1 - p_2)}{\sigma_{\bar{p}_1 - \bar{p}_2}}$$ όπου

$$\sigma_{\bar{p}_1 - \bar{p}_2} = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$

Επειδή συνήθως τα p_1 και p_2 δεν είναι γνωστά, προτιμότερον να υπολογίσουμε το τυπικό σφάλμα των διαφορών μεταξύ των δύο αναλογιών υπολογίζουμε με την χρήση των αναλογιών $\bar{p}_1$ και $\bar{p}_2$ των δύο δειγμάτων την κοινή αναλογία

$$\bar{p} = \frac{n_1 \bar{p}_1 + n_2 \bar{p}_2}{n_1 + n_2}$$

Άρα έχουμε την εκτίμηση του τυπικού σφάλματος

$$S_{\bar{p}_1 - \bar{p}_2} = \sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

οπότε

$$Z = \frac{(\bar{p}_1 - \bar{p}_2) - (p_1 - p_2)}{S_{\bar{p}_1 - \bar{p}_2}}$$

Τα υπόλοιπα είναι είναι ίδια με πριν.

ANOVA Analysis of Variance

1

Έλεγχος υπόθεσης για τη σύγκριση των μέσων τιμών 2 πληθυσμών.

Τι γίνεται με τις μέσες τιμές πολλών πληθυσμών?

Έστω K ανεξάρτητα τυχαία δείγματα (ένα δείγμα από κάθε πληθυσμό) που προέρχονται από K κανονικούς πληθυσμούς με

17

$$\sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$$

Τα ζεύγη που δημιουργούνται και πρέπει να ελεγχθούν είναι $\frac{k(k-1)}{2}$

ANOVA $\longrightarrow$ Επέκταση t-test

Null hypothesis: $H_0: \mu_1 = \mu_2 = \dots = \mu_k$

Εναλλακτική υπόθεση:

Τουλάχιστον δύο μέσες τιμές

ΔΙΑΦΕΡΟΥΝ

Προϋποθέσεις εφαρμογής ANOVA

1) Τυχαία δείγματα
2) Κανονικές κατανομές
3) Ίσες διακυμάνσεις πληθυσμών

Το <u>ANOVA</u> στηρίζεται στη σύγκριση δύο διαφορετικών εκτιμήσεων της διακύμανσης σ^2 του συνολικού πληθυσμού.

Περιλαμβάνει 3 Βήματα

- Καθορισμός μιας εκτίμησης διακύμανσης του πληθυσμού από τη διακύμανση μεταξύ των μέσων των δειγμάτων
- Καθορισμός μιας δεύτερης εκτίμησης της διακύμανσης του πληθυσμού από τη διακύμανση μέσα στα δείγματα
- Σύγκριση των τιμών των δύο εκτιμήσεων. Αν αυτές είναι περίπου ίσες γίνεται δεκτή η μηδενική υπόθεση H_0.

ΑΝΑΛΥΣΗ

Έστω k δείγματα με μέγεθος $n_1, n_2, \ldots, n_k$

Έστω $\bar{X}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} X_{ij}$ ο μέσος όρος του δείγματος j

$$\bar{\bar{X}} = \frac{1}{N} \sum_{j=1}^{k} \sum_{i=1}^{n_j} X_{ij}$$ ο γενικός μέσος όρος

όπου $N = n_1 + n_2 + \cdots + n_k$

Η ολική μεταβλητότητα στα δεδομένα x_{ij} μπορεί να μετρηθεί με το ολικό άθροισμα τετραγώνων (Total Sum of Squares)

$$SST = \sum_{j=1}^{k} \sum_{i=1}^{n_j} (x_{ij} - \bar{\bar{x}})^2 \qquad (19)$$

Η μεταβλητότητα μεταξύ των δειγμάτων ή το άθροισμα τετραγώνων μεταξύ των δειγμάτων είναι

$$SSTR = \sum_{j=1}^{k} \sum_{i=1}^{n_j} (\bar{x}_j - \bar{\bar{x}})^2$$

Η μεταβλητότητα μέσα στα δείγματα ή το άθροισμα των τετραγώνων των σφαλμάτων είναι

$$SSE = \sum_{j=1}^{k} \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)^2$$

Γενικά είναι

$$\boxed{SST = SSTR + SSE}$$

Επομένως

Η ολική μεταβλητότητα αναλύεται σε δύο συνιστώσες: στη μεταβλητότητα μεταξύ των δειγμάτων που οφείλεται στη μέση διαφορά των τιμών της που προέρχονται από τα διαφορετικά δείγματα και στη μεταβλητότητα μέσα στα δείγματα η οποία αντιπροσωπεύει την τυχαία μεταβλητότητα (πειραματικό σφάλμα).

Βαθμοί ελευθερίας

SST : $N-1$

SSTR : $k-1$

SSE : $N-k$

Στατιστική συνάρτηση ελέγχου

$$f = \frac{MSTR}{MSE} = \frac{SSTR}{k-1} \cdot \frac{N-k}{SSE}$$

Άρα η κρίσιμη περιοχή απόρριψης είναι η

$$C = \{ f : f > F_{\alpha, k-1, N-k} \}$$

F κατανομή : πίνακας F του παραρτήματος

ΠΑΡΑΔΕΙΓΜΑ

Έστω A, B, C, D 4 διαφορετικοί τύποι CPU

Χρόνοι επισκευών (συνολικά 28 καταγραφές) $\boxed{\alpha = 0.01}$

A: 62, 48, 49, 50, 63, 69, 57

B: 45, 57, 44, 40, 39, 54, 50

C: 66, 53, 47, 49, 59, 60, 67

D: 68, 52, 70, 55, 65, 53, 59

Είναι

SSTR = 700.7

SST = 2084,4

$H_0 : \mu_A = \mu_B = \mu_C = \mu_D$ H_1: Τουλάχιστον 2 μ διαφέρ[ουν]

SSE = SST − SSTR = 1383,7

$k-1 = 4-1 = 3$

$N-k = 28-4 = 24$

Άρα $f = 4.05 < 4.72 = F_{0.01,3,24}$

$\Rightarrow H_0$ ΔΕΝ απορρίπτεται