

## Έλεγχος για ισότητα μέσων τιμών: ζευγαρωτές παρατηρήσεις

Σε μια μελέτη για το αποτέλεσμα της έκθεσης των ανθέων του τριφυλλίου σε διαφορετικές περιβαλλοντικές συνθήκες, διαλέχθηκαν 10 υγιή φυτά με άνθη ελεύθερα εκτεθειμένα στην κορυφή και άνθη όσο το δυνατόν πιο καλά κρυμμένα στο κάτω μέρος. Στη συνέχεια καταμετρήθηκαν οι αριθμοί των σπόρων, που βρέθηκαν να είναι

ΦΥΤΟ	1	2	3	4	5	6	7	8	9	10
Άνθη στην κορυφή	4,0	5,2	5,7	4,2	4,8	3,9	4,1	3,0	4,6	6,8
Άνθη στο κάτω μέρος	4,4	3,7	4,7	2,8	4,2	4,3	3,5	3,7	3,1	1,9

Με βάσει τα παραπάνω δεδομένα βρείτε ένα 90% διάστημα εμπιστοσύνης για την πραγματική διαφορά  $\mu_1 - \mu_2$  στον αριθμό των σπόρων του τριφυλλίου που προέρχεται από λουλούδια στο πάνω και στο κάτω μέρος του φυτού.

## $\chi^2$ έλεγχος: καλής προσαρμογής σε θεωρητικό μοντέλο

$H_0$ : The sample comes from a population having a 9 : 3 : 3 : 1 ratio of yellow-smooth to yellow-wrinkled to green-smooth to green-wrinkled seeds.

$H_1$ : The sample comes from a population not having a 9 : 3 : 3 : 1 ratio of the above four seed phenotypes.

	<i>Yellow smooth</i>	<i>Yellow wrinkled</i>	<i>Green smooth</i>	<i>Green wrinkled</i>	<i>n</i>
$f_i$	152	39	53	6	250

$H_0$ :  $p_1=9/(9+3+3+1)$ ,  $p_2=3/16$ ,  $p_3=3/16$ ,  $p_4=1/16$

$H_1$ : κάποιο διαφέρει

$H_0$ :  $p_1=0.5625$ ,  $p_2=0.1875$ ,  $p_3=0.1875$ ,  $p_4=0.0625$

$H_1$ : κάποιο διαφέρει

## $\chi^2$ έλεγχος: ομοιογένειας και έλεγχος ποσοστού

- Σε τέσσερις πόλεις Α, Β, Γ, Δ ρωτήσαμε από 200 άτομα εάν προτιμούν την χρήση του €. Μπορούμε να συμπεράνουμε ότι το ποσοστό του ΝΑΙ είναι το ίδιο (ομοιόγενές) και για τις τέσσερις πόλεις;

κ/π/ν \ πόλη	Α	Β	Γ	Δ	σύνολο
ΝΑΙ	78 (59)	53 (59)	59 (59)	48 (59)	236
ΟΧΙ	124 (141)	147 (141)	141 (141)	152 (141)	564
σύνολο	200	200	200	200	800

β. Να ελέγξετε εάν σε επίπεδο σημαντικότητας  $\alpha=10\%$  δεχόμαστε την υπόθεση  $H_0 : p_N = 0.5$  έναντι της  $H_1 : p_N \neq 0.5$ .

γ. δώστε τα τομεογράμματα εμφανίζοντας σε αυτά τα ποσοστά του ναι και του όχι, για τις τέσσερις διαφορετικές πόλεις

# Απλό γραμμικό μοντέλο

Στα δεδομένα που ακολουθούν, δίδονται η ηλικία (σε ημέρες) και τα μήκη των φτερών (σε cm) για 13 σπουργίτες την ανάπτυξη των οποίων μελετούν μια ομάδα μεταπτυχιακών φοιτητών του τομέα Ζωολογίας του Τμήματος Βιολογίας.

Ηλικία (σε ημέρες) X	3	4	5	6	8	9	10	11	12	14	15	16	17
Μήκος φτερών (σε cm) Y	1.4	1.5	2.2	2.4	3.1	3.2	3.2	3.9	4.1	4.7	4.5	5.2	5

- α. Εάν  $\mu_Y$  και  $\sigma_Y^2$  είναι αντίστοιχα το μέσο και η διασπορά του μήκους των φτερών των σπουργιτών, να δοθεί εκτίμηση για αυτές τις παραμέτρους:  $\widehat{\mu}_Y = \underline{\hspace{2cm}}$   $\widehat{\sigma}_Y^2 = \underline{\hspace{2cm}}$
- β. Να δοθούν οι τιμές των παραμέτρων της καμπύλης παλινδρόμησης του μήκους των φτερών των σπουργιτών στην ηλικία τους:  $\widehat{\beta}_0 = \underline{\hspace{2cm}}$  και  $\widehat{\beta}_1 = \underline{\hspace{2cm}}$ . Ποια είναι η τιμή του συντελεστή προσδιορισμού του μοντέλου  $\underline{\hspace{2cm}}$  και τι συμπέρασμα βάζετε από αυτήν  $\underline{\hspace{4cm}}$ ;
- γ. Δώσετε το ιστόγραμμα συχνοτήτων για τα κατάλοιπα που προκύπτουν από το μοντέλο της παλινδρόμησης σχεδιάζοντας σε αυτό και τη συνάρτηση πυκνότητας της κατανομής που πιστεύετε ότι αυτά ακολουθούν.
- δ. Μπορούμε να δεχτούμε την υπόθεση  $H_0 : \beta_1 = 0$  έναντι της  $H_1 : \beta_1 \neq 0$  σε επίπεδο σημαντικότητας (ε.σ.) 10%;  $\underline{\hspace{2cm}}$  Γιατί;  $\underline{\hspace{4cm}}$
- ε. Ποια υπόθεση χρειάστηκε να κάνετε στο ερώτημα δ  $\underline{\hspace{4cm}}$ , και σε ποιο ε.σ. έγινε αυτή δεκτή  $\underline{\hspace{2cm}}$ ;
- στ. Να κατασκευάσετε διάστημα εμπιστοσύνης σ.ε. 90% για τη μέση πρόβλεψη του μήκους των φτερών ενός σπουργίτη ηλικίας 7 ημερών:  $(\underline{\hspace{2cm}}, \underline{\hspace{2cm}})$ .

# Ανάλυση διασποράς κατά έναν παράγοντα

Είκοσι εννέα νεογέννητα γουρουνάκια (με περίπου το ίδιο βάρος) μοιράστηκαν τυχαία σε τέσσερις διαφορετικές ομάδες, για την καθεμία εκ των οποίων επιλέχθηκε διαφορετική μέθοδος διατροφής. Μετά από κάποιο χρονικό διάστημα, τα βάρη των ζώων (σε κιλά) μετρήθηκαν και δίδονται στον παρακάτω πίνακα:

Ομάδα 1	Ομάδα 2	Ομάδα 3	Ομάδα 4
60.8	68.7	69.6	61.9
67.0	67.7	77.1	64.2
65.0	75.0	75.2	63.1
68.6	73.3	71.5	66.7
61.7	71.8	73.4	60.3
59.9	78.2		59.8
69.3	68.5		63.0
64.6	71.8		62.7

- α. Να ελεγχθεί, εάν το μέσο βάρος των γουρουνιών είναι ίδιο στις τέσσερις διαφορετικές ομάδες. Να γράψετε το μοντέλο, στο οποίο βασιστήκατε για τον έλεγχο αυτόν, για τις δύο πρώτες παρατηρήσεις όλων των ομάδων.

## Ανάλυση διασποράς κατά έναν παράγοντα (συνέχεια...)

- β. Μπορούμε να θεωρήσουμε, ότι η διασπορά του βάρους είναι η ίδια για τις τέσσερις διαφορετικές ομάδες; Σε περίπτωση αρνητικής απάντησης, είναι σωστό το συμπέρασμα στο ερώτημα α.;
- γ. Να ελέγξετε τις υποθέσεις  $H_0 : \mu_2 = \mu_3$  έναντι της  $H_1 : \mu_2 \neq \mu_3$ ,  $H_0 : \mu_1 \stackrel{(<)}{=} \mu_2$  έναντι της  $H_1 : \mu_1 > \mu_2$  και την  $H_0 : \mu_1 \stackrel{(>)}{=} \mu_4$  έναντι της  $H_1 : \mu_1 < \mu_4$ . Σε ποια ε.σ. γίνονται αποδεκτές καθεμία από αυτές;
- δ. Ποιες είναι οι θεωρητικές προϋποθέσεις για τους ελέγχους των ερωτημάτων α. και γ.; Πληρούνται για αυτά τα δεδομένα και ποιο είναι το συμπέρασμά σας για το κατά πόσο “διαφέρουν” τελικά οι δίαιτες που εφαρμόστηκαν;