

Περιγραφική Στατιστική (MEM-205)

Φυλλάδιο ασκήσεων (Μέρος 2/3)

Προθεσμία παράδοσης: 16-05-2020

Τρόπος παράδοσης: Με email στο kesmarag@uoc.gr με θέμα ΦΑ2 - AM

Οδηγίες

Θεωρείστε ως x, y, z, w τα ψηφία του αριθμού μητρώου σας AM

$$AM = 1000x + 100y + 10z + w$$

Στη συνέχεια θεωρείστε τις παραμέτρους A, B, C, D ως

$$A = x + 1$$

$$B = y + 1$$

$$C = z + 1$$

$$D = w + 1$$

Οι παράμετροι A, B, C, D θα χρησιμοποιούνται ως δεδομένα σε ασκήσεις του φυλλαδίου.

Παράδειγμα

$$AM = 7359$$

- $A = 7 + 1 = 8$
- $B = 3 + 1 = 4$
- $C = 5 + 1 = 6$
- $D = 9 + 1 = 10$

Άσκηση 1

Για το σύνολο δεδομένων

$$\{(10, 10 + A), (20, 20 - B), (30, 30 + C), (40, 40 - D)\}$$

υπολογίστε την ευθεία γραμμικής παλινδρόμησης χρησιμοποιώντας τη μέθοδο ελαχίστων τετραγώνων.

Άσκηση 2

Το σύνολο δεδομένων

$$\{(x_1, y_1), \dots, (x_N, y_N)\}$$

έχει συντελεστή γραμμικής συσχέτισης r_1 .

Για $\alpha \in \mathbb{R}$ εκφράστε τον συντελεστή γραμμικής συσχέτισης r_2 του συνόλου δεδομένων

$$\{(x_1, y_1 + \alpha), \dots, (x_N, y_N + \alpha)\}$$

συναρτήσει του r_1 .

Άσκηση 3

Για το σύνολο δεδομένων

$$\{(1, A, 1), (2, B, 3), (3, C, 3), (5, D, 4)\}$$

βρείτε χρησιμοποιώντας τη μέθοδο ελαχίστων τετραγώνων την εκτιμήτρια πολλαπλής γραμμικής παλινδρόμησης.

Άσκηση 4

Εφαρμόζοντας το μοντέλο γραμμική παλινδρόμησης για το παρακάτω σύνολο δεδομένων με 101 στοιχεία

$$\{j, Bj + (-1)^j D\}_{j=1}^{101}$$

υπολογίστε το 95% διάστημα εμπιστοσύνης για τα $\mu_{y|A}, \mu_{y|C}$ καθώς και για τις μεμονωμένες τιμές της y για $x = A$ και $x = C$.

Άσκηση 5

Δίνονται οι παρακάτω πληροφορίες για ένα σύνολο δεδομένων 50 στοιχείων:

$$\hat{y} = -C + 3x, s_e = A, SS_{xx} = 4D, \bar{X} = C$$

α) Υπολογίστε το 95% διάστημα εμπιστοσύνης για την τιμή της εξαρτημένης μεταβλητής για $x = A + C$.

β) Μπορεί να εκτιμηθεί το αντίστοιχο διάστημα εμπιστοσύνης εάν προστεθεί στο σύνολο δεδομένων η παρατήρηση $(C, 2C)$; Θα μεταβληθεί η ευθεία γραμμικής παλινδρόμησης;

Άσκηση 6

Επιλέξτε ένα κατάλληλο γενικευμένο μοντέλο γραμμικής παλινδρόμησης και βρείτε την εκτιμήτρια καμπύλη για το παρακάτω σύνολο δεδομένων:

$$\{(40, 16 + A/10), (50, 25 - B/10), (60, 36 + C/10), (70, 49 - D/10)\}$$

Ποια εκτίμηση θα δίνετε για την τιμή της εξαρτημένης μεταβλητής εάν $x = 45$;

Άσκηση 7

Για τις 6 ταινίες ενός σκηνοθέτη δίνονται τα παρακάτω στοιχεία

	Κόστος Παραγωγής(εκ.δολ.)	Είδος Ταινίας	Βαθμολογία την πρώτη εβδομάδα προβολής (0-10)	Έσοδα μετά το πρώτο έτος(εκ.δολ.)
#1	A	Κωμωδία	$\max(B,5)$	$2A+C$
#2	B	Κωμωδία	$\max(C,6)$	$2B+D$
#3	C	Περιπέτεια	$\max(D,7)$	$8C$
#4	D	Περιπέτεια	$\max(B,8)$	$10D$
#5	12-A	Κωμωδία	$\max(C,4)$	$3A$
#6	12-B	Περιπέτεια	$\max(A,7)$	$20B$

Κατασκευάστε κατάλληλο μοντέλο γραμμικής παλινδρόμησης για την εκτίμηση των εσόδων χρησιμοποιώντας τα διαθέσιμα δεδομένα. Καταστρώστε το σύστημα εξισώσεων που δίνει την εκτιμήτρια για την εξαρτημένη μεταβλητή χωρίς να το επιλύσετε.

Υπόδειξη: Χρησιμοποιήστε ψευδομεταβλητές όπου χρειάζεται.

z and t tables - <https://www.math.arizona.edu/~rsims/ma464/standardnormaltable.pdf> - <https://www.sjsu.edu/faculty/gerstman/StatPrimer/t-table.pdf>