

## Συσχέτιση και Παλινδρόμηση

1. Στον παρακάτω πίνακα δίνεται για 10 σταθμούς ο αριθμός των ημερών σε ένα χρόνο που η θερμοκρασία έπεσε κάτω από  $0^{\circ}\text{C}$  και το υψόμετρο τους.

Υψόμετρο (μ)	1000	1050	1110	1220	1320	1380	1420	1560	1670	1950
Αριθμός ημερών	32	29	36	38	43	53	52	63	73	100

- (α) Σχηματίστε το κατάλληλο διάγραμμα διασποράς, υπολογίστε τον συντελεστή συσχέτισης και σχολιάστε αν φαίνεται ο αριθμός των ημερών να εξαρτάται γραμμικά από το υψόμετρο.
- (β) Υπολογίστε τις σημειακές εκτιμήσεις  $a$  και  $b$  των παραμέτρων  $\alpha$  και  $\beta$  της ευθείας παλινδρόμησης (με τη μέθοδο των ελαχίστων τετραγώνων) και σχηματίστε στο παραπάνω διάγραμμα διασποράς την ευθεία ελαχίστων τετραγώνων.
- (γ) Υπολογίστε τη διασπορά των σφαλμάτων παλινδρόμησης και σχηματίστε στο διάγραμμα διασποράς ενδεικτικά όρια του τυπικού σφάλματος της εκτίμησης της ευθείας ελαχίστων τετραγώνων
- (δ) Με βάση το δείγμα, μπορείτε να εκτιμήσετε το μέσο αριθμό ημερών το χρόνο που η θερμοκρασία πέφτει κάτω από  $0^{\circ}\text{C}$  σε υψόμετρο 1500μ; Σε υψόμετρο 2200μ;
2. Σε περιοχές με υψηλά επίπεδα βροχοπτώσεων είναι σημαντικό να έχει το σκυροδέμα χαμηλό πορώδες, το οποίο μετριέται με το ποσοστό κενού στο υλικό. Ένας παράγοντας που επηρεάζει το πορώδες του σκυροδέματος είναι το βάρος μονάδας του σκυροδέματος που μετριέται σε pcf. Για κάθε ένα από 11 δοκίμια σκυροδέματος που συλλέχτηκαν τυχαία μετρήθηκαν το βάρος μονάδας και το πορώδες και οι παρατηρήσεις δίνονται στον παρακάτω πίνακα.

βάρος μονάδας	99.0	101.1	102.7	105.4	108.7	110.8	112.1	113.6	115.1	115.4	120.0
πορώδες	28.8	27.9	27.0	22.8	20.9	19.6	17.1	16.0	13.0	13.6	10.8

- (α) Υποθέτουμε ότι το πορώδες του σκυροδέματος  $Y$  εξαρτάται γραμμικά από το βάρος μονάδας του σκυροδέματος  $X$  ( $E(Y|X = x) = \alpha + \beta x$ ). Σχηματίστε το κατάλληλο διάγραμμα διασποράς και σχολιάστε αν αυτή η υπόθεση φαίνεται σωστή με βάση το δείγμα των παρατηρήσεων του πίνακα.
- (β) Υπολογίστε τις σημειακές εκτιμήσεις  $a$  και  $b$  των παραμέτρων  $\alpha$  και  $\beta$  της ευθείας παλινδρόμησης και σχηματίστε την ευθεία αυτή στο διάγραμμα διασποράς.
- (γ) Με βάση το δείγμα, εκτιμήστε (αν γίνεται) το μέσο πορώδες του σκυροδέματος που έχει βάρος μονάδας 110 και το ίδιο για βάρος μονάδας 130. Σχολιάστε με βάση τη διασπορά των σφαλμάτων παλινδρόμησης το τυπικό σφάλμα γύρω από όποια πρόβλεψη κάνετε.
3. Στα πλαίσια μιας έρευνας για τη μόλυνση από τους ρύπους των αυτοκινήτων διερευνήθηκε η σχέση μεταξύ της ροής κίνησης σε κάποιον κύριο αυτοκινητόδρομο (σε χιλιάδες αυτοκινήτων ανά 24 ώρες) και της περιεκτικότητας σε μόλυβδο στον φλοιό των

παρακείμενων δένδρων (σε  $\mu g/g$ ) ξηρού βάρους, δηλαδή  $\mu g$  μολύβδου ανά  $g$  ξηρού φλοιού). Οι μετρήσεις της ροής κίνησης και της περιεκτικότητας μολύβδου που έγιναν σε 10 μέρες που επιλέχτηκαν τυχαία δίνονται στον παρακάτω πίνακα.

ροή κίνησης	8.3	8.5	12.0	12.1	17.0	17.8	23.0	24.0	24.3	33.6
περιεκτικότητα	227	312	362	521	640	728	945	738	759	1263

- (α) Σχηματίστε το κατάλληλο διάγραμμα διασποράς και υπολογίστε τον αντίστοιχο συντελεστή συσχέτισης. Σχολιάστε αν υπάρχει σχέση και τι είδους μεταξύ της κίνησης ροής και της περιεκτικότητας σε μόλυβδο καθώς κι αν φαίνεται η περιεκτικότητα σε μόλυβδο να εξαρτάται γραμμικά από την κίνηση ροής. Δικαιολογήστε την απάντησή σας με βάση τον συντελεστή συσχέτισης και το διάγραμμα διασποράς.
- (β) Υπολογίστε τις σημειακές εκτιμήσεις  $a$  και  $b$  των παραμέτρων  $\alpha$  και  $\beta$  της ευθείας παλινδρόμησης (με τη μέθοδο των ελαχίστων τετραγώνων).
- (γ) Με βάση το δείγμα, εκτιμήστε τη μέση περιεκτικότητα σε μόλυβδο στο φλοιό των δένδρων όταν η ροή κίνησης στον αυτοκινητόδρομο σε ένα 24ωρο είναι 20 χιλιάδες. Εξηγήστε αν η εκτίμηση αυτή μπορεί να θεωρηθεί χρήσιμη.
4. Στα πλαίσια μιας μελέτης για την ποιότητα του νερού που απορρέει από εθνικούς δρόμους, έγινε μια έρευνα για τη σχέση της απορροής νερού από ένα δρόμο και της βροχόπτωσης. Μετρήθηκε η ποσότητα νερού από βροχόπτωση (σε  $m^3$ ), έστω  $X$ , κι η ποσότητα απορροής (σε  $m^3$ ) από το δρόμο, έστω  $Y$ , σε μια συγκεκριμένη περιοχή. Τα αποτελέσματα δίνονται στον παρακάτω πίνακα.

$X$	6	12	15	17	23	30	40	47	53	67	72	74	96	100	125
$Y$	6	10	20	15	30	30	26	46	38	40	48	53	70	85	92

- (α) Σχηματίστε το κατάλληλο διάγραμμα διασποράς και υπολογίστε τον συντελεστή αυτοσυσχέτισης. Σχολιάστε αν από το διάγραμμα διασποράς φαίνεται σωστή η υπόθεση πως η ποσότητα απορροής νερού εξαρτάται από την ποσότητα βροχόπτωσης.
- (β) Με βάση το δείγμα, υπολογίστε τη σημειακή εκτίμηση (αν γίνεται) της αληθινής μέσης ποσότητας απορροής νερού αυτού του δρόμου όταν η ποσότητα βροχόπτωσης είναι  $80m^3$ . Κάνετε το ίδιο για ποσότητα βροχόπτωσης  $180m^3$ .
5. Στα πλαίσια μελέτης για την ρύπανση του αέρα που έγιναν 10 διαφορετικές μετρήσεις της ταχύτητας του αέρα και του συντελεστή εξάτμισης σταγονιδίων καμένων καυσίμων σε μια προωθητική μηχανή. Οι μετρήσεις δίνονται στον παρακάτω πίνακα.

Ταχύτητα αέρα (cm/sec)	20	60	100	140	180	220	260	300	340	380
Εξάτμιση ( $mm^2/sec$ )	0.18	0.37	0.35	0.78	0.56	0.75	1.18	1.36	1.17	1.65

- (α) Σχηματίστε το κατάλληλο διάγραμμα διασποράς και σχολιάστε αν φαίνεται ο συντελεστής εξάτμισης σταγονιδίων καμένων καυσίμων να εξαρτάται γραμμικά από την ταχύτητα του αέρα.

- (β) Υπολογίστε τις σημειακές εκτιμήσεις  $a$  και  $b$  των παραμέτρων  $a$  και  $\beta$  της ευθείας παλινδρόμησης (με τη μέθοδο των ελαχίστων τετραγώνων) και σχηματίστε στο διάγραμμα διασποράς του ερωτήματος 5α' την ευθεία ελαχίστων τετραγώνων.
- (γ) Με βάση το δείγμα, εκτιμήστε (αν γίνεται) το μέσο συντελεστή εξάτμισης σταγονιδίων καμένων καυσίμων για ταχύτητα αέρα 190cm/sec και 500cm/sec.
6. Σε μια μελέτη για την επίδραση της βροχόπτωσης στον καθαρισμό του μολυσμένου αέρα, έγιναν 9 μετρήσεις καθημερινής βροχής (σε 0.01cm) και μολυσμένων σωματιδίων (σε mcg/cum) που απαλείφονται από την ατμόσφαιρα. Τα αποτελέσματα δίνονται στον παρακάτω πίνακα.

βροχή	2.1	3.8	4.3	4.5	5.2	5.6	5.9	6.1	7.5
μολυσμένα σωματίδια	141	132	126	121	118	118	116	114	108

- (α) Σχηματίστε το κατάλληλο διάγραμμα διασποράς και σχολιάστε αν φαίνεται να υπάρχει εξάρτηση της ποσότητας του μολυσμένου αέρα από το μέγεθος της βροχόπτωσης, και αν ναι, τι είδους εξάρτηση είναι.
- (β) Υπολογίστε τις σημειακές εκτιμήσεις  $a$  και  $b$  των παραμέτρων  $a$  και  $\beta$  της ευθείας παλινδρόμησης (με τη μέθοδο των ελαχίστων τετραγώνων) για το πρόβλημα της γραμμικής εξάρτησης της ποσότητας του μολυσμένου αέρα από το μέγεθος της βροχόπτωσης.
- (γ) Με βάση το δείγμα και το μοντέλο στο 6β', υπολογίστε τη σημειακή εκτίμηση (αν γίνεται) της αληθινής μέσης ποσότητας μολυσμένων σωματιδίων που απαλείφονται από τον αέρα όταν το μέγεθος της βροχόπτωσης είναι 5 (σε 0.01cm). Κάνετε το ίδιο για μέγεθος βροχόπτωσης 10 (σε 0.01cm).
7. Η χρήση αστικού λεωφορείου σε μια πόλη πιστεύεται ότι επηρεάζεται από το αντίτιμο του εισιτηρίου. Υπάρχουν στοιχεία για τη χρήση του αστικού λεωφορείου (σε ποσοστό μετακίνησης) μετά από αυξήσεις του εισιτηρίου. Στον παρακάτω πίνακα δίνονται 6 τιμές εισιτηρίων και το αντίστοιχο ποσοστό μετακίνησης με αστικό λεωφορείο.

Τιμή εισιτηρίου (σε €)	0.15	0.18	0.25	0.30	0.40	0.50
Ποσοστό κίνησης (σε %)	37	38	34	30	32	28

- (α) Σχηματίστε το κατάλληλο διάγραμμα διασποράς και σχολιάστε αν φαίνεται να υπάρχει εξάρτηση της χρήσης αστικού λεωφορείου από το αντίτιμο του εισιτηρίου, και αν ναι, τι είδους εξάρτηση είναι.
- (β) Υπολογίστε τις σημειακές εκτιμήσεις  $a$  και  $b$  των παραμέτρων  $a$  και  $\beta$  της ευθείας παλινδρόμησης (με τη μέθοδο των ελαχίστων τετραγώνων) για το πρόβλημα της γραμμικής εξάρτησης του ποσοστού χρήσης αστικού λεωφορείου από το αντίτιμο του εισιτηρίου. Σχηματίστε την ευθεία ελαχίστων τετραγώνων στο διάγραμμα διασποράς που σχηματίσατε στο ερώτημα 7α'.
- (γ) Με βάση το δείγμα και το μοντέλο στο ερώτημα 7β', υπολογίστε τη σημειακή εκτίμηση (αν γίνεται) της μέσου ποσοστού κίνησης με αστικό λεωφορείο αν το αντίτιμο είναι 0.35€. Κάνετε το ίδιο για τιμή εισιτηρίου 0.70€.