

ΑΣΚΗΣΕΙΣ

Περιγραφική Στατιστική, Εκτίμηση Παραμέτρων

1. Δείξτε ότι η εκτιμήτρια s^2 της διασποράς σ^2 είναι αμερόληπτη.
2. Δύο τυχαίες μεταβλητές X_1 και X_2 έχουν κοινή διασπορά σ^2 και s_1^2, s_2^2 είναι οι αμερόληπτες δειγματικές διασπορές των X_1 και X_2 , αντίστοιχα, από δείγματα μεγέθους n_1 και n_2 . Δείξτε ότι η εκτιμήτρια $s^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}$ της σ^2 είναι επίσης αμερόληπτη.
3. Για τον έλεγχο της περιεκτικότητας του χάλυβα σε ραδιενέργεια σε δύο εργοστάσια παραγωγής χάλυβα Α και Β έγιναν οι παρακάτω μετρήσεις ραδιενέργειας σε τυχαία δοκίμια χάλυβα (οι μετρήσεις είναι σε Bq/g):

Δοκίμια	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A (Bq/g)	0.37	0.00	0.54	0.59	0.16	0.86	0.86	0.49	0.60	0.55					
B (Bq/g)	0.24	0.52	0.12	0.95	0.26	0.33	0.62	0.32	0.27	0.05	0.39	0.10	0.51	0.79	0.09

Θεωρείται ότι η διασπορά της ραδιενέργειας στο χάλυβα είναι ίδια για τα δύο εργοστάσια ($\sigma_1^2 = \sigma_2^2 = \sigma^2$).

- (α) Σχηματίστε το θηκόγραμμα για τα δεδομένα ραδιενέργειας στο χάλυβα των δύο δειγμάτων και σχολιάστε αν η ραδιενέργεια στους χάλυβες των δύο εργοστασίων φαίνεται να ακολουθούν κανονική κατανομή. Σχολιάστε επίσης αν φαίνεται να διαφέρουν αυτές οι δύο κατανομές.
 - (β) Εκτιμήστε τη μέση ραδιενέργεια στο χάλυβα για το εργοστάσιο Α (σημειακή εκτίμηση και 95% διάστημα εμπιστοσύνης).
 - (γ) Κάνετε την ίδια εκτίμηση για το εργοστάσιο Β.
 - (δ) Έστω ότι το ανώτατο επιτρεπτό όριο για τη μέση ραδιενέργεια στο χάλυβα είναι 0.5 Bq/g. Με βάση τα παραπάνω διαστήματα εμπιστοσύνης σε επίπεδο εμπιστοσύνης 95% απαντήστε αν θα γινόταν αποδεκτός στην αγορά ο χάλυβας από το εργοστάσιο Α και από το εργοστάσιο Β.
 - (ε) Ελέγξτε χρησιμοποιώντας διάστημα εμπιστοσύνης σε επίπεδο εμπιστοσύνης 95% αν η μέση ραδιενέργεια στους χάλυβες των δύο εργοστασίων είναι ίδια.
4. (*) Έγιναν 15 μετρήσεις της συγκέντρωσης διαλυμένου οξυγόνου (Δ.Ο.) σ' ένα ποτάμι (σε mg/l)

Οξυγόνο Α	1.8	2.0	2.1	1.7	1.2	2.3	2.5	2.9	1.6	2.2	2.3	1.8	2.4	1.6	1.9
-----------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

Από παλιότερες μετρήσεις γνωρίζουμε ότι η διασπορά του Δ.Ο. είναι $0.1 (mg/l)^2$.

- (α) Εκτιμείστε τη διασπορά της συγκέντρωσης Δ.Ο. από το δείγμα καθώς και τα διαστήματα εμπιστοσύνης σε επίπεδο 99% και 90%. Εξετάστε και για τα δύο επίπεδα εμπιστοσύνης αν μπορούμε να δεχτούμε την εμπειρική τιμή της διασποράς για αυτό το δείγμα.
- (β) Υπολογίστε τα μέτρα θέσης και μεταβλητότητας για τα δεδομένα του δείγματος και σχηματίστε το κατάλληλο θηκόγραμμα. Σχολιάστε αν φαίνεται η συγκέντρωση Δ.Ο. στο νερό του ποταμού να ακολουθεί κανονική κατανομή.
- (γ) Εκτιμείστε τη μέση συγκέντρωση Δ.Ο. από το δείγμα και δώστε για αυτήν 95% διάστημα εμπιστοσύνης υποθέτοντας πρώτα ότι η διασπορά είναι γνωστή και μετά χρησιμοποιώντας αυτήν του δείγματος.
- (δ) Αν υποθέσουμε ότι για ένα εργοστάσιο δίπλα στο ποτάμι είναι σημαντικό η μέση συγκέντρωση Δ.Ο. να μην πέφτει κάτω από $1.8 mg/l$, θα προκαλούσαν ανησυχία αυτές οι παρατηρήσεις (χρησιμοποιήστε τη διασπορά από το δείγμα);
- (ε) Αν δε μας ικανοποιεί το πλάτος του τελευταίου παραπάνω διαστήματος και θέλουμε να το μειώσουμε σε $0.2 mg/l$ πόσες επιπρόσθετες ημερήσιες μετρήσεις πρέπει να γίνουν;
5. (*) Σε συνέχεια της Άσκησης 1, σ' έναν άλλο ποταμό έγιναν οι παρακάτω παρόμοιες μετρήσεις

Οξυγόνο Β	2.3	2.1	1.9	2.6	2.9	1.5	3.1	2.1	2.7	2.3	2.6	2.5
-----------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

- (α) Σχηματίζοντας κατάλληλο θηκόγραμμα σχολιάστε αν φαίνεται η συγκέντρωση Δ.Ο. στον ποταμό Β να ακολουθεί κανονική κατανομή. Από τα θηκογράμματα φαίνεται να διαφέρει η συγκέντρωση Δ.Ο. στο νερό των δύο ποταμών;
- (β) Βρείτε 95% διάστημα εμπιστοσύνης για τη διαφορά των μέσων συγκεντρώσεων Δ.Ο. στα δύο ποτάμια υποθέτοντας πρώτα ότι η διασπορά είναι γνωστή ($0.1 (mg/l)^2$) και ίδια για τα δύο δείγματα και μετά χρησιμοποιώντας τις εκτιμήσεις των διασπορών από τα δείγματα. Μπορούμε να πούμε πως η μέση συγκέντρωση Δ.Ο. είναι ίδια στα δύο ποτάμια (στην κάθε περίπτωση);
6. Η τάση διακοπής εναλλασσόμενου ρεύματος ενός μονωτικού υγρού δηλώνει τη διηλεκτρική ανθεκτικότητα του. Πήραμε τις παρακάτω παρατηρήσεις της τάσης διακοπής (kV) σε κάποιο κύκλωμα κάτω από ορισμένες συνθήκες.

41	46	47	47	48	50	50	50	50	50	50	50	50
48	50	50	50	50	50	50	50	52	52	53	55	
50	50	50	50	52	52	53	53	53	53	53	57	
52	52	53	53	53	53	53	53	54	54	55	68	

- (α) Σχηματίστε το ισόγραμμα των δεδομένων της τάσης διακοπής του κυκλώματος και σχολιάστε αν η τάση διακοπής φαίνεται να ακολουθεί κανονική κατανομή.
 - (β) Βρείτε 95% διάστημα εμπιστοσύνης για τη διασπορά της τάσης διακοπής του κυκλώματος.
 - (γ) Από παλιότερες μετρήσεις είχαμε βρει πως η τυπική απόκλιση της τάσης διακοπής παρόμοιου κυκλώματος ήταν περίπου 5 kV. Μπορούμε να δεχθούμε αυτήν την τυπική απόκλιση και γι αυτό το κύκλωμα ;
 - (δ) Βρείτε 95% διάστημα εμπιστοσύνης για τη μέση τάση διακοπής του κυκλώματος, υποθέτοντας
 - (i) τυπική απόκλιση της τάσης διακοπής 5 kV,
 - (ii) ότι η τυπική απόκλιση της τάσης διακοπής είναι άγνωστη.
 - (ε) Μπορούμε να αποκλείσουμε ότι η μέση τάση διακοπής είναι 52 kV για τις περιπτώσεις (i) και (ii) του προηγούμενου ερωτήματος ;
7. Πριν να συμφωνήσει την παραγγελία μιας μεγάλης παρτίδας καλύμματος πολυαιθυλενίου για την προστασία κάποιου ειδικού τύπου καλωδίου, μια εταιρία θέλει να σιγουρευτεί ότι η πραγματική τυπική απόκλιση του πάχους των καλυμμάτων είναι περίπου 0.05 mm. Για το σκοπό αυτό μετρήθηκε το πάχος σε 20 καλύμματα κι οι τιμές (σε mm) είναι

6.711	6.730	6.746	6.860	6.861	6.872	6.928	6.973	7.012	7.015
6.746	6.860	6.861	6.872	6.928	6.973	7.012	7.015	7.053	7.359

- (α) Υπολογίστε τα μέτρα θέσης και μεταβλητότητας για τα δεδομένα του δείγματος και σχηματίστε το κατάλληλο θηκόγραμμα. Σχολιάστε αν φαίνεται το πάχος των καλυμμάτων να ακολουθεί κανονική κατανομή.
 - (β) Ελέγξτε με βάση τα κατάλληλα διαστήματα εμπιστοσύνης αν η εταιρία θα δεχτεί ότι το μέσο πάχος των καλυμμάτων είναι 0.05 mm σε επίπεδο εμπιστοσύνης 95% και 99%.
8. Σε δύο τυχαία δείγματα ελασμάτων που χρησιμοποιούνται στην κατασκευή πινάκων κυκλωμάτων μετρήθηκε το ποσό της στρέβλωσης (σε mm) κάθε ελάσματος κάτω από συγκεκριμένες συνθήκες (διαφορετικές για το κάθε ένα από τα δύο δείγματα). Οι μετρήσεις έδωσαν τα παρακάτω αποτελέσματα

Δείγμα	Πλήθος ελασμάτων	Μέση τιμή (mm)	Τυπική απόκλιση (mm)
1	25	0.161	0.0165
2	22	0.150	0.0130

Υποθέστε ότι το ποσό στρέβλωσης των ελασμάτων για τις δύο διαφορετικές συνθήκες (με ή χωρίς συγκόλληση) ακολουθεί κανονική κατανομή.

- (α) Βρείτε 95% διάστημα εμπιστοσύνης για το μέσο ποσό στρέβλωσης των ελασμάτων για τις δύο διαφορετικές συνθήκες ξεχωριστά.

(β) Ελέγξτε με βάση το 95% διάστημα εμπιστοσύνης αν το μέσο ποσό στρέβλωσης των ελασμάτων είναι το ίδιο κάτω από τις δύο διαφορετικές συνθήκες.

9. Ο υπολογισμός της φόρτισης του δικτύου υπολογιστών γίνεται εύκολα από τη μέτρηση του χρόνου που κάνει ένα πακέτο πληροφορίας να πάει από μια υπολογιστική μονάδα (host) (μονάδα αποστολής) σε μια άλλη μονάδα (μονάδα παραλαβής) και να επιστρέψει πίσω (ping time). Από την κεντρική υπολογιστική μονάδα του εργαστηρίου και μέσα σε μικρό χρονικό διάστημα μετρήθηκαν οι χρόνοι επιστροφής πακέτων πληροφορίας σε 10 μακρινές μονάδες παραλαβής και δίνονται παρακάτω σε χιλιοστά του δευτερολέπτου (ms).

Μονάδα παραλαβής	1	2	3	4	5	6	7	8	9	10
Χρόνος (ms)	25	38	58	52	69	101	132	148	202	232

- (α) Εκτιμείστε σε επίπεδο εμπιστοσύνης 95% το μέσο χρόνο επιστροφής πακέτων πληροφορίας με βάση το δείγμα, υποθέτοντας πως η κατανομή του χρόνου επιστροφής είναι κανονική.
- (β) Αν το δίκτυο χαρακτηρίζεται υπερφορτωμένο όταν ο μέσος χρόνος επιστροφής πακέτων πληροφορίας ξεπερνάει τα 200 ms θα θεωρούσατε το δίκτυο υπερφορτωμένο τη χρονική περίοδο που έγιναν οι μετρήσεις;
- (γ) Πόσες μετρήσεις ακόμα πρέπει να κάνουμε για να έχουμε ακρίβεια εκτίμησης του μέσου χρόνου επιστροφής πακέτων πληροφορίας ± 20 ms (δηλαδή το εύρος του 95% διαστήματος εμπιστοσύνης να είναι 40 ms);
10. Μετρήθηκε η χωρητικότητα (σε αμπερώρες) 10 μπαταριών και τα αποτελέσματα δίνονται ως εξής:

140	136	150	144	148	152	138	141	143	151
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

Από παλιότερες μετρήσεις γνωρίζουμε ότι η χωρητικότητα της μπαταρίας ακολουθεί κανονική κατανομή με τυπική απόκλιση 6 αμπερώρες.

- (α) Εκτιμείστε σε επίπεδο εμπιστοσύνης 95% και σε επίπεδο εμπιστοσύνης 99% τη μέση χωρητικότητα της μπαταρίας.
- (β) Η μπαταρία θεωρείται μη αποδεκτή αν η μέση χωρητικότητα της είναι κάτω από 140 αμπερώρες. Με βάση τα διαστήματα εμπιστοσύνης στο (10α) θα γινόταν η μπαταρία αποδεκτή; [απαντήστε για το κάθε ένα διάστημα εμπιστοσύνης ξεχωριστά].
11. Σ ένα πείραμα μετρήθηκε η αντίσταση 14 καλωδίων τύπου A και η μέση τιμή αντίστασης καλωδίων του δείγματος βρέθηκε να είναι $\bar{x}_1 = 0.137$ Ohms. Η κατανομή του μεγέθους αντίστασης καλωδίου θεωρείται κανονική και η τυπική απόκλιση της αντίστασης καλωδίου θεωρείται γνωστή και είναι 0.01 Ohms.
- (α) Εκτιμώντας το κατάλληλο διάστημα εμπιστοσύνης, εξετάστε αν μπορούμε να δεχτούμε σε επίπεδο εμπιστοσύνης 95% ότι το καλώδιο τύπου A έχει μέση αντίσταση 1.40 Ohms.

(β) Σ ένα δεύτερο πείραμα μετρήθηκε επίσης η αντίσταση 10 καλωδίων από έναν άλλο τύπο Β. Η κατανομή και η τυπική απόκλιση της αντίστασης καλωδίου τύπου Β είναι όπως και για το καλώδιο τύπου Α. Η μέση τιμή αντίστασης των 10 καλωδίων τύπου Β του δείγματος βρέθηκε να είναι $\bar{x}_2 = 0.133$ Ohms. Εκτιμώντας το κατάλληλο διάστημα εμπιστοσύνης, εξετάστε αν μπορούμε να δεχτούμε σε επίπεδο εμπιστοσύνης 95% ότι το καλώδιο τύπου Α έχει μεγαλύτερη αντίσταση από το καλώδιο τύπου Β.

12. Σ ένα ολοκληρωμένο κύκλωμα μετρήθηκε η απολαβή κρυσταλλολυχνίας (τρανζίστορ) μεταξύ εκπομπού και δέκτη κατά τη διαδικασία εναπόθεσης. Έγιναν 14 επαναλήψεις της διαδικασίας εναπόθεσης σε διαφορετικές συνθήκες εκπομπής και οι μετρήσεις της απολαβής κρυσταλλολυχνίας (σε hFE) δίνονται παρακάτω :

1004	1636	1260	1270	1272	903	1555	1276	1321	852	1506	1269	1146	1225
------	------	------	------	------	-----	------	------	------	-----	------	------	------	------

- (α) Για τα παραπάνω δεδομένα, σχηματίστε το θηκόγραμμα, υπολογίζοντας πρώτα τα κατάλληλα συνοπτικά μέτρα, κι εξετάστε αν η κατανομή της απολαβής κρυσταλλολυχνίας φαίνεται να είναι κανονική.
- (β) Υποθέτοντας πως η κατανομή της απολαβής κρυσταλλολυχνίας είναι κανονική, εκτιμήστε σε επίπεδο εμπιστοσύνης 95% διάστημα εμπιστοσύνης για τη μέση απολαβή της κρυσταλλολυχνίας.
- (γ) Σχολιάστε πως θα εκτιμούσατε το παραπάνω διάστημα εμπιστοσύνης στο (12β), αν το δείγμα είχε μέγεθος 100 (αντί για 14).
13. Σε 144 δοκιμές σε κάποιο εργαστήριο, 48 κατέληξαν σε ανάφλεξη κάποιου συγκεκριμένου τύπου υποστρώματος από αναμμένο τσιγάρο. Έστω p η αναλογία όλων των δοκιμών που καταλήγουν σε ανάφλεξη του υποστρώματος.
- (α) Υπολογίστε το 99% διάστημα εμπιστοσύνης για την αναλογία p .
- (β) Αν θέλουμε το εύρος του 99% διαστήματος εμπιστοσύνης για την αναλογία να είναι 0.1 πόσες ακόμα δοκιμές πρέπει να κάνουμε ;
14. Σε μια μελέτη για ένα καινούριο σύστημα εκτόξευσης ρουκετών μικρού βεληνεκούς έγιναν δοκιμές με το παλιό και το καινούριο σύστημα. Σε δείγμα 60 πειραματικών εκτοξεύσεων με το παλιό σύστημα 44 ήταν πετυχημένες και σε 80 πειραματικές εκτοξεύσεις με το καινούριο σύστημα 72 ήταν πετυχημένες.
- (α) Υπολογίστε τα 95% διαστήματα εμπιστοσύνης για τις αναλογίες πετυχημένων εκτοξεύσεων με τα δύο συστήματα.
- (β) Σε επίπεδο εμπιστοσύνης 90% κάνετε στατιστικό έλεγχο για την υπόθεση ότι οι δύο αναλογίες δε διαφέρουν.
- (γ) Αν στον παραπάνω έλεγχο βρήκατε ότι υπάρχει διαφορά εκτιμήστε πόση είναι αυτή η διαφορά.
15. Μετρήθηκε ο χρόνος απόκρισης μιας υπολογιστικής μονάδας μέσου δικτύου σε επικοινωνία με την κεντρική υπολογιστική μονάδα. Παρακάτω δίνονται οι μετρήσεις σε 21 διαφορετικές χρονικές στιγμές (σε ms).

16.1	9.6	24.9	20.4	12.7	21.2	30.2	25.8	18.5	10.3	25.3
14.0	27.1	45.0	23.3	24.2	14.6	8.9	32.4	11.8	28.5	

Θεωρούμε ότι ο χρόνος απόκρισης ακολουθεί κανονική κατανομή αλλά με άγνωστη για μας διασπορά.

- (α) Υπολογίστε το 95% διάστημα εμπιστοσύνης καθώς και το 99% διάστημα εμπιστοσύνης για το μέσο χρόνο απόκρισης με βάση το δείγμα.
- (β) Με βάση τα παραπάνω διαστήματα εμπιστοσύνης στο 15α' (ή αντίστοιχα και ισοδύναμα με κατάλληλο έλεγχο υπόθεσης) εξετάστε αν μπορεί ο μέσος χρόνος απόκρισης να είναι μεγαλύτερος των 30 ms. [Δώστε απαντήσεις σε επίπεδο εμπιστοσύνης 95% και 99%.]

16. Μετρήθηκε ο χρόνος ζωής 10 λαμπτήρων τύπου A σε ώρες:

430	530	590	580	740	800	830	1000	1020	1100
-----	-----	-----	-----	-----	-----	-----	------	------	------

- (α) Για τα παραπάνω δεδομένα, σχηματίστε το θηκόγραμμα, υπολογίζοντας πρώτα τα κατάλληλα συνοπτικά μέτρα, κι εξετάστε αν η κατανομή του χρόνου ζωής των λαμπτήρων φαίνεται να είναι κανονική.
- (β) Υπολογίστε το 95% διάστημα εμπιστοσύνης για την τυπική απόκλιση του χρόνου ζωής των λαμπτήρων τύπου A.
- (γ) Μετρήθηκε επίσης ο χρόνος ζωής 19 λαμπτήρων τύπου B σε ώρες και βρέθηκαν τα εξής:
 δειγματική μέση τιμή $\bar{x}_2 = 500$
 δειγματική τυπική απόκλιση $s_2 = 150$
 Υποθέτοντας πως η κατανομή του χρόνου ζωής λαμπτήρων είναι κανονική (και για τους δύο τύπους και ανεξάρτητα από την απάντησή σας στο 16α'), εκτιμήστε σε επίπεδο εμπιστοσύνης 99% αν υπάρχει διαφορά και πόση στο χρόνο ζωής των δύο τύπων λαμπτήρων.

17. Μετρήθηκε η χωρητικότητα (σε αμπερώρες) 100 μπαταριών και υπολογίστηκε για το δείγμα αυτό η μέση χωρητικότητα και τυπική απόκλιση και βρέθηκαν να είναι $\bar{x} = 140$ αμπερώρες και $s = 30$ αμπερώρες, αντίστοιχα.

- (α) Εκτιμήστε σε επίπεδο εμπιστοσύνης 95% και σε επίπεδο εμπιστοσύνης 90% τη μέση χωρητικότητα της μπαταρίας.
- (β) Η μπαταρία θεωρείται μη αποδεκτή αν η μέση χωρητικότητά της είναι κάτω από 135 αμπερώρες. Με βάση τα διαστήματα εμπιστοσύνης που υπολογίσατε στο 17α' θα γινόταν η μπαταρία αποδεκτή; [απαντήστε για το κάθε ένα επίπεδο εμπιστοσύνης ξεχωριστά]
- (γ) Για να απαντήσουμε με μεγαλύτερη σιγουριά στο ερώτημα 17β', θέλουμε να μικρύνουμε το εύρος του 95% διαστήματος εμπιστοσύνης σε 6 αμπερώρες (δηλαδή να δηλώσουμε τη μέση χωρητικότητά με ακρίβεια ± 3 αμπερώρες). Για να το πετύχουμε αυτό σε πόσες ακόμα μπαταρίες πρέπει να μετρήσουμε τη χωρητικότητά;

18. Από παλιά μελέτη (πριν αρκετά χρόνια) για το χρόνο λειτουργίας ηλεκτρικής μηχανής μέχρι την πρώτη βλάβη που έγινε σε δείγμα 169 μηχανών, γνωρίζουμε πως το 95% διάστημα εμπιστοσύνης για το μέσο χρόνο μέχρι την πρώτη βλάβη έχει εύρος 20 μέρες.
- (α) Πόση ήταν η τυπική απόκλιση του χρόνου μέχρι την πρώτη βλάβη (σε μέρες) στο δείγμα των 169 μηχανών; Στα πλαίσια νέας μελέτης αλλά σε μικρότερη κλίμακα μετρήσαμε το χρόνο μέχρι την πρώτη βλάβη σε 61 ηλεκτρικές μηχανές ίδιου τύπου αλλά νεότερης τεχνολογίας και βρήκαμε το μέσο όρο του χρόνου μέχρι την πρώτη βλάβη $\bar{x} = 720$ μέρες και την τυπική απόκλιση $s = 100$ μέρες.
- (β) Υπολογίστε το 95% διάστημα εμπιστοσύνης για το μέσο χρόνο μέχρι την πρώτη βλάβη με βάση το νέο δείγμα.
- (γ) Θεωρούμε πως η εκτίμηση της τυπικής απόκλισης του χρόνου μέχρι την πρώτη βλάβη από το νέο δείγμα ($s = 100$) είναι ακριβής (δηλαδή είναι πολύ κοντά στην πραγματική τυπική απόκλιση). Κάνετε τους απαραίτητους υπολογισμούς για να διερευνήσετε αν μπορεί να γίνει η εκτίμηση του μέσου χρόνου μέχρι την πρώτη βλάβη με μεγαλύτερη ακρίβεια (μικρότερο εύρος του 95% διαστήματος εμπιστοσύνης) στις νέες μηχανές απ' ότι στις παλιές μηχανές (όπου είχαμε ακρίβεια 10 ημερών με 169 παλιές μηχανές).
19. Εκτενής μελέτη ενός συστήματος υπολογιστών καταμερισμού υπολογιστικού χρόνου (computer time-sharing system) έδειξε ότι ο χρόνος απόκρισης σε κάποια εντολή του λειτουργικού συστήματος ακολουθεί κανονική κατανομή με τυπική απόκλιση 15 millisecc.
- (α) Θέλουμε να προσδιορίσουμε τον πραγματικό μέσο χρόνο απόκρισης μ με ακρίβεια 3 millisecc σε επίπεδο εμπιστοσύνης 95% (δηλαδή το διάστημα εμπιστοσύνης να έχει εύρος 6 millisecc). Πόσες μετρήσεις του χρόνου απόκρισης πρέπει να κάνουμε;
- Δοκιμάστηκε ένα νέο λειτουργικό σύστημα και μετρήθηκε ο χρόνος απόκρισης του λειτουργικού συστήματος σε 61 εντολές του λειτουργικού συστήματος. Από τις 61 μετρήσεις υπολογίστηκαν τα παρακάτω μέτρα γι αυτό το δείγμα των παρατηρήσεων του χρόνου απόκρισης:
- μέση τιμή $\bar{x} = 25$ millisecc
 τυπική απόκλιση $s = 18$ millisecc.
- (β) Υπολογίζοντας κατάλληλο διάστημα εμπιστοσύνης εξετάστε αν σε επίπεδο εμπιστοσύνης 99% μπορούμε να δεχθούμε ότι η τυπική απόκλιση του χρόνου απόκρισης για το νέο λειτουργικό σύστημα παραμένει στο επίπεδο των 15 millisecc.
- (γ) Ο προμηθευτής του λειτουργικού συστήματος ισχυρίζεται ότι ο μέσος χρόνος απόκρισης είναι 20 millisecc. Υπολογίζοντας κατάλληλο διάστημα εμπιστοσύνης για το ίδιο δείγμα εξετάστε αν σε επίπεδο εμπιστοσύνης 90% μπορούμε να δεχθούμε τον ισχυρισμό του προμηθευτή.

Συσχέτιση - Παλινδρόμηση

1. (*) Η πρώτη ύλη που χρησιμοποιείται στην παραγωγή κάποιας συνθετικής ίνας αποθηκεύεται σε κάποιο χώρο χωρίς έλεγχο υγρασίας. Έγιναν μετρήσεις για 15 μέρες της σχετικής υγρασίας στον αποθηκευτικό χώρο και της περιεκτικότητας σε υγρότητα σε δείγμα της πρώτης ύλης. Τα αποτελέσματα (σε ποσοστά) είναι

Σχετική υγρασία	46	53	29	61	36	39	47	49	52	38	55	32	57	54	44
Υγρότητα	12	15	7	17	10	11	11	12	14	9	16	8	18	14	12

- (α) Σχεδιάστε το διάγραμμα διασποράς και υπολογίστε τον συντελεστή συσχέτισης της σχετικής υγρασίας στον αποθηκευτικό χώρο και της περιεκτικότητας σε υγρότητα της πρώτης ύλης. Μπορούμε να υποθέσουμε ότι η συσχέτιση είναι γραμμική;
- (β) Εκτιμήστε την ευθεία ελαχίστων τετραγώνων και τη διασπορά των σφαλμάτων.
- (γ) Προβλέψτε αν γίνεται την περιεκτικότητα σε υγρότητα της πρώτης ύλης όταν η σχετική υγρασία στον αποθηκευτικό χώρο είναι 10%, 40% και 70%.
2. Ο χρόνος ως την αποτυχία κάποιου μηχανικού στοιχείου σχετίζεται με την ηλεκτρική τάση λειτουργίας της μηχανής. Εκτελέστηκε ένα σχεδιασμένο πείραμα σ' ερευνητικό εργαστήριο που έδωσε τα παρακάτω δεδομένα

Χρόνος αποτυχίας (min)	2145	2155	2220	2225	2260	2266	2334	2340	2212	2180
Τάση λειτουργίας (αμπέρ)	110	110	110	110	120	120	120	130	115	115

- (α) Σχεδιάστε το διάγραμμα διασποράς και σχολιάστε αν μπορεί να θεωρηθεί ότι ο χρόνος ως την αποτυχία του μηχανικού στοιχείου εξαρτάται γραμμικά από την τάση λειτουργίας της μηχανής.
- (β) Υποθέτοντας γραμμική παλινδρόμηση του χρόνου αποτυχίας στην τάση λειτουργίας, εκτιμήστε την ευθεία ελαχίστων τετραγώνων και τη διασπορά των σφαλμάτων.
- (γ) Σχολιάστε αν η γραμμική εξάρτηση είναι ισχυρή υπολογίζοντας πρώτα το συντελεστή συσχέτισης.
- (δ) Μπορούμε να υποθέσουμε ότι ο μέσος χρόνος αποτυχίας για τάση λειτουργίας 115 αμπέρ είναι μικρότερος από 2000 min; Μπορεί να είναι μικρότερος από 2200 min; Ποια είναι η καλύτερη εκτίμηση του μέσου χρόνου αποτυχίας για αυτήν την τάση με βάση τη γραμμική παλινδρόμηση κι αυτό το δείγμα;
3. Το κόστος παραγωγής ισχύος (ανά κιλοβατώρα) πιστεύεται ότι εξαρτάται (εκτός άλλων παραγόντων) από το κόστος του λιγνίτη (σε cent ανά εκατομμύριο Btu). Οι παρακάτω παρατηρήσεις έγιναν από 12 σπαστήρες λιγνίτη

Κόστος λιγνίτη	14	16	22	24	20	29	26	15	29	24	25	13
Κόστος ισχύος	4.1	4.4	5.6	5.1	5.0	5.3	5.4	4.8	6.1	5.5	4.7	3.9

- (α) Σχεδιάστε το διάγραμμα διασποράς και υπολογίστε τον συντελεστή συσχέτισης του κόστους της παραγωγής ισχύος και του κόστους του λιγνίτη. Μπορούμε να υποθέσουμε ότι υπάρχει συσχέτιση;
- (β) Εκτιμείστε την ευθεία ελαχίστων τετραγώνων και τη διασπορά των σφαλμάτων.
- (γ) Με βάση το μοντέλο παλινδρόμησης, αν το κόστος λιγνίτη είναι 18 cent/millBtu θα μπορούσε το τυπικό κόστος ισχύος ανά κιλοβατώρα να είναι 6 μονάδες; Θα μπορούσε για το ίδιο κόστος λιγνίτη να είναι 5 μονάδες;
4. Θέλουμε να μελετήσουμε κατά πόσο εξαρτάται ο χρόνος επιστροφής πακέτων πληροφορίας (ping time) από τον αριθμό των ενδιαμέσων υπολογιστικών μονάδων της διαδρομής της πληροφορίας στο δίκτυο από τη μονάδα αποστολής στη μονάδα παραλαβής. Από την κεντρική υπολογιστική μονάδα του εργαστηρίου και μέσα σε μικρό χρονικό διάστημα μετρήθηκαν οι χρόνοι επιστροφής (σε χιλιοστά του δευτερολέπτου, ms) πακέτων πληροφορίας στην ίδια μακρινή μονάδα παραλαβής αλλά με διαφορετικό αριθμό ενδιαμέσων σταθμών. Τα αποτελέσματα δίνονται παρακάτω

Ενδιάμεσοι σταθμοί	1	1	2	2	3	3	4	5	6	7
Χρόνος (ms)	25	38	58	52	69	101	132	148	202	232

- (α) Σχηματίστε το κατάλληλο διάγραμμα διασποράς και σχολιάστε αν η υπόθεση, δηλαδή ότι ο χρόνος επιστροφής πακέτων πληροφορίας εξαρτάται γραμμικά από τον αριθμό των ενδιαμέσων υπολογιστικών μονάδων της διαδρομής, φαίνεται σωστή με βάση το δείγμα των παρατηρήσεων από τους δύο πίνακες.
- (β) Για το παραπάνω πρόβλημα στο ερώτημα (4α'), υπολογίστε τις σημειακές εκτιμήσεις a και b των παραμέτρων α και β της ευθείας παλινδρόμησης (με τη μέθοδο των ελαχίστων τετραγώνων).
- (γ) Προβλέψτε το χρόνο επιστροφής πακέτων πληροφορίας (αν γίνεται) όταν δεν υπάρχει ενδιάμεσος σταθμός κι όταν υπάρχουν 5 ενδιάμεσοι σταθμοί.
5. Η απευθείας μέτρηση της ποσότητας πρωτεΐνης σε δείγμα (κομμάτι) συκωτιού είναι δύσκολη κι απαιτεί πολύ χρόνο. Πιστεύεται ότι η ποσότητα της πρωτεΐνης σχετίζεται με την ποσότητα του φωτός που θα απορροφούσαν από το δείγμα συκωτιού. Γι αυτό έγινε στο εργαστήριο το ακόλουθο πείραμα. Στάλθηκε από ένα φασματόμετρο φως σε διάλυμα που περιείχε δείγμα συκωτιού και μετρήθηκε το φως που απορροφήθηκε. Αυτή η διαδικασία εφαρμόστηκε σε 5 δείγματα συκωτιού για τα οποία η ποσότητα της πρωτεΐνης ήταν γνωστή. Τα αποτελέσματα δίνονται στον παρακάτω πίνακα.

Φως που απορροφήθηκε	0.44	0.82	1.20	1.61	1.83
Ποσότητα πρωτεΐνης (mg)	2	16	30	46	55

- (α) Σχηματίστε το κατάλληλο διάγραμμα διασποράς και σχολιάστε αν η υπόθεση, ότι η ποσότητα πρωτεΐνης σε δείγμα συκωτιού εξαρτάται γραμμικά από το αντίστοιχο φως που απορροφείται από το δείγμα, φαίνεται σωστή με βάση το δείγμα των παρατηρήσεων από τον πίνακα.

(β) Για το παραπάνω πρόβλημα στο ερώτημα (5α'), υπολογίστε τις σημειακές εκτιμήσεις a και b των παραμέτρων α και β της ευθείας παλινδρόμησης (με τη μέθοδο των ελαχίστων τετραγώνων) και προβλέψτε την ποσότητα πρωτεΐνης σε δείγμα συκωτιού για το οποίο η απορρόφηση του φωτός είναι 1.5.

6. Στα πλαίσια μιας μελέτης σε πετρελαιοκίνητες μηχανές (diesel) ελαφρών φορτηγών για την έκλυση πρωτοξειδίου του αζώτου, διερευνήθηκε η εξάρτηση της από την υγρασία. Έγινε γι αυτό ένα πείραμα και μετρήθηκε σε 10 διαφορετικές χρονικές στιγμές η υγρασία (σε ποσοστό επί των κορεσμένων ατμών) και η έκλυση πρωτοξειδίου του αζώτου (σε ppm). Οι μετρήσεις δίνονται στον παρακάτω πίνακα :

Υγρασία	8.3	10.7	12.9	20.1	24.0	34.3	35.1	41.6	72.2	72.4
Άζωτο	1.15	1.00	1.10	1.03	1.07	0.96	0.89	0.91	0.77	0.90

- (α) Σχηματίστε το κατάλληλο διάγραμμα διασποράς και σχολιάστε αν η υπόθεση, ότι η έκλυση πρωτοξειδίου του αζώτου Y εξαρτάται γραμμικά από την υγρασία X φαίνεται οσωτή με βάση το δείγμα των παρατηρήσεων.
- (β) Για το παραπάνω πρόβλημα στο ερώτημα (6α'), υπολογίστε τις σημειακές εκτιμήσεις a και b των παραμέτρων α και β της ευθείας παλινδρόμησης (με τη μέθοδο των ελαχίστων τετραγώνων). Σχηματίστε την ευθεία ελαχίστων τετραγώνων στο διάγραμμα διασποράς και προβλέψτε την έκλυση μονοξειδίου του αζώτου για υγρασία 40%.
7. Πιστεύεται ότι σ' ένα ολοκληρωμένο κύκλωμα η απολαβή κρυσταλλολυχνίας (τρανζίστορ) μεταξύ εκπομπού και δέκτη κατά τη διαδικασία εναπόθεσης εξαρτάται από τη δόση εκπομπής, η οποία ελέγχεται κατά τη διαδικασία εναπόθεσης. Έγιναν 14 επαναλήψεις της διαδικασίας εναπόθεσης σε διαφορετικές συνθήκες εκπομπής και οι μετρήσεις της απολαβής κρυσταλλολυχνίας (σε hFE) και της δόσης εκπομπής (σε ιόντα) δίνονται στον παρακάτω πίνακα :

Δόση εκπομπής	Απολαβή
4.00	1004
4.00	1636
4.00	1260
4.10	1270
4.20	1272
4.30	903
4.30	1555
4.30	1276
4.30	1321
4.60	852
4.60	1506
4.60	1269
4.70	1146
4.72	1225

- (α) Σχηματίστε το κατάλληλο διάγραμμα διασποράς και σχολιάστε αν η υπόθεση, ότι η απολαβή κρυσταλλολυχνίας εξαρτάται γραμμικά από τη δόση εκπομπής φαίνεται σωστή με βάση το δείγμα των παρατηρήσεων.
- (β) Για το παραπάνω πρόβλημα στο ερώτημα (7α'), υπολογίστε τις σημειακές εκτιμήσεις a και b των παραμέτρων α και β της ευθείας παλινδρόμησης (με τη μέθοδο των ελαχίστων τετραγώνων), καθώς και την εκτίμηση s^2 της διασποράς των σφαλμάτων σ^2 . Σχολιάστε με βάση τις εκτιμήσεις αυτές, αν μπορούμε να κάνουμε ακριβείς προβλέψεις της απολαβής κρυσταλλολυχνίας όταν γνωρίζουμε τη δόση εκπομπής για τιμές δόσης εκπομπής μεταξύ 4.00 και 4.70.

8. Στον παρακάτω πίνακα δίνεται για 10 σταθμούς ο αριθμός των ημερών σ' ένα χρόνο που η θερμοκρασία έπεσε κάτω από 0°C και το υψόμετρο τους.

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

- (α) Υποθέτουμε ότι ο αριθμός των ημερών Y εξαρτάται γραμμικά από το υψόμετρο 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 α και β της ευθείας παλινδρόμησης (με τη μέθοδο των ελαχίστων τετραγώνων).
- (γ) Με βάση το δείγμα, μπορείτε να εκτιμήσετε το μέσο αριθμό ημερών το χρόνο που η θερμοκρασία πέφτει κάτω από 0°C σε υψόμετρο 1500m; Σε υψόμετρο 2200m;
9. Θέλουμε να διερευνήσουμε αν η μηνιαία κατανάλωση ηλεκτρικού ρεύματος (σε χιλιάδες ΩXB) σε μια χημική βιομηχανία εξαρτάται γραμμικά από την παραγωγή του προϊόντος (σε τόνους). Οι μετρήσεις σε 9 μήνες δίνονται παρακάτω:

Προϊόν	98	104	105	106	107	109	110	115	120
Κατανάλωση	27	30	29	43	30	31	35	35	39

- (α) Σχηματίστε το κατάλληλο διάγραμμα διασποράς και σχολιάστε αν η υπόθεση, ότι η κατανάλωση ηλεκτρικού ρεύματος εξαρτάται γραμμικά από την παραγωγή του προϊόντος, φαίνεται σωστή με βάση το δείγμα.
- (β) Για το παραπάνω πρόβλημα στο ερώτημα 9α', υπολογίστε τις σημειακές εκτιμήσεις a και b των παραμέτρων α και β της ευθείας παλινδρόμησης (με τη μέθοδο των ελαχίστων τετραγώνων).
- (γ) Προβλέψτε την κατανάλωση ρεύματος για παραγωγή 110 τόνων και παραγωγή 150 τόνων του προϊόντος.
10. Στα πλαίσια μιας μελέτης για την ποιότητα ελασμάτων που χρησιμοποιούνται στην κατασκευή πινάκων κυκλωμάτων μετρήθηκε σε 9 ελάσματα το ποσό στρέβλωσης (σε mm) κάτω από συγκεκριμένη θερμοκρασία περιβάλλοντος για το κάθε έλασμα. Οι μετρήσεις δίνονται στον παρακάτω πίνακα:

Θερμοκρασία (σε $^\circ\text{C}$)	15	20	23	28	30	30	32	36	40
Στρέβλωση (σε mm)	0.11	0.08	0.17	0.12	0.16	0.14	0.18	0.12	0.17

- (α) Σχηματίστε το κατάλληλο διάγραμμα διασποράς και σχολιάστε αν η υπόθεση, ότι η στρέβλωση ελάσματος εξαρτάται γραμμικά από την θερμοκρασία φαίνεται σωστή με βάση το δείγμα των παρατηρήσεων.
- (β) Για το παραπάνω πρόβλημα στο ερώτημα 10α', υπολογίστε τις σημειακές εκτιμήσεις a και b των παραμέτρων α και β της ευθείας παλινδρόμησης (με τη μέθοδο των ελαχίστων τετραγώνων).
- (γ) Σχηματίστε την ευθεία ελαχίστων τετραγώνων στο διάγραμμα διασποράς και προβλέψτε τη στρέβλωση ελάσματος για θερμοκρασία 30°C.

11. Σε ένα φορτηγό μηχανής diesel έγιναν μετρήσεις της εκπομπής νιτρικού οξέος και της υγρασίας σε 10 διαφορετικές συνθήκες υγρασίας, και τα αποτελέσματα δίνονται παρακάτω.

Υγρασία (σε %)	72	41	34	35	10	12	8	20	72	24
Νιτρικό οξύ (σε ppm)	0.90	0.91	0.96	0.89	1.00	1.10	1.15	1.03	0.77	1.07

- (α) Σχηματίστε το κατάλληλο διάγραμμα διασποράς και σχολιάστε αν φαίνεται να εξαρτάται, και πως, η εκπομπή νιτρικού οξέος από την υγρασία.
- (β) Υπολογίστε τις σημειακές εκτιμήσεις a και b των παραμέτρων α και β της ευθείας παλινδρόμησης (με τη μέθοδο των ελαχίστων τετραγώνων) για το πρόβλημα της γραμμικής εξάρτησης της εκπομπής νιτρικού οξέος από την υγρασία. Σχηματίστε την ευθεία ελαχίστων τετραγώνων στο διάγραμμα διασποράς.

12. Ένας οδηγός ισχυρίζεται ότι η κατανάλωση καυσίμου του αυτοκινήτου του δεν εξαρτάται από την ταχύτητα του αυτοκινήτου. Για να ελέγξει αυτήν την υπόθεση, οδήγησε το αυτοκίνητο στον αυτοκινητόδρομο με σταθερή ταχύτητα και για ταχύτητες από 80km/h μέχρι 110km/h. Μέτρησε επίσης την κατανάλωση καυσίμου (σε km/l) σε κάθε περίπτωση και τα αποτελέσματα δίνονται στον παρακάτω πίνακα.

Ταχύτητα (km/h)	80	85	90	95	100	105	110
Κατανάλωση καυσίμου (km/l)	9.1	8.7	9.2	8.5	8.0	8.7	8.3

- (α) Σχηματίστε το κατάλληλο διάγραμμα διασποράς και σχολιάστε αν η υπόθεση του οδηγού φαίνεται σωστή και αν όχι, τότε σχολιάστε πως φαίνεται να εξαρτάται η κατανάλωση καυσίμου από την ταχύτητα που τρέχει το αυτοκίνητο.
- (β) Για το παραπάνω πρόβλημα υπολογίστε τις σημειακές εκτιμήσεις a και b των παραμέτρων α και β της ευθείας παλινδρόμησης (με τη μέθοδο των ελαχίστων τετραγώνων). Σχηματίστε την ευθεία ελαχίστων τετραγώνων στο διάγραμμα διασποράς.
- (γ) Υποδείξτε αν μπορούμε να κάνουμε πρόβλεψη της κατανάλωσης καυσίμου για ταχύτητα 130 km/h, και αν ναι υπολογίστε την πρόβλεψη της μέσης κατανάλωσης καυσίμου σε αυτή την περίπτωση.