

ΑΣΚΗΣΕΙΣ**Περιγραφική Στατιστική, Εκτίμηση και Έλεγχος 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. Υπολογίστε την εκτιμήτρια της μέγιστης πιθανοφάνειας της παραμέτρου λ της εκθετικής κατανομής $f_X(x) = \lambda e^{-\lambda x}$ από ένα δείγμα της X μεγέθους n . Είναι αυτή η εκτιμήτρια αμερόληπτη;
4. Μετρήθηκε η αντοχή διάτμησης σε κάθε ένα από 10 σημεία συγκόλλησης που χρησιμοποιήθηκαν σε ένα πείραμα και τα αποτελέσματα είναι (σε psi)

392	376	401	367	389	362	409	415	358	375
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

- (α) Υποθέτοντας πως η αντοχή διάτμησης ακολουθεί κανονική κατανομή εκτιμήστε την πραγματική μέση αντοχή διάτμησης και την τυπική απόκλιση της αντοχής διάτμησης με τη μέθοδο της μέγιστης πιθανοφάνειας.
 - (β) Υποθέτουμε πως θέλουμε να ελέγξουμε την αντοχή διάτμησης ενός άλλου σημείου συγκόλλησης. Χρησιμοποιώντας τα παραπάνω δεδομένα βρείτε την πιθανότητα η αντοχή διάτμησης να είναι μικρότερη από 400 psi.
5. 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) Για το παραπάνω ερώτημα κάνετε κατάλληλο στατιστικό έλεγχο στο ίδιο επίπεδο εμπιστοσύνης. Σχολιάστε τις απαντήσεις σας από τον έλεγχο και το διάστημα εμπιστοσύνης.
- (ϛ) Ελέγξτε χρησιμοποιώντας διάστημα εμπιστοσύνης και έλεγχο υπόθεσης σε επίπεδο εμπιστοσύνης 95% αν η μέση ραδιενέργεια στους χάλυβες των δύο εργοστασίων είναι ίδια.

6. Στα πλαίσια μελέτης για την επίβλεψη της ποιότητας του νερού σε έναν ποταμό μετρήθηκε η ποσότητα των αιωρούμενων στερεών (αναλογία σωματιδίων ανά εκατομύριο) σε δείγμα νερού από τον ποταμό. Τα αποτελέσματα των μετρήσεων για 10 μέρες είναι

14	25	21	28	30	63	29	63	55	19
----	----	----	----	----	----	----	----	----	----

Από προηγούμενες μετρήσεις σε άλλους ποταμούς είναι γνωστό ότι η τυπική απόκλιση της ποσότητας των αιωρούμενων σωματιδίων στο νερό είναι 20 σωματίδια ανά εκατομύριο.

- (α) Μπορούμε να δεχτούμε την παραπάνω εμπειρική τιμή της τυπικής απόκλισης των αιωρούμενων στερεών γι αυτόν το ποταμό; Γιατί; Κάνετε τον κατάλληλο έλεγχο για να απαντήσετε αυτό το ερώτημα σε στατιστική σημαντικότητα $\alpha = 0.01$.
- (β) Υπολογίστε τα μέτρα θέσης και μεταβλητότητας για τα δεδομένα του δείγματος και σχηματίστε το κατάλληλο θηκόγραμμα. Σχολιάστε αν φαίνεται η περιεκτικότητα των αιωρούμενων σωματιδίων στο νερό του ποταμού να ακολουθεί κανονική κατανομή.
- (γ) Θεωρώντας γνωστή την τυπική απόκλιση των αιωρούμενων σωματιδίων και ίση με 20, εκτιμήστε τη μέση ποσότητα των αιωρούμενων στερεών στο νερό του ποταμού υπολογίζοντας τη σημειακή εκτίμηση και το 95% διάστημα εμπιστοσύνης.
- (δ) Το νερό του ποταμού θεωρείται κακής ποιότητας όταν η μέση ποσότητα των αιωρούμενων σωματιδίων στο νερό υπερβαίνει την ποσότητα των 50 σωματιδίων (ανά εκατομύριο). Διατυπώστε την κατάλληλη στατιστική υπόθεση και κάνετε τον έλεγχο για να απαντήσετε σε στατιστική σημαντικότητα $\alpha = 0.05$ αν το νερό του ποταμού θα χαρακτηριζόταν ως κακής ποιότητας με βάση τό δείγμα. Ποιά είναι η απάντηση στο ίδιο ερώτημα σε επίπεδο σημαντικότητας $\alpha = 0.01$;
- (ε) Πόσες επιπλέον παρατηρήσεις πρέπει να κάνουμε για να έχει το 95% διάστημα εμπιστοσύνης της μέσης ποσότητας αιωρούμενων στερεών εύρος 10 (δηλαδή η ακρίβεια της μέσης ποσότητας να δίνεται με ± 5);
7. Έχει μελετηθεί η ρωγματογενής διάβρωση σιδήρου κι ατσαλιού από καυστική τάση γιατί προκαλεί θραύσεις του μετάλλου, για παράδειγμα παρουσιάζονται θραύσεις γύρω από πιρτσίνια (καρφιά) σε χαλύβδινους λέβητες. Σε δοκιμές που έγιναν μετρήθηκε το μήκος ρωγμής (σε μm) ως αποτέλεσμα της σταθερής τάσης διάβρωσης που ασκήθηκε σε λείες ράβδους από κάποιο κράμα χάλυβα για ορισμένο χρονικό διάστημα. Έγιναν οι παρακάτω μετρήσεις μήκους ρωγμής (σε μm) σε 21 δοκίμια :

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	

- (α) Σχηματίστε το ιστόγραμμα των δεδομένων του μήκους ρωγμής και σχολιάστε αν το μήκος ρωγμής του μετάλλου φαίνεται να ακολουθεί κανονική κατανομή.

- (β) Υπολογίστε το 95% διάστημα εμπιστοσύνης καθώς και το 99% διάστημα εμπιστοσύνης για το μέσο μήκος ρωγμής του μετάλλου με βάση το δείγμα.
- (γ) Το μέταλλο θεωρείται ακατάλληλο όταν το μήκος ρωγμής υπερβαίνει τα 30 μm (επειδή πιστεύεται ότι τότε μπορεί να προκληθεί θραύση από τη ρωγματογενή διάβρωση). Με βάση στατιστικό έλεγχο υπόθεσης σε επίπεδο εμπιστοσύνης 95% μπορούμε να χαρακτηρίσουμε αυτό το κράμα χάλυβα ακατάλληλο; Θα άλλαζε η απάντηση αν κάναμε τον ίδιο έλεγχο σε επίπεδο εμπιστοσύνης 99%; Συγκρίνετε τις απαντήσεις με τα διαστήματα εμπιστοσύνης από το ερώτημα 7β'.
- (δ) Θέλουμε να εκτιμήσουμε το μέσο μήκος ρωγμής με μεγάλη ακρίβεια, $\pm 2\mu\text{m}$, δηλαδή το 95% διάστημα εμπιστοσύνης του μέσου μήκους ρωγμής να έχει εύρος 4 μm . Πόσες ακόμα μετρήσεις πρέπει να κάνουμε;
8. Η φθορά στους αγωγούς ενός αστικού δικτύου αγωγών έχει προκαλέσει σοβαρές ανησυχίες. Η αποκατάσταση ενός αγωγού με φθορές μπορεί να επιτευχθεί εφαρμόζοντας μια ευπροσάρμοστη επένδυση στον αγωγό. Για να δοκιμασθεί μια νέα τεχνική συγκόλλησης της επένδυσης στον αγωγό έγιναν μετρήσεις της αντοχής εφελκυσμού (σε μονάδα psi) σε 18 δοκίμια με επένδυση. Σε 10 από αυτά χρησιμοποιήθηκε η νέα τεχνική συγκόλλησης και στα άλλα 8 δε χρησιμοποιήθηκε. Τα αποτελέσματα είναι:

Με συγκόλληση (μέση τιμή $\bar{x}_1 = 2917.5$, τυπική απόκλιση $s_1 = 272.3$)

2748	2700	2655	2822	2511	3149	3257	3213	3220	2900
------	------	------	------	------	------	------	------	------	------

Χωρίς συγκόλληση (μέση τιμή $\bar{x}_2 = 3082.9$, τυπική απόκλιση $s_2 = 243.7$)

3027	3356	3359	3297	3125	2910	2889	2700
------	------	------	------	------	------	------	------

- (α) Σχηματίζοντας κατάλληλα γραφήματα σχολιάστε αν φαίνεται η αντοχή εφελκυσμού στους αγωγούς με την επένδυση να ακολουθεί κανονική κατανομή (με ή χωρίς συγκόλληση). Από τα γραφήματα φαίνεται να διαφέρουν οι δύο κατανομές;
- (β) Υπολογίστε το 99% διάστημα εμπιστοσύνης για τη μέση αντοχή εφελκυσμού σε κάθε μια από τις δύο περιπτώσεις επένδυσης με βάση το δείγμα.
- (γ) Αν θέλουμε να εκτιμήσουμε τη μέση αντοχή εφελκυσμού στην επένδυση χωρίς συγκόλληση με ακρίβεια ± 80 psi, δηλαδή το 99% διάστημα εμπιστοσύνης του μέσου μήκους ρωγμής να έχει εύρος 160 psi, πόσες ακόμα μετρήσεις πρέπει να κάνουμε;
- (δ) Χρησιμοποιώντας είτε διάστημα εμπιστοσύνης ή έλεγχο υπόθεσης, εξετάστε σε επίπεδο εμπιστοσύνης 95% αν η μέση αντοχή εφελκυσμού διαφέρει με στατιστική σημαντικότητα σε επενδύσεις χωρίς συγκόλληση και σε επενδύσεις με συγκόλληση.
9. 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;

10. 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 εκτιμείστε πόση είναι αυτή η διαφορά.
11. Οι παρακάτω μετρήσεις κατέγραψαν το χρόνο που περνάει (σε ώρες) ως που να ξεραθεί μια συγκεκριμένη μάρκα Α μπογιάς λατέξ σε 15 διαφορετικές περιπτώσεις:

2.5 2.8 2.8 2.9 3.0 3.3 3.4 3.6 3.7 4.0 4.4 4.8 4.8 5.2 5.6

- (α) Υπολογίστε και σχεδιάστε το κατάλληλο θηκόγραμμα από το δείγμα των 15 παρατηρήσεων. Με βάση το θηκόγραμμα σχολιάστε αν φαίνεται η κατανομή του χρόνου ξήρανσης της μπογιάς να ακολουθεί κανονική κατανομή.
- (β) Υπολογίστε το 95% διάστημα εμπιστοσύνης για την τυπική απόκλιση του χρόνου ξήρανσης της μπογιάς με βάση το δείγμα.
- (γ) Για μία άλλη μάρκα μπογιάς Β, βρήκαμε από ένα δείγμα 10 μετρήσεων πως για το χρόνο που περνάει ως που να ξεραθεί η μπογιά είναι:
δειγματική μέση τιμή $\bar{x} = 3.8$ ώρες,
τυπική απόκλιση $s = 0.6$ ώρες
Υποθέτουμε πως η κατανομή του χρόνου ξήρανσης και για τους δύο τύπους μπογιάς (μάρκας Α και Β) είναι κανονική, καθώς επίσης πως η διασπορά είναι κοινή για τους δύο τύπους αλλά άγνωστη. Χρησιμοποιώντας είτε διάστημα εμπιστοσύνης ή έλεγχο υπόθεσης, εξετάστε σε επίπεδο εμπιστοσύνης 95% αν ο χρόνος ξήρανσης διαφέρει στους δύο τύπους μπογιάς (Α και Β).
12. Μετρήθηκε το περιεχόμενο 7 όμοιων βυτιών (για αποθήκευση θειϊκού οξέος) και οι μετρήσεις δίνονται στον παρακάτω πίνακα (σε λίτρα):

9.8 10.2 10.4 9.8 10.0 10.2 9.6

- (α) Υπολογίστε και σχεδιάστε το κατάλληλο θηκόγραμμα από το δείγμα των 7 παρατηρήσεων. Με βάση το θηκόγραμμα σχολιάστε αν φαίνεται η κατανομή του περιεχόμενου των βυτιών να ακολουθεί κανονική κατανομή.
- Στη συνέχεια (κι ανεξάρτητα από την απάντησή σας στο 12α'), υποθέστε πως η κατανομή του περιεχόμενου των βυτιών ακολουθεί κανονική κατανομή.
- (β) Υπολογίστε το 95% διάστημα εμπιστοσύνης για το μέσο περιεχόμενο βυτίου με βάση το δείγμα.

- (γ) Χρησιμοποιώντας είτε διάστημα εμπιστοσύνης ή έλεγχο υπόθεσης, εξετάστε σε επίπεδο εμπιστοσύνης 99% αν το μέσο περιεχόμενο βυτίου μπορεί να είναι μικρότερο από 9.5 λίτρα.
- (δ) Θέλουμε να εκτιμήσουμε το μέσο περιεχόμενο βυτίου με ακρίβεια ± 0.05 λίτρα στο επίπεδο εμπιστοσύνης 95% (δηλαδή το εύρος του 95% διαστήματος εμπιστοσύνης για το μέσο περιεχόμενο να είναι 0.1 λίτρα). Με βάση το δείγμα που έχουμε, πόσες ακόμα μετρήσεις πρέπει να κάνουμε για να πετύχουμε αυτήν την ακρίβεια ;
13. Ένας μηχανικός θέλει να δοκιμάσει τη μεροληψία ενός μηχανήματος που μετράει PH. Έγιναν 11 μετρήσεις του PH μιας ουδέτερης ουσίας που έχει PH=7.0. Τα αποτελέσματα για το δείγμα έδωσαν:
- δειγματική μέση τιμή $\bar{x} = 7.05$,
 δειγματική τυπική απόκλιση $s = 0.05$.
 Θεωρούμε ότι η μέτρηση του PH του μηχανήματος ακολουθεί κανονική κατανομή αλλά με άγνωστη για μας διασπορά.
- (α) Υπολογίστε το 95% διάστημα εμπιστοσύνης για τη μέση μέτρηση PH στη συγκεκριμένη ουσία.
- (β) Υπολογίστε το 95% διάστημα εμπιστοσύνης για την τυπική απόκλιση της μέτρησης PH στη συγκεκριμένη ουσία.
- (γ) Είναι η πραγματική τιμή για τη συγκεκριμένη ουσία (PH=7.0) αποδεκτή ως μέση τιμή του μηχανήματος με βάση το δείγμα (σε επίπεδο εμπιστοσύνης 95%); Για να απαντήσετε στο ερώτημα κάνετε στατιστικό έλεγχο και συγκρίνετε με το αντίστοιχο διάστημα εμπιστοσύνης στο 13α'.
- (δ) Από παλιότερες μελέτες πιστεύεται πως η τυπική απόκλιση των μετρήσεων PH είναι 0.02. Είναι αυτή η τιμή αποδεκτή σε επίπεδο εμπιστοσύνης 95% με βάση το δείγμα ; Για να απαντήσετε στο ερώτημα κάνετε στατιστικό έλεγχο και συγκρίνετε με το αντίστοιχο διάστημα εμπιστοσύνης στο 13β'.
- (ε) Αν θα θέλαμε να εκτιμήσουμε τη μέση μέτρηση PH στη συγκεκριμένη ουσία με ακρίβεια ± 0.01 , δηλαδή το 95% διάστημα εμπιστοσύνης της μέσης μέτρησης PH να έχει εύρος 0.02, πόσες ακόμα μετρήσεις πρέπει να κάνουμε ;
14. Σε μια μελέτη για την επίδραση της χρήσης αζώτου στην ανάπτυξη ριζών δέντρων, φυτεύτηκαν 10 φιντάνια βαλανιδιάς επεξεργασμένα χωρίς άζωτο και 9 φιντάνια επεξεργασμένα με άζωτο σε κάποιο θερμοκήπιο. Όλες οι περιβαλλοντικές συνθήκες ήταν ίδιες. Μετά από ένα χρονικό διάστημα μετρήθηκε το βάρος των στελεχών (κοτσανιών) των φυτών και τα αποτελέσματα δίνονται στον παρακάτω πίνακα (σε γραμμάρια):

Χωρίς άζωτο	0.28	0.32	0.36	0.37	0.38	0.42	0.43	0.43	0.47	0.53
Με άζωτο	0.18	0.27	0.35	0.49	0.52	0.62	0.75	0.84	0.97	

- (α) Υπολογίστε και σχεδιάστε το κατάλληλο θηκόγραμμα από το δείγμα των 10 παρατηρήσεων που αναφέρονται στην επεξεργασία χωρίς άζωτο. Με βάση το θηκόγραμμα σχολιάστε αν φαίνεται η κατανομή του βάρους των στελεχών των φυτών (όταν η επεξεργασία γίνεται χωρίς άζωτο) να είναι κανονική.

Στη συνέχεια (κι ανεξάρτητα από την απάντησή σας στο 14α), υποθέστε πως το βάρος των στελεχών των φυτών (όταν η επεξεργασία γίνεται με ή χωρίς άζωτο) ακολουθεί κανονική κατανομή.

- (β) Υπολογίστε το 95% διάστημα εμπιστοσύνης για το μέσο βάρος των στελεχών των φυτών όταν η επεξεργασία γίνεται χωρίς άζωτο και όταν γίνεται με άζωτο.
- (γ) Κάνετε στατιστικό έλεγχο υπόθεσης για να ελέγξετε αν το μέσο βάρος των στελεχών των φυτών μπορεί να είναι μεγαλύτερο από 0.50 γραμμάρια, όταν η επεξεργασία γίνεται χωρίς άζωτο και όταν γίνεται με άζωτο. Σχολιάστε πως μπορούμε να οδηγηθούμε στα αποτελέσματα των ελέγχων από τα διαστήματα εμπιστοσύνης στο 14β'.
- (δ) Θέλουμε να διερευνήσουμε αν το μέσο βάρος των στελεχών των φυτών διαφέρει στις δύο επεξεργασίες, δηλαδή χωρίς άζωτο και με άζωτο. Περιγράψτε τον στατιστικό έλεγχο υπόθεσης που αντιστοιχεί στην παραπάνω διερεύνηση (χωρίς υπολογισμούς). Σχολιάστε αν τηρούνται οι προϋποθέσεις για να χρησιμοποιήσετε στατιστική ελέγχου από την κατανομή Student.

15. Για τον εντοπισμό μικρής ποσότητας ρύπων στο περιβάλλον έγιναν οι παρακάτω μετρήσεις περιεκτικότητας μολύβδου (σε $\mu\text{g/l}$) σε 11 δοκίμια

2.4	1.9	1.7	2.6	2.2	2.0	2.1	1.2	2.1	1.8	2.0
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

- (α) Υπολογίστε τη δειγματική διάμεσο, πρώτο και τρίτο τεταρτομόριο και σχεδιάστε το κατάλληλο θηκόγραμμα για το παραπάνω δείγμα. Με βάση το θηκόγραμμα σχολιάστε αν φαίνεται η κατανομή της περιεκτικότητας σε μόλυβδο να είναι κανονική.

Στη συνέχεια (κι ανεξάρτητα από την απάντησή σας στο 15α), υποθέστε πως η περιεκτικότητα σε μόλυβδο ακολουθεί κανονική κατανομή.

- (β) Υπολογίστε το 95% και 90% διάστημα εμπιστοσύνης για τη μέση περιεκτικότητα σε μόλυβδο.
- (γ) Αν τα δοκίμια δεν είναι ρυπογόνα θα πρέπει η μέση περιεκτικότητα του μολύβδου να είναι $1.5 \mu\text{g/l}$. Με βάση τα διαστήματα εμπιστοσύνης από το δείγμα που υπολογίσατε στο 15β', μπορείτε να δεχθείτε ότι τα δοκίμια δεν είναι ρυπογόνα (σε επίπεδο εμπιστοσύνης 95% και 90%);
- (δ) Πόσες ακόμα μετρήσεις πρέπει να κάνουμε για να εκτιμήσουμε σε 95% επίπεδο εμπιστοσύνης τη μέση περιεκτικότητα σε μόλυβδο με ακρίβεια $0.05 \mu\text{g/l}$ (δηλαδή το εύρος του 95% διαστήματος εμπιστοσύνης να είναι $0.1 \mu\text{g/l}$);

16. Για τη μεταφορά προϊόντων απαιτούνται ανθεκτικά κουτιά. Μια εταιρεία μεταφορών έκανε ένα πείραμα για να μετρήσει την ανθεκτικότητα ενός είδους κουτιών. Έγιναν μετρήσεις της αντοχής συμπίεσης (σε lb) σε 10 κουτιά που επιλέχτηκαν τυχαία και δίνονται στον παρακάτω πίνακα.

201	373	380	376	366	409	450	163	312	343
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

Θεωρούμε ότι η αντοχή συμπίεσης του κουτιού ακολουθεί κανονική κατανομή.

- (α') Από παλιότερες μετρήσεις γνωρίζουμε ότι η τυπική απόκλιση της αντοχής συμπίεσης του κουτιού είναι γύρω στις 100lb. Ελέγξτε σε επίπεδο εμπιστοσύνης 95% και χρησιμοποιώντας κατάλληλο διάστημα εμπιστοσύνης αν μπορούμε να δεχτούμε πως η τυπική απόκλιση είναι 100lb.
- (β) Υπολογίστε το 95% διάστημα εμπιστοσύνης για τη μέση αντοχή συμπίεσης του κουτιού υποθέτοντας πρώτα ότι η τυπική απόκλιση είναι γνωστή και ίση με 100lb και έπειτα υποθέτοντας ότι είναι άγνωστη. Συγκρίνετε τα δύο διαστήματα εμπιστοσύνης και σχολιάστε που οφείλεται η διαφορά στο εύρος τους.
- (γ) Θέλουμε να εκτιμήσουμε τη μέση αντοχή συμπίεσης με διπλάσια ακρίβεια, δηλαδή το εύρος του 95% διαστήματος εμπιστοσύνης της μέσης αντοχής συμπίεσης να γίνει το μισό από αυτό που βρήκαμε στο 16β' (για γνωστή τυπική απόκλιση). Πόσες ακόμα μετρήσεις πρέπει να κάνουμε για να το πετύχουμε ;

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°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 a και b 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°C σε υψόμετρο 1500μ; Σε υψόμετρο 2200μ;
2. Σε περιοχές με υψηλά επίπεδα βροχοπτώσεων είναι σημαντικό να έχει το σκυρόδεμα 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 Q ($E(U|X = x) = a + \beta x$). 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 a και b 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 $\mu\text{g}/\text{g}$) ξηρού βάρους, δηλαδή μg 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 μολύβδο καθώς κι αν φαίνεται η περιεκτικότητα σε μολύβδο να εξαρτάται γραμμικά από την κίνηση ροής. Δικαιολογείστε την απάντησή σας με βάση τον συντελεστή συσχέτισης και το διάγραμμα διασποράς.
- (β) Υπολογίστε τις σημειακές εκτιμήσεις a και b των παραμέτρων a και b της ευθείας παλινδρόμησης (με τη μέθοδο των ελαχίστων τετραγώνων).
- (γ) Με βάση το δείγμα, εκτιμείστε τη μέση περιεκτικότητα σε μολύβδο 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. Έγινε μια μελέτη για το ποσό ενός χημικού στοιχείου Y που διαλύεται σε 100 γραμμάρια νερού σε διαφορετικές θερμοκρασίες. Έγιναν 3 δοκιμές σε 6 διαφορετικές θερμοκρασίες και τα αποτελέσματα δίνονται στον παρακάτω πίνακα.

Θερμοκρασία σε $^{\circ}\text{C}$	Χημικό στοιχείο Y σε grams		
10	12	10	14
20	25	21	24
30	31	33	28
40	44	39	42
50	48	51	44
60	49	52	53

- (α) Σχηματίστε το κατάλληλο διάγραμμα διασποράς και σχολιάστε αν φαίνεται να υπάρχει εξάρτηση της διαλυτικότητας στο νερό του χημικού στοιχείου Y από τη θερμοκρασία του νερού, και αν ναι, τι είδους εξάρτηση είναι.
- (β) Υπολογίστε τις σημειακές εκτιμήσεις a και b των παραμέτρων a και b της ευθείας παλινδρόμησης (με τη μέθοδο των ελαχίστων τετραγώνων) για το πρόβλημα της γραμμικής εξάρτησης του ποσού χημικού στοιχείου Y που διαλύεται στο νερό από τη θερμοκρασία του νερού.
- (γ) Με βάση το δείγμα και το μοντέλο παλινδρόμησης, υπολογίστε τη σημειακή εκτίμηση (αν γίνεται) της αληθινής μέσης ποσότητας του χημικού στοιχείου Y που διαλύεται σε 100 γραμμάρια νερού θερμοκρασίας 45°C . Κάνετε το ίδιο όταν η θερμοκρασία του νερού είναι 80°C .
5. Έγινε μια μελέτη για το ποσό της μεταποιημένης ζάχαρης σε διάφορες θερμοκρασίες μέσω μιας συγκεκριμένης χημικής διαδικασίας. Σε κάθε επανάληψη της διαδικασίας καταγράφηκε η ποσότητα της μεταποιημένης ζάχαρης στην επιλεγμένη θερμοκρασία. Οι

μετρήσεις της θερμοκρασίας (κωδικοποιημένες σε τιμές από 1 ως 2) και της ποσότητας μεταποιημένης ζάχαρης (σε kgr) δίνονται στον παρακάτω πίνακα.

Θερμοκρασία (κωδικοποιημένη)	Ποσότητα μεταποιημένης ζάχαρης (σε kgr)
1.0	8.1
1.1	7.8
1.2	8.5
1.3	10.5
1.4	9.5
1.5	8.9
1.6	8.6
1.7	10.2
1.8	9.3
1.9	9.2
2.0	10.5

- (α') Σχηματίστε το κατάλληλο διάγραμμα διασποράς και υπολογίστε το συντελεστή συσχέτισης μεταξύ της ποσότητας μεταποιημένης ζάχαρης και της (κωδικοποιημένης) θερμοκρασίας. Σχολιάστε με βάση το διάγραμμα διασποράς και τον συντελεστή συσχέτισης αν φαίνεται να υπάρχει εξάρτηση της ποσότητας μεταποιημένης ζάχαρης Y από τη θερμοκρασία X , και αν ναι, τι είδους εξάρτηση είναι.
- (β') Υπολογίστε τις σημειακές εκτιμήσεις a και b των παραμέτρων a και b της ευθείας παλινδρόμησης (με τη μέθοδο των ελαχίστων τετραγώνων) για το πρόβλημα της γραμμικής εξάρτησης της ποσότητας μεταποιημένης ζάχαρης Y από τη (κωδικοποιημένη) θερμοκρασία X .
- (γ') Υπολογίστε επίσης το τυπικό σφάλμα $s = s_{Y|X}$ της παλινδρόμησης. Σχολιάστε αν η ποσότητα μεταποιημένης ζάχαρης 10.5kgr που παρατηρήσαμε για κωδικοποιημένη θερμοκρασία 1.3 είναι τυπική με βάση το μοντέλο παλινδρόμησης.
6. Γίνεται μια εργασία στη χημική επεξεργασία και καθαρισμό αποβλήτων που χαρακτηρίζονται εκτός άλλων από υψηλή στάθμη απαιτούμενου βιοχημικού οξυγόνου και πτητικές στερεές ουσίες. Θέλουμε να μελετήσουμε αν μπορούμε να προβλέψουμε το ποσοστό μείωσης του απαιτούμενου χημικού οξυγόνου όταν γνωρίζουμε το ποσοστό μείωσης των στερεών. Έγιναν 6 πειράματα και τα ποσοστά μείωσης (σε %) στα 6 δείγματα δίνονται στον παρακάτω πίνακα.

Στερεά ύλη	3	7	11	15	18	27
Χημικό οξυγόνο	5	11	21	16	16	28

- (α') Σχηματίστε το κατάλληλο διάγραμμα διασποράς και σχολιάστε αν φαίνεται να υπάρχει εξάρτηση της μείωσης του απαιτούμενου χημικού οξυγόνου από τη μείωση των στερεών ουσιών, και αν ναι, τι είδους εξάρτηση είναι.
- (β') Υπολογίστε τις σημειακές εκτιμήσεις a και b των παραμέτρων a και b της ευθείας παλινδρόμησης (με τη μέθοδο των ελαχίστων τετραγώνων) για το πρόβλημα της

γραμμικής εξάρτησης του ποσοστού μείωσης του απαιτούμενου χημικού οξυγόνου από το ποσοστό μείωσης των στερεών ουσιών. Σχηματίστε την ευθεία ελαχίστων τετραγώνων στο διάγραμμα διασποράς.

(γ) Με βάση το δείγμα και το μοντέλο στο 6β', υπολογίστε τη σημειακή εκτίμηση (αν γίνεται) του μέσου ποσοστού μείωσης του απαιτούμενου χημικού οξυγόνου αν το ποσοστό μείωσης των στερεών ουσιών είναι 20%. Κάνετε το ίδιο για ποσοστό μείωσης στερεών 40%.

7. Θέλουμε να διερευνήσουμε την επίδραση της θερμοκρασίας στην οποία γίνεται μια χημική αντίδραση στο ποσοστό μετατροπής της χημικής αντίδρασης. Τα αποτελέσματα από 9 πειράματα δίνονται στον παρακάτω πίνακα.

Θερμοκρασία [°C]	189	200	200	225	225	225	250	260	260
μετατροπή [%]	48	69	43	74	65	79	78	78	84

(α) Σχηματίστε το κατάλληλο διάγραμμα διασποράς και σχολιάστε με βάση το διάγραμμα διασποράς αν φαίνεται να εξαρτάται (και πως) το ποσοστό μετατροπής της χημικής αντίδρασης από τη θερμοκρασία στην οποία έγινε.

(β) Υπολογίστε τις σημειακές εκτιμήσεις a και b των παραμέτρων a και b της ευθείας παλινδρόμησης (με τη μέθοδο των ελαχίστων τετραγώνων) για το πρόβλημα της γραμμικής εξάρτησης του ποσοστού μετατροπής της χημικής αντίδρασης από τη θερμοκρασία στην οποία έγινε.

(γ) Υπολογίστε επίσης το τυπικό σφάλμα $s = s_{Y|X}$ της παλινδρόμησης (που εκτιμήσατε στο 7β'). Δώστε τα όρια για μια τυπική τιμή του ποσοστού μετατροπής της χημικής αντίδρασης που περιμένουμε όταν η χημική αντίδραση γίνεται σε θερμοκρασία 240°C.

8. Παρακάτω δίνονται 10 μετρήσεις της ταχύτητας του αέρα και του συντελεστή εξάτμισης σταγονιδίων καυσίμου που καίγεται σε έναν κινητήρα.

Ταχύτητα αέρα [cm/sec]	20	60	100	140	180	220	260	300	340	380
Συντελεστής εξάτμισης [mm ² /sec]	0.2	0.4	0.4	0.8	0.6	0.7	1.2	1.4	1.2	1.7

(α) Σχηματίστε το κατάλληλο διάγραμμα διασποράς και σχολιάστε με βάση το διάγραμμα διασποράς αν φαίνεται να συσχετίζεται (και πως) ο συντελεστής εξάτμισης σταγονιδίων καυσίμου με την ταχύτητα του αέρα στον κινητήρα.

(β) Υπολογίστε τον συντελεστή συσχέτισης μεταξύ της ταχύτητας του αέρα και του συντελεστή εξάτμισης σταγονιδίων καυσίμου. Σχολιάστε τι δηλώνει η τιμή του συντελεστή συσχέτισης που υπολογίσατε και αν συμφωνεί με τα σχόλια σας για την εξάρτηση του συντελεστή εξάτμισης από την ταχύτητα του αέρα στο 8α'.