

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. Ένας πολιτικός μηχανικός επιβλέπει την ποιότητα του νερού μετρώντας την ποσότητα των αιωρούμενων στερεών (αναλογία σωματιδίων ανά εκατομμύριο) σε δείγμα νερού από ένα 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 εκατομμύριο.

- (α) Μπορούμε να δεχτούμε την παραπάνω εμπειρική τιμή της τυπικής απόκλισης των αιωρούμενων στερεών γι αυτόν το ποταμό; Γιατί; Κάνετε τον κατάλληλο στατιστικό 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 εκατομμύριο). 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 ατσάλιού 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, $\pm 2\mu m$, δηλαδή το 95% διάστημα εμπιστοσύνης του μέσου μήκους ρωγμής να έχει εύρος $4\mu m$. 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. Στα πλαίσια μελέτης για τη μόλυνση του αέρα που έγινε σε έναν πειραματικό σταθμό μετρήθηκε η ποσότητα αιωρούμενης οργανικής ύλης διαλυτής στο βενζόλιο (σε μικρογραμμάρια ανά κυβικό μέτρο, $\mu g/m^3$) σε 12 δείγματα αέρα.

1.3	1.4	1.5	1.6	1.7	1.8	1.9	1.9	2.4	2.6	3.0	3.4
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

- (α) Υπολογίστε τη δειγματική διάμεσο, το πρώτο και τρίτο τεταρτομόριο και σχηματίστε το θηκόγραμμα για τα δεδομένα της ποσότητας μολυσματικής ύλης. Με βάση

το θηκόγραμμα σχολιάστε αν η ποσότητα αυτή φαίνεται να ακολουθεί κανονική κατανομή.

Στη συνέχεια θεωρούμε πως η ποσότητα μολυσματικής ύλης στον αέρα ακολουθεί κανονική κατανομή.

- (β) Υπολογίστε το 95% διάστημα εμπιστοσύνης για τη μέση ποσότητα μολυσματικής ύλης στον αέρα της περιοχής του σταθμού με βάση το δείγμα των παραπάνω μετρήσεων.
- (γ) Χρησιμοποιώντας είτε διάστημα εμπιστοσύνης ή έλεγχο υπόθεσης, εξετάστε σε επίπεδο εμπιστοσύνης 99% αν η μέση ποσότητα μολυσματικής ύλης στον αέρα της περιοχής του σταθμού δε ξεπερνάει τα $3 \mu\text{g}/\text{m}^3$.
- (δ) Χωρίς να κάνετε υπολογισμούς απαντήστε στο ερώτημα θγ' για επίπεδο εμπιστοσύνης 90% και δικαιολογήστε την απάντησή σας.

10. Μια εταιρεία ταξί θέλει να κάνει μια μεγάλη αγορά ελαστικών για το στόλο της και πρέπει να διαλέξει μεταξύ δύο τύπων (δηλαδή μάρκων ελαστικών) Α και Β. Για να εκτιμήσει τη διαφορά στους δύο τύπους έγινε ένα πείραμα χρησιμοποιώντας 12 ελαστικά από κάθε τύπο. Μετρήθηκαν τα χιλιόμετρα που διένυσαν τα ελαστικά μέχρι να φθαρούν. Τα αποτελέσματα για τα δύο δείγματα είναι:

Τύπος Α : μέση τιμή $\bar{x}_1 = 36300\text{km}$, τυπική απόκλιση $s_1 = 5000\text{km}$

Τύπος Β : μέση τιμή $\bar{x}_2 = 40100\text{km}$, τυπική απόκλιση $s_2 = 4100\text{km}$

Θεωρούμε ότι η «ζωή» του ελαστικού και των δύο τύπων (Α και Β) ακολουθεί κανονική κατανομή αλλά με άγνωστη για μας διασπορά, που είναι κοινή και για τους δύο τύπους.

- (α) Υπολογίστε το 99% διάστημα εμπιστοσύνης για τη μέση «ζωή» ελαστικού του τύπου Α. Κάνετε το ίδιο για τον τύπο Β.
- (β) Αν θέλουμε να εκτιμήσουμε τη μέση «ζωή» ελαστικού του τύπου Α με μεγαλύτερη ακρίβεια και συγκεκριμένα με ακρίβεια $\pm 500\text{km}$, δηλαδή το 99% διάστημα εμπιστοσύνης της μέσης «ζωής» ελαστικού του τύπου Α να έχει εύρος 1000km , πόσες ακόμα μετρήσεις πρέπει να κάνουμε;
- (γ) Χρησιμοποιώντας έλεγχο υπόθεσης εξετάστε σε επίπεδο σημαντικότητας $\alpha = 0.05$ αν η μέση «ζωή» ελαστικού διαφέρει στους δύο τύπους.

11. Θέλουμε να μελετήσουμε τη διάρκεια ενός συγκεκριμένου υπεραστικού οδικού δρομολογίου. Έγιναν 11 μετρήσεις σε διαφορετικές μέρες και ώρες και βρέθηκαν τα εξής αποτελέσματα για το δείγμα :

μέση τιμή $\bar{x} = 125\text{min}$, τυπική απόκλιση $s = 15\text{min}$

Θεωρούμε ότι η διάρκεια του δρομολογίου ακολουθεί κανονική κατανομή αλλά με άγνωστη για μας διασπορά.

- (α) Υπολογίστε το 95% διάστημα εμπιστοσύνης για τη μέση διάρκεια του δρομολογίου.
- (β) Υπολογίστε το 95% διάστημα εμπιστοσύνης για την τυπική απόκλιση της διάρκειας του δρομολογίου.

- (γ) Από προηγούμενες μελέτες πιστεύεται πως η μέση διάρκεια αυτού του δρομολογίου είναι 120min και η τυπική απόκλιση 5min. Είναι αυτές οι τιμές αποδεκτές σε επίπεδο εμπιστοσύνης 95% με βάση αυτό το δείγμα ;
- (δ) Αν θα θέλαμε να εκτιμήσουμε τη μέση διάρκεια του δρομολογίου με μεγαλύτερη ακρίβεια και συγκεκριμένα με ακρίβεια $\pm 2\text{min}$, δηλαδή το 95% διάστημα εμπιστοσύνης της μέσης διάρκειας του δρομολογίου να έχει εύρος 4min, πόσες ακόμα μετρήσεις πρέπει να κάνουμε ;
12. 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 ;
13. 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.