

Χρησιμοποίησα chatgpt και deepseek για τις απαντήσεις, πιθανόν να μην είναι σωστές κάποιες από αυτές

1. Ποιος είναι ο τύπος του δείγματος μέσης τιμής;

- α) $(\sum x_i)/n(\sum x_i)/n$
- β) $(\sum (x_i - \mu)^2)/(n-1)(\sum (x_i - \mu)^2)/(n-1)$
- γ) $\sigma/\sqrt{n}\sigma/\sqrt{n}$
- δ) $\mu + z \cdot \sigma\mu + z \cdot \sigma$

$$\bar{x} = \frac{\sum x_i}{n}$$

2. Αν ένα γεγονός έχει πιθανότητα 0.2, ποια είναι η πιθανότητα να μην συμβεί;

- α) 0.2
- β) **0.8**
- γ) 1.2
- δ) 0.5

3. Ποια κατανομή προσεγγίζεται από τη διωνυμική κατανομή όταν το πλήθος των δοκιμών είναι μεγάλο και η πιθανότητα επιτυχίας μικρή;

- α) Κανονική
- β) **Πουασσόν**
- γ) Ομοιόμορφη
- δ) χ^2

4. Το επίπεδο σημαντικότητας α σε έναν έλεγχο τοποθεσιών αντιπροσωπεύει:

- α) την πιθανότητα αποδοχής της εναλλακτικής υπόθεσης
- β) **την πιθανότητα λάθους τύπου I**
- γ) την πιθανότητα λάθους τύπου II
- δ) την πιθανότητα επιτυχίας

5. Σε ποια περίπτωση χρησιμοποιούμε το t-test αντί για το z-test;
- ο α) όταν γνωρίζουμε τη διακύμανση του πληθυσμού
 - ο β) όταν το δείγμα είναι πολύ μεγάλο
 - ο **γ) όταν το δείγμα είναι μικρό και η διακύμανση άγνωστη**
 - ο δ) ποτέ
6. Ποιος από τους παρακάτω δείκτες μέτρησης μεταβλητότητας είναι πιο ευαίσθητος σε ακραίες τιμές;
- ο α) εύρος (range)
 - ο β) ενδοτεταρτημοριακό εύρος (IQR)
 - ο **γ) διακύμανση (variance)**
 - ο δ) διάμεσος (median)
7. Σε μια απλή γραμμική παλινδρόμηση, ο συντελεστής B_1 αντιπροσωπεύει:
- ο α) τη σταθερά
 - ο β) το σημείο εκκίνησης της ευθείας
 - ο **γ) τον ρυθμό μεταβολής της εξαρτημένης μεταβλητής σε σχέση με την ανεξάρτητη**
 - ο δ) το συνολικό σφάλμα
8. Ποιο από τα παρακάτω δεν είναι προϋπόθεση της κανονικής κατανομής;
- ο α) συμμετρία γύρω από τη μέση τιμή
 - ο β) καμπύλη σε σχήμα καμπάνας
 - ο **γ) ασυμμετρία προς τα δεξιά**
 - ο δ) η μέση τιμή, διάμεσος και επικρατούσα ταυτίζονται

9. Όταν αυξάνουμε το μέγεθος του δείγματος, τι συμβαίνει στο πλάτος του διαστήματος εμπιστοσύνης;

- ο α) αυξάνεται
- ο **β) μειώνεται**
- ο γ) μένει σταθερό
- ο δ) μηδενίζεται

10. Ποια είναι η αναμενόμενη τιμή της μεταβλητής σε μια Πουασσόν κατανομή με μέσο όρο λ ;

- ο α) λ^2
- ο **β) λ**
- ο γ) $\sqrt{\lambda}$
- ο δ) 2λ

11. Η διακύμανση είναι:

- ο α) ο μέσος όρος των παρατηρήσεων
- ο β) το άθροισμα των παρατηρήσεων
- ο **γ) ο μέσος όρος των τετραγώνων των αποκλίσεων από τη μέση τιμή**
- ο δ) το γινόμενο των παρατηρήσεων

12. Η πιθανότητα της ένωσης δύο ασυμβίβαστων γεγονότων A και B είναι:

- ο **α) $P(A) + P(B)$**
- ο β) $P(A) * P(B)$
- ο γ) $P(A) - P(B)$
- ο δ) $P(A) + P(B) - P(A \cap B)$

13. Η τιμή του μέσου σε μια συμμετρική κατανομή είναι:

- ο α) μεγαλύτερη από τη διάμεσο
- ο **β) ίση με τη διάμεσο**
- ο γ) μικρότερη από τη διάμεσο
- ο δ) δεν σχετίζεται

14. Η τυπική απόκλιση:

- ο α) το τετράγωνο της διακύμανσης
- ο **β) η τετραγωνική ρίζα της διακύμανσης**
- ο γ) η μέση τιμή των παρατηρήσεων
- ο δ) η μέγιστη τιμή του δείγματος

15. Ποια είναι η βασική ιδιότητα της κανονικής κατανομής;

- ο α) είναι ασύμμετρη
- ο β) η μέση τιμή ισούται με την τυπική απόκλιση
- ο **γ) είναι συμμετρική γύρω από τον μέσο όρο**
- ο δ) έχει δύο κορυφές

16. Σε μια κανονική κατανομή, περίπου το 68% των παρατηρήσεων βρίσκονται:

- ο **α) μέσα σε 1 τυπική απόκλιση από τη μέση τιμή**
- ο β) μέσα σε 2 τυπικές αποκλίσεις
- ο γ) μέσα σε 3 τυπικές αποκλίσεις
- ο δ) σε οποιαδήποτε απόσταση

17. Ο δείκτης συσχέτισης r μετρά:

- α) τη μέση τιμή
- **β) το βαθμό γραμμικής σχέσης**
- γ) το εύρος των δεδομένων
- δ) τη διακύμανση

18. Στον έλεγχο υποθέσεων, η εναλλακτική υπόθεση συμβολίζεται συνήθως με:

- **α) H_1**
- β) H_0
- γ) μ
- δ) σ

19. Το διάστημα εμπιστοσύνης για το μέσο όρο μικρού δείγματος ($n < 30$) χρησιμοποιεί:

- α) κατανομή z
- **β) κατανομή t**
- γ) κατανομή F
- δ) κατανομή χ^2

20. Σε έναν έλεγχο υποθέσεων, αν $p\text{-value} < \alpha$, τότε:

- α) αποδεχόμαστε H_0
- **β) απορρίπτουμε H_0**
- γ) δεν παίρνουμε απόφαση
- δ) αλλάζουμε την τιμή του δείγματος

21. Ο τύπος για τον δείκτη μεταβλητότητας (CV) είναι:

- α) **$CV = \sigma / \bar{x} * 100\%$**
- β) $CV = \bar{x} / \sigma * 100\%$
- γ) $CV = \sigma * \bar{x}$
- δ) $CV = \sigma - \bar{x}$

22. Ποιο από τα παρακάτω αφορά μια δειγματοληψία χωρίς επανάθεση;

- α) η ίδια μονάδα μπορεί να επιλεγεί πολλές φορές
- β) **κάθε μονάδα επιλέγεται μόνο μία φορά**
- γ) δεν υπάρχει αντικειμενικότητα
- δ) το δείγμα είναι πάντοτε άπειρο

23. Το διάγραμμα κουτιού (boxplot) δίνει πληροφορίες για:

- α) την κανονικότητα των δεδομένων
- β) τη μέση τιμή μόνο
- γ) **τη διασπορά και τις ακραίες τιμές**
- δ) μόνο το εύρος

24. Σε μια χ^2 (chi-square) κατανομή, ποιο χαρακτηριστικό ισχύει;

- α) είναι συμμετρική
- β) **είναι πάντοτε θετική**
- γ) είναι κανονική
- δ) είναι δικόρφη

25. Το συνολικό άθροισμα τετραγώνων (SST) σε ανάλυση διακύμανσης (ANOVA) ισούται με:

- α) σφάλμα + μεταξύ ομάδων
- β) μέσος όρος + σφάλμα
- γ) ενδοομαδική διακύμανση μόνο
- δ) μέση τιμή + τυπική απόκλιση

26. Αν $r = 0$, τότε:

- α) υπάρχει ισχυρή συσχέτιση
- β) **δεν υπάρχει γραμμική συσχέτιση**
- γ) υπάρχει μη γραμμική συσχέτιση
- δ) τα δεδομένα είναι κανονικά κατανεμημένα

27. Ποιο από τα παρακάτω αφορά τη διακριτή τυχαία μεταβλητή;

- α) παίρνει συνεχείς τιμές
- β) **παίρνει συγκεκριμένες τιμές**
- γ) παίρνει άπειρως πολλές τιμές
- δ) δεν παίρνει καμία τιμή

28. Το εύρος (range) υπολογίζεται ως:

- α) μικρότερη τιμή – μεγαλύτερη τιμή
- β) **μεγαλύτερη τιμή – μικρότερη τιμή**
- γ) μέση τιμή $\times 2$
- δ) τετραγωνική ρίζα της διακύμανσης

29. Ποια από τις παρακάτω κατανομές είναι διωνυμική;

- **α) επιτυχία/αποτυχία**
- β) ύψος ανθρώπων
- γ) χρόνος αντίδρασης
- δ) πλήθος συμβάντων ανά μονάδα χρόνου

30. Ποια κατανομή χρησιμοποιούμε για καταμέτρηση αριθμού γεγονότων σε δεδομένο χρόνο ή χώρο;

- α) κανονική
- **β) Poisson**
- γ) διωνυμική
- δ) υπερδιωνυμική

31. Η μέση τιμή των δεδομένων 2,4,6,8,2,4,6,8,2,4,6,8 είναι:

- **α) 5**
- β) 6
- γ) 4
- δ) 7

32. Ποια μέθοδος δειγματοληψίας βασίζεται στην πιθανότητα;

- α) στοχευμένη
- β) συστηματική
- **γ) τυχαία**
- δ) χιονοστιβάδας

33. Το επίπεδο εμπιστοσύνης 95% αντιστοιχεί σε ποιο z-score περίπου;

- ο α) 1.28
- ο β) 1.64
- ο **γ) 1.96**
- ο δ) 2.58

34. Ποια από τις παρακάτω ποσότητες δεν είναι στατιστικό μέτρο κεντρικής τάσης;

- ο α) μέση τιμή
- ο β) διάμεσος
- ο γ) επικρατούσα τιμή
- ο **δ) τυπική απόκλιση**

35. Ο νόμος των μεγάλων αριθμών υποδεικνύει ότι:

- ο α) το δείγμα διαφέρει από τον πληθυσμό
- ο **β) το δείγμα προσεγγίζει την πραγματική μέση τιμή**
- ο γ) το δείγμα αυξάνει πάντα
- ο δ) το δείγμα αποτυγχάνει

36. Ποιο επίπεδο μέτρησης περιγράφει την κατηγορική ταξινόμηση χωρίς σειρά;

- ο **α) ονομαστικό**
- ο β) τακτικό
- ο γ) αναλογικό
- ο δ) διάστημα

37. Πότε χρησιμοποιούμε το F-test;

- ο α) για έλεγχο δύο μέσων
- ο **β) για έλεγχο δύο διακυμάνσεων**
- ο γ) για έλεγχο μιας πιθανότητας
- ο δ) για έλεγχο της διαμέσου

38. Ποια είναι η εξίσωση της απλής γραμμικής παλινδρόμησης;

- ο **$y = \beta_0 + \beta_1 \cdot x$**

39. Σε έναν έλεγχο υποθέσεων, το σφάλμα τύπου II είναι:

- ο α) απόρριψη σωστής H_0
- ο **β) μη απόρριψη λανθασμένης H_0**
- ο γ) απόρριψη σωστής H_1
- ο δ) αποδοχή σωστής H_0

40. Ο συντελεστής προσδιορισμού R^2 δείχνει:

- ο α) τη διασπορά
- ο β) την ακρίβεια του δείγματος
- ο **γ) το ποσοστό της μεταβλητότητας που εξηγείται**
- ο δ) την τυπική απόκλιση

41. Αν η διακύμανση ενός δείγματος είναι μηδέν, τότε:

- ο α) όλες οι παρατηρήσεις είναι διαφορετικές
- ο **β) όλες οι παρατηρήσεις είναι ίδιες**
- ο γ) η τυπική απόκλιση είναι 1
- ο δ) η διάμεσος είναι μηδέν

42. Το επίπεδο σημαντικότητας α σε έλεγχο υποθέσεων είναι:

- ο α) η πιθανότητα αποδοχής H_0
- ο **β) η πιθανότητα απόρριψης σωστής H_0**
- ο γ) το δείγμα
- ο δ) το μέγεθος του δείγματος

43. Ποια κατανομή πλησιάζει το n τείνει στο συν άπειρο;

- ο **α) διωνυμική**
- ο β) Πουασσόν
- ο γ) υπερδιωνυμική
- ο δ) όλες οι παραπάνω

44. Σε μια κανονική κατανομή, το 68% των δεδομένων βρίσκεται μέσα σε:

- ο **α) 1 τυπική απόκλιση από τη μέση τιμή**
- ο β) 2 τυπικές αποκλίσεις
- ο γ) 3 τυπικές αποκλίσεις
- ο δ) 4 τυπικές αποκλίσεις

45. Σε ποια περίπτωση χρησιμοποιούμε το t-test αντί για το z-test;

- ο α) όταν το δείγμα είναι πολύ μεγάλο
- ο β) όταν γνωρίζουμε την πληθυσμιακή διακύμανση
- ο **γ) όταν το δείγμα είναι μικρό και η διακύμανση άγνωστη**
- ο δ) όταν τα δεδομένα είναι διακριτά

46. Η διακύμανση ενός δείγματος υπολογίζεται διαιρώντας με:

- ο α) n
- ο **β) $n - 1$**
- ο γ) $n + 1$
- ο δ) n^2

47. Η τυπική απόκλιση είναι:

- ο α) το τετράγωνο της διακύμανσης
- ο β) η διακύμανση επί δύο
- ο **γ) η τετραγωνική ρίζα της διακύμανσης**
- ο δ) η μέση τιμή επί δύο

48. Η διάμεσος επηρεάζεται πολύ από:

- ο α) ακραίες τιμές
- ο β) το μέσο όρο
- ο γ) το εύρος
- ο **δ) δεν επηρεάζεται από ακραίες τιμές**

49. Σε ένα διάγραμμα διασποράς (scatterplot), τι παριστάνει κάθε σημείο;

- ο α) μια μέση τιμή
- ο **β) μια παρατήρηση (ζεύγος τιμών)**
- ο γ) μια γραμμική σχέση
- ο δ) την τυπική απόκλιση

50. Ποιος δείκτης μετράει τη γραμμική σχέση δύο μεταβλητών;

- α) δείκτης μεταβλητότητας
- **β) συντελεστής συσχέτισης Pearson**
- γ) δείκτης διακύμανσης
- δ) δείκτης επικινδυνότητας

51. ένα διάστημα εμπιστοσύνης για τη μέση τιμή ενός πληθυσμού με άγνωστη διακύμανση και δείγμα μεγέθους $n < 30$ βασίζεται στη:

- α) κανονική κατανομή
- β) F-κατανομή
- **γ) t-κατανομή**
- δ) χ^2 -κατανομή

52. η πιθανότητα να κάνουμε σφάλμα τύπου 2 αυξάνεται όταν:

- α) αυξάνεται το μέγεθος δείγματος
- **β) μειώνεται το επίπεδο σημαντικότητας α**
- γ) η απόκλιση από την H_0 είναι πολύ μεγάλη
- δ) το p-value είναι μικρό

53. στην ανάλυση διακύμανσης (ANOVA), αν η τιμή F είναι πολύ μεγάλη τότε:

- α) οι μέσοι όροι των ομάδων είναι ίδιοι
- β) δεν υπάρχει σημαντική διαφορά μεταξύ των ομάδων
- **γ) υπάρχει τουλάχιστον μία ομάδα με διαφορετικό μέσο**
- δ) το δείγμα είναι μικρό

54. ο έλεγχος Kolmogorov–Smirnov χρησιμοποιείται για να:

- α) εξετάσει αν δύο πληθυσμοί έχουν ίδιες διακυμάνσεις
- **β) ελέγξει την κανονικότητα μιας κατανομής**
- γ) υπολογίσει την τυπική απόκλιση
- δ) υπολογίσει το εύρος

55. το p-value εκφράζει:

- α) τη μέγιστη πιθανότητα απόρριψης της H_1
- **β) την πιθανότητα να παρατηρήσουμε ένα δείγμα τόσο ακραίο όσο αυτό αν η H_0 είναι αληθής**
- γ) τη διαφορά μεταξύ των μέσων όρων
- δ) τη δύναμη του τεστ

56. ένα μεγάλο δείγμα $n=1000$ έχει μέση τιμή 50 και τυπική απόκλιση 10. ποια είναι η πιθανότητα μια νέα παρατήρηση να απέχει πάνω από 30 μονάδες από τη μέση τιμή;

- α) (περίπου) 0.05
- **β) 0.002**
- γ) 0.01
- δ) 0.001

57. ο δείκτης R^2 στην παλινδρόμηση δείχνει:

- α) την ακρίβεια των παρατηρήσεων
- β) την πιθανότητα απόρριψης της H_0
- **γ) το ποσοστό μεταβλητότητας της εξαρτημένης μεταβλητής που εξηγείται από τις ανεξάρτητες**
- δ) τη μέση τιμή των σφαλμάτων

58. αν η τυπική απόκλιση είναι πολύ μεγαλύτερη από τη μέση τιμή, τότε ο δείκτης μεταβλητότητας (CV) θα είναι:

- α) $< 10\%$
- **β) μεγάλος**
- γ) μηδενικός
- δ) ίσος με 1

59. στη χ^2 κατανομή, η προσέγγιση στην κανονική κατανομή είναι καλύτερη όταν:

- α) οι βαθμοί ελευθερίας είναι πολύ μικροί
- **β) οι βαθμοί ελευθερίας είναι μεγάλοι**
- γ) ο αριθμός ομάδων είναι 1
- δ) τα δεδομένα είναι διακριτά

60. η μέθοδος Bonferroni χρησιμοποιείται για:

- α) έλεγχο κανονικότητας
- **β) έλεγχο υποθέσεων σε επαναλαμβανόμενες δοκιμές (διόρθωση πολλαπλών συγκρίσεων)**
- γ) υπολογισμό της διακύμανσης
- δ) εκτίμηση της τυπικής απόκλισης

61. αν δύο μεταβλητές έχουν συντελεστή συσχέτισης $r = -0.95$, τότε:

- α) δεν έχουν γραμμική σχέση
- **β) έχουν ισχυρή αρνητική γραμμική σχέση**
- γ) έχουν αδύναμη συσχέτιση
- δ) είναι στατιστικά ανεξάρτητες

62. σε έναν έλεγχο ANOVA με 4 ομάδες και συνολικά 20 παρατηρήσεις πόσοι είναι οι βαθμοί ελευθερίας μεταξύ των ομάδων (d):

- α) 4
- **β) 3**
- γ) 16
- δ) 19

63. η κατανομή Student's t τείνει στην κανονική κατανομή όταν:

- α) το δείγμα είναι πολύ μικρό
- β) η διακύμανση είναι γνωστή
- **γ) το n τείνει στο άπειρο**
- δ) το επίπεδο σημαντικότητας είναι 0.01

64. ο στατιστικός έλεγχος Levene χρησιμοποιείται για να ελέγξει:

- α) ισότητα μέσων όρων
- **β) ισότητα διασπορών**
- γ) ισότητα τυπικών αποκλίσεων
- δ) κανονικότητα

65. σε μια απλή γραμμική παλινδρόμηση, το σφάλμα e (residual) ορίζεται ως:

- α) η προβλεπόμενη τιμή
- β) η πραγματική τιμή
- **γ) η διαφορά μεταξύ πραγματικής και προβλεπόμενης τιμής**
- δ) η μέση τιμή

66. ο έλεγχος Shapiro–Wilk ελέγχει:

- α) αν η διακύμανση είναι μικρή
- β) αν τα δεδομένα είναι ανεξάρτητα
- **γ) την κανονικότητα των δεδομένων**
- δ) τη γραμμική συσχέτιση

67. αν $r^2 = 0.81$ στην παλινδρόμηση τότε:

- α) το 81% των παρατηρήσεων είναι ίδιες
- **β) το 81% της μεταβλητότητας εξηγείται από το μοντέλο**
- γ) το 81% των τιμών είναι σωστές
- δ) η παλινδρόμηση έχει μικρή ακρίβεια

68. ποιο από τα παρακάτω είναι μη παραμετρικό τεστ;

- α) t-test
- β) z-test
- **γ) Mann–Whitney U**
- δ) ANOVA

69. ποια από τις παρακάτω μεταβλητές είναι ονομαστική;

- α) ύψος
- β) ημερομηνία
- **γ) φύλο**
- δ) βάρος

70. αν έχουμε μεγάλο δείγμα και χρησιμοποιούμε t-test αντί για z-test, το αποτέλεσμα:

- α) είναι σημαντικά λανθασμένο
- β) είναι πιο συντηρητικό
- **γ) είναι περίπου ίδιο**
- δ) δεν είναι αποδεκτό

71. η συνθήκη ανεξαρτησίας των παρατηρήσεων είναι κρίσιμη για:

- α) υπολογισμό του εύρους
- **β) την εφαρμογή του t-test**
- γ) κατασκευή ραβδογράμματος
- δ) υπολογισμό της διαμέσου

72. το 95% διάστημα εμπιστοσύνης για μια μέση τιμή δεν περιέχει το μηδέν. τι συμπεραίνουμε;

- α) η μέση τιμή είναι σίγουρα 0
- β) το αποτέλεσμα δεν είναι στατιστικά σημαντικό
- **γ) η μηδενική υπόθεση $H_0: \mu=0$ απορρίπτεται**
- δ) δεν υπάρχει επαρκές δείγμα

73. η διαφορά μεταξύ παραμετρικών και μη παραμετρικών τεστ έγκειται κυρίως:

- α) στην πολυπλοκότητα
- β) στην ταχύτητα υπολογισμού
- **γ) στις υποθέσεις για την κατανομή**
- δ) στον αριθμό μεταβλητών

74. σε πολυμεταβλητή γραμμική παλινδρόμηση, το πρόβλημα της πολυσυγγραμμικότητας αφορά:

- α) τη συνάρτηση πυκνότητας
- β) τη συσχέτιση μεταξύ εξαρτημένων μεταβλητών
- **γ) τη συσχέτιση μεταξύ ανεξάρτητων μεταβλητών**
- δ) τη διακύμανση των σφαλμάτων

75. ποιο είναι το βασικό αποτέλεσμα του κεντρικού οριακού θεωρήματος (Central Limit Theorem):

- α) το δείγμα αντιπροσωπεύει τον πληθυσμό
- **β) ο μέσος δείγματος έχει κανονική κατανομή (για μεγάλο n)**
- γ) η τυπική απόκλιση αυξάνεται
- δ) η διάμεσος πλησιάζει το μηδέν

76. η μεθοδολογία bootstrap βασίζεται στη:

- α) αναλυτική λύση εξισώσεων
- β) στατιστική θεωρία F-test
- **γ) επαναληπτική δειγματοληψία από το δείγμα**
- δ) προσομοίωση με κανονική κατανομή

77. ο συντελεστής συσχέτισης r ίσου με μηδέν σημαίνει ότι:

- α) οι μεταβλητές δεν έχουν καμία σχέση
- **β) οι μεταβλητές δεν σχετίζονται γραμμικά**
- γ) οι μεταβλητές είναι ταυτόσημες
- δ) υπάρχει αιτιακή σχέση

78. σε έλεγχο χ^2 ανεξαρτησίας με $p\text{-value} = 0.03$ και $\alpha=0.05$:

- α) αποδεχόμαστε την ανεξαρτησία
- **β) απορρίπτουμε την ανεξαρτησία**
- γ) δεν μπορούμε να αποφασίσουμε
- δ) χρειαζόμαστε μεγαλύτερο δείγμα

79. αν η τυπική απόκλιση είναι 0, τότε:

- α) η διακύμανση είναι 1
- **β) όλες οι παρατηρήσεις είναι ίδιες**
- γ) το δείγμα είναι μεγάλο
- δ) η μέση τιμή είναι 0

80. η απόφαση απόρριψης της H_0 όταν $p > \alpha$:

- α) είναι σωστή
- β) είναι σφάλμα τύπου 1
- **γ) δεν είναι σωστή**
- δ) είναι σφάλμα τύπου 2

81. η τυπική απόκλιση του μέσου όρου (standard error) μειώνεται όταν:

- α) μειώνεται το μέγεθος δείγματος
- β) αυξάνεται η διακύμανση
- **γ) αυξάνεται το μέγεθος δείγματος**
- δ) αυξάνεται το επίπεδο σημαντικότητας

82. αν το p-value ενός t-test είναι 0.12 και $\alpha=0.05$ τότε:

- α) απορρίπτουμε την H_0
- β) αποδεχόμαστε την H_1
- **γ) δεν απορρίπτουμε την H_0**
- δ) πρέπει να αλλάξουμε το επίπεδο σημαντικότητας

83. στην ανάλυση παλινδρόμησης, το σφάλμα (residual) θα πρέπει να έχει:

- α) ασυμμετρία
- β) τάση
- **γ) κανονική κατανομή με μέσο 0**
- δ) τυχαίες εξάρσεις

84. η δοκιμή Bartlett χρησιμοποιείται για να ελέγξει:

- α) ισότητα μέσων όρων
- β) κανονικότητα
- **γ) ισότητα διασπορών**
- δ) ισοκατανομή τυχαίων σφαλμάτων

85. αν το 99% διάστημα εμπιστοσύνης είναι πιο ευρύ από το 95%, αυτό συμβαίνει γιατί:

- α) το δείγμα είναι μικρότερο
- β) αυξάνεται η διακύμανση
- **γ) υψηλότερη εμπιστοσύνη απαιτεί μεγαλύτερο εύρος**
- δ) μειώθηκε το α

86. ποια από τις παρακάτω περιπτώσεις αφορά την υπερεφαρμογή (overfitting) σε παλινδρόμηση:

- α) χρήση μόνο μιας ανεξάρτητης μεταβλητής
- β) κακή εκτίμηση της τυπικής απόκλισης
- **γ) χρήση παραπάνω πολλών μεταβλητών**
- δ) μη κανονικότητα στα σφάλματα

87. η δύναμη ενός στατιστικού ελέγχου (power) είναι:

- α) $1-\alpha$
- **β) $1-\beta$**
- γ) $\alpha+\beta$
- δ) $\alpha\cdot\beta$

88. η ανάλυση συσταδοποίησης (cluster analysis) είναι τεχνική:

- α) περιγραφικής στατιστικής
- β) υποθετικής στατιστικής
- **γ) εξερεύνησης μη ετικετοποιημένων δεδομένων**
- δ) γραμμικής πρόβλεψης

89. η μη κανονικότητα στα σφάλματα επηρεάζει:

- α) τη μέση τιμή
- **β) την εγκυρότητα των παραμετρικών τεστ**
- γ) το εύρος
- δ) την ονομαστική μεταβλητή

90. η ανάλυση παραγόντων (factor analysis) χρησιμοποιείται για:

- α) μείωση της διακύμανσης
- β) έλεγχο της ισότητας μέσω
- **γ) εντοπισμό υποκείμενων δομών σε δεδομένα**
- δ) ανάλυση κατηγορικών μεταβλητών

91. σε ανάλυση παλινδρόμηση αν τα σφάλματα εμφανίζουν αυτοσυσχέτιση τότε:

- α) το μοντέλο είναι υπερεκτιμημένο
- β) οι παρατηρήσεις είναι ανεξάρτητες
- **γ) παραβιάζεται βασική υπόθεση του μοντέλου**
- δ) τα υπολείμματα έχουν μέση τιμή 1

92. η δοκιμή Durbin–Watson χρησιμοποιείται για:

- α) έλεγχο της ισότητας μέσω
- **β) έλεγχο αυτοσυσχέτισης σφαλμάτων στην παλινδρόμηση**
- γ) έλεγχο κανονικότητας
- δ) έλεγχο πολυσυγγραμμικότητας

93. ποια από τις παρακάτω μεθόδους είναι κατάλληλη για κατηγορική εξαρτημένη μεταβλητή:

- α) απλή γραμμική παλινδρόμηση
- **β) λογιστική παλινδρόμηση (logistic regression)**
- γ) ανάλυση διακύμανσης
- δ) συσταδοποίηση

94. ο δείκτης VIF (Variance Inflation Factor) χρησιμοποιείται για:

- α) εκτίμηση της ετεροσκεδαστικότητας
- **β) μέτρηση της πολυσυγγραμμικότητας**
- γ) έλεγχο της τυπικής απόκλισης
- δ) ανίχνευση outlier

95. αν το p-value είναι 0.0001 και $\alpha=0.05$, τότε το αποτέλεσμα θεωρείται:

- α) μη σημαντικό
- β) οριακά σημαντικό
- **γ) εξαιρετικά σημαντικό στατιστικό**
- δ) δεν μπορούμε να αποφασίσουμε

96. σε ανάλυση διακύμανσης οι ενδοομαδικές διαφορές μετρούν:

- α) τη διακύμανση μεταξύ ομάδων
- β) τη συνολική διακύμανση
- **γ) τη διακύμανση εντός των ομάδων**
- δ) τη μεταβλητότητα της εξαρτημένης μεταβλητής

97. σε ανάλυση κύριων συνιστωσών (PCA) οι συνιστώσες:

- α) είναι γραμμικά εξαρτημένες
- β) έχουν πάντα ίδια διακύμανση
- **γ) είναι μη συσχετισμένες μεταξύ τους**
- δ) χρησιμοποιούνται μόνο για κατηγορικά δεδομένα

98. ποια από τις παρακάτω επιλογές είναι προϋπόθεση για χρήση του z-test;

- α) άγνωστη διακύμανση
- β) μικρό δείγμα
- γ) κανονικότητα των δεδομένων
- **δ) γνωστή πληθυσμιακή διακύμανση**

99. η ετεροσκεδαστικότητα (heteroscedasticity) σημαίνει ότι:

- **α) τα σφάλματα έχουν διαφορετικές διασπορές**
- β) οι μεταβλητές είναι μη κανονικές
- γ) η μέση τιμή των σφαλμάτων είναι μηδέν
- δ) υπάρχει outlier

100. η μέθοδος Bonferroni διορθώνει:

- α) ετεροσκεδαστικότητα
- **β) σφάλματα τύπου 1 σε πολλαπλούς ελέγχους**
- γ) σφάλματα τύπου 2
- δ) πολυσυγγραμμικότητα