

Γραμμικά Μοντέλα

Θέμα 1. Έστω το απλό γραμμικό μοντέλο

$$Y_i = \beta_0 + \beta_1 \frac{Z_i}{X_i} + \epsilon_i, \quad i = 1, \dots, n,$$

όπου Y_i η απαντητική μεταβλητή, X_i συνεχής ερμηνευτική μεταβλητή, Z_i ψευδομεταβλητή (δίτιμη μεταβλητή που παίρνει τις τιμές 0 και 1) και $\epsilon_i \sim N(0, \sigma^2)$ και ανεξάρτητα για $i \neq j$.

(α) Υπολογίστε τις ΕΜΠ $\hat{\beta}_0$ και $\hat{\beta}_1$ των β_0 και β_1 .

(β) Να βρείτε την κατανομή των εκτιμητριών $\hat{\beta}_0$ και $\hat{\beta}_1$ και τη συνδιακύμανσή τους.

(γ) Υπολογίστε διαστήματα εμπιστοσύνης συντελεστή $1 - \alpha$ για τις παραμέτρους β_0 και β_1 .

(δ) Υπολογίστε διάστημα μέσης πρόβλεψης συντελεστή $1 - \alpha$ για την απαντητική μεταβλητή αν η τιμή της συνεχούς ερμηνευτικής μεταβλητής είναι X_0 και η ψευδομεταβλητή παίρνει την τιμή 1.

(ε) Τι εκφράζουν οι παράμετροι β_0 και β_1 στο παραπάνω γραμμικό μοντέλο; Τι σημαίνει το να μην είναι η καθεμία από αυτές στατιστικά σημαντική; Σχεδιάστε ένα σημειόγραμμα παρατηρήσεων που θα μπορούσαν να προέρχονται από το μοντέλο αυτό με τις δύο παραμέτρους του να είναι στατιστικά σημαντικές.

Θέμα 2. Α. Ένας μηχανικός ενδιαφέρεται να κατασκευάσει μια μπαταρία για κάποια συσκευή που εκτίθεται σε μεγάλες μεταβολές θερμοκρασίας. Ο χρόνος ζωής της μπαταρίας επηρεάζεται από τον τύπο του υλικού από το οποίο θα κατασκευαστεί (τύπος I, II, III) και από τη θερμοκρασία. Ο μηχανικός επέλεξε να ελέγξει τους τρεις τύπους υλικού σε τρία επίπεδα θερμοκρασίας (10°F , 70°F , 130°F), χρησιμοποιώντας 4 μπαταρίες για κάθε συνδυασμό επιπέδων υλικού και θερμοκρασίας. Τα βασικά ερωτήματα που ενδιαφέρεται να απαντήσει ο μηχανικός είναι τα εξής.

1. Τί επιδράσεις έχουν ο τύπος υλικού και η θερμοκρασία στο χρόνο ζωής της μπαταρίας;

2. Προκαλεί το μέγιστο επίπεδο θερμοκρασίας μικρότερη διάρκεια ζωής της μπαταρίας από τα άλλα δύο επίπεδα θερμοκρασίας;

(α) Διατυπώστε μαθηματικά το ερώτημα 1. Θεωρήστε ένα κατάλληλο στατιστικό μοντέλο για αυτό το πρόβλημα, αναφέρετε τις υποθέσεις του μοντέλου και θεωρήστε κατάλληλους ελέγχους υποθέσεων που να διατυπώνουν το ερώτημα.

(β) Συμπληρώστε τον παρακάτω πίνακα ANOVA ως προς δύο παράγοντες με αλληλεπίδραση για το παραπάνω πρόβλημα και μετά χρησιμοποιήστε τον για να απαντήσετε στο ερώτημα 1 ($\alpha=5\%$).

Πηγή Μεταβλητότητας	Άθροισμα Τετραγώνων	Βαθμοί Ελευθερίας	Μέσα Τετράγωνα	F
Τύπος Υλικού (A)	10670			
Θερμοκρασία (B)	39120			
Αλληλεπίδραση AB	9610			
Σφάλματα				
Σύνολο	77640			

(γ) Διατυπώστε μαθηματικά, με τη βοήθεια ενός κατάλληλου ελέγχου υποθέσεων, και απαντήστε το ερώτημα 2 ($\alpha=5\%$). Δίνεται ότι η μέση διάρκεια ζωής της μπαταρίας (σε ημέρες) για τα τρία επίπεδα θερμοκρασίας είναι, αντίστοιχα, 120, 110, 80.

(δ) Υπολογίστε ένα 95% διάστημα εμπιστοσύνης για τη διαφορά στη μέση διάρκεια ζωής της μπαταρίας ανάμεσα στο πρώτο και το δεύτερο επίπεδο θερμοκρασίας. Είναι αυτή στατιστικά σημαντική;

Δίνονται

$$Z_{0.05} = 1.645, Z_{0.025} = 1.96,$$

$$t_{0.05}(12) = 1.78, t_{0.025}(12) = 2.18, t_{0.05}(27) = 1.70, t_{0.025}(27) = 2.05$$

$$F_{0.05}(2, 12) = 3.89, F_{0.05}(3, 12) = 3.49, F_{0.05}(4, 12) = 3.26, F_{0.05}(5, 12) = 3.11$$
$$F_{0.05}(2, 27) = 3.35, F_{0.05}(3, 27) = 2.96, F_{0.05}(4, 27) = 2.73, F_{0.05}(5, 27) = 2.57$$