

Άσκηση 1

Η κατανάλωση, σε λίτρα βενζίνης, των αυτοκινήτων θεωρητικά εξαρτάται από τη χώρα παραγωγής τους. Θέλοντας να δούμε αν αυτό ισχύει, πραγματοποιήσαμε ανάλυση διακύμανσης (ANOVA). Μερικά από τα αποτελέσματα δίνονται στον επόμενο πίνακα.

Πίνακας ANOVA

Πηγή Μεταβλητότητας	Άθροισμα Τετραγώνων -SS	Βαθμοί Ελευθερίας-df	Μέσο Τετράγωνο-MS	F
Παράγοντας (TR)	7.985,084	2	3.992,542	
Σφάλματα (E)	16.056,288	394	40,752	91,97
Ολική	24.041,372	396		

Να προσδιορίσετε το πλήθος των διαφορετικών χωρών παραγωγής, το πλήθος των αυτοκινήτων που συμμετείχαν στην έρευνα, και να ελέγξετε αν η κατανάλωση διαφοροποιείται ανάλογα με τη χώρα παραγωγής, σε επίπεδο σημαντικότητας $\alpha=0,05$

Λύση

Πλήθος διαφορετικών χωρών παραγωγής: $k-1=2$ άρα $k=3$

Πλήθος των αυτοκινήτων: $N-k=394$ άρα $N=394+k=394+3=397$ ή

$N-1=396$ επομένως $N=396+1=397$

F=91,97 θα υπολογιστεί η κριτική τιμή $F_{\alpha,k-1,N-k}$ από τον πίνακα της F κατανομής. Συγκεκριμένα η τιμή $F_{0.05,2,394}=3,00$

Αν $F > F_{\alpha,k-1,N-k}$ απορρίπτουμε την μηδενική υπόθεση της ισότητας.

Στη συγκεκριμένη περίπτωση έχουμε: $F=91,97 > 3,00=F_{0.05,2,394}$ και συνεπώς συμπεραίνουμε ότι η κατανάλωση διαφοροποιείται ανάλογα με τη χώρα παραγωγής

Άσκηση 2

Ένα χαρακτηριστικό του κόστους χρήσης των αυτοκινήτων είναι το κόστος επισκευής μετά από ατύχημα. Προσπαθώντας να γίνει πιο ανταγωνιστική μια αυτοκινητοβιομηχανία δοκιμάζει διαφορετικούς τύπους προφυλακτήρων σε ένα αυτοκίνητο μεσαίου μεγέθους. Κάθε τύπος προφυλακτήρα τοποθετήθηκε σε έναν ίδιο αριθμό όμοιων αυτοκινήτων, καθένα από τα οποία οδηγήθηκε να συγκρουστεί με ταχύτητα 10 χιλ. την ώρα πάνω σε έναν τοίχο και καταγράφηκε το κόστος επισκευής μετά τη σύγκρουση. Στον πίνακα που ακολουθεί παρουσιάζονται τα αποτελέσματα. Να προσδιοριστεί το πλήθος των διαφορετικών τύπων προφυλακτήρων, το πλήθος των αυτοκινήτων που πήραν μέρος στο πείραμα και σε επίπεδο σημαντικότητας 5% να ελέγξετε αν το μέσο κόστος επισκευής των διαφόρων τύπων προφυλακτήρων διαφέρει.

Πίνακας ANOVA

Πηγή Μεταβλητότητας	Άθροισμα Τετραγώνων-SS	Βαθμοί Ελευθερίας-df	Μέσο Τετράγωνο-MS	F
Παράγοντας (TR)	149.517,0	3	49.839,000	4,009
Σφάλματα (E)	447.589,4	36	12.431,778	
Ολική	597.106,4	39		

Άσκηση 3

Οι κάτοικοι μιας τουριστικής περιοχής αξιολόγησαν τις επιπτώσεις του τουρισμού, από πολύ αρνητικές έως πολύ θετικές. Θέλοντας να δούμε αν οι απόψεις τους διαφοροποιούνται ανάλογα με το επάγγελμά τους, πραγματοποιήσαμε ανάλυση διακύμανσης (ANOVA). Μερικά από τα αποτελέσματα δίνονται στον επόμενο πίνακα.

Πίνακας ANOVA

Πηγή Μεταβλητότητας	Άθροισμα Τετραγώνων-SS	Βαθμοί Ελευθερίας-df	Μέσο Τετράγωνο-MS	F
Παράγοντας (TR)	47,8	4	11,95	29,875
Σφάλματα (E)	178,0	445	0,4	
Ολική	225,8	449		

Να προσδιορίσετε το πλήθος των κατοίκων που συμμετείχαν στην έρευνα, το πλήθος των διαφορετικών επαγγελμάτων και να ελέγξετε αν οι απόψεις τους διαφοροποιούνται ανάλογα με το επάγγελμά τους, σε επίπεδο σημαντικότητας $\alpha=0,05$.

Άσκηση 4

Τα τελευταία χρόνια η έλλειψη εμπιστοσύνης στην ταχυδρομική υπηρεσία των ΗΠΑ έχει οδηγήσει πολλές επιχειρήσεις στη συνεργασία με ιδιωτικές εταιρείες courier. Για να επιλέξει ανάμεσα σε τέτοιες εταιρείες μια επιχείρηση έστειλε σε κάθε μια σε διαφορετικές ώρες της ημέρας μια επιστολή με σταθερό προορισμό ένα απομακρυσμένο σημείο της πόλης και κατέγραψε το συνολικό χρόνο παράδοσης της κάθε επιστολής. Τα αποτελέσματα παρουσιάζονται στον πίνακα που ακολουθεί.

Πίνακας ANOVA

Πηγή Μεταβλητότητας	Άθροισμα Τετραγώνων-SS	Βαθμοί Ελευθερίας-df	Μέσο Τετράγωνο-MS	F
Παράγοντας (TR)	204,222	2	102,111	2,048
Σφάλματα (E)	1645,333	33	49,859	
Ολική	1849,556	35		

Να προσδιοριστεί το πλήθος των εταιρειών courier, το πλήθος των συνολικών ωρών και να ελέγξετε αν υπάρχουν διαφορές στο χρόνο παράδοσης μεταξύ των εταιρειών, σε επίπεδο σημαντικότητας 5%.

Άσκηση 5

Το τμήμα προσωπικού μιας εταιρείας υπολογιστών εξετάζει αν υπάρχει διαφορά στην ικανότητα πώλησης μεταξύ των αποφοίτων διαφορετικών κατευθύνσεων. Για τον λόγο αυτό επέλεξε δείγματα από αποφοίτους διαφορετικών κατευθύνσεων και κατέγραψε τη συνολική προμήθεια από τις πωλήσεις που πραγματοποίησε ο καθένας τον προηγούμενο χρόνο. Τα αποτελέσματα παρουσιάζονται στον επόμενο πίνακα.

Πίνακας ANOVA

Πηγή Μεταβλητότητας	Άθροισμα Τετραγώνων-SS	Βαθμοί Ελευθερίας-df	Μέσο Τετράγωνο-MS	F
Παράγοντας (TR)	10,261	2	5,130	,090
Σφάλματα (E)	3247,007	57	56,965	
Ολική	3257,268	59		

Να προσδιορίσετε το πλήθος των κατευθύνσεων, το πλήθος των αποφοίτων και να ελέγξετε, σε επίπεδο 3% αν υπάρχει διαφορά στην ικανότητα πώλησης μεταξύ των αποφοίτων διαφορετικών κατευθύνσεων.

Άσκηση 6

Πόσο διαφέρουν οι ειδικότητες των ιατρών ως προς το χρόνο που αφιερώνουν στη φροντίδα των ασθενών τους; Μια έρευνα επέλεξε ένα τυχαίο δείγμα ιατρών διαφόρων ειδικοτήτων και κατέγραψε τον αριθμό των ωρών που αφιέρωσε ο καθένας στη φροντίδα των ασθενών του στη διάρκεια μιας εβδομάδας. Τα αποτελέσματα εμφανίζονται στον επόμενο πίνακα.

Πίνακας ANOVA

Πηγή Μεταβλητότητας	Άθροισμα Τετραγώνων- SS	Βαθμοί Ελευθερίας- df	Μέσο Τετράγωνο- MS	F
Παράγοντας (TR)	1406,390	4	351,598	5,169
Σφάλματα (E)	11903,578	175	68,020	
Ολική	13309,968	179		

Να προσδιορίσετε το πλήθος των ειδικοτήτων, το πλήθος των ιατρών και να ελέγξετε, σε επίπεδο 2% αν υπάρχει διαφορά στις ειδικότητες ως προς το χρόνο που αφιερώνουν στη φροντίδα των ασθενών τους.

Άσκηση 7

Κατασκευαστής μπρούντζινων γραμματοκιβωτίων δέχθηκε παράπονα ότι τα προϊόντα του σκουριάζουν πρόωρα και εντόπισε το πρόβλημα στην κακή ποιότητα της βαφής. Για να αντικαταστήσει τη βαφή που χρησιμοποιεί επέλεξε εναλλακτικές βαφές και έβαψε με κάθε μια ορισμένο αριθμό γραμματοκιβωτίων και κατέγραψε σε πόσες μέρες κάθε γραμματοκιβώτιο εμφάνισε το πρώτο ίχνος σκουριάς. Τα αποτελέσματα στον επόμενο πίνακα.

Πίνακας ANOVA

Πηγή Μεταβλητότητας	Άθροισμα Τετραγώνων- SS	Βαθμοί Ελευθερίας- df	Μέσο Τετράγωνο- MS	F
Παράγοντας (TR)	17250,512	4	4312,628	3,319
Σφάλματα (E)	155941,120	120	1299,509	
Ολική	173191,632	124		

Να προσδιορίσετε το πλήθος των βαφών, το πλήθος των γραμματοκιβωτίων και να ελέγξετε, σε επίπεδο 1% αν υπάρχει διαφορά στον χρόνο εμφάνισης σκουριάς.

Άσκηση 8 (ANOVA_8)

A, B, C και D είναι 4 διαφορετικά μοντέλα Η/Υ. Ο χρόνος επισκευής, σε λεπτά δίνεται στον παρακάτω πίνακα.

A	B	C	D
62	45	66	68
48	57	53	52
49	44	47	70
50	40	49	55
63	39	59	65
69	54	60	53
57	50	67	59

Ο μέσος χρόνος επισκευής των 4 μοντέλων κομπιούτερ, σε 5% επίπεδο σημαντικότητας, είναι διαφορετικός;

Λύση

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	700,714	3	233,571	4,051	,018
Within Groups	1383,714	24	57,655		
Total	2084,429	27			

Η συγκεκριμένη τεχνική αφορά τον έλεγχο ισότητας των μέσων τιμών

H₀: Ο μέσος χρόνος επισκευής των τεσσάρων μοντέλων είναι ίσος

H₁: Ο μέσος χρόνος επισκευής των τεσσάρων μοντέλων δεν είναι ίσος

Επειδή το Sig. της F τιμής (4,051) είναι $0,018 < 0,05$ απορρίπτεται η μηδενική υπόθεση και συνεπώς ο μέσος χρόνος δεν είναι ίσος σε όλα τα μοντέλα.

Άσκηση 9^η

Επιλέχθηκε τυχαία ένα δείγμα από ένα πεπερασμένο πληθυσμό και καταγράφηκε το εισόδημά και το επίπεδο εκπαίδευσης των ατόμων του δείγματος. Στη συνέχεια πραγματοποιήθηκε Ανάλυση Διακύμανσης (ANOVA) προκειμένου να προσδιοριστεί αν το μέσο εισόδημα των ατόμων είναι ίσο, ανεξαρτήτως του επιπέδου εκπαίδευσης. Στον παρακάτω πίνακα δίνονται μερικά από τα αποτελέσματα της Ανάλυσης.

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	74.969,471	4	18.742,36775	13,35	0,000
Within Groups	1.186.276,91	845	1.403,878		
Total	1.261.246,381				

Ζητούμενα:

1. Καθορίστε την Μηδενική (H_0) και την Εναλλακτική (H_1) Υπόθεση του παρόντος ελέγχου.
2. Προσδιορίστε το πλήθος των επιπέδων εκπαίδευσης και το μέγεθος του δείγματος;
3. Είναι ίσο το μέσο εισόδημα, ανεξαρτήτως του επιπέδου εκπαίδευσης;

Απαντήσεις

1. Η συγκεκριμένη τεχνική αφορά τον έλεγχο ισότητας των μέσων τιμών
 H_0 : Το μέσο εισόδημα όλων των επιπέδων εκπαίδευσης είναι ίσο
 H_1 : Το μέσο εισόδημα όλων των επιπέδων εκπαίδευσης δεν είναι ίσο
2. $k-1= 4$ άρα $k=5$ επίπεδα εκπαίδευσης. $N-k = 845$ επομένως $N-5= 845$ και τέλος $N = 850$ είναι το μέγεθος του δείγματος
3. Επειδή το Sig. της F τιμής είναι $<0,05$ απορρίπτεται η μηδενική υπόθεση και συνεπώς το μέσο εισόδημα δεν ίσο σε όλα τα επίπεδα εκπαίδευσης.

Άσκηση 10^η

Στην προσπάθεια να διαπιστώσουμε αν το επίπεδο εξοικείωσης με τη χρήση των υπολογιστών είναι ίδιο ανεξαρτήτως επιπέδου εκπαίδευσης πραγματοποιήθηκε Ανάλυση Διακύμανσης (ANOVA). Στους πίνακες που ακολουθούν δίνονται τα αποτελέσματα της ανάλυσης.

ANOVA

MEAN_FAM					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	10,413	4	2,603	5,561	0,023
Within Groups	241,553	516	,468		
Total	251,966				

MEAN_FAM			
Tukey HSD			
Επίπεδο εκπαίδευσης	N	Subset for alpha = 0.05	
		1	2
Μέση	44	3,1894	
Κολλέγιο/IEK	56	3,3274	
Στοιχειώδη	57	3,3743	
Πανεπιστήμιο/Τ.Ε.Ι	279	3,5006	3,5006
M.Sc/PhD	85		3,7137
Sig.		,063	,365

1. Διατυπώστε την Μηδενική και την Εναλλακτική υπόθεση του παραπάνω ελέγχου.
2. Πόσα είναι τα επίπεδα εκπαίδευσης και ποιο το μέγεθος του δείγματος;
3. Παρατηρώντας τα αποτελέσματα, ο ισχυρισμός ότι το επίπεδο εξοικείωσης, σε επίπεδο σημαντικότητας 5%, είναι το ίδιο ανεξαρτήτως επιπέδου εκπαίδευσης είναι αληθινός;
4. Σε περίπτωση αρνητικής απάντησης ποια επίπεδα εκπαίδευσης έχουν το ίδιο επίπεδο εξοικείωσης;

Απαντήσεις

1. **Μηδενική Υπόθεση (H_0):** Το επίπεδο εξοικείωσης είναι το ίδιο ανεξαρτήτως επιπέδου εκπαίδευσης

Εναλλακτική Υπόθεση (H_1): Το επίπεδο εξοικείωσης δεν είναι το ίδιο ανεξαρτήτως επιπέδου εκπαίδευσης

2. $k - 1 = 4$ και συνεπώς $k = 5$ επίπεδα εκπαίδευσης. $N - k = 516$, $N - 5 = 516$, $N = 516 + 5 = 521$ άτομα.
3. Από τον πίνακα ANOVA παρατηρούμε ότι η τιμή Sig. του Στατιστικού F είναι $0,023 < 0,05$. Συνεπώς, ο ισχυρισμός δεν ισχύει και επομένως το επίπεδο εξοικείωσης δεν είναι το ίδιο ανεξαρτήτως επιπέδου εκπαίδευσης.
4. Στον δεύτερο πίνακα μπορούμε να διακρίνουμε ότι οι κατηγορίες «Μέση εκπαίδευση», «Κολλέγιο/IEK», «Στοιχειώδη» και «Πανεπιστήμιο/Τ.Ε.Ι.» έχουν το ίδιο επίπεδο εξοικείωσης (Sig.= $0,063 > 0,05$).
Επίσης οι κατηγορίες «Πανεπιστήμιο/Τ.Ε.Ι.» και «M.Sc/PhD» έχουν το ίδιο επίπεδο εξοικείωσης (Sig.= $0,365 > 0,05$).

Άσκηση 11^η: Επιλέχθηκε τυχαία ένα δείγμα από ένα πεπερασμένο πληθυσμό και καταγράφηκε το εισόδημά και η θέση εργασίας των ατόμων του δείγματος. Στη συνέχεια πραγματοποιήθηκε Ανάλυση Διακύμανσης (ANOVA) προκειμένου να προσδιοριστεί αν το μέσο εισόδημα των ατόμων είναι ίσο, ανεξαρτήτως της θέσεως εργασίας. Στον παρακάτω πίνακα δίνονται μερικά από τα αποτελέσματα της Ανάλυσης.

Ζητούμενα:

1. Να συμπληρώσετε τα σκιασμένα κελιά.
2. Πόσες είναι οι θέσεις εργασίας;
3. Ποιο είναι το μέγεθος του δείγματος;
4. Είναι ίσο το μέσο εισόδημα, ανεξαρτήτως της θέσεως εργασίας;

ANOVA

α. Salary

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1.196,844	5	239,3688	3,2259	0,007
Within Groups	30.051,000	405	74,2		
Total	31.247,844	410			

- b. Υπάρχουν 6 θέσεις εργασίας καθώς: $k - 1 = 5$, $k = 5 + 1 = 6$.
- c. Το δείγμα αποτελείται από: $N - 1 = 410$, $N = 410 + 1 = 411$.
- d. Όχι, το μέσο εισόδημα δεν είναι ίδιο καθώς το Sig. του F- statistic είναι $< 0,05$.