

ΠΑΛΑΙΑ ΘΕΜΑΤΑ

(Α-Ι) [5 μονάδες] Μία εταιρία δοκιμάζει ένα καινούργιο προϊόν σε 30 άτομα, από τα οποία 15 δοκιμάζουν το νέο προϊόν και οι υπόλοιποι 15 το υπάρχων προϊόν. Οι καταναλωτές δεν γνωρίζουν αν τους χορηγήθηκε το νέο ή το παλιό προϊόν. Οι καταναλωτές αξιολογούν το προϊόν σε κλίμακα από 1 ως 3: **1 = Ικανοποίηση, 2 = Ουδέτερη στάση, 3 = Δυσαρέστηση**. Οι αξιολογήσεις είναι οι εξής:

Παλιό προϊόν: 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 3, 3, 3

Νέο προϊόν: 1, 1, 1, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 3, 3

Για να ελέγξουμε την υπόθεση $H_0: \mu_1 = \mu_2$ έναντι $H_1: \mu_1 \neq \mu_2$ εφαρμόζουμε τέσσερις μεθόδους: α) ένα t-test με ίσες διακυμάνσεις, β) ανάλυση διακύμανσης με έναν παράγοντα, γ) τον προσημικό βαθμολογικό έλεγχο Wilcoxon, και δ) τον έλεγχο χ^2 . Τα αποτελέσματα των μεθόδων είναι τα εξής:

α) t-test με ίσες διακυμάνσεις

Έλεγχος t δύο δειγμάτων με υποτιθέμενες ίσες διακυμάνσεις

	Μεταβλητή 1	Μεταβλητή 2
Μέσος	1.667	2.333?
Κοινή Διακύμανση	0.667	0.667
Μέγεθος δείγματος	15	15
βαθμοί ελευθερίας	28	
t	-2.236?	
P(T<=t) δίπλευρη	0.033?	
t κριτική, τιμή	2.048?	

β) ανάλυση διακύμανσης με έναν παράγοντα

Ανάλυση διακύμανσης κατά ένα παράγοντα

ΣΥΜΠΕΡΑΣΜΑ

Ομάδες	Πλήθος	Άθροισμα	Μέσος όρος	Διακύμανση
prod1	15	25	1,67	0.67
prod2	15	35	2.33	0.67

ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ

Προέλευση διακύμανσης	SS	βαθμοί ελευθερίας	MS	F	π-τιμή	F-κριτική τιμή
Μεταξύ ομάδων	3,333	1?	3.333?	5?	0,033	4.196?
Μέσα στις ομάδες	18.667?	28?	0.667?			

Σύνολο 22

γ) τον προσημικό βαθμολογικό έλεγχο Wilcoxon

Ομάδα	Αξιολόγηση	Ταξινόμηση	Βαθμός	Ομάδα 1	Ομάδα 2
1	1	1	?	?	
1	?	2	6	?	
1	?	?	6	6	
1	?	4	?	?	
1	1	?	?	?	
1	1	6	6	?	
1	?	7	?	?	
1	?	?	?	?	
2	1	9	6		?
2	?	10	?		?
2	1	11	6		?
1	2	?	15.5	?	
1	2	?	?	?	
1	?	?	?	?	
1	2	15	?	15.5	
2	?	16	15.5		15.5
2	?	?	?		?
2	2	?	?		?
2	2	?	15.5		?
1	3	20	25	?	
1	?	21	?	?	
1	3	?	25	?	
2	3	?	?		?
2	?	24	?		?
2	3	?	?		?
2	?	?	?		25
2	?	?	?		25
2	3	?	?		?
2	?	29	25		?
2	?	?	?		?
Συνολικός Βαθμός				185	280

**Wilcoxon προσημικός
βαθμολογικός έλεγχος**

Δείγμα	1	2	E(T)	232.521?
Μέγεθος Δείγματος	15	15	σ_T	24.110?
Συνολικός Βαθμός	185?	280	z-στατιστική	-1.9702?
άλφα	0.05		P(Z<=z) δίπλευρη	0.0488?
			z κριτική τιμή δίπλευρη	1.9600?

δ) τον έλεγχο χ^2

Παρατηρημένες Προϊόν	Αξιολόγηση			Σύνολο
	1	2	3	
1	8	4	3	15
2	3	4	8	15
Σύνολο	11	8	11	30

Αναμενόμενες Προϊόν	Αξιολόγηση			Σύνολο
	1	2	3	
1	5.5	4	5.5	15
2	5.5	4?	5.5?	15
Σύνολο	11	8	11	30

χ^2 -τιμή = 4.545?

Βαθμοί ελευθερίας = 2?

π-τιμή = 0.10?

(I) [15 μονάδες] Ποια από τις τέσσερις μεθόδους είναι η πιο κατάλληλη για αυτήν την ανάλυση; Εξηγήστε γιατί. Ποια είναι η επόμενη πιο κατάλληλη που θα προτιμούσατε; Εξηγήστε γιατί. Είναι κάποιες από τις μεθόδους ισοδύναμες; Εξηγήστε γιατί.

(II) [10 μονάδες] Τι συμπέρασμα εξάγεται από τις παραπάνω αναλύσεις σχετικά με το νέο προϊόν;

(III) [65 μονάδες] Στην μέθοδο γ) για τον προσημικό βαθμολογικό έλεγχο Wilcoxon, συμπληρώστε τον 1^ο πίνακα με τα κατάλληλα νούμερα όπου υπάρχουν ερωτηματικά.

(IV) [110 μονάδες] Στις τέσσερις αναλύσεις δείξτε πως υπολογίζονται τα έντονα νούμερα με ερωτηματικό στα δεξιά. Γράψτε τον κατάλληλο τύπο (όπου υπάρχει τύπος) και αντικαταστήτε τα κατάλληλα νούμερα στους τύπους. Οι μέθοδοι, οι μονάδες και τα ζητούμενα νούμερα δίνονται ακολούθως.

α) ένα t-test με ίσες διακυμάνσεις

(1) [5 μονάδες] 2.333

(2) [5 μονάδες] -2.236

(3) [5 μονάδες] 0.033(βρείτε την πλησιέστερη τιμή από τον κατάλληλο πίνακα)

(4) [5 μονάδες] 2.048

β) ανάλυση διακύμανσης με έναν παράγοντα

(1) [5 μονάδες] 18.667

(2) [5 μονάδες] 1

(3) [5 μονάδες] 28

(4) [5 μονάδες] 3.333

(5) [5 μονάδες] 0.667

(6) [5 μονάδες] 5

(7) [5 μονάδες] 4.196

γ) τον προσημικό βαθμολογικό έλεγχο Wilcoxon

(1) [5 μονάδες] 185

(2) [5 μονάδες] 232.521

(3) [5 μονάδες] 24.110

(4) [5 μονάδες] -1.9702

(5) [5 μονάδες] 0.0488

(6) [5 μονάδες] 1.96

δ) τον έλεγχο χ^2

(1) [5 μονάδες] 4

(2) [5 μονάδες] 5.5

(3) [5 μονάδες] 4.545

(4) [5 μονάδες] 2

(5) [5 μονάδες] 0.10

1. [5 μονάδες] Επιλέγουμε δείγμα μεγέθους 16 από έναν κανονικό πληθυσμό. Ο δειγματοληπτικός μέσος που προκύπτει είναι 30 και η δειγματοληπτική τυπική απόκλιση είναι 20. Ποια από τις παρακάτω θα χρησιμοποιήσετε ως κριτική τιμή για να κατασκευάσετε το 99% διάστημα εμπιστοσύνης για τον μέσο του πληθυσμού;

(α) 2.602

(β) 2.576

(γ) 2.947

(δ) 2.921

(ε) Κανένα από τα παραπάνω

ΕΞΗΓΗΣΤΕ

2. [5 μονάδες] Τυχαίο δείγμα 144 παρατηρήσεων έχει μέσο 20, διάμεσο 21, και τύπο 22. Η τυπική απόκλιση στον πληθυσμό είναι 4.8. Το 99% διάστημα εμπιστοσύνης για τον μέσο του πληθυσμού είναι με ακρίβεια πρώτου δεκαδικού

(α) [19.0 , 21.0]

(β) [19.2 , 20.8]

(γ) [19.2 , 22.8]

(δ) [21.2 , 22.8]

(ε) Κανένα από τα παραπάνω

ΕΞΗΓΗΣΤΕ

3-5. [5 μονάδες] **Εκφώνηση:** Για να ελεγχθεί η μηδενική υπόθεση $H_0: \mu = 20$ έναντι της εναλλακτικής $H_a: \mu > 20$ σε πληθυσμό με τυπική απόκλιση $\sigma = 12$ επιλέχθηκε τυχαίο δείγμα μεγέθους $n = 36$. Ο δειγματοληπτικός μέσος είναι $\bar{x} = 24.6$.

3. [5 μονάδες] Η παρατηρούμενη τιμή της στατιστικής ελέγχου είναι

(α) 2.3

(β) 0.38

(γ) -2.3

(δ) -0.38

(ε) Κανένα από τα παραπάνω

4. [5 μονάδες] Η p-τιμή του ελέγχου είναι μεταξύ

(α) 0.005 και 0.01 (β) 0.01 και 0.025 (γ) 0.025 και 0.05

(δ) 0.05 και 0.10

(ε) Κανένα από τα παραπάνω

5. [5 μονάδες] Αν ο έλεγχος διεξαχθεί σε επίπεδο εμπιστοσύνης $\alpha=0.05$, τότε συμπεραίνουμε ότι

(α) η μηδενική υπόθεση θα γίνει απορριφθεί

(β) η μηδενική υπόθεση θα αποδεκτή

(γ) η εκφώνηση δεν δίνει όλες τις απαραίτητες πληροφορίες για να διεξαχθεί ο έλεγχος

(δ) αφού $\bar{x} = 24.6 > 20$ απορρίπτουμε την μηδενική και δεν χρειάζεται η εφαρμογή του ελέγχου

(ε) π-τιμή δεν μπορεί να υπολογιστεί

6-9. [25 μονάδες] **Εκφώνηση:** Για να ελεγχθεί στατιστικά εάν ένα φάρμακο έχει αποτέλεσμα στην καταπολέμηση της γρίπης χρησιμοποιήθηκε τυχαίο δείγμα 150 ασθενών. Σε μια ομάδα χορηγήθηκε το φάρμακο, ενώ σε μια άλλη χορηγήθηκαν δισκία ζάχαρης ώστε να ελεγχθούν οι πιθανές επιδράσεις ψυχολογικών παραγόντων. Τα αποτελέσματα της μελέτης καταγράφονται στον παρακάτω πίνακα:

Ασθενείς που	Έγιναν καλά	Δεν έγιναν καλά
Έλαβαν το φάρμακο	70	10

Έλαβαν δισκία ζάχαρης	20	50
-----------------------	----	----

6. [5 μονάδες] Η αναμενόμενη συχνότητα των ασθενών που έλαβαν το φάρμακο και έγιναν καλά είναι:

- (α) 70 (β) 150 (γ) 28 (δ) 48 (ε) Κανένα από τα παραπάνω.

ΕΞΗΓΗΣΤΕ

7. [10 μονάδες] Η τιμή της στατιστικής ελέγχου είναι περίπου

- (α) 10.08 (β) 54.02 (γ) -1.96 (δ) 1.645 (ε) Κανένα από τα παραπάνω

ΕΞΗΓΗΣΤΕ

8. [5 μονάδες] Αν ο έλεγχος διεξαχθεί σε επίπεδο σημαντικότητας $\alpha=0.05$, τότε η κριτική τιμή της στατιστικής ελέγχου είναι:

- (α) 3.84 (β) 7.81 (γ) 5.99 (δ) 9.34 (ε) Κανένα από τα παραπάνω

ΕΞΗΓΗΣΤΕ

9. [5 μονάδες] Η τιμή πιθανότητας (p-τιμή) του ελέγχου είναι

- (α) $p < 0.005$ (β) $0.005 < p < 0.01$ (γ) $0.01 < p < 0.025$ (δ) $0.025 < p < 0.05$

(ε) Κανένα από τα παραπάνω

ΕΞΗΓΗΣΤΕ

ΔΙΝΕΤΑΙ ΟΤΙ:

$$\chi^2 = \sum_{i=1}^k \frac{(f_i - e_i)^2}{e_i}$$

Δίνεται ότι ακολουθεί χ^2 κατανομή με $(r-1)(c-1)$ βαθμούς ελευθερίας.

Με f και e συμβολίζουμε τις συχνότητες που παρατηρούμε και τις αναμενόμενες ((σύνολο γραμμής)×(σύνολο στήλης)/σύνολο), αντιστοίχως. Επίσης, r και c είναι ο αριθμός των γραμμών και των στηλών.

10. [150 μονάδες] **Εκφώνηση:** Για την σύγκριση τριών νέων φαρμάκων, αναθέτουμε τα τρία φάρμακα σε τρεις ομάδες των τριών ατόμων ανεξαρτήτως. Δηλαδή ο κάθε ασθενής παίρνει ένα μόνο φάρμακο και όλα τα άτομα της ίδιας ομάδας παίρνουν το ίδιο φάρμακο. Τα αποτελέσματα δίνονται στον ακόλουθο πίνακα. Η ύπαρξη αρνητικών αριθμών οφείλεται στην σύγκριση των καθαρών αποτελεσμάτων με την αναμενόμενη τιμή ενός άλλου υπαρκτού φαρμάκου.

ΠΙΝΑΚΑΣ Α

Ομάδα 1	Ομάδα 2	Ομάδα 3
-1	-3	0
0	0	2
1	0	4

Αναλύουμε τα δεδομένα με τους εξής δύο τρόπους: (α) Ανάλυση Διακύμανσης με έναν παράγοντα (ANOVA), και (β) Τον έλεγχο Kruskal-Wallis και τα αποτελέσματα καταγράφονται αντιστοίχως στους Πίνακες Β και Γ. Αναφέρεται αναλυτικά, σύντομα αλλά ξεκάθαρα, την μηδενική και την εναλλακτική υπόθεση και τα συμπεράσματα και των δύο μεθόδων αφού συμπληρώσετε τα κενά κελιά με τα ερωτηματικά στους Πίνακες Α και Γ. Εξηγήστε στο τέλος των Πινάκων Β και Γ τους

τύπους που χρησιμοποιήσατε και αντικαταστήσετε τα ανάλογα νούμερα, δεν χρειάζεται να δείξετε τις πράξεις αναλυτικά.

ΠΙΝΑΚΑΣ Β

Περιγραφική Στατιστική

Ομάδες	<i>n</i>	Μέσοι	Διακυμάνσεις
Ομάδα 1	?	0	1
Ομάδα 2	3	?	3
Ομάδα 3	3	2	4

ANOVA

Πηγή Απόκλισης	SS Άθροισμα Τετραγώνων	d.f. β.ε.	MS Μέσος Τετραγώνων	F	p-value Π-τιμή	F critical - κριτική τιμή
Αγωγές	14	?	?	?	0,15	?
Σφάλμα	?	?	?			
Σύνολο	30					

ΠΙΝΑΚΑΣ Γ

Ομάδα	Αποκρίσεις	Βαθμοί	Μέσοι	Ομάδα 1	Ομάδα 2	Ομάδα 3
			Βαθμοί			
2	?	1	1		1	
1	-1	?	?	?		
1	0	3	4.5	4.5		
2	?	?	?		?	
2	?	?	?		?	
3	0	6	4.5			?
1	1	7	7	7		
3	?	?	?			?
3	4	9	9			9
				?	?	?
				T1	T2	T3

H - Τιμή
Ελέγχου
?

Κριτική
Τιμή
?

Δίνεται ότι

$$H = \left[\frac{12}{n(n+1)} \sum_{j=1}^k \frac{T_j^2}{n_j} \right] - 3(n+1)$$

ακολουθεί χ^2 κατανομή με $\kappa-1$ βαθμούς ελευθερίας.

1) [20 μονάδες] Μεγάλη επιχείρηση με αλυσίδα καταστημάτων ειδών πολυτελείας, για να αποφασίσει αν θα ανοίξει κατάστημα σε μια περιοχή, θέλει να ελέγξει αν το μέσο διαθέσιμο οικογενειακό εισόδημα των κατοίκων της περιοχής είναι τουλάχιστον $\mu = 6,75$ €. Αν το τυχαίο δείγμα $n = 60$ οικογενειών υπολόγισε μέσο εισόδημα $\bar{x} = 6,16$ € και τυπική απόκλιση $\sigma = 4,70$ €, ελέγξτε την υπόθεση $H_0 : \mu = 6,75$ € με εναλλακτική $H_a : \mu < 6,75$ €, με επίπεδο σημαντικότητας $\alpha = 0,05$. Υπολογίστε την p-τιμή του ελέγχου.

2) [15 μονάδες] Αν ο χρόνος συναρμολόγησης ενός μηχανήματος ακολουθεί την κανονική κατανομή και σε τυχαίο δείγμα $n = 31$ προϊόντων υπολογίσαμε μέσο χρόνο συναρμολόγησης $\bar{x} = 16$ min και τυπική απόκλιση $s = 2$ min, να εκτιμηθεί διάστημα εμπιστοσύνης για τη μέση τιμή πληθυσμού με επίπεδο εμπιστοσύνης 0,95.

(3) [30 μονάδες] Από δύο διαφορετικές πόλεις παίρνουμε δύο δείγματα μεγέθους $n_1 = 150$ και $n_2 = 100$ αποσκοπώντας να βρούμε το ποσοστό αυτών που χρησιμοποιούν ένα νέο προϊόν καθαριότητας. 54 άτομα από την πρώτη πόλη και 28 άτομα από τη δεύτερη δήλωσαν ότι χρησιμοποιούν το νέο αυτό προϊόν.

(α) [10 μονάδες] Κατασκευάστε ένα 90% διάστημα εμπιστοσύνης για τη διάφορα μεταξύ των πληθυσμιακών αναλογιών. Δικαιολογήστε τον τύπο που θα χρησιμοποιήσετε ελέγχοντας τις αντίστοιχες υποθέσεις.

(β) [20 μονάδες] Να ελεγχθεί η υπόθεση της ισότητας των αναλογιών πληθυσμού σε επίπεδο σημαντικότητας $\alpha=0.1$. Υπολογίστε την ρ -τιμή.

(4) [35 μονάδες] Ο παρακάτω πίνακας παρουσιάζει τις θέσεις απασχόλησης για τους πτυχιούχους τμημάτων Διοίκησης Επιχειρήσεων και Μηχανολόγων ανά είδος βιομηχανικής μονάδας. Χρησιμοποιήστε επίπεδο σημαντικότητας $\alpha=0.01$ και ελέγξτε αν υπάρχει ανεξαρτησία μεταξύ της εξειδίκευσης σπουδών και του τύπου της βιομηχανίας.

Σπουδές	Βιομηχανία			
	Πετρέλαιο	Χημική	Ηλεκτρική	Πληροφορικής
Διοίκηση Επιχειρήσεων	30	15	15	40
Μηχανολόγοι	30	30	20	20