ΓΕΩΠΟΝΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
AGRICULTURAL UNIVERSITY OF ATHENS

Απαραμετρικές δοκιμασίες

Κατσιλέρος Αναστάσιος

2017

ΑΠΑΡΑΜΕΤΡΙΚΕΣ ΔΟΚΙΜΑΣΙΕΣ

Οι μέθοδοι που χρησιμοποιούσαμε στις προηγούμενες στατιστικές δοκιμασίες βασίζονται στην ύπαρξη κανονικότητας. Υπάρχουν όμως περιπτώσεις που τα δεδομένα μας δεν ακολουθούν την κανονική ή κάποια άλλη γνωστή κατανομή. Σε αυτές τις περιπτώσεις καταφεύγουμε στις μεθόδους της απαραμετρικής στατιστικής όπου οι απαιτήσεις για τις κατανομές είναι λιγότερες σε σχέση με τις αντίστοιχες παραμετρικές μεθόδους.

Οι μη παραμετρικές δοκιμασίες είναι γνωστές και ως δοκιμασίες που είναι ελεύθερες κατανομών (distribution-free tests).

Πειραματικοί Σχεδιασμοί

Σχεδόν για κάθε παραμετρική δοκιμασία έχει αναπτυχθεί και η αντίστοιχη μη παραμετρική. Το πλεονέκτημα τους είναι ότι είναι άπλες δοκιμασίες και απαιτούν λιγότερους υπολογισμούς. Το μειονέκτημα των απαραμετρικών δοκιμασιών είναι μικρότερη αποτελεσματικότητα σε σχέση με τις παραμετρικές δοκιμασίες.

Οι απαραμετρικές δοκιμασίες χρησιμοποιούνται:

- ✓ Όταν η κατανομή του πληθυσμού που μελετούμε αποκλίνει από την κανονική.
- ✓ Ήταν τα δεδομένα δίνονται υπό μορφή βαθμών (ranks).
- ✓ Όταν τα δεδομένα δίνονται κατά κατηγορίες.

Δοκιμασία Wilcoxon – Mann – Whitney, δείγματα ανεξάρτητα.

Όταν μας ενδιαφέρει να αξιολογήσουμε τη στατιστική σημαντικότητα της διαφοράς δύο διαμέσων από ανεξάρτητα δείγματα, τότε χρησιμοποιήσουμε τη μέθοδο *U Mann – Whitney.* Η διαδικασία είναι η εξής:

α. Τοποθετούμε σε ανιούσα σειρά τις παρατηρήσεις και από τα δύο δείγματα .
β. Αντικαθιστούμε κάθε παρατήρηση με τη σειρά της στην ακολουθία των αριθμών (βαθμό R) αλλά με βάση και τα δύο δείγματα.
γ. Υπολογίζουμε το άθροισμα των βαθμών R για κάθε δείγμα.
δ. Υπολογίζουμε τα U_1 και U_2 με βάση τους τύπους:

$$U_1 = n_1 n_2 + \frac{n_2 (n_2 + 1)}{2} - R_2 \quad \text{και} \quad U_2 = n_1 n_2 + \frac{n_1 (n_1 + 1)}{2} - R_1$$

Ο έλεγχος της σημαντικότητας γίνεται συγκρίνοντας το μεγαλύτερο των U_1 και U_2 με την τιμή που παίρνουμε από τον Πίνακα Wilcoxon-Mann-Whitney για τις τιμές n_1 και n_2 και επίπεδο σημαντικότητας α.

Παράδειγμα: Αριθμός ημερών διατήρησης χωρίς αλλοίωση, δύο προϊόντων γάλακτος.

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**Σύνολο**
$\mathbf{X_1}$	5	3	8	6	5	7	4	6	
R_1	*3,5*	*1*	*8,5*	*5,5*	*3,5*	*7*	*2*	*5,5*	*36,5*
$\mathbf{X_2}$	9	8	10	12	13	11	14	12	
R_2	*10*	*8,5*	*11*	*13,5*	*15*	*12*	*16*	*13,5*	*99,5*

$$U_1 = n_1 n_2 + \frac{n_2 (n_2 + 1)}{2} - R_2 = 8 * 8 + \frac{8 (8 + 1)}{2} - 99{,}5 = 0{,}5$$

$$U_2 = n_1 n_2 + \frac{n_1 (n_1 + 1)}{2} - R_1 = 8 * 8 + \frac{8 (8 + 1)}{2} - 36{,}5 = 63{,}5$$

Επειδή το U_2 = 63,4 είναι μεγαλύτερο από την τιμή του U = 49 του Πίνακα για n_1 = 8 και n_2 = 8 και για α = 0,05, συμπεραίνουμε ότι η διαφορά μεταξύ των δύο διαμέσων είναι στατιστικώς σημαντική.

```
> attach(Mann_Whitney)
> wilcox.test(Y~X)

Wilcoxon rank sum test with continuity correction

data:  Y by X
W = 0.5, p-value = 0.001094
alternative hypothesis: true location shift is not equal to 0

Warning message: In wilcox.test.default(x = c(5L, 3L, 8L, 6L, 5L, 7L, 4L, 6L), y = c(9L, : cannot compute exact p-value with ties

> library(exactRankTests)
> wilcox.exact(Y~X)

        Exact Wilcoxon rank sum test

data:  Y by X
W = 0.5, p-value = 0.0003108
alternative hypothesis: true mu is not equal to 0
```

Δοκιμασία Wilcoxon – προσημασμένων βαθμών, τιμές κατά ζεύγη.
(Wilcoxon signed-rank test)

Όταν τα δεδομένα προέρχονται από ζεύγη τιμών τότε χρησιμοποιούμε την βαθμολογική προσημική δοκιμασία του Wilcoxon. Η δοκιμασία αυτή είναι η αντίστοιχη της δοκιμασίας του t για ζεύγη τιμών. Η διαδικασία είναι η εξής:

α. Υπολογίζουμε τις διαφορές ανάμεσα στις δύο παρατηρήσεις για κάθε δείγμα.
β. Αγνοούμε το θετικό ή αρνητικό σημείο των διαφορών και τις βαθμολογούμε ανάλογα με τις απόλυτες τιμές. Βάζουμε στους βαθμούς το πρόσημο της διαφοράς από την οποία προήλθαν.
γ. Αθροίζουμε τους βαθμούς των θετικών διαφορών R+ και ξεχωριστά αυτούς των αρνητικών διαφορών R-.

Η στατιστική της δοκιμασίας είναι το "μικρότερο" από τα δύο R και το οποίο συμβολίζουμε με Τ. Για την αξιολόγηση της σημαντικότητας του Τ χρησιμοποιούμε τον πίνακα της δοκιμασίας Wilcoxon.

Παράδειγμα: Βαθμολογία μαθητών πριν και μετά από ενισχυτικό μάθημα.

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Σύνολο**
$\mathbf{X_1}$	12	9	14	13	16	11	15	12	9	11	
$\boldsymbol{X_2}$	10	15	14	19	15	18	19	20	17	16	
$X_2 - X_1$	*-2*	*6*	*0*	*6*	*-1*	*7*	*4*	*8*	*8*	*5*	
Βαθμός	*2*	*5,5*		*5,5*	*1*	*7*	*3*	*8,5*	*8,5*	*4*	
R_+		*5,5*		*5,5*		*7*	*3*	*8,5*	*8,5*	*4*	*39*
R_-	*2*				*1*						*3*

Επειδή το $R_+ = 3$ είναι μικρότερο από την τιμή 5 του Πίνακα των κρίσιμων τιμών της δοκιμασίας Wilcoxon, για συγκρίσεις n = 9 και για α = 0,05, συμπεραίνουμε ότι η διαφορά μεταξύ των δύο διαμέσων είναι στατιστικώς σημαντική.

```
> wilcox.test(Y~X, paired=T)

Wilcoxon signed rank test with continuity correction

data:  Y by X
V = 42, p-value = 0.02414
alternative hypothesis: true location shift is not equal to 0

Warning messages:
1: In wilcox.test.default(x = c(10L, 15L, 14L, 19L, 15L, 18L, 19L,  :
cannot compute exact p-value with ties
2: In wilcox.test.default(x = c(10L, 15L, 14L, 19L, 15L, 18L, 19L,  :
cannot compute exact p-value with zeroes

> library(exactRankTests)
> wilcox.exact(Y~X, paired=T)

Exact Wilcoxon signed rank test

data:  Y by X
V = 42, p-value = 0.01953
alternative hypothesis: true mu is not equal to 0
```

Προσημική δοκιμασία (Sign test).

Η προσημική δοκιμασία είναι μια μέθοδος για να ελεγχθεί η διαφορά μεταξύ ζευγών παρατηρήσεων. Ελέγχει εάν η παρατήρηση ενός ζεύγους τείνει να είναι μεγαλύτερη ή μικρότερη από την άλλη παρατήρηση.

Αν θέλουμε να συγκρίνουμε n ζεύγη, ελέγχουμε ποια επέμβαση έδωσε το επιθυμητό αποτέλεσμα σε r ζεύγη και το μη επιθυμητό αποτέλεσμα σε n – r. Η στατιστική για την προσημική δοκιμασία, που τη συμβολίζουμε με R, είναι το μικρότερο από τα r και n – r.

Για την αξιολόγηση της σημαντικότητας του R χρησιμοποιούμε τον αντίστοιχο πίνακα της προσημικής δοκιμασίας. Για να κριθεί σημαντική μια διαφορά θα πρέπει η τιμή του R να είναι μικρότερη από ή ίση με την αντίστοιχη τιμή του πίνακα.

Για τιμές πάνω από 30 υπολογίζουμε την τιμή του Z με βάση τον τύπο:

$$Z = (n - 2R - 1) / \sqrt{n}$$

και μετά χρησιμοποιούμε τον πίνακα της κανονικής κατανομής ή τους πίνακες του t με ∞ βαθμούς ελευθερίας.

Παράδειγμα: Βαθμολογία δύο κρασιών από δώδεκα κριτές.

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
X_1	-	+	+	-	+	+	+	-	-	+	+	+
X_2	+	-	-	+	-	-	-	+	+	-	-	-

Το X_1 κρίθηκε ως καλύτερο από τους οκτώ στους δώδεκα κριτές ($r = 8$) και το X_2 από τέσσερις ($n – r = 4$). Από τον πίνακα της προσημικής δοκιμασίας για $\alpha = 0{,}05$ χρειάζεται δύο μόνο διαφορές ($R = 2$), επειδή $n – r = 4$ είναι μεγαλύτερο $R = 2$, δεχόμαστε τη μηδενική υπόθεση ότι οι δύο επεμβάσεις δεν διαφέρουν.

```
> binom.test(8, 12)

Exact binomial test

data:  8 and 12
number of successes = 8, number of trials = 12, p-value = 0.3877
alternative hypothesis: true probability of success is not equal to 0.5
95 percent confidence interval:
0.3488755 0.9007539
sample estimates:
probability of success
0.6666667
```

Γενική δοκιμασία Kolmogorov-Smirnov δύο δειγμάτων.

(Two sample Kolmogorov-Smirnov)

Η γενική δοκιμασία δύο δειγμάτων Kolmogorov-Smirnov χρησιμοποιείται για να ελεγχθεί αν δύο δείγματα προέρχονται από την ίδια κατανομή. Η διαδικασία είναι η εξής:

α. Τοποθετούμε τις τιμές σε κάθε δείγμα σε ανιούσα σειρά.
β. Δημιουργούμε κλάσεις
γ. Για κάθε δείγμα παίρνουμε τις αθροιστικές συχνότητες μέσα σε κάθε κλάση
δ. Για κάθε κλάση παίρνουμε την απόλυτη τιμή της διαφοράς των αθροιστικών συχνοτήτων.
ε. Η στατιστική της δοκιμασίας είναι η μεγαλύτερη από τις διαφορές των αθροιστικών συχνοτήτων.

Ο έλεγχος της σημαντικότητας γίνεται συγκρίνοντας τη μεγαλύτερη από τις διαφορές με την αντίστοιχη τιμή που παίρνουμε από τον Πίνακα Kolmogorov-Smirnov.

Παράδειγμα: Προσβολή από περονόσπορο δύο υβριδίων τομάτας.

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
X_1	15	23	34	26	27	19	14	31	22	29
X_2	45	39	30	29	46	55	34	61	41	*47*

	X_1	X_2	ΑΣ X_1	ΑΣ X_2	**|Διαφορά|**
5-9			0	0	0
10-14	14		1	0	1
15-19	15, 19		3	0	3
20-24	22, 23		5	0	5
25-29	26, 27, 29	29	8	1	6
30-34	31, 34	30, 34	10	3	7
35-39		39	10	4	6
40-44		41	10	5	5
45-49		45, 46, 47	10	8	2
50-54			10	8	2
55-59		55	10	9	1
60-64		61	10	10	0

Συγκρίνοντας τη μεγαλύτερη διάφορα (7) με την αντίστοιχη τιμή (6) που παίρνουμε από τον Πίνακα κρίσιμων τιμών Kolmogorov-Smirnov για n=10 και α=0,05, συμπεραίνουμε ότι τα δύο δείγματα δεν προέρχονται από την ίδια κατανομή.

```
> ks.test(X1, X2)

Two-sample Kolmogorov-Smirnov test

data: X1 and X2
D = 0.7, p-value = 0.01489
alternative hypothesis: two-sided

Warning message:
In ks.test(X, Y) : cannot compute exact p-value with ties
```

McNemar δοκιμασία.

Η McNemar δοκιμασία είναι μια μη παραμετρική δοκιμασία για ονομαστικά (binary) δεδομένα κατά ζεύγη. Αναφέρεται και ως χ^2-McNemar, επειδή η στατιστική δοκιμή ακολουθεί την χ^2 κατανομή. Αν υποθέσουμε ότι αξιολογούμε δυο πειραματικές επεμβάσεις ανά ζεύγη και ελέγχουμε την αποτυχία ή επιτυχία, υπάρχουν τέσσερα πιθανά αποτελέσματα για κάθε ζεύγος:

Μάρτυρας	**Επέμβαση**	
	Αποτυχία	Επιτυχία
Αποτυχία	a	b
Επιτυχία	c	d

Η δοκιμασία McNemar είναι:

$$Q = \frac{(b-c)^2}{b+c} \quad \text{ή} \quad Qc = \frac{(|b-c|-1)^2}{b+c} \quad \text{(με συνεχή διόρθωση)}$$

Για μεγάλο αριθμό παρατηρήσεων ακολουθεί την κατανομή χ^2 με 1 βαθμό ελευθερίας.

Παράδειγμα: Η επίδραση της λήψης δισκίων βιταμίνης C στην ενίσχυση του ανοσοποιητικού, σε 37 άτομα (αρνητικό – μη ασθενής και θετικό – ασθενής).

	Μετά		
Πριν	Αρνητικό	Θετικό	
Αρνητικό	10	5	15
Θετικό	7	15	22
	17	20	37

$$Qc = \frac{\left(\left|b-c\right|-1\right)^2}{b+c} = \frac{\left(\left|5-7\right|-1\right)^2}{5+7} = \frac{1^2}{12} = 0{,}0833$$

Συγκρίνοντας την τιμή 0,0833 με την αντίστοιχη τιμή που παίρνουμε από τον Πίνακα κρίσιμων τιμών χ^2 για 1 ΒΕ και α=0,05, συμπεραίνουμε ότι δεν υπάρχουν στατιστικά σημαντικές διαφορές.

```
> test= matrix(c(10, 5, 7, 15),2, 2)
> mcnemar.test(test)
```

```
McNemar's Chi-squared test with continuity correction

data:  test
McNemar's chi-squared = 0.083333, df = 1, p-value = 0.7728
```

Δοκιμασία Cochran Q.

Η δοκιμασία Q του Cochran αποτελεί μια επέκταση της δοκιμασίας McNemar και χρησιμοποιείται για την αξιολόγηση τριών ή περισσότερων επεμβάσεων με δυαδικές (binary) παρατηρήσεις (π.χ. αποτυχία – επιτυχία) κατά ομάδες.

Η στατιστική δοκιμασία είναι:

$$Q = (k-1)\frac{k\sum_{i=1}^{k} X_{i.}^{2} - (\sum_{i=1}^{k} X_{i.})^{2}}{k\sum_{j=1}^{b} X_{.j} - \sum_{j=1}^{b} X_{.j}^{2}}$$

Όπου: *k* είναι ο αριθμός των επεμβάσεων, $X_{i.}$ είναι το άθροισμα των επεμβάσεων, *b* ο αριθμός ομάδων, $X_{.j}$ είναι το άθροισμα των ομάδων και *N* ο συνολικός αριθμός των παρατηρήσεων.

Ο έλεγχος της σημαντικότητας γίνεται συγκρίνοντας το Q με την τιμή που παίρνουμε από τον Πίνακα χ^2 για ΒΕ = k – 1 και επίπεδο σημαντικότητας α.

Παράδειγμα: Ρωτήθηκαν 6 άτομα να εκφράσουν τη γνώμη τους (0 αρνητική και 1 θετική) για τέσσερα προϊόντα.

	Προϊόντα				
Ομάδες	**1**	**2**	**3**	**4**	Σύνολο
1	0	1	1	0	2
2	1	1	0	1	3
3	1	0	0	0	1
4	0	1	1	0	2
5	1	1	0	1	3
6	0	0	1	0	1
Σύνολο	3	4	3	2	12

$$Q = (k-1)\frac{k\sum_{i=1}^{k} X_{i.}^{2} - (\sum_{i=1}^{k} X_{i.})^{2}}{k\sum_{j=1}^{b} X_{.j} - \sum_{j=1}^{b} X_{.j}^{2}} = 3*\frac{4*(3^2+4^2+3^2+2^2)-12^2}{4*12-(2^2+3^2+1^2+2^2+3^2+1^2)} = 1{,}2$$

Συγκρίνοντας την τιμή 1,2 με την αντίστοιχη τιμή που παίρνουμε από τον Πίνακα κρίσιμων τιμών χ^2 για 3 ΒΕ και α=0,05, συμπεραίνουμε ότι δεν υπάρχουν στατιστικά σημαντικές διαφορές.

```
> attach(Cochran_Q)
> library(RVAideMemoire)
> cochran.qtest(Y~X|Block)

Cochran's Q test

data:  Y by X, block = Block
Q = 1.2, df = 3, p-value = 0.753
alternative hypothesis: true difference in probabilities is not equal to 0
sample estimates:
proba in group A  proba in group B  proba in group C  proba in group D
       0.5000000        0.6666667        0.5000000        0.3333333
```

Δοκιμασία των Kruskal και Wallis.

Η δοκιμασία Kruskal-Wallis-*H* ελεγχθεί εάν δύο ή περισσότερα ανεξάρτητα δείγματα προέρχονται από την ίδια κατανομή. Η διαδικασία είναι η εξής:

α. Αντικαθιστούμε τις παρατηρήσεις με το βαθμό που προκύπτει γι' αυτές βάζοντας όλα τα δείγματα μαζί
β. Αθροίζουμε τους βαθμούς για κάθε δείγμα R_1, R_2, ..., $R_κ$.
γ. Για κάθε δείγμα υπολογίζουμε R_i^2/n_i.
δ. Η στατιστική της δοκιμασίας ισούται με:

$$H = \frac{12}{n(n+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(n+1)$$

Ο έλεγχος της σημαντικότητας γίνεται συγκρίνοντας το Η με την τιμή που παίρνουμε από τον Πίνακα χ^2 για ΒΕ = k – 1 και επίπεδο σημαντικότητας α.

Παράδειγμα: Αριθμός ελαττωμάτων

	Επαναλήψεις							
Επεμβάσεις	**1**	**2**	**3**	**4**	**5**	n_i	R_i	R_i^2/n
$\mathbf{X_1}$	10	14	11	12	9			
R_1	*12*	*15*	*13*	*14*	*10,5*	*5*	*64,5*	832,05
$\mathbf{X_2}$	2	4	5	3	5			
R_2	*1*	*3*	*5*	*2*	*5*	*5*	*16*	51,2
$\mathbf{X_3}$	5	9	8	7	6			
R_3	*5*	*10,5*	*9*	*8*	*7*	*5*	*39,5*	312,05
						15		1195,3

$$H=\frac{12}{n(n+1)}\sum_{i=1}^{k}\frac{R_i^2}{n_i}-3(n+1)=\frac{12}{15(15+1)}1195{,}3-3(15+1)=11{,}87$$

Συγκρίνοντας την τιμή 11,84 με την αντίστοιχη τιμή που παίρνουμε από τον Πίνακα κρίσιμων τιμών χ^2 για k -1 = 2 ΒΕ και α=0,05, συμπεραίνουμε ότι υπάρχουν στατιστικά σημαντικές διαφορές.

Πειραματικοί Σχεδιασμοί

```
> attach(Kruskal)
> library(agricolae)
> kruskal(Y, X, console = T)
```

```
Study: Y ~ X
Kruskal-Wallis test's
Ties or no Ties

Critical Value: 11.87099
Degrees of freedom: 2
Pvalue Chisq  : 0.002643912

X,  means of the ranks

     Y   r
1  12.9  5
2   3.2  5
3   7.9  5

Post Hoc Analysis

t-Student: 2.178813
Alpha    : 0.05
Minimum Significant Difference: 2.584137

Treatments with the same letter are not significantly different.

     Y   groups
1  12.9    a
3   7.9     b
2   3.2      c
```

Δοκιμασία Mood's Median.

Η Δοκιμασία Mood's Median είναι μια μη-παραμετρική δοκιμασία που ελέγχει τη διάμεσο δύο ή περισσότερων δειγμάτων. Είναι εύρωστη στην παρουσία ακραίων τιμών σε σχέση με την δοκιμασία Kruskal-Wallis test αλλά λιγότερο αποτελεσματική σε μεγάλα δείγματα.

α. Υπολογίζεται η διάμεσος όλων των παρατηρήσεων.

β. Υπολογίζεται ο αριθμός των παρατηρήσεων ανά δείγμα με μικρότερο ή ίσο και μεγαλύτερο από τη γενική διάμεσο.

γ. Υπολογίζεται η αναμενόμενη τιμή για κάθε κελί.

δ. Η στατιστική της δοκιμασίας ισούται με:

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

Ο έλεγχος της σημαντικότητας γίνεται συγκρίνοντας με την τιμή που παίρνουμε από τον Πίνακα χ^2 για ΒΕ=(αριθμός στηλών – 1)*(αριθμός γραμμών – 1) και επίπεδο σημαντικότητας α.

	1	**2**	**3**	**4**	**5**
X_1	10	14	11	12	9
X_2	2	4	5	3	5
X_3	5	9	8	7	6

	X_1	X_2	X_3	Σύνολο
Παρατηρήσεις ≤ Διάμεσο (= 7)	0	5	3	8
Αναμενόμενη τιμή	(2,66)	(2,66)	(2,66)	
Παρατηρήσεις > Διάμεσο	5	0	2	7
Αναμενόμενη τιμή	(2,33)	(2,33)	(2,33)	
Σύνολο	5	5	5	15

$$\chi^2 = \frac{(0-2{,}66)^2}{2{,}66} + \frac{(5-2{,}3)^2}{2{,}3} + \frac{(5-2{,}66)^2}{2{,}66} + \frac{(0-2{,}33)^2}{2{,}33} + \frac{(3-2{,}66)^2}{2{,}66} + \frac{(2-2{,}33)^2}{2{,}33} = 10{,}17$$

Συγκρίνοντας την τιμή 10,17 με την αντίστοιχη τιμή 5,991 που παίρνουμε από τον Πίνακα κρίσιμων τιμών χ^2 για ΒΕ=(αριθμός στηλών – 1)*(αριθμός γραμμών – 1)=(3 – 1)*(2 – 1)= 2 και α=0,05, συμπεραίνουμε ότι υπάρχουν στατιστικά σημαντικές διαφορές.

```
> Median.test(Y, X, console = T)
```

The Median Test for Y ~ X

Chi Square = 10.17857   DF = 2   P.Value 0.00616242
Median = 7

	Median	r	Min	Max	Q25	Q75
1	11	5	9	14	10	12
2	4	5	2	5	3	5
3	7	5	5	9	6	8

Post Hoc Analysis

Groups according to probability of treatment differences and alpha level.

Treatments with the same letter are not significantly different.

	Y	groups
1	11	a
3	7	b
2	4	c

Δοκιμασία Van der Waerden.

Η δοκιμασία Van der Waerden ελεγχθεί εάν δύο ή περισσότερα ανεξάρτητα δείγματα προέρχονται από την ίδια κατανομή. Η δοκιμή Van der Waerden μετατρέπει τους βαθμούς σε ποσοστά της κανονικής κατανομής.

α. Υπολογίζεται ο κανονικός βαθμός.

$$A_{ij} = \Phi^{-1}\left(\frac{R(X_{ij})}{N+1}\right)$$

β. Υπολογίζεται ο μέσος κανονικός βαθμός για κάθε επέμβαση.

$$\overline{A}_i = \frac{1}{n_i}\sum_{i=1}^{k} A_{ij}$$

γ. Η διακύμανση των κανονικών βαθμών

$$s^2 = \frac{1}{N-1}\sum_{i=1}^{k}\sum_{j=1}^{n} A_{ij}^2$$

Η στατιστική δοκιμή είναι: $$T_1 = \frac{1}{s^2}\sum_{i=1}^{k} n_k \overline{A}_i^2$$

Ο έλεγχος της σημαντικότητας γίνεται συγκρίνοντας το T_1 με την τιμή που παίρνουμε από τον Πίνακα χ^2 για ΒΕ = k – 1 και επίπεδο σημαντικότητας α.

```
> waerden.test(Y, X, console = T)

Study: Y ~ X
Van der Waerden (Normal Scores) test's

Value : 11.30606
Pvalue: 0.003506867
Degrees of Freedom:  2

X,  means of the normal score

            Y         std  r
1  0.92997983  0.4373162  5
2 -0.91084343  0.4487457  5
3 -0.01772371  0.3349149  5

Post Hoc Analysis

Alpha: 0.05 ; DF Error: 12

Minimum Significant Difference: 0.5652532

Treatments with the same letter are not significantly different.

Means of the normal score

        score  groups
1  0.92997983   a
3 -0.01772371   b
2 -0.91084343   c
```

Δοκιμασία Friedman.

Η δοκιμασία Friedman είναι μια μη παραμετρική δοκιμασία που χρησιμοποιείται για την αξιολόγηση πολλών πειραματικών επεμβάσεων σε ομάδες, με διατακτικές ή συνεχής παρατηρήσεις. Η διαδικασία είναι η εξής:

α. Μέσα σε κάθε ομάδα αντικαθιστούμε την τιμή της παρατήρησης με το βαθμό της.
β. Αθροίζουμε τους βαθμούς για κάθε επέμβαση R_1, R_2, ..., R_κ.
γ. Υπολογίζουμε την δοκιμασία R_F με τον τύπο.

$$R_F = \frac{12}{bk(k+1)} \sum_{i}^{k} R_{i.}^2 - 3b(k+1)$$

Ο έλεγχος της σημαντικότητας γίνεται συγκρίνοντας το R_F με την τιμή που παίρνουμε από τον Πίνακα χ^2 για ΒΕ = k – 1 και επίπεδο σημαντικότητας α.

Παράδειγμα

	Επεμβάσεις			
Ομάδες	**A**	**B**	**C**	**D**
1	10	2	5	4
$R_{.1}$	*4*	*1*	*3*	*2*
2	12	5	8	5
$R_{.2}$	*4*	*1,5*	*3*	*1,5*
3	9	4	9	7
$R_{.3}$	*3,5*	*1*	*3,5*	*2*
4	13	3	8	4
$R_{.4}$	*4*	*1*	*3*	*2*
Σύνολο $R_{i.}$	*15,5*	*4,5*	*12,5*	*7,5*

$$R_F = \frac{12}{bk(k+1)} \sum_i^k R_{i.}^2 - 3b(k+1) = \frac{12}{4*4*(4+1)} 473 - 3*4*(4+1) = 11{,}52$$

Συγκρίνοντας την τιμή 11,52 με την αντίστοιχη τιμή που παίρνουμε από τον Πίνακα κρίσιμων τιμών χ^2 για ΒΕ=k-1= 2 και α=0,05, συμπεραίνουμε ότι υπάρχουν στατιστικά σημαντικές διαφορές μεταξύ των επεμβάσεων.

```
> attach(Friedman)
> friedman(Block, X, Y, console = T)

Study: Y ~ Block + X
X,  Sum of the ranks

      Y    r
A  15.5   4
B   4.5   4
C  12.5   4
D   7.5   4

Friedman's Test
===============
Adjusted for ties
Critical Value: 11.52632
P.Value Chisq: 0.009195161
F Value: 73
P.Value F: 1.227533e-06

Post Hoc Analysis
Alpha: 0.05 ; DF Error: 9
t-Student: 2.262157
LSD: 1.847044

Treatments with the same letter are not significantly different.
   Sum of ranks   groups
A       15.5          a
C       12.5           b
D        7.5            c
B        4.5             d
```

Δοκιμασία Durbin.

Η δοκιμασία Durbin χρησιμοποιείται αντί της δοκιμασίας Friedman για πειράματα με ισορροπημένες μη πλήρεις ομάδες.
Η στατιστική δοκιμασία είναι:

$$T_2 = \frac{T_1/(t-1)}{(bk-b-T_1)/(bk-b-t+1)}$$

Όπου:

$$R_i = \sum_{j=1}^{b} R(X_{ij}) \qquad\qquad T_1 = \frac{(t-1)\sum_{i=1}^{k}(R_i^2 - r\frac{1}{4}bk(k+1)^2)}{\sum_{i=1}^{k}\sum_{j=1}^{b} R(X_{ij})^2 - \frac{1}{4}bk(k+1)^2}$$

t είναι ο αριθμός των επεμβάσεων, *k* είναι ο αριθμός των επεμβάσεων ανά ομάδα, *b* ο αριθμός ομάδων και r οι επαναλήψεις των επεμβάσεων. Το T_2 το συγκρίνουμε με την κρίσιμη τιμή του πίνακα του F με ΒΕ k-1 και bk-b-t+1.

Παράδειγμα

Ομάδες	**Επεμβάσεις**			
	A	**B**	**C**	**D**
1	-	2	1	1
$R_{.1}$		*3*	*1,5*	*1,5*
2	5	-	2	1
$R_{.2}$	*3*		*2*	*1*
3	8	1	-	2
$R_{.3}$	*3*	*1*		*2*
4	4	3	3	-
$R_{.4}$	*3*	*1,5*	*1,5*	
$\boldsymbol{R_{i.}}$	*9*	*5,5*	*5*	*4,5*

```
> attach(Durbin)
> durbin.test(Block, X, Y, console = T)
Study: Y ~ Block + X
Durbin Test
===========
Value      : 5.357143
DF 1       : 3
P-value    : 0.1474372
Alpha      : 0.05
DF 2       : 5
t-Student  : 2.570582
Least Significant Difference between the sum of ranks:  4.28216

Parameters BIB
Lambda     : 2
Treatmeans : 4
Block size : 3
Blocks     : 4
Replication: 3

Treatments with the same letter are not significantly different.
  Sum of ranks  groups
A      9.0        a
B      5.5        ab
C      5.0        ab
D      4.5         b
```

Συντελεστής συσχέτισης Spearman ρ

Ο συντελεστής συσχέτισης Spearman ρ, είναι μη παραμετρικό μέτρο συσχέτισης με βάση τις τάξεις δεδομένων.

Για τον υπολογισμό του συντελεστή του Spearman ακολουθούμε την εξής διαδικασία:

α. Βαθμολογούμε τις παρατηρήσεις κάθε μεταβλητής

β. Παίρνουμε την διαφορά των βαθμών για τις κατά ζεύγη παρατηρήσεις.

γ. Υψώνουμε στο τετράγωνο τις διαφορές.

δ. Υπολογίζουμε τον συντελεστή συσχέτισης χρησιμοποιώντας την παρακάτω εξίσωση:

$$r_s = 1 - \frac{6\sum d_i^{\,2}}{(n^3 - n)}$$

Για τον έλεγχο της δοκιμασίας χρησιμοποιούμε τον πίνακα κρίσιμων τιμών Spearman.

Παράδειγμα

X	R_X	**Y**	R_Y	d_i	d_i^2
4	*1*	15	*6*	-5	25
6	*3*	12	*5*	-2	4
5	*2*	9	*3*	-1	1
7	*4*	11	*4*	0	0
8	*5*	8	*2*	3	9
9	*6*	6	*1*	5	25
					64

$$r_s = 1 - \frac{6\sum d_i^{\ 2}}{(n^3 - n)} = 1 - \frac{6 * 64}{(216 - 6)} = -0,82$$

```
> attach(Spearman)
> cor.test(X, Y, method = "spearman")

Spearman's rank correlation rho

data:  X and Y
S = 64, p-value = 0.05833
alternative hypothesis: true rho is not equal to 0
sample estimates:
  rho
-0.8285714
```

Συντελεστής συσχέτισης kendall τ.

Ο συντελεστής συσχέτισης Kendall τ, είναι ένα μη παραμετρικό μέτρο συσχέτισης με βάση τον αριθμό των συμφωνιών ή μη σε ζεύγη παρατηρήσεων.

Για τον υπολογισμό του συντελεστή kendall ακολουθούμε την εξής διαδικασία:

α. Βαθμολογούμε ξεχωριστά τις X και Y και αντικαθιστούμε τις αρχικές μεταβλητές με τους αντίστοιχους βαθμούς.

β. Γράφουμε τους βαθμούς της X σε ανιούσα σειρά σε ζεύγη με το βαθμό της άλλης μεταβλητής Y.

γ. Εξετάζουμε την πρώτη τιμή στη στήλη των βαθμών R_Y και μετρούμε όλους τους βαθμούς που τον ακολουθούν και είναι μεγαλύτεροι από αυτόν.

δ. Ο βαθμολογικός συντελεστής του Kendall, (χωρίς ισοβαθμίες) υπολογίζεται από την έκφραση:

$$\tau = \frac{4\sum C_i}{n(n-1)} - 1$$

Ο έλεγχος σημαντικότητας γίνεται: $t_s = \tau / \sqrt{2(2n+5)/9n(n-1)}$

και συγκρίνουμε το t_s με το *t* των πινάκων της κατανομής του t για ∞ ΒΕ.

Παράδειγμα:

X	R_X	**Y**	R_Y
5	*1*	4	*1*
10	*6*	8	*5*
7	*3*	5	*2*
6	*2*	7	*4*
8	*4*	6	*3*
9	*5*	9	*6*

$\mathbf{R_X}$	$\mathbf{R_Y}$	Βαθμοί που ακολουθούν το $\mathbf{R_Y}$	*Αριθμός* C_i
1	*1*	*4, 2, 3, 6, 5*	*5*
2	*4*	*6, 5*	*2*
3	*2*	*3, 6, 5*	*3*
4	*3*	*6, 5*	*2*
5	*6*		*0*
6	*5*		*0*
		Σύνολο	*12*

$$\tau = 4\sum^{n} C_i / n(n-1) - 1 = 4*12/6*(6-1) - 1 = 0{,}6$$

$$t_s = \tau / \sqrt{2(2n+5)/9n(n-1)} = 1{,}69$$

```
> attach(Kendall)
> cor.test(X, Y, method = "kendall")

Kendall's rank correlation tau

data:  X and Y
T = 12, p-value = 0.1361
alternative hypothesis: true tau is not equal to 0
sample estimates:
tau
0.6
```