Σύμφωνα με Βαγενάς (2019) (σελ. 349 – 355) και τα παραδείγματα με τις εφαρμογές SPSS που βρίσκονται στο τέλος του κεφαλαίου.

Μη παραμετρικές συγκρίσεις

Καθηγητής Αντώνης Κ. Τραυλός (B.A., M.A., Ph.D.)
ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΛΟΠΟΝΝΗΣΟΥ
Τμήμα Οργάνωσης και Διαχείρισης Αθλητισμού
Σχολή Επιστημών Ανθρώπινης Κίνησης και Ποιότητας Ζωής

1

1

Παραμετρική ή μη-παραμετρική προσέγγιση?

- Κρίνεται απαραίτητο να τονισθεί ότι αν οι μετρήσεις των μεταβλητών δεν έχουν γίνει στην **κλίμακα διαστήματος** ή **λόγου** (αν δηλαδή τα αρχικά δεδομένα δεν είναι ποσοτικά) δεν μπορεί να εφαρμοστεί **παραμετρική ανάλυση.**

2

2

Η ερευνητική πρακτική έχει δείξει ότι, …

- … όταν υπάρχει σοβαρή **παραβίαση (violation)** της **παραδοχής της κανονικότητας (assumption of normality)** ή και της **παραδοχής της ομοιογένειας των διασπορών (assumption of homogeneity of variance)** σε μια σύγκριση k ανεξάρτητων ή εξαρτημένων δειγμάτων ή όταν το δείγμα είναι **σχετικά μικρό (small sample size) και η μεταβλητότητά του είναι υψηλή (high variability)**, <u>τότε αντί παραμετρικής μπορούμε να επιλέξουμε μη παραμετρική σύγκριση (non parametric comparison)</u>.

3

3

Οι μη παραμετρικοί έλεγχοι (non parametric tests) σύγκρισης δειγμάτων μιας μεταβλητής στηρίζονται σε 2 στατιστικές παραδοχές:

a) της ανεξαρτησίας των τιμών μεταξύ υποκειμένων (subjects) και μεταξύ δειγμάτων (samples) και
b) της υποφαινόμενης συνέχειας των τιμών (underlined continuity).

- Στη βάση αυτών των 2 παραδοχών διατυπώνεται η μηδενική υπόθεση (H_0) ότι τα δείγματα προέρχονται από την ίδια κατανομή έναντι στην εναλλακτική υπόθεση (H_A) ότι τα δείγματα προέρχονται από διαφορετικές κατανομές..

4

4

Οι μη παραμετρικοί έλεγχοι σύγκρισης k δειγμάτων …

- … βασίζονται στη διαφορά k αθροισμάτων σειρών (ΣR) και για τον λόγο αυτό οι αντίστοιχες τεχνικές ανάλυσης έχουν ως βασική αφετηρία τη σειρά των αρχικών τιμών.
- Στο κεφάλαιο αυτό εξετάζονται οι 4 κύριοι **μη παραμετρικοί έλεγχοι σύγκρισης**:
  1. **Mann-Whitney (U)** για 2 ανεξάρτητα δείγματα,
  2. **Wilcoxon (W)** για 2 εξαρτημένα δείγματα,
  3. **Kruskal-Wallis** για k ανεξάρτητα δείγματα,
  4. **Friedman** για k εξαρτημένα δείγματα.

5

5

Έλεγχος Mann-Whitney (U): Σύγκριση 2 Ανεξάρτητων Δειγμάτων

- Ο έλεγχος Mann-Whitney (U) είναι κατάλληλος για τη σύγκριση 2 ανεξάρτητων δειγμάτων (X_1, X_2) μιας μεταβλητής X της οποίας τα δεδομένα είναι **σειράς (ordinal)**, δηλαδή διατακτικής κλίμακας ή είναι ποσοτικά αλλά με μη κανονική κατανομή, π.χ. λόγω ακραίων τιμών.
- Ο Mann-Whitney ελέγχει αν τα 2 ανεξάρτητα δείγματα (ομάδες) προέρχονται από 2 διαφορετικούς πληθυσμούς και αποτελεί το μέγιστο μη παραμετρικό ισοδύναμο του ελέγχου t για ανεξάρτητα δείγματα.
- Ο έλεγχος M-W προϋποθέτει **(α) τυχαίο δείγμα** και **(β) ανεξαρτησία των δεδομένων μεταξύ υποκειμένων και ομάδων**.

6

6

- Στην έρευνα πολλές φορές τίθεται το ερώτημα αν 2 ανεξάρτητα δείγματα προέρχονται από τον ίδιο πληθυσμό και επομένως έχουν την ίδια δομή και θέση κατανομής (distributional shape and location).
- Για τη διερεύνηση τέτοιων προβλημάτων υπάρχου 4 κύριοι έλεγχοι (SPSS): **Mann-Whitney (U)**, **Kolmogorov-Smirnov (Z)**, **Wald-Wolfowitz (runs)**, **Moses (extreme reactions)**.
- Ο **Mann-Whitney (U)** ελέγχει αν οι 2 πληθυσμοί από τους οποίους προέρχονται τα δύο δείγματα είναι ίδιοι στην κεντρική τους θέση (π.χ. mean ranks).
  - Αν αυτό ισχύει τότε οι σειρές των τιμών Χ στα δύο δείγματα θα πρέπει να είναι τυχαία κατανεμημένες και, επομένως, θα έχουν ίδια κατανομή.

7

7

- Ο **Kolmogorov-Smirnov (Z)** και ο **Wald-Wolfowitz (runs)** ελέγχουν τα 2 δείγματα αν διαφέρουν και ως προς τη μορφή των κατανομών και ως προς τη θέση.
- Ο **Kolmogorov-Smirnov** συγκρίνει τις αθροιστικές κατανομές των 2 δειγμάτων, ενώ ο **Wald-Wolfowitz** βασίζεται στον συνδυασμό σειρών και αρχικών παρατηρήσεων από τα 2 δείγματα.
- Ο **Moses (extreme reactions)** ελέγχει ακραίες αποκρίσεις σε πειραματική παρέμβαση (experimental treatment) που προκαλεί θετικές και αρνητικές επιδράσεις σε σύγκριση με αυτές μιας ομάδας ελέγχου (control group).

8

8

- Ο έλεγχος Mann-Whitney βασίζεται σε αρχικά δεδομένα (raw data) 2 ομάδων (N_1, N_2) τιμών της ίδιας μεταβλητής ιδιότητας και περιλαμβάνει τα εξής βήματα:
- **Βήμα 1ο** - Οι αρχικές τιμές X σε κοινή σειρά μεγέθους και για τις 2 ομάδες,
- **Βήμα 2ο** - Άθροισμα σειρών για κάθε ομάδα: ΣR_1 & ΣR_2,
- **Βήμα 3ο** - Στατιστικό U για κάθε ομάδα:

$$U_1 = N_1 N_2 + \frac{N_1(N_1 - 1)}{2} - \Sigma R_1$$

$$U_2 = N_1 N_2 + \frac{N_2(N_2 - 1)}{2} - \Sigma R_2$$

9

9

- **Βήμα 4ο** - Έλεγχος σημαντικότητας του U.

(i) Για N_1 & $N_2 < 20$: Άμεσος Έλεγχος (U)

σημαντικότητα της μικρότερης U (Πίν. M): με επίπεδο πιθανότητας α (π.χ. στο 0.05 ή στο 0.01) και με πλήθη των 2 ομάδων N_1 & N_2 η κρίσιμη τιμή U είναι U_c οπότε

- αν $U < U_c$ η διαφορά είναι σημαντική,
- αν $U > U_c$ η διαφορά είναι μη σημαντική.

435

Παράρτημα M - Κρίσιμες τιμές **U** για τον έλεγχο **Mann-Whitney**

Πιθανότητες : **p = 0.05** για μονόπλευρο και **p = 0.10** για δίπλευρο έλεγχο.

↓ N_1 \ N_2 →	9	10	11	12	13	14	15	16	17	18	20
1											0
2	1	1	1	2	2	2	3	3	3	4	4
3	3	4	5	5	6	7	7	8	9	9	11
4	6	7	8	9	10	11	12	14	15	16	18
5	9	11	12	13	15	16	18	19	20	22	25
6	12	14	16	17	19	21	23	25	26	28	32
7	15	17	19	21	24	26	28	30	33	35	39
8	18	20	23	26	28	31	33	36	39	41	47
9	21	24	27	30	33	36	39	42	45	48	54
10	24	27	31	34	37	41	44	48	51	55	62
11	27	31	34	38	42	46	50	54	57	61	69
12	30	34	38	42	47	51	55	60	64	68	77
13	33	37	42	47	51	56	61	65	70	75	84
14	36	41	46	51	56	61	66	71	77	82	92
15	39	44	50	55	61	66	72	77	83	88	100
16	42	48	54	60	65	71	77	83	89	95	107
17	45	51	57	64	70	77	83	89	96	102	115
18	48	55	61	68	75	82	88	95	102	109	123
19	51	58	65	72	80	87	94	101	109	116	130
20	54	62	69	77	84	92	100	107	115	123	138

Πιθανότητες : **p = 0.025** για μονόπλευρο και **p = 0.05** για δίπλευρο έλεγχο.

↓ N_1 \ N_2 →	9	10	11	12	13	14	15	16	17	18	20
1											0
2	0	0	1	1	1	1	1	1	2	2	2
3	2	3	3	4	4	5	5	6	6	7	8
4	4	5	6	7	8	9	10	11	11	12	13
5	7	8	9	11	12	13	14	15	17	18	20
6	10	11	13	14	16	17	19	21	22	24	27
7	12	14	16	18	20	22	24	26	28	30	34
8	15	17	19	22	24	26	29	31	34	36	41
9	17	20	23	26	28	31	34	37	39	42	48
10	20	23	26	29	33	36	39	42	45	48	55
11	23	26	30	33	37	40	44	47	51	55	62
12	26	29	33	37	41	45	49	53	57	61	69
13	28	33	37	41	45	50	54	59	63	67	76
14	31	36	40	45	50	55	59	64	67	74	83
15	34	39	44	49	54	59	64	70	75	80	90

10

- **Βήμα 4ο** - Έλεγχος σημαντικότητας του U … **(συνέχεια)**

ii) Για N_1 & $N_2 > 20$: Έμμεσος Έλεγχος (z)

σημαντικότητα όποιας από τις 2 τιμές U (έμμεσα) με το στατιστικό

$$Z = \frac{U - (N_1 N_2)/2}{\sqrt{\dfrac{N_1 N_2 (N_1 + N_2 + 1)}{12}}}$$

το οποίο συγκρίνεται το κρίσιμο z_c, δηλαδή με αυτό που αντιστοιχεί στην πιθανότητα α :

- αν $z > z_c$ η διαφορά είναι σημαντική,
- αν $z < z_c$ η διαφορά είναι μη σημαντική.

- **Βήμα 5ο** - Μέγεθος Επίδρασης (Effect Size), Διαφοράς:
- Η τιμή z μετασχηματίζεται σε Point-Biserial $r_p = Z/\sqrt{N}$, όπου $N = N_1+N_2$ (συνολικό δείγμα), και το αποτέλεσμα αξιολογείται ως συσχέτιση μεταξύ της ανεξάρτητης μεταβλητής “ομάδα” και της εξαρτημένης μεταβλητής X με την κλίμακα Cohen (μικρή = 0.10, μέτρια = 0.30, μεγάλη = 0.50).

Π.χ., 16.1 - Mann-Whitney (U): σύγκριση 2 ανεξάρτητων ομάδων

- Σε 2 ομάδες μπάσκετ (A με $\mathrm{N}_1 = 8$, B με $\mathrm{N}_2 = 10$) μετρήθηκε η κόπωση (X) κ προέκυψαν οι τιμές που δίνονται στον Πίν. 16.1.

Πίνακας 16.1 - Έλεγχος Mann-Whitney: σύγκριση 2 ανεξάρτητων δειγμάτων

Ομάδα	Παίχτης	Κόπωση X	Σειρά R	Άθροιμα ΣR	Τιμή U
A N1 = 8	1	1.1	1	52	64
	2	3.2	3		
	3	4.4	4		
	4	5.3	5		
	5	7.1	7		
	6	9.1	9		
	7	13.6	11		
	8	14.7	12		
B N2 = 10	1	2.1	2	119	**16**
	2	6.2	6		
	3	8.8	8		
	4	12.1	10		
	5	15.1	13		
	6	16.2	14		
	7	17.5	15		
	8	18.4	16		
	9	19.3	17		
	10	20.2	18		

Ομάδα	Παίχτης	Κόπωση X	Σειρά R	Άθροιμα ΣR	Τιμή U
A N1 = 8	1	1.1	1	52	64
	2	3.2	3		
	3	4.4	4		
	4	5.3	5		
	5	7.1	7		
	6	9.1	9		
	7	13.6	11		
	8	14.7	12		
B N2 = 10	1	2.1	2	119	**16**
	2	6.2	6		
	3	8.8	8		
	4	12.1	10		
	5	15.1	13		
	6	16.2	14		
	7	17.5	15		
	8	18.4	16		
	9	19.3	17		
	10	20.2	18		

- Η μεταβλητή X είναι μεν συνεχής (ποσοτική), αλλά έχει <u>έντονα ασύμμετρη κατανομή</u> και <u>υψηλή μεταβλητότητα</u>, ενώ έχουμε και σχετικώς μικρό δείγμα σε κάθε ομάδα.
- Έτσι, επιλέγεται ο μη παραμετρικός έλεγχος Mann-Whitney αντί του παραμετρικού ελέγχου t, προκειμένου να ελεγχθεί αν οι 2 κατανομές διαφέρουν.
- Έτσι, ζητείται να ελεγχθεί στο $\alpha = 0.05$ αν οι 2 ομάδες αθλητών διαφέρουν μεταξύ τους σημαντικά στις σειριακές τιμές (ranks) των δεικτών κόπωσης.

Ομάδα	Παίχτης	Κόπωση X	Σειρά R	Άθροιμα ΣR	Τιμή U
A N1 = 8	1	1.1	1		
	2	3.2	3		
	3	4.4	4		
	4	5.3	5		
	5	7.1	7	52	64
	6	9.1	9		
	7	13.6	11		
	8	14.7	12		
B N2 = 10	1	2.1	2		
	2	6.2	6		
	3	8.8	8		
	4	12.1	10	119	**16**
	5	15.1	13		
	6	16.2	14		
	7	17.5	15		
	8	18.4	16		
	9	19.3	17		
	10	20.2	18		

Υπολογιστικά βήματα

- **Βήμα 1ο** - Οι αρχικές τιμές X σε κοινή σειρά μεγέθους και για τις 2 ομάδες.
- **Βήμα 2ο** - Άθροισμα σειρών: $\Sigma R_1 = 52$ & $\Sigma R_2 = 119$,
- **Βήμα 3ο** - Στατιστικό U για κάθε ομάδα:

$$U_1 = N_1 N_2 + \frac{N_1(N_1 - 1)}{2} - \Sigma R_1$$

$$U_2 = N_1 N_2 + \frac{N_2(N_2 - 1)}{2} - \Sigma R_2$$

Ομάδα	Παίχτης	Κόπωση X	Σειρά R	Άθροιμα ΣR	Τιμή U
A N1 = 8	1	1.1	1		
	2	3.2	3		
	3	4.4	4		
	4	5.3	5		
	5	7.1	7	52	64
	6	9.1	9		
	7	13.6	11		
	8	14.7	12		
B N2 = 10	1	2.1	2		
	2	6.2	6		
	3	8.8	8		
	4	12.1	10	119	**16**
	5	15.1	13		
	6	16.2	14		
	7	17.5	15		
	8	18.4	16		
	9	19.3	17		
	10	20.2	18		

- **Βήμα 3ο** - Στατιστικό U για κάθε ομάδα:

$$U_1 = N_1 N_2 + \frac{N_1(N_1 - 1)}{2} - \Sigma R_1$$

$$U_1 = (8)(10) + \frac{8(8-1)}{2} - 52$$

$$U_1 = 80 + 36 - 52 = 64$$

$$U_2 = N_1 N_2 + \frac{N_2(N_2 - 1)}{2} - \Sigma R_2$$

$$U_2 = (8)(10) + \frac{10(10-1)}{2} - 119$$

$$U_2 = 80 + 55 - 119 = 16$$

→ μικρότερη τιμή U είναι η $U_2 = 16$

435

Παράρτημα Μ - Κρίσιμες τιμές **U** για τον έλεγχο **Mann-Whitney**

Πιθανότητες : **p = 0.05** για μονόπλευρο και **p = 0.10** για δίπλευρο έλεγχο.

↓ N_1 \ N_2 →	9	10	11	12	13	14	15	16	17	18	20
1											0
2	1	1	1	2	2	2	3	3	3	4	4
3	3	4	5	5	6	7	7	8	9	9	11
4	6	7	8	9	10	11	12	14	15	16	18
5	9	11	12	13	15	16	18	19	20	22	25
6	12	14	16	17	19	21	23	25	26	28	32
7	15	17	19	21	24	26	28	30	33	35	39
8	18	20	23	26	28	31	33	36	39	41	47
9	21	24	27	30	33	36	39	42	45	48	54
10	24	27	31	34	37	41	44	48	51	55	62
11	27	31	34	38	42	46	50	54	57	61	69
12	30	34	38	42	47	51	55	60	64	68	77
13	33	37	42	47	51	56	61	65	70	75	84
14	36	41	46	51	56	61	66	71	77	82	92
15	39	44	50	55	61	66	72	77	83	88	100
16	42	48	54	60	65	71	77	83	89	95	107
17	45	51	57	64	70	77	83	89	96	102	115
18	48	55	61	68	75	82	88	95	102	109	123
19	51	58	65	72	80	87	94	101	109	116	130
20	54	62	69	77	84	92	100	107	115	123	138

Πιθανότητες : **p = 0.025** για μονόπλευρο και **p = 0.05** για δίπλευρο έλεγχο.

↓ N_1 \ N_2 →	9	10	11	12	13	14	15	16	17	18	20
1											0
2	0	0	1	1	1	1	1	1	2	2	2
3	2	3	3	4	4	5	5	6	6	7	8
4	4	5	6	7	8	9	10	11	11	12	13
5	7	8	9	11	12	13	14	15	17	18	20
6	10	11	13	14	16	17	19	21	22	24	27
7	12	14	16	18	20	22	24	26	28	30	34
8	15	$\boxed{17}$	19	22	24	26	29	31	34	36	41
9	17	20	23	26	28	31	34	37	39	42	48
10	20	23	26	29	33	36	39	42	45	48	55
11	23	26	30	33	37	40	44	47	51	55	62
12	26	29	33	37	41	45	49	53	57	61	69
13	28	33	37	41	45	50	54	59	63	67	76
14	31	36	40	45	50	55	59	64	67	74	83
15	34	39	44	49	54	59	64	70	75	80	90

- **Βήμα 4ο** - Έλεγχος σημαντικότητας:
- Επειδή N_1 & $N_2 < 20$ -- > με τον άμεσο τρόπο (Πίν. Μ): με $\alpha = 0.05$ (δίπλευρο έλεγχο), $N_1 = 8$ και $N_2 = 10$ η κρίσιμη τιμή U είναι $U_c = 17$, οπότε,
- επειδή $\mathbf{U_2 = 16 < U_C = 17}$ συμπεραίνουμε ότι η διαφορά είναι σημαντική.

17

17

Εναλλακτικά μετασχηματίζουμε όποια από τις 2 τιμές U (π.χ. την $U_2 = 16$)

$$Z = \frac{16 - ((8)(10))/2}{\sqrt{\frac{(8)(10)(8+10+1)}{12}}} = \frac{-24}{11.2546} = -2.132$$

- Πηγαίνουμε στον Πίνακα με τις τιμές z και βρίσκουμε ότι η τιμή z που αντιστοιχεί στην πιθανότητα $\alpha = 0.05$ είναι 1.64, οπότε επειδή $z = 2.1325 >$ κρίσιμο $z = 1.64$ συμπεραίνουμε ότι η διαφορά είναι στατιστικώς σημαντική και μεγάλη, όπως φαίνεται από το μέγεθος επίδρασης.

Positive Z Score Table

The chart shows the values of positive z scores which is either to the right or above the mean valu

X	0.0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	$\boxed{0.9495}$	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817

18

- **Βήμα 5ο - Μέγεθος Επίδρασης (Effect Size)**:
- Point-biserial $r_p = z\sqrt{N} = 2.1325\sqrt{18} = 0.50$.
- Η τιμή $r_p = 0.50$ δείχνει μεγάλο μέγεθος επίδρασης, σύμφωνα με την κλίμακα Cohen (0.10 = μικρό, 0.30 = μέτριο, 0.50 = μεγάλο).
- Εναλλακτικά, ως τυποποιημένη διαφορά (standardized difference), η τιμή $r_p = 0.50$ είναι

$$d = 2{*}r/\sqrt{1 - r^2} = 2 * 0.50/\sqrt{1 - 0.502} = 1.1547 \approx 1.15$$

- που επίσης δείχνει μεγάλη διαφορά (> 0.80 που δίνει ο Cohen ως μεγάλη).

'Ελεγχος Wilcoxon (W): Σύγκριση 2 Εξαρτημένων Δειγμάτων

- Ο έλεγχος Wilcoxon (W) είναι κατάλληλος για τη σύγκριση 2 εξαρτημένων δειγμάτων (X_1, X_2) μιας μεταβλητής X της οποίας τα δεδομένα είναι **σειράς (ordinal, δηλαδή διατακτικής κλίμακας)** ή **ποσοτικά αλλά με μη κανονική κατανομή (Wilcoxon, 1947)**.
- Ο Wilcoxon ελέγχει αν τα 2 εξαρτημένα δείγματα προέρχονται από 2 διαφορετικούς πληθυσμούς και αποτελεί το μέγιστο μη παραμετρικό ισοδύναμο του ελέγχου t για εξαρτημένα δείγματα.
- Προϋποθέτει (α) τυχαίο δείγμα και (β) ανεξαρτησία δεδομένων μεταξύ υποκειμένων.

- Οι έλεγχοι αυτού του είδους διερευνούν αν 2 εξαρτημένα δείγματα προέρχονται από τον ίδιο πληθυσμό.
- Η καταλληλότητα τους εξαρτάται από το είδος των δεδομένων (SPSS). Για **δεδομένα συνεχή** κατάλληλοι είναι ο **Sign** και ο **Wilcoxon**.
- Ο **Sign** συγκρίνει τις θετικές (+) με τις αρνητικές (-) διαφορές μεταξύ των τιμών των 2 δειγμάτων (X_1, X_2) κατά ζεύγη.
- Ο **Wilcoxon** είναι αναβάθμιση του Sign, καθότι ελέγχει όχι μόνο την κατεύθυνση (+, -) αλλά και το μέγεθος της διαφοράς των τιμών των 2 δειγμάτων και είναι πιο ισχυρός, γι αυτό και χρησιμοποιείται σχεδόν αποκλειστικά για τέτοιου είδους δεδομένα έναντι το ελέγχου Sign.

21

21

- Για *δεδομένα 2-κατηγορικά* (binomial, dichotomous, π.χ. 0 και 1 ή 1 και 2) κατάλληλος είναι ο **McNemar**, που εφαρμόζεται σε πειραματικούς σχεδιασμούς πριν-μετά (before, after experiments) από κάποια παρέμβαση.
- Ο έλεγχος **Marginal Homogeneity** είναι γενίκευση του **McNemar**, είναι κατάλληλος για **δεδομένα πολυκατηγορικά** (multi-categorical, multinomial data) και βασίζεται στην κατανομή χ^2.
- Τα αρχικά δεδομένα περιλαμβάνουν 1 ομάδα υποκειμένων (π.χ. ατόμων) πλήθους Ν στα οποία η ίδια μεταβλητή Χ έχει μετρηθεί σε 2 διαφορετικές περιπτώσεις (X_1 & X_2) και κατά συνέπεια περιλαμβάνουν Ν ζεύγη τιμών.

22

22

Ο έλεγχος Wilcoxon περιλαμβάνει τα εξής βήματα:

- **Βήμα 1ο** - Αλγεβρικές μεταβολές (διαφορές d) μεταξύ 1ου (X_1) 2ου (X_2) δείγματος: $d = X_2 - X_1$.
- **Βήμα 2ο** - Ταξινόμηση των μεταβολών d σε σειρά μεγέθους (R) με βάση τις απόλυτες τιμές τους (αγνοούνται τα πρόσημα).
- **Βήμα 3ο** - Αθροίσματα σειρών με θετικά (R+) και αρνητικά (R-) πρόσημα ξεχωριστά (ΣR+, ΣR-) και εντοπισμός του μικρότερου αθροίσματος (W).

- **Βήμα 4ο** - Έλεγχος **στατιστικής σημαντικότητας** του μικρότερου αθροίσματος: W = ΣR+ ή W = ΣR-.

23

23

- **Βήμα 4ο** - Έλεγχος **στατιστικής σημαντικότητας** του **μικρότερου αθροίσματος**: W = ΣR+ ή W = ΣR-.
- (i) **Για N < 50 Άμεσος Έλεγχος (W)** σημαντικότητα του W (παράρτημα Ν): με επίπεδο πιθανότητας α & πλήθος δείγματος Ν
- το κρίσιμο μικρότερο άθροισμα W είναι $W_c = W(_{a,N})$, οπότε
  - αν $W < W_c$ η διαφορά είναι σημαντική,
  - αν $W > W_c$ η διαφορά είναι μη σημαντική.

437

Παράρτημα Ν - Κρίσιμες τιμές **W** για τον έλεγχο **Wilcoxon**

(έλεγχος διπλής κατεύθυνσης)

N	0.05	0.02	0.01	N	0.05	0.02	0.01
6	1			28	117	102	92
7	2	0		29	127	111	100
8	4	2	0	30	137	120	109
9	6	3	2	31	148	130	118
10	8	5	3	32	159	141	128
11	11	7	5	33	171	151	138
12	14	10	7	34	183	162	149
13	17	13	10	35	195	174	160
14	21	16	13	36	208	186	171
15	25	20	16	37	222	198	183
16	30	24	19	38	235	211	195
17	35	28	23	39	250	224	208
18	40	33	28	40	264	238	221
19	46	38	32	41	279	252	234
20	52	43	37	42	295	267	248
21	59	49	43	43	311	281	262
22	66	56	49	44	327	297	277
23	73	62	55	45	344	313	292
24	81	69	61	46	361	329	307
25	90	77	67	47	379	345	323
26	98	85	76	48	397	362	339
27	107	93	84	49	415	380	356
				50	434	398	373

Από το : F. Wilcoxon (1947). Probability Tables for individual comparisons in Ranking Method. Biometrics, 3 : 114-122. Μετά από γραπτή άδεια της International Biometrics Society.

24

- **Βήμα 4ο** - Έλεγχος **στατιστικής σημαντικότητας** του **μικρότερου αθροίσματος**: W = ΣR+ ή W = ΣR-.
- (ii) **Για N > 50 Έμμεσος Έλεγχος (z)**
- Το μικρότερο άθροισμα (W) μετασχηματίζεται σε z

$$Z = \frac{W_{\min} - N(N+1)/4}{\sqrt{\frac{N(N+1)(2N+1)}{24}}}$$

- Και συγκρίνεται το z με το κρίσιμο z_c (δηλαδή με αυτό που αντιστοιχεί στην πιθανότητα α):
- αν $z > z_c$ η διαφορά είναι σημαντική,
- αν $z < z_c$ η διαφορά είναι μη σημαντική.

Positive Z Score Table

The chart shows the values of positive z scores which is either to the right or above the mean valu

X	0.0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817

25

- **Βήμα 5ο** - **Μέγεθος Επίδρασης (Effect Size)**:
- Η τιμή z μετασχηματίζεται σε Point-Biserial $r_p = z/\sqrt{N}$, όπου Ν ο αριθμός των παρατηρήσεων, και το αποτέλεσμα αξιολογείται ως συσχέτιση με την κλίμακα Cohen (μικρή = 0.10, μέτρια = 0.243, μεγάλη = 0.371).

26

26

Π.χ., 16.2 _ Wilcoxon (W): Σύγκριση 2 Εξαρτημένων Δειγμάτων

- Μια ομάδα 10 αθλητών αξιολογήθηκε με ερωτηματολόγιο (questionnaire) σε μια ιδιότητα συμπεριφοράς πριν (X_1) και μετά (X_2) τη συμμετοχή της σε έναν κρίσιμο αγώνα και πέτυχε τους βαθμούς που δίνονται στον Πίν. 16.2.

Πίνακας 16.2 - Έλεγχος Wilcoxon: σύγκριση 2 εξαρτημένων δειγμάτων

	Αξιολόγηση		Μεταβολή		Σειρές	
N	X_1	X_2	$d=X_2-X_1$	R	R+	R-
1	17	16	-1	-2		2
2	18	20	+2	+5	5	
3	10	13	+3	+8	8	
4	16	19	+3	+8	8	
5	13	14	+1	+2	2	
6	15	13	-2	-5		5
7	11	16	+5	+10	10	
8	15	16	+1	+2	2	
9	12	15	+3	+8	8	
10	10	12	+2	+5	5	
				Σ =	48	7

Οι τιμές R βγήκαν με βάση τις μεταβολές |d|.

27

- Υποθέστε ότι η μεταβλητή X είναι μεν ποσοτική (quantitative), αλλά έχει έντονα ασύμμετρη κατανομή (μη κανονική), ενώ το δείγμα είναι σχετικά μικρό.
- Έτσι επιλέγεται ο μη παραμετρικός έλεγχος Wilcoxon αντί του ελέγχου t και ζητείται να ελεγχθεί στο α = 0.05 αν οι σειριακές (ranks) επιδόσεις των 10 αθλητών στις 2 αξιολογήσεις διαφέρουν σημαντικά.

	Αξιολόγηση		Μεταβολή		Σειρές	
N	X_1	X_2	$d=X_2-X_1$	R	R+	R-
1	17	16	-1	-2		2
2	18	20	+2	+5	5	
3	10	13	+3	+8	8	
4	16	19	+3	+8	8	
5	13	14	+1	+2	2	
6	15	13	-2	-5		5
7	11	16	+5	+10	10	
8	15	16	+1	+2	2	
9	12	15	+3	+8	8	
10	10	12	+2	+5	5	
				Σ =	48	7

Οι τιμές R βγήκαν με βάση τις μεταβολές |d|.

28

- **Βήμα 1ο** - Οι αλγεβρικές μεταβολές μεταξύ 1ης (X_1) και 2ης (X_2) αξιολόγησης δίνονται στη στήλη d = $X_2 – X_1$.
- **Βήμα 2ο** - Οι μεταβολές d ταξινομήθηκαν σε σειρά μεγέθους με βάση τις απόλυτες τιμές (στήλη R).
- **Βήμα 3ο** - Αθροίσματα σειρών: ΣR+ = 48 = W+ και ΣR- = 7 = W-
- **Βήμα 4ο** - Στατιστική σημαντικότητα του μικρότερου W = ΣR- = 7:

	Αξιολόγηση		Μεταβολή		Σειρές	
N	X_1	X_2	$d=X_2\text{-}X_1$	R	R+	R-
1	17	16	-1	-2		2
2	18	20	+2	+5	5	
3	10	13	+3	+8	8	
4	16	19	+3	+8	8	
5	13	14	+1	+2	2	
6	15	13	-2	-5		5
7	11	16	+5	+10	10	
8	15	16	+1	+2	2	
9	12	15	+3	+8	8	
10	10	12	+2	+5	5	
				Σ =	48	7

Οι τιμές R βγήκαν με βάση τις μεταβολές |d|.

- **Βήμα 4ο** - Στατιστική σημαντικότητα του μικρότερου W = ΣR- = 7:

επειδή N < 50 γίνεται **άμεσος έλεγχος** σημαντικότητας, με επίπεδο πιθανότητας α = 0.05 (δίπλευρος έλεγχος & πλήθος δείγματος N = 10 (Πίν. N),

η κρίσιμη τιμή W είναι $W_c = W_{(0.05,\ 10)} = 8$, οπότε, επειδή $W = 7 < W_c = 8$ συμπεραίνουμε ότι η διαφορά είναι στατιστικώς σημαντική.

437

Παράρτημα Ν - Κρίσιμες τιμές **W** για τον έλεγχο **Wilcoxon**

(έλεγχος διπλής κατεύθυνσης)

N	0.05	0.02	0.01	N	0.05	0.02	0.01
6	1			28	117	102	92
7	2	0		29	127	111	100
8	4	2	0	30	137	120	109
9	6	3	2	31	148	130	118
10	$\boxed{8}$	5	3	32	159	141	128
11	11	7	5	33	171	151	138
12	14	10	7	34	183	162	149
13	17	13	10	35	195	174	160
14	21	16	13	36	208	186	171
15	25	20	16	37	222	198	183
16	30	24	19	38	235	211	195
17	35	28	23	39	250	224	208
18	40	33	28	40	264	238	221
19	46	38	32	41	279	252	234
20	52	43	37	42	295	267	248
21	59	49	43	43	311	281	262
22	66	56	49	44	327	297	277
23	73	62	55	45	344	313	292
24	81	69	61	46	361	329	307
25	90	77	67	47	379	345	323
26	98	85	76	48	397	362	339
27	107	93	84	49	415	380	356
				50	434	398	373

Από το : F. Wilcoxon (1947). Probability Tables for individual comparisons in Ranking Method. Biometrics, 3 : 114-122. Μετά από γραπτή άδεια της International Biometrics Society.

- Εναλλακτικά μετασχηματίζουμε τη μικρότερη τιμή W (εδώ είναι η W- = 7) σε Z.

$$Z = \frac{7 - 10(10+1)/4}{\sqrt{\dfrac{10(10+1)[(2)(10)+1]}{24}}} = \frac{-20.5}{9.8107} = 2.09 \approx 2.1$$

- Από το Πίν. 21.Γ βρίσκουμε ότι η τιμή z που αντιστοιχεί στην πιθανότητα α = 0.05 είναι 1.64, οπότε, επειδή z = 2.1 > κρίσιμο z = 1.64, συνάγεται ότι η διαφορά είναι στατιστικώς σημαντική και μεγάλη όπως φαίνεται από το μέγεθος επίδρασης.

Positive Z Score Table

The chart shows the values of positive z scores which is either to the right or above the mean valu

x	0.0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817

31

- **Βήμα 5ο - Μέγεθος Επίδρασης (Effect Size)**:
- Point-Biserial $r_p = z/\sqrt{N} = 2.09/\sqrt{20} = 0.47$ και δείχνει μεγάλη επίδραση, σύμφωνα με την κλίμακα Cohen (0.10 = μικρή, 0.243 = μέτρια, 0.371 = μεγάλη).
- Εναλλακτικά, ως τυπική διαφορά d, η τιμή r_p =0.47, είναι
- $d = 2*r/\sqrt{1-r^2} = 2*0.47/\sqrt{1-0.47^2} = 1.057$ ≈1.06, που επίσης δείχνει μεγάλη διαφορά (> 0.80 που δίνει ο Cohen ως μεγάλη).

32

32

Τις αναλύσεις αναφορικά με περισσότερες των 2 συγκρίσεων για εξαρτημένα και ανεξάρτητα δείγματα δεν θα τις έχετε για τις εξετάσεις.

33

33

34