Μη παραμετρικές δοκιμασίες

Γιώτα Τουλούμη

Καθηγήτρια Βιοστατιστικής και Επιδημιολογίας

Εργ. Υγιεινής, Επιδημιολογίας και Ιατρικής Στατιστικής

Ιατρική Σχολή, ΕΚΠΑ

gtouloum@med.uoa.gr

Βάνα Σύψα

Καθηγήτρια Επιδημιολογίας και Ιατρικής Στατιστικής

Εργ. Υγιεινής, Επιδημιολογίας και Ιατρικής Στατιστικής

Ιατρική Σχολή, ΕΚΠΑ

vsipsa@med.uoa.gr

Είδη μεταβλητών

Συνεχείς ή ποσοτικές (quantitative)

Π.χ. ηλικία, επίπεδα χοληστερόλης, ύψος

Ποιοτικές (qualitative, categorical)

Διχοτομικές: φύλο, κάπνισμα

Ποιοτικές με >2 κατηγορίες: φυλή, ομάδα αίματος

Στατιστικές δοκιμασίες για τη διερεύνηση της σχέσης μεταξύ 2 παραγόντων

Παράγοντας 1	Παράγοντας 2	
	Ποσοτική	Ποιοτική
Ποιοτική	**t-test** **(αν η ποιοτική έχει 2 επίπεδα)**	**X^2-test**
Ποσοτική	**Correlation coefficient (με αξιολόγηση)** **Απλή γραμμική εξάρτηση**	**t-test** **(αν η ποιοτική έχει 2 επίπεδα)**

Σημείωση: Οι περισσότερες δοκιμασίες για **ποσοτικά** χαρακτηριστικά υποθέτουν την **κανονική** κατανομή των χαρακτηριστικών

Παράδειγμα

- Σε μία έρευνα συλλέγονται, μεταξύ άλλων, τα εξής δεδομένα για τους συμμετέχοντες:
  - Ηλικία (έτη)
  - Φύλο
  - Συστολική πίεση (mm Hg)

Πιθανά ερωτήματα

- **Διαφέρουν τα επίπεδα διαστολικής πίεσης μεταξύ ανδρών και γυναικών;**
  - Πίεση: ποσοτική μεταβλητή
  - Φύλο: ποιοτική μεταβλητή με 2 επίπεδα

  → t-test

- **Διαφέρουν οι καπνιστικές συνήθειες μεταξύ ανδρών και γυναικών;**
  - Κάπνισμα: ποιοτική μεταβλητή
  - Φύλο: ποιοτική μεταβλητή

  → X^2 test

- **Υπάρχει σχέση μεταξύ ηλικίας και διαστολικής πίεσης;**
  - Ηλικία: ποσοτική μεταβλητή
  - Συστολική πίεση: ποσοτική μεταβλητή

  →Συντελεστής συσχέτισης Pearson με την αντίστοιχη αξιολόγηση (ή απλή γραμμική εξάρτηση)

Παραμετρικές δοκιμασίες

- Είναι οι δοκιμασίες που εφαρμόζονται όταν γνωρίζουμε την κατανομή των μεγεθών που αναλύουμε
  - Π.χ. t-test (κανονική κατανομή)

- Τι χρησιμοποιούμε όταν:
  - οι προϋποθέσεις για την κανονικότητα των κατανομών δεν ισχύουν,

  ή
  - όταν δεν είμαστε σίγουροι αν ισχύουν (λόγω π.χ. ιδιαίτερα μικρού αριθμού παρατηρήσεων);

→ **μη-παραμετρικές στατιστικές δοκιμασίες** (έχουν γενικά λιγότερο αυστηρές προϋποθέσεις) **<u>Δοκιμασίες σειράς</u>**

Μη παραμετρικές δοκιμασίες

- Σε **ποιοτικά** δεδομένα:

  Το X^2 τεστ είναι μη παραμετρική δοκιμασία

- Σε **διατάξιμα** δεδομένα:

  Μπορεί να χρησιμοποιηθεί το X^2 τεστ αλλά οι μη παραμετρικές δοκιμασίες έχουν μεγαλύτερη **ισχύ** **(πιθανότητα να ανιχνεύσουμε μια διαφορά/σχέση όταν υπάρχει)**

- Σε **ποσοτικά** δεδομένα **που δεν κατανέμονται κανονικά** <u>**ούτε μετά από μετασχηματισμό:**</u>

  Μη παραμετρικές δοκιμασίες

Πότε προτιμάμε μη παραμετρικές δοκιμασίες;

1. Όταν οι προϋποθέσεις για την κανονικότητα των κατανομών δεν ισχύουν (μπορούμε πρώτα να δοκιμάσουμε κάποιο μετασχηματισμό των δεδομένων)

2. Όταν το μέγεθος του δείγματος είναι μικρό

3. Όταν αναλύουμε διατάξιμα δεδομένα

Είδος δεδομένων	Στατιστική επεξεργασία
Ποσοτικά που κατανέμονται **κανονικά** (έστω προσεγγιστικά **ή μετά από μετασχηματισμό**)	**Παραμετρικές δοκιμασίες** (t-test, παραμετρικός συντελεστής συσχέτισης με αξιολόγηση κλπ)
Ποσοτικά που **δεν κατανέμονται κανονικά** ή κατανέμονται με άγνωστο τρόπο	**Μη παραμετρικές δοκιμασίες σειράς**
Διατάξιμα	• **Μη παραμετρικές δοκιμασίες σειράς** (μεγαλύτερη ισχύς) • X^2
Ποιοτικά	X^2

ΜΗ ΠΑΡΑΜΕΤΡΙΚΕΣ

Μη παραμετρικές δοκιμασίες σε **ποσοτικά** και **διατάξιμα** δεδομένα: Δοκιμασίες σειράς (rank tests)

Στατιστικές δοκιμασίες που εφαρμόζονται σε ποσοτικά δεδομένα που δεν κατανέμονται κανονικά και σε διατάξιμα χαρακτηριστικά

Αντικαθιστούν τα δεδομένα με την ταξινόμησή τους (1, 2, 3, κλπ που περιγράφουν τη θέση που έχει κάθε μία τιμή όταν διατάξουμε τα δεδομένα)

1. **Δοκιμασία σημείων (sign test)** (για παρατηρήσεις **κατά ζεύγη**)
2. **Δοκιμασία Wilcoxon για παρατηρήσεις κατά ζεύγη** (Wilcoxon matched-pairs signed-rank test) (για παρατηρήσεις **κατά ζεύγη**)
3. **Δοκιμασία Wilcoxon για παρατηρήσεις χωρίς αντιστοιχία** (Wilcoxon test for two samples, Mann-Whitney U test, Wilcoxon rank-sum test)
4. **Μη παραμετρικός συντελεστής συσχέτισης σειράς του Spearman με την αντίστοιχη αξιολόγηση** (Spearman rank correlation coefficient)

Αντιστοιχία δοκιμασιών σειράς με αντίστοιχες παραμετρικές

Παραμετρική μέθοδος	Αντίστοιχη μη παραμετρική
t-test για ανεξάρτητα δείγματα	Δοκιμασία Wilcoxon για παρατηρήσεις χωρίς αντιστοιχία (Mann – Whitney U test)
t-test κατά ζεύγη	1. Δοκιμασία των σημείων (sign test) 2. Wilcoxon signed rank test
Pearson correlation coefficient με αξιολόγηση	Spearman correlation coefficient με αξιολόγηση

Μικρότερη ισχύ → μπορεί να αποτύχουν να αναδείξουν μία διαφορά ως στατιστικά σημαντική

Πλεονεκτήματα και μειονεκτήματα των μη παραμετρικών δοκιμασιών σειράς

Πλεονεκτήματα

- Ευρύτερη εφαρμογή: Ποσοτικά/Διατάξιμα
- Πιο απλοί υπολογισμοί
- Λιγότερο ευαίσθητες σε σφάλματα μέτρησης

Μειονεκτήματα

- Μικρότερη ισχύ από την αντίστοιχη παραμετρική μέθοδο αν οι προϋποθέσεις ισχύουν (δεν αξιοποιούν πλήρως τα δεδομένα)
- Κυρίως για στατιστική αξιολόγηση χωρίς να αποτυπώνουν το είδος της σχέσης (ενώ π.χ. οι μέσες τιμές & SE που υπολογίζουμε στο t-test μας βοηθούν να κατανοήσουμε το είδος της διαφοράς)
- Δεν μπορούν να εφαρμοστούν σε σύνθετες στατιστικές αναλύσεις (π.χ. μοντέλα) – δύσκολος υπολογισμός ορίων αξιοπιστίας

Δοκιμασία σημείων (Sign test)

- n (πλήθος) παρατηρήσεις **διατάξιμου ή ποσοτικού** (μη κανονικά κατανεμημένου ή άγνωστης κατανομής) μεγέθους σε δείγμα Α και n παρατηρήσεις του ίδιου διατάξιμου ή ποσοτικού μεγέθους σε δείγμα Β.

- Οι n παρατηρήσεις παρουσιάζουν **αντιστοιχία κατά ζεύγη**

- Κάθε ζεύγος παρατηρήσεων Α, Β κατατάσσεται ως ακολούθως:
  - Α>Β→ +
  - Β>Α→ -
  - Α=Β→ 0 (τα 0 εξαιρούνται)

Συμβολίζουμε $\boxed{\text{ε = πλήθος ‘+’}}$ $\boxed{\text{ζ = πλήθος ‘–’}}$

Δοκιμασία σημείων (Sign test)

Ερευνητική υπόθεση: Είναι η θέση της κατανομής στον πληθυσμό από τον οποίο προέρχεται το δείγμα Α ίδια με την αντίστοιχη του Β;

Αλλιώς διατυπωμένη: Είναι το ε ίσο με το ζ;

Μηδενική υπόθεση	H_0: ε=ζ
Εναλλακτική υπόθεση	H_1: ε≠ζ

Δοκιμασία σημείων (Sign test)

Στατιστικά κριτήρια/δοκιμασίες για τη σύγκριση των ε & ζ:

- **Μέθοδος 1:** X^2 κατά ζεύγη
- **Μέθοδος 2:** Σύγκριση του ποσοστού ε/(ε+ζ) με το 1/2

Ελέγχουμε αν ε = ζ

ή ισοδύναμα ελέγχουμε αν $\frac{\varepsilon}{\varepsilon+\zeta} = \frac{1}{2}$

[Γιατί; Επειδή αν ε=ζ → $\frac{\varepsilon}{\varepsilon+\zeta} = \frac{\varepsilon}{\varepsilon+\varepsilon} = \frac{1}{2}$]

Μηδενική υπόθεση H_0 : ε=ζ→ ε/(ε+ζ) = 1/2

Εναλλακτική υπόθεση H_1: ε/(ε+ζ) ≠ 1/2

Άρα σύγκριση της αναλογίας p= ε/(ε+ζ) με την p_0=1/2

Δοκιμασία σημείων (Sign test)

- **Μέθοδος 3:** Απευθείας από πίνακα για τη δοκιμασία σημείων

→ Οριακές τιμές ε_κ, ε_α

- Η διαφορά είναι στατιστικά σημαντική εάν το ε (αριθμός θετικών) είναι **εκτός** των ορίων αξιοπιστίας του ε+ζ → $\boldsymbol{\varepsilon \geq \varepsilon_\alpha}$ **ή** $\boldsymbol{\varepsilon \leq \varepsilon_\kappa}$

	Επίπεδο 5%	Επίπεδο 1%
Ζεύγη χωρίς ισοβαθμίες ε+ζ	ε_κ - ε_α	ε_κ - ε_α
Π.χ. 27	7-20	6-21

Επεξηγηματικά για τη μέθοδο 2

Για τη σύγκριση της αναλογίας p= ε/(ε+ζ) με την p_0=1/2:

A. Αν 0.1<p<0.9 & n>40:

- Υπολογίζουμε την ποσότητα $\frac{|p_1 - p_0|}{SE(p)}$ όπου $SE(p) = \sqrt{\frac{p_0 q_0}{(\varepsilon + \zeta)}}$
- Ελέγχουμε σε πίνακα t-test στους ∞ Β.Ε.

B. Αν n<40

- Υπολογίζουμε το διάστημα εμπιστοσύνης της αναλογίας ½ από ειδικούς πίνακες
  - Αν το p <u>εκτός</u> των ορίων του ΔΕ → στατιστικά σημαντική η διαφορά του από το ½
  - Αν το p <u>εντός</u> των ορίων του ΔΕ → μη στατιστικά σημαντική η διαφορά του από το ½

Δοκιμασία σημείων - Παράδειγμα

Σε 60 ασθενείς που έπασχαν από εκφυλιστική σπονδυλοαρθροπάθεια, χορηγήθηκαν με εναλλασσόμενη χρονική προτεραιότητα δύο διαφορετικά αναλγητικά (Α και Β).

- 32 ασθενείς διαπίστωσαν μεγαλύτερη βελτίωση με το Α
- 16 ασθενείς διαπίστωσαν μεγαλύτερη βελτίωση με το Β
- 12 ασθενείς δεν ανέφεραν προτίμηση.

Έχουν διαφορά τα 2 φάρμακα ως προς την αναλγητική τους δράση;

Δοκιμασία σημείων - Παράδειγμα

Επεξεργασία του ερωτήματος:

- Βαθμός ανακούφισης → διατάξιμο χαρακτηριστικό από συνεχή υποκείμενη κατανομή

- Συγκρίσεις κατά ζεύγη

→ Sign test ή Wilcoxon κατά ζεύγη (αυτό θα το πούμε αμέσως μετά)

Δοκιμασία σημείων - Παράδειγμα

32 ασθενείς διαπίστωσαν μεγαλύτερη βελτίωση με το Α

16 ασθενείς διαπίστωσαν μεγαλύτερη βελτίωση με το Β

12 ασθενείς δεν ανέφεραν προτίμηση

- Α>Β: 32 άτομα → ε=32
- Β>Α: 16 άτομα → ζ=16

- Α=Β: 12 άτομα (εξαιρούνται)

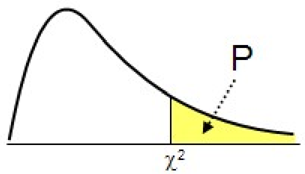
- Συνολικά ε+ζ=48 ζεύγη παρατηρήσεων

Παράδειγμα δοκιμασίας σημείων: Αξιολόγηση με μέθοδο 1 (X^2 κατά ζεύγη)

$$X^2 = \frac{(\varepsilon - \zeta)^2}{\varepsilon + \zeta} \quad \Rightarrow \quad X^2 = \frac{(|32 - 16| - 1)^2}{48} = 4.7$$

Βαθμοί ελευθερίας = 1

Chi-square distribution table



4,7>3,84

Στατιστικά σημαντικό στο 5% (p<0.05)

	P										
DF	**0.995**	**0.975**	**0.20**	**0.10**	**0.05**	**0.025**	**0.02**	**0.01**	**0.005**	**0.002**	**0.001**
1	0.0000393	0.000982	1.642	2.706	3.841	5.024	5.412	6.635	7.879	9.550	10.828
2	0.0100	0.0506	3.219	4.605	5.991	7.378	7.824	9.210	10.597	12.429	13.816
3	0.0717	0.216	4.642	6.251	7.815	9.348	9.837	11.345	12.838	14.796	16.266
4	0.207	0.484	5.989	7.779	9.488	11.143	11.668	13.277	14.860	16.924	18.467
5	0.412	0.831	7.289	9.236	11.070	12.833	13.388	15.086	16.750	18.907	20.515
6	0.676	1.237	8.558	10.645	12.592	14.449	15.033	16.812	18.548	20.791	22.458
7	0.989	1.690	9.803	12.017	14.067	16.013	16.622	18.475	20.278	22.601	24.322
8	1.344	2.180	11.030	13.362	15.507	17.535	18.168	20.090	21.955	24.352	26.124

Παράδειγμα δοκιμασίας σημείων: Αξιολόγηση με μέθοδο 2 (σύγκριση ποσοστού)

Μέθοδος 2: Σύγκριση του ποσοστού ε/(ε+ζ) με το p_0 =0,5

p=ε/(ε+ζ) = 0,67

<u>Προϋποθέσεις</u>

0,1 < p < 0,9 και n> 40

(Πράγματι p=0.67 και n=(ε+ζ) =48)

$$SE_p = \sqrt{\frac{p_0 q_0}{\varepsilon + \zeta}} = \sqrt{\frac{0,50 * 0,50}{48}} = 0,0722$$

$$\frac{|\mathrm{p}_1 - \mathrm{p}_0|}{\mathrm{SE_P}} = \frac{|0,67 - 0,50|}{0,0722} = \frac{0,17}{0.0722} = 2,36$$

Πως κρίνω αν η διαφορά είναι στατιστικά σημαντική;

Πίνακας t-test, στους ∞ Β.Ε.

ΒΕ	10%	5%	1%	0.1%
∞	1.65	1.96	2.58	3.29

2,36

$0.01 < p < 0.05$

Στατιστικά σημαντικό στο 5%

Παράδειγμα δοκιμασίας σημείων: Αξιολόγηση με μέθοδο 3 (απευθείας από πίνακα sign test)

Μέθοδος 3: Απευθείας από πίνακα

- Η διαφορά είναι στατιστικά σημαντική εάν το ε (αριθμός θετικών) είναι $ε \geq ε_α$ ή $ε \leq ε_κ$
- ε= 32

	Επίπεδο 5%	Επίπεδο 1%
Ζεύγη χωρίς ισοβαθμίες ε+ζ	$ε_κ$ - $ε_α$	$ε_κ$ - $ε_α$
48	**16-32**	**14-34**

$ε = 32 = ε_α$

Στατιστικά σημαντικό στο επίπεδο του 5%

Συμπέρασμα

Η διαφορά στην βελτίωση της εκφυλιστικής σπονδυλοαρθροπάθειας μεταξύ των αναλγητικών Α και Β είναι στατιστικά σημαντική στο επίπεδο του 5%

Ερμηνεία: Το αναλγητικό Α φαίνεται να προκαλεί μεγαλύτερη βελτίωση από το αναλγητικό Β σε ασθενείς με εκφυλιστική σπονδυλοαρθροπάθεια

Δοκιμασία σημείων: Παρατηρήσεις

- Διατάξιμα/ποσοτικά δεδομένα
- **Μικρότερη ισχύ σε σχέση με:**
  - **την μη παραμετρική Wilcoxon κατά ζεύγη (θα τη δούμε παρακάτω)**
  - **την παραμετρική t-test κατά ζεύγη (εάν μπορεί να εφαρμοστεί)**

Δοκιμασία Wilcoxon κατά ζεύγη

- n (πλήθος) παρατηρήσεις **διατάξιμου ή ποσοτικού** (μη κανονικά κατανεμημένου ή άγνωστης κατανομής) μεγέθους σε δείγμα Α και n παρατηρήσεις του ίδιου διατάξιμου ή ποσοτικού μεγέθους σε δείγμα Β

- Οι n παρατηρήσεις παρουσιάζουν **<u>αντιστοιχία κατά ζεύγη</u>**

- Είναι η θέση της κατανομής στον πληθυσμό από τον οποίο προέρχεται το δείγμα Α ίδια με την αντίστοιχη του Β;

- Το πρόβλημα είναι το ίδιο με αυτό που ελέγχουμε και στη δοκιμασία των σημείων, αλλά εδώ πρέπει οπωσδήποτε να έχουμε **ποσοτική μέτρηση**

Δοκιμασία Wilcoxon κατά ζεύγη: Διαδικασία

1. Για κάθε ζεύγος παρατηρήσεων Α, Β υπολογίζεται το Α-Β

2. Για Α-Β = 0 τα άτομα (ζεύγη) εξαιρούνται από την ανάλυση

3. Για καθένα από τα εναπομείναντα ζεύγη παρατηρήσεων Α, Β υπολογίζεται το |Α-Β|

4. Τα |Α-Β| διατάσσονται κατά αύξουσα σειρά και σημειώνεται η ‘θέση’ της καθεμιάς διαφοράς καθώς και το πρόσημο της διαφοράς

5. Αθροίζονται οι θέσεις με πρόσημο «+», έστω Τ+ και οι θέσεις με πρόσημο «-» έστω Τ-.

Δοκιμασία Wilcoxon κατά ζεύγη: Διαδικασία

Ερευνητική υπόθεση: Είναι η θέση της κατανομής στον πληθυσμό από τον οποίο προέρχεται το δείγμα Α ίδια με την αντίστοιχη του Β;

Αλλιώς διατυπωμένη: Είναι το Τ+ ίσο με το Τ-;

Μηδενική υπόθεση	**H_0: T+ = T-**
Εναλλακτική υπόθεση	**H_α T+ ≠ T-**

- Στατιστικός έλεγχος με πίνακα για $n \leq 25$
- Εάν n>25 η δοκιμασία γίνεται περίπλοκη. Σε αυτές τις περιπτώσεις μπορεί να εφαρμοστεί η δοκιμασία σημείων με μικρή απώλεια ισχύος.

Δοκιμασία Wilcoxon κατά ζεύγη: Αξιολόγηση με χρήση πίνακα

	Επίπεδο 5%	Επίπεδο 1%
Ζεύγη χωρίς ισοβαθμίες	Κατώτερο όριο T_l - Ανώτερο όριο T_r	Κατώτερο όριο T_l - Ανώτερο όριο T_r
Π.χ.19	**46-144**	**32-158**

Διαφορά <u>στατιστικά σημαντική</u> όταν η τιμή του T+ ή του T- είναι $\leq T_l$ ή $\geq T_r$ στο αντίστοιχο επίπεδο.
Όταν T+ $\leq T_l$ τότε ΥΠΟΧΡΕΩΤΙΚΑ το T- $\geq T_r$ και αντίστροφα

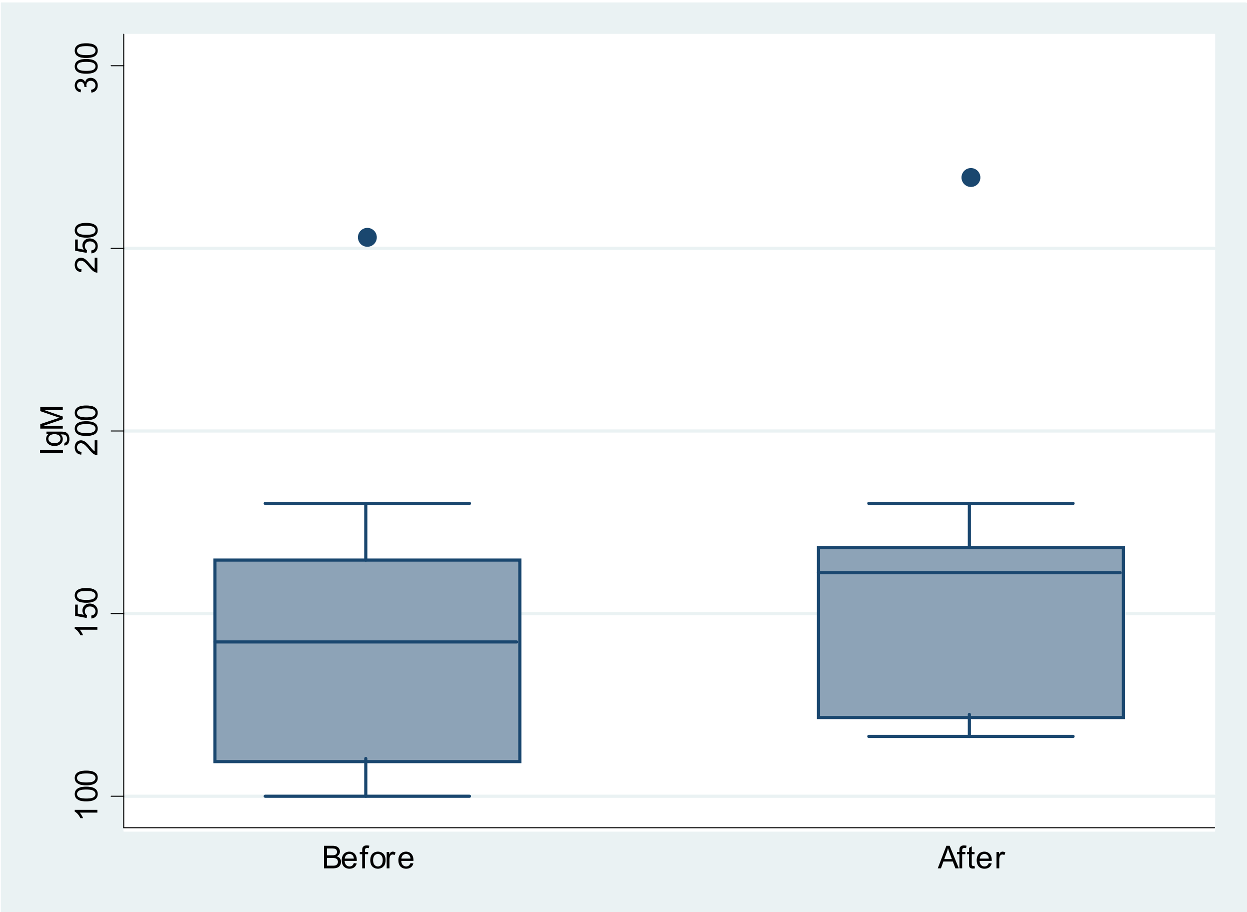
Δοκιμασία Wilcoxon κατά ζεύγη: Παράδειγμα

- Μελετήθηκαν οι ανοσοσφαιρίνες Μ του ορού 9 ενηλίκων ατόμων πριν και 15 ημέρες μετά τον αντιχολερικό εμβολιασμό.

- Να αξιολογηθεί αν ο αντιχολερικός εμβολιασμός προκαλεί μεταβολή στα επίπεδα των ανοσοσφαιρινών Μ.

Ανοσοσφαιρίνες Μ (g/100ml)	
Πριν	**Μετά**
158	168
140	168
165	161
100	120
110	122
142	134
104	116
254	270
180	180

Δοκιμασία Wilcoxon κατά ζεύγη: Παράδειγμα



Δοκιμασία Wilcoxon κατά ζεύγη: Παράδειγμα

Επεξεργασία του ερωτήματος:

- IgM → ποσοτικό χαρακτηριστικό του οποίου η κατανομή ενδέχεται να μην είναι κανονική και οι παρατηρήσεις είναι λίγες
- Οι παρατηρήσεις εμφανίζουν ατομική αντιστοιχία κατά ζεύγη

→ **Sign test ή Wilcoxon κατά ζεύγη**

Ανοσοσφαιρίνες Μ (g/100ml)	
Πριν	**Μετά**
158	168
140	168
165	161
100	120
110	122
142	134
104	116
254	270
180	180

Δοκιμασία Wilcoxon κατά ζεύγη: Παράδειγμα

Ανοσοσφαιρίνες Μ					
Πριν (A)	Μετά (B)	Διαφορά (A - B)	|A-B|	Θέση διαφοράς	Πρόσημο διαφοράς
158	168	−10	10	3	−
140	168	−28	28	8	−
165	161	4	4	1	+
100	120	−20	20	7	−
110	122	−12	12	4,5 ((4+5)/2)	−
142	134	8	8	2	+
104	116	−12	12	4,5 ((4+5)/2)	−
254	270	−16	16	6	−
180	180	0	−−	−−	−−

T+ = 1+2=3

T-=3+8+7+4,5+4,5+6=33

Αριθμός ζευγών χωρίς ισοψηφία = 9-1=8

Δοκιμασία Wilcoxon κατά ζεύγη: Παράδειγμα

Διαφορά <u>στατιστικά σημαντική</u> όταν η τιμή του T+ ή του T- είναι $\leq T_l$ ή $\geq T_r$ στο αντίστοιχο επίπεδο.

	Επίπεδο 5%	Επίπεδο 1%
Ζεύγη χωρίς ισοβαθμίες	**Κατώτ. όριο T_l – Ανώτ. όριο T_r**	**Κατώτ. όριο T_l – Ανώτ. όριο T_r**
8	**3-33**	**0-36**

T+ = 3, T-=33 **T+ = T_l, στο επίπεδο του 5%**

<u>Συμπέρασμα</u>: Οι τιμές στις ανοσοσφαιρίνες Μ του ορού πριν και 15 ημέρες μετά τον αντιχολερικό εμβολιασμό είναι στατιστικά σημαντικά διαφορετικές στο επίπεδο του 5%

<u>Ερμηνεία</u>: Οι τιμές στις ανοσοσφαιρίνες Μ του ορού αυξάνονται σε βαθμό στατιστικά σημαντικό 15 ημέρες μετά τον αντιχολερικό εμβολιασμό

Δοκιμασία Wilcoxon κατά ζεύγη: Παρατηρήσεις

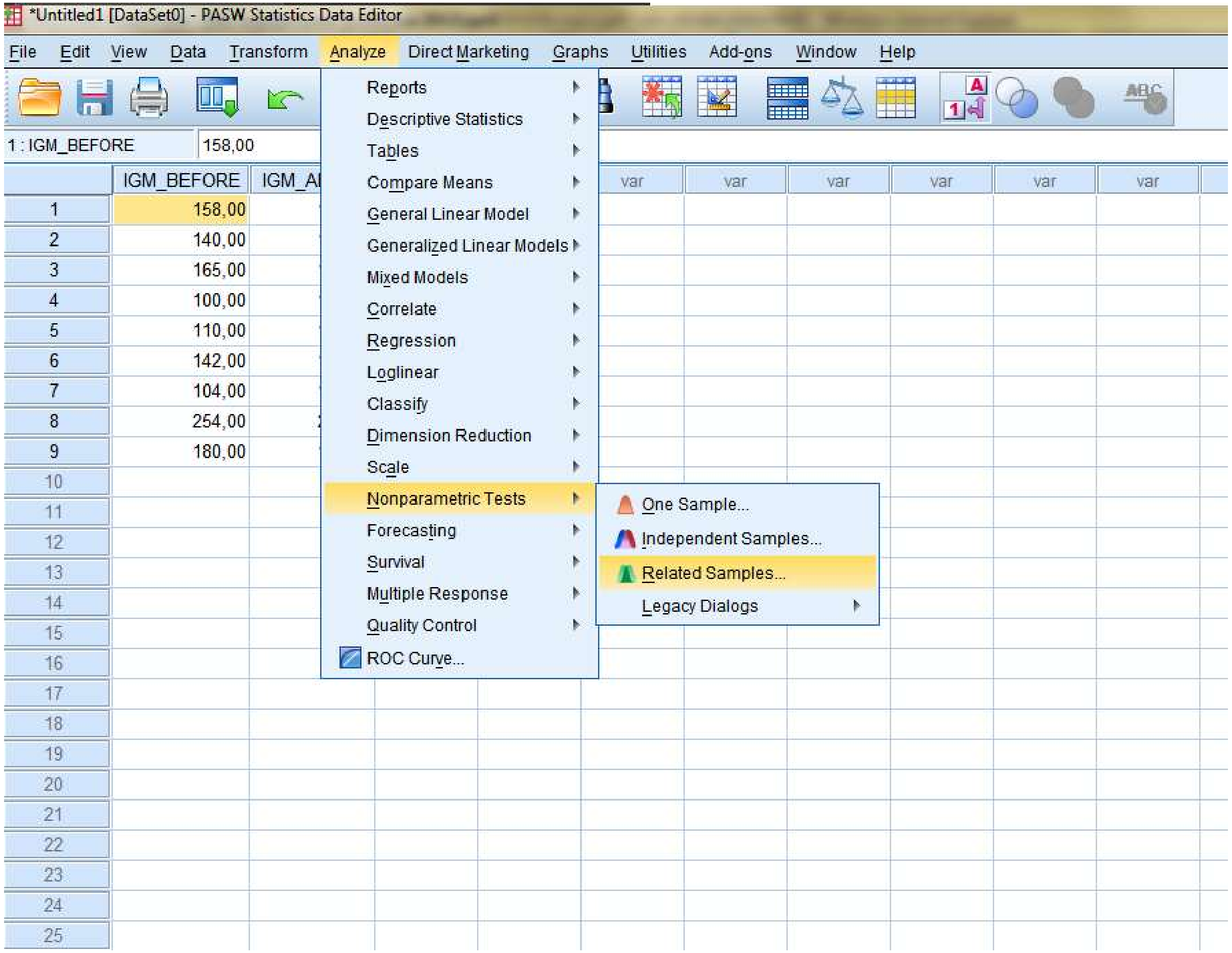
- Η ύπαρξη πολλών ζευγών με ισοβαθμίες ή πολλών απόλυτων διαφορών με την ίδια θέση στη διάταξη δεν επηρεάζουν σοβαρά την αξιοπιστία των ευρημάτων της δοκιμασίας

- Το άθροισμα των T+και T- είναι πάντοτε ίσο με n(n+1)/2

- Ο ελάχιστος αριθμός ζευγών n προκειμένου να αξιολογηθεί η δοκιμασία:
  - στο επίπεδο 5% πρέπει να είναι 6
  - στο επίπεδο 1% πρέπει να είναι 8

Δεν χρειάζεται να τα θυμάστε απέξω!

Υλοποίηση του sign test & Wilcoxon signed rank test στο SPSS στο παράδειγμα ανοσοσφαιρινών



Υλοποίηση του sign test & Wilcoxon signed rank test στο SPSS

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The median of differences between IGM_AFTER and IGM_BEFORE equals 0.	Related-Samples Sign Test	.289[1]	Retain the null hypothesis.
2	The median of differences between IGM_BEFORE and IGM_AFTER equals 0.	Related-Samples Wilcoxon Signed Ranks Test	.035	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

[1]Exact significance is displayed for this test.

Δοκιμασία Wilcoxon για παρατηρήσεις χωρίς αντιστοιχία (2 δειγμάτων)

Έστω:

- n_1 παρατηρήσεις διατάξιμου ή ποσοτικού μεγέθους στην ομάδα Α και n_2 παρατηρήσεις του ίδιου μεγέθους στην ομάδα Β, $n_1 \leq n_2$

- ΠΡΟΥΠΟΘΕΣΕΙΣ:
  - Ίδια περίπου κατανομή στις δύο ομάδες
  - Συνεχής υποκείμενη κατανομή

Δοκιμασία Wilcoxon για παρατηρήσεις <u>χωρίς αντιστοιχία-</u> Διαδικασία

1. Διάταξη όλων των (n_1+n_2) παρατηρήσεων κατά αύξουσα σειρά

2. Θέση της κάθε παρατήρησης

3. Άθροιση των θέσεων που καταλαμβάνουν οι παρατηρήσεις κάθε ομάδας ξεχωριστά (T_1, T_2)

4. Για **$n_1 \leq 25$, $n_2 \leq 50$** στατιστικός έλεγχος με ειδικό πίνακα

5. Συγκρίνεται η τιμή T_1 με τις τιμές T_l ή T_r. Εάν $T_1 \leq T_l$ ή $T_1 \geq T_r$ → στατιστικά σημαντικό αποτέλεσμα

	n_1		
n_2	4	5	….
	T_l - T_r	T_l - T_r	
4	**10-26**	**--**	…
5	**11-29**	**17-38**	…
…..	…	…	…

Στατιστικά σημαντικό αποτέλεσμα:

αν $T_1 \leq$ κάτω όριο διαστήματος ή

$T_1 \geq$ άνω όριο διαστήματος

Δοκιμασία Wilcoxon για παρατηρήσεις χωρίς αντιστοιχία για n1>25 ή n2>50 - Διαδικασία (<u>εκτός ύλης</u>)

- Για **n_1>25 ή n_2>50** υπολογισμός του ακόλουθου κριτηρίου

$$\frac{\left| T_1 - \frac{n1}{2}(n1 + n2 + 1) \right|}{\sqrt{\frac{n1 n2 (n1 + n2 + 1)}{12}}}$$

και αξιολόγηση με τις οριακές τιμές της t-κατανομής σε άπειρους βαθμούς ελευθερίας (θεωρούμε ότι το T_1 κατανέμεται κανονικά για μεγάλα n_1 και n_2)

ΒΕ	10%	5%	1%
500+	1,65	1,96	2,58
…..	…	…	…

Δοκιμασία Wilcoxon για παρατηρήσεις χωρίς αντιστοιχία – Παράδειγμα

Μία δοκιμασία με την οποία εκτιμάται η ικανότητα αντίληψης αριθμητικών δεδομένων υποβλήθηκε σε 11 απόφοιτους τεχνικών σπουδών και 8 απόφοιτους κλασσικών σπουδών. Τα αποτελέσματα, που δηλώνουν τον αριθμό των ασκήσεων που επιλύθηκαν σε σύνολο 36 που δόθηκαν για επίλυση ήταν:

Απόφοιτοι κλασσικών σπουδών: 17, 30, 25, 15, 11, 17, 22, 18

Απόφοιτοι τεχνικών σπουδών: 21, 21, 18, 32, 27, 20, 33, 30, 24, 29, 17

Να αξιολογηθεί στατιστικά η διαφορά επίδοσης μεταξύ των αποφοίτων κλασσικών και τεχνικών σπουδών

Δοκιμασία Wilcoxon για παρατηρήσεις χωρίς αντιστοιχία – Παράδειγμα

<u>Επεξεργασία του ερωτήματος</u>

- Η επίδοση στη δοκιμασία είναι ποσοτικό χαρακτηριστικό με άγνωστη – ενδεχομένως μη κανονική- κατανομή
- Δεν υπάρχει αντιστοιχία μεταξύ των παρατηρήσεων (οι δύο ομάδες είναι διαφορετικές)

→ Wilcoxon για παρατηρήσεις χωρίς αντιστοιχία

Ομάδα Α (n_1=8)	Ομάδα Β (n_2=11)
17	21
30	21
25	18
15	27
11	32
17	20
22	33
18	30
	24
	29
	17

Κοινή αύξουσα σειρά	Θέση ομάδας Α	Θέση ομάδας Β
11	1	
15	2	
17	(3+4+5)/3=4	
17	(3+4+5)/3=4	
17		(3+4+5)/3=4
18	(6+7)/2=6,5	
18		(6+7)/2=6,5
20		8
21		(9+10)/2=9,5
21		(9+10)/2=9,5
22	11	
24		12
25	13	
27		14
29		15
30	(16+17)/2=16,5	
30		16,5
32		18
33		19

T_1=(1+2+4+4+ 6,5+11+13+16,5)=58

T_2 = 132

Δοκιμασία Wilcoxon - Παράδειγμα

Αξιολόγηση:

	n_1=8		
	Επίπεδο 5%	Επίπεδο 1%	
n_2			
	T_l - T_r	T_l - T_r	
11	55-105	49-111	

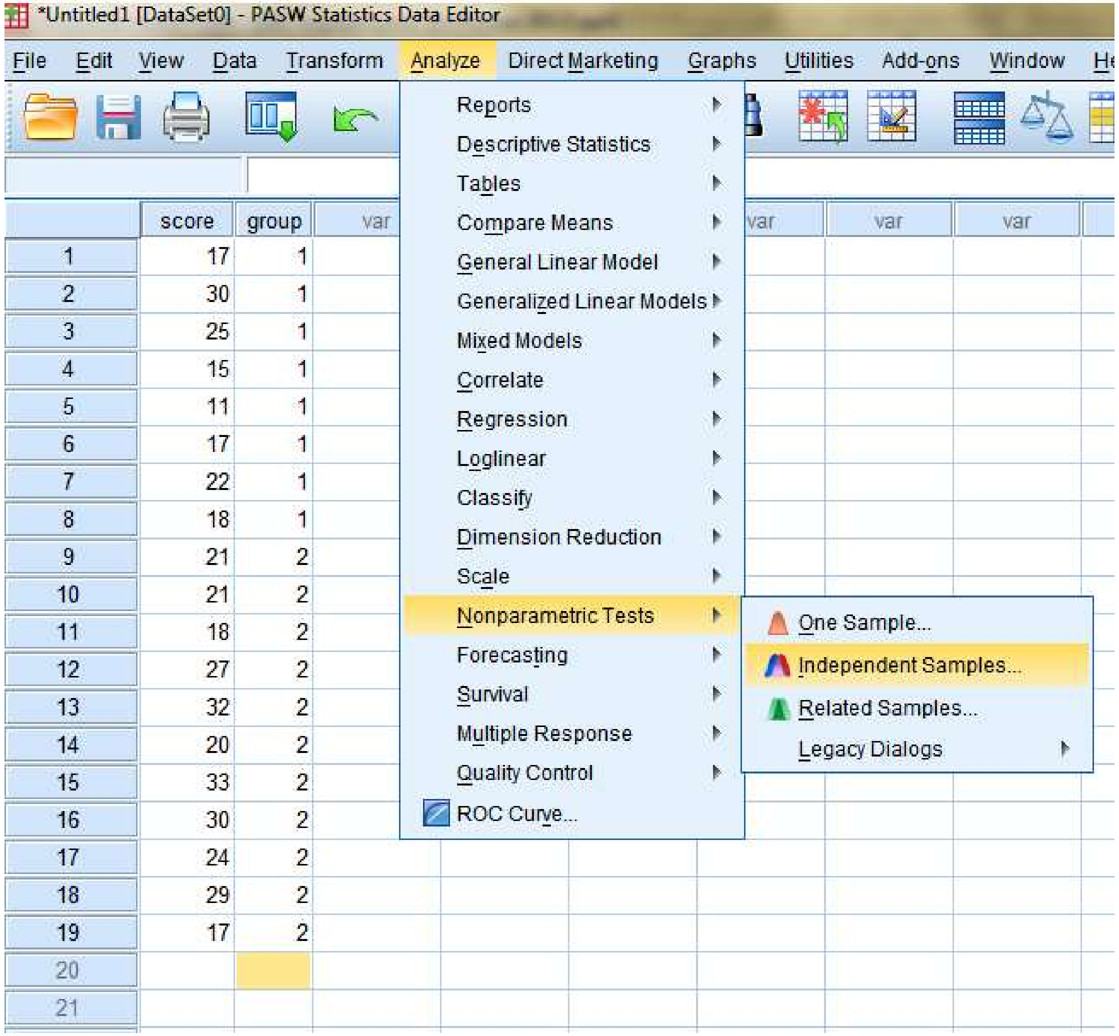
T1= 58

55 < 58 < 105

$p>0.05$

<u>Συμπέρασμα</u>: Η ικανότητα αντίληψης αριθμητικών δεδομένων δεν είναι στατιστικά σημαντικά διαφορετική μεταξύ αποφοίτων κλασσικών και τεχνικών σπουδών

Υλοποίηση του Wilcoxon rank sum test στο SPSS στο προηγούμενο παράδειγμα



Υλοποίηση του Wilcoxon rank sum test στο SPSS στο προηγούμενο παράδειγμα

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of score is the same across categories of group.	Independent-Samples Mann-Whitney U Test	.068	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

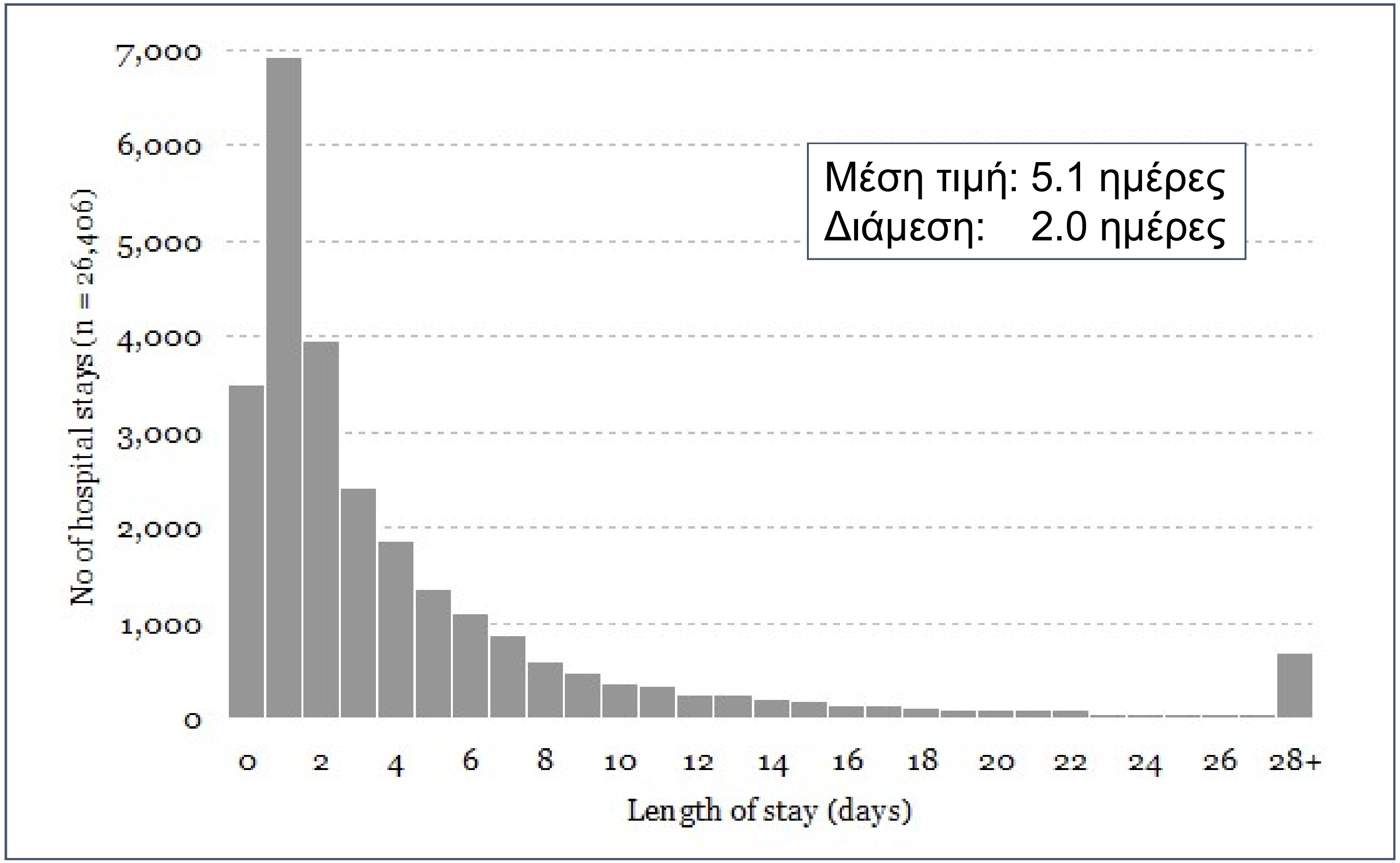
Table 1. Baseline characteristics and outcomes

	Septic patients (n = 32)	Control patients (n = 29)	*p*
Age (years), mean (SD)	58.0 (19.5)	58.5 (13.6)	.91
Gender (male), n (%)	19 (59.4)	15 (51.7)	.55
ICU type			
SICU, n (%)	19 (59.4)	12 (41.4)	.16
MICU, n (%)	13 (40.6)	17 (58.6)	
Reason for ICU admission, n (%)			.14
Cardiovascular disease	4 (12.5)	5 (17.2)	
Respiratory disease	6 (18.8)	9 (31.0)	
Gastrointestinal or renal disease	8 (25.0)	6 (20.7)	
Postoperative	4 (12.5)	7 (24.1)	
Trauma	10 (31.3)	2 (6.9)	
APACHE II, mean (SD)	15.7 (5.5)	15.0 (5.8)	.64
Mechanical ventilation, n (%)	24 (75.0)	12 (41.4)	.008
Vasopressors, n (%)	8 (25.0)	5 (17.2)	.46
In-hospital mortality, n (%)	9 (28.1)	2 (6.9)	.03
ICU LOS (days), median (IQR)	15.0 (9.5, 22.0)	4.0 (3.0, 4.0)	<.001

SD, standard deviation; ICU, intensive care unit; MICU, medical intensive care unit; SICU, surgical intensive care unit; APACHE, Acute Physiology and Chronic Health Evaluation; WBC, white blood cell; LOS, length of stay; IQR, 25%, 75% interquartile range

Drewery et al. 2013

Length of stay



Μη παραμετρικός συντελεστής συσχέτισης Spearman

Έστω:

- n παρατηρήσεις διατάξιμου ή ποσοτικού μεγέθους X και n παρατηρήσεις διατάξιμου ή ποσοτικού μεγέθους Y πάνω στις ίδιες ερευνητικές μονάδες (π.χ. στα ίδια άτομα, n πλήθος, μετράμε το X και το Y)

- ΠΡΟΥΠΟΘΕΣΕΙΣ:
  - Συνεχής υποκείμενη κατανομή για το X και για το Y

- Αποτελεί μέτρο της συσχέτισης των X και Y [όχι απαραίτητα γραμμικής συσχέτισης όπως ο συντελεστής συσχέτισης Pearson)

Συντελεστής Spearman: Διαδικασία

1. Για κάθε παρατήρηση του X διατάσσονται κατά ανιούσα σειρά οι n παρατηρήσεις και σημειώνεται η θέση
2. Για κάθε παρατήρηση του Y διατάσσονται κατά ανιούσα σειρά οι n παρατηρήσεις και σημειώνεται η θέση
3. Για κάθε ερευνητική μονάδα υπολογίζεται η διαφορά **d** ανάμεσα στις θέσεις ως προς X και Y
4. Υπολογίζεται ο Spearman συντελεστής συσχέτισης ως ακολούθως:

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d^2}{n(n^2-1)}$$

Τιμές του συντελεστή: [-1,..,1]
<0: αρνητική συσχέτιση
>0: θετική συσχέτιση
=0: απουσία σχέσης

Συντελεστής Spearman: Στατιστική αξιολόγηση

Το ερώτημα είναι αν υπάρχει συσχέτιση μεταξύ 2 μεταβλητών στον πληθυσμό

ρ → συντελεστής συσχέτισης των 2 μεταβλητών στον πληθυσμό

r → συντελεστής συσχέτισης των 2 μεταβλητών στο δείγμα

Αξιολογούμε το r βασιζόμενοι σε ένα δείγμα π.χ. ηλικία και συστολική πίεση σε δείγμα 10 ατόμων

→ r=0.70

Είναι το εύρημα αυτό στατιστικά σημαντικό;

H_0: ρ=0

H_1: ρ≠0

Συντελεστής Spearman: Στατιστική αξιολόγηση

Ελέγχουμε την τιμή |r| στον ειδικό πίνακα με τις οριακές τιμές (συνήθως με την τιμή που αντιστοιχεί στο 5% επίπεδο σημαντικότητας) σε n ζεύγη παρατηρήσεων

Αν $|r| \geq$ οριακή τιμή → απορρίπτω H_0 και συμπεραίνω ότι η σχέση είναι στατιστικά σημαντική

Αν $|r| <$ οριακή τιμή → η σχέση δεν είναι στατιστικά σημαντική

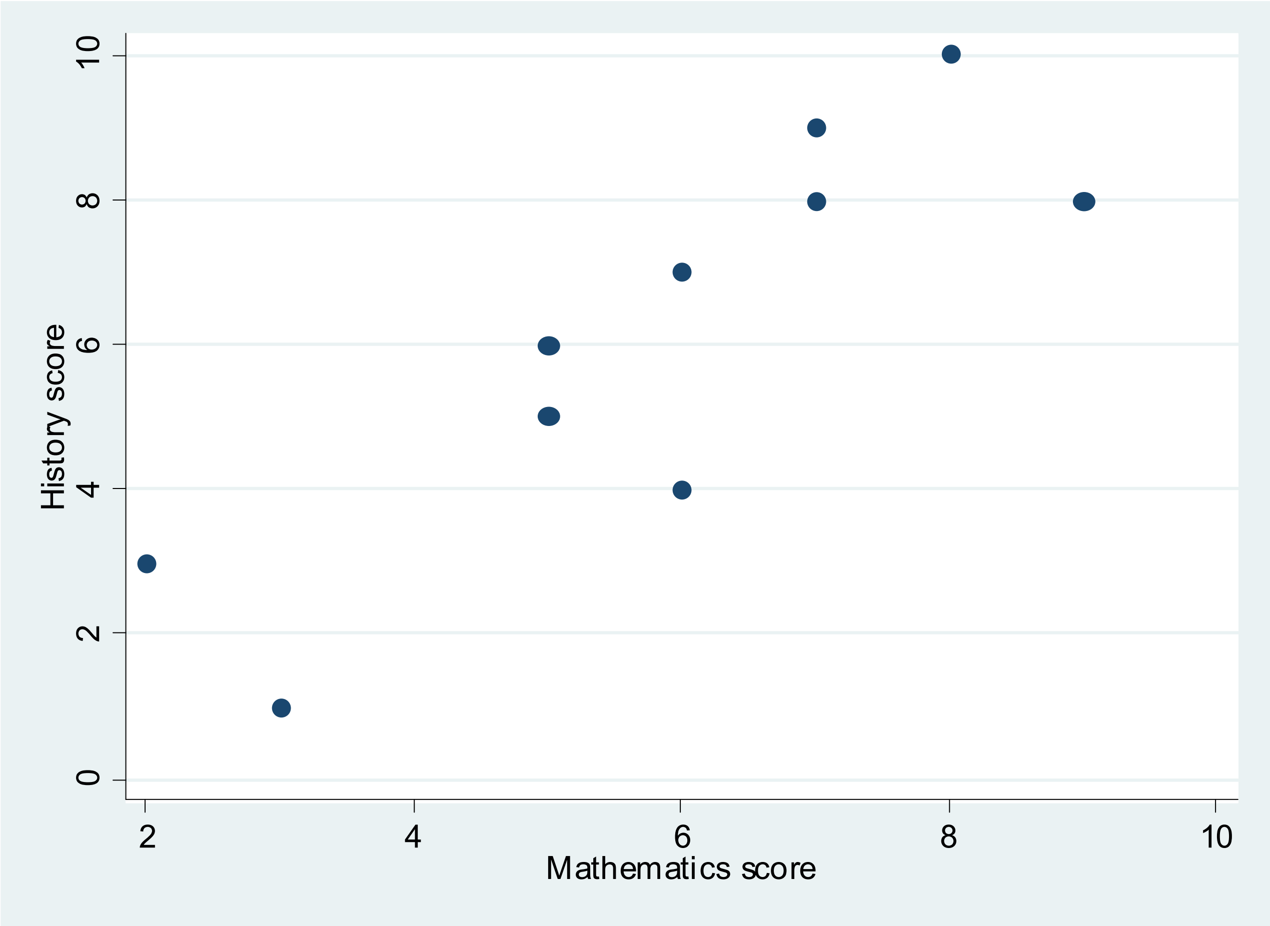
Spearman συντελεστής συσχέτισης: Παράδειγμα

- Έστω η βαθμολογία 10 σπουδαστών στα Μαθηματικά (X) και στην Ιστορία (Y)

- Υπάρχει σχέση της βαθμολογίας των σπουδαστών μεταξύ των 2 μαθημάτων;

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

Spearman συντελεστής συσχέτισης: Παράδειγμα



Spearman συντελεστής συσχέτισης: Παράδειγμα

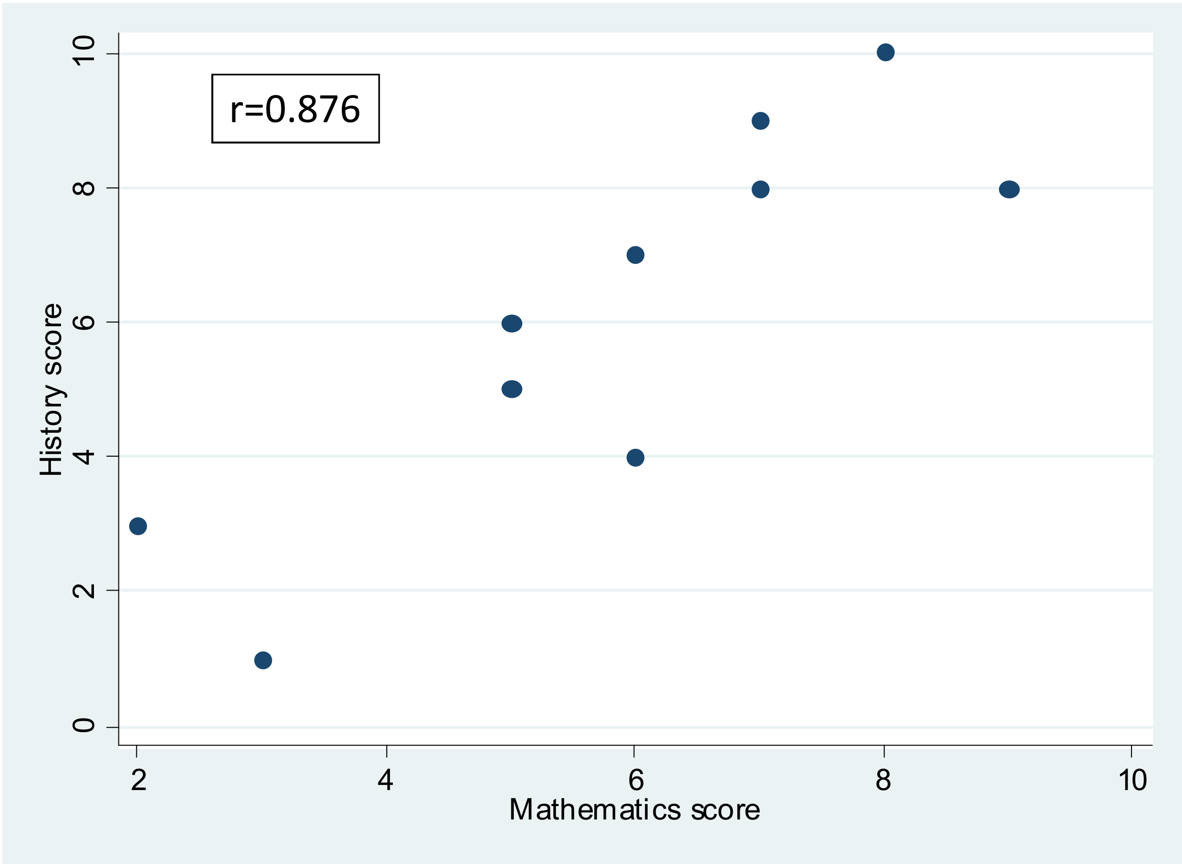
X	Θέση X	Y	Θέση Y	d	d²
8	9	10	10	-1	1
7	7,5	9	9	-1,5	2,25
9	10	8	7,5	2,5	6,25
3	2	1	1	1	1
6	5,5	7	6	-0,5	0.25
6	5,5	4	3	2,3	6,25
2	1	3	2	-1	1
5	3,5	5	4	-0,5	0,25
7	7,5	8	7,5	0	0
5	3,5	6	5	-1,5	2,25

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d^2}{n(n^2-1)}$$

$$= 1 - \frac{6 x 20,50}{10(10^2-1)} = 0,876$$

Σd^2 = 20,5

Spearman συντελεστής συσχέτισης: Παράδειγμα



Spearman συντελεστής συσχέτισης: Παράδειγμα

Αριθμός ζευγών

n=10

$$r_s = 0{,}876 \Rightarrow$$

$$r_s > 0{,}794$$

$$\Rightarrow p < 0.01$$

Συμπέρασμα: Υπάρχει στατιστικά σημαντική σχέση μεταξύ βαθμολογίας στα Μαθηματικά και την Ιστορία – φαίνεται πως οι σπουδαστές που είναι καλοί στο ένα τείνουν να είναι καλοί και στο άλλο μάθημα

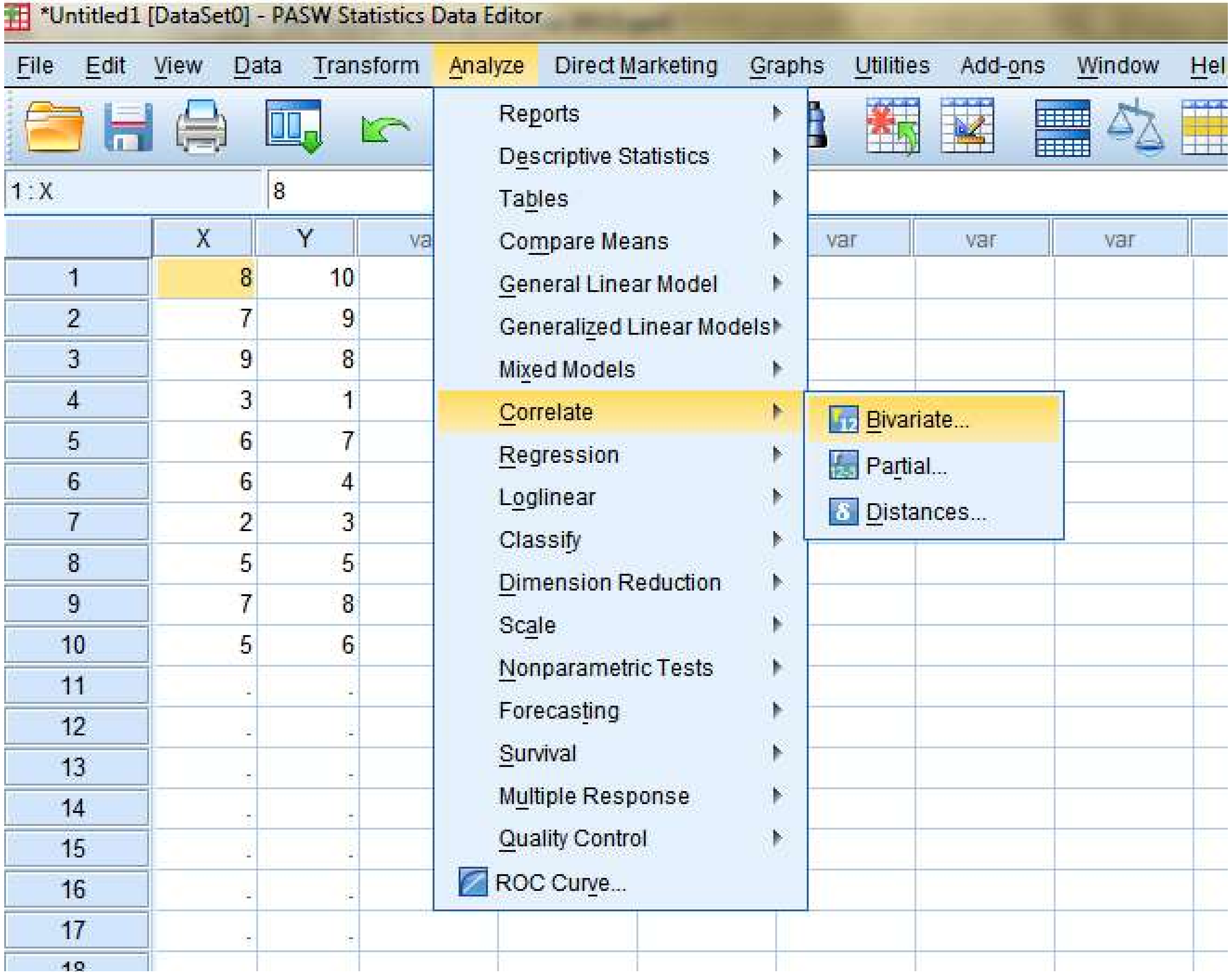
n	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.02$	$\alpha = 0.01$
5	.900	—	—	—
6	.829	.886	.943	—
7	.714	.786	.893	.929
8	.643	.738	.833	.881
9	.600	.700	.783	.833
$\boxed{10}$	$\boxed{.564}$	$\boxed{.648}$	$\boxed{.745}$	$\boxed{.794}$
11	.536	.618	.709	.755
12	.503	.587	.678	.727
13	.484	.560	.648	.703
14	.464	.538	.626	.679
15	.446	.521	.604	.654
16	.429	.503	.582	.635
17	.414	.485	.566	.615
18	.401	.472	.550	.600
19	.391	.460	.535	.584
20	.380	.447	.520	.570
21	.370	.435	.508	.556
22	.361	.425	.496	.544
23	.353	.415	.486	.532
24	.344	.406	.476	.521
25	.337	.398	.466	.511
26	.331	.390	.457	.501
27	.324	.382	.448	.491
28	.317	.375	.440	.483
29	.312	.368	.433	.475
30	.306	.362	.425	.467

Συντελεστής Spearman: Παρατηρήσεις

- Η ύπαρξη ισοβαθμιών στην ανοδική (ή καθοδική) διάταξη των παρατηρήσεων ως προς τις μεταβλητές X και Y ελάχιστα επηρεάζει την αξιολόγηση του Spearman συντελεστή συσχέτισης

- Το άθροισμα των διαφορών d είναι πάντα ίσο με το 0.

Υλοποίηση του Spearman correlation coefficient στο SPSS στο προηγούμενο παράδειγμα



Υλοποίηση του Spearman correlation coefficient στο SPSS στο προηγούμενο παράδειγμα

Correlations

			X	Y
Spearman's rho	X	Correlation Coefficient	1,000	,874**
		Sig. (2-tailed)	.	,001
		N	10	10
	Y	Correlation Coefficient	,874**	1,000
		Sig. (2-tailed)	,001	.
		N	10	10

**. Correlation is significant at the 0.01 level (2-tailed).

Άσκηση

Σε μια προοπτική έρευνα σε 23 έγκυες γυναίκες μελετήθηκαν τα επίπεδα της οιστριόλης (ορμόνη εγκυμοσύνης) κατά την 16η και 27η εβδομάδα κύησης.

(α) Διαφέρουν τα επίπεδα οιστριόλης μεταξύ 16ης και 27ης εβδομάδας κύησης;

(β) Διαφέρουν τα επίπεδα οιστριόλης κατά την 27η εβδομάδα κύησης στις γυναίκες με 1 και 2 εγκυμοσύνες;

Γυναίκα	Εγκυμοσύνη	Οιστριόλη (nmol/l)-$16^{η}$	Οιστριόλη (nmol/l)-$27^{η}$
1	$1^{η}$	24,9	>50
2	$1^{η}$	13,2	38,7
3	$1^{η}$	<10	26,7
4	$1^{η}$	<10	24,7
5	$1^{η}$	<10	32,2
6	$1^{η}$	23,7	>50
7	$1^{η}$	7,5	34,0
8	$1^{η}$	10,5	20,2
9	$1^{η}$	17,4	37,6
10	$1^{η}$	13,0	28,5
11	$1^{η}$	22,7	36,9
12	$2^{η}$	12,9	47,9
13	$2^{η}$	<10	31,9
14	$2^{η}$	10,9	46,3
15	$2^{η}$	14,0	26,6
16	$2^{η}$	11,5	26,0
17	$2^{η}$	10,4	40,0
18	$2^{η}$	11,5	>50
19	$2^{η}$	11,2	28,9
20	$2^{η}$	15,7	>50
21	$2^{η}$	<10	46,2
22	$2^{η}$	14,9	41,3
23	$2^{η}$	10,3	30,9

Ερωτήματα

1.Διαφέρουν τα επίπεδα οιστριόλης μεταξύ 16ης και 27ης εβδομάδας κύησης ;

Με αυτά που έχουμε μάθει μέχρι τώρα:
t-test κατά ζεύγη

2.Διαφέρουν τα επίπεδα οιστριόλης κατά την 27 εβδομάδα κύησης στις γυναίκες με 1 και 2 εγκυμοσύνες;

Με αυτά που έχουμε μάθει μέχρι τώρα:
t-test για ανεξάρτητα δείγματα

Ερώτημα

Διαφέρουν τα επίπεδα οιστριόλης μεταξύ 16ης και 27ης εβδομάδας κύησης ;

Πρόβλημα σε αυτά τα δεδομένα:

Γυναίκα	Οιστριόλη (nmol/l) 16η	Οιστριόλη (nmol/l) 27η
1	24,9	>50
2	13,2	38,7
3	<10	26,7
4	<10	24,7
5	<10	32,2
6	23,7	>50
7	7,5	34,0
8	10,5	20,2
9	17,4	37,6
10	13,0	28,5
11	22,7	36,9
12	12,9	47,9
13	<10	31,9
14	10,9	46,3
15	14,0	26,6
16	11,5	26,0
17	10,4	40,0
18	11,5	>50
19	11,2	28,9
20	15,7	>50
21	<10	46,2
22	14,9	41,3
23	10,3	30,9

Διαφέρουν τα επίπεδα οιστριόλης μεταξύ 16ης και 27ης εβδομάδας κύησης ;

Δεδομένα: Αντιστοιχία κατά ζεύγη- Μέγεθος ποσοτικό

↓

Δοκιμασία σημείων (Sign test)

Με <u>3 τρόπους</u> η αξιολόγηση

1. X^2
2. Κατασκευή ορίων αξιοπιστίας και έλεγχος αν το ½ βρίσκεται εντός
3. Χρήση έτοιμων πινάκων

Γυναίκα	Οιστριόλη (nmol/l) $16^{η}$ (A)	Οιστριόλη (nmol/l) $27^{η}$ (B)	Σημείο Διαφοράς (A)-(B)
1	24,9	>50	-
2	13,2	38,7	-
3	<10	26,7	-
4	<10	24,7	-
5	<10	32,2	-
6	23,7	>50	-
7	7,5	34,0	-
8	10,5	20,2	-
9	17,4	37,6	-
10	13,0	28,5	-
11	22,7	36,9	-
12	12,9	47,9	-
13	<10	31,9	-
14	10,9	46,3	-
15	14,0	26,6	-
16	11,5	26,0	-
17	10,4	40,0	-
18	11,5	>50	-
19	11,2	28,9	-
20	15,7	>50	-
21	<10	46,2	-
22	14,9	41,3	-
23	10,3	30,9	-

Παράδειγμα (συνέχεια)

- ε: το πλήθος των + → ε= 0
- ζ: το πλήθος των - → ζ=23

- Προϋποθέσεις
  - Η μορφή της κατανομής περίπου η ίδια στις 2 συγκρινόμενες ομάδες

- **Σύγκριση ε με ζ :**
  - 1ος τρόπος: Κατά ζεύγη X^2
  - 2ος τρόπος: Χρήση πινάκων sign test

Κατά ζεύγη X²

$$X^2 = \frac{(\varepsilon - \zeta)^2}{\varepsilon + \zeta}$$

Και με διόρθωση κατά Yates:

$$X^2_{B.E.=1} = \frac{(|\varepsilon - \zeta|-1)^2}{\varepsilon + \zeta} = \frac{(|0 - 23|-1)^2}{23} = \frac{484}{23} = 21.04$$

Αξιολόγηση του X^2

ΒΕ	10%	5%	1%	0,1%
1	2,71	3,84	6,64	10,83

21.04

21,04>10,83

p<0,1%

- **<u>Συμπέρασμα:</u>** Τα επίπεδα οιστριόλης μεταξύ 16ης και 27ης εβδομάδας κύησης διαφέρουν σε βαθμό πολύ στατιστικά σημαντικό.

- **<u>Ερμηνεία:</u>** Τα επίπεδα οιστριόλης την 27η εβδομάδα είναι υψηλότερα από τα αντίστοιχα κατά την 16η εβδομάδα κύησης και η διαφορά αυτή είναι στατιστικά σημαντική.

Εναλλακτικά: Χρήση πινάκων sign test

ε=0

ε+ζ	5%	1%
	ε_{κ}-ε_{α}	ε_{κ}-ε_{α}
23	6 – 17	4 - 19

0<4 → Στατιστικά σημαντική διαφορά στο 1%

Σημείωση: Για να χρησιμοποιήσουμε πίνακες πρέπει:

ε+ζ> 5 για 5% ή > 7 για 1%

Μειονεκτήματα του sign test

1. Μικρή ισχύς

2. Επηρεάζεται από τις ισοβαθμίες

Ερώτημα

Διαφέρουν τα επίπεδα οιστριόλης κατά την 27 εβδομάδα κύησης στις γυναίκες με 1 και 2 εγκυμοσύνες;

Διαφέρουν τα επίπεδα οιστριόλης κατά την 27η εβδομάδα κύησης στις γυναίκες με 1 και 2 εγκυμοσύνες;

Παρατηρήσεις χωρίς αντιστοιχία

↓

Wilcoxon 2 δειγμάτων

Προϋποθέσεις

- Ίδια περίπου κατανομή στις 2 συγκρινόμενες ομάδες.
- Συνεχής υποκείμενη κατανομή.

Γυναίκα	Εγκυμοσύνη	Οιστριόλη (nmol/l)-27η
1	1η	>50
2	1η	38,7
3	1η	26,7
4	1η	24,7
5	1η	32,2
6	1η	>50
7	1η	34,0
8	1η	20,2
9	1η	37,6
10	1η	28,5
11	1η	36,9
12	2η	47,9
13	2η	31,9
14	2η	46,3
15	2η	26,6
16	2η	26,0
17	2η	40,0
18	2η	>50
19	2η	28,9
20	2η	>50
21	2η	46,2
22	2η	41,3
23	2η	30,9

Αρχικά διατάσσω τις παρατηρήσεις των 2 ομάδων (1 ή 2 εγκυμοσύνες) σε ενιαία ανοδική (ή καθοδική) σειρά

	Ομάδα	Θέση
20,2	1^{η}	1
24,7	1^{η}	2
26,0	2^{η}	3
26,6	2^{η}	4
26,7	1^{η}	5
28,5	1^{η}	6
28,9	2^{η}	7
30,9	2^{η}	8
31,9	2^{η}	9
32,2	1^{η}	10
34,0	1^{η}	11
36,9	1^{η}	12
37,6	1^{η}	13
38,7	1^{η}	14
40,0	2^{η}	15
41,3	2^{η}	16
46,2	2^{η}	17
46,3	2^{η}	18
47,9	2^{η}	19
>50	1η/2^{η}	
>50	1η/2^{η}	21.5
>50	1η/2^{η}	
>50	1η/2^{η}	

Αθροίζω τις θέσεις ανά ομάδα:

T_1 = 1+2+5+6+10+11+12+13+14+21.5+21.5=117

T_2 = 3+4+7+8+9+15+16+17+18+19+21.5+21.5= 159

n_1=11

n_2=12

Επειδή **n1≤25, n2≤50** στατιστικός έλεγχος με Πίνακα

	5%	1%
(T_1 - Tr) →	(99 – 165)	(90 – 174)

Στατιστικά σημαντικό αποτέλεσμα:
αν T1 <κάτω όριο διαστήματος ή
T1 > άνω όριο διαστήματος

(1η ομάδα είναι η ομάδα με το μικρότερο αριθμό παρατηρήσεων)

	5%	1%
$(T_1 - Tr) \rightarrow$	(99 – 165)	(90 – 174)

T_1=117

99<117<165

P> 5%

Συμπέρασμα: Τα επίπεδα οιστριόλης κατά την 27^{η} εβδομάδα κύησης δε διαφέρουν ανάμεσα σε γυναίκες με 1 ή 2 εγκυμοσύνες.