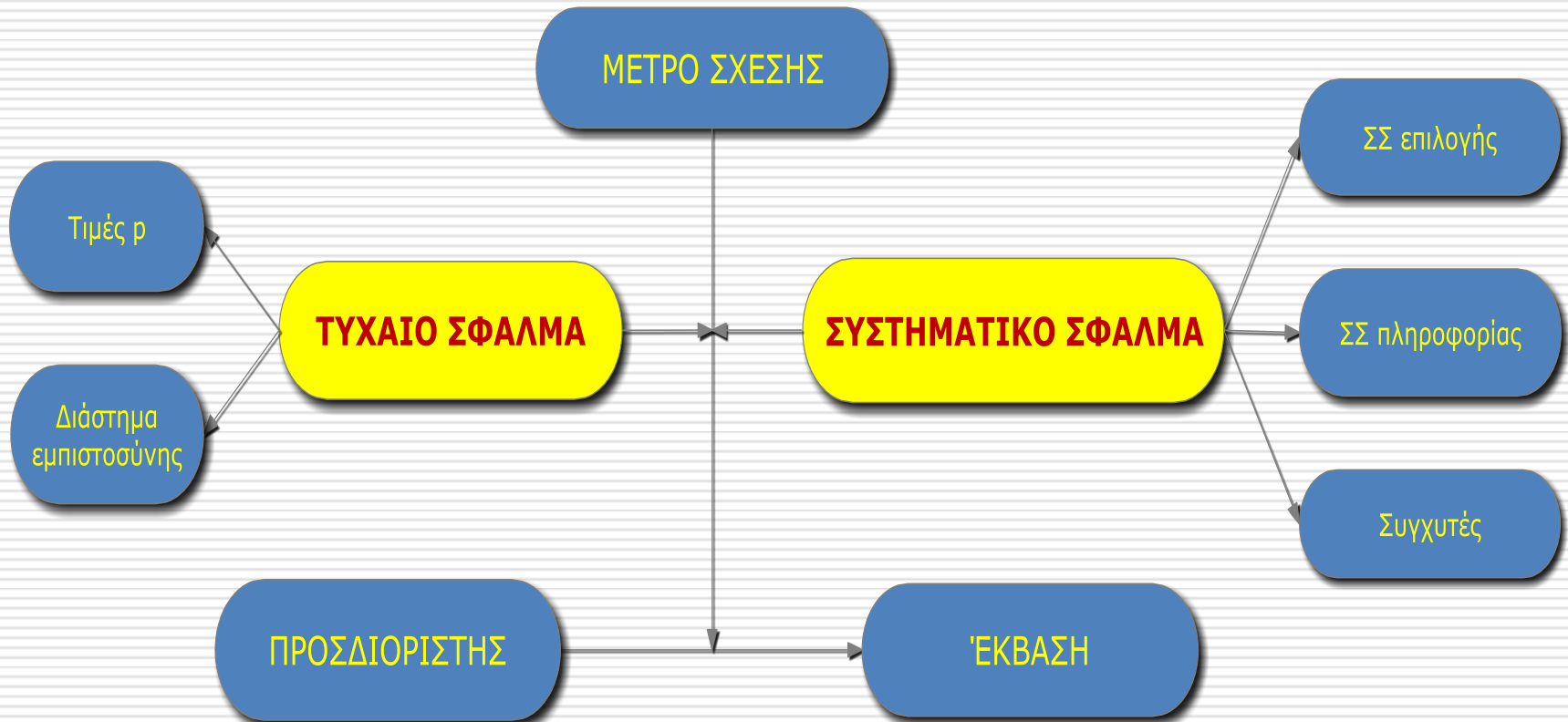
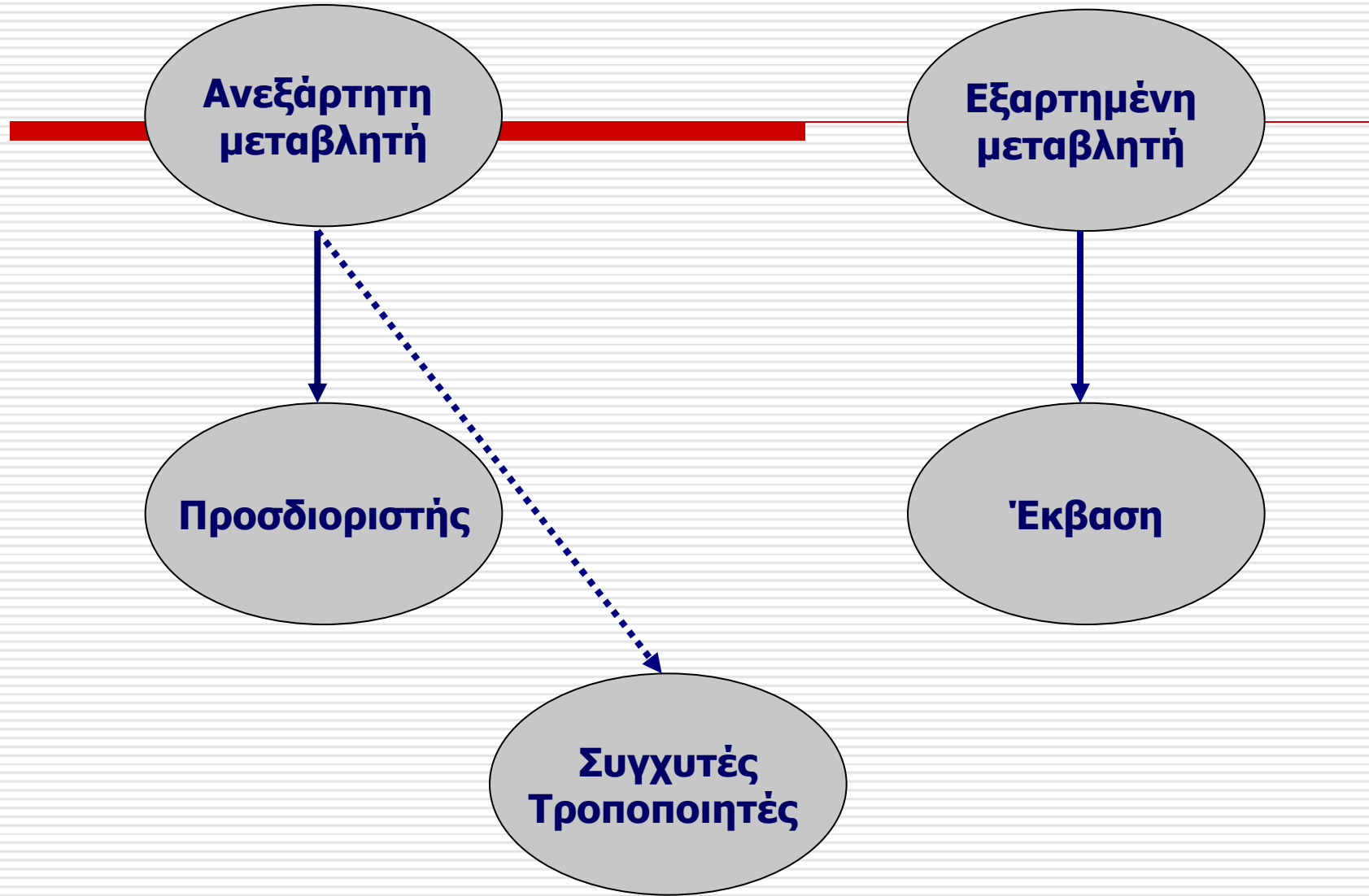
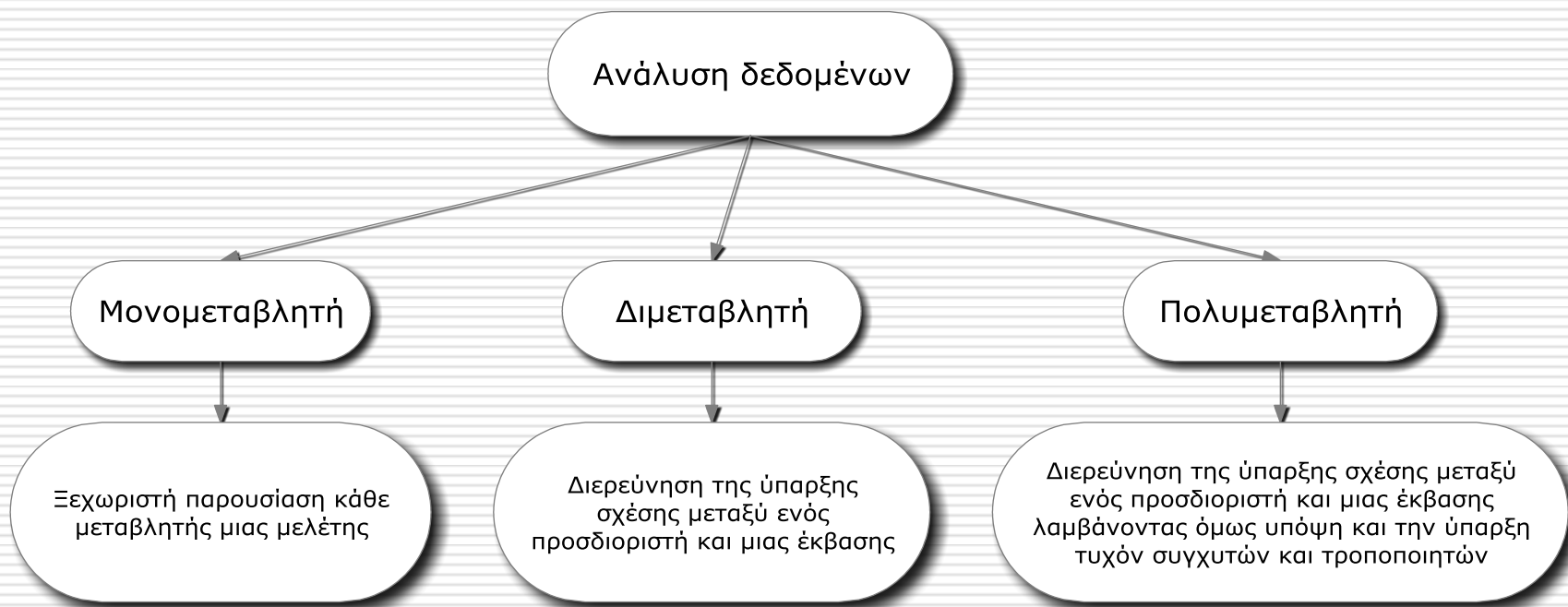

Διμεταβλητή ανάλυση

Αντικείμενο των επιδημιολογικών μελετών







Διμεταβλητή ανάλυση

- Διερεύνηση της ύπαρξης σχέσης μεταξύ προσδιοριστή και έκβασης
 - Παρουσιάζεται το **αδρό μέτρο σχέσης** (crude measure of association) που δηλώνει τη σχέση μεταξύ προσδιοριστή και έκβασης
 - Για την εκτίμηση του τυχαίου σφάλματος στον υπολογισμό του μέτρου σχέσης παρουσιάζεται το διάστημα εμπιστοσύνης
-

Πολυμεταβλητή ανάλυση

- Διερεύνηση της ύπαρξης σχέσης μεταξύ προσδιοριστή και έκβασης λαμβάνοντας υπόψη και τη δράση συγχυτών
- Παρουσιάζεται το **διορθωμένο ή, αλλιώς, σταθμισμένο μέτρο σχέσης** (adjusted measure of association) που δηλώνει τη σχέση μεταξύ προσδιοριστή και έκβασης, εξουδετερώνοντας τη δράση των συγχυτών
- Για την εκτίμηση του τυχαίου σφάλματος στον υπολογισμό του μέτρου σχέσης παρουσιάζεται το διάστημα εμπιστοσύνης

Διμεταβλητή ανάλυση



Πολυμεταβλητή ανάλυση

Διορθωμένη σχέση

- εξουδετερώνονται οι συγχυτές
- εκτιμάται η ύπαρξη τροποποίησης

ΠΡΟΣΔΙΟΡΙΣΤΗΣ

ΕΚΒΑΣΗ

Διορθωμένο μέτρο σχέσης

Στατιστικοί έλεγχοι

Εξαρτημένη μεταβλητή

Διχοτόμος

Ανεξάρτητα "δείγματα"

Ανεξάρτητη μεταβλητή

Διχοτόμος

Ονομαστική
(>2 κατηγορίες)

Διατάξιμη

Ποσοτική

"Δείγματα" ανά ζεύγη

Ανεξάρτητη μεταβλητή

Διχοτόμος

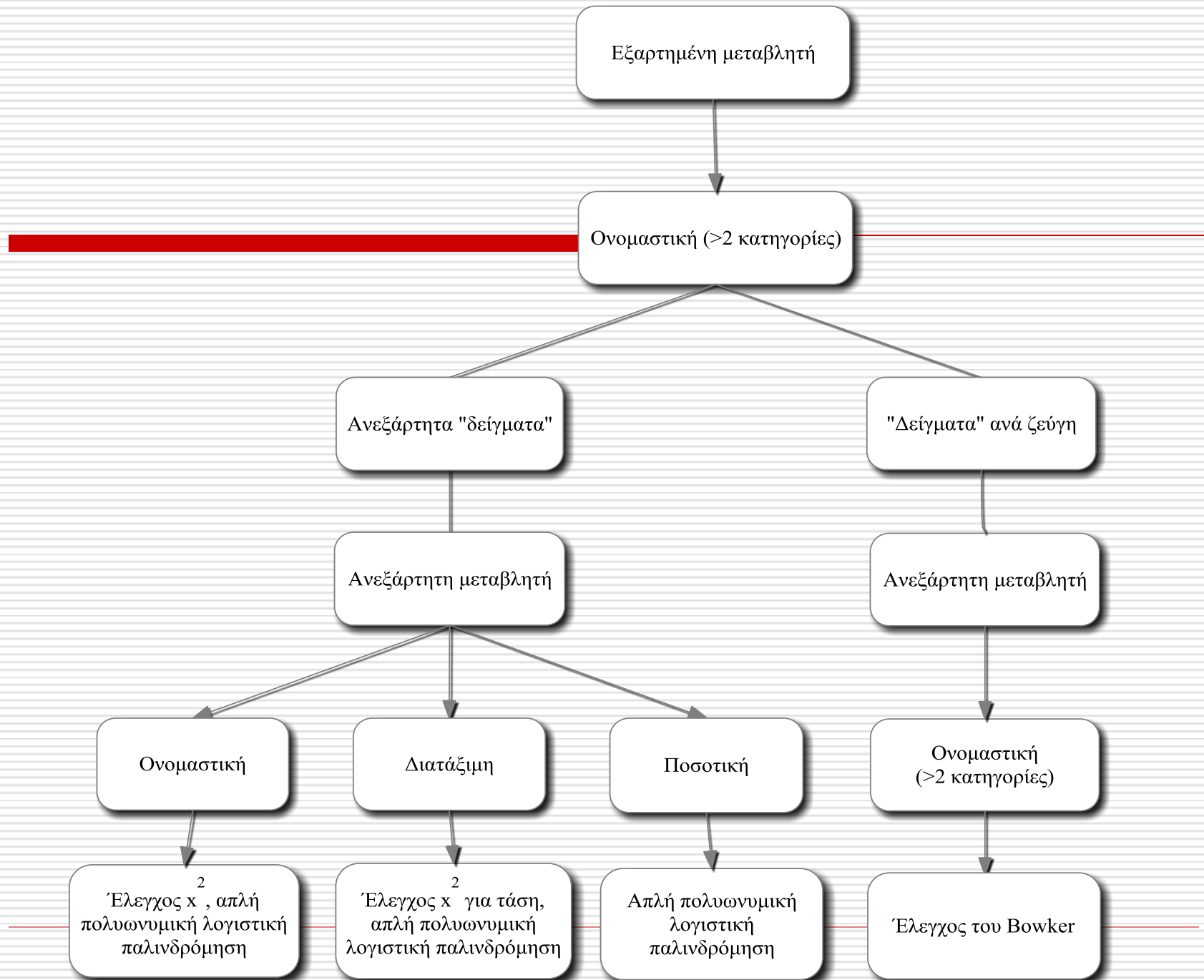
χ^2
Έλεγχος χ^2 , ακριβής
έλεγχος του Fisher,
απλή λογιστική
παλινδρόμηση

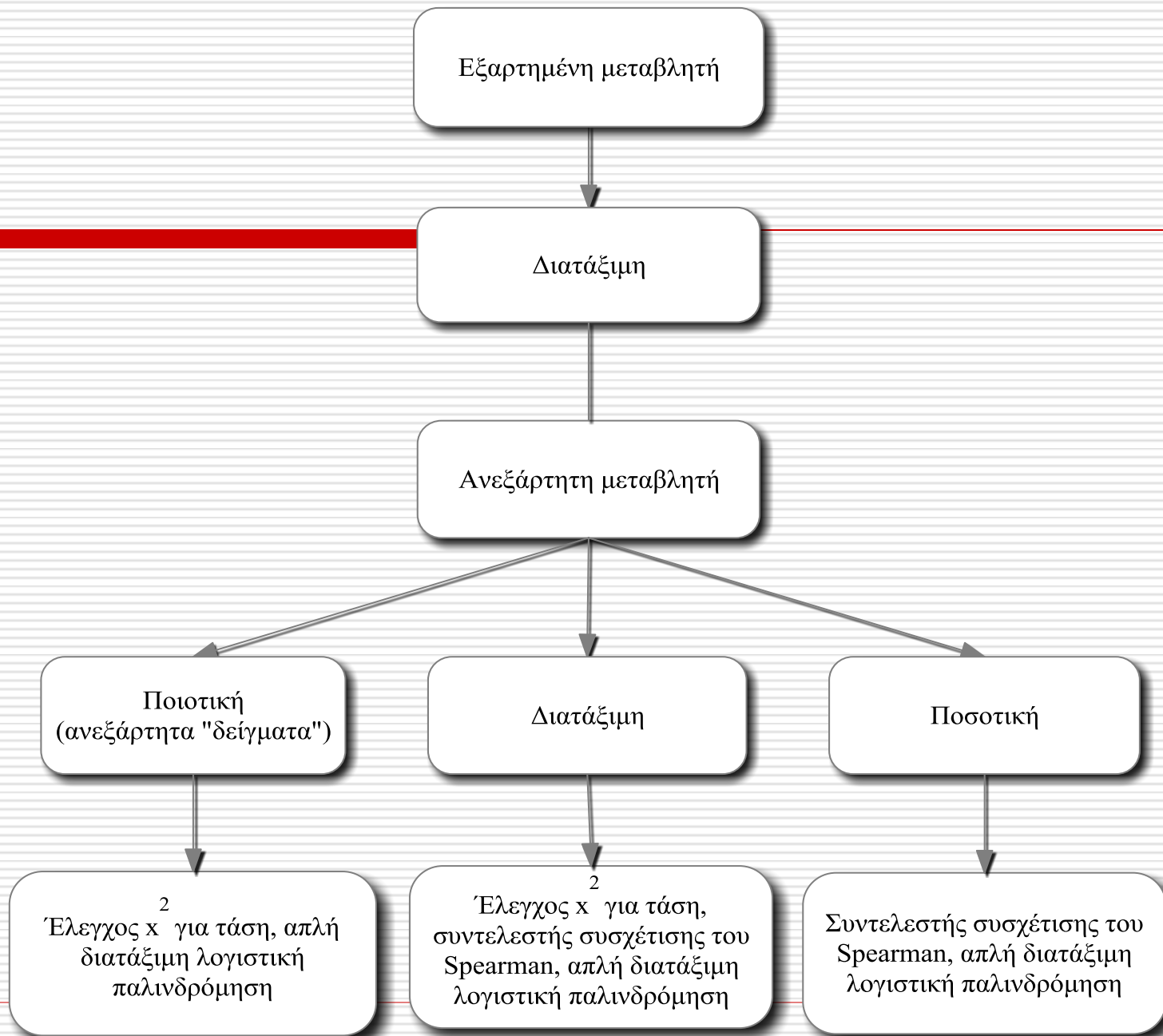
χ^2
Έλεγχος χ^2 , ακριβής
έλεγχος Monte-Carlo,
απλή λογιστική
παλινδρόμηση

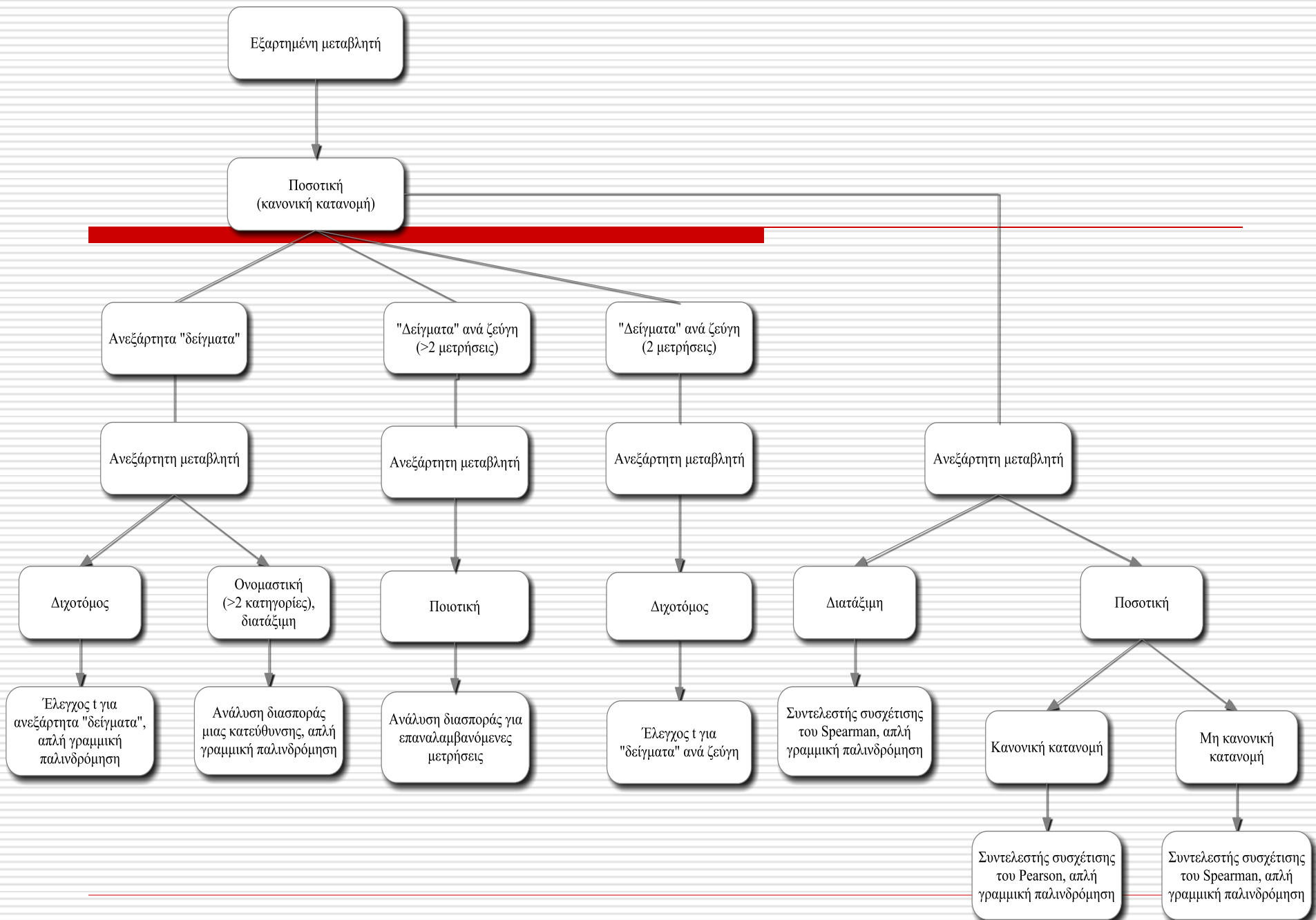
χ^2
Έλεγχος χ^2 για τάση,
απλή λογιστική
παλινδρόμηση

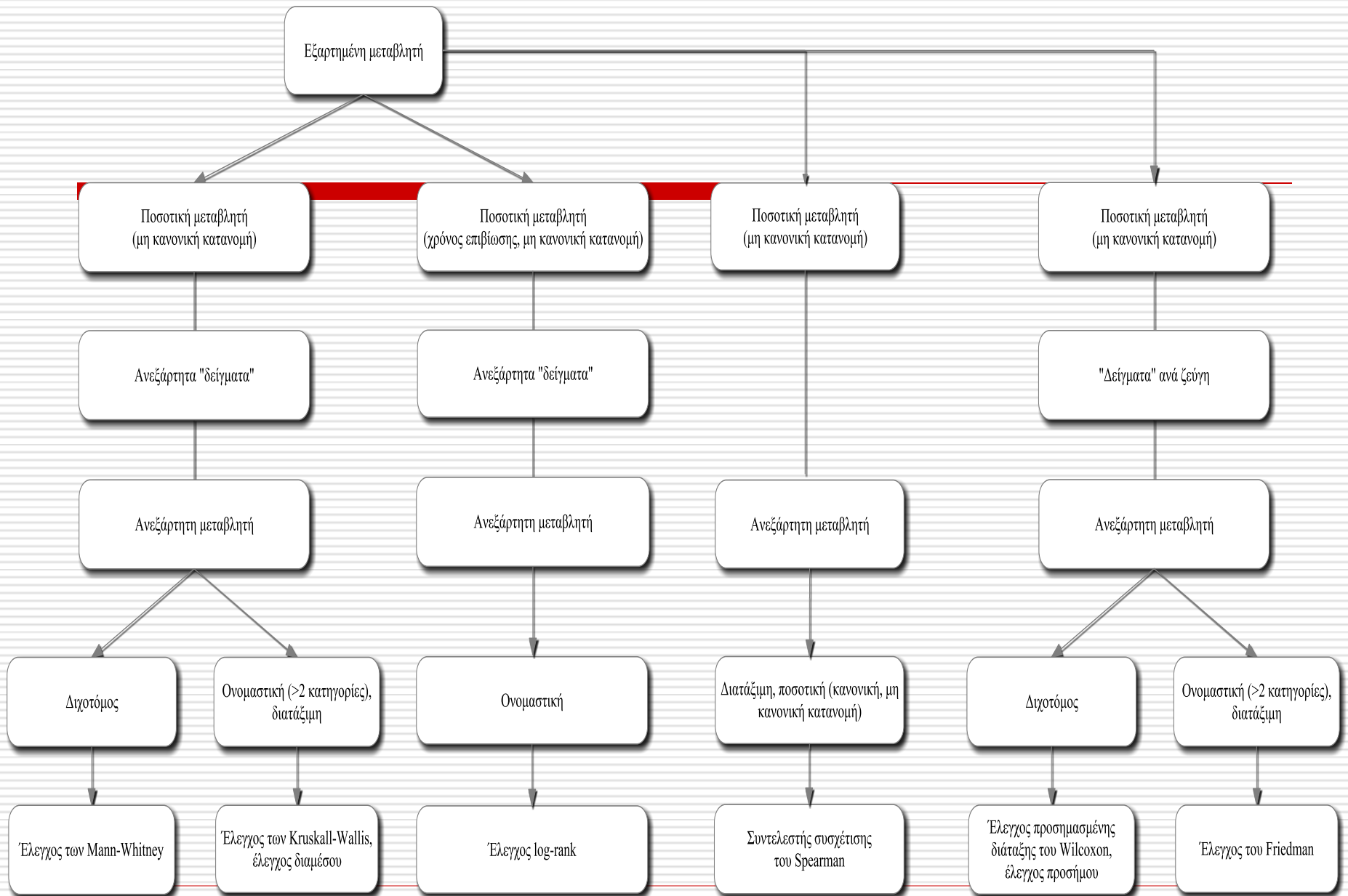
Απλή λογιστική
παλινδρόμηση

Έλεγχος του McNemar









Έλεγχος χ^2

- Ανεξάρτητα δείγματα
 - Προσδιοριστής: ονομαστική μεταβλητή
 - Έκβαση: ονομαστική μεταβλητή
 - Σχέση μεταξύ φύλου και οικογενειακής κατάστασης
 - Σχέση μεταξύ φύλου και αρτηριακής πίεσης (υπερτασικοί και μη)
-

Έλεγχος χ^2

- Στην περίπτωση που ο αριθμός των παρατηρήσεων σε ορισμένες κατηγορίες είναι μικρός χρησιμοποιούνται ακριβείς έλεγχοι
 - Πιο συγκεκριμένα, στην περίπτωση που η ανεξάρτητη μεταβλητή είναι διχοτόμος χρησιμοποιείται ο **ακριβής έλεγχος του Fisher** (Fisher's exact test), ενώ στην περίπτωση που η ανεξάρτητη μεταβλητή είναι ονομαστική με >2 κατηγορίες χρησιμοποιείται ο **ακριβής έλεγχος Monte-Carlo** (Monte-Carlo exact test)
-

Λόγος των odds

- Υπολογίζεται ο λόγος των odds της έκβασης στους εκτεθειμένους σε σχέση με τους μη εκτεθειμένους
 - Όταν ο λόγος των odds είναι ίσος με 1, τότε δεν υπάρχει σχέση μεταξύ προσδιοριστή και έκβασης
 - Όταν ο λόγος των odds είναι >1 , τότε οι εκτεθειμένοι εμφανίζουν συχνότερα την έκβαση σε σχέση με τους μη εκτεθειμένους
 - Όταν ο λόγος των odds είναι <1 , τότε οι μη εκτεθειμένοι εμφανίζουν συχνότερα την έκβαση σε σχέση με τους εκτεθειμένους
-

Λόγος των odds

$$odds = \frac{p}{1 - p}$$

$$\text{Λόγος των odds} = \frac{\text{odds στους εκτεθειμένους}}{\text{odds στους μη εκτεθειμένους}} = \frac{\frac{P(\text{έκβασης/έκθεση})}{1 - P(\text{έκβασης/έκθεση})}}{\frac{P(\text{έκβασης/μη έκθεση})}{1 - P(\text{έκβασης/μη έκθεση})}}$$

Μηδενική και εναλλακτική υπόθεση

- ❑ Μια μελέτη διερευνά τη σχέση μεταξύ της χρήσης κινητών τηλεφώνων (διχοτόμος μεταβλητή: χρήστες και μη χρήστες) και της συχνότητας εμφάνισης όγκων στον εγκέφαλο (διχοτόμος μεταβλητή: εμφάνιση ή όχι όγκων)
 - ❑ H_0 : δεν υπάρχει σχέση μεταξύ της χρήσης κινητών τηλεφώνων και της συχνότητας εμφάνισης όγκων στον εγκέφαλο
 H_0 : το ποσοστό όγκων στους χρήστες κινητών είναι ίσο με το ποσοστό όγκων στους μη χρήστες κινητών
 - ❑ H_A : υπάρχει σχέση μεταξύ της χρήσης κινητών τηλεφώνων και της συχνότητας εμφάνισης όγκων στον εγκέφαλο
 H_A : το ποσοστό όγκων στους χρήστες κινητών είναι διαφορετικό από το ποσοστό όγκων στους μη χρήστες κινητών
-

Έλεγχος χ^2

- Η κατανομή χ^2 έχει τους εξής βαθμούς ελευθερίας:

$$(\text{γραμμές} - 1) \times (\text{στήλες} - 1)$$

- Όταν ο προσδιοριστής και η έκβαση είναι διχοτόμες μεταβλητές, τότε η κατανομή χ^2 έχει έναν βαθμό ελευθερίας
-

Παρατηρούμενες απόλυτες συχνότητες

		Έκβαση (εξαρτημένη μεταβλητή)		Σύνολο
		Ναι	Όχι	
Προσδιοριστής (ανεξάρτητη μεταβλητή)	Ενδεικτική κατηγορία (εκτεθειμένοι)	a	b	$a+b$
	Κατηγορία αναφοράς (μη εκτεθειμένοι)	c	d	$c+d$
Σύνολο		$a+c$	$b+d$	$a+b+c+d$

Αναμενόμενες απόλυτες συχνότητες

		Έκβαση (εξαρτημένη μεταβλητή)		Σύνολο
		Ναι	Όχι	
Προσδιοριστής (ανεξάρτητη μεταβλητή)	Ενδεικτική κατηγορία (εκτεθειμένοι)	$\frac{(a+b) \times (a+c)}{n}$	$\frac{(a+b) \times (b+d)}{n}$	$a+b$
	Κατηγορία αναφοράς (μη εκτεθειμένοι)	$\frac{(c+d) \times (a+c)}{n}$	$\frac{(c+d) \times (b+d)}{n}$	$c+d$
	Σύνολο	$a+c$	$b+d$	$n = a+b+c+d$

Λόγος των odds

$$OR = \frac{odds_{\text{εκτεθειμένοι}}}{odds_{\text{μη εκτεθειμένοι}}} = \frac{\frac{a}{b}}{\frac{c}{d}} = \frac{a \times d}{b \times c}$$

$$SE(\ln OR) = \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}$$

$$95\% \Delta E_{OR} = e^{\ln OR \pm 1,96 \times \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}}$$

Έλεγχος χ^2

$$\chi^2 = \sum_{i=1}^{\text{αριθμός κελιών}} \frac{(\text{παρατηρούμενη συχνότητα}_i - \text{αναμενόμενη συχνότητα}_i)^2}{\text{αναμενόμενη συχνότητα}_i}$$

$$\text{βαθμοί ελευθερίας} = (\text{γραμμές} - 1) \times (\text{στήλες} - 1)$$

Κατανομή χ^2

Στον πίνακα παρουσιάζονται οι επιφάνειες στην μια πλευρά της κατανομής

**Στην πρώτη στήλη
βρίσκονται οι
βαθμοί ελευθερίας**

df (β.ε.)	Επιφάνεια στην Άνω Πλευρά				
	0,100	0,050	0,025	0,010	0,001
1	2,71	3,84	5,02	6,63	10,83
2	4,61	5,99	7,38	9,21	13,82
3	6,25	7,81	9,35	11,34	16,27
4	7,78	9,49	11,14	13,28	18,47
5	9,24	11,07	12,83	15,09	20,52
6	10,64	12,59	14,45	16,81	22,46
7	12,02	14,07	16,01	18,48	24,32
8	13,36	15,51	17,53	20,09	26,12
9	14,68	16,92	19,02	21,67	27,88
10	15,99	18,31	20,48	23,21	29,59
11	17,28	19,68	21,92	24,72	31,26
12	18,55	21,03	23,34	26,22	32,91
13	19,81	22,36	24,74	27,69	34,53
14	21,06	23,68	26,12	29,14	36,12
15	22,31	25,00	27,49	30,58	37,70
16	23,54	26,30	28,85	32,00	39,25
17	24,77	27,59	30,19	33,41	40,79
18	25,99	28,87	31,53	34,81	42,31
19	27,20	30,14	32,85	36,19	43,82
20	28,41	31,41	34,17	37,57	45,31
21	29,62	32,67	35,48	38,93	46,80
22	30,81	33,92	36,78	40,29	48,27
23	32,01	35,17	38,08	41,64	49,73
24	33,20	36,42	39,36	42,98	51,18
25	34,38	37,65	40,65	44,31	52,62

**Στην πρώτη σειρά
βρίσκονται οι
πιθανότητες p**

**Εντός του πίνακα
βρίσκονται οι τιμές χ^2**

Έλεγχος χ^2

- ❑ Χρησιμοποιώντας την κατανομή χ^2 και τους βαθμούς ελευθερίας της κατανομής υπολογίζεται η τιμή p
 - ❑ Εάν $p < 0,05$, τότε απορρίπτεται η μηδενική υπόθεση και υπάρχει στατιστικά σημαντική σχέση μεταξύ προσδιοριστή και έκβασης
 - ❑ Εάν $p \geq 0,05$, τότε δεν απορρίπτεται η μηδενική υπόθεση και δεν υπάρχει στατιστικά σημαντική σχέση μεταξύ προσδιοριστή και έκβασης
 - ❑ Τέλος, υπολογίζεται το 95% διάστημα εμπιστοσύνης για τον λόγο των odds
-

Έλεγχος χ^2

- Δεδομένα μιας μελέτης για τη διερεύνηση της ύπαρξης σχέσης μεταξύ της κατανάλωσης οινοπνεύματος (προσδιοριστή) και της συχνότητας εμφάνισης τροχαίου ατυχήματος (έκβασης)

		Τροχαίο ατύχημα		Σύνολο
		Ναι	Όχι	
Κατανάλωση οινοπνεύματος	Ναι	50	90	140
	Όχι	10	90	100
Σύνολο		60	180	240

Κατανάλωση οινοπνεύματος * Τροχαίο ατύχημα Crosstabulation

			Τροχαίο ατύχημα		
			Ναι	Όχι	Total
Κατανάλωση οινοπνεύματος	Ναι	Count	50	90	140
		% within Κατανάλωση οινοπνεύματος	35,7%	64,3%	100,0%
		% of Total	20,8%	37,5%	58,3%
	Όχι	Count	10	90	100
		% within Κατανάλωση οινοπνεύματος	10,0%	90,0%	100,0%
		% of Total	4,2%	37,5%	41,7%
	Total	Count	60	180	240
		% within Κατανάλωση οινοπνεύματος	25,0%	75,0%	100,0%
		% of Total	25,0%	75,0%	100,0%

Chi-Square Tests					
	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	20,571^a	1	,000		
Continuity Correction ^b	19,223	1	,000		
Likelihood Ratio	22,412	1	,000		
Fisher's Exact Test				,000	,000
Linear-by-Linear Association	20,486	1	,000		
N of Valid Cases	240				
a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 25,00.					
b. Computed only for a 2x2 table					



Risk Estimate			
	Value	95% Confidence Interval	
		Lower	Upper
Odds Ratio for Κατανάλωση οινόπνεύματος (Ναι / Όχι)	5,000	2,388	10,471
For cohort Τροχαίο ατύχημα = Ναι	3,571	1,905	6,696
For cohort Τροχαίο ατύχημα = Όχι	,714	,621	,821
N of Valid Cases	240		



Έλεγχος t

- Ανεξάρτητα δείγματα
 - Προσδιοριστής: διχοτόμος μεταβλητή
 - Έκβαση: ποσοτική μεταβλητή που ακολουθεί την κανονική κατανομή στις δυο κατηγορίες του προσδιοριστή
 - Σχέση μεταξύ φύλου και δείκτη μάζας σώματος
 - Σχέση μεταξύ φύλου και συστολικής αρτηριακής πίεσης
-

Μηδενική και εναλλακτική υπόθεση

- Μια μελέτη διερευνά τη σχέση μεταξύ του φύλου (διχοτόμος μεταβλητή: άντρες και γυναίκες) και της συστολικής αρτηριακής πίεσης (ποσοτική μεταβλητή που ακολουθεί την κανονική κατανομή στον πληθυσμό των αντρών και των γυναικών)
 - H_0 : δεν υπάρχει σχέση μεταξύ του φύλου και της συστολικής αρτηριακής πίεσης
 H_0 : η μέση συστολική αρτηριακή πίεση στους άντρες είναι ίση με τη μέση συστολική αρτηριακή πίεση στις γυναίκες
 - H_A : υπάρχει σχέση μεταξύ του φύλου και της συστολικής αρτηριακής πίεσης
 H_A : η μέση συστολική αρτηριακή πίεση στους άντρες είναι δεν είναι ίση με τη μέση συστολική αρτηριακή πίεση στις γυναίκες
-

Έλεγχος t

- Όταν οι διασπορές των πληθυσμών από τους οποίους προέρχονται τα δείγματα είναι ίσες

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

$$\mu_1 - \mu_2 = 0$$

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$\text{βαθμοί ελευθερίας} = n_1 + n_2 - 2$$

Έλεγχος t

- Όταν οι διασπορές των πληθυσμών από τους οποίους προέρχονται τα δείγματα είναι άνισες

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$\mu_1 - \mu_2 = 0$$

$$\text{βαθμοί ελευθερίας} = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \right)^2}{\frac{\left(\frac{s_1^2}{n_1} \right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2} \right)^2}{n_2 - 1}}$$

Κατανομή t
Στον πίνακα παρουσιάζονται οι επιφάνειες στην μια πλευρά της κατανομής

Επιφάνεια στη Μία Πλευρά

df (β.ε.)	0,10	0,05	0,025	0,01	0,005	0,0005
1	3,078	6,314	12,706	31,821	63,657	636,619
2	1,886	2,920	4,303	6,965	9,925	31,599
3	1,638	2,353	3,182	4,541	5,841	12,924
4	1,533	2,132	2,776	3,747	4,604	8,610
5	1,476	2,015	2,571	3,365	4,032	6,869
6	1,440	1,943	2,447	3,143	3,707	5,959
7	1,415	1,895	2,365	2,998	3,499	5,408
8	1,397	1,860	2,306	2,896	3,355	5,041
9	1,383	1,833	2,262	2,821	3,250	4,781
10	1,372	1,812	2,228	2,764	3,169	4,587
11	1,363	1,796	2,201	2,718	3,106	4,437
12	1,356	1,782	2,179	2,681	3,055	4,318
13	1,350	1,771	2,160	2,650	3,012	4,221
14	1,345	1,761	2,145	2,624	2,977	4,140
15	1,341	1,753	2,131	2,602	2,947	4,073
16	1,337	1,746	2,120	2,583	2,921	4,015
17	1,333	1,740	2,110	2,567	2,898	3,965
18	1,330	1,734	2,101	2,552	2,878	3,922
19	1,328	1,729	2,093	2,539	2,861	3,883
20	1,325	1,725	2,086	2,528	2,845	3,850
21	1,323	1,721	2,080	2,518	2,831	3,819
22	1,321	1,717	2,074	2,508	2,819	3,792
23	1,319	1,714	2,069	2,500	2,807	3,768
24	1,318	1,711	2,064	2,492	2,797	3,745
25	1,316	1,708	2,060	2,485	2,787	3,725
26	1,315	1,706	2,056	2,479	2,779	3,707
27	1,314	1,703	2,052	2,473	2,771	3,690
28	1,313	1,701	2,048	2,467	2,763	3,674
29	1,311	1,699	2,045	2,462	2,756	3,659
30	1,310	1,697	2,042	2,457	2,750	3,646
40	1,303	1,684	2,021	2,423	2,704	3,551
50	1,299	1,676	2,009	2,403	2,678	3,496
60	1,296	1,671	2,000	2,390	2,660	3,460
70	1,294	1,667	1,994	2,381	2,648	3,435
80	1,292	1,664	1,990	2,374	2,639	3,416
90	1,291	1,662	1,987	2,368	2,632	3,402
100	1,290	1,660	1,984	2,364	2,626	3,390
110	1,289	1,659	1,982	2,361	2,621	3,381
120	1,289	1,658	1,980	2,358	2,617	3,373
∞	1,282	1,645	1,960	2,327	2,576	3,291

Στην πρώτη στήλη
βρίσκονται οι
βαθμοί ελευθερίας

Στην πρώτη
γραμμή
βρίσκονται οι
πιθανότητες p

Εντός του πίνακα
βρίσκονται οι
τιμές t

Έλεγχος t

- ❑ Χρησιμοποιώντας την κατανομή t και τους βαθμούς ελευθερίας της κατανομής υπολογίζεται η τιμή p
 - ❑ Εάν $p < 0,05$, τότε απορρίπτεται η μηδενική υπόθεση και υπάρχει στατιστικά σημαντική σχέση μεταξύ προσδιοριστή και έκβασης.
 - ❑ Εάν $p \geq 0,05$, τότε δεν απορρίπτεται η μηδενική υπόθεση και δεν υπάρχει στατιστικά σημαντική σχέση μεταξύ προσδιοριστή και έκβασης
 - ❑ Τέλος, υπολογίζεται το 95% διάστημα εμπιστοσύνης για τη διαφορά των μέσων τιμών
-

Έλεγχος t

- Διερεύνηση της ύπαρξης σχέσης μεταξύ καπνιστικής συνήθειας (διχοτόμου μεταβλητής: καπνιστές και μη καπνιστές) και συστολικής αρτηριακής πίεσης (ποσοτική μεταβλητή που ακολουθεί την κανονική κατανομή στον πληθυσμό των καπνιστών και των μη καπνιστών)
-

Group Statistics

Καπνιστική συνήθεια	N	Mean	Std. Deviation	Std. Error Mean
Συστολική αρτηριακή πίεση Καπνιστές	84	129,3810	6,80876	,74290
Μη καπνιστές	139	128,9784	7,20051	,61074



Independent Samples Test											
		Levene's Test for Equality of Variances				t-test for Equality of Means					
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference		
Συστολική αρτηριακή πίεση	Equal variances assumed	,815	,368	,413	221	,680	,40254	,97512	-1,51920	2,32427	
	Equal variances not assumed			,419	182,866	,676	,40254	,96172	-1,49495	2,30002	



Έλεγχος Mann-Whitney

- Ανεξάρτητα δείγματα
 - Προσδιοριστής: διχοτόμος μεταβλητή
 - Έκβαση: ποσοτική μεταβλητή που δεν ακολουθεί την κανονική κατανομή στις δυο κατηγορίες του προσδιοριστή
 - Σχέση μεταξύ φύλου και δείκτη μάζας σώματος
 - Σχέση μεταξύ φύλου και συστολικής αρτηριακής πίεσης
-

Μηδενική και εναλλακτική υπόθεση

- ❑ Μια μελέτη διερευνά τη σχέση μεταξύ του φύλου (διχοτόμος μεταβλητή: άντρες και γυναίκες) και της συστολικής αρτηριακής πίεσης (ποσοτική μεταβλητή που δεν ακολουθεί την κανονική κατανομή στον πληθυσμό των αντρών και των γυναικών)
 - ❑ H_0 : δεν υπάρχει σχέση μεταξύ του φύλου και της συστολικής αρτηριακής πίεσης
 H_0 : η διάμεση συστολική αρτηριακή πίεση στους άντρες είναι ίση με τη διάμεση συστολική αρτηριακή πίεση στις γυναίκες
 - ❑ H_A : υπάρχει σχέση μεταξύ του φύλου και της συστολικής αρτηριακής πίεσης
 H_A : η διάμεση συστολική αρτηριακή πίεση στους άντρες είναι δεν είναι ίση με τη διάμεση συστολική αρτηριακή πίεση στις γυναίκες
-

Έλεγχος Mann-Whitney

- ❑ Χρησιμοποιώντας την κατανομή Z υπολογίζεται η τιμή p
 - ❑ Εάν $p < 0,05$, τότε απορρίπτεται η μηδενική υπόθεση και υπάρχει στατιστικά σημαντική σχέση μεταξύ προσδιοριστή και έκβασης
 - ❑ Εάν $p \geq 0,05$, τότε δεν απορρίπτεται η μηδενική υπόθεση και δεν υπάρχει στατιστικά σημαντική σχέση μεταξύ προσδιοριστή και έκβασης
-

Έλεγχος Mann-Whitney

- Διερεύνηση της ύπαρξης σχέσης μεταξύ καπνιστικής συνήθειας (δихοτόμος μεταβλητή: καπνιστές και μη καπνιστές) και συστολικής αρτηριακής πίεσης (ποσοτική μεταβλητή που δεν ακολουθεί την κανονική κατανομή στον πληθυσμό των καπνιστών και των μη καπνιστών)
-



Ranks				
	Καπνιστική συνήθεια	N	Mean Rank	Sum of Ranks
Συστολική αρτηριακή πίεση (mmHg)	Μη καπνιστές	19	10,00	190,00
	Καπνιστές	11	25,00	275,00
	Total	30		



Test Statistics ^b	
Συστολική αρτηριακή πίεση (mmHg)	
Mann-Whitney U	,000
Wilcoxon W	190,000
Z	-4,500
Asymp. Sig. (2-tailed)	,000
Exact Sig. [2*(1-tailed Sig.)]	,000 ^a
a. Not corrected for ties.	
b. Grouping Variable: Καπνιστική συνήθεια	

Ανάλυση διασποράς μιας κατεύθυνσης (ANOVA)

- ❑ Ανεξάρτητα δείγματα
 - ❑ Προσδιοριστής: ποιοτική μεταβλητή με >2 κατηγορίες
 - ❑ Έκβαση: ποσοτική μεταβλητή που ακολουθεί την κανονική κατανομή στις >2 κατηγορίες του προσδιοριστή
 - ❑ Σχέση μεταξύ εκπαιδευτικού επιπέδου (κατώτερο, μέσο και ανώτερο) και εισοδήματος
 - ❑ Σχέση μεταξύ ομάδας αίματος και συστολικής αρτηριακής πίεσης
-

Μηδενική και εναλλακτική υπόθεση

- Μια μελέτη διερευνά τη σχέση μεταξύ του εκπαιδευτικού επιπέδου (κατώτερο, μέσο και ανώτερο) και του εισοδήματος (ποσοτική μεταβλητή που ακολουθεί την κανονική κατανομή στον πληθυσμό των ατόμων με κατώτερο, μέσο και ανώτερο εκπαιδευτικό επίπεδο)
 - H_0 : δεν υπάρχει σχέση μεταξύ του εκπαιδευτικού επιπέδου και του εισοδήματος
 H_0 : το μέσο εισόδημα στα άτομα των τριών κατηγοριών εκπαιδευτικού επιπέδου είναι ίσο
 - H_A : υπάρχει σχέση μεταξύ του εκπαιδευτικού επιπέδου και του εισοδήματος
 H_A : το μέσο εισόδημα στα άτομα των τριών κατηγοριών εκπαιδευτικού επιπέδου δεν είναι ίσο
-

Ανάλυση διασποράς μιας κατεύθυνσης (ANOVA)

- ❑ Χρησιμοποιώντας την κατανομή F και τους βαθμούς ελευθερίας της κατανομής υπολογίζεται η τιμή p
 - ❑ Εάν $p < 0,05$, τότε απορρίπτεται η μηδενική υπόθεση και υπάρχει στατιστικά σημαντική σχέση μεταξύ προσδιοριστή και έκβασης. Στην περίπτωση αυτή εξετάζονται οι διαφορές των μέσων τιμών ανά δυο κατηγορίες
 - ❑ Εάν $p \geq 0,05$, τότε δεν απορρίπτεται η μηδενική υπόθεση και δεν υπάρχει στατιστικά σημαντική σχέση μεταξύ προσδιοριστή και έκβασης
 - ❑ Τέλος, υπολογίζεται το 95% διάστημα εμπιστοσύνης για τις διαφορές των μέσων τιμών στις διάφορες κατηγορίες του προσδιοριστή
-

Ανάλυση διασποράς μιας κατεύθυνσης (ANOVA)

- Διερεύνηση της ύπαρξης σχέσης μεταξύ του εκπαιδευτικού επιπέδου (κατώτερο, μέσο και ανώτερο) και του εισοδήματος (ποσοτική μεταβλητή που ακολουθεί την κανονική κατανομή στον πληθυσμό των ατόμων με κατώτερο, μέσο και ανώτερο εκπαιδευτικό επίπεδο)
-



Descriptives								
Μηνιαίο εισόδημα								
95% Confidence Interval for								
Mean								
	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Minimum	Maximum
Κατώτερο	30	881,7000	147,85364	26,99426	826,4905	936,9095	567,00	1123,00
Μέσο	30	1156,8000	160,58565	29,31879	1096,8363	1216,7637	985,00	1450,00
Ανώτερο	30	1341,5000	176,17385	32,16480	1275,7156	1407,2844	1000,00	1565,00
Total	90	1126,6667	248,45559	26,18952	1074,6287	1178,7047	567,00	1565,00



ANOVA					
Μηνιαίο εισόδημα					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	3212101,400	2	1606050,700	61,233	,000
Within Groups	2281884,600	87	26228,559		
Total	5493986,000	89			

Multiple Comparisons							
Dependent Variable:Μηνιαίο εισόδημα							
						95% Confidence Interval	
	(I) Εκπαιδευτικό επίπεδο	(J) Εκπαιδευτικό επίπεδο	Mean Difference (I-J)	Std. Error	Sig.	Lower Bound	Upper Bound
Bonferroni	Κατώτερο	Μέσο	-275,10000*	41,81591	,000	-377,1789	-173,0211
		Ανώτερο	-459,80000*	41,81591	,000	-561,8789	-357,7211
	Μέσο	Κατώτερο	275,10000*	41,81591	,000	173,0211	377,1789
		Ανώτερο	-184,70000*	41,81591	,000	-286,7789	-82,6211
	Ανώτερο	Κατώτερο	459,80000*	41,81591	,000	357,7211	561,8789
		Μέσο	184,70000*	41,81591	,000	82,6211	286,7789
Dunnett T3	Κατώτερο	Μέσο	-275,10000*	39,85325	,000	-372,9951	-177,2049
		Ανώτερο	-459,80000*	41,99124	,000	-563,0151	-356,5849
	Μέσο	Κατώτερο	275,10000*	39,85325	,000	177,2049	372,9951
		Ανώτερο	-184,70000*	43,52202	,000	-291,6123	-77,7877
	Ανώτερο	Κατώτερο	459,80000*	41,99124	,000	356,5849	563,0151
		Μέσο	184,70000*	43,52202	,000	77,7877	291,6123

*. The mean difference is significant at the 0.05 level.

Έλεγχος Kruskal-Wallis

- Ανεξάρτητα δείγματα
 - Προσδιοριστής: ποιοτική μεταβλητή με >2 κατηγορίες
 - Έκβαση: ποσοτική μεταβλητή που δεν ακολουθεί την κανονική κατανομή στις >2 κατηγορίες του προσδιοριστή
 - Σχέση μεταξύ εκπαιδευτικού επιπέδου (κατώτερο, μέσο και ανώτερο) και εισοδήματος
 - Σχέση μεταξύ ομάδας αίματος και συστολικής αρτηριακής πίεσης
-

Μηδενική και εναλλακτική υπόθεση

- ❑ Μια μελέτη διερευνά τη σχέση μεταξύ του εκπαιδευτικού επιπέδου (κατώτερο, μέσο και ανώτερο) και του εισοδήματος (ποσοτική μεταβλητή που δεν ακολουθεί την κανονική κατανομή στον πληθυσμό των ατόμων με κατώτερο, μέσο και ανώτερο εκπαιδευτικό επίπεδο)
 - ❑ H_0 : δεν υπάρχει σχέση μεταξύ του εκπαιδευτικού επιπέδου και του εισοδήματος
 H_0 : το διάμεσο εισόδημα στα άτομα των τριών κατηγοριών εκπαιδευτικού επιπέδου είναι ίσο
 - ❑ H_A : υπάρχει σχέση μεταξύ του εκπαιδευτικού επιπέδου και του εισοδήματος
 H_A : το διάμεσο εισόδημα στα άτομα των τριών κατηγοριών εκπαιδευτικού επιπέδου δεν είναι ίσο
-

Έλεγχος Kruskal-Wallis

- ❑ Χρησιμοποιώντας την κατανομή χ^2 και τους βαθμούς ελευθερίας της κατανομής υπολογίζεται η τιμή p
 - ❑ Εάν $p < 0,05$, τότε απορρίπτεται η μηδενική υπόθεση και υπάρχει στατιστικά σημαντική σχέση μεταξύ προσδιοριστή και έκβασης. Στην περίπτωση αυτή εξετάζονται οι διαφορές των διάμεσων τιμών ανά δυο κατηγορίες
 - ❑ Εάν $p \geq 0,05$, τότε δεν απορρίπτεται η μηδενική υπόθεση και δεν υπάρχει στατιστικά σημαντική σχέση μεταξύ προσδιοριστή και έκβασης
-

Έλεγχος Kruskal-Wallis

- Διερεύνηση της ύπαρξης σχέσης μεταξύ του εκπαιδευτικού επιπέδου (κατώτερο, μέσο και ανώτερο) και του εισοδήματος (ποσοτική μεταβλητή που δεν ακολουθεί την κανονική κατανομή στον πληθυσμό των ατόμων με κατώτερο, μέσο και ανώτερο εκπαιδευτικό επίπεδο)
-



Ranks			
	Εκπαιδευτικό επίπεδο	N	Mean Rank
Μηνιαίο εισόδημα	Κατώτερο	10	7,00
	Μέσο	10	16,25
	Ανώτερο	10	23,25
	Total	30	



Test Statistics^{a,b}

Μηνιαίο εισόδημα

Chi-Square	17,164
df	2
Asymp. Sig.	,000

a. Kruskal Wallis Test

b. Grouping Variable: Εκπαιδευτικό επίπεδο

Συντελεστής συσχέτισης

- Το είδος και το μέγεθος της γραμμικής σχέσης μεταξύ δυο ποσοτικών ή διατάξιμων μεταβλητών εκτιμάται με τους συντελεστές συσχέτισης
 - Στην περίπτωση αυτή, είναι απαραίτητη η δημιουργία και του αντίστοιχου διαγράμματος σημείων, με τη μεταβλητή που θεωρείται ως ανεξάρτητη (προσδιοριστής) να τοποθετείται στον οριζόντιο άξονα ή, αλλιώς, άξονα των x και τη μεταβλητή που θεωρείται ως εξαρτημένη (έκβαση) να τοποθετείται στον κάθετο άξονα ή, αλλιώς, άξονα των y
-

Συντελεστής συσχέτισης

- ❑ Συμβολίζεται με ρ όταν αφορά ολόκληρο τον πληθυσμό και με r όταν υπολογίζεται για ένα συγκεκριμένο δείγμα, αποτελώντας έτσι εκτίμηση για τον πληθυσμό από τον οποίο προέρχεται το δείγμα
 - ❑ Δεν έχει μονάδες μέτρησης και λαμβάνει τιμές από -1 έως +1
 - ❑ Οι τιμές -1 και +1 συμβαίνουν όταν υπάρχει μια τέλεια, αρνητική ή θετική αντιστοίχως, γραμμική σχέση μεταξύ των δυο μεταβλητών
 - ❑ Όταν λαμβάνει την τιμή 0 σημαίνει πως δεν υπάρχει γραμμική σχέση μεταξύ των δυο μεταβλητών, αλλά δεν αποκλείεται η ύπαρξη μιας άλλης μη γραμμικής σχέσης
-

Συντελεστής συσχέτισης

- Εάν το μέγεθος της μιας μεταβλητής τείνει να αυξάνεται όπως αυξάνεται το μέγεθος και της άλλης μεταβλητής, τότε υπάρχει **θετική συσχέτιση** (positive correlation) μεταξύ των δυο μεταβλητών και ο συντελεστής συσχέτισης είναι >0
 - Αντιθέτως, εάν το μέγεθος της μιας μεταβλητής τείνει να αυξάνεται όπως ελαττώνεται το μέγεθος της άλλης μεταβλητής, τότε υπάρχει **αρνητική συσχέτιση** (negative correlation) μεταξύ των δυο μεταβλητών και ο συντελεστής συσχέτισης είναι <0
-

Συντελεστής συσχέτισης

- Τιμές μεταξύ 0 έως 0,3 ή μεταξύ 0 έως -0,3 υποδηλώνουν μικρή συσχέτιση
 - Τιμές μεταξύ 0,31 έως 0,6 ή μεταξύ -0,31 έως -0,6 υποδηλώνουν μέτρια συσχέτιση
 - Τιμές μεταξύ 0,61 έως 0,8 ή μεταξύ -0,61 έως -0,8 υποδηλώνουν ισχυρή συσχέτιση
 - Τιμές μεταξύ 0,81 έως 1 ή μεταξύ -0,81 έως -1 υποδηλώνουν πολύ ισχυρή συσχέτιση
-

Μηδενική και εναλλακτική υπόθεση

- Μια μελέτη διερευνά τη σχέση μεταξύ του δείκτη μάζας σώματος και της συστολικής αρτηριακής πίεσης
 - H_0 : δεν υπάρχει γραμμική σχέση μεταξύ του δείκτη μάζας σώματος και της συστολικής αρτηριακής πίεσης
 H_0 : ο συντελεστής συσχέτισης ισούται με μηδέν
 - H_A : υπάρχει σχέση μεταξύ του δείκτη μάζας σώματος και της συστολικής αρτηριακής πίεσης
 H_A : ο συντελεστής συσχέτισης δεν ισούται με μηδέν
-

Συντελεστής συσχέτισης

- ❑ Χρησιμοποιώντας την κατανομή t και τους βαθμούς ελευθερίας της κατανομής υπολογίζεται η τιμή p
 - ❑ Εάν $p < 0,05$, τότε απορρίπτεται η μηδενική υπόθεση και υπάρχει στατιστικά σημαντική σχέση μεταξύ προσδιοριστή και έκβασης.
 - ❑ Εάν $p \geq 0,05$, τότε δεν απορρίπτεται η μηδενική υπόθεση και δεν υπάρχει στατιστικά σημαντική σχέση μεταξύ προσδιοριστή και έκβασης
-

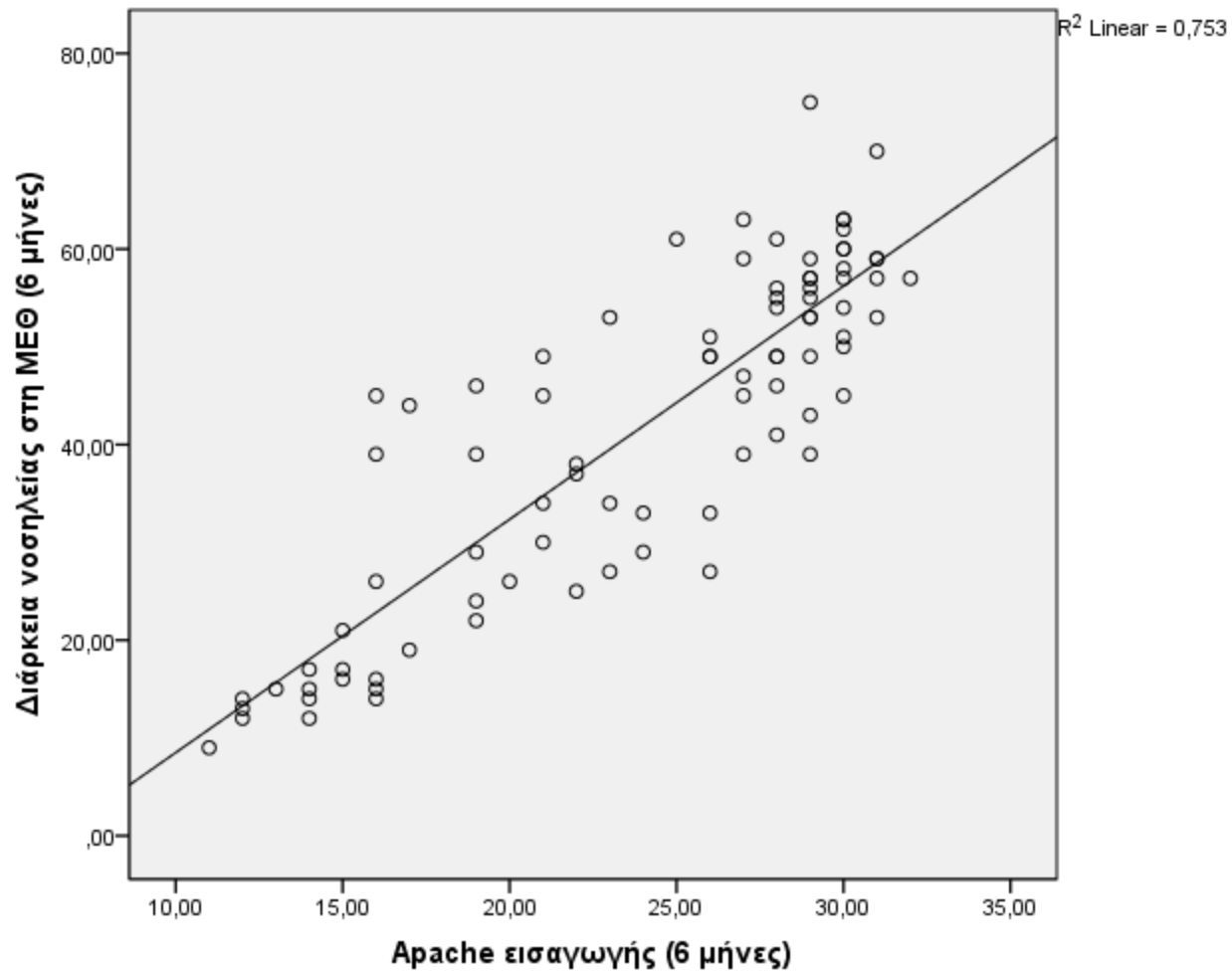
Συντελεστής συσχέτισης Pearson

- ❑ Σχέση μεταξύ δυο ποσοτικών μεταβλητών που ακολουθούν την κανονική κατανομή
 - ❑ Σχέση μεταξύ δείκτη μάζας σώματος και συστολικής αρτηριακής πίεσης
 - ❑ Δημιουργία του διαγράμματος σημείων
-

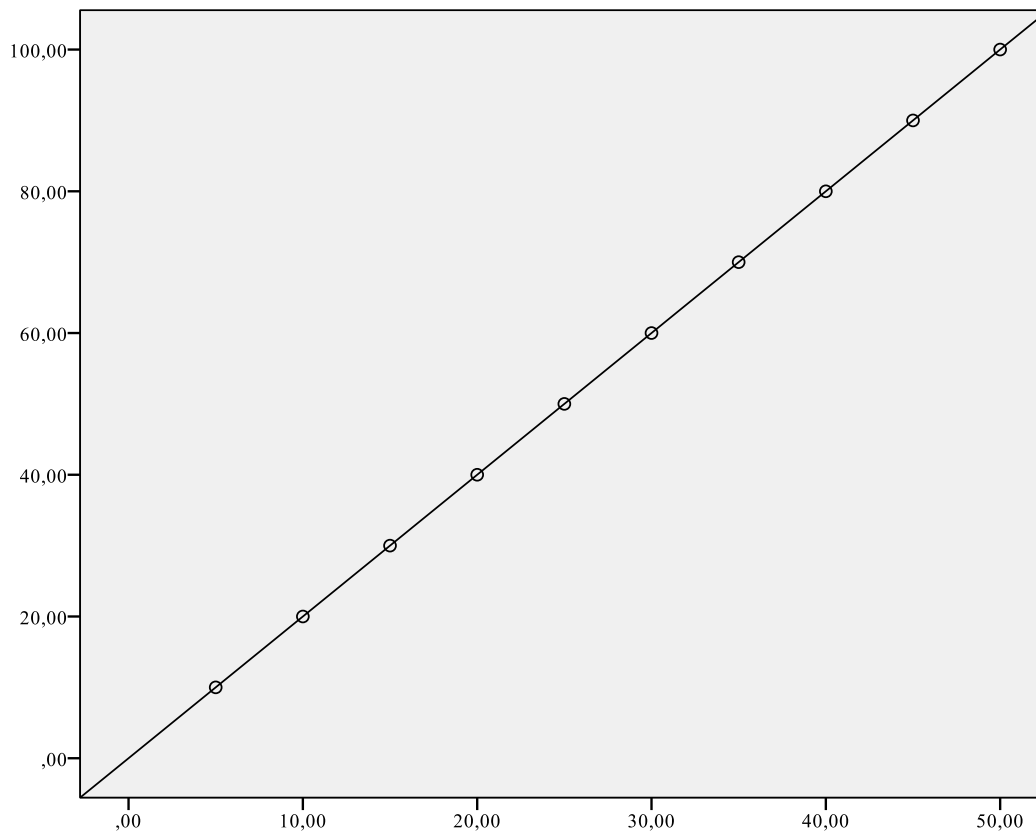
Συντελεστής συσχέτισης Spearman

- ❑ Σχέση μεταξύ δυο ποσοτικών μεταβλητών που η μια τουλάχιστον δεν ακολουθεί την κανονική κατανομή
 - ❑ Σχέση μεταξύ μιας ποσοτικής και μιας διατάξιμης μεταβλητής
 - ❑ Σχέση μεταξύ δυο διατάξιμων μεταβλητών
 - ❑ Σχέση μεταξύ δείκτη μάζας σώματος και συστολικής αρτηριακής πίεσης
 - ❑ Δημιουργία του διαγράμματος σημείων
-

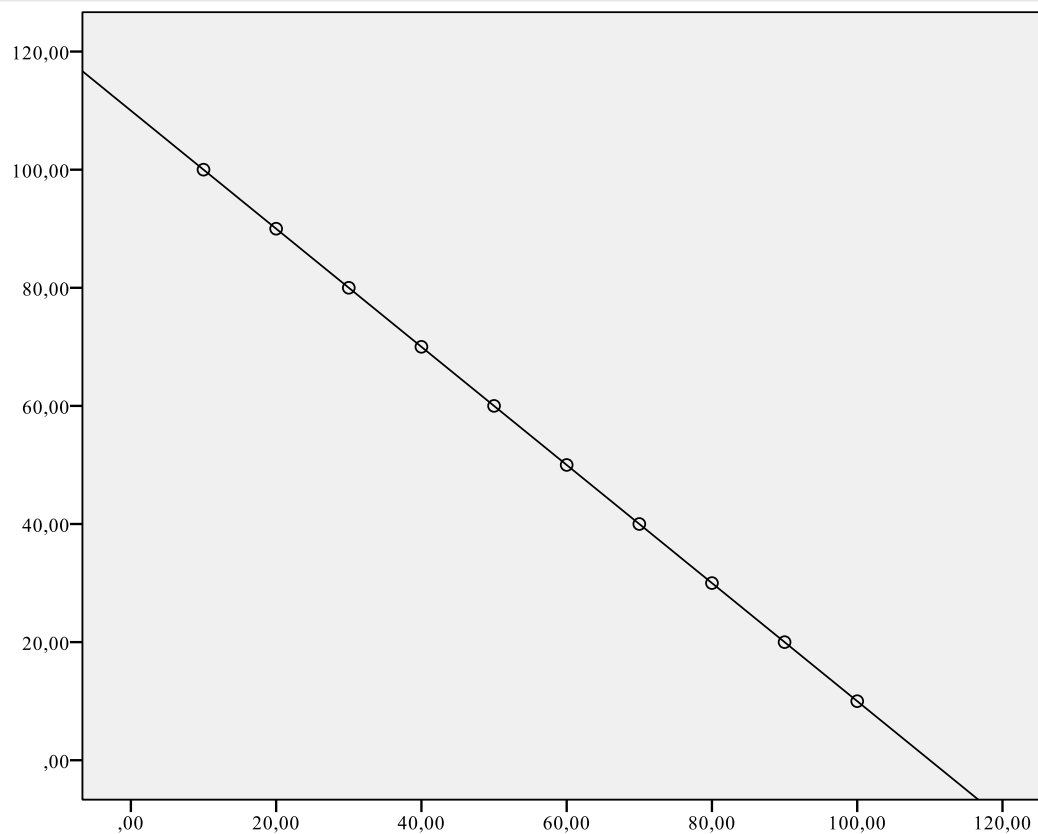
Διάγραμμα σημείων



Πλήρης θετική γραμμική συσχέτιση ($r=1$) μεταξύ δυο ποσοτικών μεταβλητών



Πλήρης αρνητική γραμμική συσχέτιση ($r=-1$) μεταξύ δυο ποσοτικών μεταβλητών



Συντελεστής συσχέτισης

- Διερεύνηση της ύπαρξης γραμμικής σχέσης μεταξύ της συστολικής αρτηριακής πίεσης, του δείκτη μάζας σώματος και της ηλικίας
-

Correlations				
		Συστολική αρτηριακή πίεση	Δείκτης μάζας σώματος	Ηλικία
Συστολική αρτηριακή πίεση	Pearson Correlation	1	,246**	,282**
	Sig. (2-tailed)		,000	,000
	N	223	223	223
Δείκτης μάζας σώματος	Pearson Correlation	,246**	1	,083
	Sig. (2-tailed)	,000		,217
	N	223	223	223
Ηλικία	Pearson Correlation	,282**	,083	1
	Sig. (2-tailed)	,000	,217	
	N	223	223	223

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