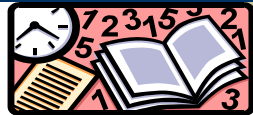


ΒΙΟΣΤΑΤΙΣΤΙΚΗ II

ΜΑΘΗΜΑ 5 ΕΡΓΑΣΤΗΡΙΟ 1 ΜΕΤΡΑ ΚΙΝΔΥΝΟΥ & ΣΥΜΠΕΡΑΣΜΑΤΟΛΟΓΙΑ ΜΕ ΤΗΝ ΧΡΗΣΗ SPSS

I. Μαρκάκη

Από την θεωρία



στην πράξη....



2

Στόχος μαθήματος:

- ♦ Περιγραφή δίτιμων κατηγορικών δεδομένων
- ♦ Υπολογισμός βασικών μέτρων κινδύνου (αποδιδόμενος κίνδυνος, σχετικός κίνδυνος, λόγος σχετικών πιθανοτήτων)
- ♦ Έλεγχος ανεξαρτησίας των δύο μεταβλητών (χ^2 του Pearson, έλεγχος μέγιστης πιθανοφάνειας, ακριβής έλεγχος ανεξαρτησίας του Fisher)

ΧΡΗΣΗ ΔΥΟ ΠΡΟΓΡΑΜΜΑΤΩΝ: SPSS & StatXact

3

Παράδειγμα 1: μελέτη ασθενών-μαρτύρων

Μυοκαρδιακή Ανεπάρκεια και Αντισύλληψη (Μελέτη Μαρτύρων-Ασθενών)

Τα δεδομένα του 2x2 Πίνακα που ακολουθούν, προέρχονται από την ερευνητική δουλειά των Mann *et al* (1975, Brit.J.Med.). Στη έρευνα αυτή 58 γυναίκες κάτω των 45 χρονών με μυοκαρδιακή ανεπάρκεια εξετάστηκαν ως προς τη χρήση ή όχι αντισυλληπτικού χαπιού. Το δείγμα προερχόταν από δύο νοσοκομεία της Αγγλίας και της Ουαλίας ενώ από τα ίδια νοσοκομεία επιλέχθηκαν 166 μάρτυρες που δεν είχαν την νόσο της μυοκαρδιακής ανεπάρκειας, και ερωτήθηκαν ως προς τη χρήση ή όχι αντισυλληπτικού χαπιού.

(Δεδομένα: Categorical Data Analysis, Agresti (1990) σελ. 11-13).

4

Παράδειγμα 1 (συνέχεια): μελέτη ασθενών-μαρτύρων

Πίνακας 1: Πίνακας διπλής εισόδου παραδείγματος 1: Μυοκαρδιακή Ανεπάρκεια και Αντισυλλληπτικό χάπι (Mann *et al.*, 1975)

X: Αντισυλ. Χάπι	Y: Μυοκαρδιακή Ανεπάρκεια		Περ. Κατ. X
	1: <i>Ναι</i>	2: <i>Όχι</i>	
1: <i>Ναι</i>	23	34	57
2: <i>Όχι</i>	35	132	167
Περ. Κατ Y	58	166	224

5

1. Εισαγωγή δεδομένων στο SPSS

- Ο πρώτος τρόπος εισαγωγής των στοιχείων στο SPSS είναι εισάγοντας για κάθε άτομο τις τιμές των δύο μεταβλητών

6

2. «Προετοιμασία» δεδομένων

Απαραίτητα / Χρήσιμα διευκρινιστικά στοιχεία:

- Όνομα μεταβλητής
- Τύπος μεταβλητής
- Μήκος-δεκαδικά
- «Ταμπέλα» μεταβλητής & κωδικοποίησή της (SOS!)

Name	Type	Width	Decimals	Label	Values
1 contrac	Numeric	8	2	«Ταμπέλα» μεταβλητής	None
2 myocard	Numeric	8	2		None
3 counts	Numeric	8	2		None

8

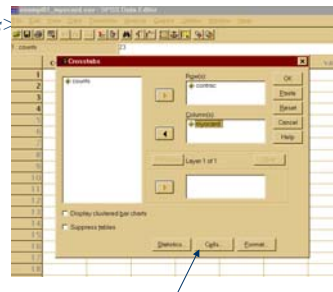
- Ο δεύτερος τρόπος εισαγωγής, εφόσον έχουμε στη διάθεσή μας τον πίνακα 2x2, είναι μέσω της χρήσης του “weight cases” θέτοντας την μεταβλητή ‘counts’ ως ‘βάρος’

7

3α. Ανάλυση πίνακα διπλής εισόδου

Analyze > Descriptive Stats > Crosstabs

1. Χρήση Αντισυλ. ως γραμμές
2. Νόσος ως στήλες



9

3γ. Δεσμευμένη συνάρτηση πιθανότητας

- ♦ **Στόχος:** Σύγκριση της κατανομής της νόσου ανάμεσα στις γυναίκες που χρησιμοποιούν αντισύλληψη, με την αντίστοιχη κατανομή ανάμεσα στις γυναίκες που δεν χρησιμοποιούν αντισύλληψη. Αλλιώς: επιθυμούμε να συγκρίνουμε τις πιθανότητες (ή ποσοστά) εμφάνισης της νόσου μεταξύ των δύο ομάδων

CONTRAC * MYOCARD Crosstabulation

		MYOCARD		Total
		1.00 yes	2.00 no	
CONTRAC	1.00 yes	Count 23	34	57
		% within CONTRAC	40.4%	59.6%
2.00 no	Count	35	132	167
	% within CONTRAC	21.0%	79.0%	100.0%
Total	Count	58	166	224
	% within CONTRAC	25.9%	74.1%	100.0%

11

3β. Από κοινού συνάρτηση πιθανότητας

Από το cells επιλέγουμε την εμφάνιση «percentages: Total»

CONTRAC * MYOCARD Crosstabulation

		MYOCARD		Total
		1.00	2.00	
CONTRAC	1.00	Count 23	34	57
		% of Total 10.3%	15.2%	25.4%
2.00	Count	35	132	167
	% of Total	15.6%	59.9%	74.6%
Total	Count	58	166	224
	% of Total	25.9%	74.1%	100.0%

5

3δ. Δεσμευμένη συνάρτηση πιθανότητας

- ♦ **Όμως:** Στο παραπάνω πρόβλημα η κατανομή της μεταβλητής απόκρισης (μυοκαρδιακής ανεπάρκειας) είναι σταθερή (προκαθορισμένη) από το σχεδιασμό του πειράματος – μελέτης (αναδρομική μελέτη ή αλλιώς «μελέτη ασθενών-μαρτύρων»). Αυτό όμως που μπορούμε να ελέγξουμε, χωρίς να μπορεί να μας μεταφέρει όλη την επιθυμητή πληροφορία, είναι οι κατανομές $P(X|Y)$, εφόσον το X είναι όντως τυχαίο.

12

6

3ε. Δεσμευμένη συνάρτηση πιθανότητας

Στόχος, επομένως, στις μελέτες ασθενών-μαρτύρων είναι ο έλεγχος της κατανομής της X (δηλ. του παράγοντα), με δεδομένη την μεταβλ. απόκρισης Y (δηλ. τη νόσο) $P(X|Y)$.

13

3ζ. Δεσμευμένη συνάρτηση πιθανότητας

Εν συνεχεία, με την χρήση του θεωρήματος του Bayes μπορούμε να υπολογίσουμε τις δεσμευμένες κατανομές $P(Y|X)$ που μας ενδιαφέρουν, με απαραίτητη προϋπόθεση να γνωρίζουμε το ποσοστό της ασθένειας στον πληθυσμό (δείκτη επιπολασμού).

15

3στ. Δεσμευμένη συνάρτηση πιθανότητας

Επιθυμούμε δηλαδή να συγκρίνουμε την κατανομή της χρήσης χαπιού στις γυναίκες που νοσούν, με την αντίστοιχη κατανομή στις γυναίκες που δεν νοσούν

Ο συγκεκριμένος πίνακας συνάφειας που μας ενδιαφέρει, επομένως, είναι αυτός που παίρνουμε από το SPSS με την επιλογή των ποσοστών ως προς την στήλη (Y) (δεσμευμένες κατανομές της χρήσης χαπιού ως προς την εμφάνιση της νόσου)

CONTRAC * MYOCARD Crosstabulation

		MYOCARD		Total
		1.00 yes	2.00 no	
CONTRAC 1.00 yes	Count	23	34	57
	% within MYOCA	(39.7%)	(20.5%)	25.4%
2.00 no	Count	35	132	167
	% within MYOCA	60.3%	79.5%	74.6%
Total	Count	58	166	224
	% within MYOCA	100.0%	100.0%	100.0%

14

4α. Εκτίμηση ΛΣΠ και σχετικού κινδύνου

- ♦ Από το statistics επιλέγουμε το Risk



16

4β. Εκτίμηση ΛΣΠ –ερμηνεία 1

	Risk Estimate		
	Value	95% Confidence Interval	
		Lower	Upper
Odds Ratio for CONTRAC (1.00 yes / 2.00 no)	2.551	1.336	4.873
For cohort MYOCARD = 1.00 yes	1.925	1.250	2.965
For cohort MYOCARD = 2.00 no	.755	.601	.947
N of Valid Cases	224		

OR= odds(X=1)/odds(X=2) = 2.55

Ερμηνεία 1: η σχετική πιθανότητα μια γυναίκα να έχει την νόσο (Y=1) όταν παίρνει το χάπι (X=1) είναι 2.55 την αντίστοιχη πιθανότητα μιας γυναίκας που δεν λαμβάνει το χάπι (X=2)

17

4γ. Εκτίμηση ΛΣΠ –ερμηνεία 2

Ερμηνεία 2: η σχετική πιθαν. μη εμφάνιση της νόσου σε μία γυναίκα που δεν λαμβάνει το χάπι είναι 2.55 την αντίστοιχη πιθανότητα μιας γυναίκας που λαμβάνει το χάπι

	Risk Estimate		
	Value	95% Confidence Interval	
		Lower	Upper
Odds Ratio for CONTRAC (1.00 / 1.00 yes)	2.551	1.336	4.873
For cohort MYOCARD = 1.00	1.325	1.056	1.663
For cohort MYOCARD = 1.00 yes	.519	.337	.800
N of Valid Cases	224		

19

Παρατήρηση: Η ταυτόχρονη αντιστροφή των κατηγοριών των δύο μεταβλητών δεν αλλάζει το αποτέλεσμα

«Πείραμα»: Αλλάξτε την κωδικοποίηση της κατηγορίας «όχι» από 2 σε 0 και παρακολουθείστε τι θα γίνει. Στη συνέχεια υπολογίστε τον κίνδυνο (Risk)

		MYOCARD		Total	
		00	1.00 yes		
CONTRAC	00	Count	132	35	167
		% within MYOCARD	79.5%	60.3%	74.8%
1.00 yes	Count	34	23	57	
	% within MYOCARD	20.5%	39.7%	25.4%	
Total	Count	166	58	224	
	% within MYOCARD	100.0%	100.0%	100.0%	

18

4δ. Συμπέρασμα για το ΛΣΠ

Οποιαδήποτε ερμηνεία και αν ακολουθήσουμε συμπεραίνουμε ότι η χρήση αντισυλληπτικού χαπιού αυξάνει την (σχετική) πιθανότητα (ή τον κίνδυνο εμφάνισης) μυοκαρδιακής ανεπάρκειας, δηλαδή η χρήση του χαπιού αποτελεί παράγοντα κινδύνου για την νόσο.

Παρατήρηση: 95% ΔΕ του OR – στατιστικά σημαντικό (δεν συμπεριλαμβάνεται η μονάδα)

20

4ε. Εκτίμηση σχετικού κινδύνου

	Value	95% Confidence Interval	
		Lower	Upper
Odds Ratio for CONTRAC (1.00 yes / 2.00 no)	2.551	1.336	4.873
For cohort MYOCARD = 1.00 yes	1.925	1.250	2.965
For cohort MYOCARD = 2.00 no	755	601	947
Ni of Valid Cases	224		

$$RR(\text{myocard}=1) = \frac{23/57}{35/167} = \frac{0.404}{0.210} = 1.93 \text{ (διαφάνεια 3γ.)}$$

Ερμηνεία: Η πιθανότητα μυοκαρδιακής ανεπάρκειας για μια γυναίκα που χρησιμοποιεί αντισυλληπτικό χάπι είναι 93% υψηλότερη από την αντίστοιχη πιθανότητα μιας γυναίκας που δεν χρησιμοποιεί αυτό το είδος αντισύλληψης

21

4στ. Εκτίμηση σχετικού κινδύνου

$$RR(\text{myocard}=2) = \frac{34/57}{132/167} = \frac{0.596}{0.790} = 0.759 \text{ (διαφάνεια 3γ.)}$$

Ερμηνεία: Η πιθανότητα να μην εμφανίσει την νόσο μυοκαρδιακής ανεπάρκειας μια γυναίκα που χρησιμοποιεί αντισυλληπτικό χάπι είναι 24% μικρότερη από την αντίστοιχη πιθανότητα μιας γυναίκας που δεν χρησιμοποιεί αυτό το είδος αντισύλληψης

22

4ζ. Εκτίμηση σχετικού κινδύνου

Τι κάναμε λάθος στην παραπάνω ανάλυση???

- ♦ Ανάλυση *σχετικού κινδύνου* → ΛΑΘΟΣ

Μελέτη ασθενών-μαρτύρων: η κατανομή της νόσου (Y) είναι προκαθορισμένη.

Καλή προσέγγιση του σχετικού κινδύνου: εκτίμηση του *ΑΣΠ (OR)*, εάν δείκτης επιπολασμού της νόσου στο γενικό πληθυσμό μικρός

23

5. Έλεγχος ανεξαρτησίας χ^2 του Pearson

- ♦ Έλεγχος της υπόθεσης:

H_0 «Ανεξαρτησία μεταξύ *Νόσου (Y)* και *Αντισυλληπτικού χαπιού (X)*»,

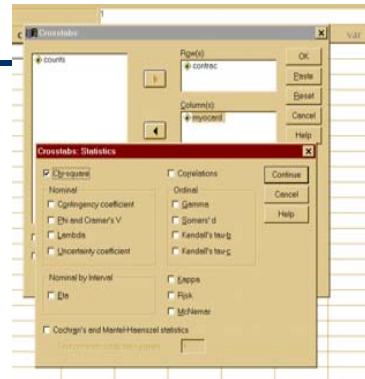
Vs.

H_1 «*Νόσος (Y)* και *Αντισυλληπτικό χάπι (X)* εξαρτημένες μεταβλητές»

24

5α. Έλεγχος Ανεξαρτησίας χ^2 του Pearson (στο SPSS)

- ◆ Analyze > Descriptive Stats > Crosstabs
- ◆ Και μέσα από το: Statistics → Chi-square



25

5γ. Επιπλέον έλεγχοι ανεξαρτησίας

Στον ίδιο πίνακα βλέπουμε και:

- a) χ^2 με διόρθωση του Yates
- b) Έλεγχο μέγιστης πιθανοφάνειας (LRT) [προϋπόθεση $n/(I \cdot J) \geq 5$]

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	8.329 ^a	1	.004		
Continuity Correction ^b	7.349	1	.007		
Likelihood Ratio	7.866	1	.005	.005	.004
Fisher's Exact Test					
Linear-by-Linear Association	8.292	1	.004		
N of Valid Cases	224				

a. Computed only for a 2x2 table
b. 0 cells (.0%) have expected count less than 5. The minimum expected count is 14.76.

27

5β. Αποτελέσματα ελέγχου ανεξαρτησίας

- ◆ P-value < 0.05 => απορρίπτουμε την υπόθεση της ανεξαρτησίας των δύο μεταβλητών (προϋπόθεση για Pearson's χ^2 : αναμενόμενες τιμές των κελιών ≥ 5)

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	8.329 ^a	1	.004		
Continuity Correction ^b	7.349	1	.007		
Likelihood Ratio	7.866	1	.005	.005	.004
Fisher's Exact Test					
Linear-by-Linear Association	8.292	1	.004		
N of Valid Cases	224				

a. Computed only for a 2x2 table
b. 0 cells (.0%) have expected count less than 5. The minimum expected count is 14.76.

26

5γ. Επιπλέον έλεγχοι ανεξαρτησίας

...Στον ίδιο πίνακα βλέπουμε και:

- c) Ακριβή έλεγχο ανεξαρτησίας του Fisher (όταν έχουμε αναμενόμενες τιμές κελιών < 5)

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	8.329 ^a	1	.004		
Continuity Correction ^b	7.349	1	.007		
Likelihood Ratio	7.866	1	.005	.005	.004
Fisher's Exact Test				.004	
Linear-by-Linear Association	8.292	1	.004		
N of Valid Cases	224				

a. Computed only for a 2x2 table
b. 0 cells (.0%) have expected count less than 5. The minimum expected count is 14.76.

28

6. Συμπεράσματα Παραδείγματος 1

- ♦ Από το OR: η χρήση αντισυλληπτικού χαπιού αυξάνει, στατιστικά σημαντικά, τον κίνδυνο εμφάνισης μυοκαρδιακής ανεπάρκειας κατά 155%.
- ♦ Από τους ελέγχους ανεξαρτησίας: η χρήση αντισυλληπτικού χαπιού συσχετίζεται στατιστικά σημαντικά με την μυοκαρδιακή ανεπάρκεια

29

Ανάλυση του παραδείγματος 1 με τη χρήση του StatXact

(αλλαγή αρχείου)

Παράδειγμα 2: προοπτική μελέτη

Επίδραση καπνίσματος στην επιβίωση ασθενών με καρδιακό επεισόδιο

Σε μια προοπτική μελέτη εξετάστηκαν 368 άνδρες, καπνιστές, ηλικίας κάτω των 60 ετών οι οποίοι έπαθαν καρδιακή προσβολή και επιβίωσαν. (Daly et al 1991, p 185)

Μετά από 2 έτη εξετάστηκαν πόσοι από αυτούς είχαν επιβιώσει και χωρίστηκαν ανάλογα με το εάν είχαν σταματήσει το τσιγάρο ή όχι. Εδώ μας ενδιαφέρει να εξετάσουμε αν το σταμάτημα του καπνίσματος (X) είχε ευνοϊκή επίδραση στην επιβίωση μετά από δύο έτη (Y). Τα δεδομένα δίνονται στον 2x2 Πίνακα που ακολουθεί

31

Παράδειγμα 2: δεδομένα

Πίνακας διτλής εισόδου παραδείγματος 2: Κάπνισμα και επιβίωση σε ασθενείς καρδιακής ανεπάρκειας (Daly et al., 1983)

X: Συνέχιση καπνίσματος	Y: Επιβίωση σε δύο χρόνια		
	1: Απεβίωσε	2: Εν ζωή	Περ. Κατ. X
1: Ναι	19	135	154
2: Όχι	15	199	214
Περ. Κατ Y	34	334	368

32

1/ Εισαγωγή στοιχείων στο SPSS

1. Δεδομένα
2. Ονομασία μεταβλητών
3. Ονομασία κατηγοριών

33

2/ Προβολή 2x2 πίνακα

Επιλογή σωστών ποσοστών (υπενθύμηση: το είδος της μελέτης)

SMOKE Continuation of smoking * SURVIVAL Survival in 2 yrs Crosstabulation

		SURVIVAL Survival in 2 yrs		Total
		1 dead	2 alive	
SMOKE Continuation of smoking	1 yes	Count 19	Count 135	154
		% within SMOKE Continuation of smoking 12.3%	% within SMOKE Continuation of smoking 87.7%	100.0%
	2 no	Count 15	Count 199	214
		% within SMOKE Continuation of smoking 7.0%	% within SMOKE Continuation of smoking 93.0%	100.0%
Total		Count 34	Count 334	368
		% within SMOKE Continuation of smoking 9.2%	% within SMOKE Continuation of smoking 90.8%	100.0%

35

Δεδομένα στο SPSS

	smoke	survival	count	var	var
1	1	1	19		
2	1	2	135		
3	2	1	15		
4	2	2	199		
5					
6					
7					

34

3α/ Στατιστική Ανάλυση

α. Αποδιδόμενος ή Αποδοτέος κίνδυνος

(διαφορά μεταξύ των ρυθμών επίπτωσης -ή δεικτών θνησιμότητας- των ομάδων με άτομα εκτεθειμένα και μη εκτεθειμένα σε ένα παράγοντα κινδύνου)

Ποσοστό εμφάνισης της νόσου όταν *υπάρχει* έκθεση στον κίνδυνο – ποσοστό εμφάνισης της νόσου όταν *δεν υπάρχει* έκθεση στον κίνδυνο = $12.3\% - 7\% = 5.3\%$

ΣΥΜΠΕΡΑΣΜΑ: ΑΣΘΕΝΕΙΣ ΠΟΥ ΣΥΝΕΧΙΖΟΥΝ ΝΑ ΚΑΠΝΙΖΟΥΝ ΜΕΤΑ ΤΟ ΚΑΡΔΙΑΚΟ ΕΠΕΙΣΟΔΙΟ ΠΑΡΟΥΣΙΑΖΟΥΝ ΜΕΓΑΛΥΤΕΡΗ ΠΙΘΑΝΟΤΗΤΑ ΘΑΝΑΤΟΥ ΣΕ ΣΧΕΣΗ ΜΕ ΟΣΟΥΣ ΔΙΕΚΟΨΑΝ ΤΟ ΚΑΠΝΙΣΜΑ, ΚΑΤΑ 5.3 ΠΟΣΟΣΤΙΑΙΕΣ ΜΟΝΑΔΕΣ.

36

3β/ Στατιστική Ανάλυση

β. Ποσοστιαίος αποδοτέος κίνδυνος

[«Πιθανότητα νόσου/έκθεση στον παράγοντα» - «πιθαν νόσου/ όχι έκθεση στον παράγοντα»] / «Πιθανότητα νόσου/έκθεση στον παράγοντα»]

Εδώ: $(0.053/0.123)*100=43.1\%$

ΣΥΜΠΕΡΑΣΜΑ: ΤΟ 43% ΤΟΥ ΚΙΝΔΥΝΟΥ ΘΑΝΑΤΟΥ ΠΟΥ ΔΙΑΤΡΕΧΕΙ ΕΝΑΣ ΑΣΘΕΝΗΣ ΠΟΥ ΣΥΝΕΧΙΖΕΙ ΤΟ ΚΑΠΝΙΣΜΑ, ΜΠΟΡΕΙ ΝΑ ΑΠΟΔΟΘΕΙ ΣΤΗΝ ΣΥΝΕΧΙΣΗ ΤΟΥ ΚΑΠΝΙΣΜΑΤΟΣ.

ΑΛΛΙΩΣ, ΤΟ 43% ΤΩΝ ΘΑΝΑΤΩΝ (2 ΕΤΗ ΜΕΤΑ ΤΟ 1ο ΕΠΕΙΣΟΔΙΟ) ΘΕΩΝ ΣΥΝΕΧΙΖΑΝ ΝΑ ΚΑΠΝΙΖΟΥΝ (δηλ. $19*43\%=8$ άτομα) ΘΑ ΜΠΟΡΟΥΣΕ ΝΑ ΕΙΧΕ ΑΠΟΦΕΥΓΘΕΙ ΑΝ ΕΙΧΑΝ ΣΤΑΜΑΤΗΣΕΙ ΤΟ ΚΑΠΝΙΣΜΑ

37

3γ/ Στατιστική Ανάλυση (συνέχεια)

- ♦ Από τα δύο RR που υπολογίζει το SPSS μας ενδιαφέρει ο σχετικός κίνδυνος της νόσου (δηλαδή το πρώτο RR)
- ♦ Όταν η ασθένεια είναι σπάνια, αναμένουμε το 2ο RR να είναι κοντά στη μονάδα, καθώς επίσης οι τιμές του σχετικού κινδύνου της νόσου (το 1ο RR) και του ΛΣΠ να είναι πολύ κοντά.

39

3γ/ Στατιστική Ανάλυση

γ. Σχετικός Κίνδυνος

$RR = 0.1234 / 0.0701 = 1.760$

Κίνδυνος θανάτου για ασθενείς που συνέχισαν το κάπνισμα είναι 1.76 φορές (ή 76% μεγαλύτερος από) τον ίδιο κίνδυνο ασθενών που σταμάτησαν το κάπνισμα.

	Risk Estimate		
	Value	95% Confidence Interval	
		Lower	Upper
Odds Ratio for SMOKE Continuation of smoking (1 yes / 2 no)	1.867	.917	3.803
For cohort SURVIVAL Survival in 2 yrs = 1 dead	1.760	.924	3.353
For cohort SURVIVAL Survival in 2 yrs = 2 alive	.943	.879	1.011
N of Valid Cases	368		

38

3δ/ Στατιστική Ανάλυση

δ. Λόγος Σχετικών Πιθανοτήτων

$$OR = \frac{19/135}{15/199} = (19 \times 199) / (15 \times 135) = 1.867$$

Η σχετική πιθανότητα θανάτου για τους ασθενείς που συνέχισαν το κάπνισμα είναι 86% μεγαλύτερη από την πιθανότητα θανάτου ασθενών που σταμάτησαν το κάπνισμα

	Risk Estimate		
	Value	95% Confidence Interval	
		Lower	Upper
Odds Ratio for SMOKE Continuation of smoking (1 yes / 2 no)	1.867	.917	3.803
For cohort SURVIVAL Survival in 2 yrs = 1 dead	1.760	.924	3.353
For cohort SURVIVAL Survival in 2 yrs = 2 alive	.943	.879	1.011
N of Valid Cases	368		

40

3ε/ Έλεγχοι ανεξαρτησίας

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	3.032 ^a	1	.082		
Continuity Correction ^b	2.430	1	.119		
Likelihood Ratio	2.985	1	.084		
Fisher's Exact Test				.100	.060
Linear-by-Linear Association	3.024	1	.082		
N of Valid Cases	368				

a. Computed only for a 2x2 table

b. 0 cells (.0%) have expected count less than 5. The minimum expected count is 14.23.

Μη στατιστικά σημαντικό p-value, άρα δεν απορρίπτουμε ανεξαρτησία μεταβλητών

41

ΑΣΚΗΣΗ ΓΙΑ ΤΟ ΣΠΙΤΙ

- ♦ Με βάση την Ιατρική βιβλιογραφία υπάρχουν ενδείξεις ότι ο καρκίνος του μαστού στις γυναίκες σχετίζεται με την ηλικία κατά την πρώτη γέννα. Συγκεκριμένα έχει αναφερθεί ότι ο κίνδυνος της νόσου αυξάνει με την αύξηση της ηλικίας της εγκύου κατά την 1η κύηση. Το 1970 πραγματοποιήθηκε διεθνής πολυκεντρική μελέτη (MacMachon et al, 1970) (ΗΠΑ, Ελλάδα, Γιουγκοσλαβία, Βραζιλία, Ταϊβάν, Ιαπωνία).

43

4/Συμπεράσματα Παραδείγματος 2

- ♦ **Από το OR:** Η συνέχιση του καπνίσματος μετά από την επιβίωση από καρδιακό επεισόδιο, αυξάνει τον κίνδυνο θανάτου τα επόμενα δύο χρόνια κατά 86%, χωρίς όμως να είναι στατιστικά σημαντικός
- ♦ **Από τον έλεγχο ανεξαρτησίας:** η πιθανότητα επιβίωσης μετά από δύο έτη δεν σχετίζεται στατιστικά σημαντικά με την συνέχιση του καπνίσματος μετά την επιβίωση από καρδιακό επεισόδιο

42

Άσκηση για το σπίτι (συνέχεια)

Περιπτώσεις καρκινοπαθών επιλέχθηκαν από τα συγκεκριμένα νοσοκομεία στις παραπάνω χώρες. Οι μάρτυρες επιλέχθηκαν από τα ίδια νοσοκομεία έτσι ώστε να είναι συγκρίσιμης ηλικίας και να μην έχουν καρκίνο μαστού ή άλλα σχετικά νοσήματα. Όλες οι γυναίκες χωρίστηκαν σε δύο ομάδες κινδύνου ανάλογα με την ηλικία τους κατά την πρώτη κύηση (E: $X \geq 30$, \bar{E} : $X \leq 29$). Λεπτομέρειες σχετικά με το παράδειγμα βλ Rosner 1994, σελ 346. Τα δεδομένα βρίσκονται στον ακόλουθο πίνακα.

44

Άσκηση για το σπίτι (συνέχεια)

Καρκίνος του μαστού	Ηλικία εγκόου κατά την 1η κήση		
	(1) ≥ 30	(2) ≤ 29	Περ. Κατ. Υ
1: Ασθενής	683	2537	3220
2: Μάρτυρας	1498	8747	10245
Περ. Κατ X	2181	11284	13465

45

Άσκηση για το σπίτι (συνέχεια)

1. Τι μελέτη είναι και γιατί;
2. Να υπολογίσετε όλα τα μέτρα κινδύνου μόνοι σας και να τα ερμηνεύσετε. Ποια από αυτά είναι κατάλληλα για τη συγκεκριμένη μελέτη και γιατί;
3. Χρησιμοποιείστε επιπλέον για την ανάλυσή σας το SPSS και το StatXact. Εκτυπώστε τα αποτελέσματά σας. Διατυπώστε τα συμπεράσματά σας.

46