

ΚΡΗΤΙΚΟΥ ΚΑΤΕΡΙΝΑ Α.Μ.143

ΝΙΚΑΚΗ ΚΑΤΕΡΙΝΑ Α.Μ.154

McNemar test

Εργαστήριο σε θέματα στατιστικής

McNemar Test

Απαραμετρικό κριτήριο για 2 συσχετισμένα δείγματα

- ✓ Χρησιμοποιείται για να ελέγξουμε κατά πόσο έχουμε σημαντικές αλλαγές των χαρακτηριστικών ενός πληθυσμού σε 2 χρονικές περιόδους με τη μεσολάβηση κάποιου γεγονότος.
- ✓ Επειδή τα ίδια άτομα χρησιμοποιούνται σαν δείγμα μαρτύρων, θα έχουμε συσχετισμένες παρατηρήσεις.
- ✓ Το κριτήριο αυτό μπορεί να χρησιμοποιηθεί και με κατηγορικά δεδομένα.

Παράδειγμα:

Ρωτώνται n άτομα να δηλώσουν ποιον από τους 2 υποψηφίους δημάρχους A,B προτιμούν. Μετά την πραγματοποίηση τηλεοπτικής αναμέτρησης τα ίδια άτομα ξαναρωτήθηκαν για την προτίμησή τους. Στον πίνακα απεικονίζεται ο αριθμός των ατόμων που άλλαξαν προτίμηση από τον έναν υποψήφιο στον άλλο ή παρέμειναν σταθεροί.

		ΜΕΤΑ	
		A(+)	B(-)
ΠΡΙΝ	A(+)	b	a
	B(-)	d	c

Όπου:

- a: άτομα που άλλαξαν προτίμηση από τον A υποψήφιο στον B
- d: άτομα που άλλαξαν προτίμηση από τον B στον A
- b: άτομα που παρέμειναν σταθεροί στον A υποψήφιο
- c: άτομα που παρέμειναν σταθεροί στον B υποψήφιο

***Η υπόθεση H_0 είναι:** “Δεν υπάρχουν σημαντικές αλλαγές προτίμησης των 2 υποψηφίων μετά την τηλεοπτική αναμέτρηση”.*

- ✓ (a+d): συνολικός αριθμός ατόμων που άλλαξαν προτίμηση
- ✓ (a+d)/2: αναμενόμενος αριθμός ατόμων που άλλαξαν προτίμηση από τον A στον B.
- ✓ (a+d)/2: αναμενόμενος αριθμός ατόμων που άλλαξαν προτίμηση από τον B στον A.

Επειδή ο αριθμός των ατόμων που παρέμειναν σταθεροί στην προτίμηση τους δεν μας ενδιαφέρει, το κριτήριο που θα χρησιμοποιήσουμε είναι το χ^2 με 2 κελιά τα: (A,B) και (B,A) με

O_i : παρατηρηθείσες τιμές

E_i : αναμενόμενες τιμές

Για το κελί (A,B): $O_1 = a$, $E_1 = (a+d)/2$

Για το κελί (B,A): $O_2 = d$, $E_2 = (a+d)/2$

Συνεπώς το κριτήριο είναι:

$$\chi^2 = \frac{\sum_{i=1}^2 (O_i - E_i)^2}{E_i} = \frac{\left(a - \frac{a+d}{2}\right)^2}{\frac{a+d}{2}} + \frac{\left(d - \frac{a+d}{2}\right)^2}{\frac{a+d}{2}} = \frac{(a-d)^2}{(a+d)}$$

► Το κριτήριο χ^2 ακολουθεί την χ^2_1 .

- Αν ισχύει $\chi^2 > \chi^2_{1,\alpha}$ → απόρριψη της H_0
- Διαφορετικά → αποδοχή της H_0

ΠΑΡΑΤΗΡΗΣΗ:

Στην περίπτωση που όλες οι αναμενόμενες συχνότητες είναι πολύ μικρές, τότε στη ποσότητα χ^2 κάνουμε τη διόρθωση της συνέχειας του Yates και υπολογίζουμε την ποσότητα χ'^2 που ακολουθεί κι αυτή την χ^2_1 και παρέχεται από τον τύπο:

$$\chi'^2 = \frac{(|a - d| - 1)^2}{(a + d)}$$

ΕΦΑΡΜΟΓΗ: Ένα δείγμα 15 ατόμων ρωτήθηκε αν ήταν υπέρ ή κατά ενός υποψηφίου πριν και μετά την τηλεοπτική του εμφάνιση. Τα αποτελέσματα αυτής της δημοσκοπήσης δίνονται στον παρακάτω πίνακα:

Άτομο	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Πριν	+	+	-	-	+	+	-	-	-	-	+	-	+	-	+
Μετά	-	+	+	+	-	-	-	-	+	+	+	-	+	-	+

(+:υπέρ) (-:κατά)

Ποιο συμπέρασμα βγαίνει;
Υπήρξε μεταστροφή του εκλογικού σώματος;

Λύση:

Εφαρμόζοντας το κριτήριο McNemar σχηματίζουμε τον ακόλουθο 2*2 πίνακα.

A=3 B=4 C=4
D=4.Υπολογίζουμε την ποσότητα:

$$X^2 = \frac{(A - D)^2}{A + D} = 0,143$$

Παρατηρούμε ότι:
 $X^2=0,143 < 3,84 = X^2_{1,0,05}$.

Αυτό σημαίνει ότι σε επίπεδο σημαντικότητας $\alpha=0,05$ γίνεται δεκτή η H_0 , δηλ. δεν υπάρχει μεταστροφή του εκλογικού σώματος μετά την τηλεοπτική εμφάνιση του υποψηφίου.

		Μετά	
		+	-
Πριν	+	4	3
	-	4	4

McNemar.test

ΧΡΗΣΗ: `mcnemar.test(x, y=NULL, correct=T)`
`mcnemar.test(x, y=NULL, correct=F)`

Όπου:

► x είναι τετραγωνικός πίνακας με μη αρνητικά στοιχεία ή factor ή κατηγορικά δεδομένα.

► y είναι factor ή κατηγορικό δεδομένο. Αν x είναι πίνακας το y αγνοείται.

Διαφορετικά το y πρέπει να έχει την ίδια διάσταση με το x .

ΣΕ S-PLUS

```
> x_c("+", "+", "-", "-", "+", "+", "-", "-", "-", "-", "+", "-", "+", "-  
      ", "+")
```

```
> x
```

```
[1] "+" "+" "-" "-" "+" "+" "-" "-" "-" "-" "+" "-" "+" "-"  
     "+"
```

```
> y_c("-", "+", "+", "+", "-", "-", "-", "-", "+", "+", "+", "-", "+", "-  
      ", "+")
```

```
> table(x,y)
```

```
  + -  
+ 4 3  
- 4 4
```

> mcnemar.test(table(x,y))

McNemar's chi-square test with continuity correction

data: table(x, y)

McNemar's chi-square = 0, df = 1, p-value = 1

> table(x,y)

+ -

+ 4 3

- 4 4

> mcnemar.test(table(x,y),correct=F)

McNemar's chi-square test without continuity correction

data: table(x, y)

McNemar's chi-square = 0.1429, df = 1, p-value = 0.7055

Στην περίπτωση 3*3

Στην περίπτωση 3*3:

Σε ένα τουριστικό πρακτορείο διεξήχθη ένα μικρό γκάλοπ σε 131 παραθεριστές κατά το οποίο ρωτήθηκαν για την προτίμηση τους σχετικά με το ποιο νησί θα επιθυμούσαν να επισκεφθούν για παραπάνω από μια φορά. Τα επιλεγμένα νησιά ήταν :Ρόδος, Μύκονος και Σαντορίνη. Τα συγκεντρωτικά αποτελέσματα καταγράφονται στον διπλανό πίνακα:

	Μετά			
Πριν		Ρόδος	Μύκονος	Σαντορίνη
	Ρόδος	15	3	9
	Μύκονος	5	7	20
	Σαντορίνη	4	8	60

ΑΡΙΘΜΗΤΙΚΗ ΕΠΙΛΥΣΗ

H_0 : “επιλογή του ίδιου νησιού για παραπάνω από μία φορά”

$$X^2 = \frac{(9-4)^2}{(9+4)} + \frac{(5-3)^2}{(5+3)} + \frac{(20-8)^2}{(20+8)} = 7.565$$

Όμως $X^2_{1,0.05} = 3.84 < 7.565$ άρα απόρριψη της H_0 .

Σε S-PLUS

```
> s_matrix(c(15,5,4,3,7,8,9,20,60),3,3,byrow=F)
```

```
> s
```

```
      [,1] [,2] [,3]  
[1,]  15   3   9  
[2,]   5   7  20  
[3,]   4   8  60
```

```
> mcnemar.test(s,correct=F)
```

McNemar's chi-square test without continuity correction

data: s

McNemar's chi-square = 7.5659, df = 3, p-value = 0.0559

Άρα απόρριψη της H_0