ΙΑΤΡΙΚΗ ΣΤΑΤΙΣΤΙΚΗ

ΑΝΑΛΟΓΙΕΣ
Δοκιμασία X^2 στους τετράπτυχους πίνακες

Κατανομή 2239 γυναικών καρκινοπαθών του μαστού και 1370 μη καρκινοπαθών κατά αναμνηστικό προκλητής ή φυσιολογικής εμμηνόπαυσης			
Τρόπος εμμηνόπαυσης	Καρκινοπαθείς	Μη-καρκινοπαθείς	Σύνολο
Προκλητή εμμ.	a= 469 **(21%)**	b= 473 **(35%)**	a+b= 942
Φυσιολογική εμμ.	c= 1770 **(79%)**	d= 897 **(65%)**	c+d= 2667
Σύνολο	a+c=2239	b+d= 1370	n= 3609

το X^2 συγκρίνει τις αναλογίες της μίας κατηγορίας της μιας μεταβλητής μεταξύ των κατηγοριών της άλλης

ΑΝΑΛΟΓΙΕΣ – ΔΙΑΣΤΗΜΑΤΑ ΑΞΙΟΠΙΣΤΙΑΣ

- **Παραδείγματα αναλογιών**

- Ο αριθμός των αγοριών στο σύνολο μιας τάξης είναι ή 40% ή 0,40

- Από 50 άτομα που αρρώστησαν από μία νόσο πέθαναν οι 10. Η αναλογία (**θνητότητα**) είναι 20%

    - Υπό ορισμένες συνθήκες ερμηνεύεται ως: η πιθανότητα να πεθάνει κάποιος αν αρρωστήσει από τη συγκεκριμένη νόσο είναι ίση με 0,20.

ΑΝΑΛΟΓΙΕΣ – ΔΙΑΣΤΗΜΑΤΑ ΑΞΙΟΠΙΣΤΙΑΣ (2)

- Μία αναλογία p=r/n που υπολογίστηκε από κάποιο δείγμα υπόκειται σε **τυχαία δειγματοληπτική διακύμανση**. Π.χ. η θνητότητα στον πληθυσμό των ασθενών στο παράδειγμα 2 μπορεί να είναι μεγαλύτερη ή μικρότερη από 0,20.

- Δειγματοληπτική διακύμανση: **πιθανό σφάλμα**.

- Επομένως μπορούν να υπολογιστούν τα όρια αξιοπιστίας της αναλογίας (κατ' αναλογία με τα όρια αξιοπιστίας μέσης τιμής)

- Ο υπολογισμός πρέπει να λαμβάνει υπόψη ότι οι αναλογίες (και συνεπώς τα όρια αξιοπιστίας τους) δεν μπορούν να είναι μικρότερες του 0 ή μεγαλύτερες του 1.

Μέθοδοι υπολογισμού ορίων αξιοπιστίας

1η περίπτωση:
Προϋπόθεση:
α) Ο αριθμός των παρατηρήσεων είναι αρκετά μεγάλος (**n>40**) και η υπολογισθείσα αναλογία μεταξύ 0,10 και 0,90 **(0,10 <p< 0,90)**

Σε αυτές τις περιπτώσεις το πιθανό σφάλμα υπολογίζεται από τον τύπο της **διωνυμικής** κατανομής:

$$SE_p = \sqrt{\frac{p * q}{n}}$$

όπου q=1-p = συμπληρωματική πιθανότητα της p και n ο αριθμός των παρατηρήσεων.

Μέθοδοι υπολογισμού ορίων αξιοπιστίας

1η περίπτωση:

95% όρια αξιοπιστίας μίας αναλογίας:

$$p \pm 1,96 * SE_p = p \pm 1,96\sqrt{\frac{p \cdot q}{n}}$$

99% όρια αξιοπιστίας μίας αναλογίας :

$$p \pm 2,58 * SE_p = p \pm 2,58\sqrt{\frac{p \cdot q}{n}}$$

όπου 1,96 η οριακή τιμή του επιπέδου 5% της κανονικής κατανομής και 2,58 η αντίστοιχη τιμή του 1%

Παράδειγμα 1 – 1η μέθοδος

Ανάμεσα σε 240 άτομα τα 34 βρέθηκαν να είναι Rhesus αρνητικά. Ποια είναι η αναλογία των ατόμων αυτών και ποια είναι τα 95% όρια αξιοπιστίας της;

$$p = \frac{34}{240} = 0{,}142 \qquad q = 1 - p = 1 - 0{,}142 = 0{,}858$$

$$SE_p = \sqrt{\frac{p \cdot q}{n}} = \sqrt{\frac{0{,}142 \cdot 0{,}858}{240}} = \sqrt{0{,}000508} = 0{,}0226$$

Τα 95% όρια αξιοπιστίας της υπολογισμένης αναλογίας: $p = p \pm 1{,}96 SE_p$

$$p \pm 1{,}96 * (0{,}0226) \rightarrow 0{,}142 + 0{,}044 = 0{,}186$$

$$p \pm 1{,}96 * (0{,}0226) \rightarrow 0{,}142 - 0{,}044 = 0{,}098$$

Άρα η πραγματική αναλογία των Rhesus αρνητικών στον αντίστοιχο πληθυσμό περιλαμβάνεται με πιθανότητα 95% ανάμεσα στα όρια 0,098 (9,8%) και 0,186 (18,6%), η δε πιθανότερη τιμή της αναλογίας αυτής είναι εκείνη που υπολογίστηκε, δηλαδή 0,142 (14,2%).

Μέθοδοι υπολογισμού ορίων αξιοπιστίας

2η περίπτωση:
Προϋπόθεση:

α) Ο αριθμός των παρατηρήσεων είναι αρκετά μεγάλος (**n>40**) και η αναλογία είναι πολύ μικρή (**p <0,10**) ή πολύ μεγάλη (**p> 0,90**).

α1) Αν $p*n>5$ ή $q*n>5$ (*p* ή *q* είναι η μικρότερη από τις δύο αναλογία) τότε πάλι εφαρμόζουμε την κανονική προσέγγιση όπως προηγουμένως.

95% όρια αξιοπιστίας μίας αναλογίας:

$$p \pm 1{,}96 * SE_p = p \pm 1{,}96\sqrt{\frac{p \cdot q}{n}}$$

α2) Αν *p*n* $\leq$ 5 και *q*n* $\leq$ 5 τότε οι παραπάνω τύποι δεν ισχύουν, το διάστημα αξιοπιστίας δεν είναι είναι συμμετρικό ως προς την p και ο προσδιορισμός του απαιτεί την χρήση πολύπλοκων μαθηματικών εξισώσεων ή έτοιμων πινάκων

Μέθοδοι υπολογισμού ορίων αξιοπιστίας

2η περίπτωση: Παράδειγμα

Σε μία έρευνα για τη διερεύνηση διαταραχών της χρωματικής όρασης σε παιδιά σχολικής ηλικίας διαπιστώθηκαν 13 περιπτώσεις με διαταραχές σε σύνολο 239 αγοριών που εξετάστηκαν. Ποιο είναι το ποσοστό (αναλογία) των διαταραχών αυτών σε αγόρια και ποια τα 95% όρια αξιοπιστίας του ποσοστού αυτού;

$$p = \frac{13}{239} = 0,054 \qquad q = 1 - p = 1 - 0,054 = 0,946$$

Εφόσον $p * n = 0,054 * 239 = 13 > 5$ εφαρμόζεται η 1η μέθοδος

$$SE_p = \sqrt{\frac{p \cdot q}{n}} = 0,0146$$

$$p = p \pm 1,96 * SE_p$$

$$0,083$$

$$0,025$$

Μέθοδοι υπολογισμού ορίων αξιοπιστίας

3η περίπτωση:

Προϋπόθεση: Ο αριθμός των παρατηρήσεων είναι μικρός. Το πιθανό σφάλμα σχετικά μεγάλο. Η συμμετρική εφαρμογή του πιθανού σφάλματος γύρω από την εκτιμώμενη αναλογία (1η μέθοδος) πρέπει να αποφεύγεται.

95% όρια αξιοπιστίας: αναφορά σε ειδικούς πίνακες.

Μέθοδοι υπολογισμού ορίων αξιοπιστίας

Μέθοδος 3: Παράδειγμα για n-10

		n=10	
κ	*κ/n*	95% Ο.Α.	99% Ο.Α.
0	0,000	0,0000 – 0,3085	0,0000 – 0,4113
1	0,100	0,0025 – 0,4450	0,0005 – 0,5443
2	0,200	0,0252 – 0,5561	0,0109 – 0,6482
3	0,300	0,0667 – 0,6525	0,0370 – 0,7351
4	0,400	0,1216 – 0,7376	0,0768 – 0,8091
5	0,500	0,1871 – 0,8129	0,1283 – 0,8717
6	0,600	0,2624 – 0,8784	0,1909 – 0,9232
7	0,700	0,3475 – 0,9333	0,2649 – 0,9630
8	0,800	0,4439 – 0,9748	0,3518 – 0,9891
9	0,900	0,5550 – 0,9975	0,4557 – 0,9995
10	1,000	0,6915 – 1,0000	0,5887 – 1,000

Μέθοδοι υπολογισμού ορίων αξιοπιστίας

3η μέθοδος:
Εξετάσθηκαν 38 δείγματα κατεψυγμένου κρέατος από τα οποία τα 6 βρέθηκαν θετικά για σαλμονέλες. Ποια είναι τα 95% όρια αξιοπιστίας του ποσοστού (αναλογίας) μόλυνσης με σαλμονέλα των κρεάτων αυτών;

Από τον αντίστοιχο πίνακα για n=38 και k=6 προκύπτει η αναλογία :

$p = k/n = 0,158$

και τα 95% *Ο.Α.*: (0,060 – 0,313)

Άρα η αναλογία που βρέθηκε είναι 15,8%, αλλά η πραγματική αναλογία στο γενικότερο σύνολο από το οποίο προέρχεται το δείγμα βρίσκεται με 95% πιθανότητα μεταξύ 6,0% και 31,3%.

ΕΛΕΓΧΟΣ ΥΠΟΘΕΣΕΩΝ ΠΟΥ ΑΦΟΡΟΥΝ ΑΝΑΛΟΓΙΕΣ

Ο τρόπος αντιμετώπισης εξαρτάται από τις αναλογίες που εκτιμώνται

Σύγκριση παρατηρηθείσας αναλογίας με δοθείσα

Περίπτωση 1 ή 2 (σχετικά με την παρατηρηθείσα αναλογία)

ΠΑΡΑΔΕΙΓΜΑ 1

Ανάμεσα σε 325 παντρεμένες γυναίκες ηλικίας 30-34 ετών, οι 88 είχαν αναμνηστικό προκλητής έκτρωσης. Αν είναι γνωστό ότι στον πληθυσμό μιας άλλης μεγάλης χώρας έχουν αναμνηστικό προκλητής έκτρωσης το 50% των γυναικών ηλικίας 30-34 ετών, διαφέρει η αναλογία του δείγματός μας από αυτή την αναλογία;

Ho : p1=p0
HA: p1≠p0

$$p_1 = r_1 / n_1 \, , \, p_0 : \delta o \theta \varepsilon \acute{\iota} \sigma \alpha = 0,50$$

Σύγκριση παρατηρηθείσας αναλογίας με δοθείσα

Περίπτωση 1 ή 2 (σχετικά με την παρατηρηθείσα αναλογία)

ΜΕΘΟΔΟΣ

1) Εάν ίσχυε η μηδενική υπόθεση η πραγματική αναλογία στον πληθυσμό από τον οποίο πήραμε το δείγμα (π.χ. των 325 γυναικών) θα ήταν p_0 (π.χ. $p_0 = 50\%$). Το πιθανό σφάλμα εάν ίσχυε η μηδενική υπόθεση θα ήταν :

$$SE = \sqrt{\frac{p_0 * q_0}{n}}, q_0 = 1 - p_o$$

2)Υπολογίζουμε την τιμή του στατιστικού κριτηρίου $t = \frac{|p_1 - p_0|}{SE_{p_0}}$

3) Αξιολόγηση στους πίνακες της *t* κατανομής στους άπειρους Β.Ε (κανονική κατανομή)

Σύγκριση παρατηρηθείσας αναλογίας με δοθείσα

Περίπτωση 1 ή 2 (σχετικά με την υπολογισθείσα αναλογία)

ΠΑΡΑΔΕΙΓΜΑ

1) p_1=88/325=0,271, p_0=0,50, q_0=0,50

$$SE = \sqrt{\frac{0.50 * 0,50}{325}} = 0,0277$$

2) $$t = \frac{|p_1 - p_0|}{SE_p} = \frac{|0,271 - 0,50|}{0,0277} = \frac{0,229}{0,0277} = 8,27$$

3)

ΒΕ	10%	5%	1%	0,1%
άπειρο	1,65	1,96	2,58	3,29

8,27>3,29, p<0,1%

Σύγκριση παρατηρηθείσας αναλογίας με δοθείσα

Περίπτωση 1 ή 2 (σχετικά με την υπολογισθείσα αναλογία)

Συμπέρασμα:
Η αναλογία στο δείγμα μας διαφέρει σε βαθμό πολύ στατιστικά σημαντικό από τη δοθείσα αναλογία

Ερμηνεία:
Η αναλογία των γυναικών ηλικίας 30-34 ετών με προκλητή έκτρωση στο δείγμα μας είναι σημαντικά μικρότερη από αυτή στον αντίστοιχο πληθυσμό της άλλης χώρας (50%).

,

Σύγκριση δύο παρατηρηθεισών αναλογιών

ΠΑΡΑΔΕΙΓΜΑ 2

Σε μία κλινική δοκιμή για την αξιολόγηση της αποτελεσματικότητας του φαρμάκου *Α* έναντι του παλαιότερου φαρμάκου Β, τυχαιοποιήθηκαν 501 ασθενείς. Από τους 257 που πήραν το *Α* φάρμακο πέθαναν σε ένα έτος 41, από τους 244 που πήραν το Β πέθαναν 64. Διαφέρει η αποτελεσματικότητα των δύο φαρμάκων;

Ho : p1=p2→ p1-p2=0

HA: p1≠p2→ p1-p2≠0

$$p_1 = r_1 / n_1 \qquad p_2 = r_2 / n_2$$

Και οι δύο αναλογίες είναι παρατηρηθείσες άρα υπόκεινται σε τυχαία δειγματοληπτική διακύμανση. Το πιθανό σφάλμα της διαφοράς 2 αναλογιών δίνεται από την παρακάτω συνάρτηση:

$$SE\left(p_1 - p_2\right) = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$$

Σύγκριση δύο παρατηρηθεισών αναλογιών

ΘΑΝΑΤΟΣ	ΦΑΡΜΑΚΟ Α	ΦΑΡΜΑΚΟ Β	ΣΥΝΟΛΟ
ΝΑΙ	41	64	105
ΌΧΙ	216	180	396
ΣΥΝΟΛΟ	257	244	501

Το 95% ΟΑ της διαφοράς των 2 αναλογιών είναι

$$p_1 = \frac{41}{257} = 0{,}16 \qquad p_2 = \frac{64}{244} = 0{,}26 \qquad p_1 - p_2 = 0{,}16 - 0{,}26 = -0{,}10$$

$$SE(p_1 - p_2) = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}} = \sqrt{\frac{(0{,}16 \times 0{,}84)}{257} + \frac{(0{,}26 \times 0{,}74)}{244}} = \sqrt{0{,}0013} = 0{,}036$$

$$-0{,}10 \pm 1{,}96 \times 0{,}036$$

-0.03

-0.17

Σύγκριση δύο παρατηρηθεισών αναλογιών

ΘΑΝΑΤΟΣ	ΦΑΡΜΑΚΟ Α	ΦΑΡΜΑΚΟ Β	ΣΥΝΟΛΟ
ΝΑΙ	41	64	105
ΌΧΙ	216	180	396
ΣΥΝΟΛΟ	257	244	501

Ένας τρόπος για να συγκρίνουμε στατιστικά τις 2 αναλογίες είναι πραγματοποιώντας ένα x^2 στον αντίστοιχο τετράπτυχο πίνακα

X^2 = 7,37, ΒΕ=1

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

7,37>6,64, $p<0.01$

,

Σύγκριση δύο παρατηρηθεισών αναλογιών

ΣΥΜΠΕΡΑΣΜΑ

Η αποτελεσματικότητα του φαρμάκου Α (νέου) (που ορίζεται με βάση τον ποσοστό θανάτων επί των ατόμων που πήραν την θεραπεία αυτή) είναι στατιστικά σημαντικά διαφορετική από αυτήν του φαρμάκου Β (παλιού)

ΕΡΜΗΝΕΙΑ

Το ποσοστό των ατόμων που πέθαναν υπό το νέο φάρμακο Α είναι μικρότερο από το ποσοστό των ατόμων που πέθαναν υπό το φάρμακο Β κατά 10 ποσοστιαίες μονάδες , 95% ΟΑ (3 - 17)

Ποια τιμή δεν εμπεριέχεται στο διάστημα αυτό?
Εάν υπολογίζαμε το 99% ΟΑ θα ήταν η τιμή 0 μέσα στο διάστημα?