Στατιστικοί Έλεγχοι Υποθέσεων

Σαλαντή Γεωργία
Εργαστήριο Υγιεινής και Επιδημιολογίας
Ιατρική Σχολή

Τι θέλουμε να συγκρίνουμε;

- **Δύο δείγματα**
  - Μέση αρτηριακή πίεση σε δύο ομάδες
  - Πιθανότητα θανάτου με δύο διαφορετικά είδη αντικαταθλιπτικών
- **Την τιμή ενός μεγέθους και αν αυτή ισούται με μια θεωρητική τιμή**
  - Η διαφορά δύο μέσων είναι 0
  - Το RR είναι 1

Μηδενική υπόθεση

- Μία υπόθεση που θέλουμε να ελέγξουμε
- Αναφέρεται σε κάποια χαρακτηριστικά των δειγμάτων που εξετάζουμε εκφρασμένα σε στατιστικές ποσότητες
- H_0: $\mathbf{M_1=M_2}$ (η μέση αρτηριακή πίεση είναι ίδια σε δύο ομάδες)
- H_1: $\mathbf{M_1>M_2}$
- H_1: $\mathbf{M_1<M_2}$
- H_1: $\mathbf{M_1>M_2}$ ή $\mathbf{M_1<M_2}$

Ποιο τέστ θα χρησιμοποιήσω;

- Συνεχή ή διχότομα δεδομένα;
- Αν είναι συνεχή, ακολουθούν **κανονική κατανομή** ή όχι;
- **Για κανονική κατανομή:** z-test, t-test
- **Για μη κανονική κατανομή:** Mann–Whitney test, Wilcoxon test
- **Για διχότομα δεδομένα:** X^2 test, Fisher test. McNemar test

Τεστ για συνεχή δεδομένα

Παράδειγμα

- Σε ένα δείγμα 100 ατόμων 50 ετών μετρήθηκε η συστολική πίεση
- Μέση τιμή (δείγματος) μ =124 mmHg
- SD =15 mmHg (υποθέτουμε ότι είναι ίδια στον πληθυσμό και το δείγμα…)
- α =120
- Είναι πολύ υψηλά αυτά τα επίπεδα;
- Δεν μπορούμε να πούμε αν το 124 είναι μεγάλο σε σχέση με το 120, και γι αυτό θα μετασχηματίσουμε τα δεδομένα σε ένα μέγεθος για το οποίο μπορούμε να αποφανθούμε εαν είναι μεγάλο ή μικρό
- **Μετασχηματισμός Z**

Z-test

- Μέση τιμή του πληθυσμού **M**
- Μέση τιμή του δείγματος **μ**
- Μέγεθος δείγματος **n**
- Τυπική απόκλιση SD
- **H_0: M=α**
- **H_1: α>M ή α<M**

Μετασχηματισμός Z

$$z=\frac{|\mu-\alpha|}{\frac{SD}{\sqrt{n}}}=\frac{|\mu-\alpha|}{SE}$$

Η τιμή του z ακολουθεί κανονική κατανομή με μέσο 0 και SD 1 και άρα ξέρουμε αν είναι μεγάλη ή μικρή

Παράδειγμα

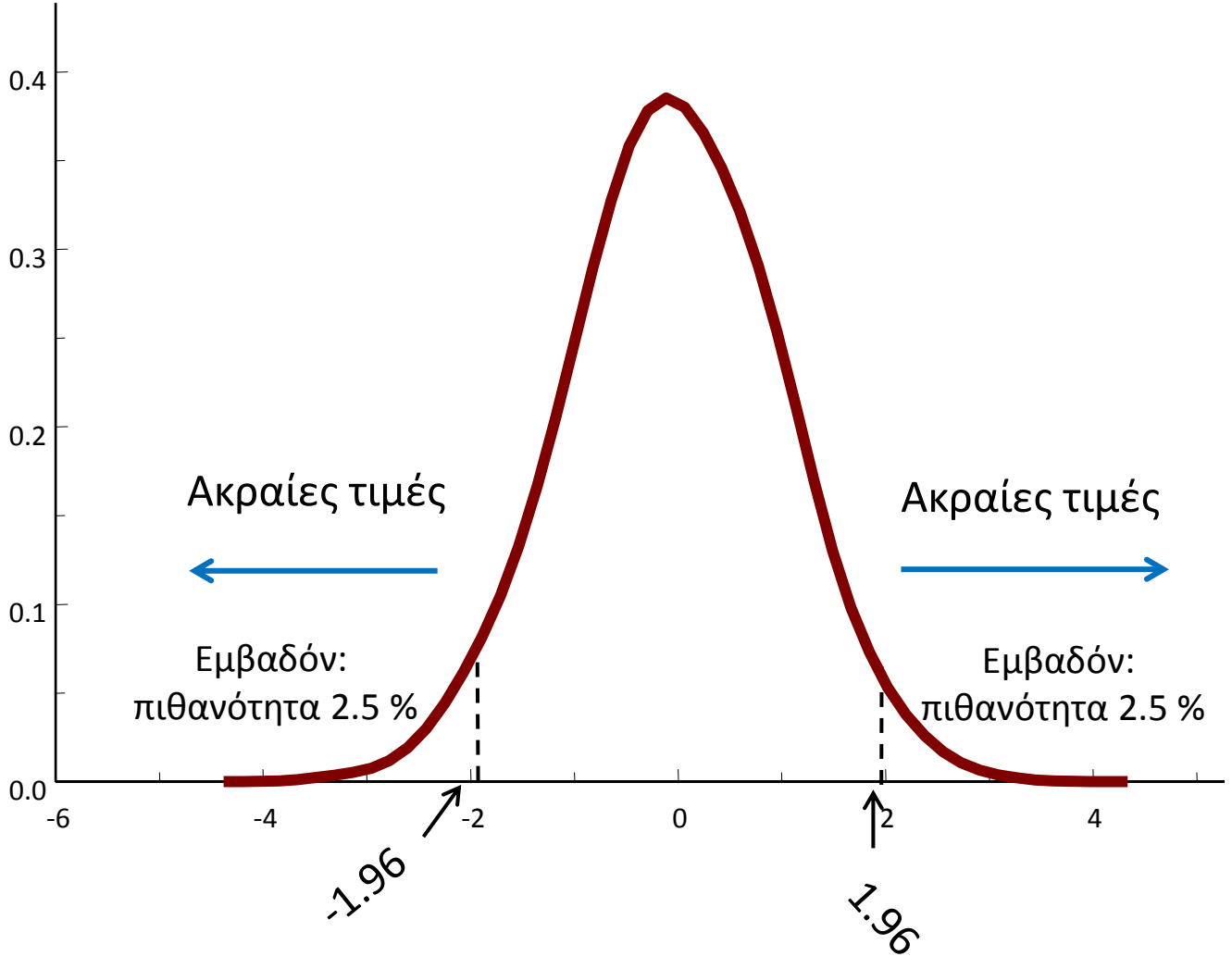
- Σε ένα δείγμα 100 ατόμων 50 ετών μετρήθηκε η συστολική πίεση
- Μέση τιμή (δείγματος) μ =124 mmHg
- SD =15 mmHg (υποθέτουμε ότι είναι ίδια στον πληθυσμό και το δείγμα…)
- α =120
- Θα μετασχηματίσουμε τα δεδομένα σε ένα μέγεθος για το οποίο μπορούμε να αποφανθούμε εαν είναι μεγάλο ή μικρό
- **H_0:Μ=α, η μέση πίεση στο δείγμα δεν διαφέρει από αυτή στον γενικό πληθυσμό**

Παράδειγμα: υπολογισμός Z-test

$$z = \frac{|124 - 120|}{\frac{15}{\sqrt{100}}} = 2.67$$

Το 2.67 θα το συγκρίνουμε με την πρότυπη κανονική κατανομή

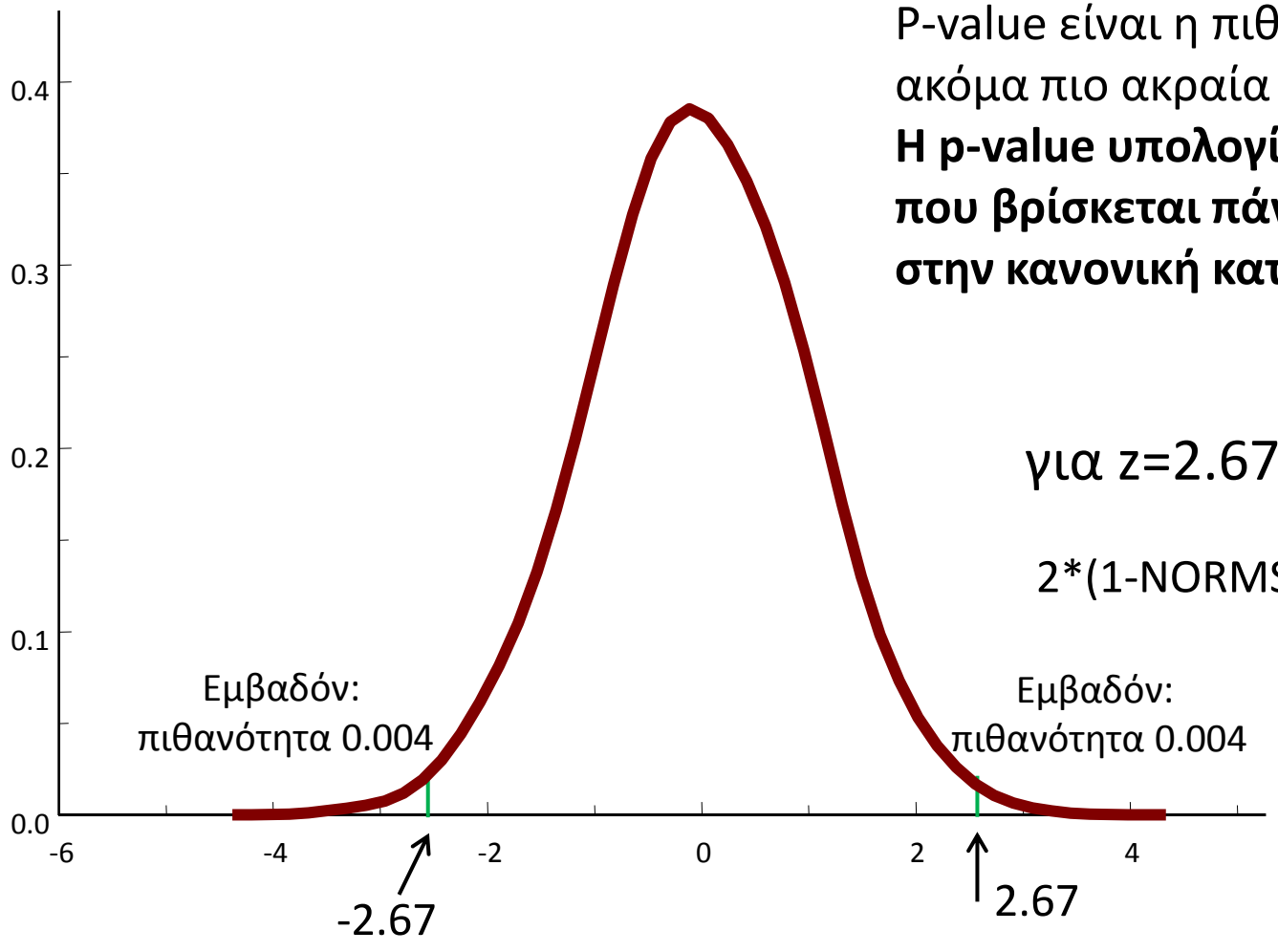
Πρότυπη κανονική κατανομή N(0,1)



Z-τεστ

- Εάν η τιμή του z είναι μεγαλύτερη του 1.96 σημαίνει οτι έιναι ‘ακραία’ ή μεγάλη τιμή
- Άρα το εύρημα είναι «**στατιστικά σημαντικό**»
- Άρα **απορρίπτω την μηδενική υπόθεση α=Μ**
- Συνεπώς, το 124 δεν είναι ‘στατιστικώς’ συγκρίσημο με το 120

P-value του z-test



Βρίκαμε z=2.67

P-value είναι η πιθανότητα να βρω μιά ακόμα πιο ακραία τιμή

Η p-value υπολογίζεται σαν το εμβδαδόν που βρίσκεται πάνω απο z και κάτω απο -z στην κανονική κατανομή

για z=2.67 η p-value είναι 0.8%

2*(1-NORMSDIST(**z**)) (στο excel)

P-value: ερμηνεία

- Εάν επαναλάβω αυτή τη μέτρηση σε άλλα 100 άτομα θα βρω το ίδιο;
- Εάν το επαναλάβω πολλές φορές (πχ 1000 φορές) σε δείγματα του ιδίου μεγέθους, πόσες φορές θα βρω κάτι τόσο ‘ακραίο’ όπως μέση πίεση 124;
- x/1000 φορές = p-value
- p-value=το εμβαδόν στην πρότυπη κανονική κατανομή που είναι έξω από το διάστημα (-z,z)

P-value: ερμηνεία

- Εάν η τιμή του δείγματος δεν είναι στο 5% των πιο ακραίων τιμών, τότε το εύρημα θεωρείται '**στατιστικά μη σημαντικό**'και **δεν απορρίπτουμε την H_0**
- Εάν είναι στο 5%, **τότε απορρίπτουμε την H_0**
- P-value<0.05: απορρίπτουμε την H_0, το εύρημα που έχουμε (δηλ. μια μέση πίεση 124) δεν μπορεί να εξηγηθεί «από την τύχη»

Πώς υπολογίζουμε p-values

- 2*(1-NORMSDIST(z)) (στο excel)
- 2*(1-NORMSDIST(2.67)) = 0.00759
- Αφού η p-value είναι μικρότερη του 5% λέμε οτι έχουμε ένα «στατιστικά σημαντικό» αποτέλεσμα
- Απορρίπτουμε τη μηδενική υπόθεση

Αμφίπλευρα και μονόπλευρα τεστ

Τα μονόπλευρα τεστ κοιτάνε μόνο ένα απο τα δύο άκρα της κατανομής

- H_0: α=Μ με H_1: α>Μ ή α<Μ

  z=1.7, **p-value=0.09 στατιστικά μή σημαντικό**
- H_0: α=Μ με H_1: α>Μ

  z=1.7, **p-value=0.045 στατιστικά σημαντικό**

Τα αμφίπλευρα τεστ είναι πιο **'συντηρητικά'**

(= δεν απορρίπτουν τόσο εύκολα τη μηδενική υπόθεση)

Z-τεστ: Προϋποθέσεις εφαρμογής

- Η κατανομή του μεγέθους που μετράμε πρέπει να είναι κανονική
- Η τυπική απόκλιση στον πληθυσμό να είναι η ίδια με αυτή στο δείγμα
  - Πχ στο παράδειγμα ήταν SD=15 που είναι μάλλον μικρή τ.α.
  - Ποιο θα ήταν το αποτέλεσμα του τεστ αν η SD ήταν μεγαλύτερη;

Παράδειγμα

Υπάρχει διαφορά οσον αφορά στον πόνο ανάμεσα στις δύο μεθόδους;

Πόνος	**Ναι**	**Όχι**	**Σύνολο**
Ενδοσκοπική	10	90	100
Ανοιχτή	14	86	100
	24	176	200

Μηδενική υπόθεση H_0: RR=1 ή lnRR=0

Παράδειγμα: λύση

$$RR=\frac{\frac{10}{100}}{\frac{14}{100}}=0.71,\ \ln RR=\ln(0.71)=-0.34$$

$$SE(\ln RR)=\sqrt{\frac{1}{10}-\frac{1}{100}+\frac{1}{14}-\frac{1}{100}}=0.39$$

$$z=\frac{|\mu-M|}{SE}=\frac{|-0.34-0|}{0.39}=0.87$$

P-value για z=0.87 είναι 2*(1-NORMSDIST(0.87))=**0.38**. Άρα το έυρημα είναι **στατιστικά μή σημαντικό** και δεν μπορούμε να ισχυριστούμε οτι υπάρχει διαφορά μεταξύ των δύο ομάδων

Δεν απορρίπτουμε τη μηδενική υπόθεση

Το 95% ΔΕ για το RR ήταν (0.33, 1.52)

t-test για ανεξάρτητα δείγματα

- Όταν θέλουμε να συγκρίνουμε **δυο μέσους από δύο δείγματα**
- **Προϋπόθεση:** ότι οι δύο πληθυσμοί από όπου προέρχονται τα δείγματα έχουν τις ίδιες διασπορές και ότι τα δείγματα προέρχονται από την κανονική κατανομή
- H_0: $M_1=M_2$
- H_1: $M_1>M_2$ ή $M_1<M_2$ (αμφίπλευρος έλεγχος)

t-test: τύποι

$$t=\frac{\mu_1-\mu_2}{s}$$

Δείγμα 1: μέσος $\boldsymbol{\mu_1}$ με $\mathbf{SD_1}$ ($\mathbf{n_1}$ μετρήσεις)

Δείγμα 2: μέσος $\boldsymbol{\mu_2}$ με $\mathbf{SD_2}$ ($\mathbf{n_1}$ μετρήσεις)

s είναι μια ‘απο κοινού’ τυπική απόκλιση

$$s=\sqrt{\frac{SD_1^2+SD_2^2}{n_1}},n_1=n_2$$

$$s=\sqrt{\frac{(n_1-1)SD_1^2+(n_2-1)SD_2^2}{n_1+n_2-2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

t κατανομή με n_1+n_2-2 βαθμούς ελευθερίας

Παράδειγμα

- μ_1=15, SD_1^2 διασπορά$_1$=4.6, n_1=8
- μ_2=14, SD_2^2 διασπορά$_1$=6.6, n_2=8

$$t=\frac{15-14}{\sqrt{\frac{4.6+6.6}{8}}}=0.85$$

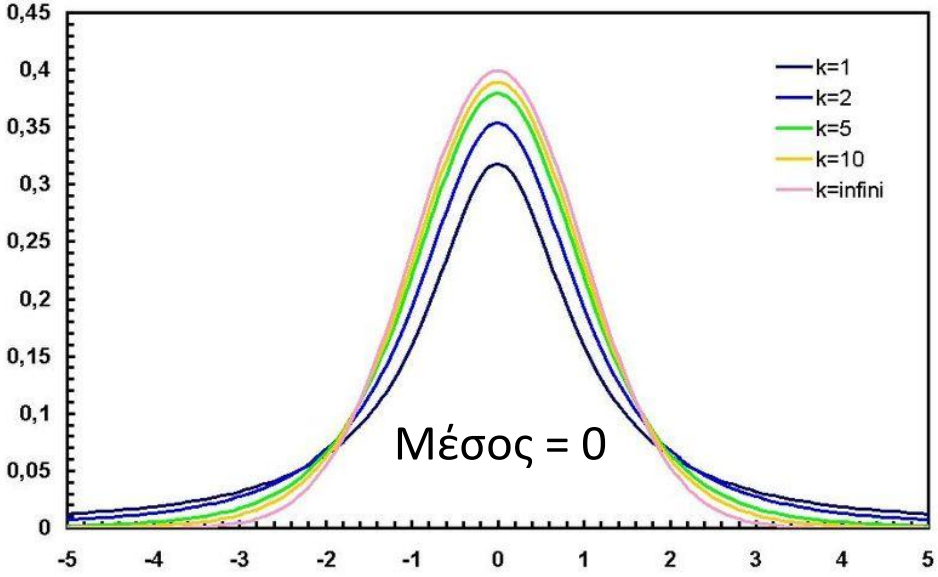
Το 0.85 είναι μεγάλο ή μικρό;

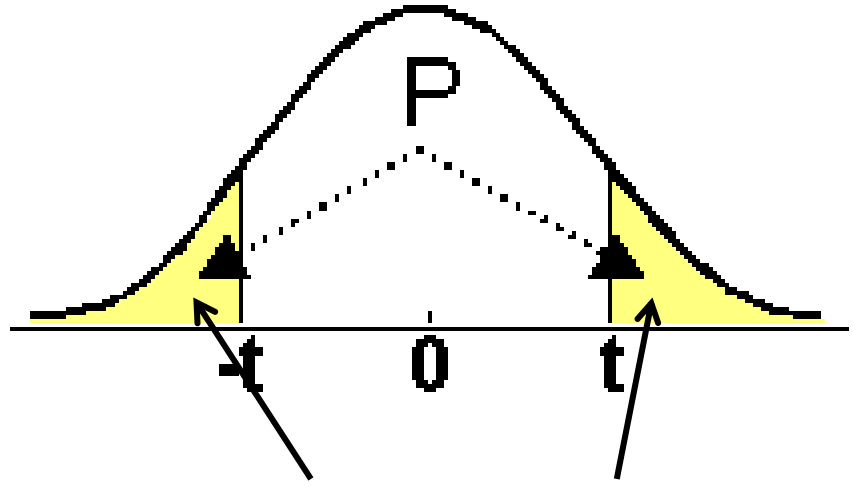
Πάμε στην t κατανομή με $n_1+n_2-2=14$ βαθμούς ελευθερίας

t κατανομή (student)

κ οι βαθμοί ελευθερίας (degrees of freedom **df**) ορίζονται από το μέγεθος του συνολικού δείγματος. Όσο πιο πολλοί βαθμοί ελευθερίας, τόσο πιο «ψιλόλιγνη» είναι

Για μέγεθος δείγματος>20 η t κατανομή είναι παρόμοια με την N(0,1)





p-value = TDIST(t,df,2)=TDIST(0.85,14,2)= **0.41** Εμβαδον=0.205

Τι σημαίνει αυτό;

Δεν απορρίπτουμε την H_0

t-test: προϋποθέσεις εφαρμογής

- Είναι τα δείγματα από κανονική κατανομή; Πώς το ελέγχω;
  - Κοιτάμε τις κατανομές των δεδομένων να «μοιάζουν» με κανονικές
  - Ελέγχουμε: Kolmogorov-Smirnov test ή shapiro-wilk (H_0: το δείγμα προέρχεται από την κανονική κατανομή)
- Είναι οι διασπορές ίδιες; Πώς το ελέγχω;
  - Ελέγχουμε: με το levene F τεστ (H_0: οι δύο διασπορές είναι ίδιες)
- Όταν έχω μεγάλα δείγματα, μπορώ να το υπολογίσω σαν z-test

Παράδειγμα t-test και έλεγχος προυποθέσεων

- μ_1=15, $SD^2{}_1$=4.6, n_1=8
- μ_1=14, $SD^2{}_2$=6.6, n_2=8

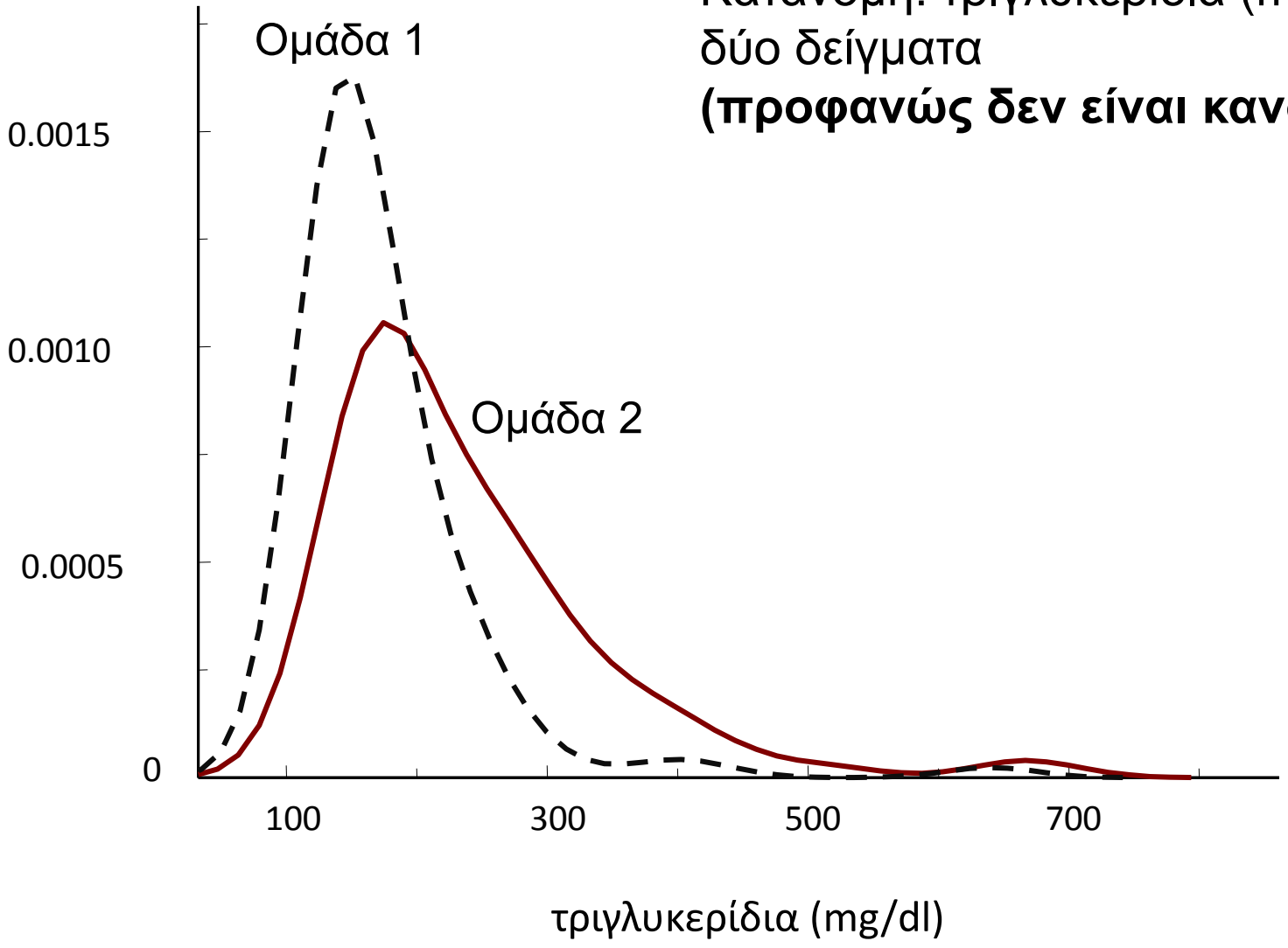
1. Ελέγχουμε εάν τα δείγματα αναφέρονται σε κανονική κατανομή
   - P-value (K-S test)>0.05 για να μην απορρίψω την υπόθεση της κανονικότητας
2. Ελέγχουμε εάν οι διασπορές είναι ίδιες
   - P-value (F test)>0.05 για να μην απορρίψω την υπόθεση της ισότητας
3. Υπολογίζουμε το t-test
   - P-value (t test)<0.05 απορρίπτουμε την υπόθεση ισότητας

Mann–Whitney test

- Όταν τα δεδομένα **δεν** προέρχονται από την κανονική κατανομή
  - Τριγλυκερίδια
- Η μέση τιμή δεν αντιπροσωπεύει επιτυχώς τα δεδομένα - καλύτερη η διάμεσος
- Για να ελέγξω αν η ομάδα 1 είναι συγκρίσιμη με την ομάδα 2 στα τριγλυκερίδια ελέγχω την

H_0: τα δύο δείγματα προέρχονται από την ίδια κατανομή

Κατανομή: τριγλυκερίδια (mg/dl) σε δύο δείγματα
(προφανώς δεν είναι κανονική)



Γιατί όχι t-test;

- Ομάδα 1: διάμέσος=136, μέσος=237
- Ομάδα 2: διαμέσος=191, μέσος=442
- t-test: p-value=0.009
- Mann–Whitney: p-value=0.14

Ζευγαρωτές παρατηρήσεις (Paired data)

- Ζευγαρωτές παρατηρήσεις: μετράμε τα ίδια άτομα ‘πριν’ και ‘μετά’
  - π.χ. μέση πίεση πριν την αγωγή 154, μετά την αγωγή 145
- Χρησιμοποιούμε παραλλαγές των τεστ (paired t-test and Wilcoxon)

Ασθενής	Πρίν	Μετά
1	160	158
2	174	172
3	168	172
4	180	165
5	182	170
6	175	175

t-test για ζευγαρωτές παρατηρήσεις

Υπολογίζουμε τις διαφορές (πρίν και μετά)

Ασθενής	Πρίν	Μετά	Διαφορά
1	160	158	2
2	174	172	2
3	168	172	-4
4	180	165	15
5	182	170	12
6	175	175	0

Μέση διαφορά μ_D=4.5 mmHg

Μηδενική υπόθεση M_D =0

Paired t-test $$t=\frac{\mu_D-0}{SE_D}$$

t κατανομή με n-1 βαθμούς ελευθερίας

Ισχύς ενός τεστ

Η ‘δύναμη’ που έχει ένα τεστ να απορρίπτει τη μηδενική υπόθεση (=να δίνει p-value<0.05) όταν αυτή δεν ισχύει

		Πραγματικότητα	
τεστ		H_0 σωστή	H_0 λάθος
	H_0 απορρίπτεται	Σφάλμα τύπου I =0.05	Ισχύς
	H_0 δεν απορρίπτεται		Σφάλμα τύπου II

Ισχύς

- Εξαρτάται από το μέγεθος του δείγματος
  - Όσο πιο μεγάλο, τόσο πιο μεγάλη η ισχύς
- Εξαρτάται από τη διαφορά ανάμεσα στις δύο ομάδες
  - Όσο πιο μεγάλη η διαφορά, τόσο πιο μεγάλη η ισχύς

Ισχύς και μέγεθος δείγματος

- $\mu_1 = 5$ in $n_1=10$ and $\mu_2=6$ in $n_2=10$
  - t-test **p-value=0.13**
- $\mu_1 = 5$ in $n_1=20$ and $\mu_2=6$ in $n_2=20$
  - t -test **p-value=0.03**
- Άρα, όταν έχουμε μια p-value **στατιστικά μη σημαντική**, πρέπει να έχουμε στο νου μας ότι **αυτό μπορεί να οφείλεται στο μικρό δείγμα**!

Υπολογισμός μεγέθους δείγματος

- Πρέπει από πριν να καθορίσουμε
  - Το τεστ που θα χρησιμοποιήσουμε
  - Την προσδοκώμενη διαφορά ανάμεσα στις δύο ομάδες
  - Την ισχύ που θέλουμε (Χ%, π.χ. 80% πιθανότητα αν η H_0 είναι λάθος το τεστ να μπορέσει να την απορρίψει)
- Έπειτα υπολογίζουμε το απαιτούμενο μέγεθος δείγματος για να έχουμε ισχύ Χ%

Παράδειγμα

- Από μια κλινική δοκιμή που τυχαιοποίησε τους ασθενείς σε 2 ομάδες των 100 ατόμων προέκυψαν τα παρακάτω αποτελέσματα
- Δίαιτα 1, μέσο βάρος 80 κιλά, SD 30
- Δίαιτα 2, μέσο βάρος 83 κιλά, SD 31

Διαφέρουν στατιστικώς οι δύο δίαιτες;
Απαντήστε την ερώτηση με 4 τρόπους!

Παράδειγμα:Λύση

Λύση 1:Υπολογισμός δύο 95% CI

- 95% CI για δίαιτα 1 (74.12, 85.88)
- 95% CI για δίαιτα 2 (76.92, 89.08)

Λύση 2:Υπολογισμός 95% CI για MD

- MD=3, SE(MD)=4.31
- 95% CI για MD (-5.46, 11.46,)

Λύση 3:Υπολογισμός z-test

- Z=0.70
- Z<1.96 άρα p-value>0.05

Λύση 4: Υπολογισμός t-τεστ

- t=0.70, df βαθμοί ελευθερίας=198
- Άρα η t κατανομή είναι σχεδόν η κανονική κατανομή
- Άρα p-value>0.05

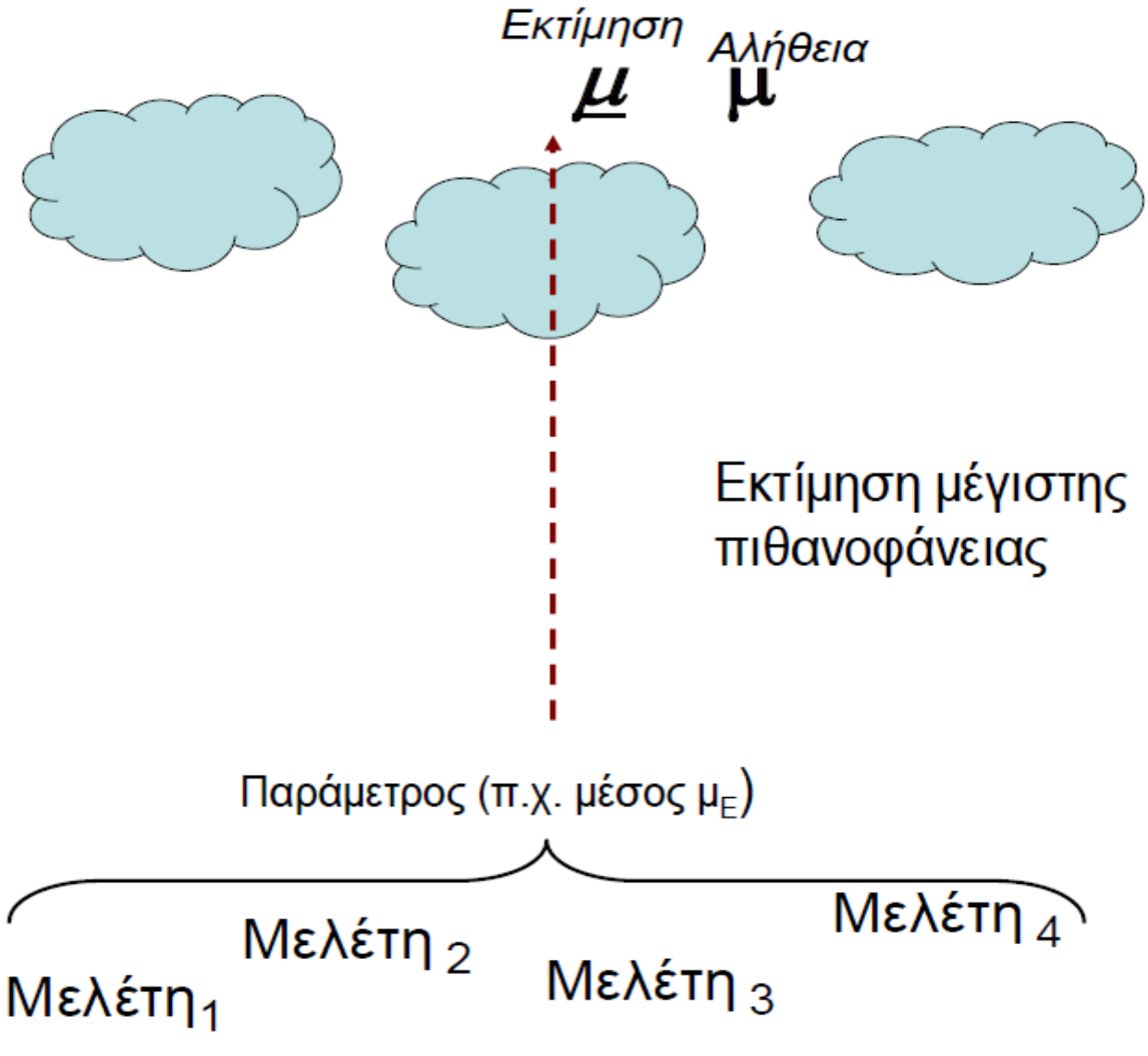
Δεν απορίπτουμε την μηδενική υπόθεση οτι οι δύο δίαιτες είναι εξίσου αποτελεσματικές

Τεστ για διχότομα και διακριτά δεδομένα

Γεωργία Σαλαντή
Κώστας Τσιλίδης

Βιβλίο Pagano: κεφ. 15, 16

Γιατί χρησιμοποιούμε στατιστικές μεθόδους;



Διακριτές μεταβλητές

- Διχότομα δεδομένα (δύο πιθανές εκβάσεις)
  - Φύλο, νοσήματα: ναι/οχι
- Διακριτά δεδομένα (πάνω από δύο εκβάσεις)
  - Ναι/όχι/ίσως, βελτίωση/επιδείνωση/σταθερός

Αναιμία και φύλο: 2x2 πίνακας

		Αναιμία		Σύνολο
		Όχι	Ναι	
Φύλο	Γυναίκες	47 (82%)	10 (18%)	57
	Άνδρες	59 (91%)	6 (9%)	65
		106	16	

Είναι πιο συχνή η αναιμία στους άντρες ή στις γυναίκες;

Μηδενική υπόθεση

Η κατανομή των παρατηρήσεων στα κελιά του 2x2 πίνακα είναι τυχαία

(H_0: η πιθανότητα αναιμίας στους άνδρες = πιθανότητα αναιμίας στις γυναίκες)

χ^2 τεστ για διχότομα δεδομένα: Υπολογισμός αναμενόμενων τιμών

Υπολογίζουμε τον **αναμενόμενο αριθμό** σε κάθε κελί ***εάν*** η πιθανότητα αναιμίας στους άνδρες και τις γυναίκες ήταν ίδια

		Αναιμία		Σύνολο
		Όχι	Ναι	
Φύλο	Γυναίκες	47 57·106/122 =49.5	10	57
	Άνδρες	59	6	65
		106	16	122

Αναμενόμενος αριθμός=
Σύνολο κολώνας×Σύνολο γραμμής/Γενικό Σύνολο

χ^2 τεστ για διχότομα δεδομένα: Υπολογισμός αναμενόμενων τιμών

Υπολογίζουμε τον αναμενόμενο αριθμό σε κάθε κελί ***εάν*** η πιθανότητα αναιμίας στους άνδρες και τις γυναίκες ήταν ίδια– **ταιριάζουν με τις παρατηρήσεις;**

		Αναιμία		Σύνολο
		Όχι	Ναι	
Φύλο	Γυναίκες	O=47 E=49.5	O=10 E=7.5	57
	Άνδρες	O=59 E=56.5	O=6 E=8.5	65
		106	16	122

χ^2 τεστ για διχότομα δεδομένα

Υπολογίζουμε τον αναμενόμενο αριθμό σε κάθε κελί ***εάν*** η πιθανότητα αναιμίας στους άνδρες και τις γυναίκες ήταν ίδια– ταιριάζουν με τις παρατηρήσεις;

$$\chi^2 = \sum_{i=1,2,3,4} \frac{\left(O_i - E_i\right)^2}{E_i}$$

O_i οι παρατηρήσεις, E_i οι προσδοκώμενες τιμές στα 4 κελιά

Ακολουθεί χ^2 κατανομή με 1 βαθμό ελευθερίας

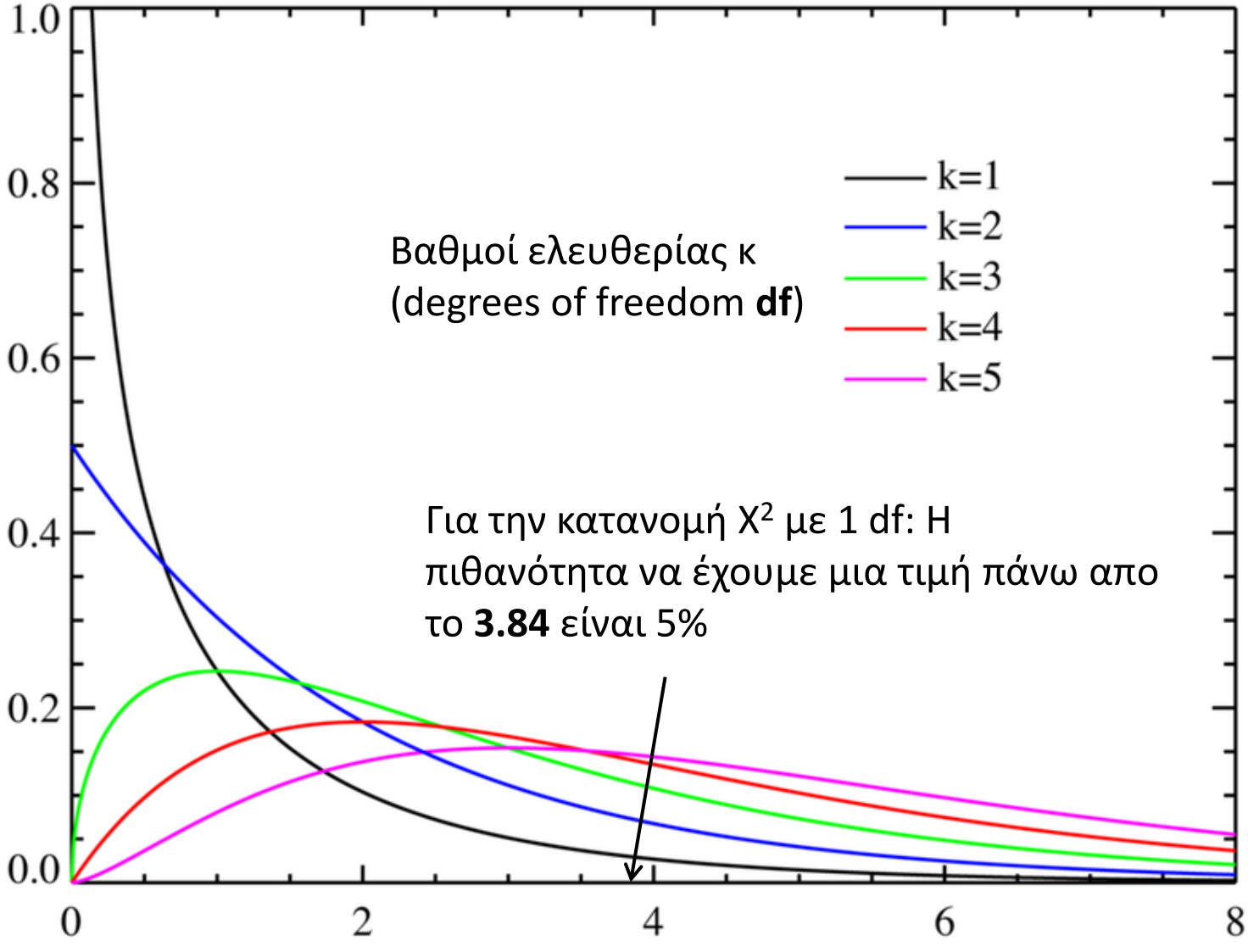
χ^2 τεστ για διχότομα δεδομένα

Υπολογίζουμε τον αναμενόμενο αριθμό σε κάθε κελί ***εάν*** η πιθανότητα αναιμίας στους άνδρες και τις γυναίκες ήταν ίδια– ταιριάζουν με τις παρατηρήσεις;

		Αναιμία		Σύνολο
		Όχι	Ναι	
Φύλο	Γυναίκες	O=47 E=49.5	O=10 E=7.5	57
	Άνδρες	O=59 E=56.5	O=6 E=8.5	65
		106	16	122

$$\chi^2=1.81$$

Κατανομή X^2

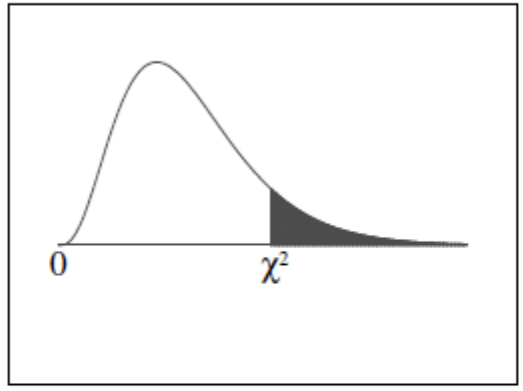

1.0
0.8
0.6
0.4
0.2
0.0
0
2
4
6
8
k=1
k=2
k=3
k=4
k=5
Βαθμοί ελευθερίας κ
(degrees of freedom **df**)
Για την κατανομή X^2 με 1 df: Η πιθανότητα να έχουμε μια τιμή πάνω απο το **3.84** είναι 5%

P-value του X^2 test

- Η πιθανότητα να βρούμε μια τόσο ακραία τιμή όσο αυτή του τέστ (=1.81) εάν το πείραμα επαναληφθεί με το ίδιο μέγεθος δείγματος
- 1.81 <3.84
- Για X^2=1.81 η p-value είναι =0.18
- Στο Excel "`1-CHISQ.DIST(3.84;1;1)`"
- Τι σημαίνει αυτό;
  - Ότι η H_0 δεν απορρίπτεται
  - Ότι η συσχέτιση δεν είναι στατιστικά σημαντική

P-value του X^2 test

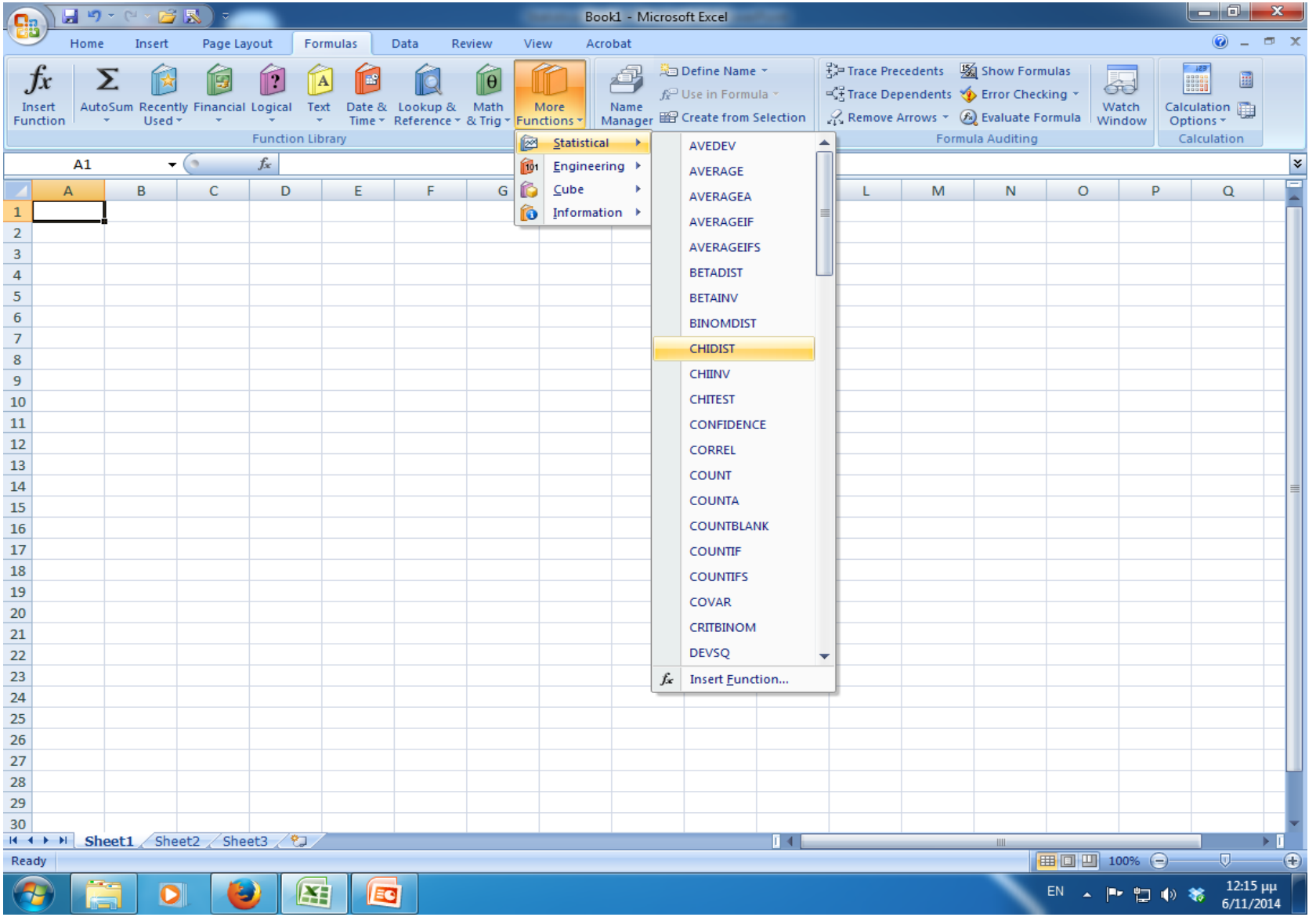
Chi-Square Distribution Table



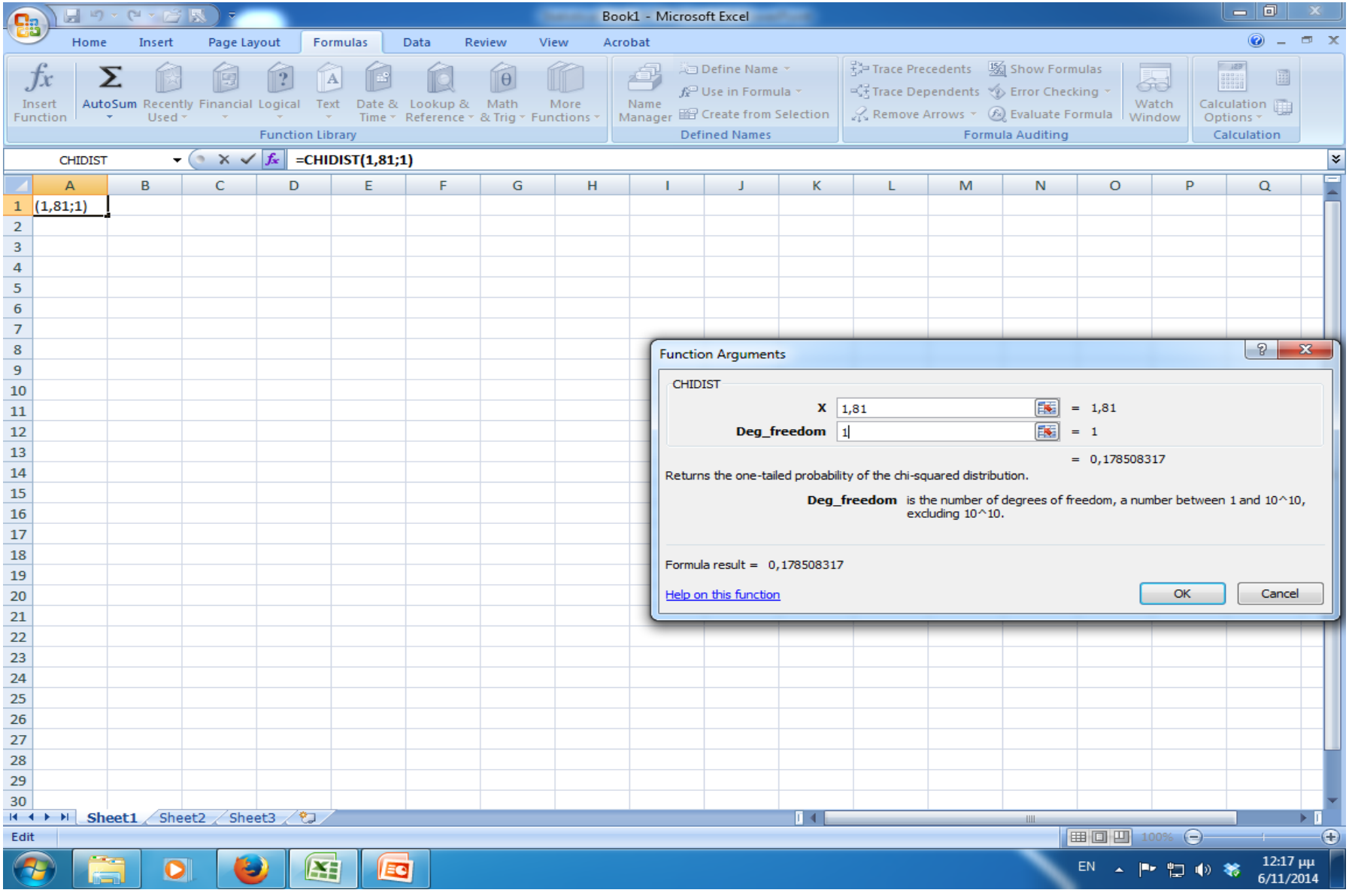
The shaded area is equal to α for $\chi^2 = \chi^2_\alpha$.

df	$\chi^2_{.995}$	$\chi^2_{.990}$	$\chi^2_{.975}$	$\chi^2_{.950}$	$\chi^2_{.900}$	$\chi^2_{.100}$	$\chi^2_{.050}$	$\chi^2_{.025}$	$\chi^2_{.010}$	$\chi^2_{.005}$
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750

P-value του X^2 test



P-value του X^2 test



Λόγος αναλογιών (λόγος odds, OR)

		Αναιμία		Σύνολο
		Όχι	Ναι	
Φύλο	Γυναίκες	47 (β1)	10 (α1)	57 (N1)
	Άνδρες	59 (β0)	6 (α0)	65 (N0)
		106	16	122

- OR = $\dfrac{\frac{a1}{n1}/1-\frac{a1}{n1}}{\frac{a0}{n0}/1-\frac{a0}{n0}} = \dfrac{a1/b1}{a0/b0} = \dfrac{a1\cdot b0}{a0\cdot b1}$
- SE(lnOR) = $\sqrt{\dfrac{1}{a1}+\dfrac{1}{a0}+\dfrac{1}{b1}+\dfrac{1}{b0}}$
- 95% CI του lnOR = lnOR ± 1.96·SE(lnOR)
- 95% CI του RR = exp{lnOR ± 1.96·SE(lnOR)}

χ^2 τεστ για διακριτά δεδομένα

		Αναιμία		Σύνολο
		Όχι	Ναι	
Γονότυπος	AA	53	10	63
	AC	45	5	50
	CC	8	1	9
		106	16	122

χ^2 τεστ για διακριτά δεδομένα

- Υπολογίζω τις προσδοκώμενες τιμές όπως πριν
  - Σύνολο κολώνας×Σύνολο γραμμής/Γενικό Σύνολο
- Χρησιμοποιώ τον ίδιο τύπο για το χ^2 όπως πριν
- Κατανομή χ^2 με **2** βαθμούς ελευθερίας
- P-value=0.64
- Τι σημαίνει αυτό;

Υπολογισμός βαθμών ελευθερίας df σε ένα X^2 τεστ

Αν ξέρω τα συνολικά νούμερα (σε γραμμές και στήλες), πόσα κελιά χρειάζομαι για να συμπληρώσω **όλον τον πίνακα;**

		Αναιμία		Σύνολο
		Όχι	Ναι	
Φύλο	Γυναίκες	•		57
	Άνδρες			65
		106	16	122

		Αναιμία		Σύνολο
		Όχι	Ναι	
Γονότυπος	AA	•		63
	AC	•		50
	CC			9
		106	16	122

2x2 πίνακας έχει 1 df
3x2 πίνακας έχει 2 df
3x3 πίνακας έχει 4 df
4x2 πίνακας έχει 3 df

Προϋποθέσεις για την εφαρμογή του X^2 τεστ

Όλες οι προσδοκώμενες τιμές να είναι πάνω από 5

Το τεστ δεν είναι αξιόπιστο…

		Αναιμία		Σύνολο
		Όχι	Ναι	
Γονότυπος	AA	53	10	63
	AC	45	5	50
	CC	8	1 1.18	9
		106	16	122

Fisher exact test

Μπορούμε να το εφαρμόσουμε όταν έχουμε λίγες ή σπάνιες παρατηρήσεις με κάποια κελιά να έχουν λίγες ή και καθόλου παρατηρήσεις

p-value = 0.77

McNemar test

Είναι η παραλλαγή του X^2 για ζευγαρωτές πρατηρήσεις

Πρίν την εγχείρηση

Μετά την εγχείρηση

	Ναι	**όχι**	**Συνολο**
Ναι	*a*	*b*	*a+b*
Οχι	*c*	*d*	*c+d*
	a+c	*b+d*	*n*

H_0 η πιθανότητα πρίν και μετά την εγχείρηση είναι η ίδια

$$X^2 = \frac{(b-c)^2}{b+c}$$

Κατανομή με 1 βαθμό ελευθερίας

Διαστρωμάτωση

- **Λεπτοσπίρωση:** Πιο πολλά κρούσματα στις αγροτικές περιοχές;
- Πόσο επί τις εκατό του πληθυσμού έχει αντισώματα στις πόλεις και πόσο στις αγροτικές περιοχές;

Αντισώματα	Αγροτικές	Αστικές	Σύνολο
Ναι	60	60	120
Οχι	140	140	280
Σύνολο	200	200	400

Επιπολασμός αντισωμάτων για λεπτοσπίρωση στις δυτικές Ινδίες (αστικές και αγροτικές περιοχές)

Kirkwood, Medical Statistics

Ανάλυση

- Τι τεστ θα κάνουμε;
- Πόσο θα είναι η p-value X^2 test?
- p-value > 0.05
- $X^2 = 0$, df = 1, p-value = 1
- Τι σημαίνει αυτό;

Άνδρες

Αντισώματα	Αγροτικές	Αστικές	Σύνολο
Ναι	36	50	86
Οχι	14	50	64
Σύνολο	50	100	150

Γυναίκες

Αντισώματα	Αγροτικές	Αστικές	Σύνολο
Ναι	24	10	34
Οχι	126	90	216
Σύνολο	150	100	250

2 Στρώματα

X^2=5.73, df=1, **p-value<0.025**

X^2=1.36, df=1, **p-value=0.25**

X^2 Mantel-Haenszel=7.09, df=1, p-value<0.01

Mantel-Haenszel (MH) τεστ

- Το χρησιμοποιούμε για να ελέγξουμε δεδομένα που υποψιαζόμαστε ότι είναι **στρωματοποιημένα**
  - Π.χ. διαφορετικά φύλα, διαφορετικές ομάδες ηλικίας
- Το MH ελέγχει κάθε ένα από τα στρώματα και τα συνοψίζει τους ελέγχους σε μία μόνο p-value
- Γιατί δεν κάνουμε δύο ξεχωριστά τεστ;
- Γιατί να συνοψίσουμε τους ελέγχους;
- Για να κερδίσουμε **ισχύ και να δούμε εάν υπάρχει διαφορά μεταξύ των συσχετίσεων στα δύο στρώματα.**

Άσκηση

	Βελτίωση	Σταθερός
Βελονισμός	12	15
Φάρμακα	6	14

Δείξτε με δύο διαφορετικούς τρόπους αν ο βελονισμός σχετίζεται με βελτίωση του πόνου της ράχης

Άσκηση

Λύση 1: Υπολογισμός 95% CI για το OR

OR=1.87 lnOR=0.62 var(lnOR)=0.39 SE(lnOR)=0.62

95% ΔΕ για lnOR: 0.62-1.96·0.62, 0.62+1.96·0.62

95% ΔΕ για lnOR: -0.60, 1.84

95% ΔΕ για OR: (0.55, 6.3)

Λύση 2: Υπολογισμός X^2 test

X^2 test: 2.696 με 1 df

p-value=0.101

Άρα δεν απορρίπτω την υπόθεση οτι ο βελονισμός είναι το ίδιο αποτελεσματικός με το φάρμακο

Διαστήματα εμπιστοσύνης (confidence intervals)

- Τα μέτρα σχέσης ή εκτιμητές σχέσης (π.χ. OR) δεν αντιπροσώπευουν την αλήθεια, αλλά αποτελούν μόνο μία εκτίμησή της
- Το αληθινό αποτέλεσμα μπορεί να είναι μεγαλύτερο ή μικρότερο από αυτό που παρατηρήσαμε
- Όρισμός 95% CI: εάν επαναλάβουμε ένα πείραμα ή μία μελέτη πολλές φορές (π.χ. 100) και πάρουμε ένα διαφορετικό μέτρο σχέσης και το αντίστοιχο διάστημα εμπιστοσύνης του από κάθε επανάληψη, τότε 95 από τα 100 αυτά διαστήματα εμπιστοσύνης θα περιέχουν την αλήθεια (αληθινό μέτρο σχέσης)
- Το διάστημα εμπιστοσύνης μας λέει μέσα στα όρια της αληθοφάνειας πόσο μεγαλύτερο ή μικρότερο μπορεί να είναι το αληθινό αποτέλεσμα

Μεταβλητή 1/ Μεταβλητή 2	**Συνεχής**	**Διακριτή**
Συνεχής	Συντελεστής συσχέτισης	t-test, paired t-test$ Z-test Mann–Whitney* Wilcoxon*$
Διακριτή		X^2 test Fisher's exact@ McNeamar$ Mantel-Haenszel

*: το τεστ δεν απαιτεί κανονική κατανομή των δεδομένων
$: το τεστ είναι κατάλληλο για ζευγαρωτά δεδομένα
@: το τεστ δεν απαιτεί όλες οι προσδοκόμενες τιμές στα κελιά να ειναι πάνω απο 5