Στατιστικές δοκιμές

Συνεχή δεδομένα

Γεωργία Σαλαντή

Τι θέλουμε να συγκρίνουμε;

- Δύο δείγματα
  – Μέση αρτηριακή πίεση σε δύο ομάδες
  – Πιθανότητα θανάτου με δύο διαφορετικά είδη αντικαταθλιπτικών
- Την μέση τιμή ενός μεγέθους σε ένα δείγμα με μια θεωρητική τιμή
  – Πτώση της LDL χοληστερίνης κάτω από 90ml/lt

Μηδενική υπόθεση

- Μία υπόθεση που θέλουμε να ελέγξουμε
- Αναφέρεται σε κάποια χαρακτηριστικά των δειγμάτων που εξετάζουμε εκφρασμένα σε στατιστικές ποσότητες
- H_0: $\overline{x}_1 = \overline{x}_2$ (η μέση αρτηριακή πίεση είναι ίδια σε δύο ομάδες)
- H_1: $\overline{x}_1 > \overline{x}_2$
- H_1: $\overline{x}_1 < \overline{x}_2$
- H_1: $\overline{x}_1 > \overline{x}_2$ ή $\overline{x}_1 < \overline{x}_2$

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: Wilcoxon, Mann-Witney test

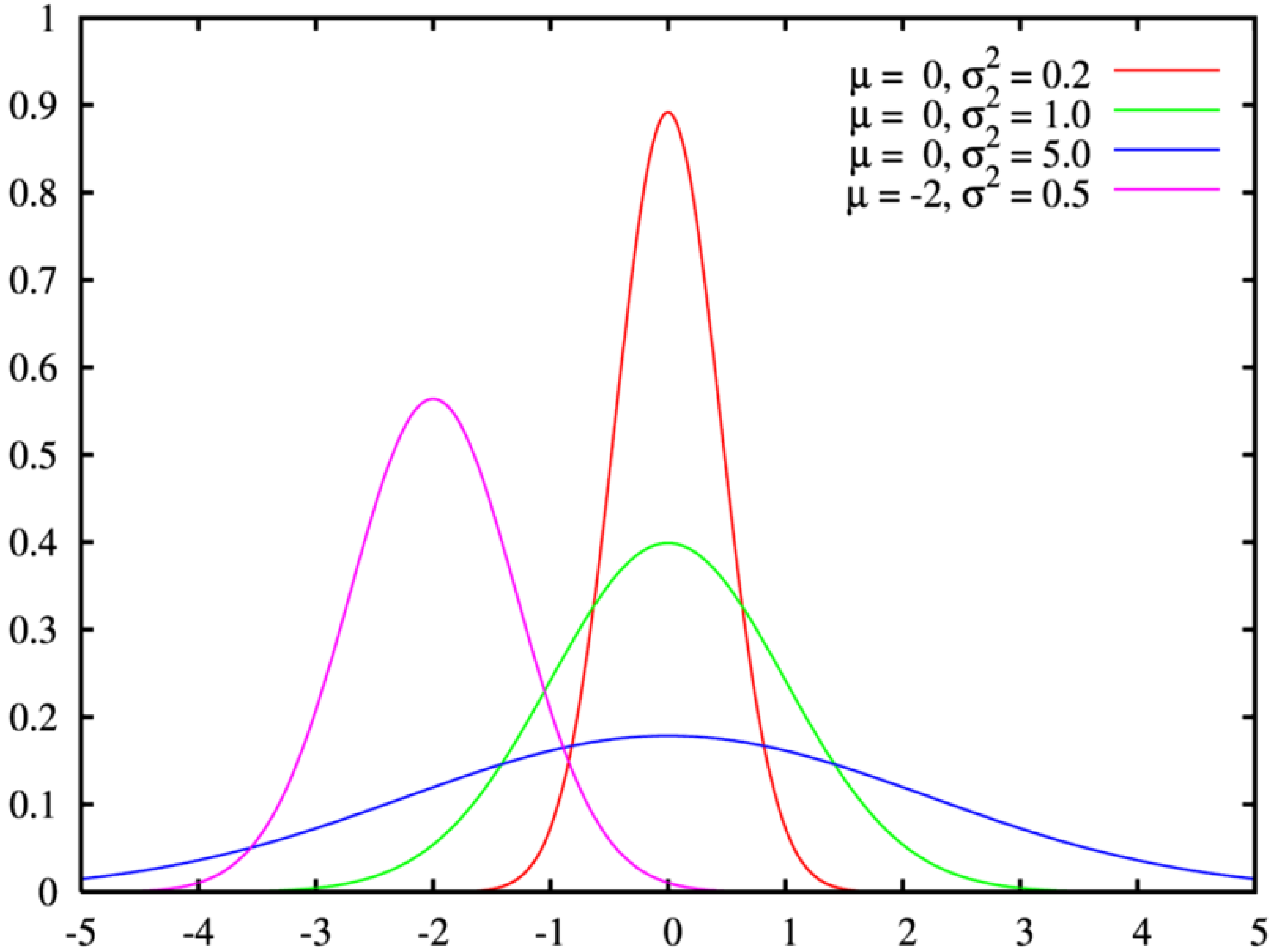
Κατανομές

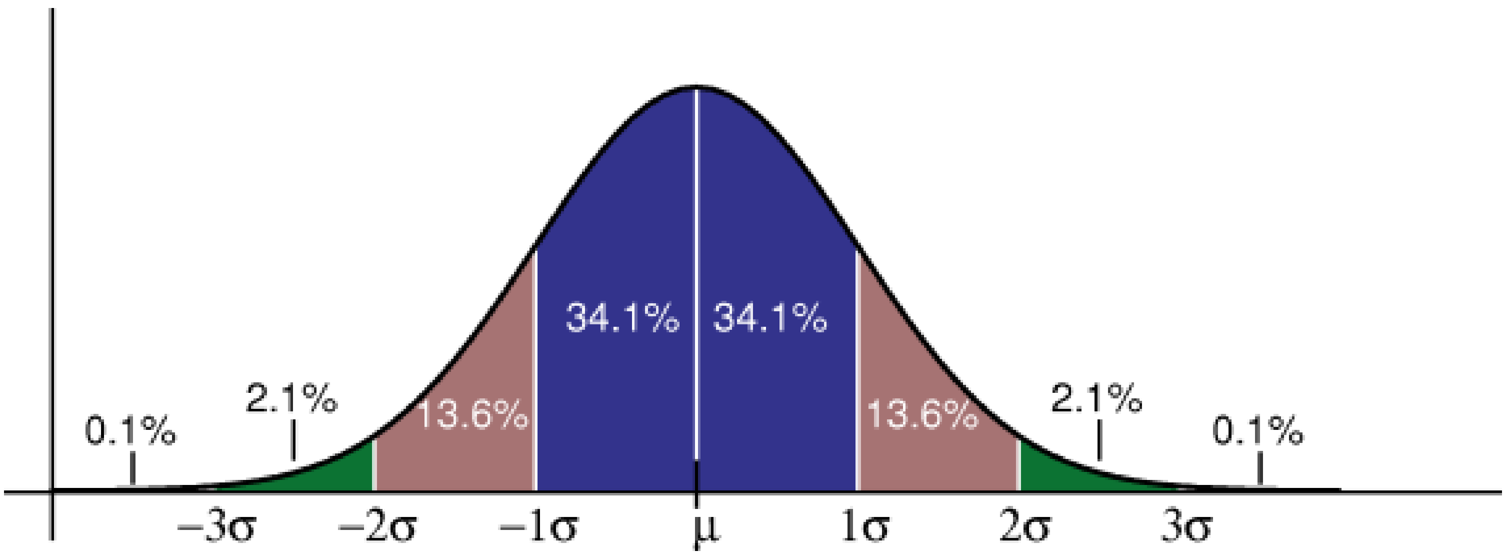
- Το ιστόγραμμα ίσως να μοιάζει με κάποια από τις θεωρητικές κατανομές
- Συνεχείς κατανομές:
  - Κανονική
  - t – student
  - X^2
  - F

Κανονική κατανομή

- $N(\mu,\sigma^2)$: μέσος, διασπορά
- Όσο πιο μεγάλο το σ^2, τόσο πιο ‘απλωτή’ είναι η κατανομή
- Η πρότυπη κανονική κατανομή είναι $N(0,1)$
- Πολλά μεγέθη ακολουθούν την κανονική κατανομή (βάρος, ύψος, πίεση…)

Κανονική κατανομή




34.1%
34.1%
13.6%
13.6%
2.1%
2.1%
0.1%
0.1%
−3σ
−2σ
−1σ
μ
1σ
2σ
3σ

Z-test – δοκιμασία Z

- Μέση τιμή του πληθυσμού μ
- Μέση τιμή του δείγματος $\bar{x}$
- Μέγεθος δείγματος ν
- Διασπορά (variance) σ^2, τυπική απόκλιση σ
- H_0: x=μ
- H_1: x>μ ή x<μ

Z-test – δοκιμασία Z

$$z = \frac{\left|\bar{x} - \mu\right|}{\dfrac{\sigma}{\sqrt{\nu}}}$$

Παράδειγμα

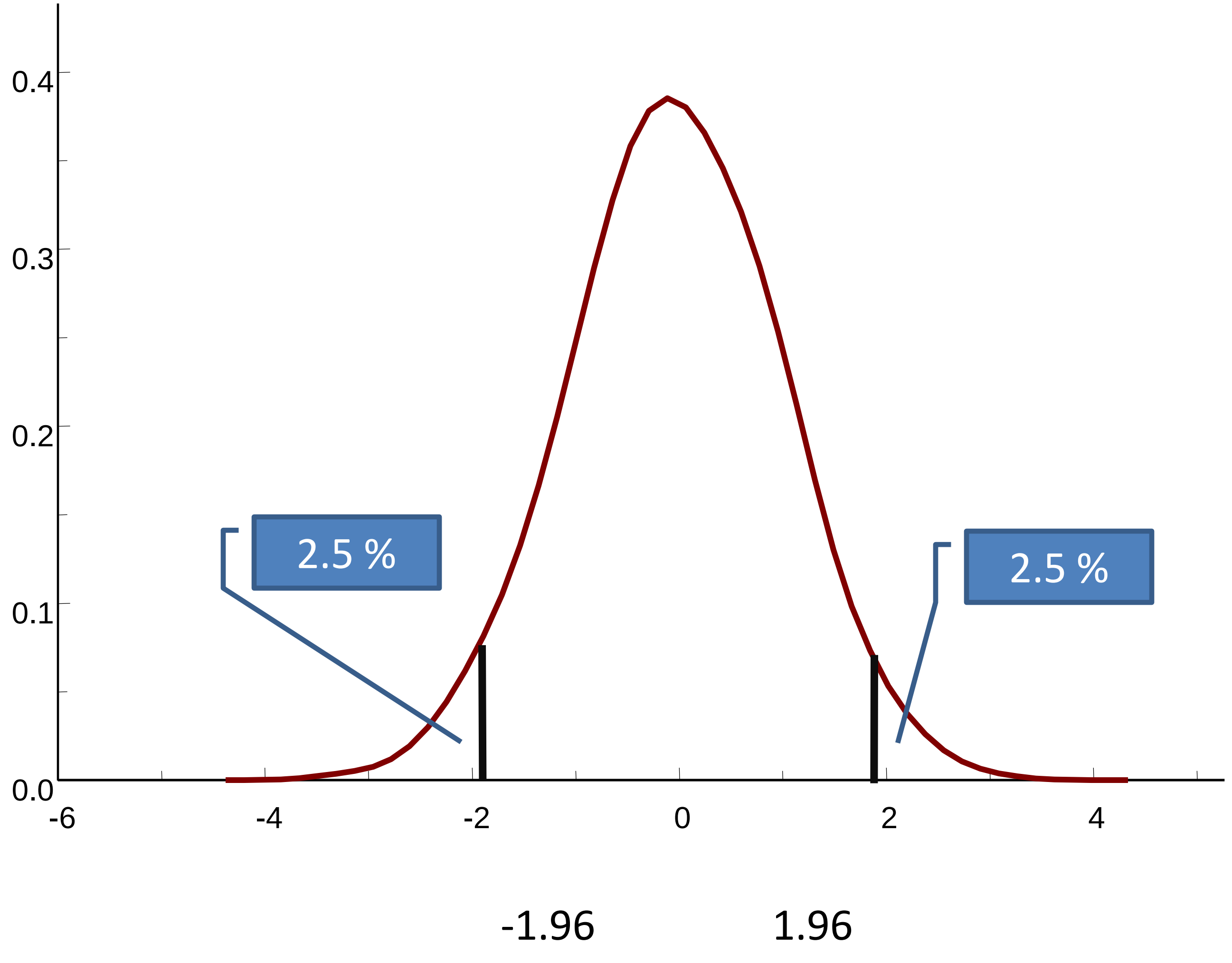
- Σε ένα δείγμα 100 ατόμων 50 ετών μετρήθηκε η συστολική πίεση
- Μέση τιμή (δείγματος) $\overline{\mathrm{X}}$ =126 mmHg
- σ =15 mmHg (υποθέτουμε ότι είναι ίδια στον πληθυσμό και το δείγμα…)
- μ =120
- Είναι πολύ υψηλά αυτά τα επίπεδα;
- H_0: $\overline{\mathrm{X}} = \mu$, η μέση πίεση στο δείγμα δεν διαφέρει από αυτή στον γενικό πληθυσμό

Z-test – δοκιμασία Z

$$z = \frac{126 - 120}{\dfrac{15}{\sqrt{100}}} = 4$$

Το 4 θα το συγκρίνουμε με την πρότυπη κανονική κατανομή

Πρότυπη κανονική κατανομή



P-value

- 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 126;
- π/1000 φορές = p-value
- p-value=το εμβαδόν στην πρότυπη κανονική κατανομή που είναι έξω από το διάστημα (-z,z)

P-value

- Εάν η τιμή του δείγματος δεν είναι στο 5% των πιο ακραίων τιμών, τότε το εύρημα θεωρείται ‘στατιστικά μη σημαντικό’
- 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 126) δεν μπορεί να εξηγηθεί από την τύχη

Πώς υπολογίζουμε p-values

- 2*(1-NORMSDIST(z)) (στο excel)
- 2*(1-NORMSDIST(4)) = 0.00006334248
- Στατιστικά σημαντικό αποτέλεσμα
- Απορρίπτουμε την μηδενική υπόθεση

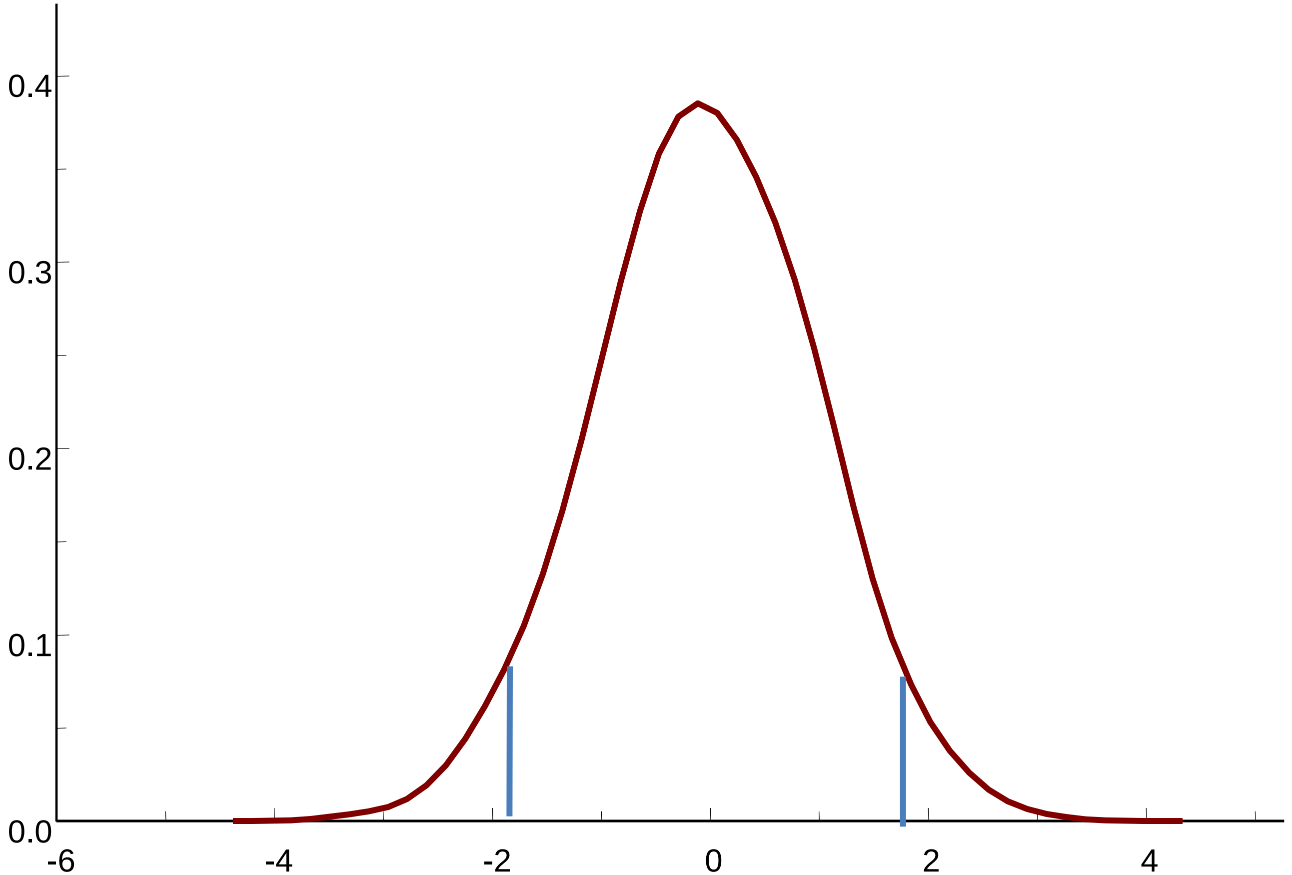
Αμφίπλευρα και μονόπλευρα τεστ

- H_0: x=μ με H_1: x>μ ή x<μ
  z=1.7, p-value=0.09
- H_0: x=μ με H_1: x>μ
  z=1.7, p-value=0.045

Τα αμφίπλευρα τεστ είναι πιο ‘συντηρητικά’

(= δεν απορρίπτουν τόσο εύκολα τη μηδενική υπόθεση)

Πρότυπη κανονική κατανομή



-1.7 1.7

Z-τεστ: Υποθέσεις

- Ότι η κατανομή του που μετράμε είναι όντως κανονική
- Ότι η τυπική απόκλιση σ στον πληθυσμό είναι η ίδια με αυτή στο δείγμα
  - Πχ εδώ σ=15 είναι μάλλον μικρή τ.α.
  - Ποιο θα ήταν το αποτέλεσμα του τεστ αν η τ.α. ήταν μεγαλύτερη;

t-test για ανεξάρτητα δείγματα

- Όταν θέλουμε να συγκρίνουμε δυο μέσους από δύο δείγματα
- Προϋπόθεση: ότι έχουνε τις ίδιες διασπορές και ότι τα δείγματα προέρχονται από κανονική κατανομή
- H_0: $\mu_1=\mu_2$
- H_1: $\mu_1>\mu_2$ (μονόπλευρος έλεγχος)
- H_1 :$\mu_1<\mu_2$ (μονόπλευρος έλεγχος)
- H_1: $\mu_1>\mu_2$ ή $\mu_1<\mu_2$ (αμφίπλευρος έλεγχος)

t-test

$$\boxed{t = \frac{\bar{x}_1 - \bar{x}_2}{s_{\bar{x}_1 - \bar{x}_2}}}$$

$$s_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_1^2 + s_2^2}{\nu_1}}, \nu_1 = \nu_2$$

$$s_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{(\nu_1 - 1)s_1^2 + (\nu_2 - 1)s_2^2}{\nu_1 + \nu_2 - 2}\left(\frac{1}{\nu_1} + \frac{1}{\nu_2}\right)}$$

Παράδειγμα

- μ_1=15, s_1^2 διασπορά$_1$=4.6, ν_1=8
- μ_2=14, s_2^2 διασπορά$_1$=6.6, ν_2=8

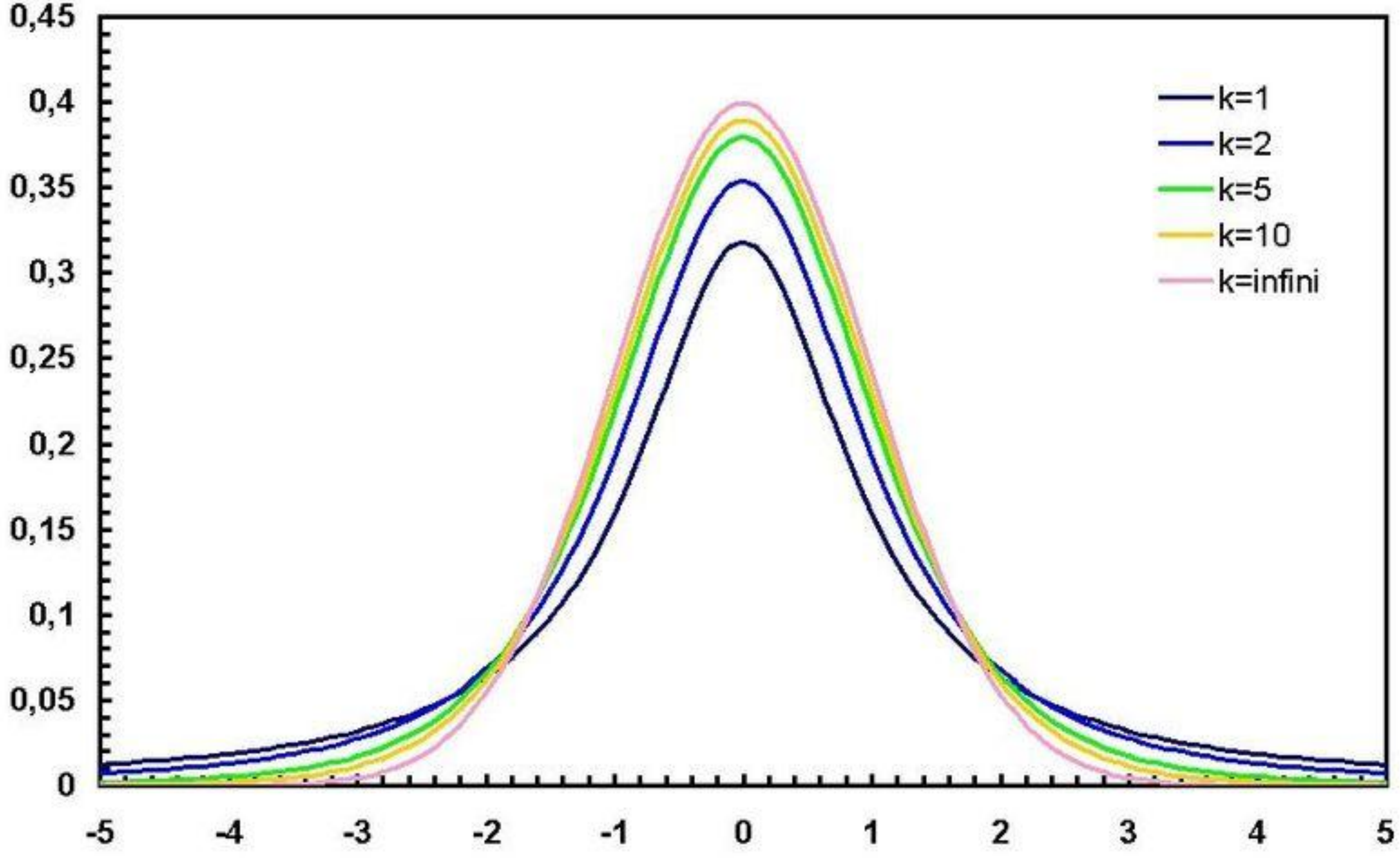
$$t = \frac{\mu_1 - \mu_2}{\sqrt{\dfrac{\mathrm{var}_1 + \mathrm{var}_2}{\nu_1}}} = \frac{15 - 14}{\sqrt{\dfrac{4.6 + 6.6}{8}}} = 0.85$$

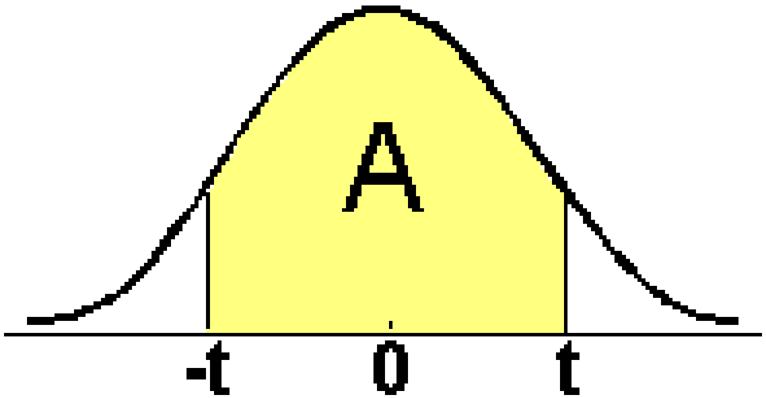
- Πώς βγήκε αυτός ο τύπος;
- Το 0.85 είναι μεγάλο ή μικρό;
- Πάμε στην t κατανομή με ν_1+ν_2-2=14 βαθμούς ελευθερίας

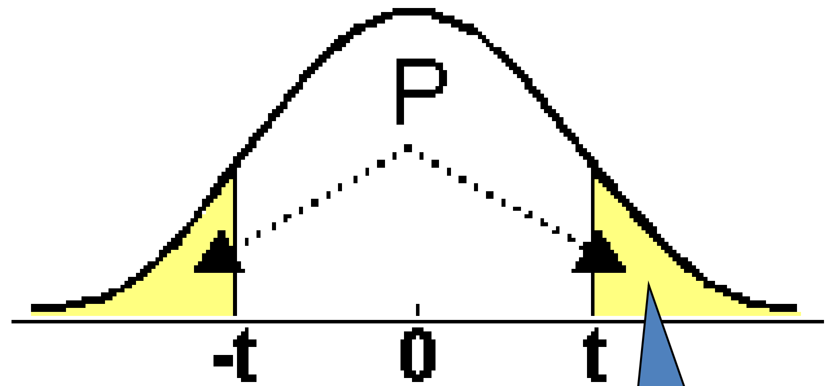
t κατανομή (student)

- Για μεγάλο μέγεθος δείγματος η κατανομή t είναι ίδια με την κανονική
- Μέσος = 0
- Όσο πιο πολλοί βαθμοί ελευθερίας, τόσο πιο «ψιλόλιγνη» είναι

κ οι βαθμοί ελευθερίας που ορίζονται από το μέγεθος του συνολικού δείγματος







P-value = TDIST(t,df,1)=

=TDIST(0.85,14,1)= 0.41

Τι σημαίνει αυτό;

Δεν απορρίπτουμε την H_0

Εμβαδον=0.205

T-τεστ και Z-test

$$t = \frac{\overline{x}_1 - \overline{x}_2}{s_{\overline{x}_1 - \overline{x}_2}}$$

$$s_{\overline{x}_1 - \overline{x}_2} = \sqrt{\frac{s_1^2 + s_2^2}{\nu_1}}, \nu_1 = \nu_2$$

$$\boxed{Z = \frac{\kappa\acute{\alpha}\tau\iota}{SE(\kappa\alpha\tau\iota)} \sim N(0,1)}$$

$$MD = \bar{x}_1 - \bar{x}_2$$

$$SE(MD) = \sqrt{\frac{s^2{}_1 + s^2{}_2}{\nu_1}} \text{ όταν } \nu_1 = \nu_1$$

t-test: υποθέσεις

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

Παράδειγμα

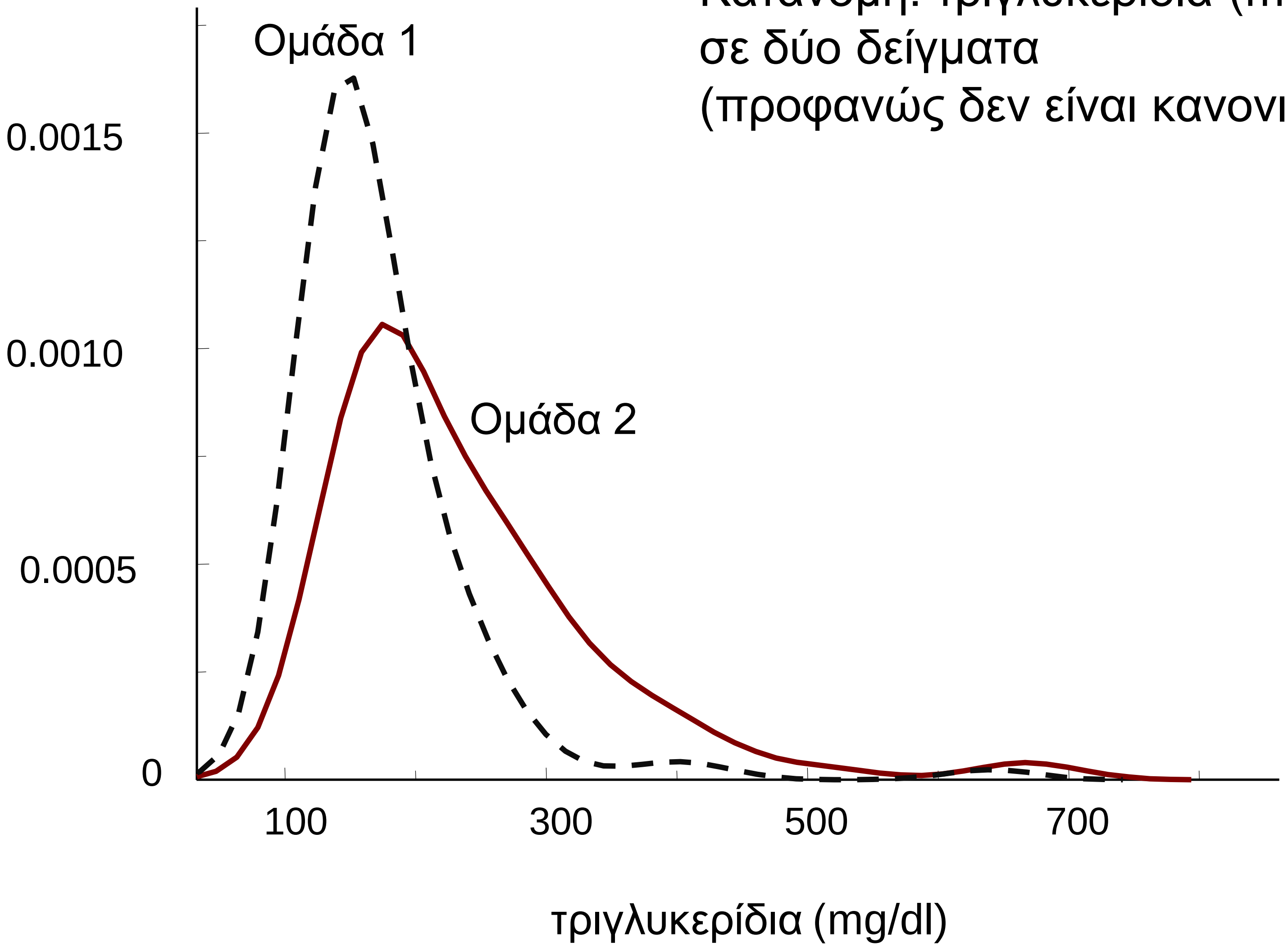
- $μ_1$=15, s^2 διασπορά$_1$=4.6, $ν_1$=8
- $μ_1$=14, s^2 διασπορά$_1$=6.6, $ν_2$=8

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

Wilcoxon, Mann-Witney test

- Όταν τα δεδομένα δεν προέρχονται από κανονική κατανομή
  - Τριγλυκερίδια
- Η μέση τιμή δεν αντιπροσωπεύει επιτυχώς τα δεδομένα - καλύτερη η διάμεσος
- H_0: τα δύο δείγματα προέρχονται από την ίδια κατανομή ($\delta_1=\delta_2$)

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
- Wilcoxon: p-value=0.14

Ζευγαρωτές παρατηρήσεις

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

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	
		H_0 σωστή	H_0 λάθος
τεστ	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

Υπολογισμός μεγέθους δείγματος

- Πρέπει από πριν να καθορίσουμε
  - Το τεστ που θα χρησιμοποιήσουμε
  - Την προσδοκώμενη διαφορά ανάμεσα στις δύο ομάδες
  - Την ισχύ που θέλουμε (X%, πχ 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 X%

Παράδειγμα

- Από μια ΤΚΔ που τυχαιοποίησε τους ασθενείς σε 2 ομάδες των 100 ατόμων προέκυψαν τα παρακάτω αποτελέσματα
- Δίαιτα 1, μέσο βάρος 80 κιλά, τ.α 30
- Δίαιτα 2, μέσο βάρος 83 κιλά, τ.α 31

Ποια δίαιτα είναι στατιστικά καλύτερη; Απαντήστε την ερώτηση με δύο τρόπους

95% ΔΕ για τους δύο μ.ο.

- 95% ΔΕ για δίαιτα 1

(74.12, 85.88)

- 95% ΔΕ για δίαιτα 2

(76.92, 89.08)

Mean difference

$$MD = \bar{x}_1 - \bar{x}_2$$

$$SE(MD) = \sqrt{\frac{s^2{}_1 + s^2{}_2}{\nu_1}} \ \text{όταν}\ \nu_1 = \nu_2$$

$$SE(MD) = \sqrt{\frac{s^2{}_1}{\nu_1} + \frac{s^2{}_2}{\nu_2}}$$

95% ΔΕ για την διαφορά δύο μέσων

- MD=-3
- 95% ΔΕ (-11.46, 5.46)

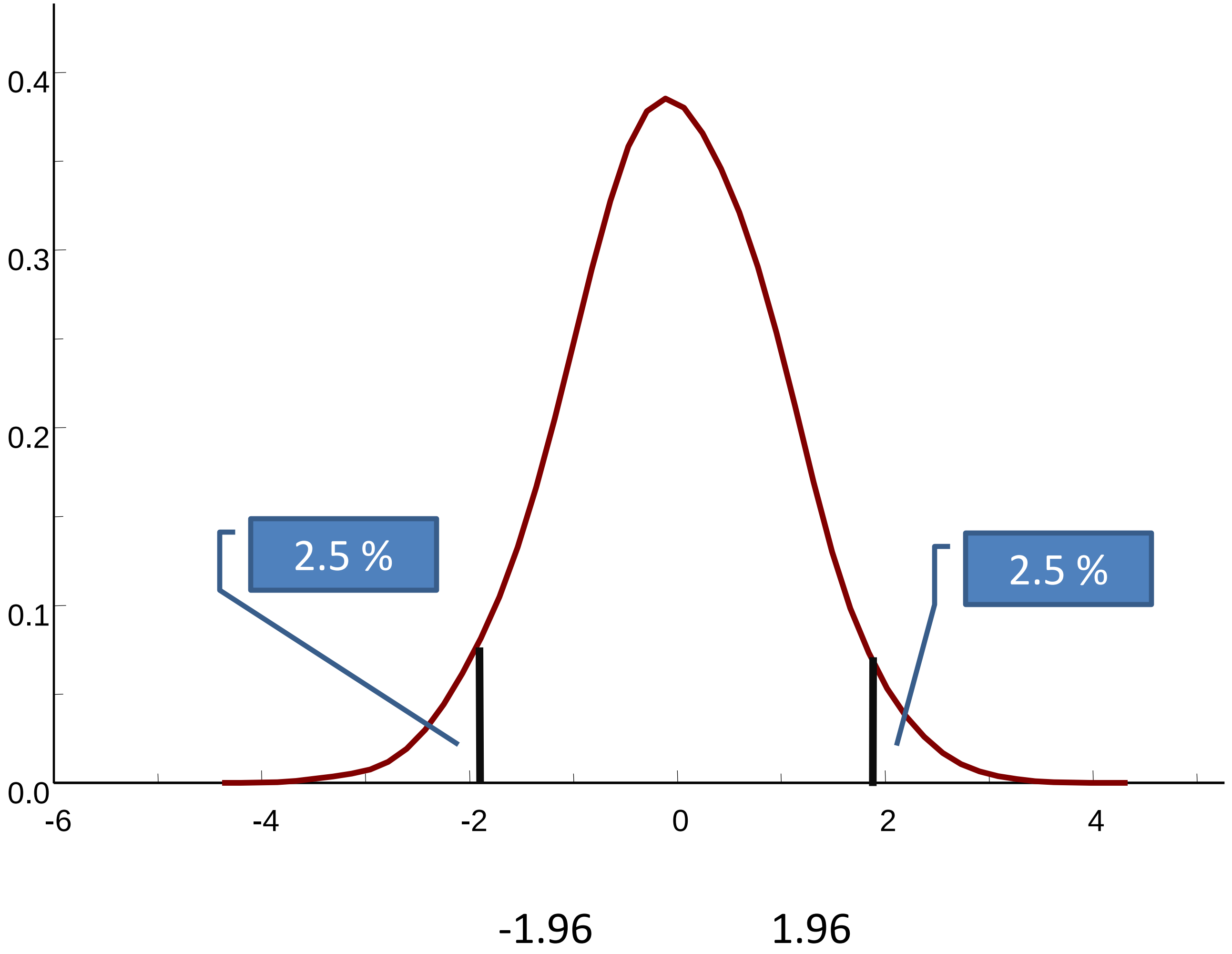
t-test

- t=-0.70
- Df=198
- Άρα η t κατανομή είναι σχεδόν η κανονική κατανομή
- Άρα p-value>0.05

Z-test

- MD=3
- SE(MD)=4.31
- Z=0.70
- Z<1.96 άρα p-value>0.05

Πρότυπη κανονική κατανομή



Λόγος κινδύνων και λόγος αναλογιών

$$\text{ΛΚ} = \frac{\text{\# Κίνδυνος Ομάδα 1}}{\text{\# Κίνδυνος Ομάδα 2}}$$

Risk Ratio (RR)

$$\text{ΛΑ} = \frac{\text{\# Αναλογία Ομάδα 1}}{\text{\# Αναλογία Ομάδα 2}}$$

Odds Ratio (OR)

	Γεγονός	Όχι γεγονός	Συνολο
Νέα θεραπεία	*a*	*b*	*a*+*b*
Παλιά θεραπεία	*c*	*d*	*c*+*d*
	a+*c*	*b*+*d*	*n*

$$\text{ΛΚ} = \frac{a\,(c+d)}{c\,(a+b)}$$

$$\text{ΛΑ} = \frac{a\,d}{c\,b}$$

Risk ratio: λόγος κινδύνων

$$\ln RR = \ln\left(\frac{a/(a+b)}{c/(c+d)}\right) = \ln\left(\frac{a(c+d)}{c(a+b)}\right)$$

$$\operatorname{var}(\ln RR) = \frac{1}{a} - \frac{1}{a+b} + \frac{1}{c} - \frac{1}{c+d}$$

$$SE(\ln(RR)) = \sqrt{\operatorname{var}(\ln(RR))}$$

Odds Ratio: Λόγος αναλογιών

$$\ln \mathrm{OR} = \ln\left(\frac{a/b}{c/d}\right) = \ln\left(\frac{ad}{bc}\right)$$

$$\mathrm{var}\left(\ln \mathrm{OR}\right) = \frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}$$

$$SE(\ln(OR)) = \sqrt{\mathrm{var}(\ln(OR))}$$