Σύγκριση μέσου όρου πληθυσμού με τιμή ελέγχου

One-Sample *t*-Test

Μια σύντομη αναδρομή

Στα τέλη του $19^{ου}$ αιώνα μια μεγάλη αλλαγή για την επιστήμη ζυμώνονταν στην ζυθοποιία Guinness.

Ο William Gosset ανακάλυψε μια νέα μέθοδο για να εκτιμήσει τη δοκιμασία γεύσης της μπύρας από ένα μικρό δείγμα δοκιμαστών.

Το αποτέλεσμα ήταν μια στατιστική μέθοδος που άλλαξε την επιστήμη – και ίσως την μπύρα

- Το 1908, ο Gosset δημοσίευσε τη μέθοδό του στο περιοδικό Biometrika με το ψευδώνυμο «μαθητής» (‘student’). Και για αυτό συχνά η μέθοδος λέγεται και ‘student’s t test’.

Πότε χρησιμοποιούμε αυτόν το στατιστικό έλεγχο;

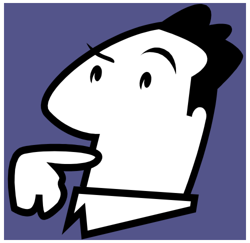
Όταν θέλουμε να ελέγξουμε εάν δυο σύνολα δεδομένων διαφέρουν σημαντικά μεταξύ τους.

Πως δουλεύει το one-sample t-test για ΈΝΑ σύνολο?

- Ελέγχει εάν ο μέσος όρος του πληθυσμού από τον οποίο πήραμε δείγμα διαφέρει από μία τιμή ελέγχου.
  - Ποια είναι η τιμή ελέγχου;
- Η τιμή ελέγχου συνήθως καθορίζεται από την πρότερη γνώση και από το ερώτημα που καλούμαστε να απαντήσουμε.

Παράδειγμα

- Θεωρητικά, τα τεστ ευφυΐας (που μετρούν το IQ) κατασκευάζονται έτσι ώστε ο μέσος όρος μεταξύ των ενηλίκων να είναι **100** μονάδες.
- Έστω ότι παίρνουμε ένα δείγμα έξι προπτυχιακών φοιτητών του πανεπιστημίου μας (N = 6), και ελέγχουμε εάν ο μέσος όρος του IQ των φοιτητών στο πανεπιστήμιο είναι υψηλότερος από 100.
- Με απλά λόγια, είναι οι φοιτητές μας εξυπνότεροι από το μέσο άνθρωπο;



- Έστω ότι οι τιμές IQ των έξι φοιτητών είναι

Φοιτητής 1: 110
Φοιτητής 2: 118
Φοιτητής 3: 110
Φοιτητής 4: 122
Φοιτητής 5: 110
Φοιτητής 6: 150

Ερευνητικό ερώτημα

Κατά μέσο όρο, ο πληθυσμός των προπτυχιακών φοιτητών του Πανεπιστημίου μας έχουν υψηλότερο δείκτη ευφυϊας από το μέσο άνθρωπο (IQ = 100);

- Πρώτα πρέπει να υπολογίσουμε το μέσο δείκτη ευφυϊας του δείγματος μας

$$\overline{X} \qquad \frac{\sum X}{n} = \frac{110+118+110+122+110+150}{6} = 120$$

- Προφανώς αυτοί οι έξι έχουν μέσο όρο IQ υψηλότερο από 100, αλλά τι ισχύει στον πληθυσμό;

- Πόσο πιθανό είναι αυτό το αποτέλεσμα;
- Θα ίσχυε εάν είχαμε πάρει 6 διαφορετικούς φοιτητές;
- Τι θα συνέβαινε εάν παίρναμε πολλά σύνολα από 6 φοιτητές;

Το one-sample *t*-test θα μας απαντήσει αυτά τα ερωτήματα

Θα μας πει εάν το αποτέλεσμα είναι ‘σημαντικό’, ή με άλλα λόγια εάν είναι πιθανό να επαναληφθεί σε άλλο δείγμα του αυτού πληθυσμού.

Υπολογίζουμε τη διακύμανση των τιμών του δείγματος

IQ	Μέσος όρος	$X-\overline{X}$	$(X-\overline{X})^2$
110	120	-10	100
118	120	- 2	4
110	120	-10	100
122	120	2	4
110	120	-10	100
150	120	30	900

$$s_x^2 = \frac{\sum (X-\overline{X})^2}{n} = 201.\overline{3}$$

Για να βρούμε πόσο σημαντικά απέχει ο μέσος όρος από την τιμή ελέγχου χρειαζόμαστε ένα μέτρο να μετρήσουμε αυτή την απόσταση

$$\overline{\mathrm{X}} - \mu$$

Αλλά πρέπει να σταθμίσουμε αυτή την απόσταση με τη διακύμανση του φαινομένου που μελετάμε
Αυτό το σταθμισμένο μέτρο είναι η ποσότητα t

$$t = \frac{\bar{X} - \mu}{\sqrt{\frac{s_x^2}{n-1}}}$$

Υπολογίζοντας το t στο παράδειγμα

$$t = \frac{\bar{X} - \mu}{\sqrt{\frac{s_x^2}{n-1}}} = \frac{120 - 100}{\sqrt{\frac{201.333}{6-1}}} \approx 3.152$$

Έχουμε μετρήσει 6 άτομα άρα οι βαθμοί ελευθερίας είναι $df = n - 1 = 5$

Οι βαθμοί ελευθερίας είναι σημαντικοί για τον υπολογισμό των πιθανοτήτων στο επόμενο βήμα

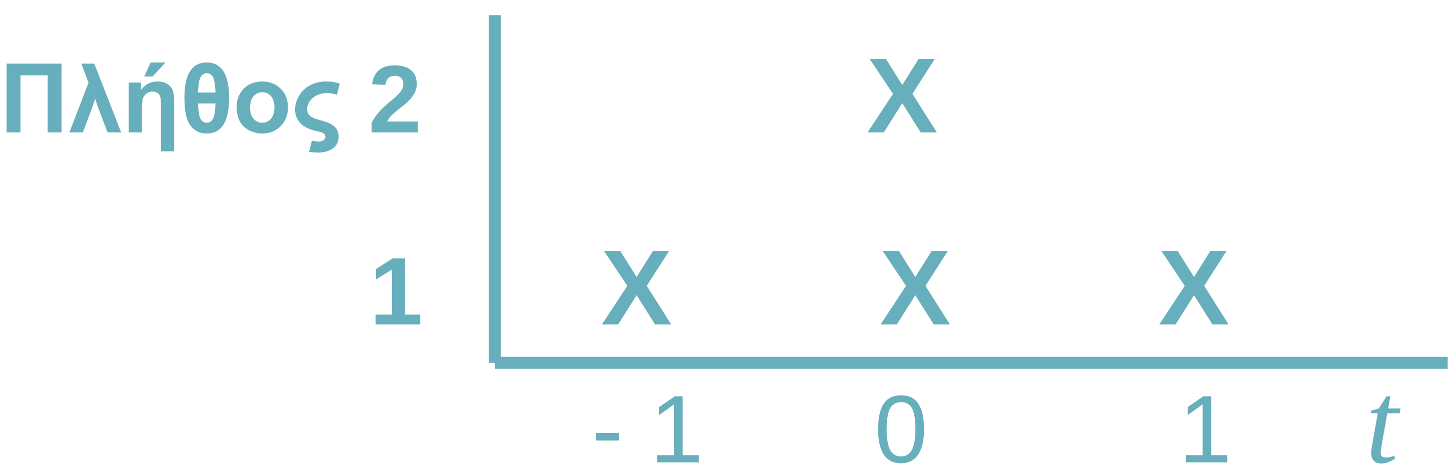
Και τώρα για κάτι διαφορετικό

Τώρα ας κάνουμε ένα διάλειμμα από τους τύπους και ας σκεφθούμε την ευρύτερη εικόνα.

Και αυτό είναι το δύσκολο μέρος.

Η κατανομή t

- Όπως οι καταγραφές, έτσι και η ποσότητα ‘*t*’ ακολουθεί μια κατανομή.
- Εάν όντως η μέση τιμή IQ του πληθυσμού είναι ακριβώς 100, και εμείς πάρουμε πολλαπλά δείγματα 6 ατόμων, και υπολογίσουμε την τιμή t τότε τι περιμένουμε;
- Για παράδειγμα έστω 4 επαναλήψεις



Η κατανομή t

- Η πιο συχνή τιμή του ‘***t***’ θα είναι 0. Γιατί; Γιατί ο αριθμητής στον τύπο υπολογισμού του ***t*** υπολογίζει τη διαφορά του μέσου όρου του δείγματος από την τιμή ελέγχου.
- Και εάν ο μέσος όρος είναι όντως ίσως με την τιμή ελέγχου, τότε τα περισσότερα δείγματα θα έχουν αυτόν το μέσο όρο.

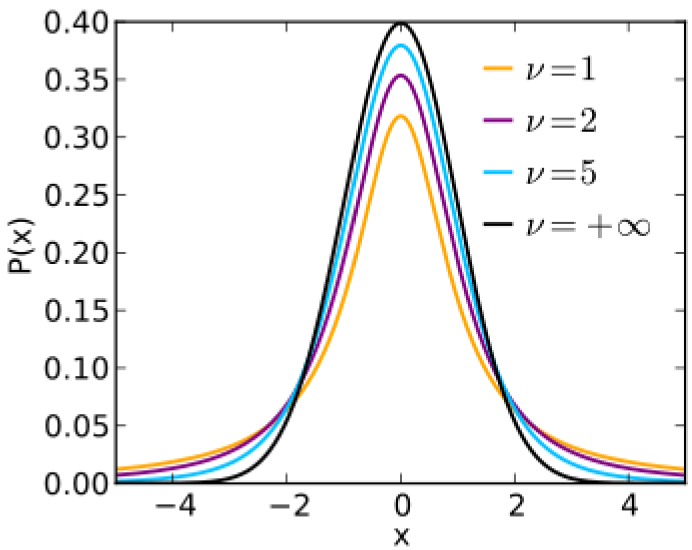
$$t = \frac{100 - 100}{\sqrt{S_x^2 / n-1}} = 0$$

Η κατανομή t

- Φυσικά πολλά δείγματα 6 ατόμων δε θα έχουν μέσο όρο ίσο με τον μέσο όρο του πληθυσμού
- Κάποια θα έχουν μεγαλύτερο μέσο όρο ($t>0$) και κάποια μικρότερο ($t<0$).
- Η συχνότητα με την οποία βρίσκουμε τιμές t μεγαλύτερες από 0 (ή μικρότερες από 0) είναι ακριβώς αυτό που εκτιμάει η κατανομή t.

Η κατανομή t

$$f(t) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\,\Gamma(\frac{\nu}{2})}\left(1+\frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \qquad \Gamma(t) = \int_0^\infty x^{t-1}e^{-x}\,dx$$



Και τώρα προσοχή αυτό είναι το κρίσιμο μέρος

- Στο παράδειγμά μας το t είναι 3,152. Με άλλα λόγια είναι μεγαλύτερο από 0.
- Άρα δε φαίνεται το δείγμα μας να προήλθε από πληθυσμό με μέσο IQ 100.
- Εάν είχε έρθει θα περιμέναμε τιμή t κοντά στο 0.

- *Αλλά εάν είχε έρθει από τέτοιο πληθυσμό πόσο πιθανό θα ήταν να είχε καταγράψει τέτοια τιμή;*

Και τώρα προσοχή αυτό είναι το κρίσιμο μέρος

- Πως μπορούμε να υπολογίσουμε αυτή την πιθανότητα;
- Η κατανομή t μας εξυπηρετεί γιατί ακριβώς υπολογίζει πόσο συχνά θα παρατηρηθεί μια συγκεκριμένη τιμή **t** σε έναν πληθυσμό δειγμάτων με **μέση τιμή t = 0**.

Η κατανομή t

- Η κατανομή t που θα πρέπει να χρησιμοποιήσουμε εξαρτάται από τους βαθμούς ελευθερίας (*df* = 5).
- Γιατί;
  - Εάν μετρήσουμε μόνο λίγα άτομα, υπάρχει πιθανότητα να πετύχουμε ανάμεσα σε αυτά κάποια ακραία τιμή που θα επηρεάσει έντονα το δειγματικό μέσο όρο και άρα θα προκύψει μεγάλη τιμή της μετρικής t.
  - Γενικά, **μικρά δείγματα είναι λιγότερο αντιπροσωπευτικά του πληθυσμού.**

Και επίσης κρίσιμο

- Πότε θεωρούμε μια διαφορά **σημαντική**;
- Η πιο απλά πότε μπορούμε να απαντήσουμε το ερευνητικό μας ερώτημα με ναι ή όχι;

Ερευνητικό ερώτημα

Κατά μέσο όρο, ο πληθυσμός των προπτυχιακών φοιτητών του Πανεπιστημίου μας έχουν υψηλότερο δείκτη ευφυϊας από το μέσο άνθρωπο (IQ = 100);

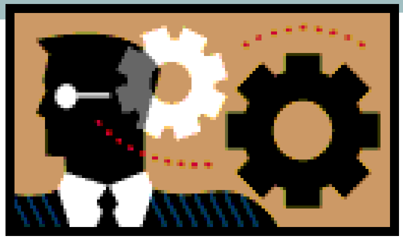
Μονόπλευρος ή αμφίπλευρος έλεγχος

- Το ερώτημα μας είναι διατυπωμένο ως μονόπλευρος έλεγχος, γιατί ρωτάει εάν ο μέσος όρος είναι μεγαλύτερος από την τιμή ελέγχου.
- Εάν το ερώτημα ήταν αν **διαφέρει** ο μέσος όρος από την τιμή ελέγχου (δηλαδή είναι είτε μικρότερος είτε μεγαλύτερος) τότε θα μιλήσουμε για τον αμφίπλευρο έλεγχο που είναι και η πιο συνηθισμένη περίπτωση
- Στην περίπτωση του μονόπλευρου ελέγχου για 5 βαθμούς ελευθερίας η κρίσιμη τιμή του t είναι **2,02.**

Απάντηση στο παράδειγμα

Στο παράδειγμά μας $t = 3{,}152$
άρα μεγαλύτερο από την κρίσιμη τιμή 2,02.
Άρα η διαφορά είναι σημαντική και μπορούμε να απαντήσουμε το ερευνητικό ερώτημα θετικά.

Ή με πιο τεχνικούς όρους μπορούμε να απορρίψουμε τη μηδενική υπόθεση ότι το δείγμα μας προέρχεται από πληθυσμό με μέσο όρο 100.



Μηδενική υπόθεση

- Στη στατιστική δοκιμασία που περιγράψαμε υπάρχει μια μηδενική και μια εναλλακτική υπόθεση.
- Η μηδενική υπόθεση ήταν ότι ο μέσος όρος του πληθυσμού είναι ίσος ή μικρότερος από την τιμή ελέγχου.
- Η εναλλακτική υπόθεση ήταν ότι ο μέσος όρος του πληθυσμού είναι μεγαλύτερος από την τιμή ελέγχου.

Τι μας είπε η κατανομή t

- Η κατανομή t μας είπε ότι στο 95% των δειγμάτων (μεγέθους 6 ατόμων) που θα πάρουμε από πληθυσμό με μέση ευφυϊα 100 η τιμή t θα είναι μικρότερη από 2,02

- Και μόνο σπάνια ($p<0{,}05$) θα παρατηρούνται τιμές t μεγαλύτερες του 2,02

Άρα τι σημαίνει ότι εμείς είχαμε t=3.152

- Άρα με πιθανότητα σφάλματος μικρότερη από 5% ($p<0.05$) μπορούμε να ισχυριστούμε ότι το δείγμα μας δεν προήλθε από πληθυσμό με μέση ευφυΐα 100.
- Με βάση το ερώτημα μπορούμε να ισχυριστούμε ότι προήλθε από πληθυσμό με μέση ευφυΐα μεγαλύτερη από 100.
- Πόσο συχνά θα είναι εσφαλμένος αυτός ο ισχυρισμός;
- Αυτό είναι το επίπεδο στατιστικής σημαντικότητας

Πως το διατυπώνουμε αυτό επιστημονικά;

Ο μέσος όρος του δείκτη ευφυΐας των φοιτητών του πανεπιστημίου μας είναι σημαντικά μεγαλύτερος από τον αναμενόμενο μέσο όρο (τιμή ελέγχου 100) σύμφωνα με τη μονόπλευρη στατιστική δοκιμασία t για έναν πληθυσμό ($t(5)=3{,}15$ $p=0{,}012$).

Στατιστική σημαντικότητα

- Σύμφωνα με την επιστημονική παράδοση του εικοστού αιώνα, τιμές p οι οποίες είναι χαμηλότερες από 0,05 θεωρούνται σημαντικές.

Παράδειγμα

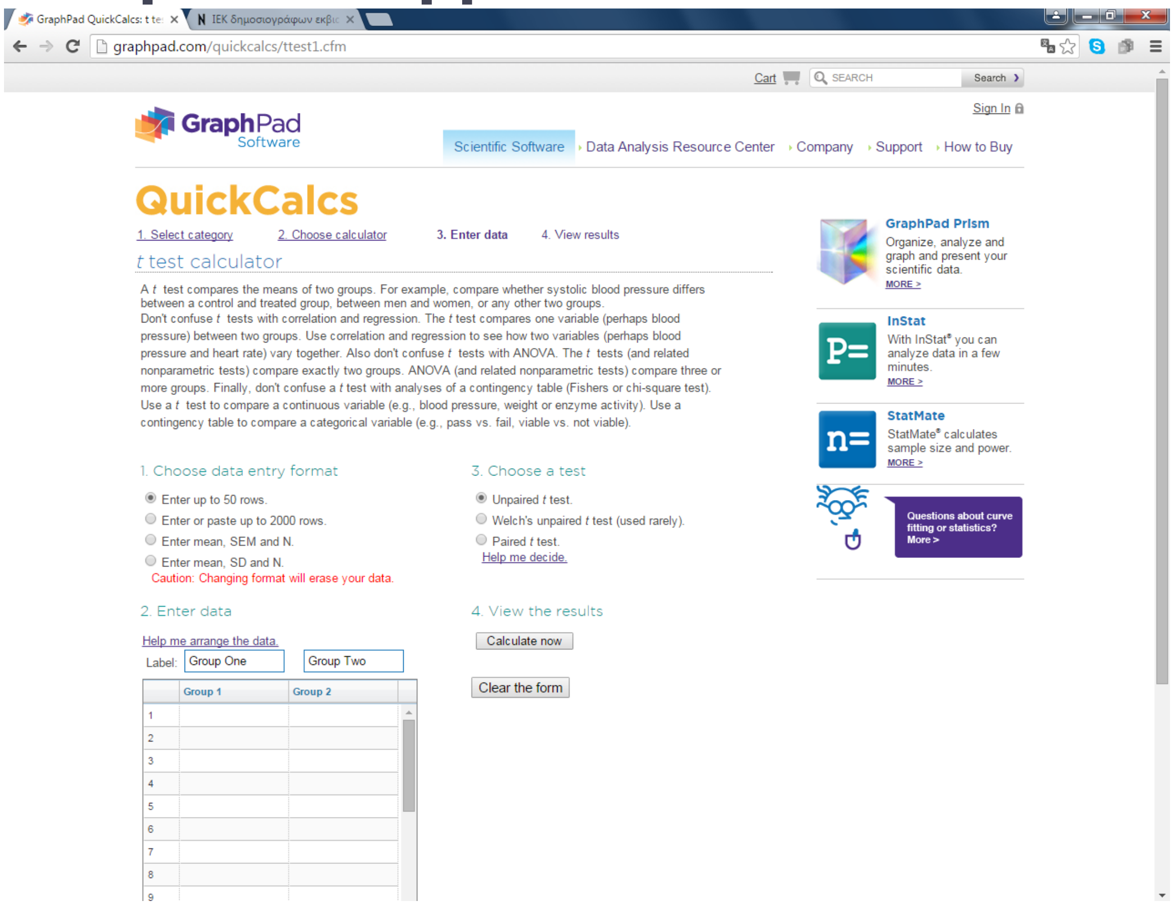
One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Intelligence Quotient	6	120.0000	15.54349	6.34560

One-Sample Test

	Test Value = 100					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Intelligence Quotient	3.152	5	.025	20.00000	3.6881	36.3119

Παράδειγμα


graphpad.com/quickcalcs/ttest1.cfm
Cart
SEARCH
Search
Sign In
GraphPad Software
Scientific Software
Data Analysis Resource Center
Company
Support
How to Buy
QuickCalcs
1. Select category
2. Choose calculator
3. Enter data
4. View results
t test calculator
A t test compares the means of two groups. For example, compare whether systolic blood pressure differs between a control and treated group, between men and women, or any other two groups.
Don't confuse t tests with correlation and regression. The t test compares one variable (perhaps blood pressure) between two groups. Use correlation and regression to see how two variables (perhaps blood pressure and heart rate) vary together. Also don't confuse t tests with ANOVA. The t tests (and related nonparametric tests) compare exactly two groups. ANOVA (and related nonparametric tests) compare three or more groups. Finally, don't confuse a t test with analyses of a contingency table (Fishers or chi-square test).
Use a t test to compare a continuous variable (e.g., blood pressure, weight or enzyme activity). Use a contingency table to compare a categorical variable (e.g., pass vs. fail, viable vs. not viable).
1. Choose data entry format
Enter up to 50 rows.
Enter or paste up to 2000 rows.
Enter mean, SEM and N.
Enter mean, SD and N.
Caution: Changing format will erase your data.
3. Choose a test
Unpaired t test.
Welch's unpaired t test (used rarely).
Paired t test.
Help me decide.
2. Enter data
Help me arrange the data.
Label: Group One Group Two
Group 1 Group 2
4. View the results
Calculate now
Clear the form
GraphPad Prism
Organize, analyze and graph and present your scientific data.
MORE >
InStat
With InStat® you can analyze data in a few minutes.
MORE >
StatMate
StatMate® calculates sample size and power.
MORE >
Questions about curve fitting or statistics? More >

Στατιστική συμπερασματολογία

- Καταλήξαμε σε ένα συμπέρασμα για τον πληθυσμό των φοιτητών στηριζόμενοι σε ένα μικρό δείγμα.

Τελειώσαμε;

Τελειώσαμε;

- ΟΧΙ
- Κάθε στατιστική δοκιμασία ισχύει μόνο κάτω από κάποιες προϋποθέσεις.

Προϋποθέσεις εφαρμογής του one sample t-test

- Αντιπροσωπευτικό δείγμα (αν το δείγμα είναι στρεβλό, κάθε συμπέρασμα είναι αμφίβολο)
- Κανονική κατανομή (η κατανομή t προκύπτει για επαναλαμβανόμενη δειγματοληψία σε πληθυσμό που ακολουθεί την κανονική κατανομή)