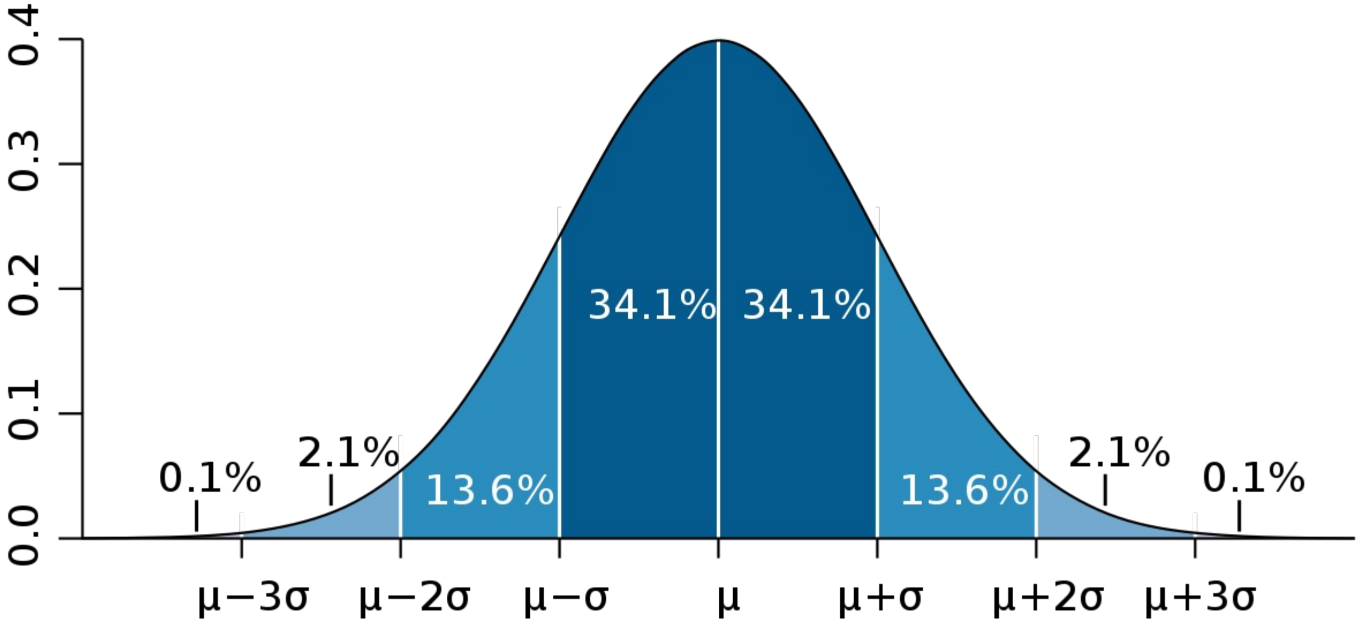
Θεωρία Πιθανοτήτων και Στατιστική



Διδάσκων: Επαμεινώνδας Διαμαντόπουλος
Επικοινωνία: epdiaman@ee.duth.gr

9η διάλεξη

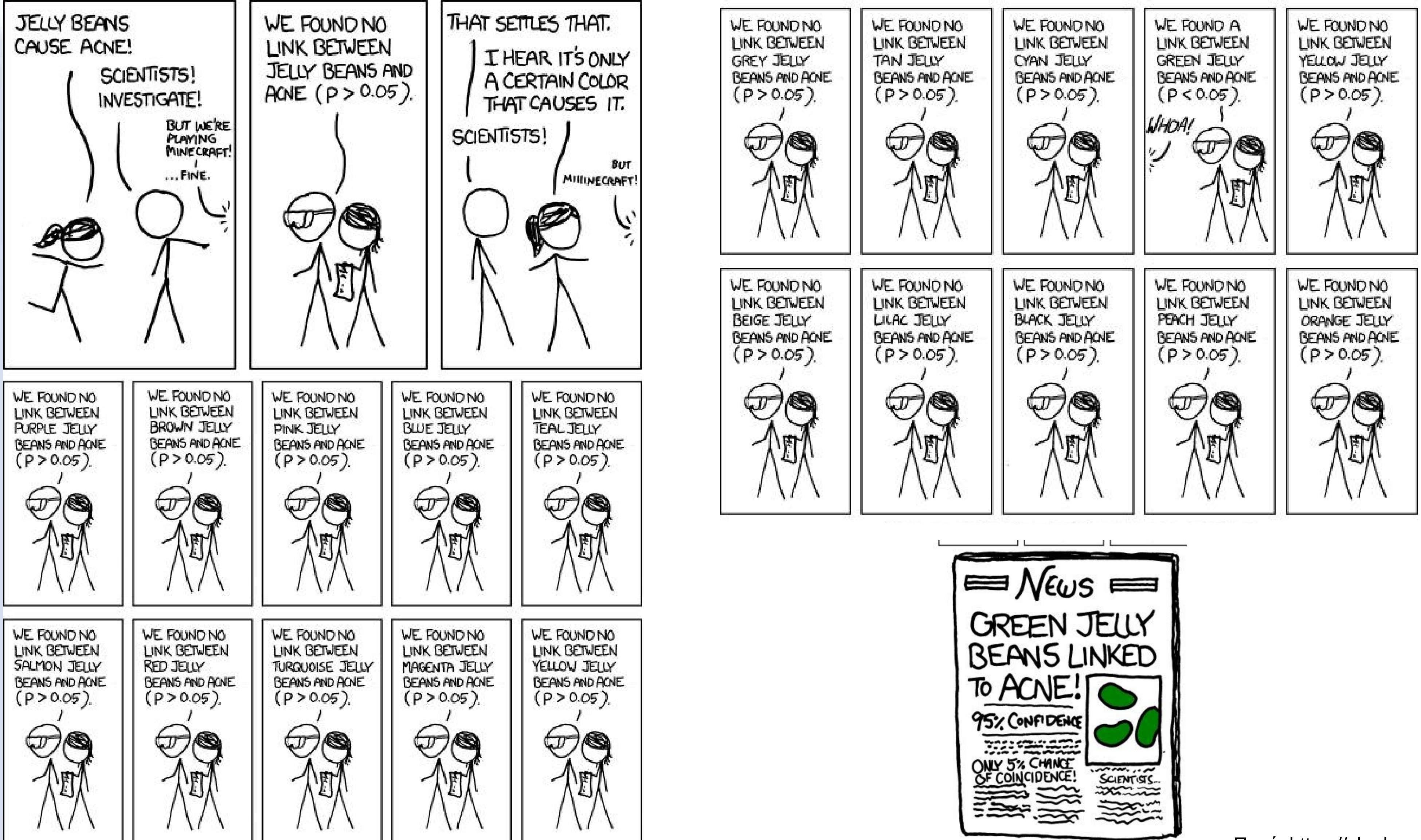
Περιεχόμενα 9ου μαθήματος

- Δοκιμασία χ^2 (συνέχεια)
- Δοκιμασία t – test για μία μέση τιμή.
- Δοκιμασία t – test για μία αναλογία.
- Δοκιμασία t – test για δύο ανεξάρτητα δείγματα (μέσες τιμές)
- Δοκιμασία t – test για δύο ανεξάρτητα δείγματα (αναλογίες)
- Δοκιμασία t – test για ζεύγη παρατηρήσεων.

Γνωστικοί στόχοι μαθήματος

Ο φοιτητής θα πρέπει:

- Να γνωρίζει τα βήματα μίας στατιστικής έρευνας και να αντιλαμβάνεται την ηθική σειρά των βημάτων που πρέπει να ακολουθήσει.
- Να γνωρίζει τις προϋποθέσεις της δοκιμασίας χι-τετράγωνο και του t-test.
- Να αντιλαμβάνεται τη διαφορά μεταξύ ενός t-test και ενός z-test.
- Να μπορεί να καταγράψει την στατιστική και την ερευνητική υπόθεση μίας δοκιμασίας χι-τετράγωνο ή ενός t-test
- Να μπορεί να υλοποιήσει τους υπολογισμούς μίας δοκιμασίας.



Πηγή: https://xkcd.com/882/

“If you torture data enough it will confess”

Ronald Coase (1910 – 2013)

Δοκιμασία χι-τετράγωνο ως έλεγχος προσαρμογής

Έστω ότι υπάρχει μία ποιοτική μεταβλητή με k επίπεδα τιμών και έστω, p_i το ποσοστό εμφάνισης στον πληθυσμό της τιμής i, i = 1, 2, …., k. Λαμβάνουμε ένα δείγμα μεγέθους N και βρίσκουμε τις συχνότητες $n_1, n_2, \ldots, n_k$, $(n_1 + n_2 + \ldots + n_k = N)$.

Για να ελέγξουμε τη στατιστική υπόθεση **H_0: $p_1 = p_{1,0}$, $p_2 = p_{2,0}$, …, $p_k = p_{k,0}$,** έναντι της **H_1: όχι η H_0.**

(α) Συμπληρώνουμε τον πίνακα

Συχνότητες \ Τιμή	1	2	...	k
Αναμενόμενες σύμφωνα με την H_0	$p_{1,0}\cdot N$	$p_{2,0}\cdot N$	...	$p_{k,0}\cdot N$
Παρατηρούμενες στο Δείγμα	n_1	n_2	...	n_k

(β) Υπολογίζουμε το στατιστικό $$\chi_0^2 = \sum_{i=1}^{k} \frac{(n_i - p_{i,0}\cdot N)^2}{p_{i,0}\cdot N} = \sum_{i=1}^{k} \frac{\delta_i^2}{p_{i,0}\cdot N}.$$

(γ) Βρίσκουμε την πιθανότητα $p = P(\chi^2 > \chi_0^2)$.

- Αν $p < 0{,}05$ λέμε (και γράφουμε) ότι η H_0 απορρίπτεται ($\chi^2(k-1) = \chi_0^2$, p = …)
- Αν $p \geq 0{,}05$ λέμε (και γράφουμε) ότι η H_0 δεν απορρίπτεται ($\chi^2(k-1) = \chi_0^2$, p = …)

(δ) Το όριο απόρριψης α = 0,05 ορίζεται στην αρχή της διαδικασίας.

Προϋποθέσεις εφαρμογής δοκιμασίας χ^2

Προϋποθέσεις εφαρμογής δοκιμασίας χ^2

Οι προϋποθέσεις της εφαρμογής της δοκιμασίας χ^2, αναδεικνύονται μέσα από την απόδειξη του Fisher. Αν ελέγχεται η H_0: $p_i = p_{i,0}$, $i = 1, 2, \dots, k$, έναντι της H_1: όχι η H_0, τότε

Απόδειξη (Fisher, 1922)
"Αν N_i, το πλήθος εμφανίσεων της κατηγορίας i, τότε η H_0 δηλώνει πως $\mathbf{N_i \sim Poisson(\lambda_i) \approx N(\lambda_i, \lambda_i)}$, με $\lambda_i = p_{i,0}\cdot N$ και $\lambda_1 + \lambda_2 + \dots + \lambda_k = N$. Άρα,

$$(N_i - p_{i,0}\cdot N)/(p_{i,0}\cdot N)^{0,5} \sim N(0, 1) \rightarrow (N_i - p_{i,0}\cdot N)^2/(p_{i,0}\cdot N) \sim \chi^2(1) \rightarrow \Sigma_{i = 1, \dots, k} (N_i - p_{i,0}\cdot N)^2/(p_{i,0}\cdot N) \sim \chi^2(k-1)$$

Οι βαθμοί ελευθερίας είναι k – 1 γιατί τα N_i περιορίζονται από τη σχέση $N_1 + N_2 + \dots + N_k = N$."

Η απόδειξη βασίζεται στην υπόθεση $\mathbf{N_i \sim Poisson(\lambda_i) \approx N(\lambda_i, \lambda_i)}$, με $\mathbf{\lambda_i = p_{i,0}\cdot N = E_i}$. ($E_i$: Expected frequency – αναμενόμενη συχνότητα). Συνεπώς, **πρέπει:**

1. **Οι παρατηρήσεις να είναι ανεξάρτητες μεταξύ τους** (ώστε να είναι έγκυρη η $N_i \sim Poisson(\lambda_i)$)
2. **Τα E_i να είναι αρκούντως μεγάλα** (ώστε να είναι έγκυρη η $Poisson(\lambda_i) \approx N(\lambda_i, \lambda_i)$).

Fisher, R. A., 1922. On the interpretation of χ2 from contingency tables, and the calculation of P. Journal of the Royal Statistical Society 1, 87–94, https://www.jstor.org/stable/2340521

Προϋποθέσεις εφαρμογής δοκιμασίας χ^2

Βρήκαμε ότι πολύ μικρές αναμενόμενες συχνότητες ή μεγάλο πλήθος από μικρές αναμενόμενες συχνότητες αντιστοιχεί σε μειωμένη αξιοπιστία της μεθόδου. Πιο συγκεκριμένα, σύμφωνα με τον Cochran (1952) η δοκιμασία χ^2 έχει μειωμένη ισχύ αν οι αναμενόμενες τιμές είναι μικρότερες από 1 ή αν το 20% των αναμενόμενων κελιών είναι μικρότερες από το 5.

Άρα, πρέπει να ισχύουν οι εξής προϋποθέσεις:
Α. Οι παρατηρήσεις να είναι ανεξάρτητες μεταξύ τους.
Β. Οι E_i να είναι μεγαλύτερες του 5.
Γ. Αν υπάρχουν E_i μικρότερες του 5, αυτές να είναι λιγότερες από το 20% του συνόλου.
Δ. Καμία E_i να μην είναι μικρότερη του 1.

Είναι αξιοσημείωτο ωστόσο πως νεότερες αναφορές καταδεικνύουν πως τα όρια σχετικά με το μέγεθος των E_i, είναι συντηρητικά και πως ο ερευνητής μπορεί να θεωρεί αξιόπιστο το αποτέλεσμα της μεθόδου όταν οι αναμενόμενες συχνότητες είναι μεγαλύτερες από 0,5 και οι περισσότερες να είναι μεγαλύτερες από τη μονάδα (Conover (1999, σ. 202)).

Cochran, W. G. (1952) The χ^2 test of goodness of fit, *Annals of Mathematical Statistics*, 23:315-345.

Conover, W. J. (1999) *Practical nonparametric statistics*, Third Edition, John Wiley & Sons, Inc., New York, New York, USA.

$H_0: p_1 = p_2 = \dots = p_{12} = \frac{1}{12}$

Δοκιμασία χι-τετράγωνο ως έλεγχος προσαρμογής

Άσκηση 1

Αξιοποιώντας τα δεδομένα του πίνακα σχετικά με τις γεννήσεις ανά μήνα, βρείτε αν οι 600 γεννήσεις που συνέβησαν σε ένα έτος, κατανέμονται ομοιόμορφα στους 12 μήνες του χρόνου

Μήνας	1	2	3	4	5	6	7	8	9	10	11	12	Σύνολο
Πλήθος γεννήσεων	60	44	45	50	49	56	46	41	69	49	44	47	600
Expected	50	50	50	50	50	50	50	50	50	50	50	50	600
Δ_k	-10	6	5	0	1	-6	4	9	-19	1	6	3	-

$$\chi_0^2 = \sum_{k=1}^{12} \frac{\Delta_k^2}{E_k} = \frac{(-10)^2}{50} + \frac{6^2}{50} + \frac{5^2}{50} + \frac{0^2}{50} + \frac{1^2}{50} + \frac{(-6)^2}{50} + \dots + \frac{3^2}{50} = \frac{702}{50} = 14,04$$

$\chi_0^2 \sim \chi^2(12-1) = \chi^2(11)$, $p = P(\chi^2 > \chi_0^2) = P(\chi^2 > 14,04) = 0,231 > 0,05$: Η H_0 Δεν απορρίπτεται!

Ασκήσεις

Άσκηση 2

Ένα νέο φάρμακο για την αντιμετώπιση της υπέρτασης δόθηκε πειραματικά σε 200 άτομα που πάσχουν από υπέρταση. Το αποτέλεσμα της φαρμακευτικής αγωγής για κάθε ασθενή ταξινομήθηκε σε μια από τέσσερις κατηγορίες:

Α: Βαθμιαία μείωση Β: Μέτρια μείωση Γ: Μικρή μείωση Δ: Μικρή αύξηση.

Οι συχνότητες των τεσσάρων κατηγοριών στα 200 άτομα φαίνονται στον πίνακα που ακολουθεί.

	Α	Β	Γ	Δ
Συχνότητα	120	50	20	10

Από σχετικές μελέτες είναι γνωστό ότι ένα αντίστοιχο φάρμακο που ήδη κυκλοφορεί και χρησιμοποιείται, έχει την εξής αποτελεσματικότητα: Α:50%, Β:30%, Γ:19% και Δ:1%.

Να βρεθεί αν το νέο φάρμακο διαφέρει ως προς την αποτελεσματικότητά του από το φάρμακο που ήδη κυκλοφορεί σε επίπεδο σημαντικότητας 0,05.

Ασκήσεις

Λύση 2

	Α	Β	Γ	Δ
Συχνότητα	120	50	20	10
Αναμενόμενη σύμφωνα με την H_0				

Ασκήσεις

Άσκηση 3

Σε ένα σακί υπάρχουν βόλοι τριών χρωμάτων, Κόκκινο, Πράσινο και Κίτρινο. Κάποιος ισχυρίζεται ότι το ποσοστό κάθε χρώματος στο σάκο είναι 40% (Κόκκινο), 35% (Πράσινο), 25% (Κίτρινο). Εμείς επιλέγουμε με επανάθεση 100 βόλους και βρίσκουμε 32 Κόκκινους, 45 Πράσινους και 23 Κίτρινους. Ελέγξτε αν η κατανομή του δείγματος είναι σημαντικά διαφορετική από τη θεωρητικώς αναμενόμενη, αν ο ισχυρισμός είναι αληθής ($\alpha = 0,05$).

Ασκήσεις

Άσκηση 4

Υπάρχει η πεποίθηση ότι είναι πιο πιθανό να γίνει ατύχημα με Ι.Χ. το Σαββατοκύριακο παρά τις Καθημερινές. Στον πίνακα παρουσιάζονται οι συχνότητες των ατυχημάτων σε μία ελληνική πόλη κατά τη διάρκεια ενός έτους. Ελέγξτε αν οι συχνότητες κατανέμονται ομοιόμορφα στις ημέρες της εβδομάδας (α = 0,05).

Ημέρα	Δευτέρα	Τρίτη	Τετάρτη	Πέμπτη	Παρασκευή	Σάββατο	Κυριακή	Σύνολο
Αριθμός ατυχημάτων	20	20	22	22	29	26	31	170

Ασκήσεις

Άσκηση 5

Στον πίνακα δίνεται το πλήθος εργατικών ατυχημάτων που συνέβησαν ανά ημέρα σε μια βιομηχανική ζώνη τα έξι περίπου τελευταία χρόνια (1500 εργάσιμες ημέρες). Να βρείτε αν το πλήθος των ατυχημάτων περιγράφεται ικανοποιητικά από την κατανομή Poisson με λ = 1 (α = 0,05).

Αριθμός ατυχημάτων ανά ημέρα	0	1	2	3	4	5	Σύνολο
Συχνότητα	549	555	273	93	24	6	1500

X: πλήθος ατυχημάτων/ημέρα $\quad H_0: X \sim Poisson(1): P(X=k) = e^{-1}\cdot\frac{1^k}{k!}$

H_0: $P(X=0) = e^{-1} = 0{,}368$, $P(X=1) = e^{-1} = 0{,}368$, $P(X=2) = e^{-1}\cdot\frac{1^2}{2!} = 0{,}184$

$P(X=3) = e^{-1}\cdot\frac{1^3}{3!} = 0{,}061$, $P(X=4) = e^{-1}\cdot\frac{1^4}{4!} = 0{,}015$, $P(X=5) = e^{-1}\frac{1^5}{5!} = 0{,}003$.

Σχόλιο: Η άσκηση αυτή δείχνει τη δυνατότητα χρήσης της δοκιμασία χι-τετράγωνο για τον έλεγχο προσαρμογής μίας εμπειρικής κατανομής σε μία θεωρητικώς αναμενόμενη.

Αρ. Ατ.	0	1	2	3	4	5	Σύνολο
Observed	549	555	273	93	24	6	1.500
Expected	552	552	276	92	23	5	1.500
Δ_k	3	-3	3	-1	-1	-1	—

$$X_0^2 = \frac{3^2}{552} + \frac{(-3)^2}{552} + \frac{3^2}{276} + \frac{(-1)^2}{92} + \frac{(-1)^2}{23} + \frac{(-1)^2}{5} = 0,321.$$

$$X_0^2 \sim X^2(6-1) = X^2(5) \quad , \quad p = P(X^2 > X_0^2) = P(X^2 > 0,321) = 0,997.$$

Η H_0 δεν απορρίπτεται γιατί $p = 0,997 > 0,05$.

$$X_0^2 = \sum_{j=1}^{3}\sum_{i=2}^{2} \frac{(E_{ij}-O_{ij})^2}{E_{ij}} = \frac{(7,8-10)^2}{7,8} + \frac{(11,2-8)^2}{11,2} + \dots + \frac{(23-24)^2}{24} = 3,078.$$

Ασκήσεις

Άσκηση 6

Ένα μαθηματικό θεώρημα διδάχθηκε σε δύο τμήματα, στο ένα (42 μαθητές) με την παραδοσιακή μέθοδο και στο άλλο (44 μαθητές) με μία νέα εναλλακτική μέθοδο. Μετά καταγράφηκε η κατανόηση των μαθητών σε τρεις κατηγορίες (Χαμηλή, Μέτρια και Υψηλή). Τα αποτελέσματα παρουσιάζονται στον παρακάτω πίνακα. Να βρείτε αν η μέθοδος διδασκαλίας διαφοροποίησε την κατανόηση των μαθητών (α = 0,05).

	Κατανόηση			
Μέθοδος	Χαμηλή	Μέτρια	Υψηλή	Σύνολο
Παραδοσιακή	10 /7,8	8 /11,2	24 /23	42
Νέα	6 /8,2	15 /11,8	23 /24	44
Σύνολο	16	23	47	86

H_0: Κατανόηση ανεξάρτητη με τη Μέθοδο

H_1: όχι η H_0.

$p = P(X^2 > 3,078) = 0,215 > 0,05$

$X^2 \sim X^2((3-1)(2-1)) = X^2(2)$

η H_0 ΔΕΝ απορρίπτεται!

Ασκήσεις

Άσκηση 7

Ένα νέο αντιγριπικό εμβόλιο χορηγείται σε δύο δόσεις. Για τον έλεγχο της αποτελεσματικότητάς του και τον υπολογισμό της βελτίωσης από κάθε δόση επιλέχθηκαν 1000 εθελοντές στους οποίους έγιναν καμία, μία ή δύο δόσεις του εμβολίου και καταγράφηκε μετά από 6 μήνες το γεγονός της ασθένειας από γρίπη. Τα αποτελέσματα παρουσιάζονται στον παρακάτω πίνακα. Να βρεθεί αν το πλήθος των δόσεων συσχετίζεται με την ανθεκτικότητα στη γρίπη (α = 0,05).

	Δόσεις		
Ανθεκτικότητα	0	1	2
Ασθένησε	24	9	13
Δεν ασθένησε	289	100	565

Ασκήσεις

Λύση 7

	Δόσεις			
Ανθεκτικότητα	0	1	2	Σύνολο
Ασθένησε	24 ( )	9 ( )	13 ( )	
Δεν ασθένησε	289 ( )	100 ( )	565 ( )	
Σύνολο				

Δοκιμασία Student (t - test) για ένα δείγμα

Η έννοια της υπόθεσης στο t – test

Μία υπόθεση είναι μία δήλωση για έναν πληθυσμό, την οποία επιθυμούμε να ελέγξουμε, χρησιμοποιώντας δεδομένα από ένα δείγμα. Στα πλαίσια του ελέγχου t – test, η υπόθεση αποκτά δύο εκδοχές:

- Την **μηδενική** ή **στατιστική H_0** με την οποία θα “συγκριθεί” το δείγμα και θα υπολογιστεί η “απόκλιση” του από αυτή.

  Παραδείγματα: (α) H_0: $\mu = \mu_0$, (β) H_0: $\mu \leq \mu_0$, (γ) H_0: $\mu \geq \mu_0$

- Την **εναλλακτική** ή **ερευνητική H_1**, η οποία είναι η συμπληρωματική δήλωση της H_0.

  Παραδείγματα: (α) H_1: $\mu \neq \mu_0$, (β) H_1: $\mu > \mu_0$, (γ) H_1: $\mu < \mu_0$

- Η H_1 αναδεικνύεται ως ερευνητικό αποτέλεσμα αν “απορριφθεί” η H_0 εξαιτίας της μεγάλης “απόκλισης” του δείγματος από αυτήν.

Εκδοχή 1η: “Δίπλευρος” έλεγχος

Παράδειγμα υπόθεσης: Το μέσο βάρος όλων των συσκευασιών γάλατος από μία αλυσίδα παραγωγής είναι 500 γραμμάρια.

- Στατιστική υπόθεση: H_0: $\mu = 500$.
- Ερευνητική υπόθεση: H_1: $\mu \neq 500$.

Η H_0 θα απορριφθεί έναντι της H_1, αν η μέση τιμή του δείγματος είναι “σημαντικά” μεγαλύτερη ή “σημαντικά” μικρότερη από τα 500 γραμμάρια.

Εκδοχή 2η: “Μονόπλευρος” έλεγχος

Παράδειγμα υπόθεσης: Το μέσο βάρος όλων των συσκευασιών γάλατος από μία αλυσίδα παραγωγής δεν είναι μικρότερο από τα 500 γραμμάρια.

- Στατιστική υπόθεση: H_0: $\mu \geq 500$.
- Ερευνητική υπόθεση: H_1: $\mu < 500$.

Η H_0 θα απορριφθεί έναντι της H_1, αν η μέση τιμή του δείγματος είναι “σημαντικά” μικρότερη από τα 500 γραμμάρια.

Εκδοχή 2η: “Μονόπλευρος” έλεγχος

Παράδειγμα υπόθεσης: Το μέσο βάρος όλων των συσκευασιών γάλατος από μία αλυσίδα παραγωγής δεν είναι μεγαλύτερο από τα 500 γραμμάρια.

- Στατιστική υπόθεση: H_0: $\mu \leq 500$.
- Ερευνητική υπόθεση: H_1: $\mu > 500$.

Η H_0 θα απορριφθεί έναντι της H_1, αν η μέση τιμή του δείγματος είναι “σημαντικά” μεγαλύτερη από τα 500 γραμμάρια.

Το στατιστικό t

$\sigma^2 = \frac{1}{n}\sum_{k=1}^{n}(x_i - \bar{x})^2$ $\quad x_1, x_2, \ldots, x_n \quad$ $\sum_{k=1}^{n}(x_i - \mu)^2$

Το στατιστικό t

Έστω $s^2 = \frac{1}{n-1}\sum_{k=1}^{n}(x_i - \bar{x})^2$, ο αμερόληπτος εκτιμητής για την διακύμανση.

Γνωρίζουμε ότι $Z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$ K.O.Θ.

Αποδεικνύεται ότι

$$(\alpha)\quad V = (n-1)\frac{s^2}{\sigma^2} \sim \chi^2(n-1)$$

$$(\beta)\quad t = \frac{Z}{\sqrt{V/(n-1)}} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \sim t(n-1)\ \left(\text{διότι } N(0,1)/\chi^2(n) = t(n)\right)$$

(α) https://math.stackexchange.com/questions/47009/proof-of-fracn-1s2-sigma2-sim-chi2-n-1

(β) Η απόδειξη και διάφορα χρήσιμα σχόλια είναι διαθέσιμα εδώ:
https://stats.stackexchange.com/questions/151854/a-normal-divided-by-the-sqrt-chi2s-s-gives-you-a-t-distribution-proof

Το στατιστικό t

$H_0 : \mu = \mu_0$

Αν $\bar{x}$ και s είναι η μέση τιμή και η δειγματική τυπική απόκλιση[1] ενός δείγματος μεγέθους n, τότε το στατιστικό

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \sim t(n-1),$$

μετρά τη διαφορά μεταξύ του μέσου όρου του δείγματος και του υποτιθέμενου μέσου όρου μ_0, σε μονάδες $s/n^{0,5}$.

Όσο μακρύτερα από το 0 είναι η τιμή του t_0 τόσο μεγαλύτερη είναι η απόκλιση του δείγματος από την υπόθεση H_0.

Η ποσότητα $s/n^{0,5}$, ονομάζεται **τυπικό σφάλμα του δείγματος** και είναι ένα μέτρο της μεταβλητότητας του μέσου όρου του δείγματος (δηλαδή δείχνει το πόσο πολύ αναμένεται να μεταβάλλεται η μέση τιμή ενός δείγματος από n στοιχεία).

$$(1)\quad s^2 = \frac{1}{n-1}\sum_{k=0}^{n}\left(x_i - \bar{x}\right)^2$$

Το στατιστικό t

Το στατιστικό

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

αντανακλά την “απόσταση” του δείγματος από την στατιστική υπόθεση H_0 σε όλες τις εκδοχές της

- H_0: $\mu = \mu_0$ (δίπλευρος έλεγχος)
- H_0: $\mu \leq \mu_0$ (μονόπλευρος έλεγχος)
- H_0: $\mu \geq \mu_0$ (μονόπλευρος έλεγχος).

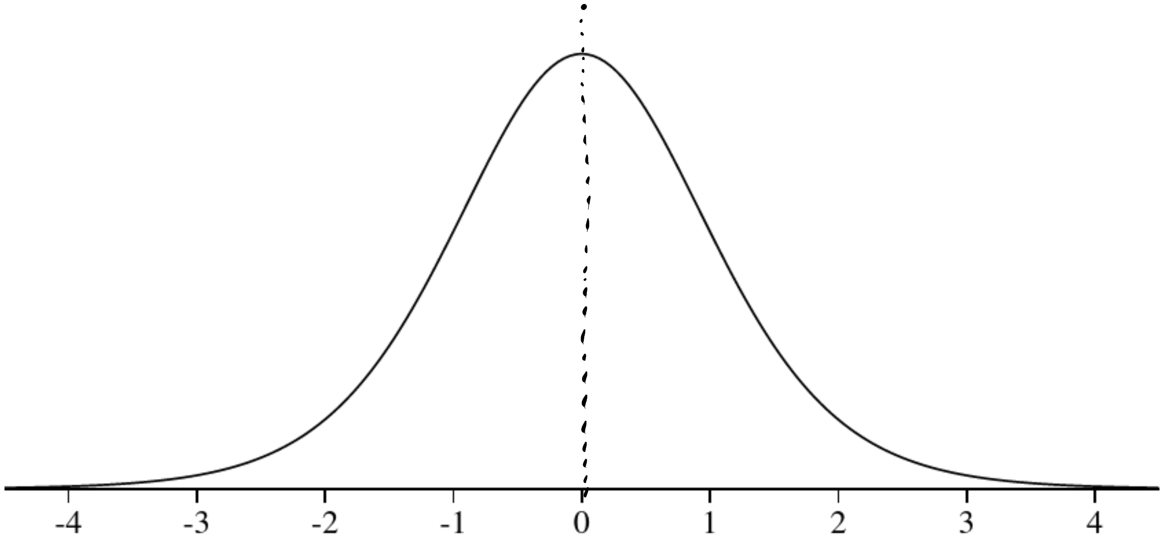
Η κατανομή του στατιστικού t

Η Θεωρία Πιθανοτήτων διαβεβαιώνει ότι, για μέγεθος δείγματος n, το στατιστικό

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \sim t(n-1)$$

π.χ. για n = 10, είναι t ~ Student t(9), η οποία περιγράφεται από την καμπύλη του σχήματος.

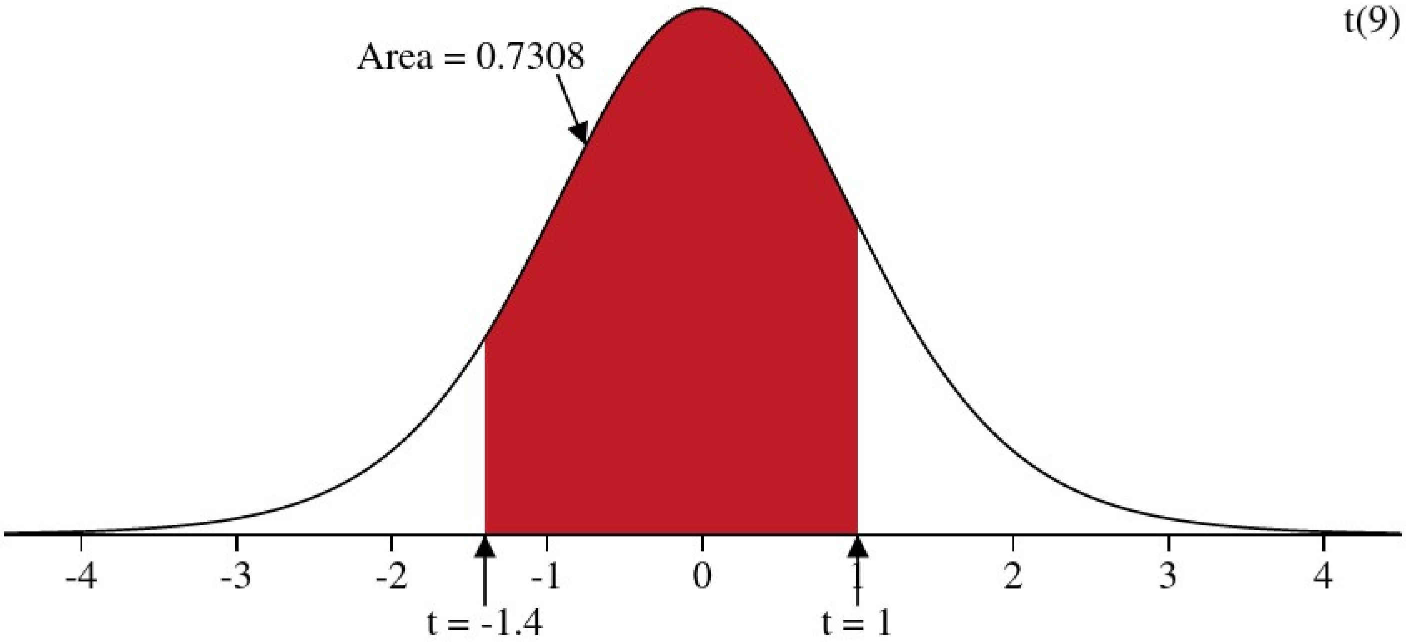
$H_0: \mu = \mu_0$



Η κατανομή του στατιστικού t

Το εμβαδόν της καμπύλης κατανομής αντιστοιχεί σε πιθανότητα.

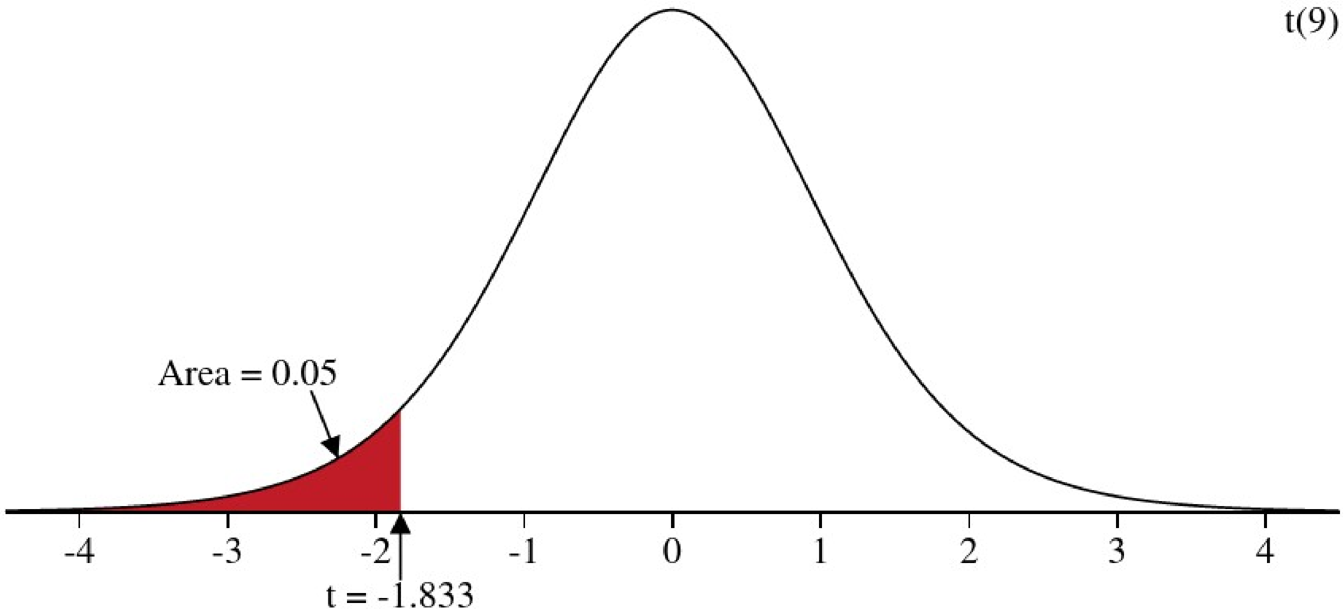
π.χ. αν t ~ Student t(9), τότε P(-1,4 < t < 1) = 0,7308 ≈ 73,1%.



Η κατανομή του στατιστικού t

Γνωρίζοντας ότι t ~ t(9) μπορούμε να υπολογίσουμε τις εξής (κρίσιμες) τιμές:

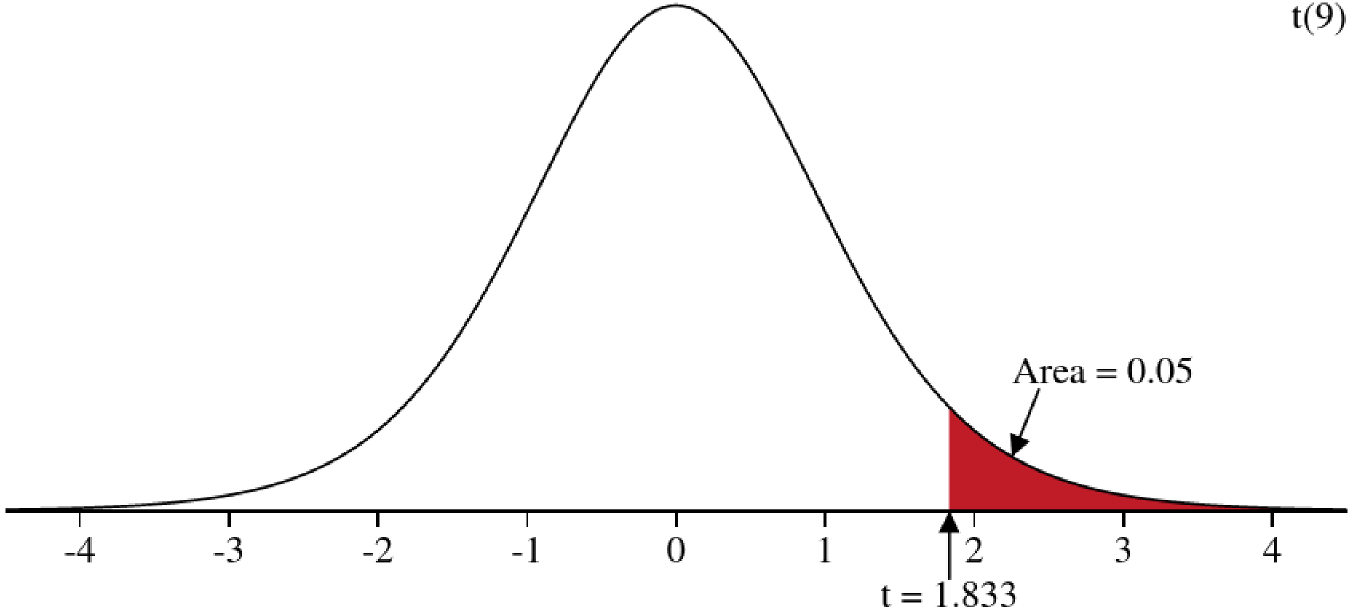
- $t_{critical1}$ = -1,833, κάτω από την οποία, η πιθανότητα εμφάνισης ενός δείγματος με τέτοια τιμή είναι μικρότερη του 5%. Ισοδύναμα, στο 95% των δειγμάτων με n = 10, αντιστοιχεί στατιστικό t > $t_{critical1}$ = -1,833.



Η κατανομή του στατιστικού t

Γνωρίζοντας ότι t ~ t(9) μπορούμε να υπολογίσουμε τις εξής (κρίσιμες) τιμές:

- $t_{critical2}$ = 1,833, πάνω από την οποία, η πιθανότητα εμφάνισης ενός δείγματος με τέτοια τιμή είναι μικρότερη του 5%. Ισοδύναμα, στο 95% των δειγμάτων με n = 10, αντιστοιχεί στατιστικό t < $t_{critical2}$ = 1,833.

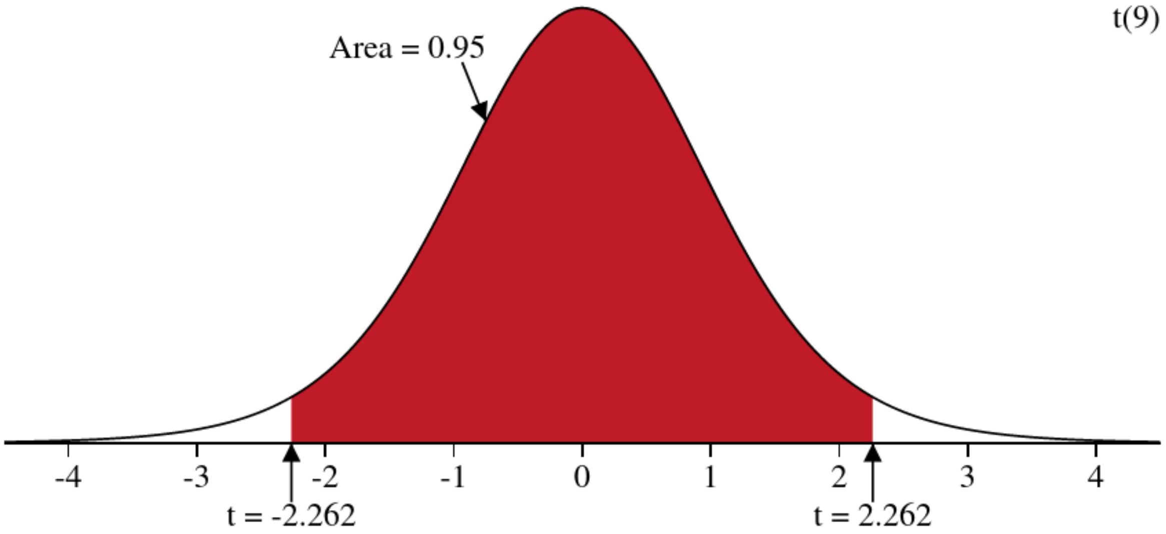


Η κατανομή του στατιστικού t

Γνωρίζοντας ότι t ~ t(9) μπορούμε να υπολογίσουμε τις εξής (κρίσιμες) τιμές:

- $t_{critical3}$ = 2,262, πέρα από την οποία (είτε προς τα πάνω είτε προς τα κάτω), η πιθανότητα εμφάνισης ενός δείγματος με τέτοια τιμή είναι μικρότερη του 5%. Ισοδύναμα, στο 95% των δειγμάτων με n = 10, αντιστοιχεί στατιστικό

$$-2{,}262 = t_{critical3} < t < t_{critical3} = 2{,}262.$$



Κρίση της H_0: Δίπλευρος έλεγχος

Κρίση της H_0: Δίπλευρος έλεγχος

Εκδοχή 1η: Δίπλευρος έλεγχος με δείγμα μεγέθους n = 10.

Στατιστική υπόθεση: H_0: $\mu = \mu_0$.

Ερευνητική υπόθεση: H_1: $\mu \neq \mu_0$.

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

Είναι t ~ Student t(9).

Αν το t_0 βρίσκεται έξω από το διάστημα $(-t_{critical3}, t_{critical3}) = (-2.262, 2.262)$, τότε **απορρίπτουμε την H_0 έναντι της H_1**.

Αν το t_0 βρίσκεται μέσα στην περιοχή $(-t_{critical3}, t_{critical3}) = (-2.262, 2.262)$, τότε **δεν απορρίπτουμε την H_0 έναντι της H_1**.

Πιθανότητα p: Δίπλευρος έλεγχος

Τα στατιστικά λογισμικά (και η γλώσσα R), αντί της κρίσιμης τιμής αναφέρουν την πιθανότητα p να εμφανιστεί μία περισσότερο ακραία δειγματική μέση τιμή από αυτή που παρατηρήθηκε στο δείγμα:

$$p_{\text{δίπλευρο}} = P(t < -t_0 \text{ ή } t > t_0).$$

Η σύγκριση της p με το 0,05 οδηγεί στα δύο πιθανά αποτελέσματα:

Αν $p < 0{,}05$, τότε **απορρίπτουμε την H_0: $\mu = \mu_0$, έναντι της H_1: $\mu \neq \mu_0$.**

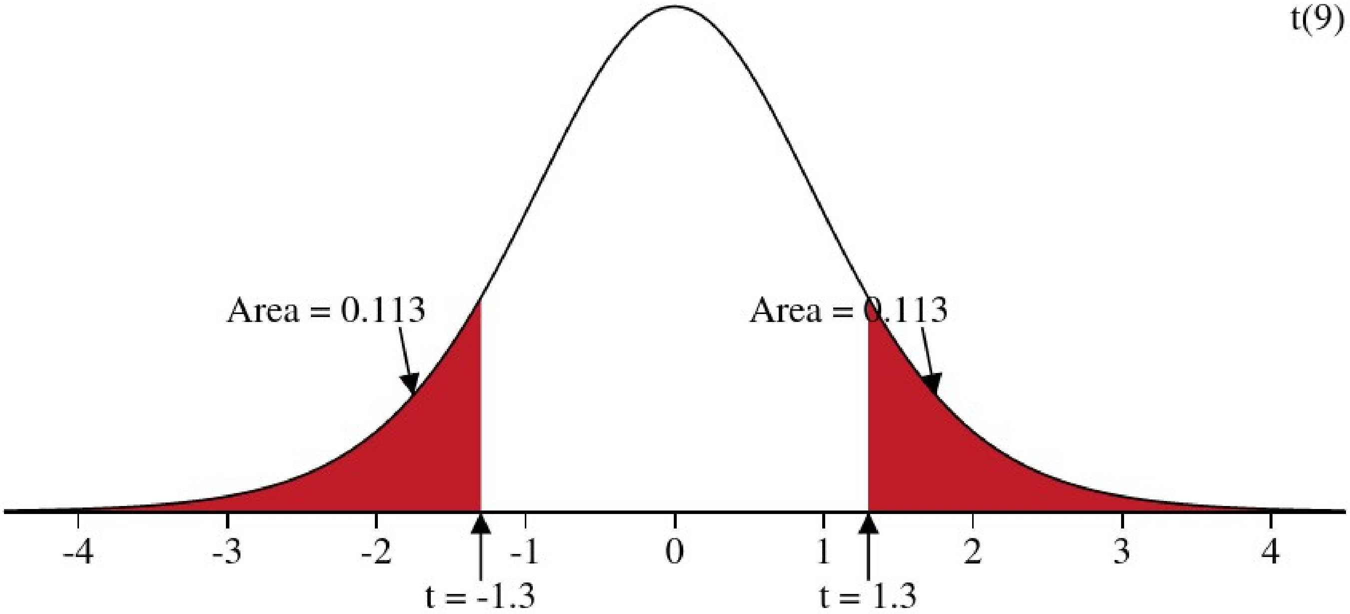
Αν $p \geq 0{,}05$, τότε **δεν απορρίπτουμε την H_0: $\mu = \mu_0$, έναντι της H_1: $\mu \neq \mu_0$.**

Κρίση της H_0: Δίπλευρος έλεγχος

Παράδειγμα (δίπλευρος έλεγχος με δείγμα μεγέθους n = 10).

Δίπλευρος έλεγχος H_0: $\mu = \mu_0$ έναντι H_1: $\mu \neq \mu_0$ και $t = 1,3$ ή $t = -1,3$.

Είναι $p_{δίπλευρο} = 0,226 > 0,05$, και **δεν απορρίπτουμε την H_0 έναντι της H_1**.

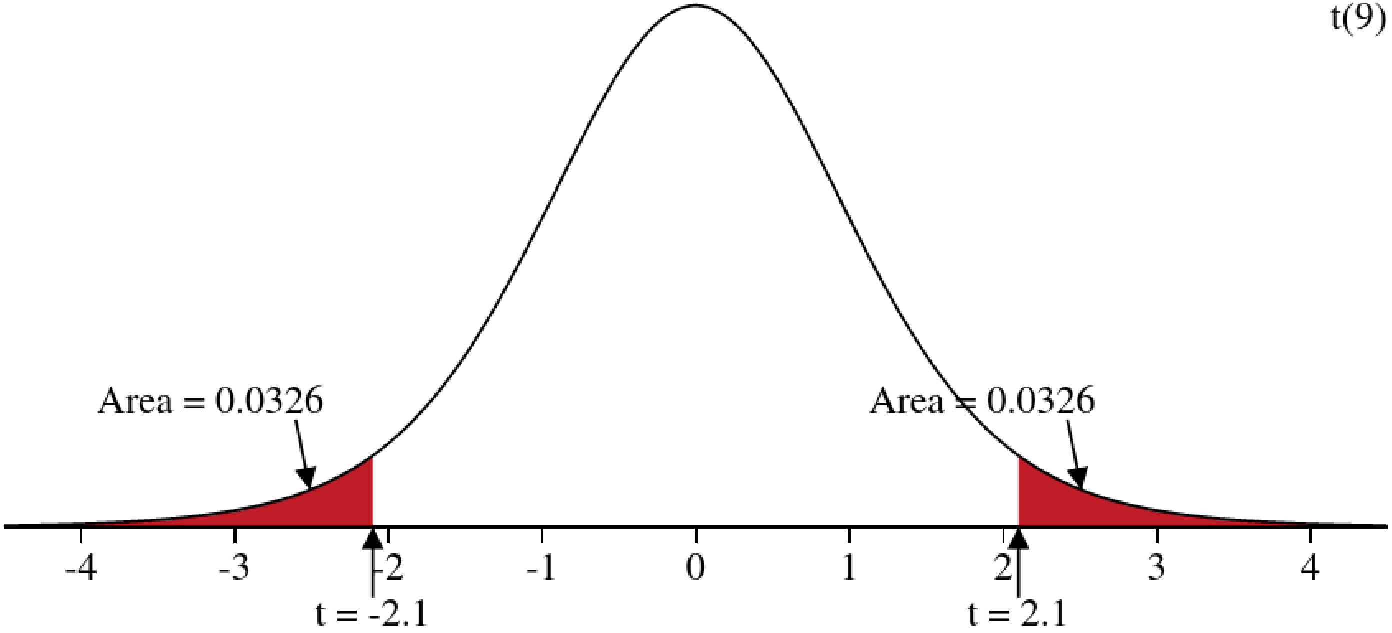


Κρίση της H_0: Δίπλευρος έλεγχος

Παράδειγμα (δίπλευρος έλεγχος με δείγμα μεγέθους n = 10).

Δίπλευρος έλεγχος H_0: $\mu = \mu_0$ έναντι H_1: $\mu \neq \mu_0$ και $t = 2{,}1$ ή $t = -2{,}1$.

Είναι $p_{\text{δίπλευρο}} = 0{,}0652 > 0{,}05$ και **δεν απορρίπτουμε την H_0 έναντι της H_1**.

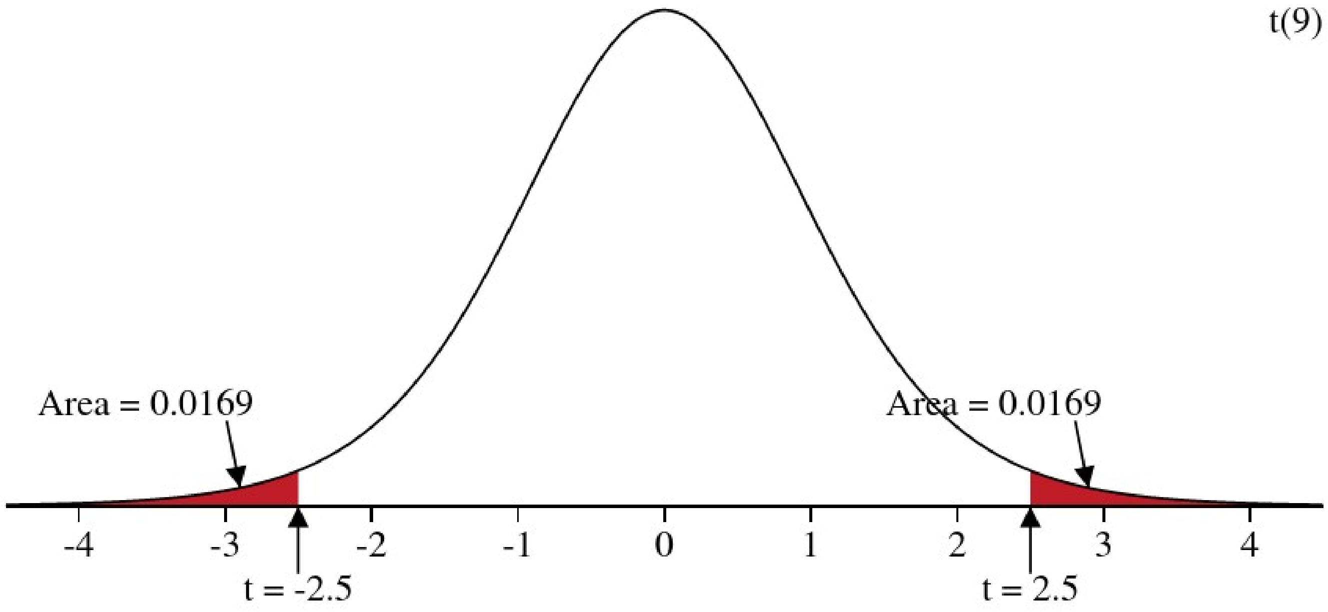


Κρίση της H_0: Δίπλευρος έλεγχος

Παράδειγμα (δίπλευρος έλεγχος με δείγμα μεγέθους n = 10).

Δίπλευρος έλεγχος H_0: $\mu = \mu_0$ έναντι H_1: $\mu \neq \mu_0$ και $t = 2{,}5$ ή $t = -2{,}5$.

Είναι $p_{\text{δίπλευρο}} = 0{,}0388 < 0{,}05$ και **απορρίπτουμε την H_0 έναντι της H_1**.



Κρίση της H_0: Μονόπλευρος έλεγχος

Κρίση της H_0: Μονόπλευρος έλεγχος

Εκδοχή 2η: Μονόπλευρος έλεγχος με δείγμα μεγέθους n = 10.

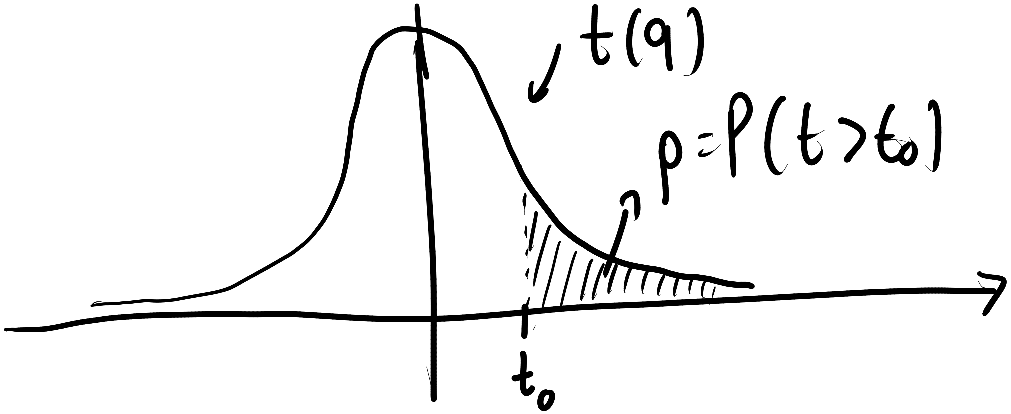
Στατιστική υπόθεση: H_0: $\mu \leq \mu_0$.

Ερευνητική υπόθεση: H_1: $\mu > \mu_0$.

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

Αν $t \geq t_{critical2} = 1{,}832$, τότε **απορρίπτουμε την H_0 έναντι της H_1**.

Αν $t < t_{critical2} = 1{,}832$, τότε **δεν απορρίπτουμε την H_0 έναντι της H_1**.



Κρίση της H_0: Μονόπλευρος έλεγχος

Εκδοχή 2η: Μονόπλευρος έλεγχος με δείγμα μεγέθους n = 10.

Στατιστική υπόθεση: H_0: $\mu \geq \mu_0$.

Ερευνητική υπόθεση: H_1: $\mu < \mu_0$.

Αν $t \leq t_{critical1} = -1{,}832$, τότε **απορρίπτουμε την H_0 έναντι της H_1**.

Αν $t > t_{critical1} = -1{,}832$, τότε **δεν απορρίπτουμε την H_0 έναντι της H_1**.

Πιθανότητα p: Μονόπλευρος έλεγχος

Στην περίπτωση όπου διεξάγεται μονόπλευρος έλεγχος τότε η πιθανότητα που μας ενδιαφέρει είναι η

$p_{\text{μονόπλευρο}} = P(t < -t_0)$ αν H_1: $\mu < \mu_0$ ή

$p_{\text{μονόπλευρο}} = P(t > t_0)$ αν H_1: $\mu > \mu_0$.

Από τη συμμετρία της κατανομής Student t(n) γίνεται αντιληπτό ότι:

$$p_{\text{μονόπλευρο}} = p_{\text{δίπλευρο}} / 2.$$

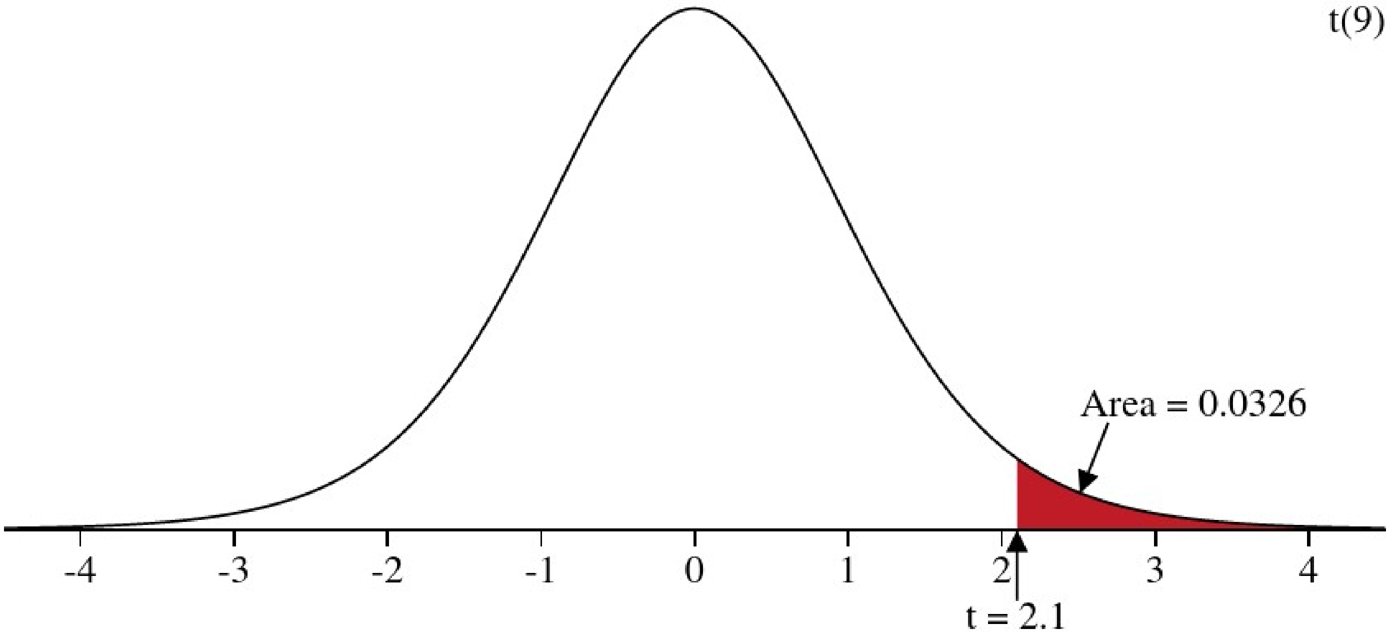
Κάποια στατιστικά προγράμματα αναφέρουν πάντα την πιθανότητα p για τον δίπλευρο έλεγχο. Συνεπώς, στην περίπτωση αυτή, **αν διεξάγουμε μονόπλευρο έλεγχο πρέπει να υπολογίζουμε το μισό της αναφερόμενης τιμής.**

Κρίση της H_0: Μονόπλευρος έλεγχος

Παράδειγμα (μονόπλευρος έλεγχος με δείγμα μεγέθους n = 10).

Μονόπλευρος έλεγχος H_0: $\mu \leq \mu_0$ έναντι H_1: $\mu > \mu_0$ και $t = 2,1$.

Είναι $p_{μονόπλευρο} = 0,0326 < 0,05$ και **απορρίπτουμε την H_0 έναντι της H_1**.

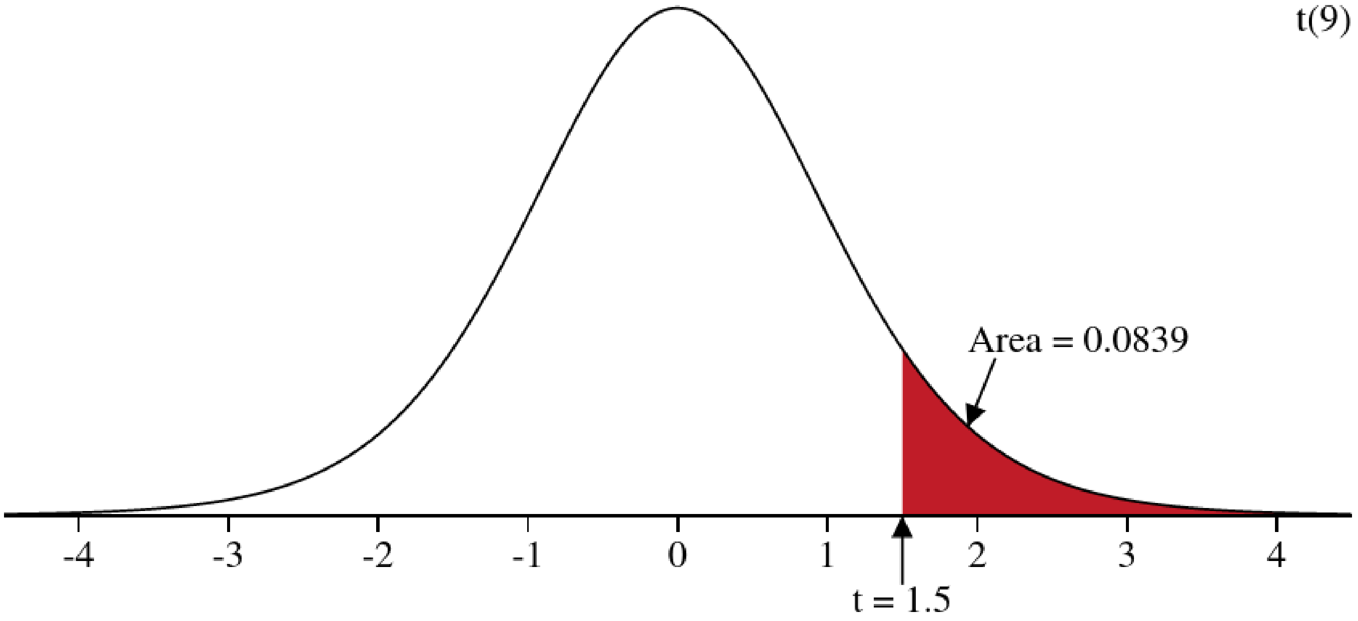


Κρίση της H_0: Μονόπλευρος έλεγχος

Παράδειγμα (μονόπλευρος έλεγχος με δείγμα μεγέθους n = 10).

Μονόπλευρος έλεγχος H_0: $\mu \leq \mu_0$ έναντι H_1: $\mu > \mu_0$ και t = 1,5.

Είναι $p_{μονόπλευρο}$ = 0,0839 > 0,05 και **δεν απορρίπτουμε την H_0 έναντι της H_1**.



Σύνοψη

Δείγμα μεγέθους n με μέση τιμή $\bar{x}$ και τυπική απόκλιση s.

1. Υπολογίζουμε το $t_0 = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$

2. Είναι t ~ t(n – 1).

3. Υπολογίζουμε το $p = P(|t| > t_0)$ ή $p = P(t > t_0)$ ή $p = P(t < -t_0)$ αντίστοιχα για δίπλευρο ή μονόπλευρο έλεγχο.

4. Αν p < 0,05 τότε απορρίπτουμε την μηδενική υπόθεση και δεχόμαστε αντίστοιχα $\mu \neq \mu_0$ ή $\mu > \mu_0$ ή $\mu < \mu_0$.

5. Όταν n > 30 τότε t(n) ~ N(0, 1).

Παραδείγματα

Παράδειγμα 1

Θέλουμε να ελέγξουμε αν μία γραμμή συσκευασίας γάλατος παράγει συσκευασίες με μέσο βάρος διαφορετικό από 500 γραμμάρια. Παίρνουμε δείγμα 10 συσκευασιών και βρίσκουμε βάρος:

490, 503, 499, 492, 500, 501, 489, 478, 498, 508,

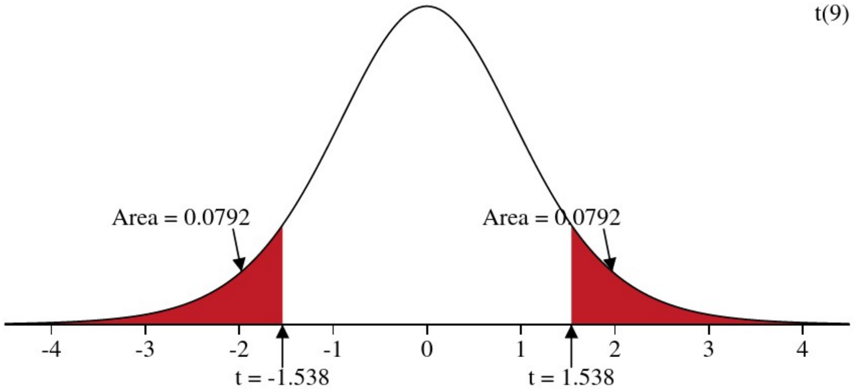
Στατιστική υπόθεση **H_0: μ = 500** έναντι της εναλλακτικής υπόθεσης **H_1: μ ≠ 500.**

Υπολογίζουμε: $\bar{x} = 495{,}8$, s = 8,64, και $t_0 = \frac{\bar{x} - 500}{s/\sqrt{n}} = \frac{495{,}8 - 500}{8{,}64/\sqrt{10}} = -1{,}538$

Είναι $p_{\text{δίπλευρο}}$ = 0,1584 > 0,05, άρα η υπόθεση H_0: μ = 500, **δεν απορρίπτεται** έναντι της H_1: μ ≠ 500.

Γράφουμε:

Το μέσο βάρος του δείγματος των 10 συσκευασιών δεν είναι σημαντικά διαφορετικό από τα 500 γραμμάρια (t(9) = 1,538, p = 0,158).



Παράδειγμα 1

Υλοποίηση στην R

```
>> one.sample.data = c(490, 503, 499, 492, 500, 501, 489, 478, 498, 508)
>> t.test(one.sample.data, mu = 500)
```

Output:

```
One Sample t-test

data:  one.sample.data
t = -1.5375, df = 9, p-value = 0.1585
alternative hypothesis: true mean is not equal to 500
95 percent confidence interval:
 489.6204 501.9796
sample estimates:
mean of x
    495.8
```

Παράδειγμα 2

Μία εταιρεία που παράγει μπάρες δημητριακών υποθέτει ότι η ποσότητα πρωτεΐνης σε κάθε μία μπάρα είναι κατά μέσο όρο μεγαλύτερη από τα 4 γραμμάρια. Σε δείγμα 20 προϊόντων βρέθηκαν οι εξής ποσότητες πρωτεΐνης:

4.1, 3.9, 4.2, 3.8, 4, 4, 4.1, 4.3, 4.4, 4.9, 3.9, 4, 4.6, 4.7, 4.7, 4.6, 4, 4, 3.8, 4.1

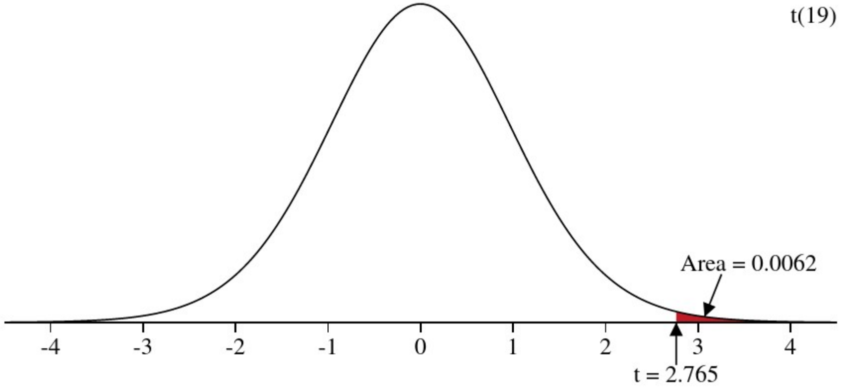
Στατιστική υπόθεση **H_0: μ ≤ 4** έναντι της εναλλακτικής υπόθεσης **H_1: μ > 4.**

Υπολογίζουμε: $\bar{x} = 4{,}205$, s = 0,3316, και $t_0 = \dfrac{\bar{x} - 4}{s/\sqrt{n}} = \dfrac{4{,}205 - 4}{0{,}3316/\sqrt{20}} = 2{,}765$

Είναι $p_{\text{μονόπλευρο}} = 0{,}0062 < 0{,}05$, άρα η **$H_0$: μ ≤ 4, απορρίπτεται έναντι της H_1: μ > 4**.

Γράφουμε:

Η μέση ποσότητα πρωτεΐνης του δείγματος των 20 προϊόντων είναι σημαντικά μεγαλύτερη από τα 4 γραμμάρια (t(19) = 2,765, p = 0,0062).



Παράδειγμα 2

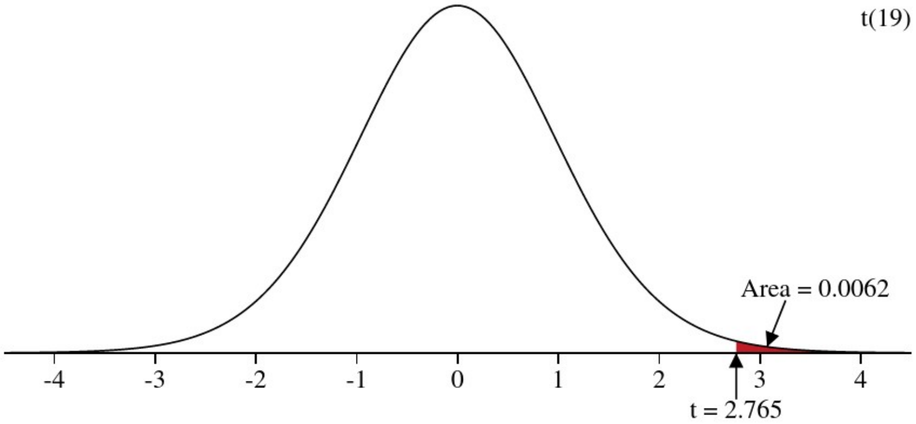
Υλοποίηση στην R

```
>> protein = c(4.1, 3.9, 4.2, 3.8, 4, 4, 4.1, 4.3, 4.4, 4.9, 3.9, 4, 4.6, 4.7, 4.7, 4.6, 4, 4, 3.8, 4.1)
>> t.test(protein, mu = 4, alternative = 'greater')
```

Output:

```
One Sample t-test

data:  protein
t = 2.7646, df = 19, p-value = 0.00617
alternative hypothesis: true mean is greater than 4
95 percent confidence interval:
 4.076779      Inf
sample estimates:
mean of x
   4.205
```



Παράδειγμα 2

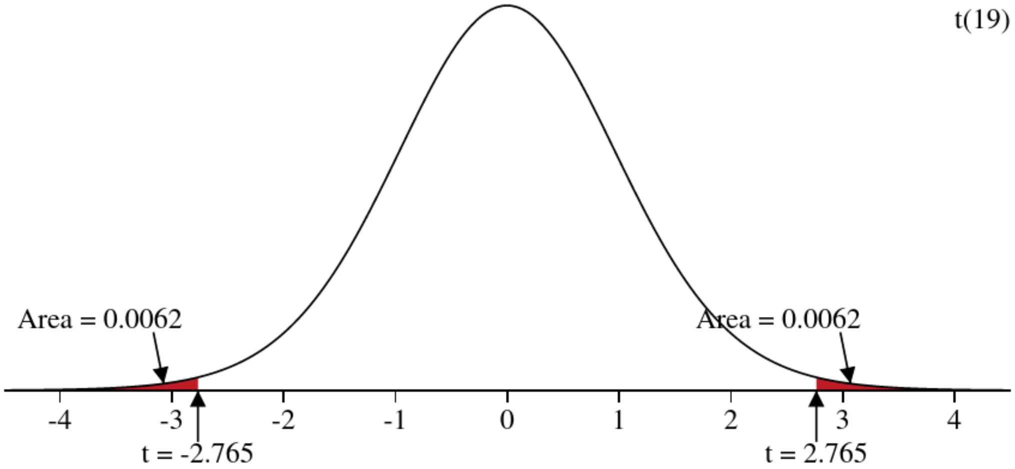
Υλοποίηση στην R

```
>> protein = c(4.1, 3.9, 4.2, 3.8, 4, 4, 4.1, 4.3, 4.4, 4.9, 3.9, 4, 4.6, 4.7, 4.7, 4.6, 4, 4, 3.8, 4.1)
>> t.test(protein, mu = 4)
```

Προσοχή!
Αν δεν οριστεί alternative = ‘greater’, τότε η R θα αναφέρει το p για το δίπλευρο έλεγχο. Στην περίπτωση που διεξάγουμε μονόπλευρο έλεγχο, θα πρέπει να υπολογίσουμε το μισό της αναφερόμενης τιμής!

```
Output:
One Sample t-test

data:  protein
t = 2.7646, df = 19, p-value = 0.01234
alternative hypothesis: true mean is not equal to 4
95 percent confidence interval:
 4.049796 4.360204
sample estimates:
mean of x
   4.205
```



Παράδειγμα 3

Ένας ερευνητής υποθέτει ότι το ChatGPT έχει επίδοση στο γνωστικό αντικείμενο της Φυσικής κατά μέσο όρο μεγαλύτερη από 18 (στα 20). Σε δείγμα 30 ασκήσεων καταγράφηκαν οι εξής επιδόσεις:

17, 16, 16, 18, 19, 19, 20, 20, 20, 20, 17, 18, 19, 19, 19, 19, 20, 19, 17, 18, 18, 17, 18, 17, 19, 20, 17, 17, 18, 19.

Στατιστική υπόθεση **H_0: $\mu \leq 18$** έναντι της εναλλακτικής υπόθεσης **H_1: $\mu > 18$.**

Υπολογίζουμε: $\bar{x} = 18{,}33$, s = 1,24, και

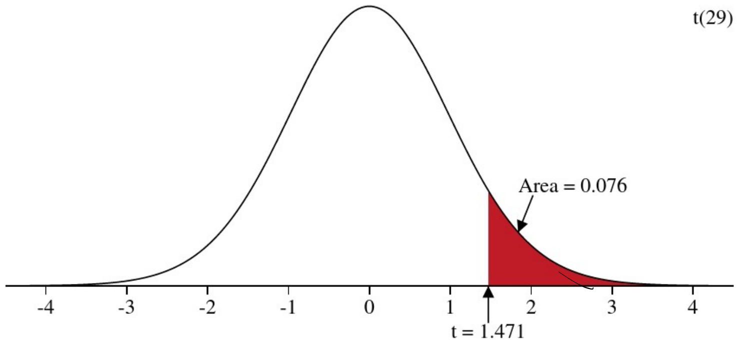
$$t_0 = \frac{\bar{x} - 18}{s/\sqrt{n}} = \frac{18{,}33 - 18}{1{,}24/\sqrt{30}} = 1{,}471$$

Είναι $p_{μονόπλευρο} = 0{,}076 > 0{,}05$, άρα η H_0: $\mu \leq 18$, **δεν απορρίπτεται** έναντι της H_1: $\mu > 18$.

Γράφουμε:

Η μέση επίδοση στο δείγμα των 30 ασκήσεων δεν είναι σημαντικά μεγαλύτερη από το 18

(t(29) = 1,471, p = 0,076).



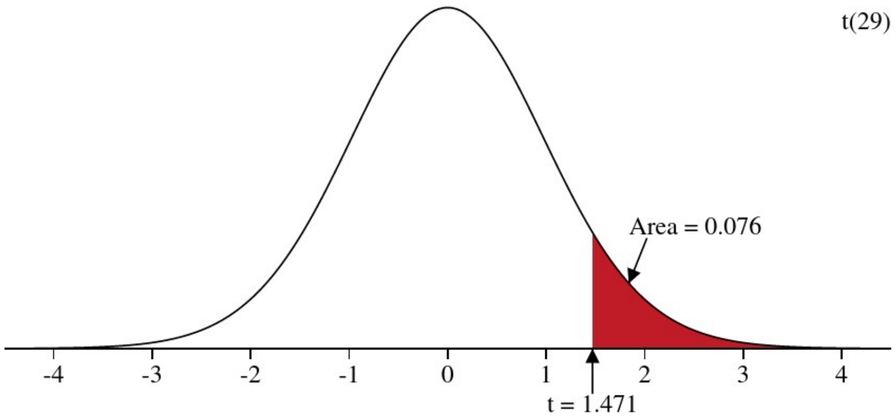
Παράδειγμα 3

Υλοποίηση στην R

```
>> physics = c(17, 16, 16, 18, 19, 19, 20, 20, 20, 20, 17, 18, 19, 19, 19, 19, 20, 19, 17, 18, 18, 17, 18,
   17, 19, 20, 17, 17, 18, 19)
>> t.test(physics, mu = 18, alternative = 'greater')
```

Output:

```
data:  physics
t = 1.4711, df = 29, p-value = 0.07601
alternative hypothesis: true mean is greater than 18
95 percent confidence interval:
 17.94834      Inf
sample estimates:
mean of x
 18.33333
```



Δραστηριότητες

Δραστηριότητα 1

Ένας μαθητής εκτιμά πως ένας καθηγητής λέει κατά μέσο όρο, περισσότερα από 3 ανέκδοτα σε κάθε μάθημα. Σε δείγμα 15 μαθημάτων καταγράφει το πλήθος των ανεκδότων:

4, 5, 3, 2, 1, 0, 5, 4, 2, 3, 3, 3, 5, 4, 2, 1

Ο μαθητής ελέγχει την υπόθεση **H_0: μ ≤ 3** έναντι της εναλλακτικής υπόθεσης **H_1: μ > 3.**

α) Εκτελέστε στην R τον παρακάτω κώδικα:

```
jokes = c(4, 5, 3, 2, 1, 0, 5, 4, 2, 3, 3, 3, 5, 4, 2, 1)
t.test(jokes, mu = 3, alternative = 'greater')
```

$$t_0 = \frac{\bar{x}-\mu_0}{s/\sqrt{n}} = \frac{2.94-3}{\sqrt{2.33}/\sqrt{15}} = -0.164$$

$$t \sim t(14). \quad p = P(t > t_0) = 0.564$$

β) Συμπληρώστε την παρακάτω πρόταση

Το μέσο πλήθος ανεκδότων στο δείγμα των 15 μαθημάτων, (είναι/δεν είναι)______ σημαντικά μεγαλύτερο από το 3 (t(14) = 0.099, p = 0.491). (μονόπλευρος έλεγχος). 2.33

$$\bar{x} = \frac{4+5+..+1}{15} = \frac{47}{15} = 2.94. \quad s^2 = \frac{1}{14}\sum_{i=1}^{15}(x_i-\bar{x})^2 = \frac{1}{14}\left[(4-3.13)^2+(5-3.13)^2+..+(1-3.13)^2\right] =$$

Δραστηριότητα 2

Ο ίδιος μαθητής εκτιμά πως ένας άλλος καθηγητής λέει κατά μέσο όρο, λιγότερα από 3 ανέκδοτα σε κάθε μάθημα. Σε δείγμα 15 μαθημάτων καταγράφει το πλήθος των ανεκδότων:

1, 2, 3, 2, 1, 0, 1, 2, 1, 3, 3, 3, 5, 2, 2, 1

Ο μαθητής ελέγχει την υπόθεση **H_0: μ ≥ 3** έναντι της εναλλακτικής υπόθεσης **H_1: μ < 3.**

α) Εκτελέστε στην R τον παρακάτω κώδικα:

```
jokes2 = c(1, 2, 3, 2, 1, 0, 1, 2, 1, 3, 3, 3, 5, 2, 2, 1)
t.test(jokes2, mu = 3, alternative = 'less')
```

β) Συμπληρώστε την παρακάτω πρόταση

Το μέσο πλήθος ανεκδότων στο δείγμα των 15 μαθημάτων, (είναι/δεν είναι)______ σημαντικά μικρότερο από το 3 (t(___) = _______, p = ______).

Σημείωση: Για πολύ μικρές τιμές του p γράφουμε p < 0,001.

Δραστηριότητα 3

Ένας οδηγός εκτιμά πως, σε κάθε φόρτιση, το νέο του Tesla καλύπτει κατά μέσο όρο περισσότερα από 300 χλμ. Καταγράφει το πλήθος χιλιομέτρων για 17 φορτίσεις και βρήκε τα εξής:

290, 310, 320, 280, 270, 306, 301, 298, 330, 320, 290, 270, 305, 315, 275, 280, 310.

Ο οδηγός ελέγχει την **H_0: μ ≤ 300** έναντι της **H_1: μ > 300**.

α) Εκτελέστε στην R τον παρακάτω κώδικα:

```
distance = c(290, 310, 305, 280, 270, 306, 301, 290, 315, 320, 290, 270, 305, 305, 270, 280, 295)
t.test(distance, mu = 300, alternative = 'greater')
```

β) Συμπληρώστε την παρακάτω πρόταση

Το μέσο πλήθος χιλιομέτρων στο δείγμα των 17 φορτίσεων, (είναι/δεν είναι)______ σημαντικά μεγαλύτερο από το 300 (t(___) = _______, p = ______).

Δραστηριότητα 4

Ένας μετεωρολόγος εκτιμά πως σε μία περιοχή πέφτουν περισσότερα από 200mm βροχής το χρόνο. Καταγράφει τη βροχόπτωση για 13 χρόνια και βρήκε τα εξής:

190, 170, 202, 210, 180, 170, 205, 185, 188, 195, 208, 190, 192.

Ο μετεωρολόγος ελέγχει την **H_0: μ ≤ 200** έναντι της **H_1: μ > 200**.

α) Εκτελέστε στην R τον παρακάτω κώδικα:

```
rain = c(190, 170, 202, 210, 180, 170, 205, 185, 188, 195, 208, 190, 192)
t.test(rain, mu = 200, alternative = 'greater')
```

β) Συμπληρώστε την παρακάτω πρόταση

Το μέσο ύψος βροχής στο δείγμα των 13 ετών, (είναι/δεν είναι)______ σημαντικά μεγαλύτερο από τα 200 mm (t(___) = _______, p = ______).

Δραστηριότητα 5

Ένας δήμαρχος εκτιμά πως το μέσο πλήθος παιδιών στην περιοχή του ανά οικογένεια είναι διαφορετικό από το μέσο πλήθος όλων των οικογενειών στην Ελλάδα που είναι 1,4. Καταγράφει τη πλήθος παιδιών για 17 οικογένειες και βρήκε τα εξής:

5, 0, 1, 2, 2, 3, 4, 2, 1, 1, 0, 0, 3, 3, 2, 2, 2

Ο δήμαρχος ελέγχει την **H_0: μ = 1,4** έναντι της **H_1: μ ≠ 1,4**.

α) Εκτελέστε στην R τον παρακάτω κώδικα:

```
children = c(5, 0, 1, 2, 2, 3, 4, 2, 1, 1, 0, 0, 3, 3, 2, 2, 2)
t.test(children, mu = 1.4)
```

β) Συμπληρώστε την παρακάτω πρόταση

Το μέσο πλήθος παιδιών στο δείγμα των 17 οικογενειών, (είναι/δεν είναι)______ σημαντικά διαφορετικό από το 1,4 (t(___) = _______, p = ______).

Δοκιμασία Student (t - test) για ένα δείγμα

Παρατηρήσεις

1. Αν το μέγεθος του δείγματος είναι μεγαλύτερο από 30 τότε t(n – 1) ~ N(0, 1) και στη θέση της κατανομής Student, μπορούμε χωρίς σημαντικό σφάλμα να χρησιμοποιούμε την κανονική κατανομή:

$$n > 30 \Rightarrow t = \frac{\left(\overline{X} - \mu_0\right)}{s/\sqrt{n}} \sim N(0\ ,1)$$

2. Έχουμε δει ότι το 95% διάστημα εμπιστοσύνης για την άγνωστη μέση τιμή του πληθυσμού είναι

$$\left(\overline{X} - 1{,}96\frac{\sigma}{\sqrt{n}}\ ,\ \overline{X} + 1{,}96\frac{\sigma}{\sqrt{n}}\right)$$

Παρατηρούμε ότι:

Η στατιστική υπόθεση H_0: μ = μ_0, απορρίπτεται έναντι της H_1: μ ≠ μ_0, σε επίπεδο σημαντικότητας 0,05 αν και μόνο αν το t_0 δεν ανήκει στο 95% διάστημα εμπιστοσύνης για την άγνωστη μέση τιμή μ.

Δοκιμασία Student (t - test) για ένα δείγμα

Παρατηρήσεις

3. Στην Στατιστική, η τυπική απόκλιση του δειγματικού μέσου, ονομάζεται **τυπικό σφάλμα (Standard Error).**

Το τυπικό σφάλμα δίνει τη δυνατότητα συντομότερης έκφρασης του 95% και του 99% διαστήματος εμπιστοσύνης για τη μέση τιμή ενός πληθυσμού.

Ενδεικτικά, αν n > 30, τότε $\overline{X} \sim N\left(\mu_0, \frac{\sigma^2}{n}\right), \; SE_N = \frac{\sigma}{\sqrt{n}}$

και το 95% διάστημα εμπιστοσύνης είναι το $\left(\overline{X} - 1{,}96 \cdot SE_N, \; \overline{X} + 1{,}96 \cdot SE_N\right)$

Αντίστοιχα αν $\overline{X} \sim N\left(\mu_0, \frac{s^2}{n}\right)$ τότε $SE_t = \frac{s}{\sqrt{n}}$ και $95\%\ \delta.\varepsilon. = \left(\overline{X} - t_{n;\,0.025} \cdot SE_t, \; \overline{X} + t_{n;\,0.025} \cdot SE_t\right)$

Δοκιμασία Student (t - test) για ένα δείγμα

4. Καταγραφή δοκιμασίας t - test για ένα δείγμα.

Στοιχεία που πρέπει να αναφερθούν

- τα στοιχεία του δείγματος (πλήθος παρατηρήσεων, μέση τιμή και τυπική απόκλιση),
- η τιμή του στατιστικού t ως απόλυτη τιμή,
- οι βαθμοί ελευθερίας,
- το είδος του στατιστικού ελέγχου (μονόπλευρος ή δίπλευρος),
- η στατιστική σημαντικότητα p της διαφοροποίησης από τη στατιστική υπόθεση H_0

Προϋποθέσεις t - test για ένα δείγμα

Απαραίτητη προϋπόθεση είναι η δυνατότητα να αποδεχθούμε την κανονικότητα της κατανομής του εκτιμητή του αριθμητικού μέσου.

Στην πράξη, αυτήν την προϋπόθεση μπορούμε να την δεχθούμε όταν:

Η ίδια η μεταβλητή γνωρίζουμε ότι ακολουθεί την κανονική κατανομή

είτε

Το μέγεθος του δείγματος είναι μεγάλο (μεγαλύτερο του 30).

Επιπλέον, καθώς η δοκιμασία βασίζεται στο Κεντρικό Οριακό Θεώρημα και αυτό αφορά άθροισμα ανεξάρτητων και ισόνομων τυχαίων μεταβλητών, συνάγεται ότι:

Οι παρατηρήσεις πρέπει να είναι ανεξάρτητες μεταξύ τους.

Δοκιμασία Student (t - test) για ένα δείγμα

Άσκηση 1

Μία έρευνα ήθελε να αξιολογήσει τη γνώμη των καταναλωτών για μία νέα μπύρα. Δείγμα 55 ατόμων δοκίμασε την μπύρα και απάντησε σε κλίμακα Likert από -5 έως 5 τη γνώμη της για τη νέα μπύρα με την κωδικοποίηση να σημαίνει 0 να είναι ουδέτερο, οι θετικές βαθμολογίες να δείχνουν ότι τους άρεσε η μπύρα και οι αρνητικές βαθμολογίες να δείχνουν ότι δεν τους άρεσε. Βρέθηκε μέση τιμή στις κριτικές 1,1 με τυπική απόκλιση 0,4. Να ελεγχθεί η υπόθεση ότι η μέση κριτική των καταναλωτών διαφοροποιείται σημαντικά από το 0 (α = 0,05).

Λύση

$n = 55$, $\bar{x} = 1{,}1$, $s = 0{,}4$ $\quad H_0 : \mu = 0$ έναντι $H_1 : \mu \neq 0$.

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{1{,}1 - 0}{0{,}4/\sqrt{55}} = 0{,}371$$

$p = P(t > 0.371$ ή $t < -0.371) = 0.712 > 0{,}05 \Rightarrow$ η H_0 δεν απορρίπτεται.

$t_0 \sim t(54) = N(0,1)$

Δοκιμασία Student (t - test) για ένα δείγμα

Άσκηση 2

Η τελική εξέταση στη Στατιστική συνήθως έχει μέση βαθμολογία 78 / 100. Το προηγούμενο εξάμηνο, οι 120 φοιτητές που έδωσαν την εξέταση πέτυχαν μέση βαθμολογία 82 / 100 με τυπική απόκλιση 3,2. Ένας καθηγητής ισχυρίζεται ότι οι φοιτητές αυτού του έτους είναι καλύτεροι από τους προηγούμενους. Ελέγξτε τον ισχυρισμό του καθηγητή (α = 0,05).

Λύση

$n = 120 \quad \bar{x} = 0,82 \;,\; s = 3,2$

$H_0: \mu \leq 0,78$ έναντι $H_1: \mu > 0,78$

$$t_0 = \frac{\bar{x} - 0.78}{s/\sqrt{120}} = \frac{0.82 - 0.78}{3.2/\sqrt{120}} = 0,137 \;,\; t \sim t(119) \approx N(0,1)$$

$p = P(t > t_0) = P(t > 0,137) = 0,446 > 0,05 \Rightarrow$ η H_0 δεν απορρίπτεται

Δοκιμασία Student (t - test) για ένα δείγμα

Άσκηση 3

Ένας μηχανικός της Citroen ισχυρίζεται ότι ο νέος σχεδιασμός ψεκασμού καυσίμου αυξάνει τη μέση χιλιομετρική απόσταση στο C4 πάνω από το τρέχον επίπεδο των 30 χλμ ανά λίτρο. Ελέγχθηκαν είκοσι από τις νέες μηχανές και ο μέσος όρος καταγράφηκε ως 32 χλμ ανά λίτρο με τυπική απόκλιση 3,87. Αξιολογήστε αυτόν τον ισχυρισμό.

Λύση

$\bar{x} = 32 \quad s = 3,87$, $H_0: \mu \leq 30$ έναντι $H_1: \mu > 30$

$$t_0 = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{32 - 30}{3,87/\sqrt{20}} = 2,311 \quad \text{και} \quad t \sim t(19)$$

$p = P(t > t_0) = 0,016 < 0,05 \Rightarrow$ η H_0 απορρίπτεται έναντι της $H_1: \mu > 30$

Δοκιμασία z - test για μία αναλογία

Δοκιμασία z - test για μία αναλογία

Με τον όρο z – test περιγράφονται όλες οι δοκιμασίες που βασίζονται στην κανονική κατανομή για τον υπολογισμό της πιθανότητας p.

Στην πράξη αυτό σημαίνει ότι έχουμε τη δυνατότητα να υπολογίσουμε την τυπική απόκλιση του δειγματικού μέσου (αυτό που ονομάζεται και τυπικό σφάλμα standard error).

Μία ενδεικτική δοκιμασία z – test είναι η δοκιμασία z – test για μία αναλογία (one sample z – test for a proportion) η οποία χρησιμοποιείται όταν θέλουμε να ελέγξουμε την ισότητα μίας αναλογίας με μία προκαθορισμένη τιμή σε κάποιον πληθυσμό παρατηρώντας ένα αντιπροσωπευτικό δείγμα του.

Προϋποθέσεις z - test για μία αναλογία

Η δοκιμασία βασίζεται στην υπόθεση πως η διωνυμική κατανομή B(n, p) μπορεί να προσεγγιστεί από την κανονική κατανομή N(np, npq).

Ως εκ τούτου, προϋπόθεση εφαρμογής της μεθόδου είναι να μπορεί να υποτεθεί ότι

$np > 5$ και $nq = n(1 – p) > 5$.

Για παράδειγμα, αν $p \approx 0,1$ τότε απαιτείται μέγεθος δείγματος τουλάχιστον 50, ενώ αν $p \approx 0,5$, τότε η διαδικασία μπορεί να εφαρμοστεί με αξιόπιστα αποτελέσματα σε δείγμα μεγέθους $n \approx 15$.

Δοκιμασία z - test για μία αναλογία

Υπολογισμός p (στατιστική σημαντικότητα ελέγχου)

Σύμφωνα με την στατιστική υπόθεση H_0: $p = p_0$, θα πρέπει να είναι ($q_0 = 1 – p_0$)

$$\hat{p} \sim N\left(p_0, \frac{p_0 q_0}{n}\right) \Leftrightarrow \frac{\sqrt{n}\left(\hat{p} - p_0\right)}{\sqrt{p_0 q_0}} \sim N(0, 1).$$

Τώρα, αρκεί να υπολογίσουμε την πιθανότητα, αν η υπόθεση H_0 ήταν αληθής, να πάρουμε μία σημαντικά διαφορετική τιμή για το στατιστικό

$$z_0 = \frac{\sqrt{n}\left(\hat{p} - p_0\right)}{\sqrt{p_0 q_0}}$$

Αυτή η πιθανότητα συγκρινόμενη με το όριο απόρριψης α (= 0,05) οδηγεί και στην απόρριψη ή μη απόρριψη της υπόθεσης

H_0: $p = p_0$, έναντι της **H_1: $p \neq p_0$** (δίπλευρος έλεγχος), ή

H_0: $p \leq p_0$, έναντι της **H_1: $p > p_0$,** ή **H_0: $p \geq p_0$,** έναντι της **H_1: $p < p_0$,** (μονόπλευρος έλεγχος)

Δοκιμασία z - test για μία αναλογία

Παράδειγμα

Ένα τυχαίο δείγμα βρήκε ότι μεταξύ 25.468 νεογέννητων παιδιών γεννήθηκαν 13.173 αγόρια. Η αναλογία του δείγματος των αγοριών ήταν 0,5172. Ελέγξτε αν είναι πιο πιθανό για τα νεογέννητα μωρά να είναι αγόρια παρά κορίτσια (α = 0,05);

p : ποσοστό αγοριών στον πληθυσμό.

$H_0: p \leq 0.5$ έναντι $H_1: p > 0.5$

Δοκιμασία z - test για μία αναλογία

Παράδειγμα

Ένα τυχαίο δείγμα βρήκε ότι μεταξύ 25.468 νεογέννητων παιδιών γεννήθηκαν 13.173 αγόρια. Η αναλογία του δείγματος των αγοριών ήταν 0,5172. Ελέγξτε αν είναι πιο πιθανό για τα νεογέννητα μωρά να είναι αγόρια παρά κορίτσια (α = 0,05);

Λύση

Αν p είναι η (άγνωστη) αναλογία των αγοριών σε όλον τον πληθυσμό, ελέγχουμε τη στατιστική υπόθεση

H_0: p = 0,5 (= p_0), έναντι της ερευνητικής υπόθεσης: H_1: p ≠ 0,5.

Υπολογίζουμε:

$$z_0 = \frac{\sqrt{25.468}\left(0{,}5172 - 0{,}5\right)}{\sqrt{0{,}5 \cdot 0{,}5}} = 5{,}49$$

Είναι p = P(|z| > z_0) = P(|z| > 5,49) < 0,001. Συνάγουμε ότι η μηδενική υπόθεση απορρίπτεται, δηλαδή, ότι η αναλογία των νεογέννητων αγοριών είναι σημαντικά διαφορετική από το 50%.

z - test vs δοκιμασία χ^2 ως έλεγχος προσαρμογής

z - test vs δοκιμασία χ^2 ως έλεγχος προσαρμογής

Η δοκιμασία χ^2 ως έλεγχος προσαρμογής είναι ισοδύναμη με το z – test για τον έλεγχο της υπόθεσης πως η αναλογία είναι ίση με κάποια συγκεκριμένη τιμή. Ενδεικτικά, αν
H_0: $p = p_0$, τότε διαμορφώνουμε τον πίνακα:

	Χαρακτηριστικό		
	Ναι	Όχι	Σύνολο
Παρατηρήσεις	n_1	n_2	N
Αναμενόμενες H_0	p_0N	q_0N	N

Υπολογίζουμε, ότι:

$$\chi_0^2 = \frac{(n_1 - p_0 \cdot N)^2}{p_0 \cdot N} + \frac{(n_2 - q_0 \cdot N)^2}{q_0 \cdot N} = \frac{N^2(\hat{p} - p_0)^2}{p_0 \cdot N} + \frac{N^2(1 - \hat{p} - (1 - p_0))^2}{q_0 \cdot N} = \frac{N(\hat{p} - p_0)^2}{p_0 q_0} = z_0^2.$$

Στην πράξη οι δύο δοκιμασίες δίνουν το ίδιο αποτέλεσμα, με την επισήμανση πως η δοκιμασία χ^2 είναι από τη φύση της δίπλευρος έλεγχος, ενώ το z – test δίνει την δυνατότητα μονόπλευρης εναλλακτικής υπόθεσης.

Ασκήσεις

(β)

	Α	Θ	Σύνολο
O_i	11	39	50
E_i	7,5	42,5	50
	−3,5	3,5	

$$\chi_0^2 = \frac{(-3,5)^2}{7,5} + \frac{3,5^2}{42,5} = 1,922 \;,\; \chi^2 \sim \chi^2(1)$$

$$p = P(\chi^2 > 1,922) = 0.166$$

Άσκηση 1

Ένας δάσκαλος εκτιμά πως το 85% των μαθητών επιθυμεί να επισκεφθεί το ζωολογικό κήπο. Σε δείγμα 50 μαθητών οι 39 αποκρίθηκαν θετικά στην ερώτηση που τους έκανε ο δάσκαλος. Να ελεγχθεί ο ισχυρισμός του δασκάλου σε επίπεδο εμπιστοσύνης 95% (δηλαδή α = 0,05),

(α) Με δοκιμασία z-test για την αναλογία πληθυσμού.

(β) Με τη δοκιμασία χι-τετράγωνο ως έλεγχο ομοιογένειας.

Λύση

$n = 50 \;,\; \hat{p} = \frac{39}{50} = 0,78$. $H_0 : p = 0,85$ έναντι $H_1 : p \neq 0,85$.

(α) $$z_0 = \frac{\hat{p} - p_0}{\sqrt{p_0 q_0}/\sqrt{n}} = \frac{0,78 - 0,85}{\sqrt{0,85 \cdot 0,15}/\sqrt{50}} = -1,386$$

$z \sim N(0,1)$, $p = P(z > |z_0|$ ή $z < -|z_0|) = P(z > 1,386$ ή $z < -1,386)$

$= 0,166 > 0,05$ άρα η H_0 δεν απορρίπτεται!

Ασκήσεις

Άσκηση 2

Μία εταιρεία που εισάγει κινητά τηλέφωνα υποστηρίζει ότι το ποσοστό των οικογενειών που έχουν τρία κινητά τηλέφωνα είναι 30%. Μια εταιρεία κινητής τηλεφωνίας έχει λόγους να πιστεύει ότι το ποσοστό δεν είναι 30%. Πριν ξεκινήσουν μια μεγάλη διαφημιστική καμπάνια, πραγματοποιούν μία έρευνα σε 150 νοικοκυριά και βρίσκουν 43 από αυτά να έχουν τρία κινητά τηλέφωνα. Να ελεγχθεί ο ισχυρισμός της εισαγωγικής εταιρείας σε επίπεδο εμπιστοσύνης 95%,

(α) Με δοκιμασία z-test για την αναλογία πληθυσμού.

(β) Με τη δοκιμασία χι-τετράγωνο ως έλεγχο ομοιογένειας.

Λύση

Ασκήσεις

Άσκηση 3

Μία εμπορική εταιρεία θεωρεί πως περισσότερο από το 70% του πληθυσμού είναι σύμφωνο με την απαγόρευση της πώλησης αεροζόλ. Σε μια έρευνα, 813 από τους 1084 ερωτηθέντες υποστήριξαν την απαγόρευση των αεροζόλ. Ελέγξτε τον ισχυρισμό της εταιρείας.

Λύση

$H_0: p \leq 0.7$ έναντι της $H_1: p > 0.7$. $\hat{p} = \frac{813}{1.084} = 0,75$

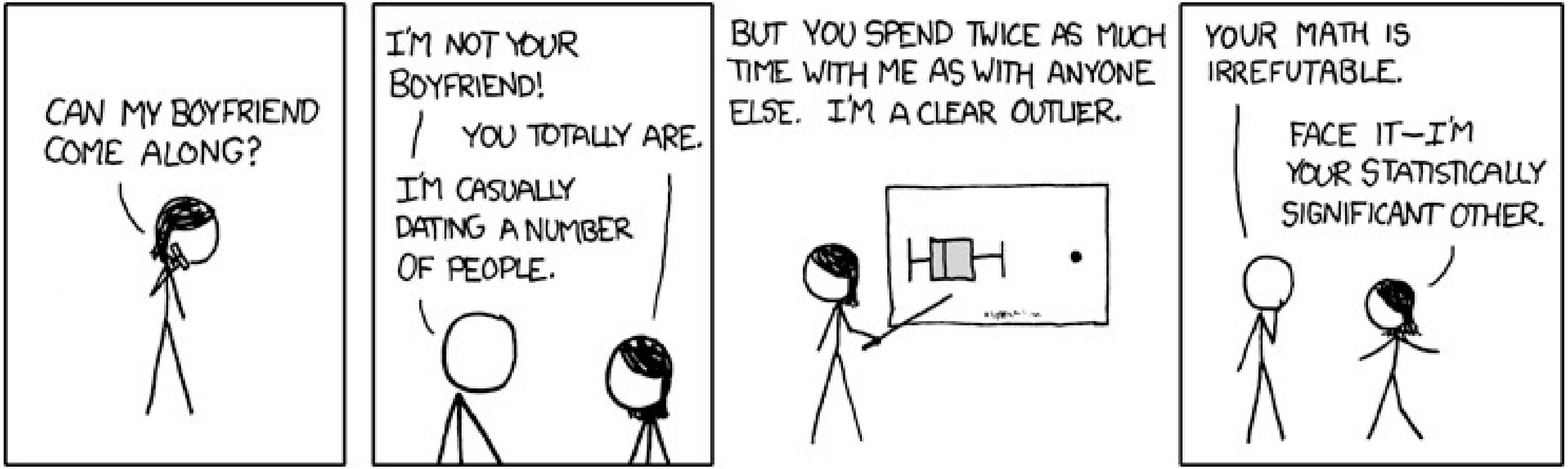
$p_0 = 0.7$, $\sqrt{p_0 q_0}/\sqrt{n} = \sqrt{0.7 \cdot 0.3}/\sqrt{1.084}$

$$z_0 = \frac{\hat{p} - p_0}{\sqrt{p_0 \cdot q_0}/\sqrt{n}} = 3,592.$$

$p = P(z > 3,592) < 0,001$ άρα η $H_0: p \leq 0.7$ απορρίπτεται έναντι της $H_1: p > 0.7$.

Ασκήσεις

Άσκηση 4
Υπάρχει η υπόνοια ότι το νόμισμα που χρησιμοποιήθηκε για την έναρξη του τοπικού αγώνα ποδοσφαίρου δεν είναι αμερόληπτο νόμισμα. Παίρνετε το ύποπτο νόμισμα και το πετάτε 100 φορές από τις οποίες οι 65 είναι Κορώνα. Είναι αυτό το αποτέλεσμα σημαντικά υψηλότερο από αυτό που θα έπρεπε να παράγεται από ένα αμερόληπτο νόμισμα;
Λύση



https://xkcd.com/539/

Δοκιμασία Student (t - test) για δύο ανεξάρτητα δείγματα

Δοκιμασία Student (t - test) για δύο ανεξάρτητα δείγματα

Η δοκιμασία t-test για δύο ανεξάρτητα δείγματα (independent samples t - test) χρησιμοποιείται όταν θέλουμε να ελέγξουμε την υπόθεση πως δύο πληθυσμοί δεν διαφέρουν σημαντικά ως προς τη μέση τιμή μίας συνεχούς μεταβλητής μετρώντας από ένα αντιπροσωπευτικό δείγμα τους.

Απαραίτητες προϋποθέσεις είναι

(α) η κανονικότητα της κατανομής της συνεχούς μεταβλητής που μελετούμε σε κάθε μία από τις δύο ανεξάρτητες ομάδες (ή μεγάλο μέγεθος δείγματος για κάθε μία ομάδα),

(β) η ομοιογένεια των δύο ομάδων (δηλαδή να έχουν τυπική απόκλιση που δεν διαφέρει σημαντικά).

Δοκιμασία Student (t - test) για δύο ανεξάρτητα δείγματα

Ομάδα 1: Μέγεθος n_1, μέση τιμή $\overline{x}_1$ και τυπική απόκλιση s_1.

Ομάδα 2: Μέγεθος n_2, μέση τιμή $\overline{x}_2$ και τυπική απόκλιση s_2.

Ελέγχουμε τη στατιστική υπόθεση

$$H_0: \mu_1 = \mu_2,$$

έναντι της ερευνητικής υπόθεσης:

- $H_1: \mu_1 \neq \mu_2$ (δίπλευρος) ή
- $H_1: \mu_1 < \mu_2$ ή
- $H_1: \mu_1 > \mu_2$ (μονόπλευρος)

Δοκιμασία Student (t - test) για δύο ανεξάρτητα δείγματα

Με την υπόθεση (β) πως οι δύο ομάδες είναι ομοιογενείς ($\overset{\sigma}{\sigma_1}$ = σ_2), είναι δυνατή η εκτίμηση της άγνωστης ~~διακύμανσης~~ τυπικη αποκλιση του κοινού πληθυσμού από τον τύπο

$$s_P = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

s_p: Pooled standard variation: Κοινή τυπική απόκλιση

Στην περίπτωση αυτή, αποδεικνύεται ότι

$$t_0 = \frac{\overline{X}_1 - \overline{X}_2}{s_p\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim t(n_1 + n_2 - 2)$$

ή ότι $t_0 = \frac{\overline{X}_1 - \overline{X}_2}{SE_p} \sim t(n_1 + n_2 - 2)$, $SE_p = s_p\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$

Απόδειξη: https://www.math.arizona.edu/~jwatkins/ttest.pdf

Δοκιμασία Student (t - test) για δύο ανεξάρτητα δείγματα

Από το σημείο αυτό, η διαδικασία ακολουθεί τα ίδια βήματα με το t-test για ένα δείγμα. Καθώς

$$t_0 = \frac{\overline{x}_1 - \overline{x}_2}{SE_p} \quad \text{και} \quad t = \frac{\overline{X}_1 - \overline{X}_2}{SE_p} \sim t(n_1 + n_2 - 2)$$

αρκεί να υπολογίσουμε την πιθανότητα να παίρναμε μία σημαντικά διαφορετική τιμή από το t_0, για το στατιστικό t αν η υπόθεση H_0: $\mu_1 = \mu_2$, ήταν αληθής.

Μονόπλευρος έλεγχος H_0: $\mu_1 \geq \mu_2$ έναντι H_1: $\mu_1 < \mu_2$: $p = P(t < t_0)$.

Μονόπλευρος έλεγχος H_0: $\mu_1 \leq \mu_2$ έναντι H_1: $\mu_1 > \mu_2$: $p = P(t > t_0)$.

Δίπλευρος έλεγχος H_0: $\mu_1 = \mu_2$ έναντι H_1: $\mu_1 \neq \mu_2$: $p = P(t < |t_0|)$.

Συγκρίνοντας το p με το επίπεδο απόρριψης α, καταλήγουμε και στο αντίστοιχο ερευνητικό συμπέρασμα.

Δοκιμασία Student (t - test) για δύο ανεξάρτητα δείγματα

Επιπλέον, γνωρίζοντας ότι

$$t = \frac{\overline{X}_1 - \overline{X}_2}{SE_p} \sim t\left(n_1 + n_2 - 2\right),\ SE_p =$$

μπορούμε να υπολογίσουμε ένα (1-α)% διάστημα εμπιστοσύνης ως εξής:

$$\left(\overline{X}_1 - \overline{X}_2 - t_{\alpha/2} s_p \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)},\ \overline{X}_1 - \overline{X}_2 - t_{\alpha/2} s_p \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}\right)$$

Ειδικότερα, αν $n_1 + n_2 > 28$, τότε $t(n_1 + n_2 – 2) \sim N(0, 1)$ και μπορούμε να γράψουμε πως το 95% δ.ε. είναι

$$\left(\overline{X}_1 - \overline{X}_2 - 1{,}96\, s_p \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)},\ \overline{X}_1 - \overline{X}_2 - 1{,}96\, s_p \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}\right)$$

Δοκιμασία Student (t - test) για δύο ανεξάρτητα δείγματα

Παράδειγμα

Μία αλυσίδα supermarket στην προσπάθεια της να ενισχύσει τις πωλήσεις επιθυμεί να ελέγξει ποια από τις δύο παρακάτω προτεινόμενες πρακτικές είναι περισσότερο αποδοτική.

Πρακτική Α: επιστροφή 4% στις αγορές των πελατών όταν αυτοί χρησιμοποιούν πιστωτική κάρτα μίας τράπεζας που συνεργάζεται με την εταιρεία ή

Πρακτική Β: προσφορά κουπόνια για έκπτωση σε αγορές των προϊόντων που εμπορεύεται.

Για να συγκρίνει την αποτελεσματικότητα των δύο αυτών διαφορετικών πρακτικών επιλέγει τυχαία 21 συχνούς πελάτες τους οποίους χωρίζει με τυχαίο τρόπο σε δύο ομάδες (Ομάδα Α και Β) των 10 και 11 ατόμων αντίστοιχα και παρακολουθεί τις αγορές τους για ένα έτος. Τα αποτελέσματα είναι (σε ευρώ):

Ομάδα Α : 2233, 2327, 1280, 1477, 1461, 1495, 1950, 1857, 1471, 1567, 1627

Ομάδα Β: 1404, 1514, 1730, 1610, 1854, 1107, 2145, 784, 1410, 2226

Δοκιμασία Student (t - test) για δύο ανεξάρτητα δείγματα

Ομάδα Α: $n_1 = 11$, $\bar{x}_1 = 1.704,1$ €, $s_1 = 341,3$ €

Ομάδα Β: $n_2 = 10$, $\bar{x}_2 = 1.578,4$ €, $s_2 = 441,8$ €

Η μέση τιμή των αγορών της ομάδας Α που χρησιμοποίησε την πιστωτική κάρτα είναι 125,7 ευρώ μεγαλύτερη από την μέση τιμή της ομάδας Β στην οποία είχαν σταλεί κουπόνια.
Είναι όμως η διαφορά των 125,7 ευρώ τόσο σημαντική ώστε να συνάγουμε πως η πιστωτική κάρτα είναι περισσότερο αποτελεσματική στο σύνολο των πελατών ή μήπως αυτή η διαφορά οφείλεται στο σφάλμα που αντιστοιχεί στην τυχαία επιλογή των πελατών κάθε ομάδας;

Περισσότερο μεθοδολογικά, αν μ_A και μ_B είναι οι μέσες καταναλώσεις που αναμένουμε να έχουν οι πελάτες των ομάδων Α και Β αντίστοιχα, αναζητούμε την απάντηση στην ερώτηση σχετικά με το αν απορρίπτεται ή όχι η στατιστική υπόθεση:

H_0: $\mu_A = \mu_B$, έναντι της ερευνητικής υπόθεσης: **H_1: $\mu_A \neq \mu_B$.**

Δοκιμασία Student (t - test) για δύο ανεξάρτητα δείγματα

Ομάδα A: $n_1 = 11$, $\overline{x}_1 = 1.704,1$ €, $s_1 = 341,3$ €
Ομάδα B: $n_2 = 10$, $\overline{x}_2 = 1.578,4$ €, $s_2 = 441,8$ €

Υπολογίζουμε:

$$SE_p = \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} = \sqrt{\left(\frac{1}{11} + \frac{1}{10}\right)\frac{(11 - 1)341,3^2 + (10 - 1)441,8^2}{11 + 10 - 2}} = 171,3$$

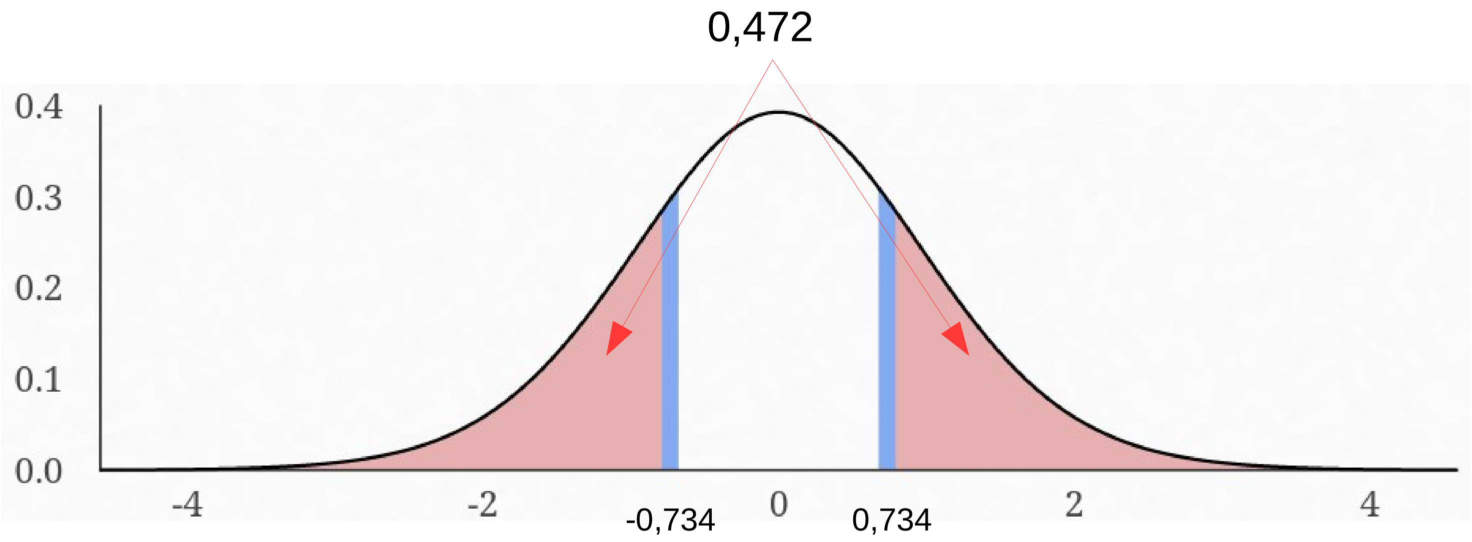
$$t_0 = \frac{\overline{x}_1 - \overline{x}_2}{SE_p} = \frac{1.704,1 - 1.578,4}{171,3} = 0,734 \quad \text{και} \quad t = \frac{\overline{X}_1 - \overline{X}_2}{171,3} \sim t(19)$$

Δίπλευρος έλεγχος H_1: $\mu_1 \neq \mu_2$: $p = P(|t| > t_0) = P(|t| > 0,734) = 0,472 = 47,2\%$. Δηλαδή: Αν λαμβάναμε 100 διαφορετικά δείγματα από 11 και 10 άτομα αντίστοιχα από τις δύο ανεξάρτητες ομάδες τότε στα 47,2 από αυτά θα είχαμε διαφορά τουλάχιστον τόσο μεγάλη όσο αυτή που παρατηρήθηκε στο δείγμα μας. Συμπερασματικά:

Βρήκαμε ότι $p = 0,472 > 0,05$, άρα η στατιστική υπόθεση H_0: $\mu_A = \mu_B$, δεν απορρίπτεται έναντι της ερευνητικής υπόθεσης: H_1: $\mu_A \neq \mu_B$.

Δοκιμασία Student (t - test) για δύο ανεξάρτητα δείγματα

Δίπλευρος έλεγχος H_1: $\mu_1 \neq \mu_2$: $p = P(|t| > t_0) = P(|t| > 0{,}734) = 0{,}472 = 47{,}2\%$.



Δοκιμασία Student (t - test) για δύο ανεξάρτητα δείγματα

Κώδικας R για την επαλήθευση των υπολογισμών

```
purchases.A = c(2233, 2327, 1280, 1477, 1461, 1495, 1950, 1857, 1471, 1567, 1627)
purchases.B = c(1404, 1514, 1730, 1610, 1854, 1107, 2145, 784, 1410, 2226)

n1 <- length(purchases.A)
n2 <- length(purchases.B)

var1 <- var(purchases.A)
var2 <- var(purchases.B)

pooled <- ((n1-1)*var1 + (n2-1)*var2) / (n1+n2-2)

s_p = ((1/n1 + 1/n2)*pooled)^0.5

s_p

t.test(purchases.A, purchases.B, var.equal = TRUE)
```

$p = 0,031 < 0,05 \Rightarrow$ η $H_0 : \mu_1 = \mu_2$ απορρίπτεται

Δοκιμασία Student (t - test) για δύο ανεξάρτητα δείγματα

Άσκηση 1

Ένας δάσκαλος θέλει να δει αν υπάρχει σημαντική διαφοροποίηση μεταξύ αγοριών και κοριτσιών στη γλώσσα. Εξετάζει τους μαθητές του (16 αγόρια και 16 κορίτσια) σε μία τυποποιημένη γραπτή δοκιμασία και βγάζει τα εξής αποτελέσματα:

Ομάδα Αγοριών: $n_1 = 16$, $x_1 = 74,8$, $s_1 = 6,8$. Ομάδα Κοριτσιών: $n_2 = 16$, $x_2 = 81,4$, $s_2 = 9,5$.

Να βρείτε αν οι δύο ομάδες διαφέρουν σημαντικά σε επίπεδο σημαντικότητας 0,05.

Λύση

$n_1 = 16$ $\quad n_2 = 16$

$\bar{x}_1 = 74,8$ $\quad \bar{x}_2 = 81,4$

$s_1 = 6,8$ $\quad s_2 = 9,5$

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

$$t_0 = \frac{\bar{x}_1 - \bar{x}_2}{SE_p} = \frac{-6,6}{2,921} = -2,26$$

$$SE_p = \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \cdot \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2}} = 2,921$$

$t \sim t(30)$

$p = P(t < -2,26 \text{ ή } t > 2,26) = 0.031$

Δοκιμασία Student (t - test) για δύο ανεξάρτητα δείγματα

Άσκηση 1

Ένας δάσκαλος θέλει να δει αν υπάρχει σημαντική διαφοροποίηση μεταξύ αγοριών και κοριτσιών στη γλώσσα. Εξετάζει τους μαθητές του (16 αγόρια και 16 κορίτσια) σε μία τυποποιημένη γραπτή δοκιμασία και βγάζει τα εξής αποτελέσματα:

Ομάδα Αγοριών: $n_1 = 16$, $x_1 = 74{,}8$, $s_1 = 6{,}8$. Ομάδα Κοριτσιών: $n_2 = 16$, $x_2 = 81{,}4$, $s_2 = 9{,}5$.

Να βρείτε αν οι δύο ομάδες διαφέρουν σημαντικά σε επίπεδο σημαντικότητας 0,05.

Λύση

$$SE_p = \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}} = \sqrt{\left(\frac{1}{16} + \frac{1}{16}\right)\frac{(16-1)6{,}8^2 + (16-1)9{,}5^2}{16+16-2}} = 2{,}92$$

$$t_0 = \frac{\overline{x}_1 - \overline{x}_2}{SE_p} = \frac{74{,}8 - 81{,}4}{2{,}92} = -2{,}260 \quad \text{και} \quad t = \frac{\overline{X}_1 - \overline{X}_2}{2{,}92} \sim t(30) = N(0\,,1)$$

Δίπλευρος έλεγχος H_1: $\mu_1 \neq \mu_2$: $p = P(|t| > t_0) = P(|t| > 2{,}26) = 0{,}031 = 3{,}1\%$. Βρήκαμε ότι $p = 0{,}031 < 0{,}05$, άρα η στατιστική υπόθεση H_0: $\mu_1 = \mu_2$, **απορρίπτεται** έναντι της ερευνητικής υπόθεσης: H_1: $\mu_1 \neq \mu_2$. Ιδιαίτερα, συμπεραίνουμε ότι η επίδοση των κοριτσιών είναι σημαντικά διαφορετική από την επίδοση των αγοριών.

Σημείωση: Προφανώς, απορρίπτεται και ο μονόπλευρος έλεγχος, άρα μπορούμε να δηλώσουμε ότι η η επίδοση των κοριτσιών είναι σημαντικά μεγαλύτερη από την επίδοση των αγοριών

Δοκιμασία Student (t - test) για δύο ανεξάρτητα δείγματα

Άσκηση 2

Η επίδοση δειγμάτων φοιτητών δύο διαδοχικών εξαμήνων στις εξετάσεις στην Στατιστική καταγράφηκε μεταξύ μέσων και βρέθηκαν τα εξής στοιχεία: Κανονική εξέταση: $n_1 = 130$, $\bar{x}_1 = 52{,}1$, $s_1 = 11{,}1$. Επαναληπτική: $n_2 = 35$, $\bar{x}_2 = 46{,}2$, $s_2 = 16{,}7$. Να βρείτε αν οι δύο ομάδες φοιτητών διαφέρουν σημαντικά ως προς την επίδοση σε επίπεδο σημαντικότητας 0,05.

Λύση

$n_1 = 130$, $\bar{x}_1 = 52{,}1$, $s_1 = 11{,}1$ $\qquad H_0: \mu_1 = \mu_2$ έναντι $H_1: \mu_1 \neq \mu_2$

$n_2 = 35$, $\bar{x}_2 = 46{,}2$, $s_2 = 16{,}7$

$$SE_p = \sqrt{\frac{1}{130} + \frac{1}{35}} \cdot \sqrt{\frac{129 \cdot 11{,}1^2 + 34 \cdot 16{,}7^2}{130 + 35 - 2}} = 2{,}376$$

$$t_0 = \frac{52{,}1 - 46{,}2}{2{,}376} = 2{,}483 \quad \text{και} \quad p = P(t < -2{,}483 \text{ ή } t > 2{,}483) = 0{,}019$$

$p = 0{,}019 < 0{,}05 \Rightarrow$ η H_0 απορρίπτεται

Δοκιμασία Student (t - test) για δύο ανεξάρτητα δείγματα

Άσκηση 3

Σε ένα γηροκομείο με 60 ηλικιωμένους, μία ομάδα ερευνητών ανέθεσε στους μισούς (30) την φροντίδα ενός φυτού. Τριάντα ασθενείς επιλέγονται τυχαία για να συμμετάσχουν στη μελέτη. Στη συνέχεια καταγράφηκε ο αριθμός παραπόνων στη διάρκεια μίας εβδομάδας.

- Ηλικιωμένοι που φρόντιζαν ένα φυτό: M = 16,6, SD = 5,8
- Ηλικιωμένοι που δεν φρόντιζαν ένα φυτό: M = 27,1, SD = 5,6

Να βρείτε αν οι δύο ομάδες διαφέρουν σημαντικά ως προς το πλήθος παραπόνων (α = 0,05).

Λύση

$n_1 = 30$, $\bar{x}_1 = 16{,}6$, $s_1 = 5{,}8$, $n_2 = 30$, $\bar{x}_2 = 27{,}1$, $s_2 = 5{,}6$

$$SE_p = \sqrt{\frac{1}{30} + \frac{1}{30}} \cdot \sqrt{\frac{29 \cdot 5{,}8^2 + 29 \cdot 5{,}6^2}{30+30-2}} = 1{,}485$$

$$t_0 = \frac{27{,}1 - 16{,}6}{1{,}485} = 7{,}073, \quad p = P(t > 7{,}073 \text{ ή } t < -7{,}073) < 0{,}001$$

άρα απορρίπτεται η H_0.

t - test για δύο δείγματα: ανομοιογενείς πληθυσμοί (Welch's t-test)

Όπως είδαμε οι απαραίτητες προϋποθέσεις για τη σωστή εφαρμογή της μεθόδου είναι

(α) η κανονικότητα της κατανομής της συνεχούς μεταβλητής που μελετούμε σε κάθε μία από τις δύο ανεξάρτητες ομάδες (ή μεγάλο μέγεθος δείγματος για κάθε μία ομάδα),

(β) η ομοιογένεια των δύο ομάδων (δηλαδή να έχουν τυπική απόκλιση που δεν διαφέρει σημαντικά).
Αν η προϋπόθεση (β) δεν τηρείται (άγνωστες διακυμάνσεις ή ελλιπής πληροφόρηση) τότε υπάρχει μία εναλλακτική εκδοχή της δοκιμασίας με αξιόπιστα αποτελέσματα. Στην περίπτωση αυτή το στατιστικό που υπολογίζεται είναι το

$$t_0 = \frac{\overline{x}_1 - \overline{x}_2}{s_c} \quad \text{και} \quad t = \frac{\overline{X}_1 - \overline{X}_2}{s_c} \sim t(df)$$

όπου

$$s_c = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \quad \text{και} \quad df = \frac{s_c^4}{\frac{1}{n_1 - 1} \cdot \frac{s_1^4}{n_1^2} + \frac{1}{n_2 - 1} \cdot \frac{s_2^4}{n_2^2}}$$

z - test για δύο αναλογίες

z - test για δύο αναλογίες

Μελετούμε δύο ανεξάρτητες ομάδες για την παρουσία ενός χαρακτηριστικού και **θέλουμε να ελέγξουμε την υπόθεση H_0: $p_1 = p_2$ έναντι της H_1: $p_1 \neq p_2$.**

Στην περίπτωση αυτή, αποδεικνύεται ότι $\hat{p}_1 - \hat{p}_2 \sim N\left(p_1 - p_2,\, s_D^2\right)$

όπου $s_D = \sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$, και $\hat{p} = \frac{n_1 \hat{p}_1 + n_2 \hat{p}_2}{n_1 + n_2}$

Μετά, αρκεί να υπολογιστεί το στατιστικό

$$z_0 = \frac{\hat{p}_1 - \hat{p}_2}{s_D} \sim N(0,\, 1)$$

Και από αυτό να υπολογιστεί η πιθανότητα p = $P(|z| > z_0)$ ή $P(z < z_0)$ ή $P(z > z_0)$

Σημείωση: Στη βιβλιογραφία αναφέρεται και η επιλογή $s_D = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$ η οποία ωστόσο θεωρείται υποδεέστερη ως προς την επίδοσή της.

Σχετικά: https://stats.stackexchange.com/questions/17195/is-there-a-reference-that-legitimises-the-use-of-the-unpooled-z-test-to-compare?rq=1

Ασκήσεις

1. Σε δείγμα 958 ανδρών βρέθηκαν 358 καπνιστές, ενώ σε δείγμα 869 γυναικών βρέθηκαν 267 καπνίστριες. Να ελεγχθεί η υπόθεση ότι το ποσοστό των καπνιστών είναι υψηλότερο στους άνδρες από ότι στις γυναίκες (α = 0,05).

Ασκήσεις

2. Θέλουμε να συγκρίνουμε την απόδοση δύο φαρμάκων Α και Β για τη γρίπη. Σε δείγμα 195 ατόμων με το φάρμακο Α, παρατηρήθηκαν μειωμένα συμπτώματα σε 41 άτομα, ενώ σε δείγμα 605 ατόμων με το φάρμακο Β, παρατηρήθηκαν μειωμένα συμπτώματα σε 351 άτομα. Έχουν τα δύο φάρμακα ανάλογη απόδοση; (α = 0,05).

Ασκήσεις

3. Μια έρευνα επέλεξε από 1000 άτομα σε δύο ελληνικές πόλεις Α και Β. Στη Α βρέθηκε πως το ποσοστό των ατόμων που είχαν γεννηθεί σε άλλη χώρα ήταν 6,5%, ενώ στην πόλη Β, το ποσοστό ήταν 1,7%. Να ελέγξετε την υπόθεση πως η αναλογία του πληθυσμού που έχει γεννηθεί σε άλλη χώρα είναι σημαντικά διαφορετική μεταξύ των δύο πόλεων ($\alpha = 0{,}05$).

z - test vs δοκιμασία ανεξαρτησίας χ^2

Η δοκιμασία χ^2 είναι εξίσου αποδεκτή και έγκυρη για τον έλεγχο δύο αναλογιών στην περίπτωση όπου υπάρχει πλήρως τυχαιοποιημένη κατανομή των παρατηρήσεων στις δύο ομάδες (αυτές που έχουν το χαρακτηριστικό και αυτές που δεν το έχουν).

	Ομάδα 1	Ομάδα 2	Σύνολο
Ναι	n_{11}	n_{21}	$n_{11} + n_{21}$
Όχι	n_{12}	n_{22}	$n_{21} + n_{22}$
	$N_1 = n_{11} + n_{12}$	$N_2 = n_{21} + n_{22}$	N

Η απόδειξη είναι ανάλογη με αυτήν που ακολουθήσαμε για να δείξουμε ότι το z – test για μία αναλογία είναι ισοδύναμο με τη δοκιμασία χ^2 ως έλεγχο ομοιογένειας. Είναι ωστόσο σημαντικό να τονιστεί πως η χ^2 είναι ισοδύναμος με το z – test μόνο στην περίπτωση όπου τόσο τα αθροίσματα των στηλών όσο και αυτά των γραμμών ήταν ελεύθερα να διαφοροποιούνται κατά τη δειγματοληψία. Για το λόγο αυτό δεν είναι κατάλληλη επιλογή στην περίπτωση ενός τυχαιοποιημένου πειραματικού σχεδίου (RCT) όπου το πλήθος κάθε μίας ομάδας είναι προκαθορισμένο. Στην περίπτωση αυτή, το z-test δεν μπορεί να αντικατασταθεί από τη δοκιμασία ανεξαρτησίας χ^2, η οποία δεν αρμόζει να εφαρμοστεί σε αυτήν την περίπτωση.

Σχετικοί σύνδεσμοι:

https://stats.stackexchange.com/questions/173415/at-what-level-is-a-chi2-test-mathematically-identical-to-a-z-test-of-propo (απόδειξη)

https://stats.stackexchange.com/questions/81975/the-z-test-vs-the-chi2-test-for-comparing-the-odds-of-catching-a-cold-in-2 (αναφορά RCT)

Έλεγχος προϋποθέσεων

Έλεγχος προϋποθέσεων – F test

Δοκιμασία F για τον έλεγχο της ισότητας των τυπικών αποκλίσεων

Έστω δύο ομάδες παρατηρήσεων με πλήθος στοιχείων n, m. Η δοκιμασία F, χρησιμοποιείται για τον έλεγχο της υπόθεσης πως οι δύο ομάδες είναι ομοιογενείς ($\sigma_1 = \sigma_2$). Ελέγχεται η ισοδύναμη υπόθεση πως $\sigma^2_1 / \sigma^2_2 = 1$.

H_0: $\sigma^2_1 / \sigma^2_2 = 1$ έναντι

H_1: $\sigma^2_1 / \sigma^2_2 \neq 1$ (δίπλευρο) ή H_1: $\sigma^2_1 / \sigma^2_2 > 1$ ή H_1: $\sigma^2_1 / \sigma^2_2 < 1$ (μονόπλευρο)

Θεωρούμε την τυχαία μεταβλητή $F = \frac{S_1^2}{S_2^2} = \left[\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}_1\right)\right] / \left[\frac{1}{m-1}\sum_{i=1}^{m}\left(X_i - \overline{X}_2\right)\right]$

Το S^2_1, ακολουθεί τη $\chi^2(n-1)$ κατανομή, ενώ το S^2_2, ακολουθεί τη $\chi^2(m-1)$ κατανομή, άρα το στατιστικό $F = S^2_1 / S^2_2$, ακολουθεί τη $F(n-1, m-1)$ κατανομή.

Σημείωση

Προϋπόθεση για την εφαρμογή του F test είναι η κανονικότητα των κατανομών των τιμών των δύο ομάδων.

Έλεγχος προϋποθέσεων – F test

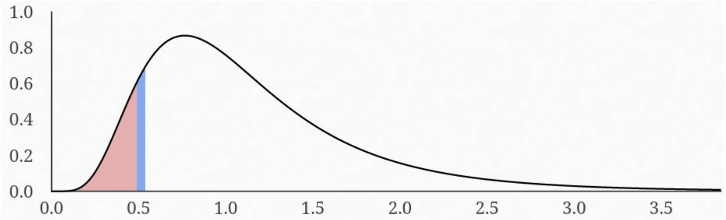
Ομάδα Αγοριών: $n_1 = 16$, $x_1 = 74{,}8$, $s_1 = 6{,}8$.
Ομάδα Κοριτσιών: $n_2 = 16$, $x_2 = 81{,}4$, $s_2 = 9{,}5$.

$$F = \frac{S_1^2}{S_2^2} \sim F(15, 15)$$

Υπολογίζουμε τα στατιστικά ελέγχου: $F_0 = \frac{s_1^2}{s_2^2} = \frac{6{,}8^2}{9{,}5^2} = 0{,}512$ και $F_1 = \frac{s_2^2}{s_1^2} = \frac{9{,}5^2}{6{,}8^2} = 1{,}952$.

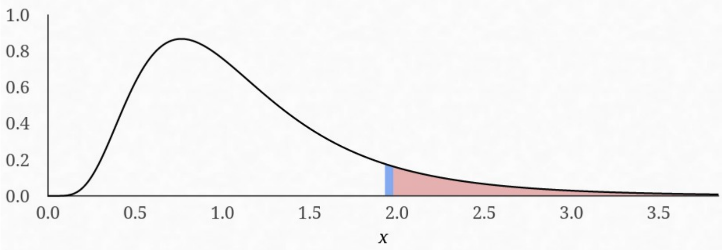
Α. Ελέγχουμε την H_0: $\sigma^2_1 / \sigma^2_2 = 1$ έναντι της H_1: $\sigma^2_1 / \sigma^2_2 < 1$ σε επίπεδο $\alpha = 0{,}05 / 2 = 0{,}025$.

$p = P(F < F_0) = P(F < 0{,}512) = 0{,}1032$



Β. Ελέγχουμε την H_0: $\sigma^2_1 / \sigma^2_2 = 1$ έναντι της H_1: $\sigma^2_1 / \sigma^2_2 > 1$ σε επίπεδο $\alpha = 0{,}05 / 2 = 0{,}025$.

$p = P(F > F_1) = P(F > 1{,}952) = 0{,}1034$.



Συμπεραίνουμε ότι οι δύο διακυμάνσεις δεν διαφοροποιούνται σημαντικά σε επίπεδο σημαντικότητας 95% (η υπόθεση H_0 δεν απορρίπτεται)

Έλεγχος προϋποθέσεων – Κατανομή τιμών

Ο έλεγχος της κατανομής από την οποία προέρχεται ένα δείγμα είναι απαραίτητος για όλες τις παραμετρικές στατιστικές δοκιμασίες. Βεβαιωνόμαστε πως μία μεταβλητή ακολουθεί κάποια θεωρητική κατανομή με τους παρακάτω τρόπους:

1. Προηγούμενες έρευνες.

2. Δειγματοληψία και παρατήρηση του ιστογράμματος και του διαγράμματος q - q.

3. Θεωρητικά επιχειρήματα που να καταδεικνύουν πως η μεταβλητή έχει τα χαρακτηριστικά που αντιστοιχούν σε μία από τις βασικές κατανομές.

4. Στατιστικοί έλεγχοι που ελέγχουν την αντίστοιχη στατιστική υπόθεση:

- Kolmogorov Smirnov
- Anderson-Darling
- Χι-τετράγωνο καλής προσαρμογής
- Shapiro-Wilk

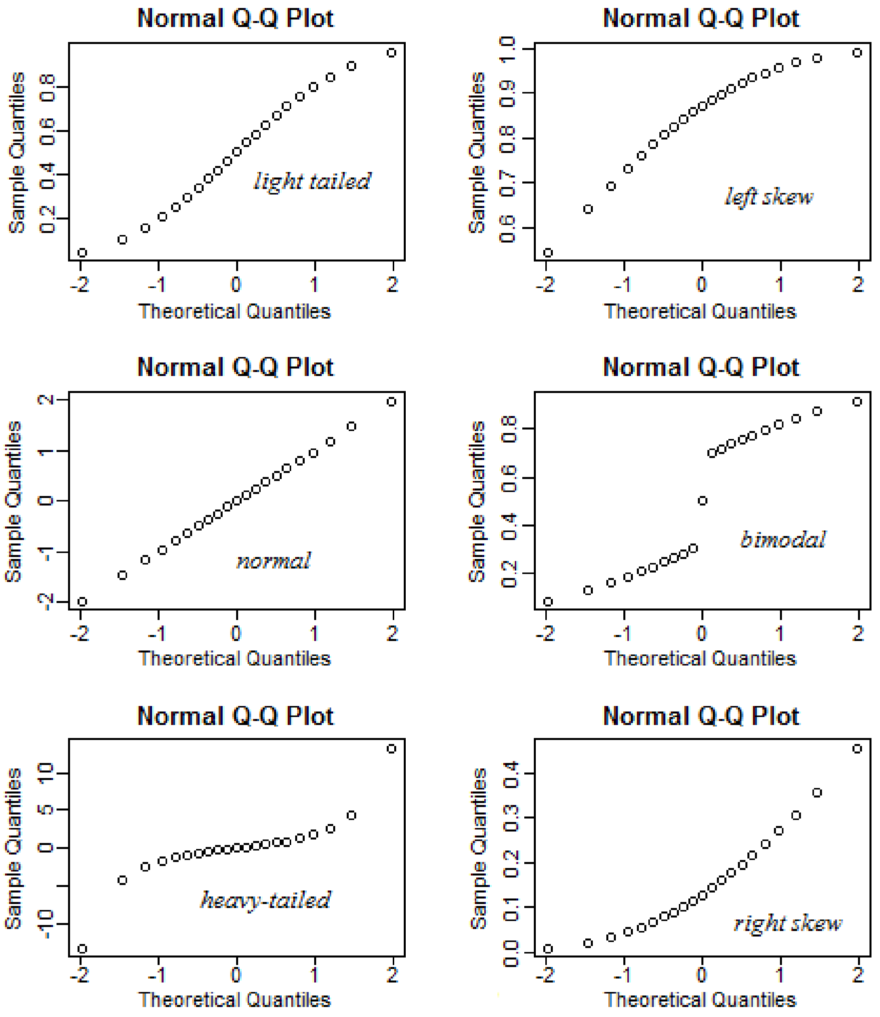
Διάγραμμα q - q

Το διάγραμμα Quantile – Quantile είναι ένα απλό διάγραμμα που προσφέρει τη δυνατότητα οπτικής επιβεβαίωσης της υπόθεσης πως το δείγμα μας ακολουθεί την κανονική (ή και όποια άλλη) κατανομή.

Τα δεδομένα σχεδιάζονται με τέτοιο τρόπο ώστε αν ακολουθούν μια θεωρητική κατανομή να πρέπει να σχηματίζουν περίπου μια ευθεία γραμμή.

Οι αποκλίσεις από αυτήν την ευθεία γραμμή υποδηλώνουν αποκλίσεις από την υποτιθέμενη κατανομή.

Ενδεικτικά στο διάγραμμα παρουσιάζεται το p – p διάγραμμα 100 τιμών που προσαρμόζονται στην κανονική κατανομή. Προσέχουμε ότι το σύνολο των σημείων δεν έχει σημαντικές αποκλίσεις από την ευθεία.



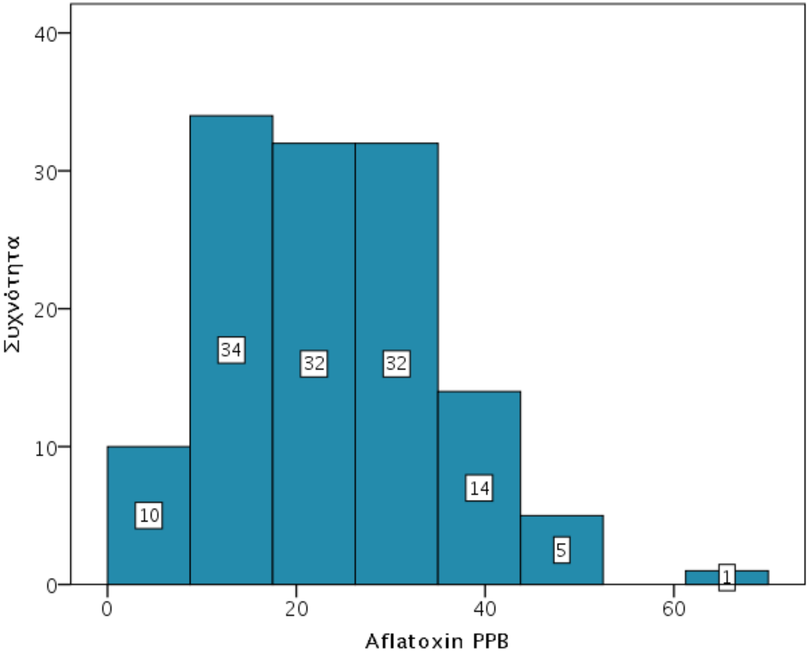
Πηγή διαγράμματος: https://stats.stackexchange.com/questions/101274/how-to-interpret-a-qq-plot/101290#101290

Συσχέτιση ιστογράμματος και q – q διαγράμματος: https://xiongge.shinyapps.io/QQplots/

Έλεγχος κανονικότητας Kolmogorov Smirnov (SPSS)

Δοκιμασία Kolmogorov Smirnov όπως υλοποιείται στο SPSS για τον έλεγχο των υποθέσεων πως ένα δείγμα ακολουθεί την κανονική ή την ομοιόμορφη κατανομή.

Έλεγχος σε 128 δείγματα από σοδιές ως προς τα επίπεδα αφλατοξίνης.



Descriptives

			Statistic	Std. Error
Aflatoxin PPB	Mean		23,8	1,0
	95% Confidence Interval for Mean	Lower Bound	21,8	
		Upper Bound	25,8	
	5% Trimmed Mean		23,2	
	Median		23,5	
	Variance		128,0	
	Std. Deviation		11,3	
	Minimum		4,0	
	Maximum		68,0	
	Range		64,0	
	Interquartile Range		16,0	
	Skewness		,7	,2
	Kurtosis		1,1	,4

Έλεγχος κανονικότητας Kolmogorov Smirnov (SPSS)

Δοκιμασία Kolmogorov Smirnov όπως υλοποιείται στο SPSS για τον έλεγχο των υποθέσεων πως ένα δείγμα ακολουθεί την κανονική ή την ομοιόμορφη κατανομή.

One-Sample Kolmogorov-Smirnov Test

		Aflatoxin PPB
N		128
Normal Parameters[a,b]	Mean	23,8203
	Std. Deviation	11,31192
Most Extreme Differences	Absolute	,070
	Positive	,070
	Negative	-,040
Kolmogorov-Smirnov Z		,797
Asymp. Sig. (2-tailed)		,549

a. Test distribution is Normal.

b. Calculated from data.

Είναι p = 0,549, άρα η
H_0: Το δείγμα προέρχεται από κανονικό πληθυσμό δεν απορρίπτεται σε επίπεδο σημαντικότητας 0,05 έναντι της H_1: όχι η H_0.

One-Sample Kolmogorov-Smirnov Test 2

		Aflatoxin PPB
N		128
Uniform Parameters[a,b]	Minimum	4,00
	Maximum	68,00
Most Extreme Differences	Absolute	,414
	Positive	,414
	Negative	-,016
Kolmogorov-Smirnov Z		4,685
Asymp. Sig. (2-tailed)		,000

a. Test distribution is Uniform.

b. Calculated from data.

Είναι p < 0,001, άρα η
H_0: Το δείγμα προέρχεται από ομοιόμορφο πληθυσμό απορρίπτεται σε επίπεδο σημαντικότητας 0,05 έναντι της H_1: όχι η H_0.

Αδυναμίες των ελέγχων Kolmogorov-Smirnov και Shapiro-Wilk

Η στατιστική σημαντικότητα p οποιουδήποτε στατιστικού ελέγχου είναι αντιστρόφως ανάλογη με το μέγεθος του/των δείγματος/ων από τα οποία υπολογίστηκε.

Στα πλαίσια ενός ελέγχου κανονικότητας αυτή η ιδιαιτερότητα σημαίνει πως μία πολύ μικρή και μη ουσιαστική απόκλιση από την κανονικότητα μπορεί να προκύψει στατιστικά σημαντική αν το μέγεθος του δείγματος είναι εξαιρετικά μεγάλο ενώ μία σημαντική απόκλιση από την κανονική κατανομή μπορεί να μην ανιχνευθεί αν το μέγεθος του δείγματος είναι μικρό.

Στην πράξη, ο ερευνητής είναι καλό να διεξάγει έναν έλεγχο κανονικότητας για να τεκμηριώνει τη χρήση μίας δοκιμασίας σε κάποιον εξωτερικό παρατηρητή όπως ο αναγνώστης ενός περιοδικού στο οποίο θα δημοσιευθεί το αποτέλεσμά του ή στην διατριβή του ωστόσο θα πρέπει να έχει πείσει τον εαυτό του για την κανονικότητα της κατανομής με στέρεα θεωρητικά επιχειρήματα τα οποία θα καταδεικνύουν πως από τη φύση της η μεταβλητή ακολουθεί την κανονική κατανομή δηλαδή πως οι τιμές της επηρεάζονται από πολλά ανεξάρτητα και αθροιστικά μεταξύ τους σφάλματα.

Δοκιμασία Student (t - test) για ζεύγη παρατηρήσεων

Δοκιμασία Student (t - test) για ζεύγη παρατηρήσεων

Η δοκιμασία t - test για ζεύγη παρατηρήσεων (paired samples t - test) χρησιμοποιείται όταν θέλουμε να ελέγξουμε τη μεταβολή της μέσης τιμής μεταξύ δύο μετρήσεων ύστερα από μία παρέμβαση σε μία συνεχή παράμετρο ενός πληθυσμού παρατηρώντας τι συμβαίνει σε ένα δείγμα του πληθυσμού.

Απαραίτητες προϋποθέσεις είναι

(α) η κανονικότητα της κατανομής των διαφορών των τιμών της συνεχούς μεταβλητής που μελετούμε (ή μεγάλο μέγεθος παρατηρήσεων),

(β) η απουσία ιδιαζόντων ή ακραίων παρατηρήσεων

Δοκιμασία Student (t - test) για ζεύγη παρατηρήσεων

Μέγεθος ομάδας: n

Πριν την παρέμβαση: Μέση τιμή x_1 και τυπική απόκλιση s_1.

Μετά την παρέμβαση: Μέση τιμή x_2 και τυπική απόκλιση s_2.

Ελέγχουμε τη στατιστική υπόθεση

$$H_0: \mu_1 = \mu_2,$$

έναντι της ερευνητικής υπόθεσης:

$$H_1: \mu_1 \neq \mu_2 \text{ (δίπλευρος) ή } H_1: \mu_1 < \mu_2 \text{ ή } H_1: \mu_1 > \mu_2 \text{ (μονόπλευρος)}$$

Θεωρούμε την τυχαία μεταβλητή $X = X_1 – X_2$. Με την υπόθεση πως η μηδενική υπόθεση H_0: $\mu_1 = \mu_2$, είναι αληθής και πως δεν μεταβάλλεται η διακύμανση πριν και μετά, έχουμε:

$$\overline{X}_1 - \overline{X}_2 \sim N\left(0, \frac{\sigma_D^2}{n}\right)$$

όπου σ_D είναι η διακύμανση των διαφορών των τιμών. Στη συνέχεια:

$$\overline{X}_1 - \overline{X}_2 \sim N\left(0, \sigma_D^2\right) \Leftrightarrow \frac{\overline{X}_1 - \overline{X}_2}{\sigma_D/\sqrt{n}} \sim N(0, 1)$$

Δοκιμασία Student (t - test) για ζεύγη παρατηρήσεων

Καταλήγουμε, ότι:

$$\frac{\overline{X}_1 - \overline{X}_2}{\sigma_D/\sqrt{n}} \sim N(0,\, 1) \Leftrightarrow \frac{\overline{X}_1 - \overline{X}_2}{s_D/\sqrt{n}} \sim t(n-1)$$

Από το σημείο αυτό, η διαδικασία ακολουθεί τα ίδια βήματα με το t-test για ένα δείγμα.

Καθώς

$$t_0 = \frac{\overline{x}_1 - \overline{x}_2}{s_D/\sqrt{n}} \quad \text{και} \quad t = \frac{\overline{X}_1 - \overline{X}_2}{s_D/\sqrt{n}} \sim t(n-1)$$

αρκεί να υπολογίσουμε την πιθανότητα να παίρναμε μία σημαντικά διαφορετική τιμή από το t_0, για το στατιστικό t αν η υπόθεση H_0: $\mu_1 = \mu_2$, ήταν αληθής.

- Μονόπλευρος έλεγχος H_0: $\mu_1 \geq \mu_2$ έναντι H_1: $\mu_1 < \mu_2$: $p = P(t < t_0)$.
- Μονόπλευρος έλεγχος H_0: $\mu_1 \leq \mu_2$ έναντι H_1: $\mu_1 > \mu_2$: $p = P(t > t_0)$.
- Δίπλευρος έλεγχος H_0: $\mu_1 = \mu_2$ έναντι H_1: $\mu_1 \neq \mu_2$: $p = P(t < |t_0|)$.

Συγκρίνοντας το p με το επίπεδο απόρριψης α, καταλήγουμε και στο αντίστοιχο ερευνητικό συμπέρασμα.

Δοκιμασία Student (t - test) για ζεύγη παρατηρήσεων

Παράδειγμα

Με σκοπό την αξιολόγηση μίας δίαιτας ως προς τη μείωση του βάρους όσων την εφαρμόζουν, επιλέχθηκαν 16 ενήλικες εθελοντές οι οποίοι την ακολούθησαν πιστά για 6 μήνες. Το βάρος μετρήθηκε στην αρχή και στο τέλος της διαδικασίας και τα αποτελέσματα ήταν τα εξής:

Πριν:89.8, 107.5, 105.7, 81.2, 99.3, 76.7, 100.7, 75.7, 90.3, 105.7, 81.2, 71.7, 71.2, 98, 116.6,68.5

Μετά: 87.1, 102.1, 102.5, 78, 97.1, 73, 95.3, 73, 87.5, 102.5, 78.5, 69.9, 64.9, 93.4, 112.9, 63.5

Να βρεθεί αν η δίαιτα επηρέασε το βάρος των εθελοντών σε επίπεδο σημαντικότητας 0,05.

Δοκιμασία Student (t - test) για ζεύγη παρατηρήσεων

Παράδειγμα

Με σκοπό την αξιολόγηση μίας δίαιτας ως προς τη μείωση του βάρους όσων την εφαρμόζουν, επιλέχθηκαν 16 ενήλικες εθελοντές οι οποίοι την ακολούθησαν πιστά για 6 μήνες. Το βάρος μετρήθηκε στην αρχή και στο τέλος της διαδικασίας και τα αποτελέσματα ήταν τα εξής:

Πριν:89.8, 107.5, 105.7, 81.2, 99.3, 76.7, 100.7, 75.7, 90.3, 105.7, 81.2, 71.7, 71.2, 98, 116.6,68.5

Μετά: 87.1, 102.1, 102.5, 78, 97.1, 73, 95.3, 73, 87.5, 102.5, 78.5, 69.9, 64.9, 93.4, 112.9, 63.5

Να βρεθεί αν η δίαιτα επηρέασε το βάρος των εθελοντών σε επίπεδο σημαντικότητας 0,05.

Λύση

Οι διαφορές του βάρους πριν και μετά τη δίαιτα είναι

-2.7 -5.4 -3.2 -3.2 -2.2 -3.7 -5.4 -2.7 -2.8 -3.2 -2.7 -1.8 -6.3 -4.6 -3.7 -5.0

Η μέση τιμή των διαφορών είναι -3,663 Kg με τυπική απόκλιση 1,302 Kg. Υπολογίζουμε:

$$t_0 = \frac{\overline{x}_1 - \overline{x}_2}{s_D/\sqrt{n}} = \frac{-3,663}{1,302/\sqrt{16}} = -11,253 \quad \text{και} \quad t \sim t(15)$$

Μονόπλευρος έλεγχος H_1: $\mu_1 < \mu_2$: $p = P(t < t_0) = P(t < -11,253) = 1,039 \cdot 10^{-8} < 0,001$.

Συμπέρασμα: Η στατιστική υπόθεση H_0 απορρίπτεται ($p < 0,001$). Η δίαιτα μειώνει σημαντικά το βάρος όσων την ακολουθούν σε επίπεδο σημαντικότητας 0,05.

Δοκιμασία Student (t - test) για ζεύγη παρατηρήσεων

Κώδικας R για την επαλήθευση των υπολογισμών

```
weight1 = c(89.8, 107.5, 105.7, 81.2, 99.3, 76.7, 100.7, 75.7, 90.3, 105.7, 81.2, 71.7, 71.2, 98,
116.6, 68.5)
weight2 = c(87.1, 102.1, 102.5, 78, 97.1, 73, 95.3, 73, 87.5, 102.5, 78.5, 69.9, 64.9, 93.4, 112.9,
63.5)

diff = weight2 - weight1

mean(diff)
sd(diff)

t.test(weight1, weight2, paired = TRUE)
```

Δοκιμασία Student (t - test) για ζεύγη παρατηρήσεων

Άσκηση 1
Ένας καθηγητής θέλει να συγκρίνει τη δυσκολία δύο θεμάτων. Για το λόγο αυτό, εξετάζει την ίδια ομάδα από 50 φοιτητές και στα δύο θέματα και καταγράφει τις βαθμολογίες. Βρίσκει:
Εξέταση 1: Μέση επίδοση 78,1, Τυπική απόκλιση 12,7
Εξέταση 2: Μέση επίδοση 79,5, Τυπική απόκλιση 9,1

Διαφορά εξετάσεων 1 και 2: Μέση διαφορά 1,4 βαθμοί με τυπική απόκλιση 0,2 βαθμοί.
Να βρεθεί αν υπάρχει σημαντική διαφορά στη δυσκολία των θεμάτων σε επίπεδο εμπιστοσύνης 95%.

$\bar{X}_1 - \bar{X}_2 = 1,4$, $S_D = 0,2$ $H_0 : \mu_1 = \mu_2$ έναντι $H_1 : \mu_1 \neq \mu_2$

$$t_0 = \frac{\bar{X}_1 - \bar{X}_2}{S_D/\sqrt{50}} = \frac{1,4}{0,2/\sqrt{50}} = 49,5 \qquad t \sim t(49)$$

$p = P(t > 49,5 \text{ ή } t < -49,5) < 0,001 \Rightarrow$ η H_0 : απορρίπτεται

Δοκιμασία Student (t - test) για ζεύγη παρατηρήσεων

Άσκηση 2

Ένας ερευνητής υποθέτει ότι η ηλεκτρική διέγερση μία ορισμένης περιοχής του εγκεφάλου σε αρουραίους θα έχει ως αποτέλεσμα τη μείωση της πρόσληψης τροφής σε αρουραίους. Τριάντα (30) αρουραίοι υποβάλλονται σε χειρουργική επέμβαση και ένα ηλεκτρόδιο εμφυτεύεται στο κατάλληλο σημείο. Μετά από μια περίοδο ανάρρωσης δέκα ημερών, οι αρουραίοι ελέγχονται για τον αριθμό των κομματιών σοκολάτας που καταναλώθηκαν κατά τη διάρκεια μιας χρονικής περιόδου 10 λεπτών, τόσο με όσο και χωρίς ηλεκτρική διέγερση. Βρέθηκε ότι:

Κατανάλωση χωρίς διέγερση: Μέσο πλήθος 8,6 με τυπική απόκλιση 1,3
Κατανάλωση με διέγερση: Μέσο πλήθος 7,6 με τυπική απόκλιση 1,2.

Διαφορά καταναλώσεων: Μέση διαφορά: 1 με τυπική απόκλιση 0,15.

Να βρείτε αν υπάρχει σημαντική διαφορά στην ποσότητα τροφής σε επίπεδο εμπιστοσύνης 95%.

Συνοψίζοντας...

Student (t – test)

- t – test για ένα δείγμα (H_0: $\mu = \mu_0$)
- t – test για δύο ανεξάρτητα δείγματα (H_0: $\mu_1 = \mu_2$)
- t – test για ζεύγη παρατηρήσεων (H_0: $\mu_{ΠΡΙΝ} = \mu_{ΜΕΤΑ}$)

Δοκιμασία χ^2

- χ^2 ως έλεγχος καλής προσαρμογής (H_0: $p_1 = p_{1,0}$, $p_1 = p_{2,1}$, …, $p_k = p_{k,0}$) $\xrightarrow{k = 2}$ z – test για μία αναλογία
- χ^2 ως έλεγχος ανεξαρτησίας (H_0: Οι X (k τιμές) και Y (m τιμές) είναι στοχαστικά ανεξάρτητες) $\xrightarrow{k = m = 2}$ z – test για ισότητα δύο αναλογιών

z – test

- z – test για μία αναλογία (H_0: $p = p_0$)
- z – test για ισότητα δύο αναλογιών (H_0: $p_1 = p_2$)

Σφάλματα Στατιστικών Ελέγχων

Σφάλματα Στατιστικών Ελέγχων

Ο έλεγχος της υπόθεσης H_0 έναντι της εναλλακτικής H_1, μπορεί να καταλήξει σε δύο πιθανά αποτελέσματα

1. Να απορριφθεί η H_0 γιατί θα έχει υπολογιστεί $p < α$.
2. Να μην απορριφθεί η H_0 γιατί θα έχει υπολογιστεί $p > α$.

Ωστόσο, την αλήθεια δεν την ξέρουμε. Οπότε ανακύπτουν οι εξής περιπτώσεις:

Απόρριψη της H_0 ενώ στην πραγματικότητα αυτή ισχύει: **Σφάλμα τύπου Ι**
Μη απόρριψη της H_0 ενώ στην πραγματικότητα αυτή δεν ισχύει: **Σφάλμα τύπου ΙΙ**

Από αυτά τα σφάλματα, μπορούμε να ελέγξουμε το Σφάλμα τύπου Ι το οποίο είναι απλά το όριο απόρριψης α, που ορίζουμε πριν τον έλεγχο. Το σφάλμα τύπου ΙΙ δεν μπορούμε να το προσδιορίσουμε επακριβώς στην γενική περίπτωση. Η ισχύς του στατιστικού ελέγχου είναι

Ισχύς = 1 - Σφάλμα τύπου ΙΙ

Η ισχύς εκφράζει την πιθανότητα ορθής απόρριψης της H_0, όταν αυτή πράγματι δεν ισχύει.

Σφάλματα Στατιστικών Ελέγχων

Οι πιθανές περιπτώσεις παρουσιάζονται στον παρακάτω πίνακα:

Αποτέλεσμα στατιστικού ελέγχου	Τι συμβαίνει στην πραγματικότητα (και δεν θα το μάθουμε ποτέ)	
	Η H_0 είναι αληθής	Η H_0 είναι ψευδής
Απόρριψη H_0	Σφάλμα τύπου I	√
Αδυναμία απόρριψης H_0	√	Σφάλμα τύπου II

Η πιθανότητα του σφάλματος τύπου I είναι το όριο απόρριψης α, που το ορίζουμε στην αρχή της έρευνας. **Η πιθανότητα του σφάλματος τύπου II συμβολίζεται με β.**

Πιθανότητες των τεσσάρων περιπτώσεων	Τι συμβαίνει στην πραγματικότητα (και δεν θα το μάθουμε ποτέ)	
	Η H_0 είναι αληθής	Η H_0 είναι ψευδής
Απόρριψη H_0	α	1 – α
Αδυναμία απόρριψης H_0	1 – β Ισχύς (power) του ελέγχου	β

Ποιο σφάλμα είναι πιο σημαντικό;

Η απάντηση σε αυτό το ερώτημα έχει να κάνει με τη φύση του προβλήματος που μελετούμε. Για παράδειγμα, αν ο στατιστικός έλεγχος είναι

H_0: Το σχοινί αναρρίχησης αντέχει τουλάχιστον 300 κιλά.
H_1: Το σχοινί αναρρίχησης αντέχει το πολύ 300 κιλά.

τότε,

Σφάλμα τύπου I:
Το σχοινί τύπου Χ αντέχει 300 κιλά αλλά εμείς απορρίπτουμε την H_0.

Σφάλμα τύπου II:
Το σχοινί τύπου Χ δεν αντέχει 300 κιλά αλλά εμείς δεν απορρίπτουμε την H_0.

Φανερά, στην περίπτωση αυτή το σφάλμα τύπου II είναι πιο σημαντικό.

Ποιο σφάλμα είναι πιο σημαντικό;

Αν πάλι, υπάρχει ο ισχυρισμός ότι η λήψη ενός παρασκευάσματος αυξάνει την πιθανότητα εγκυμοσύνης μίας γυναίκας σε αγόρι, τότε

H_0: Παρασκεύασμα $\rightarrow$ $p_{ΑΓΟΡΙ} = p_{ΚΟΡΙΤΣΙ}$, έναντι της H_1: Παρασκεύασμα $\rightarrow$ $p_{ΑΓΟΡΙ} > p_{ΚΟΡΙΤΣΙ}$,

τότε,

Σφάλμα τύπου Ι:
Το παρασκεύασμα δεν επηρεάζει την πιθανότητα, αλλά εμείς συμπεραίνουμε ότι την επηρεάζει.

Σφάλμα τύπου ΙΙ:
Το παρασκεύασμα επηρεάζει την πιθανότητα, αλλά εμείς συμπεραίνουμε ότι δεν την επηρεάζει

Στην περίπτωση αυτή το σφάλμα τύπου Ι είναι πιο σημαντικό, γιατί θα το εμπιστευτούν εσφαλμένα τα ζεύγη που επιθυμούν αγόρι.

Σφάλματα Στατιστικών Ελέγχων

Παράδειγμα 1
Η μέση τιμή των οικογενειακών αυτοκινήτων είναι 27.000 €. Διενεργείται ένας στατιστικός έλεγχος για να διαπιστωθεί εάν ο ισχυρισμός είναι αληθής.

Να αναφέρετε:
(α) Την μηδενική (στατιστική) και την εναλλακτική (ερευνητική) υπόθεση.
(β) Τα σφάλματα Τύπου Ι και Τύπου ΙΙ, περιφραστικά ως πλήρεις προτάσεις.

Λύση

(α) H_0: μ = 27.000 €, H_1: μ ≠ 27.000 €.

(β)
Σφάλμα τύπου Ι: Η μέση τιμή είναι πράγματι 27.000 €, αλλά εμείς απορρίπτουμε την H_0.
Σφάλμα τύπου ΙΙ: Η μέση τιμή δεν είναι 27.000 €, αλλά εμείς δεν καταφέρνουμε να απορρίψουμε την H_0.

Σφάλματα Στατιστικών Ελέγχων

Παράδειγμα 2

«Κόκκινη παλίρροια» ονομάζεται η παρουσία δηλητηριωδών φυκιών. Αυτά, παρουσιάζονται και ευημερούν κάτω από ορισμένες καιρικές συνθήκες και συνθήκες νερού. Στην περίπτωση αυτή τα οστρακοειδή που ζουν στην περιοχή αναπτύσσουν επικίνδυνα επίπεδα μιας τοξίνης που προκαλεί παράλυση. Στη Μασαχουσέτη, το Τμήμα Θαλάσσιας Αλιείας (DMF) παρακολουθεί τα επίπεδα της τοξίνης στα οστρακοειδή με τακτική δειγματοληψία οστρακοειδών κατά μήκος της ακτογραμμής. Εάν το μέσο επίπεδο τοξίνης στα μύδια υπερβαίνει τα 800 μg τοξίνης ανά κιλό, η συγκομιδή τους απαγορεύεται μέχρι να τελειώσει το φαινόμενο και να υποχωρήσουν τα επίπεδα τοξίνης στα μύδια.

Να αναφέρετε:

(α) Την μηδενική (στατιστική) και την εναλλακτική (ερευνητική) υπόθεση.

(β) Τα σφάλματα Τύπου Ι και Τύπου ΙΙ, περιφραστικά ως πλήρεις προτάσεις.

Λύση

(α) H_0: μ $\leq$ 800 μg, H_1: μ > 800 μg.

(β) Σφάλμα τύπου Ι: Η μέση τιμή είναι $\leq$ 800 μg, αλλά εμείς απορρίπτουμε την H_0 καταλήγοντας ότι δεν πρέπει να καταναλώνουμε τα όστρακα.

Σφάλμα τύπου ΙΙ: Η μέση τιμή είναι > από 800 μg, αλλά εμείς δεν απορρίπτουμε την H_0.

$H_0: p \geq 0,88$, $H_1: p < 0,88$

Σφάλματα Στατιστικών Ελέγχων

Άσκηση 1

Έστω ο ισχυρισμός H_0: Το ποσοστό των ενηλίκων που εργάζονται είναι τουλάχιστον 88%. Προσδιορίστε τα σφάλματα τύπου I και τύπου II μεταξύ των παρακάτω ενδεχομένων:

E_1: Δεν απορρίπτουμε την υπόθεση ότι το ποσοστό των ενηλίκων που εργάζονται είναι τουλάχιστον 88% όταν αυτό το ποσοστό είναι στην πραγματικότητα μικρότερο από 88%. : Σφάλμα τύπου II.

E_2: Δεν απορρίπτουμε την υπόθεση ότι το ποσοστό των ενηλίκων που εργάζονται είναι τουλάχιστον 88% όταν αυτό το ποσοστό είναι στην πραγματικότητα τουλάχιστον 88%.

E_3: Απορρίπτουμε την υπόθεση ότι το ποσοστό των ενηλίκων που εργάζονται είναι τουλάχιστον 88% όταν αυτό το ποσοστό είναι στην πραγματικότητα μικρότερο από 88%.

E_4: Απορρίπτουμε την υπόθεση ότι το ποσοστό των ενηλίκων που εργάζονται είναι τουλάχιστον 88% όταν αυτό το ποσοστό είναι στην πραγματικότητα τουλάχιστον 88%. = Σφάλμα τύπου I.

Υπόδειξη: Βρείτε την H_1.

Μη παραμετρικές δοκιμασίες
(Non parametric tests)

Μη παραμετρικές δοκιμασίες

Πολλές φορές ωστόσο, η κατανομή του δείγματος δεν επιβεβαιώνεται με σαφήνεια, γεγονός που καθιστά απαγορευτική τη χρήση αυτών των μεθόδων.

Στην περίπτωση αυτή καταφεύγουμε σε μία οικογένεια επαγωγικών δοκιμασιών οι οποίες ονομάζονται μη παραμετρικές και δεν έχουν προϋποθέσεις για την εφαρμογή τους έχοντας ωστόσο μικρότερη ισχύ από τις παραμετρικές καθώς αντικαθιστούν τις παρατηρήσεις της μεταβλητής με τις αντίστοιχες τάξεις αυτών (δηλαδή απλά τη σειρά τους στα ταξινομημένα δεδομένα) γεγονός το οποίο αλλάζει (ενδεχομένως) δραστικά το είδος και την ποιότητα της πληροφορίας που μας δίνει το δείγμα.

Αρμόζει, ωστόσο να αναφερθεί πως ένα μέρος της επιστημονικής στατιστικής κοινότητας προτείνουν τις μη παραμετρικές δοκιμασίες ως ισοδύναμες ή ακόμα και περισσότερο αξιόπιστες επιλογές έναντι των παραμετρικών μεθόδων καθώς στις περισσότερες περιπτώσεις πραγματικών ερευνών τα δεδομένα δεν έχουν ιδιαίτερες ανωμαλίες και η μετατροπή τους από παρατηρήσεις σε τάξεις δεν αλλάζει την ουσιαστική πληροφορία τους ενώ παράλληλα η εγκυρότητά τους ως υπολογιστικές διαδικασίες δεν εξαρτάται από την κανονικότητα ή μη κάποιας κατανομής.

Μη παραμετρικές δοκιμασίες

Αν δεν συντρέχουν οι απαραίτητες προϋποθέσεις για τη χρήση μίας παραμετρικής μεθόδου τότε δεν χρησιμοποιούμε μη παραμετρικές μεθόδους πριν να ανατρέξουμε στη βιβλιογραφία και να ελέγξουμε αν παρά ταύτα θεωρείται έγκυρη η εφαρμογή της για τα δεδομένα του δείγματος μας.

Για παράδειγμα το t - test για 1 δείγμα μπορεί να χρησιμοποιηθεί ακόμα και για μη κανονική κατανομή όταν το μέγεθος του δείγματος είναι μεγάλο.

Αν δεν συμβαίνει κάτι τέτοιο τότε αναγκαστικά εφαρμόζουμε μία μη παραμετρική μέθοδο.

Μη παραμετρικό t - test για ένα δείγμα

Αποφασίζουμε τη χρήση μη παραμετρικής δοκιμασίας για ένα δείγμα στις εξής περιπτώσεις:

Η μεταβλητή δεν είναι συνεχής αλλά διατακτική (π.χ. παίρνει τις τιμές 1, 2, 3, 4, 5 όπως σε ένα ερωτηματολόγιο) και έχουμε μικρό δείγμα.

Υπάρχει μεγάλη απόκλιση από την κανονική κατανομή.

Μία εναλλακτική είναι ο προσημικός έλεγχος (sign test) με τον οποίο ελέγχεται η υπόθεση πως η διάμεσος των τιμών είναι ίση με μία υποτιθέμενη τιμή. Ενδεικτικά, η υλοποίησή του με την R, για τα δεδομένα παραδείγματος με τη βιομηχανία γάλακτος θα ήταν η εξής:

```
library(BSDA)

SIGN.test(one.sample.data, md = 500)
```

Μη παραμετρικό t - test για ένα δείγμα

Με την παραπάνω εντολή, αν δ είναι η (άγνωστη) διάμεσος του βάρους όλης της παραγωγής, ελέγχεται η στατιστική υπόθεση: H_0: δ = 500, έναντι της ερευνητικής υπόθεσης: H_1: δ ≠ 500.

One-sample Sign-Test

data: one.sample.data

s = 3, p-value = 0.5078

alternative hypothesis: true median is not equal to 500

95 percent confidence interval:

 489.3   502.4

sample estimates:

median of x

      498.5

Από τα αποτελέσματα της δοκιμασίας συμπεραίνουμε πως η υπόθεση H_0: δ = 500, δεν απορρίπτεται έναντι της H_1: δ ≠ 500 σε επίπεδο σημαντικότητας 0,05 (p = 0,508 > 0,05).